NEAR EAST UNIVERSITY
INSTITUTE OF GRADUATE STUDIES
ECONOMICS PROGRAM

Generating Hybrid Machine Learning Algorithm to Predict Exchange Rate Fluctuations in Volatile Macroeconomic and Political Environments

HÜSEYİN İLKER ERÇEN

Ph.D. Thesis

NICOSIA
2021

Generating Hybrid Machine Learning Algorithm to Predict Exchange Rate Fluctuations on Volatile Macroeconomic and Political Environments

HÜSEYİN İLKER ERÇEN

NEAR EAST UNIVERSITY INSTITUTE OF GRADUATE STUDIES
ECONOMICS PROGRAM

Ph.D. Thesis

THESIS SUPERVISORS:
PROF. DR. HÜSEYİN ÖZDEŞER (Supervisor)
ASSOC. PROF. DR. TURGUT TÜRSOY (Co-Supervisor)

NICOSIA
2021

ACCEPTANCE/APPROVAL

We as the jury members certify the ‘Generating Hybrid Machine Learning Algorithm to Predict Exchange Rate Fluctuations on Volatile Macroeconomic and Political Environments’ prepared by Hüseyin İlker Erçen defended on 20/08/2021 has been found satisfactory for the award of the degree of PhD in Economics.

JURY MEMBERS

Professor. Dr. Hüseyin Özdeşer (Supervisor) (Head of Jury)
Near East University
Faculty of Economics and Administrative Sciences, Department of Economics

Associate Professor Dr. Turgut Türsoy (Co-Supervisor)
Near East University
Faculty of Economics and Administrative Sciences, Department of Banking and Finance

Assistant Professor Dr. Andisheh Saliminezhad
Near East University
Faculty of Economics and Administrative Sciences, Department of Economics

Assistant Professor Dr. Mehdi Seraj
Near East University
Faculty of Economics and Administrative Sciences, Department of Economics

Assistant Professor Dr. Hasan Rüstemoğlu
Cyprus International University
Faculty of Economics and Administrative Sciences, Department of Business Administration

Assistant Professor Dr. Husam Rjoub
Cyprus International University
Faculty of Economics and Administrative Sciences, Department of Accounting and Finance

Professor Dr. K. Hüsnü Can Başer
Institute of Graduate Studies
Director

DECLARATION

I, Hüseyin Ilker Erçen hereby declare that this dissertation entitled: Generating Hybrid Machine Learning Algorithm to Predict Exchange Rate Fluctuations in Volatile Macroeconomic and Political Environments has been prepared by myself under the guidance and supervision of Prof. Dr. Hüseyin Özdeşer (Supervisor) and Assoc. Prof. Dr. Turgut Türsoy (Co-Supervisor) in partial fulfilment of the Near East University, Graduate School of Social Sciences regulations and does not to the best of my knowledge breach and Law of Copyrights and has been tested for plagiarism and a copy of the result can be found in the Thesis.

- The full extent of my Thesis can be accessible from anywhere.
- My Thesis can only be accessible from Near East University.
- My Thesis cannot be accessible for two (2) years. If I do not apply for extension at the end of this period, the full extend of my Thesis will be accessible from anywhere.

Date:

Signature:

Name Surname: Hüseyin Ilker Erçen

ACKNOWLEDGEMENTS

I sincerely would like to express my sincere gratitude to my supervisors, Prof. Dr. Hüseyin Özdeşer and Assoc. Prof. Dr. Turgut Türsoy for sharing their immense knowledge and expertise, continuous support, patience, enthusiasm, and guidance throughout this programme. It was an honour having them as my supervisors, and benefitting from their wisdom. Their contributions have added a considerable value to my study.

Beside my supervisors, I would like to express my appreciations to my thesis committee members, Assist. Prof. Dr. Andisheh Saliminezhad and Assist. Prof. Dr. Mehdi Seraj for their valuable comments and suggestions, which guided me through the critical points, and realize my weaknesses.

I would like to express my gratefulness to my friend Mustafa Arıcı, who patiently and enthusiastically shared his expertise and vast knowledge on machine learning algorithms and artificial intelligence to guide me to construct the algorithm flawlessly. I also would like to express my gratefulness to my friend Ersin Aytaç, who mentored, enlightened and guide me to lay the foundations of my intellectual being. I couldn’t have the mindset that I proudly have today without Mustafa and Ersin’s contributions.

Finally, I am extremely grateful to my parents, who always prioritize me, and my education beyond their life. I can’t thank them enough for their selfless love, care, endless support, and sacrifices to be able to provide me better education and life.

ABSTRACT

Generating Hybrid Machine Learning Algorithm to Predict Exchange Rate Fluctuations in Volatile Macroeconomic and Political Environments

This thesis aims to pioneer the paradigm shift on the prediction methodology of the field of economics. By bringing a novel multidisciplinary approach, thesis is going to implement qualitative and subjective data alongside traditionally preferred quantitative and objective data so as to augment the performance of exchange rate predictions. To achieve mentioned goal, the author has generated a complex hybrid machine learning algorithm based on the computational techniques of artificial intelligence, namely; Natural Language Processing, Fuzzy Logic, and Support Vector Machines, alongside Prospect Theory of behavioral economics. The thesis aims to impersonate the subjective perception and initial instinctive behaviors of the investors to the macroeconomic and political breaking news via robust artificial intelligence that comprehends the investors' evaluation process of real world scenarios, and targets to simulate the initial responsiveness of exchange rate under the influence of varying complicated macroeconomic and political situations and environments.

Turkey has a strategic influence on global energy trade, geopolitical obligations in the Middle East as a NATO member. Geographical complications on the region caused by ongoing oil wars trigger the political tension in the region, and have adverse impact on Turkish economy. The most obvious proof of this is the remarkable 152% depreciation on the value of TRY against USD in relatively short date interval besides the sharp downtrend of macroeconomic indicators, namely; consumer price index, interest rate, unemployment rate, balance of trade, and credit ratings, from 29 December 2017 to 01 November 2019. Thus, the thesis studies USD/TRY exchange rate during mentioned date interval in daily basis, alongside mentioned macroeconomic indicators. The impact of ongoing political tension on Turkish economy has comprehended by examining the news feed

regarding to political tension, and considering the degrees of political certainty and consensus that the news articles accommodate. In order to perceive aforementioned goals, nearly 1000 news articles regarding to macroeconomic variables and political environment have been examined.
Consequently, the usability of macroeconomic and political news articles to derive the subjective perception of investors, thus their initial responses and behaviors has been confirmed, while achieving a significant prediction performance.

Keywords: Hybrid machine learning, Fuzzy logic, Support vector regression, Exchange rate forecasting, Behavioral economics, Political economics

ÖZ

Değişken Makroekonomik ve Politik Ortamlarda Döviz Kuru Dalgalanmalarını Öngörmek için Hibrit Makine Öğrenimi Algoritması Oluşturma

Bu tez, ekonomi alanının tahmin metodolojisindeki paradigma değişimine öncülük etmeyi amaçlamaktadır. Tez, yeni bir multidisipliner yaklaşım getirerek, döviz kuru tahminlerinin performansını artırmak için geleneksel olarak tercih edilen nicel ve nesnel verilerin yanı sıra nitel ve öznel verileri de kullanmaktad. Söz konusu hedefe ulaşmak için yazar, yapay zekanın hesaplama tekniklerine dayanan ve Beklenti Teorisi, Doğal Dil İşleme, Bulanık Mantık ve Destek Vektör Makineleri'nden oluşan bir hibrit makine öğrenme algoritması oluşturmuştur. Tez, yatırımcıların gerçek dünya senaryolarını değerlendirme sürecini kavrayan güçlü bir yapay zeka aracılığıyla, makroekonomik ve politik son dakika haberlerine karşı yatırımcıların öznel algılarını ve ilk içgüdüsel davranışlarını taklit etmeyi ve döviz kurunun karmaşık makroekonomik ve politik ortamları belirten haberler karşısındaki anlık tepkisini simüle etmeyi amaçlamaktadır.

Türkiye'nin küresel enerji ticaretindeki stratejik etkisi ve bir NATO üyesi olarak Orta Doğu'daki yükümlülükleri büyüktür. Halehazırda komşu ülkelerinde devam etmekte olan petrol savaşlarının bölgede yarattığı coğrafi karışıklıklar, bölgedeki siyasi gerilimi tetikleyerek Türkiye ekonomisini olumsuz yönde etkilemektedir. Bunun en bariz kanıtı, makroekonomik göstergelerin keskin düşüş eğiliminin yanı sıra, TL'nin USD karşısında nispeten kısa bir tarih aralığı olan 29 Aralık 2017 - 01 Kasım 2019 tarihleri arasında %152'lik olağandışı değer kaybıdır. Bu nedenlerden dolayi bu tez, tüketici fiyat endeksi, faiz oranı, işsizlik oranı, ticaret dengesi ve kredi notları gibi makroekonomik göstergelerin yanı sıra söz konusu tarih aralığında USD/TL kurunu günlük bazda incelemektedir. Süregelen siyasi gerilimin Türkiye ekonomisi üzerindeki etkisi, siyasi gerginliğe ilişkin haber akışı incelenerek ve haberlerin barındırdığı siyasi kesinlik ve uzlaşı dereceleri dikkate alınarak kavranmıştır.

Söz konusu hedeflerin algılanabilmesi için makroekonomik değişkenler ve siyasi ortamla ilgili 1000'e yakın haber incelenmiştir.
Sonuç olarak, makroekonomik ve politik haber makalelerinin yatırımcıların öznel algılarını, dolayısıyla ilk tepkilerini ve davranışlarını türetmede kullanılabilirliği doğrulanırken, oldukça yüksek bir tahmin performansı elde edilmiştir.

Anahtar Kelimeler: Hibrit makine öğrenmesi, Bulanık mantık, Destek vektör regresyonu, Döviz kuru tahmini, Davranışsal ekonomi, Politik ekonomi

TABLE OF CONTENTS

CHAPTER 3

OUTLOOK ON THE GEOPOLITICAL BACKGROUND OF TURKEY

CHAPTER 4

METHODOLOGY

CHAPTER 5

THE INSTANTANEOUS IMPACT OF RECENTLY PUBLISHED MACROECONOMIC NEWS ON EXCHANGE RATE

CHAPTER 6

IMPACT OF PARTICULAR POLITICAL EVENT CATEGORIES ON EXCHANGE RATE

LIST OF TABLES

LIST OF FIGURES

ABBREVIATIONS

AI	Artificial Intelligence
ANN	Artificial Neural Network
ARCH	Autoregressive Conditional Heteroskedasticity
ARIMA	Auto Regressive Integrated Moving Average
BOT	Balance of Trade
CV	Cross Validation
DL	Deep Learning
EMH	Efficient Market Hypothesis
EVS	Explained Variance Score
FL	Fuzzy Logic
FN	False Negative
FP	False Positive
FWA	Fuzzy Weighted Average
GARCH	Generalized Conditional Heteroskedasticity
GMO	Generalized Mean Operator
GS	Grid Search
INF	Inflation Rate
INT	Interest Rate
LA	Learning Automata
MAE	Mean Absolute Error
MAPE	Mean Absolute Percentage Error
MCCV	Monte Carlo Cross Validation
ML	Machine Learning
MO	Mean Operator
MSE	Mean Square Error
NLP	Natural Language Processing
NN	Neural Network
OH	Overreaction Hypothesis
PCA	Principal Component Analysis
PCR	Principal Component Regression
PLS	Partial Least Square Regression

PT	Prospect Theory
RMSE	Root Mean Square Error
R^2	Coefficient of Determination
SAD	Seasonal Affective Disorder
SI	Scatter Index
SV	Support Vector
SVM	Support Vector Machine
SVR	Support Vector Regression
TN	True Negative
TP	True Positive
TRY	Turkish Lira
UIH	Uncertain Information Hypothesis
UNEMP	Unemployment Rate
US	United States
USD	US Dollar
USD/TRY	US Dollar/Turkish Lira
VMS	V Measure Score
WFL	Weighted Fuzzy Logic
WPT	Weighted Prospect Theory
WSVR	Weighted Support Vector Regression

Chapter 1
Basis of the Study and Methodology

1.1 Introduction

Economic condition and policy of a country is highly dependent on exchange rates (Kandil et al., 2007; Cehreli et al., 2017) due to its significant role on international trade, investment determination, risk management, and balance of payments (Sharma et al, 2016; Korol, 2014). One of the most challenging, yet important difficulties concerning time series analysis is exchange rate and stock market prediction due to their non-stationary nature (Henrique et al., 2019). The nonlinearity of price time series triggered by their dynamic, chaotic, and noisy nature (Bezerra and Albuquerque, 2017; Kumar and Thenmozhi, 2014) that are remarkably dependent on the information obtained from macroeconomic indicators and political events, psychology and expectations of investors derived from obtained information, besides industrial characteristics (Kara et al., 2011; Chen et al., 2017; Zhong and Enke, 2017).

Analyzing the impact of economic uncertainty and political tension on macroeconomic variables alongside the fluctuating exchange rates is not something new. Various participations to literature had been made in 1980s and 1990s on this particular area by notable authors such as; Ben S. Bernanke (1983), Dani Rodrik (1991), and Luigi Guiso and Giuseppe Parigi (1999). However, especially insufficient technological advancements, besides to the accessibility of publicly available information had been main limiting factors for parallel participations to the literature.

Researches on macroeconomics and finance broadly argue that uncertainty is a restrictive factor for investments. For investors, formulation and implementation of economic policies, and commitment on coping with political instability and tension under uncertain environment plays a vital role on their investment decision (Antonakakis et al., 2013).

Nowadays, popularity of researches on uncertainty reappeared through the encouragement by great enhancements on computing power, accessibility of information, and experiencing economic volatilities besides the recessions highly derived by political and economic uncertainty (Bloom, 2014). Furthermore, retroactive researches indicate economic policy uncertainty and political instability such as geopolitical conflicts, terrorism and immigrant crisis as the main reason for slow global economic growth over past decade (Gholipour, 2019).

Hereby, the significant impact of publicly available political and financial information on financial markets and economics are broadly accepted. Alongside to objective and quantitative variables such as historical price data, subjective and qualitative information recently obtained an essential role in economic and financial analyzing and forecasting. Recent technological advancement unlocked an unlimited potential and provides data scientists an opportunity to deeply analyze the complex combinations of historical datasets (Chiang et al., 2016) by intense computational use of intelligent forecasting models that has been established in the literature under the title of machine learning (Sobreiro et al, 2019). More recently, machine-learning techniques are superseding traditional time series prediction techniques (Wang et al., 2012) by their nature of design to particularly cope with non-linearity, uncertainty, and randomness (Chen et al., 2017), while promising an area of research in forecasting time series (Wang and Zhang, 2017). Lately, machine-learning (ML) algorithms such as; Support Vector Machines (SVM), Fuzzy Logic (FL), and Artificial Neural Network (ANN) have been used for stock market prediction (Weng et al., 2017; Oliveira et al., 2017; Bollen et al., 2011; Schumaker and Chen, 2009; Kim, 2003), exchange rates prediction (Sharma et al., 2016; Cehreli et al., 2017; Yasir et al., 2019), and to examine the effects of political tension and instability on economics (Aisen and Veiga, 2013; Chuku et al., 2019).

This thesis aims to bring a novel approach to exchange rate prediction by investigating the impacts of recently published news and officials' statements regarding on the macroeconomic fundamentals, in addition to political environment, which is going to be monitored deeply by examining the political tension through the degrees of certainty and consensus. A hybrid machine-learning algorithm is going to be generated by the author to unveil the correlation and the significance of political and economical statements on exchange rates fluctuations, and predict the price of the exchange rates on a daily basis.

News regarding to the governments' official statements, expectations on the attitude on politics, international relations, economic policies, and publication of macroeconomic data plays a crucial role on informing the community about recent politic and economic situation that country is currently in. The information provided by the news plays a significant role on the behavior of investors' in a short run.

Due to their extremely responsive nature on domestic and foreign economic and political events, emerging economies (Mehdian et al., 2008) are perfect fit to study. Turkey is classified as an emerging economy. The substantial influence of Turkey on its western borders neighbor European Union on energy trading, alongside the eastern borders on ongoing oil wars in the Middle East cannot be ignored. Thus, E.U. and U.S. have applied a dramatic and also perpetual political and economical pressure to undermine Turkey's empowerment, growth, hereby independence, thus interfere to Turkey's foreign policies to match their expectations. The political pressure that has been applied by U.S. caused political tension, instability, and uncertainty. Besides those causes that has been triggered by political pressure, classification of Turkey as emerging economy have triggered a momentous impact on investors' rational expectations, thus prevent foreign and domestic investments to Turkey, which impacts Turkey's macroeconomic indicators, and currency negatively. If the aforementioned components deeply analyzed, the causality between political tension, political uncertainty, and excessive economic volatility may be illuminated in logical and scientific terms.

Turkey is recently having consecutive political events that triggered political instability and uncertainty, which results in abruptly dramatic fluctuations in exchange rate, thus macroeconomic variables relatively. Based on the exact specific reasons that have been stated above, Turkey has selected to be studied.

In order to examine the correlation between news and announcements considering the political and economic environment, and exchange rate; recently published news are going to be observed, and responsiveness of US Dollar/Turkish Lira exchange rate is going to be examined in daily basis from 29 December 2017 to 01 November 2019. The main reason to select this particular date interval is the strong and sharp fluctuation of the USD/TRY exchange rate that caused by numerous political and economic factors. During this date interval; the value of TRY depreciate by an exceptional 152% to USD, and hit record braking lows while becoming the worst performing currency of emerging markets. In the making of that tremendous fall, various political cases, which predominantly comprise Pastor Andrew Brunson case, S-400 crises, eventful Istanbul Mayoral elections, Presidential elections, and Parliamentary elections played a significant role. However, what all political events have in common is; uncertainty and tension. Selected date interval provides a unique opportunity to study the effects of different kinds of political tension and uncertainty on diverse issues in every aspect. Interaction between above-mentioned diversified political issues, their effect on TRY's value, and macroeconomic fundamentals of Turkey cannot be ignored, as exchange rate and macroeconomic variables causes each other. For that reason; Consumer Price Index (CPI), Interest Rate (IntR), Unemployment Rate (UnempR), Balance of Trade (BoT), and Credit Ratings (CR) are going to be examined as fundamental macroeconomic variables. Published news and recent announcements of officials' regarding to each mentioned macroeconomic variable between previously stated date interval are going to be examined to comprehend the sudden impact of macroeconomic news on expectations of the market and/or official statements on the value aside from

historical macroeconomic data, which is going to be observed in monthly basis due to their nature.
By figuring out the significance of the correlation between political and economic news on exchange rate and macroeconomic variables, thesis is aiming to predict exchange rate with smallest margin of error possible in daily basis. Literature supports the superiority of hybrid machine learning techniques that accommodate fuzzy logic and/or support vector machines on traditional econometrics models. Therefore, this thesis is going to adopt robust amalgamation of machine learning techniques to create a hybrid machine-learning algorithm. A complex combination of natural language processing (NLP), fuzzy logic (FL), and supports vector machines (SVM) soft computing methods in addition to prospect theory (PT) of behavioral economics are going to be used in order to accomplish the purpose of this thesis.

1.2 Brief Introduction to Epistemological Approach

Economics considers that humans behave in line with the information at hand intending to maximize profit and minimize the cost. To do so, it has been taught to contemplate quantitative information.
Eugene Fama had laid the foundations of theoretical and empirical search on information driven financial markets with efficient market hypothesis (EMH) in 1970. Hypothesis is supporting the argument that investors are rational beings, and may react to newly driven information on their decisions. However, it was not possible to test this hypothesis with existing techniques. Thus, hypothesis resulted as financial markets are unpredictable, and follow a random walk. Later, extensions to EMH have been added, such as overreaction hypothesis (OH) (DeBont and Thaler, 1985), and uncertain information hypothesis (UIH) (Brown et al., 1988, 1993). These hypotheses not only support the EMH, but also extend the hypothesis to investors' degree of response on uncertain markets.

Notional segregation between decisions under risk and decisions under uncertainty was first through by the economist Frank Knight in his book *Risk, Uncertainty, and Profit* (1921). Knight defined risk as a state, which probability of every possible outcome can statistically be calculated by the decision maker with certainty. On the other hand, uncertainty is a state that probabilities of possible outcomes are statistically unpredictable. Decision maker is expected to make a decision without sufficient information provided, but should be evaluating the probabilities of possible outcomes with approximate degree of vagueness (Fox and Poldrack, 2009).
In a situation of information deficiency on probabilities, traditional economic analysis with rational roots adopts the attitude that downgrades uncertain states to risky states in order to fit all possible values into a range between 0 and 1 by excluding any possible extreme cases (Tobler and Weber, 2013).

Understanding individuals' behavior under risk and uncertainty plays a tremendous role on unraveling several glitches in economics, as well as in other social sciences (Gregory et al., 2008). Expected utility theory (von Neumann and Morgenstern, 1947) had been accepted as a normative model of rational decisions under uncertainty (Gu et al., 2019). However, after a few decades, a general agreement claimed that it did not describe individuals' process of decision making adequately (Tversky and Kahneman, 1992).
Kahneman and Tuersky (1979) observed human behaviors on decision-making process. Their psychological researches indicate the likelihood of miscalculation on simple financial decisions. Studies reported that, participants avoided risk when expecting profitable results, while prefer to take risk when expected result is negative, which results with a loss that has twice as much impact than a profit of the same rate (Kahneman and Tuersky, 1979). Furthermore, expected utility theory's major deficiencies were addressed by Kahneman and Tversky's (1979) prospect theory (PT).
Psychological researches prove that intuitive biases distinguish real people from the ones who expected to be making rational decisions by evaluating the neo-classical economic theory (Ferguson, 2011).

The dilemma of human mind between conflicting forces of emotion and reason has been started to be debating by Plato, and emphasized by Immanuel Kant, and Sigmund Freud. All aforementioned philosophers had suggested the idea of irrational choices driven by emotions, rather than logical reasons, which shaped today's models of decision-making (Cohen, 2005).
The phrase *emotion* may be defined as a distinct set of psychology driven reactions on a particular internal or external event. Additionally, emotional disorders caused by seasonal affective disorder (SAD), and emotional changes, and fluctuations arise from the biological structure of individuals' influence their risk preferences, and the investor behavior over the masses accordingly (Kamstra et. al., 2003). Emotions may influence choices as a result of provoked action tendencies (Scherer, 2000). In contrast, the phrase *mood* may be defined as a relatively volatile state of subjective mindset that may or may not be related to particular event (Lempert and Phelps, 2013).

In recent years, a field of study has been formed by combining the disciplines of psychology and economics that attempts to clarify the decision making process of individuals and organizations, which is called behavioral economics (Belsky and Gilowich, 1999). To do so, behavioral economics concentrate on the effects of social, cognitive, and emotional preconception. In this context, causality of individuals' and organizations' decision making that resulted in fluctuating financial markets and economic variables are expecting to be comprehended. In addition to psychology, behavioral economics encapsulates neoclassical economics approach, sociology, and decisions that have been taken which affect aforementioned fields (Kurtoglu, 2016).
As a matter of fact, psychological sciences deal with the behavior of humans, and the basic elements that guide them. On the other hand, economics is among the human behaviors that fall under the scope of 'economic behavior' and 'financial behavior'. In other words, it deals with economic and financial decisions of individuals and the reasons beneath these decisions. Therefore, it is not wrong to say that there is a direct relationship between economics and psychology (Rabin, 1998).

Humans are inclinable to replicate the decisions of others without further investigation under uncertain circumstances (Kurtoglu, 2016). When the market's expectations diverge from the foundations of the economics, the likelihood of an emotional shocks depending on expectations increases (Peterson, 2012). Extensive researches indicate that, when individuals' subconscious level of prejudicial or biased thinking tendencies come to the same line, the market prices suddenly become predictable (Peterson, 2012).
For further understanding on the reasoning behind the fluctuations occur on financial markets, psychological aspects of individuals should be considered. Daily emotional state of individuals such as concern, panic, fear, confidence, hope, greed, etc. plays crucial role on global-neoliberal free market (Kurtoglu, 2016). Latest researches have also indicated that, investment decisions of an individual may vary based on his/her current mood. These findings reveal the correlation between psychology and finance, and encourage further studies (Peterson, 2012). Moreover, global analytics and advice firm Gallup claims that the firms implementing behavioral economics principles on their strategies are having higher average sales and profit margin than their rivals.

The greatest contribution of neoclassical perspective to economic sciences has been 'marginal benefits'. 'Marginal benefits' is an example of psychology-economics synthesis that had been developed within the framework of hedonist psychology assumption. It is noticeably observable that, in addition to microeconomic framework, macroeconomic framework is also benefitting from psychological sciences (Eser & Davletkan, 2010). Irving Fisher (1928) contributed to macroeconomics-psychology synthesis by examining the money illusion- quantity theory of money paradox and referring to concepts of perception and error for the first time, while John Maynard Keynes examined the projections of individuals' preferences in the direction of psychological motives such as speculation and involves psychology in his macroeconomic analysis (Eser & Davletkan, 2010).

Nowadays, claims of neo-classical approach on rational behavior of individuals have been falsified. At best, human behavior may be considered as 'bounded rationality'. The reason for this is; individuals are not always rational in their decisions. On the contrary, they generally exhibit behaviors parallel with their emotional status, feelings, and past experiences. Researches confirm that individuals do not act rationally in the face of economic events. Different reactions of individuals on same or similar economic events that took place on different times have been observed as an outcome of altered inner mentality against the various influential variables during the decision making phase (Soydal, 2010).

1.3 Statement of the Research Problem

Fundamentals of economic theory are comprehensively based on mathematical modeling (Lainè, 2014). However, features of cognitive processes had not even been taken into consideration in any phase of construction of economic theory (Camerer et al., 2005), as cognitive practices were comprehended as a mystery (Boudon, 2007). Outweigh of mathematical framework excluded rationality and decision making interpretations within the cognitive frame (Lainè, 2014).

Unfortunately, Adam Smith did not take behavioral factors into consideration while constructing the classical economics, and adopt Newtonian determinism instead. In his book *The Theory of Moral Sentiments* (1759), Smith defended that an invisible hand directs all human behavior, which apparently supports that economic behavior was determined (Slawson, 1981). Smith had emphasized the matter of individuals' behavior in his aforementioned book, however couldn't perceive qualitative measures of psychology into quantitative analytical sciences (Çalık and Düzü, 2009). As a positivist and materialist, Smith believed that human sentiments are objective and measurable. In his book *The Wealth of Nations* (1776), Smith did not separate desire for 'gain' or 'profit' from other human sentiments, and considered them as objectively observable and also quantitatively measurable (Slawson, 1981). Smith's positivist perspective causes utopic notions like Homo Economicus, which represents rational human being that makes logical decisions in every possible scenario in accordance to maximize profit (Çalık and Düzü, 2009).

In his time, it was hard to falsify Smith's deterministic view on economic behavior, because economics was actually functioning in that way. While market forces determining the economic behavior, sellers were almost ineffective on the market due to their limited sources such as; television, radio, widely distributed newspapers and advertisement facilities (Slawson, 1981).

The foundation of classical economics relies on objective structure. Thus, due to nature of positivist approach, classical economics is assumed to be affected neither from personal nor political controversy (Slawson, 1981).

John Maynard Keynes (1938) opposed to positivistic conception of economics, and distinguished economics from natural sciences into his own perspective of moral sciences. Unlike unchanging laws of natural sciences, moral sciences emphasized people's psychological tendencies and reactions under conditions of uncertainty. Keynes debated that considering individuals' psychological propensities emasculates the deterministic characteristic of economics (Slawson, 1981).

Current quantitative analysis methods that have been using to analyze economies for almost 300 years are incapable to enlighten academics and professionals about complicatedly bonded social, emotional, cognitive, and cultural differences of society. While aforementioned complexities may affect societies behavior dramatically, quantitative models are unfortunately ignoring the causality (Appleby, 2012).

Despite to today's comprehensive developments, science of economics and finance still relies on quantitative historical data, and expecting the history to repeat itself in exactly the same way. This insubstantial expectation obviously revels how obsolete and deficient previous methods are (Soydal, 2010).

Understanding socio-economic events is a highly complex task. To do so, theory is needed to define perspective of reality and vision. For this reason, considering what, where, and when socio-economic theory will transform variables like political and social institutions, technology based interchanged business structures and functions and behavior of individuals' from abstract to concrete has a significant importance especially in today's mutable society.

On the other hand, neoclassical-neoliberal economists considers society not as a subject of social sciences, but only as a parameter, and prefers to only rely on historical data, and to not consider aforementioned variables into their analysis (Stretton, 1999). Recent studies have also specified the significance of psychological state of decision maker on the financial decisions, the correlation between psychology and finance, and encourage further studies. Due to aforementioned findings of the recent studies, the necessity of using subjective and qualitative information aside from objective and quantitative variables for economic and financial analyzing and forecasting have arisen significantly in recent years.

1.4 Objective of the Study

Today, scientists who contribute in social, behavioral, psychological, neurological, and computational fields also investigating for the alternative paradigm to contribute in economic theory (Glimcher, 2004). Testing distinctive techniques to predict economic and financial data is a common approach due to their uncertain, imprecise, non-linear, underlying dependencies, and highly difficultly level of predictability (Johnson et al., 2016).

Historical data on it's own is not sufficient to fully understand the dependency, causality, and behavior of financial and economic variables to create an efficiently working forecasting method for imprecise, uncertain, chaotic, and nonlinear environment with extreme amount of underlying dependencies. Existing econometric forecasting methods should be reformed and improved in modernist and innovative way in order to overcome aforementioned circumstances. Roots of the circumstances need to be considered, investigated, observed, simulated, and computed to compose more accurate forecasting method, which is capable to comprise underlying dependencies that essentially causing the uncertainty and nonlinearity to minimize the error term as much as possible. Understanding the causality and dependency between uncertain and desired to be forecasted data is crucial for design the model.

Since machine learning techniques’ and time series’ paths crossed, numerous research papers have been published in both economics and finance fields. Dominant majority of the comparison papers had emphasized the superior performance of the soft computing algorithms, and exposed that; machine-learning algorithms have outperformed econometric models on prediction with a huge margin.

Literature comprises effects of political statements, instability, uncertainty, international relations, and macroeconomic variables besides associated publicly available information on exchange rate separately. Moreover, either qualitative or quantitative data has preferred to be analyzed on leading studies. However, there is not any study that widely accommodates and examines all the aforementioned variables and types of data collectively in a single study. Additionally, superior performances of hybrid machine learning algorithms compared to specific machine learning technique have been proven. These studies have preferred to combine two machine-learning techniques to compose a hybrid algorithm.

This thesis is aiming to contribute to literature by being the first study to use both qualitative and quantitative data during the learning process of the machine learning algorithm, and examine the instantaneous impacts of officials’ statements and recently published news concerning politics, international relations, and macroeconomics on exchange rate fluctuations by using Artificial Intelligence architecture on the hybrid machine learning algorithms, which composed by using three machine learning techniques alongside to economic theories simultaneously.

The main objectives of this thesis include:

1. To pioneer the paradigm shift on the prediction methodology of the fields of economics and finance that has arisen by the advancements on computing power, and soft computing algorithms.
2. Generating a sophisticated Artificial Intelligence algorithm, composed by hybrid machine learning techniques to excel previous prediction techniques.
3. Implementing qualitative and subjective data alongside the quantitative and objective data to enhance the accuracy of exchange rate prediction.
4. Emphasizing the significance of political sciences and international relations on economic fluctuations.
5. Investigating the instantaneous impacts of officials' statements and recently published news concerning politics, international relations, and macroeconomics on exchange rate fluctuations driven by the sentiments of investors.
6. Accentuating on reformist methods to clarify the complications on economic forecasting caused by vague environments.
7. Underlining the significance of multidisciplinary approach by merging the disciplines of economics, finance, political sciences, psychology, and soft computing to enhance the prediction performance, and achieving superior prediction results by lowering the prediction error as much as possible.

1.5 Research Motivations

The new economic era has overcome the traditional methods that have been accepted since the industrial revolution, and expanded to today’s novel perspective and the dimension that sits on the new hybrid behavioral and econometric parameters. The era of behavioral economics brought the theorems within that neoliberal capitalism now accepts, such as chaos theorem, and butterfly effect, which believed to reshape the global economy and financial system and lead to a new structure (Kurtoglu, 2016).

In his book *The Structure of Scientific Revolutions* (1962), Thomas Kuhn expressed paradigms as the scientific accomplishments, which accepted universally that stipulate solutions for existing problems of researchers. Over time, paradigm drifts emerges as current methods obviate to solve recently arisen questions by discoveries made during existing paradigm. Considering the deficient pieces of existing paradigm enforces radical differences in current paradigm to revolutionize. Transformation of paradigm that arises by concentrating on deeper questioning and reasoning is called paradigm shifts (Delanty, 2005). Kuhn advocates paradigm shifts as two-stage process. Shifts occurs either by poor fit of data into existing paradigm, or superiority of alternative frameworks that aims to reconcile broader data into a single conceptual approach (Glimcher, 2004).

Time series analysis tools, such as; moving averages, autoregressive models, discriminating analyses, and correlations (Kumar and Thenmozhi, 2014; Wang et al., 2012) have been using among the classic forecasting techniques, such as; technical analysis with standards of support and resistance and indicators (Chen et al., 2014) to predict economic indicators and financial markets (Sobreiro et al, 2019).

In the existing literature, few studies have compared the prediction performance of ML algorithms with traditional econometric models such as; Multiple Regression (MR), Vector Auto-Regression (VAR), Auto Regressive Integrated Moving Average (ARIMA), and Autoregressive Conditional

Heteroskedasticity /Generalized Autoregressive Conditional Heteroskedasticity (ARCH/GARCH) models. Testing distinctive techniques to predict economic and financial data is a common approach due to their nature of being uncertain, imprecise, non-linear, underlying dependencies, and highly difficultly level of predictability (Johnson et al., 2016). Meanwhile, artificial intelligence algorithms are precisely targeting to cope with the aforementioned complications by nature (Chen et al., 2017), and promising area of research in forecasting time series (Wang and Zhang, 2017). Today's technological advancement unlocked an unlimited potential and provides data scientists an opportunity to deeply analyze the complex combinations of historical datasets (Chiang et al., 2016) by intense computational use of intelligent forecasting models that has been established in the literature under the title of machine learning (Sobreiro et al, 2019).

Since machine learning techniques' and time series' paths crossed, numerous research papers have been published in both economics and finance fields. Dominant majority of the comparison papers had exposed that; machine-learning techniques are outperforming traditional econometric models on prediction.

As literature supports the superiority of hybrid machine learning techniques compared to traditional econometric methods on addressing the aforementioned problems caused by the nature of the data the field dealing with, thus running the prediction accordingly and more accurately, it is possible to verify the paradigm shift on the prediction methods. The urge for scientific development, expanding the boundaries of limitations, and working on a study that may pioneer the paradigm shift on the prediction methodology of the fields of economics and finance that has arisen by the advancements on computing power, and soft computing algorithms has motivated author for this research.

1.6 Brief Overview of the Methodological Approach

The thesis has adopted a multidisciplinary methodology to meet the predetermined objectives of the study that has been mentioned in Chapter 1.4. While key elements of machine learning is going to construct the foundations of Artificial Intelligence algorithm, behavioral economics theories are going to support the algorithm by observing the psychological states of investors under varying pressures caused by complex environments, and examining their impact on investors' decision making process.

A sophisticated combination of natural language processing (NLP), fuzzy logic (FL), and supports vector machines (SVM) soft computing methods in addition to prospect theory (PT) of behavioral economics are going to be used in order to accomplish a realistic prediction by overcoming the real life scenarios, through simulating the real life decision maker's behaviors, which is the fundamental goal of the thesis.

The methodology is constructed as three phases, named as; Sentiment Analysis, Fuzzy Logic, and Support Vector Regression respectively. The first phase is going to be benefitted from NLP and PT, followed by the second and third phases focuses solely on the methodologies of FL and SVM respectively, whereas all phases are interconnected, as the obtained outcome of a phase is going to be feed into the next phase as an input. The first two empirical chapters, chapter 5 and 6 have benefitted from NLP, PT, FL, and SVM, while last two empirical chapters chapter 7 and 8 have benefitted from NLP, besides the advanced variations of PT, FL, and SVM (going to be introduced as; WPT, WFL, and WSVR) in order to meet the requirements of the research and the expectations from each chapter.

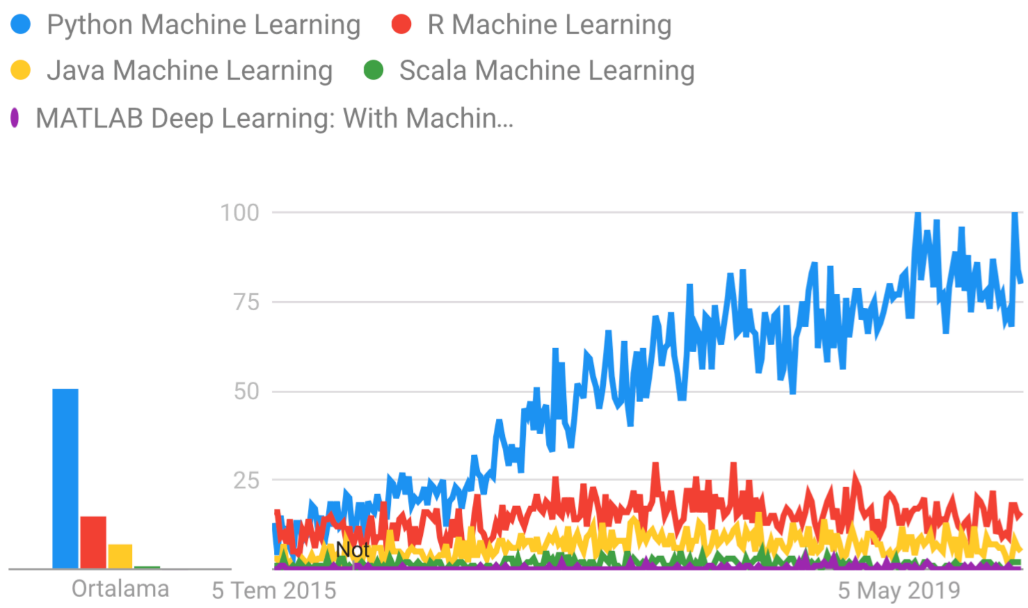
The workflow mechanism of methodology has been constructed uniquely for each chapter in order to match the requirements of the diverse datasets, and explained in depth through every chapter.

Obviously, software and hardware selection plays a crucial role on suitability, applicability, combinability, and efficiency of composed hybrid machine learning algorithm. Selected software and hardware for this thesis has introduced below.

1.6.1 Software Selection

Programming language Python has been adopted in order to construct the hybrid machine learning algorithms that forms the backbone of this thesis. Python language was found by Guido van Rossum in 1989, and released in 1991. During the last decade, Python has practiced a rapid acceptance and became one of the most widespread programming languages, especially in scientific fields, such as; artificial intelligence, data sciences, astronomy, and computational biology. In recent years, Python has also started to be accepted and adopted by the field of economics as it offers epochal possibilities. Machine learning is one of the most important of these, as it differs from studies supported by traditional software, which was introduced previously in this chapter. Deeper research and understanding is essential in order to run a machine learning model. Python is an essential tool, and the most popular programming language to implement machine learning to a data driven field. A global, and independent developer research program Developer Economics have surveyed more than 2000 data scientist in 2017 on programming languages for machine learning applications. According to their survey, 57% of data scientists are using Python for machine learning, while 33% are prioritizing it. R programming language, which is statistical software, and has a reputation among statisticians and data miners is often compared to Python. However, survey results indicated that R has scored as fourth popular language for machine learning as only 31% of data scientists are using, while only 5% are prioritizing it. Moreover, Python's popularity for machine learning purposes has keep on rising in recent years according to Google Trends.

Fig. 1 Worldwide popularity of programming languages for machine learning purposes between 01.07.2015 and 08.02.2020. Comparison has achieved from Google Trends.



(Source: Google Trends)

Python has also established a ground on academic literature, as it became the most common programming language used for artificial intelligence projects, and researches insist of machine learning. Superiority of implementing machine learning algorithms by using Python language for forecasting purposes has also beens proven. Uber Technologies data scientists; Slawek Smyl, Jai Ranganathan, and Andrea Pasqua had won the prestigious M4 Competition held by International Institute of Forecasters by identifying the most accurate forecasting method for a data set of 100,000 time series with a superior accuracy of 95% for the hourly, daily, weekly, monthly, quarterly, and yearly predictions made, while top 10 performing algorithms of the M4 Competition have achieved MAPE between 3% and 10% (Agathangelou et al., 2020). Winner of the M4 forecasting competition have adopted hybrid algorithms coded in high-level programming language, Python.

Briefly, Python is an academically accepted programming language, and the most popular language for machine learning. Its superiority on machine learning and prediction has also been crowned by being used to be winning the most prestigious forecasting competition. By considering aforementioned reasons, Python has adopted to construct a robust hybrid machine learning algorithm to achieve a realistic prediction by overcoming the real life scenarios, through simulating the real life decision maker's behaviors, which is the fundamental goal of the thesis.

Domain knowledge plays a crucial role on predicting time series. However, hybrid machine learning algorithms provide an ability to encode the necessary knowledge, and achieve superior prediction outcomes well matched to real world scenarios (Smyl et al., 2018).

1.6.2 Hardware Selection

In order to overcome the hardware requirements of processing robust hybrid machine learning algorithms, altered hardware were selected carefully to fulfill the needs of computing power on particular phases of the methodology used, as some phases need multi-cores to perform fast, and on superior levels while running complex calculations, while others can only perform with a single core. For the algorithms that perform faster and superior on multi-core architectures; IBM Supercomputer of Near East University, which is located in the Innovation and Information Technologies Center has been used. The Cluster 1 of IBM Supercomputer that has been benefitted by the complex algorithms of the thesis has specifications with; IBM Blade HS21 XM, 1280 Cores, Intel Xeon 2.33GHz Processor, 25.6TB total Memory as 2GB per Core, Infiniband 4x 20GB per second Interconnection, and the performance value of Rmax (Gflops) 9243.89 – Rpeak (Gflops) 11945. Near East University is ranked as 13th for the computation speed and capacity on global rankings. For the algorithms that can only be performed on a single-core architecture; Apple MacBook Pro with specifications of Intel i7 3.0GHz Processor, 16GB Memory, and Intel Iris graphics has been used.

1.7 Brief Overview of the Thesis

The thesis comprises eight chapters apart from preliminary and finalizing sections. The first chapter is going to be dedicated to the basis of the study and methodology. Theoretical and empirical review of the literature is going to be presented in Chapter 2. In Chapter 3, geopolitical background of Turkey is going to be introduced. The influence of Turkey on global energy trade, significant role on the ongoing oil and natural gas wars on the Middle East, and the economic impact of locating on a geopolitically strategic region is going to be studied throughout this chapter.

The methodology is going to be presented exhaustively in Chapter 4. Methods and theories that formed the hybrid machine learning algorithm, namely; natural language processing, learning automata, prospect theory, fuzzy logic, and support vector machines are going to be explained scientifically, and the principles of the hybrid algorithm are going to be clarified in this chapter.

Chapter 5 is going to investigate the initial impacts of statements and announcements on economic policy, official briefings on macroeconomic fundamentals, published news regarding to government officials' statements, expectations on the attitude on economic policies, and experts' interpretation on macroeconomic fundamentals of the country on exchange rate. By examining aforementioned macroeconomic news contents, chapter aims to comprehend the correlation between each macroeconomic indicator and exchange rate, and predict exchange rate in daily manners accordingly.

On Chapter 6, the effectiveness of officials' statements regarding to politics and international relations, publicly published news on various kinds of political cases and events, and their degree of certainty and consensus on exchange rate are going to be investigated. This chapter predominantly concentrated on four main political events that have carefully selected to be homogenously distributed as national and international events, namely; Pastor Andrew Brunson case, Parliamentary and Presidential elections, S-400 crisis, and eventful Istanbul Mayoral elections. Each case is going to be investigated separately in daily manners to comprehend how exchange rate behaves on different kinds of political events, and the degree of certainty and consensus exclusively. The initial impact of the publication of news regarding to each political case and event, on exchange rate has been examined through this chapter, and exchange rates predicted in daily manners by using obtained dependency data.

Chapter 7 is constructed on the foundations of Chapter 6 concerning specific political events, cases, and categories. This chapter aims to emphasize the correlation between officials' statements, speeches addressing the nation, speeches addressing government agencies and officials, recently published political news on media regardless of the news' political category, event, or case involving Turkey. This chapter intends to measure the initial responsiveness of exchange rate to government officials' statements, speeches, and recently published political under the degrees of certainty and consensus considered from the perspective of investors.

The objective of Chapter 8 is to simulate the actual frequency and contents of news feed, impersonate the subjective viewpoint and instinct behaviors of the investors to the breaking news via robust AI, and achieve a highly accurate exchange rate prediction. The chapter is going to be benefitted from the findings of the chapters 5, 6 and 7 to weight the macroeconomic and political breaking news parallel to their real world hierarchical order of importance on investors' subjective perspective. The chapter finalizes the thesis by developing a pioneer prediction method that concentrates on multiple disciplines, and reflects the real life behaviors and scenarios.

Chapter 2
General Review of Theoretical and Empirical Literature

This chapter is dedicated to briefly present the pioneering academic studies that advocate the significance of political studies, sentiment analysis, alongside to macroeconomic indicators on predicting economic and financial time series. Recent literature emphasizes the impact of subjective qualitative data, such as; news articles and officials' statements on political and economic uncertainty, instability, tension on objective quantitative variables like macroeconomic variables and exchange rates. Researches on macroeconomics and finance broadly argue that uncertainty is a restrictive factor for investments. For investors, formulation and implementation of economic policies, and commitment on coping with political instability and tension under uncertain environment plays a vital role on their investment decision (Antonakakis et al., 2013). Today, popularity of researches on uncertainty regained through the reinforcement by excessive augmentations on computing power, accessibility to knowledge, and experiencing economic volatilities besides the recessions highly derived by political and economic uncertainty (Bloom, 2014). Besides, retrospective researches specify economic policy uncertainty and political instability like geopolitical conflicts and terrorism is the main reason for slowing global economic growth over past decade (Gholipour, 2019). Testing distinctive techniques to predict economic and financial data is a common approach due to their uncertain, imprecise, non-linear, underlying dependencies, and highly difficultly level of predictability (Johnson et al., 2016). However, analyzing subjective and qualitative data is a complex task for traditional prediction methods. Hence, to be able to apply advanced analysis methods in pursuit of achieving higher prediction accuracy, multidisciplinary approaches have started to be observed and adopted in the field of economics. To be able to manage complex multidisciplinary approach that combined fields of political sciences, psychology, and economics, intense computational use of intelligent forecasting models that has been established in the literature under the title of machine learning (Sobreiro et al, 2019).

This chapter concentrates on the performance of time series prediction achieved by variations of machine learning algorithms composed to analyze both subjective qualitative and objective quantitative data in the existing literature. Research papers that identifies the significance of political tension and instability on economic variables, publicly available information, role of sentiment analysis on predicting uncertain environments/throughout events causing uncertainty, methodological approaches on varying data analysis, and predictive performance of machine learning algorithms are going to be observed throughout existing literature, and findings are going to be clarified.

To begin with, researches on predictive performance of machine learning algorithms are going to be present, which is going to be followed by the studies that emphasize the significance of political studies, uncertain environments, and their impact of macroeconomic variables. Finally, researches that accommodate the sentiment analysis as essence to achieve a successful prediction are going to be presented.

Kim (2003) has written the most cited article in the Scopus database on his field. Research aimed to use 12 technical indicators to predict the direction of movement of Korean Stock Market Index (KOSPI) in daily basis. 2928 trading days between January 1989 and December 1998 had been observed. Support Vector Machines methodology has been used from the software library LIBSVM, and the polynomial kernel and the Gaussian radial basis functions have been selected as the kernel function of SVM. Successful predictive results obtained by application of SVM indicate the sensitivity of the value of parameters. The results of the study been compared with the results obtained with Case-Based Reasoning (CBR), and Neural Networks (NN).

Implementation of structural risk minimization principle positioned SVM superior to other methods, and higher prediction accuracy has been observed by SVM among all the methods used. Moreover, SVM referred as a promising alternative for financial time-series forecasting.

Korol's study (2014) is one of the world's first attempts on exchange rate forecasting that combines fundamental analysis with machine learning. Besides the macroeconomic variables, many external factors such as political and psychological may also lead to speculative trading which has a direct effect on foreign or direct currencies' pricing. Fuzzy logic model have been selected due to its adjustability and modeling to include external factors to analysis. Fuzzy logic model has been implemented to forecast exchange rates JPY/USD, GBP/USD, and CHF/USD with a fundamental approach by macroeconomic indicators inflation rate, interest rate, gross domestic product, trade balance, and income level. In order to examine the accuracy, applicability, and effectiveness of the method, both prosperity (2005-2007) and recession (2009-2011) years had been studied. Author also compared his results that had been obtained with Fuzzy Logic model with the results of ARCH, GARCH, and Artificial Neural Network models. Fuzzy Logic outperformed ARCH, GARCH and ANN with 2.5, 2.08, and 1.28 times smaller mean absolute percentage error respectively. Results are promising for further studies to combine fundamental approach with fuzzy logic model in order to predict with lower error term.

Weng et al. (2017) developed a hybrid machine-learning algorithm to forecast movements of next day's Apple Inc. stock prices and volume by analyzing publicly available knowledge. To begin with, authors investigated the literature on Knowledge based Artificial Intelligence (AI) algorithms on stock market prediction, and decided to use combination of three machine learning techniques, which are; Support Vector Machines (SVM), Neural Networks (NN), and Decision Trees (DT) to achieve higher prediction accuracy. Historical market data, technical indicators, web search traffic on targeted company, and related news from Google News have been collected for 37 months of period. Outcomes of this article were; hybrid machine learning algorithms can predict on higher accuracy rates 85% accuracy rate has been obtained, which is higher than pioneering SVM prediction performance of Kim (2003), online sources improve predictive accuracy by a margin, and used method has a significant advantage on traditional methods in uncertain scenarios.

Bollen et al. (2011) have studied behavioral economics with machine learning techniques. In order to predict Dow Jones Industrial Average (DJIA), authors examined public mood states derived from Twitter feeds. Tools such as 'OpinionFinder' and 'Google-Profile of Mood States (GPOMS)' have been used to measure the public mood by identifying the reactions on particular events. Fuzzy logic (FL) and Neural networks (NN) hybrid machine learning techniques have been used to investigate the correlation between public mood and DJIA price. Authors' model predicted daily changes on closing value with 86.7% accuracy, while Mean Average Percentage Error (MAPE) decreased by 6.6%.

Schumaker and Chen (2009) pioneered the predictive machine learning on finance with their approach on textual analysis. Authors illustrate the importance of breaking financial news on predicting stock market movements. Financial news and stock quotes covering S&P 500 index examined for five weeks to predict the stock price twenty minutes after the news article published. Authors noted a superior accuracy of support vector machine (SVM) technique compared to neural networks (NN). Article terms, publication time, and stock price have been trained to SVM. Their application had a dominating performance in metrics of closeness (0.04261), directional accuracy, and simulated trading returns.

On their research, He, Ulf, and Wang (2017) measured the effects of political instability on stock returns by measuring political tension. Historical events happening since 1995 had been observed alongside with public opinion surveys available since 2006 to measure the political tension and associate it with existing and expected political instability. Historical pricing data of approximately 2000 publicly listed Taiwanese and Chinese firms had also been analyzed to comprehend the effects of political climate on stock returns. Their results confirm that, events that elevates political tension cause strong adverse effects on stock returns. Moreover, expecting political tension in the future may also lead to substantial weakening on daily stock returns.

Yasir et al. (2019) proposed a hybrid machine learning technique to forecast daily foreign exchange rate by implementing event-based sentiments and heavily dependent, yet highly volatile variables. Three exchange rates from different regions; British Pound to US Dollar (GBP/USD), Hong Kong Dollar to US Dollar (HKD/USD), and Pak Rupee to US Dollar (PKR/USD) been selected to test the validity of the model. Crude oil and gold prices also have been selected as depended variables of foreign exchange rates. Daily data from April 2008 to January 2019 have been used. In terms of sentiment analysis, twitter posts regarding to local and global events, such as; Brexit, Hong Kong Protest (2014), Lahore Blast (2016), and US Election (2012) have been collected for each event on different durations, and analyzed. Due to poor performance of traditional econometrics methods, Support Vector Machines (SVM) and Deep Learning (DL) as machine learning techniques have been adopted for their proven superior performance on nonlinear patterns. Results of the study prove that the exchange rates are heavily dependent on sociopolitical issues, and have to be taken into account with a weighted margin in order to achieve higher accuracy on exchange rate prediction.

Aisen and Veiga (2013) examined the effects of political instability on economic growth. 169 countries have been observed in 5-year intervals between 1960 and 2004. Government instability/ regime instability indicates forthcoming cabinet changes, while cabinet changes means policy uncertainty for investors. That is why; expected negative coefficient would lead to greater uncertainty concerning future economic policies. For that reason, authors preferred to measure political instability by cabinet changes. The hypothesis of adverse impact of political instability on economic growth supported empirically through the study. Results claim that, additional untimely cabinet change diminishes annual real GDP growth rate by 2.39%.

Chuku et al. (2019) argued that, it is harder to predict the economic movements in developing countries due to uncertainty of political, financial, and economical policies. The idiosyncrasies generate unstable economy by sudden stops, reversals, turning points, and sudden big jumps that result as nonlinearity. Authors claimed the volatility and chaotic behavior of input variables from developing countries are the key challenge that needed to overcome for accurate prediction. Structural econometric model, ARIMA, and machine learning technique Neural Network have been compared to figure out the best performing model to predict economic indicators of developing countries that facing aforementioned scenarios. Interest rates, inflation rates, and volume of trade of Kenya, Nigeria, and South Africa have been observed quarterly from 1970 to 2016 (188 observations) to forecast their GDP. Results claimed that Neural Network model outperformed structural econometric model, and ARIMA in such a chaotic and nonlinear environment. Direct influence of political uncertainty, external and financial shocks on developing countries' economies have also been clearly observed.

Mehdian et al. (2008) investigated significance of Overreaction Hypothesis (OH) and Uncertain Information Hypothesis (UIH), which are the two extensions to Efficient Market Hypothesis (EMH) on investors' reaction to unexpected economic and political events in Turkey. Istanbul Stock Index National-100 (ISE-100), and Istanbul Stock Index All-Share (ISE-All-Share) have been investigated daily under 28 substantial economic and political events that distributed equally on both positive and negative impacts between 1997 and 2004. Cumulative abnormal returns (CAR) have been used to test any consistency between the investors' reaction and OH or UIH. No overreaction on ISE-100 had been determined due to arrival of unexpected information. However, significant correlation has been observed between arrival of unexpected information and UIH regarding to investors' behavior. Turkey has an uncertain economy and politics as its second nature. That's why, unexpected information triggers uncertainty and markets' prices the risk of new information.

Oliveira et al. (2017) examined micro blogging data to predict returns, volatility, and trading volume. Multiple Regression (MR) and Vector Auto-Regression (VAR) methods have been commonly used to predict financial based sentiments, and frequently for text sentiment (Oliveira et al., 2017). Authors aimed to improve the accuracy of prediction on nonlinear environment. In order to find out the best fitting methodology, Multiple Regression (MR), Support Vector Machine (SVM), Random Forest (RF), Neural Networks (NN), and Ensemble Averaging (EA) have been compared on predicting the values of S&P 500, RSL, DJIA, NASDAQ 100, RMRF, SMB, HML, MOM, VIX, PSize, and PInd by investigating microblogging sentiment indicator and posting volume. SVM outperformed other methods with a superior accuracy specially on predicting returns. Implementing textual data to prediction method found informative, and did perform well on predicting specially technology, energy, and telecommunication stocks.

Sharma et al. (2016) designed a sophisticated prediction model to overcome the nonlinearity and external dependencies of foreign exchange rate movements. Authors particularly emphasized the influence of economic conditions and the foreign policy of a country on the foreign exchange volatility, and their nonlinear relationship. Artificial Neural Network (ANN) – Fuzzy Logic hybrid machine learning technique had been selected due to its superior performance on nonlinear patterns compared to traditional techniques like ARIMA. Hybrid method also compared with ANN results for further understanding of accuracy rate of hybrid method. Research aimed to predict foreign exchange rates of three Asian countries; China, Japan, and India with respect to the US Dollar. 1502 samples had been observed between 4 January 2010 and 31 December 2015. Approximately 80% of the samples trained, while 20% of the samples tested to prevent any biased orientation. Hybrid machine learning method successfully managed to forecast foreign exchange rates, and also dominated ANN results with the lowest MAPE of 0.0868. Authors argue that, prediction may be improved by implementing technical indicators and the time delay on exchange rate data.

Cehreli et al. (2017) studied the influence of speculations and manipulations on the volatile USD/TRY exchange rate. 1250 daily data that is equal to 3.4 years have been observed to comprehend the impacts of manipulation and speculation on USD/TRY volatility. Adjusted interest rates by Central Banks, and inflation rate have been taken into account as well as manipulative demand-supply theory. Manipulative behaviors of the big investors, and speculative behaviors of individuals, and their formation process regarding to Central Bank's policy adjustments have been examined. Results prove that; panic environment and uncertainty in Turkey contributes serious exchange rate volatility, even more than huge manipulations.

Shapiro et al. (2020) compared the predictive accuracy of sentiment analysis tools, and investigate the response of macroeconomic variables to news sentiments by using machine learning techniques accordingly. Large set of economic sentiments that had been derived from 238,685 economic and financial newspaper articles between the years 1980 and 2015 analyzed with the existing lexicon, and also with a new lexicon built specifically by the authors'. Lexicon that has been developed by authors' specifically to capture the sentiment in aforementioned articles predicts with highest accuracy compared to existing lexicons. The authors have provided two economic research applications of sentiment-based analysis regarding to their findings. Accurate prediction of daily news sentiment has been obtained regarding to survey based consumer sentiment. Accurately predicting daily news sentiment leads the research to investigate the impulse response of macroeconomic variables. Impulse response of key macroeconomic variables (interest rate, inflation, industrial production, and consumption) to sentiment shocks has been analyzed. Results indicate increase in industrial production, consumption, and interest rates, but decrease in inflation along with 90% and 68% confidence band by the positive news sentiment shock.

Angeletos and Werning (2006) examined the impact of endogenous information during crisis periods by building upon the theory of Morris and Shin (1998, 2001, 2003) based on unique equilibrium currency crisis, multiple equilibrium macroeconomic modeling, and global coordination games theory accordingly. Authors' investigated endogenous public information, and exogenous private information and found a positive correlation. Therefore, coordinating the market may easily be obtained by endogenous public information by forecasting others' reactions. Moreover, crises situations cause nonfundamental volatility by increasing noise on financial markets.

Angeletos and La'O (2013) investigated the effects of market sentiments and and animal spirits on rational expectations, macroeconomic models, and unique equilibrium. To do so, trading between randomly selected multiple islands that specialized in producing diverse goods have been randomized and decentralized by restricting the communication, which caused idiosyncratic trading uncertainty. By aforementioned situation, the equilibrium of islands' economy can only be determined by fixed output and expectation, while endogenous variables are substantially responsive to any type of exogenous shocks. Expectations and output have been investigated during aggregate fluctuations under perfect and imperfect communication situation to comprehend the effects of sentiment shocks on key macroeconomic variables. Results indicates that, conflicts about the state of economy causes sentiment shocks that effects market expectations and economic outcomes, which creates unique equilibriums for certain situation. Shock may be spread out via communication and upsurge its effects on economy.

Barsky and Sims (2012) emphasized the importance of consumer confidence on predicting future fluctuations of macroeconomic variables. On their research, authors used Michigan Survey of Consumers to measure the impulse response of consumption and income on confidence, and found a significant and permanent correlation. On the other hand, Granger causality had not been determined between confidence and income or consumption. In the light of these findings, authors claimed that consumer confidence data is a reflection of associated information and variables provided.

Literature clearly identifies the impact of publicly available information on exchange rates and macroeconomic indicators. Furthermore, sentiment analysis plays a crucial role on understanding the initial response of the investors on the published news articles that specifies an uncertain and/or tense political and/or economic environment, which may cause a significant impact on the volatility of exchange rate and macroeconomic indicators. It has been shown that hybrid machine learning algorithms suits perfectly to extract the sentiment of investors and evaluate the subjective manners of correlation. Moreover, when compared to traditional prediction techniques, superiority of hybrid machine learning algorithms as a multidisciplinary prediction tool that performs analysis on combined quantitative and qualitative data has been emphasized, and scientifically proven.

Chapter 3
Outlook on the Geopolitical Background of Turkey

3.1 Introduction

Turkey is geographically located as one of the most unique and influential locations of the world. As Turkey is geographically located between Europe and Asia Continents, where the world’s current and rising politically and economically influential forces are allocated respectively, Turkey has a substantial geopolitical influence, which makes Turkey a significant geostrategic player in world affairs. The main considerations that touted Turkey as geostrategic player and geopolitical power are the ability to consolidate the straits between Europe and Asia Continents, as well as the Black Sea region (Cakar, 1998). Due to aforementioned reason, Turkey has always attracted the attention of world’s economically and politically superpowers throughout the history, especially United States, Europe, and Russia. Traditionally and geographically, Turkey has relied closer on European countries for trade and investment opportunities, Russia and Iran for energy imports, and the United States and NATO for defence cooperation (Congressional Research Service, 2020).

The main purpose of this chapter is to provide an insight on the background of geopolitical situation of Turkey, how and why Turkey matters to global superpowers, and how and why geopolitical significance, tension, and uncertainty impacts investors’ interpretation, thus Turkish economy.

In order to achieve this purpose, the chapter is going to interpret the rationale reasoning behind the Turkey's geopolitical importance for United States, Europe, and Middle East, and the economic impact of allocating on a geopolitically significant location respectively.

The chapter is structured as follows: Section 2 introduces Turkey's role on the global energy market, and how and why Turkey matters geographically and strategically, especially for European Union, Section 3 introduces NATO membership of Turkey, and why Turkey plays a significant role on the ongoing oil and natural gas wars on the Middle East, Section 4 briefly mentions the economic impacts of locating on a geopolitically strategic region by considering the political tension and uncertainty on perspective of investors, and Section 5 concludes the chapter by remarking the main objectives covered through the chapter, and briefly clarifies the differentiating elements of studying Turkey compared to other countries, and the importance of multidisciplinary approach in order to achieve broader perspective and understanding on the reasoning of economic fluctuations.

3.2 Turkey's Function on Global Energy Trade

The Middle East is home to a vast majority of depleted oil reserves, while Russia, Caspian region, Nigeria, and the Near East is accommodating the large portion of the natural gas reserves of the world (Karatas, 2010).

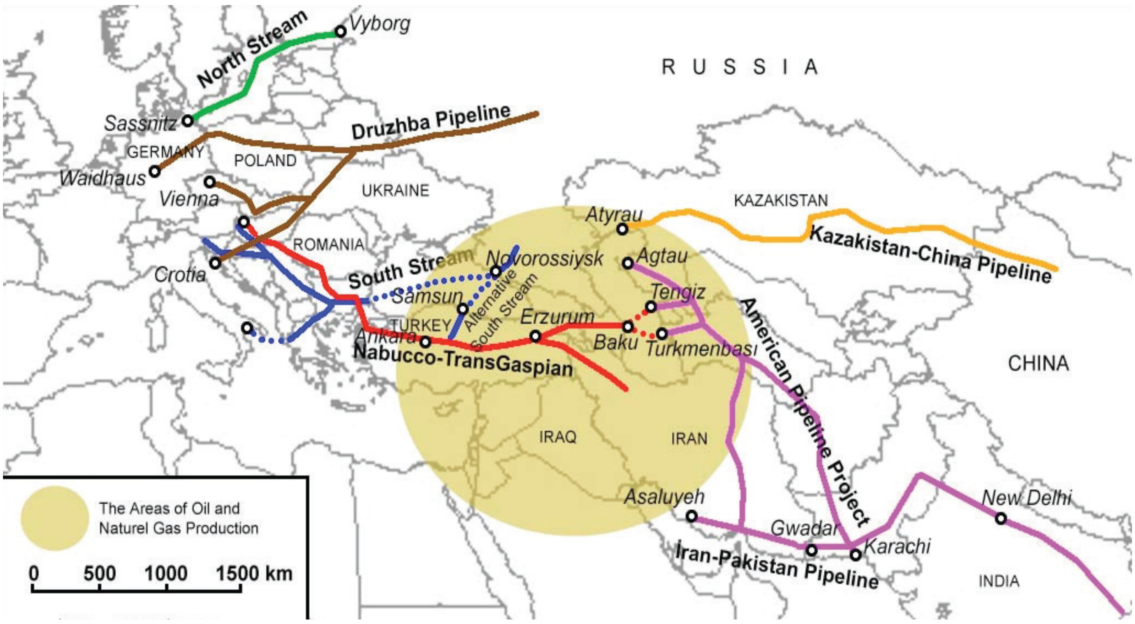
The European Union is the second leading energy consumer of the world, listing right after the United States. Throughout the history, significant portion of Europe's natural gas import has been supplied through the pipelines that are passing across the former Soviet Union countries (Akdemir, 2010).

Specifically OPEC and OPEC+ regions are the most economically practicable for European Union in terms of transportation to satisfy their energy needs, whereas the options for viable sources are reasonably inadequate. According to 2020 Eurostat database, extra-EU imports of natural gas and petroleum oil provides a significant portion of EU's energy demand. To be more precise, 25.5% of EU's petroleum oil demand has been supplied by Russia, 9.0% by

Kazakhstan, 8.5% Nigeria, and 6.6% by Saudi Arabia, while 43.4% of EU's natural gas demand has been supplied by Russia, and 12.0% by Algeria in 2020.

Geographically, Turkey borders Europe, as well as vast majority of the aforementioned countries and regions. For that reason, the significance of Turkey on supplying the energy sources into the European countries cannot be ignored, as Turkey can be used as a transit route for energy imports to European Union. Furthermore, Turkey has always been a country that interested and preferred country to 10 of natural gas producers with a share of 35.5% of natural gas reserves in global scale to be used as a transit country, which makes Turkey indispensable for European Union (Tekin and Walterova, 2007). Besides, Turkey is also undertaking a significant role, as a transit country on global oil market too. The natural gas and oil production, and gas pipeline structure between Europe and West Asia is presented in Figure 2.

Fig. 2. Oil and Natural Gas Production, and Gas Pipeline Structure between Europe and West Asia.



(Source: Akdemir, 2010)

The geopolitical significance of Turkey to European Union is clearly obvious, as Turkey accommodates both sea and overland transit ways for energy resources within its borders. However, having geopolitically significant advantage applies diverse pressures on Turkey as well. As managing Asia's oil and gas infrastructure can be interpreted as obtaining a huge power on the region, conflict of common interest on energy resources causing political pressure in addition to military pressure on the region, specially on the eastern borders of Turkey (Akdemir, 2010). Involvement of external forces into power conflict on the region transforms Asia into a shatter belt area, which triggers foreign and domestic pressures dramatically (Cohen, 2003).

3.3 Turkey as a NATO Member

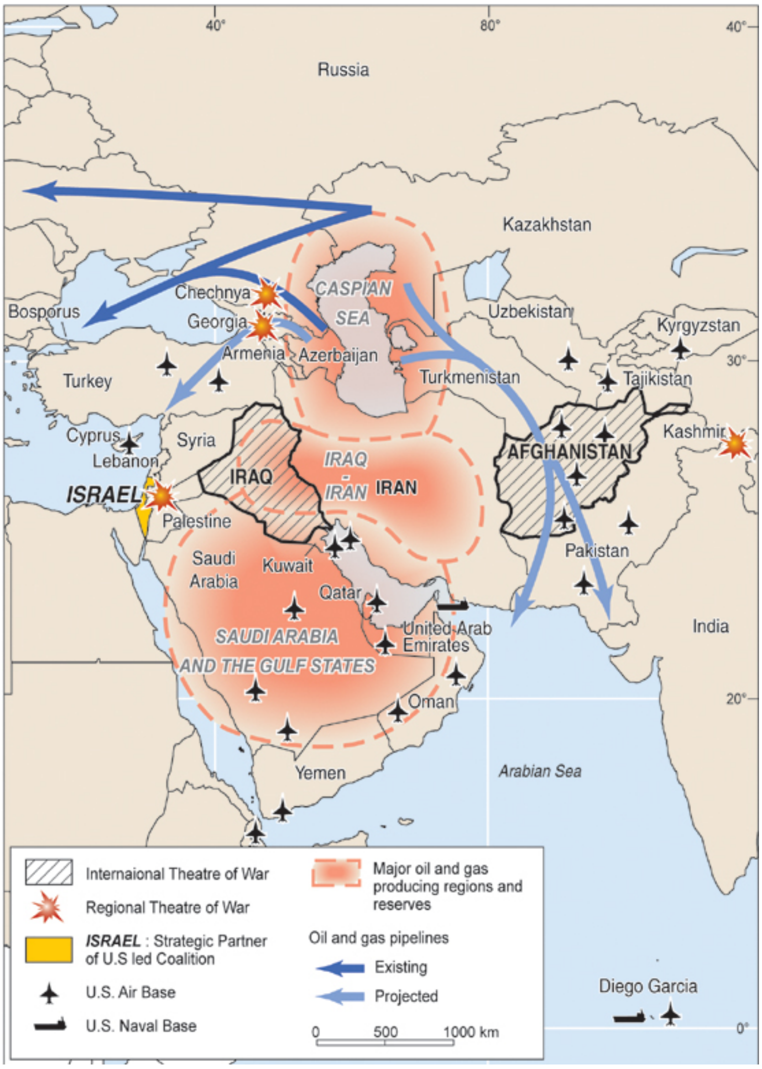
Turkey has become a NATO member in 18 February 1952. Turkish armed force, which is the second largest after United States' played a huge role on the endorsement of Turkey, as Turkish armed forces' contribution on sustaining the security and stability in the Euro-Atlantic region is substantially supportive for the NATO response forces. Moreover, Turkey's geographical proximity on numerous locations that attract global attention provides possibility and availability of stationing and transportation bases for United Sates and NATO on tense, geopolitically important, yet far regions (Congressional Research Service, 2020). Occurrence of United States and NATO Military in Turkey is presented in Figure 3.

Fig. 3. Occurrence of United States and NATO Military in Turkey

(Source: Congressional Research Service, 2020)

Considering the geographical distance between U.S. and Middle East, accessibility and usability of Turkey’s military bases to combat for the energy resources in that region is strategically vital, in addition of being cost effective for the U.S. military.

Fig. 4 United States' Middle East Policy, and Oil Wars in Recent History



(Source: American Foreign Policy, 2008)

Figure 4 demonstrates the international and regional theatre of war, U.S. military bases, major oil and gas producing regions, and existing and projected oil and gas pipelines. Figure 4 clearly illustrates that international and regional theater of war is taking place either on the borders or directly on the oil and natural gas producing soils. For that reason, accessibility and usability of allies' military bases is crucial for U.S. dominance on the battlefield, ruling the oil and natural gas production, and obtaining the monetary and regional power that has been delivered by the controlling the energy policy of the region, as well as supply and pricing strategy of energy resources.

3.4 Economic Impacts of Locating on a Geopolitically Significant Region

Ruling a region that's having crises consecutively may consider the location as strategically significant soil at first sight, however it brings political and economical complications along. Being admitted as a high-risk area undermines the economy of the country by being considered as an unstable economy, which caused by undergoing conflicts and possible threats (Bagci and Kardas, 2003).

The interaction concerning economics and the geopolitical reputation is exceedingly opposing. While rising geopolitical significance delivers power on the region, it also demands economical strength that may challenge country's capitals. When the geographical location of Turkey is considered, it may easily be seen that Turkey is surrounded by treats that are highly concerning security of the country, and causes instability, tension, and uncertainty frequently in both military and monetary manners. Geopolitical vectors of Turkey have been presented in Figure 5 below.

Fig. 5. Geopolitical Vectors of Turkey

(Source: Geopolitical Intelligence Services, 2018)

Turkey is persistently fighting against terrorism as a consequence of the geopolitical advantage on the region that the soil possesses, which costs enormous amounts in monetary terms by itself, while correspondingly frightening foreign investors to invest in the country, thus provide capital inflows. Immense costs of continuous fight against terrorism and lack of foreign investments triggering intense monetary crises, and capital shortage (Bagci and Kardas, 2003).

Economic obligations of Turkey that has been occurred by excessive monetary costs of fighting against terrorism restricts country’s liberal judgments and foreign policies (Bagci and Kardas, 2003). Therefore, investors hesitate to invest in Turkey as a result of rational expectations that has been formed under the hood of political, economical and regional tension, and uncertainty that Turkey and the region is in.

It may clearly be observed that Turkey’s currency and macroeconomic indicators are highly responsive on domestic and foreign political tension, and uncertainty throughout the history, which has believed to be triggered by the rational expectations of the domestic and foreign investors. Furthermore, earlier reports that have been presented during the United Nations Conference on Trade and Development revealed that foreign investors do not prefer to invest in Turkey due to instability of macroeconomic indicators, extensive corruption, and complicated nature of transactions (Turkish Probe, 2002), which caused a poor performance in foreign capital inflow and ranked Turkey as 122nd among 137 countries on the aforementioned report (Radikal, 2002). On the contrary side of flow of investments, when the investment outflows of Turkey have studied, it has clearly been observed that huge amount of Turkish capital is investing abroad, remarkably in Luxembourg and Switzerland. The outflow of capital from Turkey by domestic investors is also caused by the similar reasons with the scarcity of preference of foreign investors that have been mentioned above, as domestic investors lost their confidence on the domestic economy (Haberturk, 2001). When the interpreted, and expected causes and effects of political, economical, and regional tension and uncertainty have been considered, the results of studies, which have been mentioned above are not a coincidence.

3.5 Conclusion

This chapter has accentuated the geopolitical and strategic importance of Turkey. The chapter has briefly emphasized Turkey's strategic role on global energy trade, geopolitical obligations due to being a NATO member, complications on the region caused by ongoing oil wars, and their impact on Turkish economy.

Natural gas producers with a share of 35.5% of natural gas reserves in global scale is using Turkey as a transit country for natural gas trade, while also undertaking a transit country role on oil trade, which makes Turkey indispensable for European Union. Furthermore, international and regional theater of oil war is taking place on the borders of Turkey. As U.S. is the main actor on the oil war, accessibility and usability of NATO member Turkey's military bases is crucial in order to rule the region, thus the oil and natural gas production. Consequently, U.S. may obtain the monetary and regional power that has been delivered by the controlling the energy policy of the region, as well as supply and pricing strategy of energy resources.

Turkey is a significant influence for western border neighbor European Union for energy trade, besides eastern borders for ongoing oil wars in the Middle East. Therefore, a huge perpetual political and economical pressure has been applied by E.U. and U.S. that is targeting to undermine Turkey's empowerment, growth, hereby independence, thus interfere to Turkey's foreign policies to match their interests. Consequently, applied political pressure causes political tension, instability, and uncertainty alongside of classifying as a developing country impacts investors' rational expectations that prevent foreign and domestic investments to Turkey, which impacts Turkey's macroeconomic indicators, and currency negatively. When aforementioned components deeply observed, the causality between political tension, political uncertainty, and excessive economic volatility can be clarified in logical terms.

In order to clarify this claim scientifically, upcoming chapter is going to concentrate on particular political events that accommodate political tension and uncertainty between Turkey and U.S., and their impact on USD/TRY exchange rate fluctuations in daily manners by examining the complexion of selected political events individually to comprehend the responsiveness of exchange rate to varieties of political categories and grades of tension and certainty.

Chapter 4
Methodology

4.1 Introduction

The nature that surrounds us is not precise, certain, and systematic. Soft Computing is the discipline that endeavors to mimic the intelligence found in nature, to cope with uncertainty. In order to make decisions, human brain process millions of sensory data collected by series of experiences, generalizes awareness achieved, and recognizes the previous patterns (Kecman, 2001). However, natural sciences insist on linear principles for decades. Meanwhile, not a single element of life relies on linear principles. Meanwhile, even human brain process the information nonlinearly, thus decision makers response chaotic fluctuations and uncertainty of economics in nonlinear way (Wood and Grant, 2004). Conversely, most of econometric prediction methods that applied to economics and finance time series data prediction are designed to efficiently identify linear dynamics. Yet, complex structure of financial and economic data that relies on nonlinearity and underlying dependencies had been overlooked for decades.

Nowadays, arising expectations from sciences unveil the soft computing techniques that also study qualitative, informal, and approximate information besides the traditional quantitative, formal, and precise theories. The principal elements of soft computing are experimental data, which also known as statistical learning, and fuzzy logic methods. Neural Networks and Support

Vector Machines uses statistical learning to learn from feed data, while Fuzzy Logic methods entrenches human knowledge into an analytical model. Besides the difference on their learning and analyzing the data, all three core soft computing techniques directly address the result through nonlinear approximation and interpolation in high-dimensional space (Kecman, 2001).

The thesis has adopted a multidisciplinary methodology to meet the predetermined objectives of the study that has been mentioned in Chapter 1.4. While key elements of machine learning is going to construct the foundations of Artificial Intelligence algorithm, behavioral economics theories are going to support the algorithm by observing the psychological states of investors under varying pressures caused by complex environments, and examining their impact on investors' decision making process.

A sophisticated combination of soft computing methods, namely; natural language processing (NLP), fuzzy logic (FL), and supports vector machines (SVM) in addition to prospect theory (PT) of behavioral economics are going to be used in order to accomplish a realistic prediction by overcoming the real life scenarios, through simulating the real life decision maker's behaviors, which is the fundamental goal of the thesis.

This chapter is dedicated to exhaustively introduce the robust methodology that has been adopted for this thesis. The methodology is constructed as three phases, named as; Sentiment Analysis, Fuzzy Logic, and Support Vector Regression respectively. Following three sections of this chapter are dedicated to in depth introduction of basic principles of the methods that constructs the three phases of the methodology of empirical chapters 5 and 6, while weighted variations that are going to construct the phases of the methodology of empirical chapters 7 and 8 are going to be introduced in the final section.

This chapter is structured as follows: Section 2 introduces Lexicons, and illuminates Learning Automata, and Prospect Theory; Section 3 deeply introduces Fuzzy Logic; Section 4 introduces Support Vector Machines, and elucidates Support Vector Regression, and Kernel Function in detail; Section 5 is going to express the ways of splitting training and test data that is necessary for hyperparameter selection, kernel function selection, and investigating the consistency of the selected parameters. Cross validation methods, grid search methods, and the statistical tests that's going to be used to verify the performance of the mentioned methods; Section 6 presents statistical tests that are going to be used to measure the performance of the hybrid machine learning algorithm that has been constructed; and finally, Section 7 introduces the adjustments needed on the previously presented methodologies in order to meet the advanced requirements of the final chapters. To do so, weights need to be added, and transform the previously mentioned methodologies into; weighted prospect theory, weighted fuzzy logic, and feature weighted support vector regression.

4.2 Sentiment Analysis

Today, significant effects of publicly available political and financial information on financial markets and economics are broadly accepted. Alongside to objective and quantitative variables such as historical price data, subjective and qualitative information started to be used as a supplementary variable by recently developed analyzing tools.

Survey based analysis tools have been widely using since last decade. In order to estimate the economic expectations and behavior of the financial markets, judge the psychological situation of investors, and understanding the forecast changes, consumer sentiment plays a crucial role. University of Michigan Consumer Sentiment Index is one of the survey-based indexes that provide crucial subjective information to Bureau of Economic Analysis for further understanding of the direction of markets. While these methods enhance the forecasting abilities of the model (Souleles, 2004), limitations such as; cost, scope, sample size, low frequency, and time lag can't be ignored (Nyman et al., 2018).

To overcome aforementioned limitations, text sentiment analysis tool of natural language processing have been adopted lately to overcome economic policy uncertainty (Shapiro and Wilson, 2019). Text sentiment analysis is also known as opinion mining, or emotional artificial intelligence. This method aims to classify the subjective information or emotions regarding to texts' efficiency. To do so, libraries that classify the words according to their outcome of emotions are being used. These libraries are called Lexicons.

4.2.1 Lexicon

Lexicon based sentiment analysis technique uses domain specific or domain independent libraries to classify the words individually as either positive or negative with their semantic orientation value (Turney, 2002). In order to define the semantic orientation of the text, polarity value that defines the intensity measure for each word should be obtained. Text's polarity may simply be calculated by averaging the polarity value of the constituent words (Taboada et al., 2011), while length and/or the number of positive/negative words may be considered for more complex analysis (Gezici and Yanikoglu, 2018).

4.2.2 Learning Automata

Learning automata (LA) is a machine-learning (ML) model for natural language processing (NPL) that can adapt into changing and/or unknown environments. The main purpose of adopting the LA model is to build effective lexicons that analyze sentiments efficiently on stochastic environments. Learning automata is an adaptive decision making algorithm that aims to figure out optimal action in unknown and/or randomly fluctuating environments from finite set of actions (Pradigis et al., 2020). As it is capable to describe human behaviors, Learning automata has widely been accepted by disciplines that deal with highly uncertain of human behavior, such as psychology and biology due to its ability of classification. Purposely, LA exchanges information bidirectionally with the environment by picking precise action from a pool of possible actions to produce a matching output with environment's reaction. To do so, LA crosschecks its selection from the pool with the actual outcome repetitively until it finds the best possible match. Whereupon, LA unveils the response of the environment with the highest possible accuracy with which it interacts, with the enhancement provided by the customizability regarding to parameterization (Pradigis et al., 2020).

In order to create a unique lexicon with learning automata algorithm for sentiment analysis, annotated text should be feed into system, and word polarity function should be defined for each word. Polarity function defines the degree of positivity/negativity of each word in the formed lexicon. Higher polarity value indicates highly positive word, vice versa. Probability reinforcement scheme applied for each word recorded, which can be formulated as;

(1)

$$p_i(n+1) = p_i(n) - \big(1-\beta(n)\big)g_i\big(p(n)\big) + \beta(n)h_i\big(p(n)\big) \ \ if\ a(n) \neq a_i$$

$$p_i(n+1) = p_i(n) + (1 - \beta(n))\sum_{j\neq i} g_i\big(p(n)\big) - \beta(n)\sum_{j\neq i} h_j\big(p(n)\big) \ \ if\ a(n) = a_i$$

where, g_i and h_i denotes positive and negative annotation respectively for the nominated action a_i with received environmental response $\beta(n)$ at cycle n, normalized in the interval [0,1]. Favorable response can be achieved as the value of $\beta(n)$ decreases (Sarigiannidis et al., 2018). $p(i)$ denotes the score of each existing word in the training set, and word polarity vector $p(x)$ can be identified as;

(2)

$$p(i) = \frac{1}{m} \text{ for each word } i, 1 \leq i \leq m$$

where *m* denotes the total number of words appeared in the training set. Based on the aforementioned word polarity function, polarity algorithm that is going to applied for training can be formulated as;

$p(x) = p(x) + R/L \cdot p(x)$ for every word found in the sentence (3)

$p(x) = p(x) + S/(m - w) \cdot p(x)$ for every other word, not in the sentence (4)

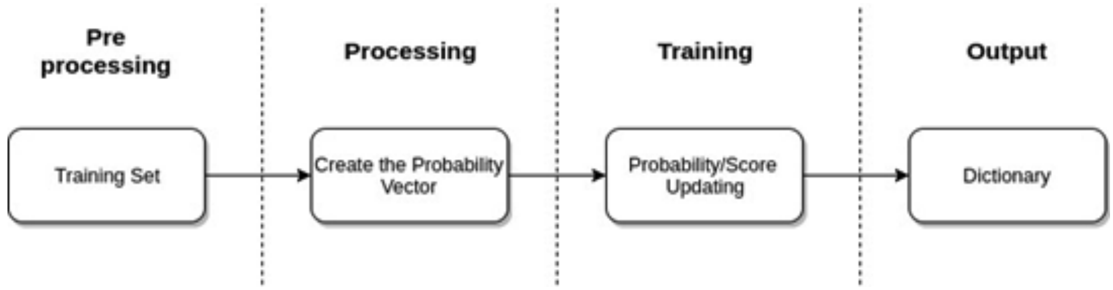
where *R*, *L,* S, and *w* denotes the score of the examined sentence, coverage speed of the automaton, the produced overall value of the sentence based on the polarity vector, and number of the words found in the particular sentence respectively. As parameter *S* states all the value received by the words appeared in the training set, it can be calculated by $S = S_1 + S_2 + \cdots + S_w$.

However, as it had been mentioned above, word polarity vector is only capable of expressing positive values. Polarity vector values needed to be normalized in order to achieve sentiment analysis in the range of [-1, 1]. Normalized polarity vector can be formulated as;

$$P_N(x) = \frac{p(x) - min}{max - min} \tag{5}$$

where the smallest value *(min)* is the most negative recorded polarity value, while the largest value *(max)* is the most positive recorded polarity value in the lexicon (Sarigiannidis et al., 2018). Workflow of the Learning Automata operation is presented below;

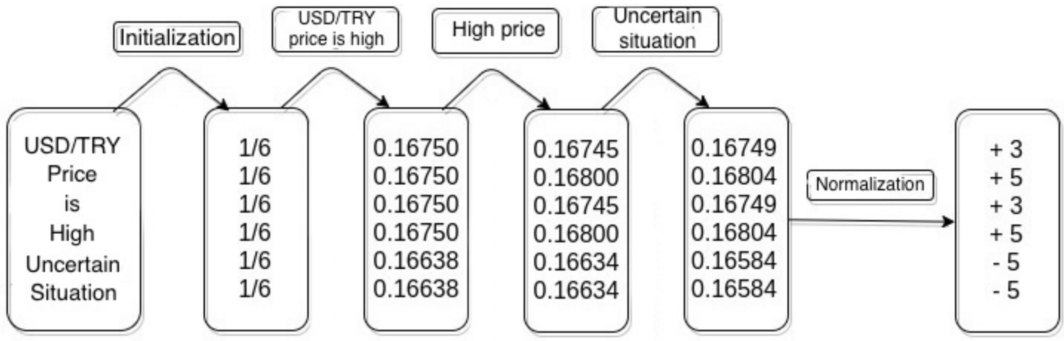
Fig. 6 Workflow of the Learning Automata



(Source: Sarigiannidis et al., 2018)

Operating mechanism of learning automata is numerically be exemplified below. To do so, three random sentences are going to form the annotated text, which contains; ‘USD/TRY price is high’ in the first sentence, ‘high price’ in the second, and ‘uncertain situation’ in third. As there are 6 words in total for annotated text, the world probability can be presented as; $p(USD/TRY) = p(price) = p(is) = p(high) = p(uncertain) = p(situation) = 1/6$. Each word, thus every sentence that take place in the annotated text will be scored as;

Fig. 7. Numerical exemplification Learning Automata operating mechanism



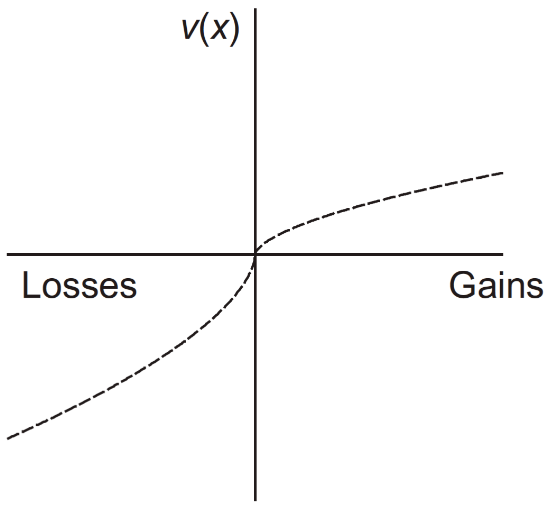
(Source: Sarigiannidis et al., (2018) modified by the author)

4.2.3 Prospect Theory

Kahneman and Tversky (1979) proposed Prospect theory in order to qualify individual’s actual decision behavior on risky situations. Since then, Prospect theory has been considered as the most popular theory of the field (Fan et al., 2013). Prospect theory has developed by addressing the inadequate rationalization of Expected Utility Theory (Neumann and Morgenstern, 1947) on Certainty Effect, Isolation Effect, and Reflection Effect (Gu et al., 2020). Prospect theory eliminates aforementioned issues by describing decision makers’ subjective outlook by weighting and value functions, and developed a novel decision model under uncertainty Gu et al., 2020).

Prospect theory replaced the expected utility theory's utility function *u(x)* that defines states of wealth with value function *v(x)* that represents gains and losses relative to the reference point with respect to psychophysics of diminishing sensitivity (Fox and Poldrack, 2009). Marginal impact diminishes as value gets farther away form the reference point. That's why; value function is represented as concave for gains, and convex for losses (Fan et al., 2013). Risk aversion of utility function contributed by the concavity of gains, while risk seeking contributed by the convexity of losses (Fox and Poldrack, 2009).

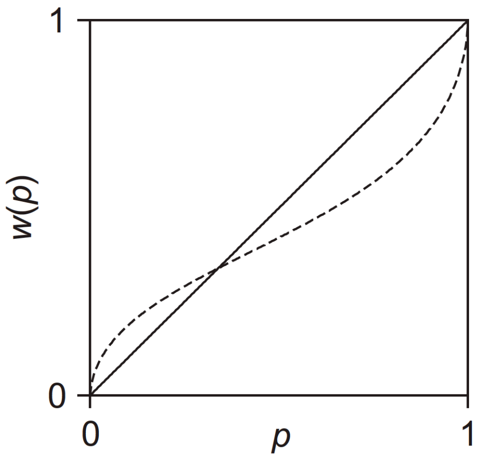
Fig. 8 Risk aversion of utility function



(Source: Kahneman and Tversky, 1991)

As Figure 8 illustrated, value function for losses is steeper than gains. This is also known as loss aversion. The reason behind this is; individual's tend to avoid risk when expecting profitable results, while prefer to take risk when expected result is negative, which results with a loss that has twice as much impact than a gain of the same rate (Kahneman and Tversky, 1979).

Fig. 9 Loss aversion of utility function



(Source: Kahneman and Tversky, 1991)

Prospect theory weighting function has been illustrated in Figure 9. Theory argues that, weight function *w(x)* denotes the impact of relevant probability on the valuation. As Figure 9 suggests, lower probabilities have higher chance to be over-weighted, while larger probabilities have lower chance to be under-weighted. Diminishing sensitivity on fluctuating probabilities has aimed to be captured by weight function (Fox and Poldrack, 2009).

Decision making process is distinguished into two phases, which are; framing and valuation. Likelihood of the possible acts of decision maker, and possible outcomes relevant to their choices composes the framing phase, while value assessment of the decision maker for each prospect composes the valuation phase (Tversky and Kahneman, 1986).

Prospect theory may be formulated as;

$$V = \sum_{i=1}^{k} (w(p_i)v(\Delta\chi_i)) \tag{6}$$

where *V* and *w(p)* denotes prospect value and probability weight function respectively, and $v(\Delta\chi)$ denotes the value derived from subjective feelings of the decision maker, where χ_0 is a certain reference point and $\Delta\chi_i$ is the deviation. Gains and losses may be determined by the $\Delta\chi_i$. While positive $\Delta\chi_i$ value indicates gains, negative value indicates losses (Liu et. al., 2011).

Kahneman and Tversky (1992) presented the power function of the value function as;

$$v(\chi) = \begin{cases} \chi^{\alpha}, & \chi \geq 0 \\ -\theta(-\chi)^{\beta}, & \chi < 0 \end{cases} \tag{7}$$

where α and β denotes concave-convex degree of gains and losses respectively regarding to the power function value of χ. Higher α and β values indicate that decision maker tends to risk, which is presented by θ. θ indicates higher possibility for the losses rather than gains, while $\theta > 1$ indicates the loss aversion (Liu et al., 2011).

Probability of weight function denotes the probability of the event outcome *p* regarding to the subjective judgment of the decision maker. Weight function represents the corresponding weight on the probability, and cannot be assessed as linear function of the probability due to its responsive nature to change of the probability. Probability weight function may be formulized as;

$$w^{+}(p) = \frac{p^{\gamma}}{\left(p^{\gamma} + (1-p)^{\gamma}\right)^{1/\gamma}} \tag{8}$$

$$w^{-}(p) = \frac{p^{\delta}}{\left(p^{\delta} + (1-p)^{\delta}\right)^{1/\delta}} \tag{9}$$

where $w^{+}(p)$ and $w^{-}(p)$ denotes the nonlinear weight function of the gains and losses respectively, while *p*, γ, *and* δ denotes probability of the event outcome, risk gain attitude coefficient, and risk loss attitude coefficient respectively (Liu et al., 2011). In other words, γ is the curvature parameter, while δ is the elevation parameter (Lattimore et al., 1992). Degree of curvature and elevation of the weighing function can be measured as $\gamma > 0$ and $\delta > 0$ respectively. As δ increase, risk aversion for losses increase while risk aversion for gains decrease (Fox and Poldrack, 2009). As the value of p denotes the probability of the event outcome, $w\,(p) > p$ indicates that the decision maker overvalued the event, while $w\,(p) < p$ indicates the undervaluation of the event (Liu et al., 2011).

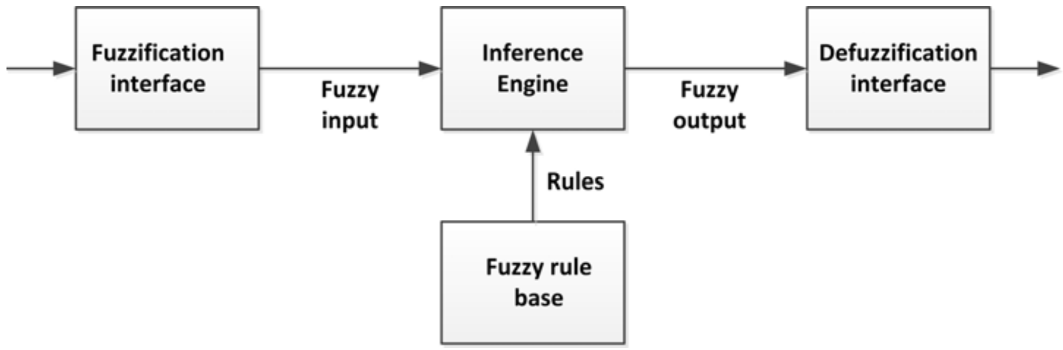
4.3 Fuzzy Logic

Human way of observing, examining, and understanding the world around them is not binary. There are many states between zeros and ones, which the meaning may vary depending on the user's experience and knowledge, just like Wittgenstein's (1953) argument on his book Philosophical Investigation. Language plays the key role on transfer of knowledge, while numbers only preferred to be used when words are not sufficiently precise. Human knowledge may be defined as fuzzy, and stated in imprecise terms without quantitative meaning in general. In order to comprehend humans' imprecise, deliberate, and uncertain thoughts that can only be expressed in linguistic terms, imitating the non-quantitative, approximate, dispositional, and linguistic way of human reasoning is necessary. This ideology is the foundations of Fuzzy Logic.

Fuzzy logic had been introduced by Lofti Zadeh (1965) as an instrument that transforms linguistically articulated information into practicable mathematical algorithms (Kecman, 2001). Fuzzy logic is a theory that supports the elasticity and matter of degree (Zahed, 1973) that aimed to deal with complexity of real life scenarios, rather than classical methods based on bivalent logic and probability theory (Kecman, 2001). The perception and understanding of an obviously unsophisticated concept is not exactly that simple no matter how unsophisticated it may seem. The possibility of analyzing a concept that should be examined in both qualitative and quantitative ways has been provided by using fuzzy set theory that tolerate transitional grades of membership.

Fuzzy logic system consists of four main components, which are; rule base, fuzzification, interference engine, and defuzzification. Architecture of fuzzy logic system presented below.

Fig. 10 Workflow of Fuzzy Logic

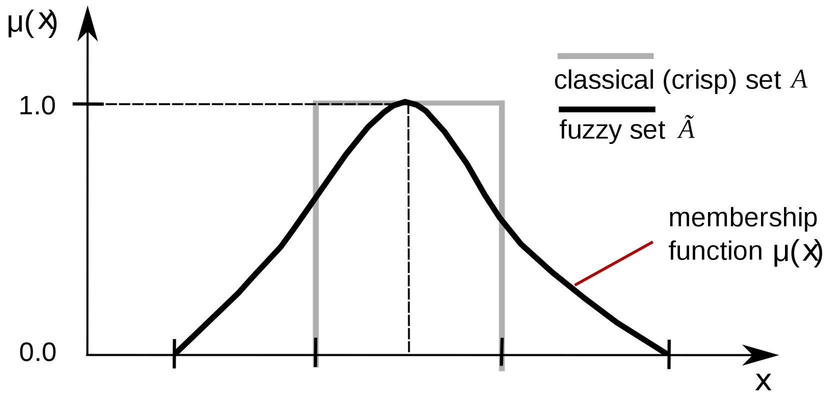


(Source: Nilashi M., et al., 2011)

Fuzzification process takes place to obtain fuzzy sets from crisp variables. Membership function has been using here to transform the crisp value into a corresponding linguistic fuzzy variable. Interference engine assign the correlation between fuzzy input and the rules, which obtained by IF-THEN conditions proposed to mimics the human decision making process. Obtained fuzzy sets gets into reverse fuzzification process (Cheng and Roy, 2011) to be converted into crisp output that is named defuzzification (Klir and Yuan, 1995).

The fuzziness can be characterized by the membership function. Membership function established the correlation between quantitative variable values and qualitative linguistic variables that take place in IF-THEN rules. Imprecise and subjective statements define classes of membership. Membership function associated with fuzzy sets may be described by $\mu(x) \rightarrow [0,1]$.

Fig. 11 Classical (crisp) set vs Fuzzy set

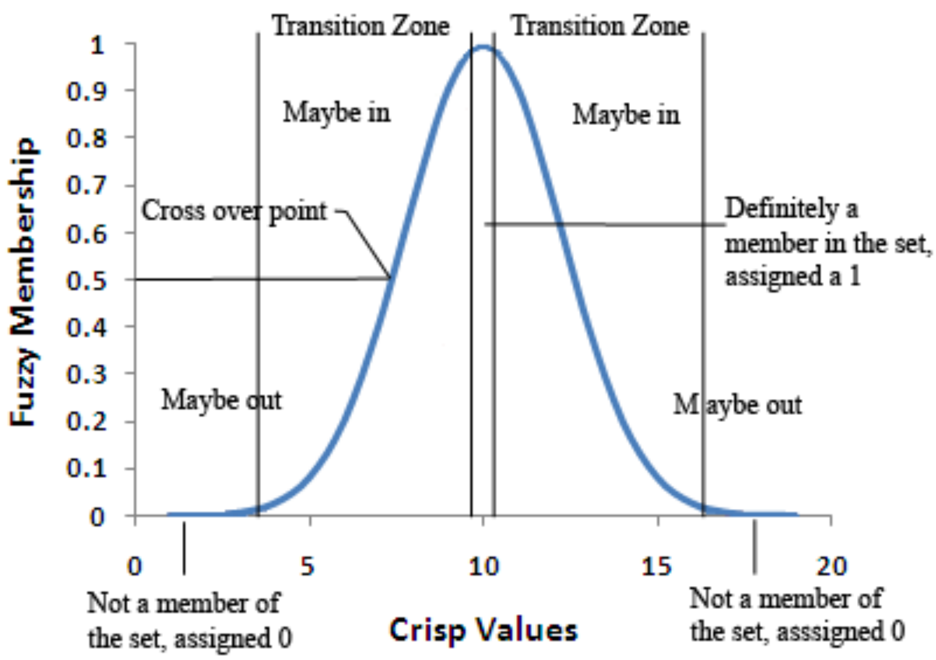


(Source: Christodoulou S., et al., 2012)

Model aims to mimic human decision-making route by utilizing approximate reasoning logic (Zahed, 1965). IF-THEN rules forms linguistic form of knowledge that can practically express the structured human knowledge entirely. That is why, IF-THEN rules play an crucial role on Fuzzy Logic applications (Zahed, 1973).

The main purpose of fuzzy logic is to model human knowledge, or human understanding and concepts about the world. Unlike the Boolean algebra, which values the truth-value of a variable as either 0 or 1, fuzzy logic values the truth with an infinite number of possible numbers between 0 and 1. This provides much broader spectrum of statement opportunity on vague states, concepts, and situations, rather than defining it as utterly false, or absolutely true (0-1). Degree of membership of element x in set A represented by the value between 0 and 1. The transition from x to $\mu_A(x)$ is known as fuzzification.

Fig. 12 Illustration of Fuzzy Membership on Vague States



(Source: Nasehi S., et al., 2017)

In this case, membership to a fuzzy set is highly subjective and problem-dependent (Kecman, 2001). Membership of a fuzzy set may be given by the function;

Partial membership is acceptable for a fuzzy set, (degrees of truth)

$$\mu_A(x): X \longrightarrow [0,1] \tag{10}$$

where degree of membership denoted as μ, membership function for x in fuzzy set *A* denoted as $\mu_A(x)$, while universal set denoted as *X* (Yagar and Zahed 1992).

For this study, three membership functions have been selected due to their compatible nature, namely; Sigmoidal membership function, Trapezoidal membership function, and Gaussian membership function. These membership functions are presented below;

A sigmoidal MF is inherently open right or left & thus, it is appropriate for representing concepts such as “very large” or “very negative”.

Right shoulder sigmoidal function; (11)

$$\mu = \frac{1}{1 + e^{-\beta(x-\alpha)}}$$

Left shoulder sigmoidal function;

(12)

$$\mu = \frac{1}{1 + e^{\beta(x-\alpha)}}$$

where β controls the steepness of sigmoid, and α denotes the point of crossover point (Robinson, 2003). Right shoulder and left shoulder of sigmoidal membership function has been illustrated in Figure 13 and Figure 14 respectively.

Fig.13 Right Shoulder of Sigmoidal Function

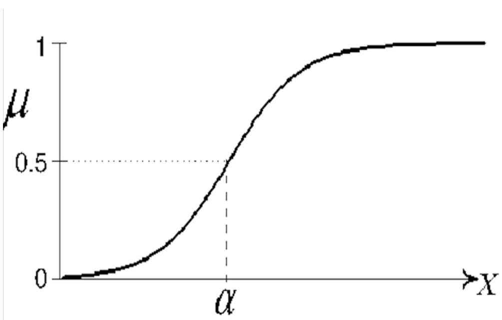
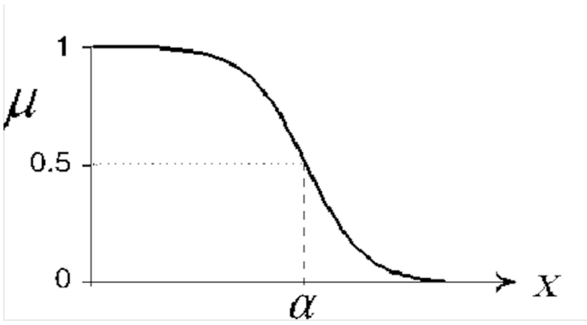


Fig.14 Left Shoulder of Sigmoidal Function.



(Source: Robinson V., 2003)

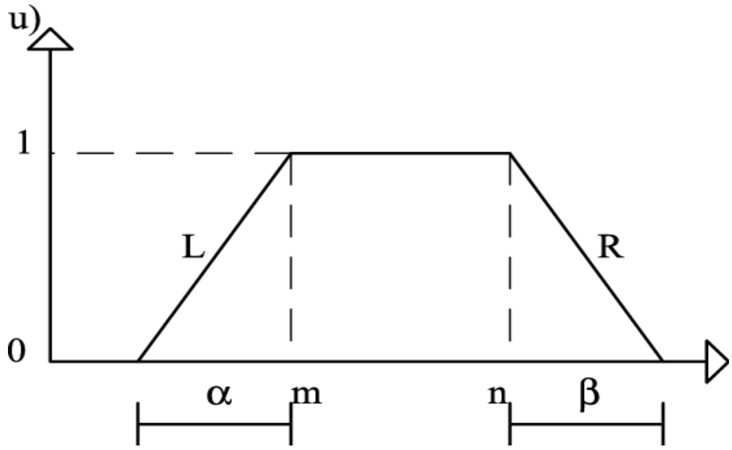
Trapezoid membership function can be formulated as;

(13)

$$\mu(x) = \begin{cases} 0 & if\ x < a \\ \frac{x-a}{m-a} & if\ x \in [a, m] \\ 1 & if\ x \in [m, n] \\ \frac{b-x}{b-n} & if\ x \in [n, b] \\ 0 & if\ x > b \end{cases}$$

where upper and lower bounds, and center of triangle denoted as *a*, *b*, and *c* respectively, while coordinates of tolerance denoted as *m* and *n*. Trapezoid membership function has been illustrated in Figure 15.

Fig. 15 Illustration of Trapezoidal Membership Function



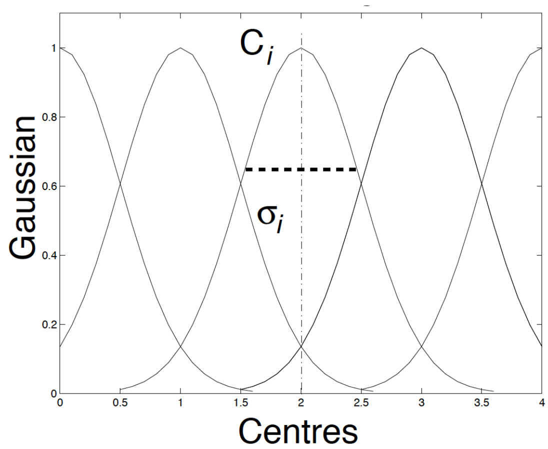
(Source: Kecman, 2001)

Gaussian membership function can be formulated as;

$$\mu_{S^i}(x) = \exp\left(-\frac{(c_i - x)^2}{2\sigma_i^2}\right) \tag{14}$$

where c_i and σ represents the center and width respectively of the ith fuzzy set S^i. Change in c_i and σ shapes membership function as distinct form. Gaussian membership function has been illustrated in Figure 16.

Fig. 16 Illustration of Gaussian Membership Function



(Source: Boostan E., et al., 2015)

Unlike crisp sets, it is possible to be a member of a fuzzy set up to a degree. To do so, fuzzy set operation is necessary. Complement, intersection, and union are the most important fuzzy set operations. Complement operator is theoretically regarded as operator NOT;

$$\mu_B(x) = 1 - \mu_A(x) \tag{15}$$

Intersection operator is theoretically regarded as operator MIN;

$$MIN\ (\mu_A(x), \mu_B(x)) \tag{16}$$

which interpreted as logical AND, while Union operator theoretically regarded as operator MAX;

$$MAX\ (\mu_A(x), \mu_B(x) \tag{17}$$

which interpreted as logical OR. Lower $\mu_{A \wedge B} = MIN\ (\mu_A, \mu_B)$ and upper $\mu_{A \vee B} = MAX\ (\mu_A, \mu_B)$ intersection points of membership functions may be represented accordingly.

The connections between different sets expressed as fuzzy relations. While the presence or absence of association, interaction, or interconnectedness between the elements of two or more sets represents the crisp relation (Klir and Folger, 1988), various degrees or strengths of relations between elements forms fuzzy relations (Kecman, 2001). As relation of the element to itself is a set, all operations may be applied to it without any adjustment required. The set of all ordered pairs

$$X \times Y = \{(x, y) \mid x \in X\ and\ y \in Y\} \tag{18}$$

defined as Cartesian products or product sets by (Kecman, 2001). Cartesian product is the simplification of n-tuples (relation matrix), characterized by a function

(19)

$$\mu_R : X_1 \times \cdots \times X_m \rightarrow [0, 1]$$

where μ_R, X_i and $X_1 \times \cdots \times X_m$ denotes the membership function of a multidimensional fuzzy set, universes of discourse, and the product space respectively (Robinson, 2003). Cartesian product $X \times Y$ can be defined as $R\ (X, Y)$ or simply *R*, where *R* represents the fuzzy relation between two sets (Kecman, 2001).

Composition can be obtained by combining different product spaces of fuzzy relations with each other. Fuzzy set that had been acquired by the composition is also fuzzy due to fuzzy relations of combined sets. The MAX-MIN composition of two fuzzy relations

(20)

$$R_1(x, y), (x, y) \in X \times Y$$

and

$$R_2(y, z), (y, z) \in Y \times Z$$

with a membership function of $\mu_{R_1 \circ R_2}$ may be represented as;

(21)

$$R_1 \circ R_2(x, z) = \left\{ \left[(x, z), \begin{matrix} max \\ y \end{matrix} \left\{ min \left\{ \mu_{R_1}(x, y), \mu_{R_2}(y, z) \right\} \right\} \right] \middle| \, x \in X, y \in Y, z \in Z \right\}$$

Interpretation of the resulting relation matrix *R* in linguistic way may be obtained by IF-THEN rules. IF-THEN rules are widely using to model the structured human knowledge (Kecman, 2001).

R: IF (fuzzy criteria) THEN (fuzzy conclusion)

Casual relation between measurements that is correlated with a membership function and control values may be expressed by aforementioned fuzzy rule

(Robinson, 2003). Relational matrix *R* needed to be calculated to perform fuzzy inference, while *R* can only be calculated after IF-THEN rules transformed into relational matrices (Kecman, 2001).

Clarification of the sets of rule-based values in the input vector to an output vector is called fuzzy inference. Most widely used fuzzy inference methods to calculate the relational matrix are Mamdani and Larsen methods.

Mamdani method uses the minimum operator as fuzzy implication and the max-min operator for the composition, while Larsen method uses the product operation as a fuzzy implication and the max-product operator for the composition.

Mamdani method is formulated as;

$$\mu_{A\Rightarrow B}(x,y) = MIN(\mu_A(x), \mu_B(y)) \tag{22}$$

Larsen method is formulated as;

$$\mu_{A\Rightarrow B}(x,y) = \mu_A(x)\mu_B(y) \tag{23}$$

where input variables $\mu_A(x)\ and\ \mu_B(y)$ varies between 0 and 1. The main objective of this method is to transform qualitative knowledge into the form of IF-THEN rules by using membership function for linguistic variables *low, medium,* and *high*.

$$R_1\text{: IF } x_1 = low \text{ AND } x_2 = medium \text{ THEN } y = high \tag{24}$$

$$\mu_R(x_1, x_2, y) = MIN(\mu_L(x_1), \mu_M(x_2)\mu_H(y))$$

MIN operator should be used while calculating rule R_1 as the association between antecedents *small* and *medium* be made by AND operator.

$$H = MIN(\mu_L(x_1'), \mu_M(x_2')) \tag{25}$$

where *H* denotes fuzzy inference and x' denotes fuzzified crisp value or transformation into a membership vector $\mu_{L'}$.

At this point, defuzzification method should be applied in order to transfer fuzzy value of *y* into a crisp value.

Defuzzification can be formulated as;

$$y' = \frac{\sum_{i=1}^{N} y_i \, H_i}{\sum_{i=1}^{N} H_i} \tag{26}$$

where *N* denotes the number of membership functions.

4.4 Support Vector Machines

Support Vector Machines (SVM) is a novel approach developed by Vapnik (1995) to solve non-linear pattern recognition (classification) and functional approximation (regression) problems in time series analysis (Grigoryan, 2016). Unlike empirical risk minimization based traditional neutral network models, SVM is based on the structural risk minimization principle that provides ability to estimate a function by minimizing an upper bound of generalization error (Vapnik, 1998). SVM has been successfully used for time series prediction due to its higher generalization performance and testing accuracy (Grigoryan, 2016).

Support Vector Machines transforms the nonlinear input into higher dimensional feature space when the input cannot be separated linearly. The transformation constructs linear model in the new space, which signifies the nonlinear decision boundary of the original space. Subsequently, the linear optimal separating hyperplane featured in new space (Kim, 2003). The algorithm that composes the optimum location for the decision boundaries is called support vector machines.

4.4.1 Support Vector Regression

Support vector machines had been improved by Drucker et al. (1996) to overcome prediction problems as well. This method is called support vector regression. The main objective of support vector regression is to minimize the prediction error and the risk of over-fitting by using ε-intensive zones (Huang, 2012). Support vector regression can be formulated as;

$$f(x) = w^T \phi(x) + b \tag{27}$$

where $f(x)$, *w*, $\phi_i(x)$, and *b* denotes function of support vectors, weight parameters that determine the hyperplane, mapping function, and bias term respectively.

Vapnik introduced ε-intensive loss function (ε-tube) to measure the margin between an optimal separating hyperplane and support vectors. Loss function with ε-intensivity zones formulated as;

$$|y - f(x,w)|_\varepsilon = \begin{cases} 0 & if\ |y - f(x,w)| \leq \varepsilon \\ |y - f(x,w)| - \varepsilon & otherwise \end{cases} \tag{28}$$

where, $f(x,w)$, and ε denotes function of the weights *w* that are the subjects of learning, and margin of tolerance respectively. The loss is zero if the predicted value is within the ε-deviation. However, for all other predicted points remained outside of the ε-tube, kernels, which will be presented later on should be applied to construct a nonlinear regression hyper surface. In that case, new empirical risk R_{emp} need to be introduced, which can be formulated as;

$$R_{emp}^{\varepsilon}(w,b) = \frac{1}{l}\sum_{i=1}^{l} |y_i - w^T x_i - b|_\varepsilon \tag{29}$$

As the main objective of support vector regression is to minimize the empirical risk R_{emp}^{ε} and the width and flatness of ε-tube $\|w\|^2$ simultaneously, risk function *R* should be minimized.

$$R = \frac{1}{2}\|w\|^2 + C\left(\sum_{i=1}^{l}|y_i - f(x_i, w)|_{\varepsilon}\right) \tag{30}$$

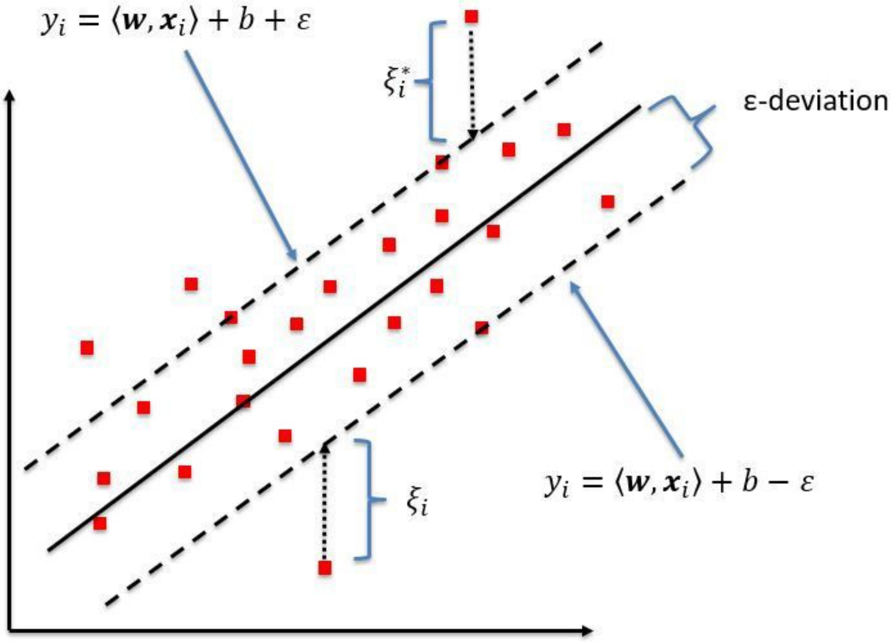
However, if the training data cannot be separated linearly, slack variables ξ and ξ^* that measures above and below of the ε-tube are introduced to tackle with unfeasible constraints of the optimization problem (Grigoryan, 2016). In that case, risk minimization (optimization) problem can be reformulated as;

$$R_{w,\xi,\xi^*} = \left[\frac{1}{2}\|w\|^2 + C\left(\sum_{i=1}^{l}\xi_i + \sum_{i=1}^{l}\xi_i^*\right)\right] \tag{31}$$

$$Subject\ to \begin{cases} y_i - w^T x_i - b \leq \varepsilon + \xi_i \\ w^T x_i + b - y_i \leq \varepsilon + \xi_i^* \\ \xi_i \geq 0 \\ \xi_i^* \geq 0 \end{cases}$$

where *l* denotes the number of training data, and *C* is the unit of regularization that controls the trade-off between the empirical risk and misclassification. The user will select parameter *C* with trade-off between an approximation error and the weights vector norm $\|w\|$. Increase in *C* value will larger errors ξ and ξ^* while decreasing the approximation error, vice versa. On the other hand, generalization performance of the model will be negatively affected in that scenario due to increase in weights vector norm $\|w\|$ (Kecman, 2001).
In order to overcome optimization problem, another design parameter should be chosen to define the size of the ε-tube. For further consideration, implementation of ε-tubes has been illustrated in Figure 17.

Fig. 17 Visual Demonstration of Support Vector Regression and ε-Tubes



(Source: Kleynhans et al., 2017)

Lagrange multipliers α_i and α_i^* are the unique minimum and maximum values of the function equivalent to ξ and ξ^* that are above and below of the ε-tube. Lagrange function can be formulated as;

(32)

$L_p(w, \xi, \xi^*, \alpha_i, \alpha_i^*, \beta_i, \beta_i^*)$:

$$L = \frac{1}{2}w^T w + C\left(\sum_{i=1}^{l} \xi_i + \sum_{i=1}^{l} \xi_i^*\right) - \sum_{i=1}^{l} \alpha_i^*[y_i - w^T x_i - b + \varepsilon + \xi_i^*]$$
$$- \sum_{i=1}^{l} \alpha_i[w^T x_i + b - y_i + \varepsilon + \xi_i^*] - \sum_{i=1}^{l}(\beta_i^* \xi_i^* + \beta_i \xi_i)$$

Primal Lagrangian variables $L_p(w, \xi, \xi^*, \alpha_i, \alpha_i^*, \beta_i, \beta_i^*)$ has to be minimized regarding to primal variables $w, b, \xi,$ and ξ^* and maximized regarding to non-negative Lagrange multipliers $\alpha_i, \alpha_i^*, \beta_i,$ and β_i^*. Minimization and maximization of Lagrangian variables has been illustrated in Figure 18.

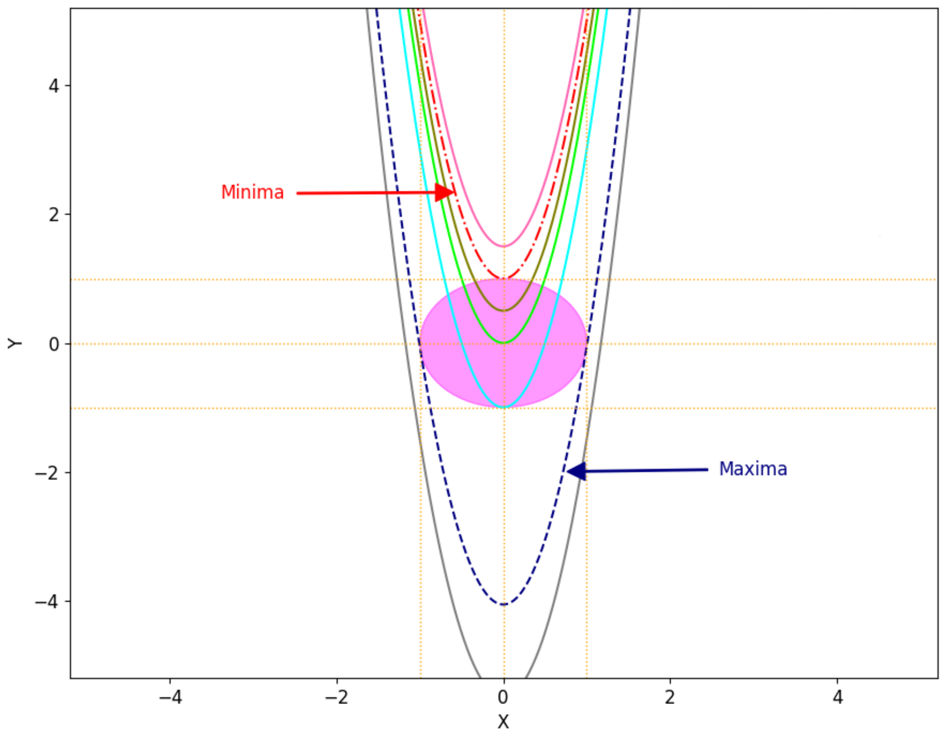
Lagrange multipliers $L_d(\alpha, \alpha^*)$ can be maximized by applying Karush-Kuhn-Tucker (KKT) conditions for regression, which may be formulized as;

(33)

$$L_d(\alpha, \alpha^*) = -\varepsilon \sum_{i=1}^{l} (\alpha_i^* + \alpha_i) + \sum_{i=1}^{l} (\alpha_i^* - \alpha_i) y_i - \frac{1}{2} \sum_{i,j=1}^{l} (\alpha_i^* - \alpha_i)(\alpha_j^* + \alpha_j) x_i^T x_j$$

$$Subject\ to \begin{cases} \sum_{i=1}^{l} \alpha_i^* = \sum_{i=1}^{l} \alpha_i \\ 0 \le \alpha_i^* \le C \\ 0 \le \alpha_i \le C \end{cases}$$

Fig. 18 Visual Demonstration of Lagrange Multiplier Method



(Source: Bhattacharyya, S., (2018) Support Vector Machine: Complete Theory. Towardsdatascience)

As Lagrange multipliers $L_d(\alpha, \alpha^*)$ obtained, optimal desired weights vector *w* of the regression hyperplane can be solved, where *w* may be obtained by;

(34)

$$w = \sum_{i=1}^{l} (\alpha_i^* - \alpha_i) x_i$$

while optimal bias *b* of the regression hyperplane may be found as;

(35)

$$b = \frac{1}{l}\left(\sum_{i=1}^{l} y_i - x_i^T w\right)$$

and the best linear regression hyperplane may be obtained by;

(36)

$$z = f(x, w) = w^T x + b$$

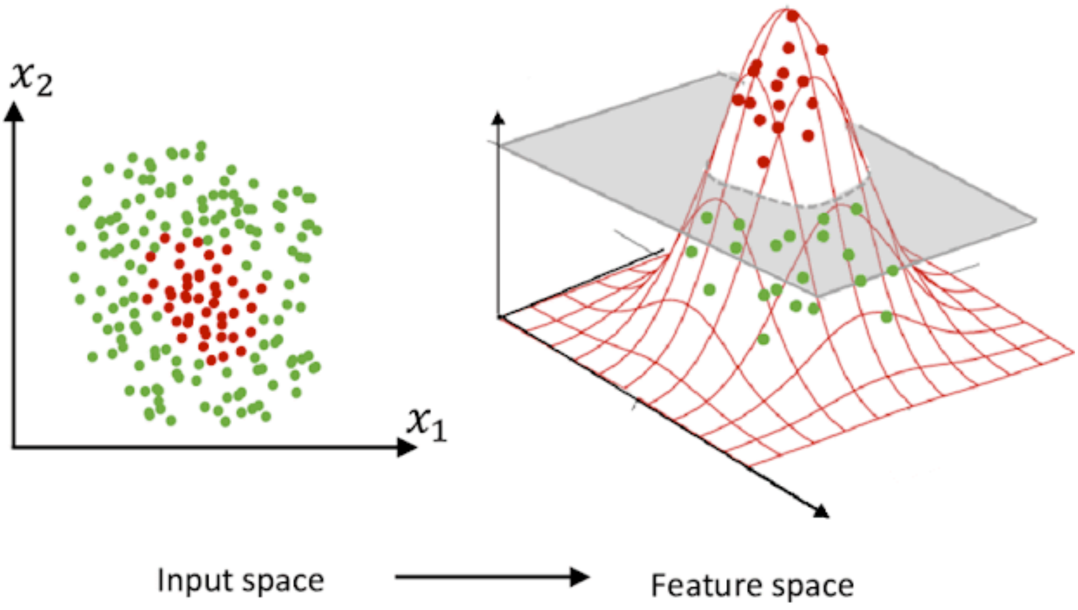
4.4.1.1 Kernel Function

In order to solve nonlinear regression, kernel function should be applied. Nonlinear input data can be separated linearly by mapping $\phi(x_i)$ the inputs into the high-dimensional feature space by using kernel function (Kim, 2003). Function $K(x_i, x_j)$ defines kernel function, where the value of kernel function is equal to the inner products x_i and x_j in the feature space $\phi(x_i)$ and (x_j) (Grigoryan, 2016). Application of kernel function in order to map nonlinear input data into hyperspace to be able to separate the data linearly has been illustrated in Figure 19. Nonlinear relationship between input and output data may be formulated by theoretically mapping the input data in the high dimensional feature space as linear function (Zhang and Li, 2012).

(37)

$$K(x, y) = \sum_{i=1}^{\infty} \lambda_i \phi_i(x) \phi_i(y)$$

Fig. 19 Mapping Inputs into High-Dimensional Feature Space by Using Kernel Function



(Source: Jenis Algoritma Pada Machine Learning, 2019)

Diverse kernels may be adopted to generate inner products to construct support vectors machines with nonlinear decision surface in the input space (Huang, 2012). Most commonly used kernel functions are Gaussian and Polynomial kernel functions.

(38)

$$K_{gaussian} = e^{-1/\sigma^2[(x-x_i)^T \Sigma^{-1}(x-x_i)]}$$

(39)

$$K_{polynomial} = [(x^T x_i) + 1]^d$$

where d and σ^2 denotes the degree and the bandwidth of the kernel respectively. After selecting the kernel function, nonlinear regression hyperplane may be obtained by;

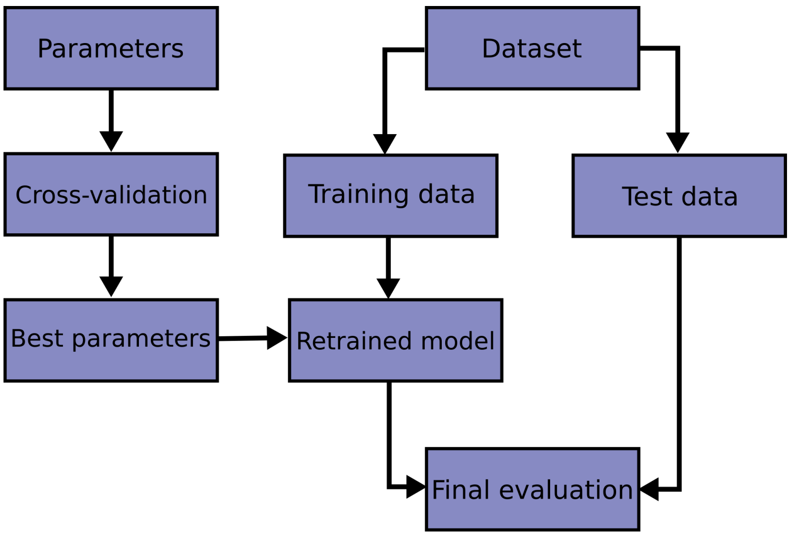
(40)

$$f(x) = \sum_{i=1}^{l} (\alpha_i^* - \alpha_i)\, K(x_i, x) + b$$

4.5 Splitting Training and Test Data

Alongside the appropriate data selection, splitting data from the presented data pool into training and test groups plays a vital role on successful regression analysis. Splitting data into training and test groups is essential to appraise the performance of the regression model. In real life scenarios where only finite number of sample data exists, most stable and lowest rate of error can only be achieved by validation approaches. This approach is articulated in literature as Cross Validation. The accommodation of cross validation alongside to training and test data separation is presented below;

Fig. 20 Accommodation of Cross Validation and Data Separation on Methodological Workflow

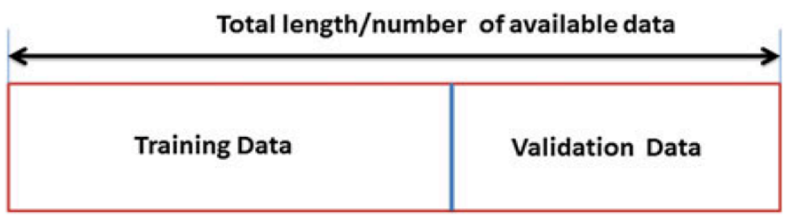


(Source: Scikit-learn, Cross-Validation: Evaluating Estimator Performance)

4.5.1 Cross Validation

Cross validation (CV) is a method for dividing data sets into test and training groups to evaluate the accuracy of the predictive machine learning model on unseen data. The main intention of cross validation methods is to achieve the best model fit by finding the equilibrium between underfitting and overfitting. To do so, it is necessary to maximize training data to achieve better learning, while maximizing test data to achieve accurate validation. The distinguishing feature of the CV technique compared to traditional methods is; While traditional methods are reducing the count of data points by using separate validation set and perform a single training and validation run, cross validation methods maximizes the count for both sets by splitting the dataset into training and tests sets, and perform of training and validation runs for several times (experiments), while using different portions of the dataset for each run to determine the performance of the model overall by averaging the attained scores. Traditional way of separating training set and validation set is visually presented as Figure 21 below (Remesan and Mathew, 2014).

Fig. 21 Traditional Way of Separating Training and Validation Groups



(Source: Dubitzky et al., 2007)

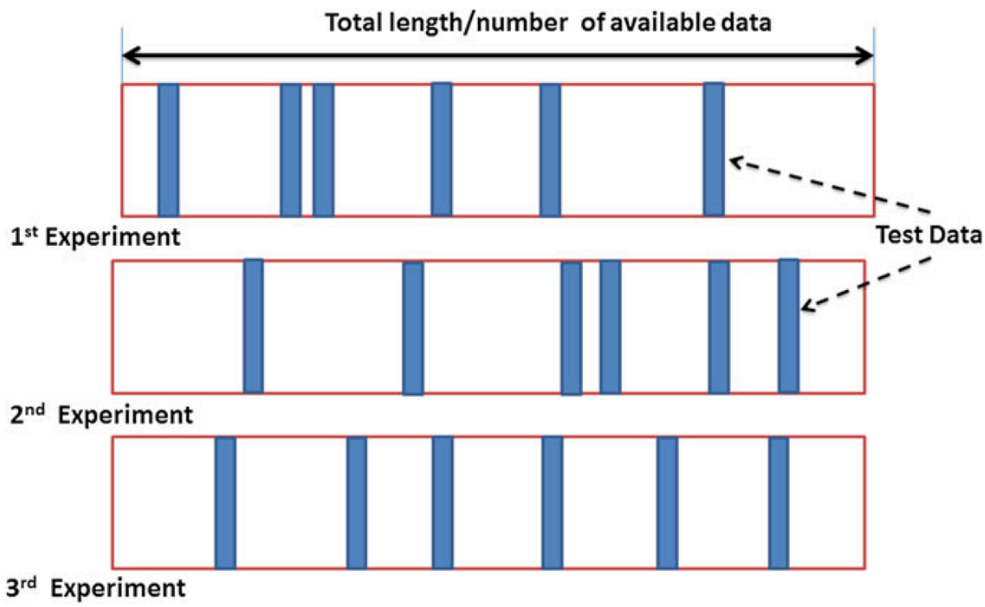
Conversely, splitting train and test set is a risky process because of inhomogeneous distribution of data. Inhomogeneous data distribution may cause regression model to overfit or underfit by accommodating way too much, or missing the trends in the training data and predict too well or too poor. In order to overcome these issues, cross validation methods with multiple validation sets had been developed. In this study, Monte Carlo Cross

Validation, and K-Fold Cross Validation methods are going to be used. CV technique is going to be applied through the Support Vector Regression to validate the statistical test results, and their sustainability. Applied CV method, and the experiment count for CV applied for each empirical chapter is going to be denoted during the workflow of the empirical chapters.

4.5.1.1 Monte Carlo Cross-Validation

The Monte Carlo cross validation (MCCV) was introduced by Richard R Picard and R Dennis Cook (1984). This method randomly splits reserved portion of data into sub-samples and assign them as test sets. The process of selecting random independent partitions repeated for multiple times. Data splitting in Monte Carlo cross validation is presented in Figure 22 (Remesan and Mathew, 2014).

Fig. 22 Monte Carlo Cross Validation

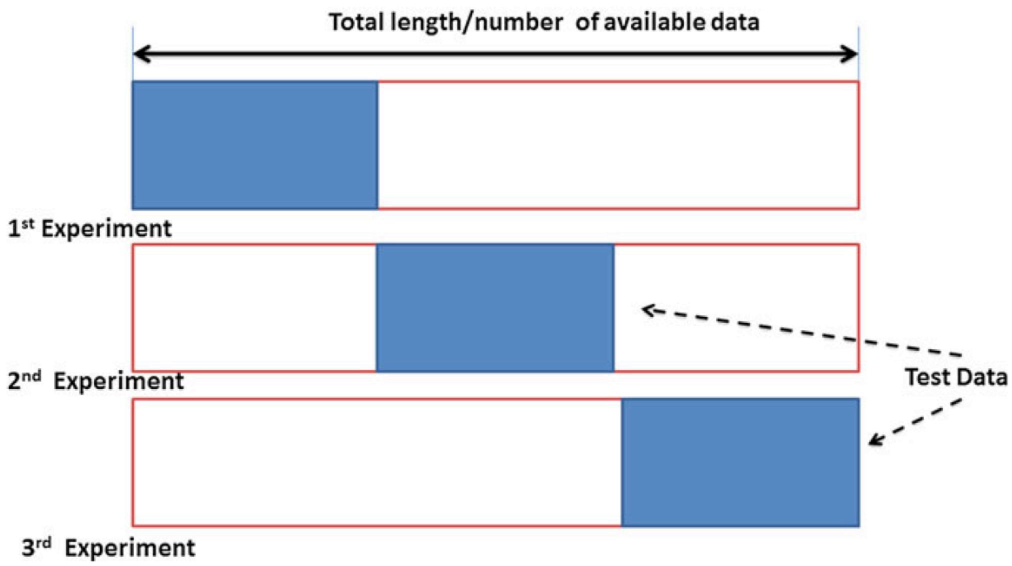


(Source: Dubitzky et al., 2007)

4.5.1.2 K-Fold Cross-Validation

The k-fold cross validation was introduced by Seymour Geisser (1975). This method splits feed data into identified number of corresponding mutually exclusive folds. Parameter *k* refers the number of folds that feed data is going to be split into. One set at a time preserved as hold out set, and trains the model *k-1* times. This process continues until each fold is used as the validation set. Smaller datasets can be benefitted from k-fold cross validation method regarding to its competent practice of limited data (SAS, 2017). Data splitting in k-fold cross validation is presented in Figure 23 (Remesan and Mathew, 2014).

Fig. 23 K-Fold Cross Validation



(Source: Dubitzky et al., 2007)

While both cross validation methods shares the ideology of separating feed data into identified percentage of training and test sets, methods have major differences. The main difference between two methods is; K-fold cross validation use each data only once in precise, while there is a possibility of same data to appear multiple times on Monte Carlo cross validation due to its random selection on individual partitions.

4.5.2 Obtaining Hyperparameters for Support Vector Regression

4.5.2.1 Grid Search

Grid Search (GS) is a traditional method for hyperparameter optimization, which intended to be used in support vector machines and kernel functions. For grid search, it is necessary to apply the guidance of cross validation as a performance metrics. In order to use GS method to tune the hyperparameters to be used in support vector machines and kernel functions, finite set of values for each hyperparameter has to be provided to the SVM. GS algorithm assesses finite set of values individually by training the SVM on every possible combination of hyperparameters, aiming to find the best performing combination when compared to data preserved as test set during the validation procedure. In order to evaluate the performance of hyperparameter combinations, training and test data has to be validated by using the confusion matrix. Confusion matrix indicates the performance of the algorithm by comparing the reserved test data with training data. Confusion matrix is presented in the Figure 24 below.

Fig. 24. Confusion Matrix

Confusion Matrix		Predicted Value	
		Positive	Negative
Actual Value	Positive	True Positive	False Negative
	Negative	False Positive	True Negative

**Green indicates the correctly predicted values.

**Red indicates the error in prediction. Red values should be minimized to achieve higher accuracy in prediction.

where True Positive (TP) and True Negative (TN) indicate correctly predicted positive and negative values respectively. False Positive (FP) and False Negative (FN) indicate the prediction error occurs when actual value and predicted value contradicts with each other.

4.5.2.2 Grid Search with Cross Validation

As it has been mentioned before, GS has to be guided by the CV technique to obtain highest scoring hyperparameter combination possible. There are four main indicators to evaluate the performance of the Grid Search with Cross Validation, which are; Accuracy, Precision, Recall, and F1 Score. Achieved values are going to be between 0 and 1 and going to be interpreted in percentages, where 1 denotes 100%. Koch et al., (1977), and Grubbs (1973) scale the value ranges from 0 to 0.20 as ‘slight agreement’, 0.21 to 0.40 as ‘fair agreement’, 0.41 to 0.60 as ‘moderate agreement’, 0.61 to 0.80 as ‘substantial agreement, and 0.81 to 1.00 as ‘almost perfect agreement’. The main purpose of this evaluation process is to find out the highest scoring hyperparameters, and assign them to the model accordingly.

Accuracy

The ratio of correctly predicted observation (TP and TN) compared to total observations (TP, TN, FP, and FN) provides the insight about the performance, in other words Accuracy of the grid search with cross validation. Accuracy can be calculated as;

$$Accuracy = \frac{True\ Positives + True\ Negatives}{True\ Positives + False\ Positives + True\ Negatives + False\ Negatives} \tag{41}$$

Precision

Precision indicates the ratio of correctly predicted positive observations (TP) compared to the sum of positive observations (TP and FP). The formula for Precision is;

$$Precision = \frac{True\ Positives}{True\ Positives + False\ Positives} \tag{42}$$

Recall

The ratio of correctly predicted positive observations (TP) compared to all classified samples (TP and FN) is named Recall. Recall is also known as the sensitivity value. Recall can be calculated as;

$$Recall = \frac{True\ Positives}{True\ Positives + False\ Negatives} \tag{43}$$

F1 Score

The weighted average of Precision and Recall can be expressed as F1 Score. F1 Score considers both false positives and false negatives. For that reason, on occasions like uneven distribution into classes, F1 Score is preferred metric instead of Accuracy. Yet, Accuracy is still the preferred metric when false positives and false negatives are similar in count. F1 Score can be calculated as;

$$F1\ Score = \frac{2 \times (Precision \times Recall)}{Precision + Recall} \tag{44}$$

4.6 Statistical Tests

Performance, and prediction accuracy of the generated hybrid machine-learning algorithm is going to be evaluated by the statistical tests that are widely accepted as the key metrics to appraise machine learning model's accuracy, namely; V Measure Score (VMS), Explained Variance Score (EVS), Mean Square Error (MSE), Root Mean Square Error (RMSE), Mean Absolute Error (MAE), Mean Absolute Percentage Error (MAPE), and Coefficient of Determination (R^2).

V Measure Score (VMS)

V Measure plays a significant role on overcoming the primary weakness of clustering techniques, which is poor evaluation of performance. V measure is developed to evaluate the performance of clustering tasks. To do so, technique requires two measures, which are; homogeneity and completeness. Homogeneity is a measure that identifies the number of data points clustered from the same class. More the number of data points clustered from the same class, more homogenous the cluster is, hence successful the performance of the algorithm. Homogeneity is formulated as;

(45)

$$h = \begin{cases} 1 & \text{if } H(C,K) = 0 \\ 1 - \dfrac{H(C|K)}{H(C)} & else \end{cases}$$

where;

$$H(C|K) = -\sum_{k=1}^{|K|}\sum_{c=1}^{|C|} \frac{a_{ck}}{N} \log \frac{a_{ck}}{\sum_{c=1}^{|C|} a_{ck}}$$

$$H(C) = -\sum_{c=1}^{|C|} \frac{\sum_{k=1}^{|K|} a_{ck}}{n} \log \frac{\sum_{k=1}^{|K|} a_{ck}}{n}$$

where *K, C, N, n,* a_{ij}, c_i, and k_j denotes set of clusters, set of classes, data points, number of classes, number of data points, member of class, and the element of cluster.

Completeness is a measure that identifies clustering the data points from the same class into same cluster. More the number of data points from the same class clustered to same cluster, more complete the cluster is, hence successful the performance of the algorithm. Completeness is formulated as;

(46)

$$c = \begin{cases} 1 & \text{if } H(K,C) = 0 \\ 1 - \dfrac{H(K|C)}{H(K)} & else \end{cases}$$

where;

$$H(K|C) = -\sum_{c=1}^{|C|}\sum_{k=1}^{|K|}\frac{a_{ck}}{N}\log\frac{a_{ck}}{\sum_{k=1}^{|K|} a_{ck}}$$

$$H(K) = -\sum_{k=1}^{|K|}\frac{\sum_{c=1}^{|C|} a_{ck}}{n}\log\frac{\sum_{c=1}^{|C|} a_{ck}}{n}$$

V measure scores the harmonic mean between mentioned variables, homogeneity and completeness. In the light of obtained homogeneity and completeness values, weighted harmonic mean of homogeneity and completeness can be calculated as;

(47)

$$V_\beta = \frac{(1+\beta) * h * c}{(\beta * h) + c}$$

where β, *h, and c* denotes the weight, homogeneity, and completeness. Strength and the weight of completeness in the calculation is justified if the β

is higher than 1. On the flip side, strength and weight of homogeneity is justified if the β is lower than 1 (Rosenberg and Hirschberg, 2007).
While introducing the V-Measure Score, Rosenberg and Hirschberg (2007) identified the score between 0 and 1, where 0 denotes completely irrelevant clustering, and 1 denotes that the harmonic mean between homogeneity and completeness achieved perfectly. Rosenberg and Hirschberg (2007) indicated the procedure of 'higher score is better' to interpret achieved scores. However, the criterion for threshold of acceptance varies, and not specified by the authors. Later on, the tolerance to test multicollinearity has been identified by Hair et al. (2010), as clustering task is unswervingly related to collinearity by nature. Authors have specified that the higher values are indicating lower collinearity, and identified the score of threshold as ≤ 0.70, significance as ≤ 0.80, and highly significant at ≤ 0.90, which are going to be denoted as *, **, and *** respectively throughout the empirical chapters.

Explained Variance Score (EVS)

Explained Variance Score aims to measure inconsistency between the true values and predicted values. Besides, it clarifies total variance of the model, which can be explained by the factors presented by the actual data and not a part of variance error. Explained variance score shares the same principle with coefficient of determination. Conversely, there is a difference between these two methods. While explained variance uses the biased variance, coefficient of determination uses the raw sum of squares. Explained variance score and coefficient of determination can be different only if the mean of the residuals is not equal to zero, and both scores can be equal only if the error of prediction is unbiased, in other words the mean error equals to zero. Explained variance score can be formulated as;

$$explained\ variance(y, \hat{y}) = 1 - \frac{Var\{y - \hat{y}\}}{Var\{y\}} \tag{48}$$

where $y, \hat{y}, \text{and } Var$ denotes true value, predicted value, and variance (squared standard deviation) respectively.

The metric EVS can also be used as a measure of homocedasticity, as it clarifies the dispersion of errors of the selected dataset, where less dispersion indicates the homocedasticity. As the score is identified between 0 and 1, it is going to be interpreted in percentages, where 1 denotes 100%. Hair et al., (2014) have indicated the threshold for imprecise datasets as ≤ 0.60, level of significance as ≤ 0.80, and highly significant as ≤ 0.90, which are going to be denoted as *, **, and *** respectively throughout the empirical chapters.

Mean Squared Error (MSE)

Mean squared error evaluates the performance of prediction by averaging the set of errors obtained by measuring the distance between data points and regression line. Squaring plays a vital role on weight distribution on points with larger differences. Mean squared error formulated as;

(49)

$$MSE = \frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2$$

where *n,* y_i, and $\hat{y}_i$ denotes number of data points, values observed, and predicted value of the i-th sample respectively.

The main objective of calculating MSE is to achieve RMSE, alongside to measure the variance of the residuals. MSE and RMSE both have a reputation on performing more accurate compared to MAE when the dataset includes outliers. It is accepted that MSE closer to 0 indicates better model fit. However, there isn't any acceptable range set for MSE as it does not shares the same unit as the original values.

Root Mean Square Error (RMSE)

Root mean square error measures the dissimilarity between the predicted vale and corresponding actual value. RMSE is very likely to contribute relatively high weight to greater error values due to squaring the errors prior to averaging, which makes RMSE advantageous on cases where high errors are exceptionally unfavorable. RMSE is formulated as;

(50)

$$RMSE = \sqrt{\frac{\frac{1}{n}\sum_{i=1}^{N}(x_i - \hat{x}_i)^2}{N}}$$

where *N*, x_i, and $\hat{x}_i$ represents number of data points, actual values, and predicted value of the i-th sample respectively.

Square rooting the MSE value provides the output in the same unit as the data plotted on the vertical axis, and specifies the absolute distance between the actual and predicted values. As RMSE output shares the identical unit with predicted values, it is hard to interpret if the value is significant, acceptable or inacceptable. To do so, Scatter Index (SI) is going to be used to convert the RMSE value into interpretable value (Mentaschi et al., 2013). SI is formulated as;

(51)

$$SI = \sqrt{\frac{\sum_{i=1}^{N}[(S_i - \bar{S}) - (O_i - \bar{O})]^2}{\sum_{i=1}^{N} O_i^2}}$$

where $\bar{O}$ and $\bar{S}$ denotes the values of average observation and simulation respectively. Attained SI value is always between 0 and 1, where lower value indicates better performance. Mentaschi et al., (2013) denoted the acceptable SI value as ≤0.10, and significant SI value as ≤ 0.05, which are going to be denoted as *, and ** throughout the empirical chapters.

Mean Absolute Error (MAE)

Mean absolute error assesses prediction error in time series analysis. It can be obtained by calculating arithmetic average of the absolute errors, where comparing the dissimilarity between true value and predicted value can attain absolute error. Mean absolute error formulated as;

$$MAE = \frac{1}{n}\sum_{i=1}^{n}|y_i - \hat{y}_i| \tag{52}$$

where *n,* y_i, and $\hat{y}_i$ denotes number of data points, true value, and predicted value of the i-th sample respectively.

MAE value is going to be compared with RMSE value due to MAE's relative insensitive nature to outliers compared to RMSE. RMSE values are always higher than MAE values, as RMSE magnifies and penalizes the bigger error values during the squaring process, while ignoring the smaller error values. For that reason, it is possible to say that RMSE emphasizes larger errors better than MAE. The difference between MAE and RMSE values indicates the variance in individual errors, where larger difference indicates the greater inconsistency in the error size. The acceptable difference between RMSE and MAE has denoted as $\leq 5\%$, while $\leq 1\%$ signifies the perfect fit of the model, which are going to be denoted as *, and ** throughout the empirical chapters.

Mean Absolute Percentage Error (MAPE)

Mean absolute percentage error computes the accuracy of the prediction made through calculating the average of absolute error in percentage terms for each time period. MAPE calculates each variable individually, and achieve normalized absolute error before averaging. Due to being intuitively interpret; MAPE frequently used a loss function on regression analysis. Mean absolute percentage error formulated as;

(53)

$$MAPE = \frac{1}{n}\sum_{t=1}^{n}\left|\frac{A_t - F_t}{A_t}\right|$$

where *n,* A_t, and F_t represents number of times the summation iteration happens, actual value, and forecast value respectively.
Where lower values indicates higher prediction rates. For the presented methodology Liu and Lin (2010), for presented discipline Swanson (2015) stated the threshold for MAPE value as ≤ 0.25, ≤ 0.10 indicates low but acceptable accuracy, and ≤ 0.05 indicates highly acceptable accuracy, which are going to be denoted as *, **, and *** throughout the empirical chapters.

Coefficient of Determination

The main purpose of coefficient of determination (CoD/ R^2) is probabilistic investigation. It indicates the percentage of the variance for a dependent variable from the regression model. Coefficient of determination indicates the percentage of predicted data points that are on the produced regression line. In other words, it specifies the strength of relationship between variables. Coefficient of Determination can be formulated as;

(54)

$$CoD = \left(\frac{\sum_{i=1}^{n}[(\tilde{y}_i - \bar{\tilde{y}}) \cdot (y_i - \bar{y})]}{\sqrt{\sum_{i=1}^{n}(\tilde{y}_i - \bar{\tilde{y}})^2 \cdot \sum_{i=1}^{n}(y_i - \bar{y})^2}}\right)^2$$

where $y, \tilde{y}, \text{and}\ \bar{\tilde{y}}$ denotes result variable, meta-model result variable, and mean variable respectively, while i represents count variables and v denotes validation. Moreover, cross validation CoD (CoD_{CV}) can be formulated as;

(55)

$$CoD_{CV} = \left(\frac{\sum_{i=1}^{n}[(\tilde{y}_{iv} - \bar{\tilde{y}}_v) \cdot (y_{iv} - \bar{y}_v)]}{\sqrt{\sum_{i=1}^{n_v}(\tilde{y}_{iv} - \bar{\tilde{y}}_v)^2 \cdot \sum_{i=1}^{n_v}(y_{iv} - \bar{y}_v)^2}}\right)^2$$

The value for R^2 always stands between 0 and 1. For that reason, it is possible to be stated in percentages, from 0% to 100%. Higher R^2 value is always preferred as it refers to more accurate measurements (Sarstedt and Mooi, 2014). The multidisciplinary approach, adopted methodology, and the size of data for this study is considered to delegate the level of acceptance and significance for attained R^2 results. Based on the recommendations and findings of Henseler et al., (2009), Hair et al., (2013), and Sarstedt and Mooi, (2014), the values 0.50 assigned as 'weak', 0.75 as 'moderate', and 0.90 as 'substantial', which are going to be denoted as *, **, and *** throughout the empirical chapters.

4.7 Assigning Weights to Prospect Theory, Fuzzy Logic and Support Vector Regression Structure for Multi-Criteria Decision-Making

Methodology that is going to be presented below is supplementary to the structures of abovementioned methodologies; prospect theory, fuzzy logic and support vector regression. The purpose of the supplementary techniques is the adaptation of methodology to the complex multi-criteria structure of Chapters 7 and 8, and enhancement of algorithms' performance. The incentive for the use of supplementary techniques, and the necessity for enhancement of the algorithms' are going to be comprehensively clarified in the aforementioned chapters where necessary. Below mentioned formulas will be supplementing as a sequel of the previously mentioned methodologies, or rewritten where enhancement required. Generated algorithm is going to be regulated accordingly.

4.7.1 Weighted Prospect Theory

Linguistic terms are not always precise. Words that are uncertain by their nature may cause varying actions by decision makers', based on their psychological state, and preferred range of risk taking. This gives uncertain linguistic variables a subjective level of understanding, in other words make them fuzzy. Fuzzy variables play a substantial role on weighting the probability function properly to achieve higher accuracy rates on further calculations. Liu, Jin, Zhang, Su, and Wang (2011) contributed to literature by transforming uncertain linguistic variables into fuzzy numbers to weight the probability function regarding to decision making reference point of each attribute. To minimize the error on understanding the behavior of decision makers', interval probability *p* on the probability weight function should be transformed into interval probability weights $\left(w_j^1, w_j^2, \ldots, w_j^{l_i}\right)$ of the *j*th attribute under the l_ith status;

(56)

$$\left[w\left(\bar{p}_j^{L1}\right), w\left(\bar{p}_j^{U1}\right)\right]\left[w\left(\bar{p}_j^{L2}\right), w\left(\bar{p}_j^{U2}\right)\right] \cdots \left[w\left(\bar{p}_j^{Ll_i}\right), w\left(\bar{p}_j^{Ul_i}\right)\right]$$
$$(j = 1, 2, \ldots, n)$$

Besides that, potential uncertain linguistic variables should be considered while weighting the probability. Uncertain linguistic variables are going to be denoted as $\left[x_{ij}^{L0}, x_{ij}^{U0}\right]$ for referring the points of decision makers for varying attributes. As these linguistic variables are uncertain by their nature, they should be transformed into fuzzy numbers for further processes. Transformed fuzzy numbers are going to be denoted as $\left[a_j^{L0}, a_j^{ML0}, a_j^{MU0}, a_j^{U0}\right]$. In this instance, the prospect value function of fuzzy number on *k*th status of *j*th attribute under *i*th alternative can be denoted as $\left[x_{ij}^{Lk}, x_{ij}^{Uk}\right]$ for the uncertain variable. This will renovate the fuzzy numbers into $\left[a_{ij}^{Lk}, a_{ij}^{MLk}, a_{ij}^{MUk}, a_{ij}^{Uk}\right]$, and the prospect value function to;

(57)

$$z_{ij}^k = \left[v\left(a_{ij}^{Lk} - a_j^{U0}\right), v\left(a_{ij}^{MLk} - a_j^{MU0}\right), v\left(a_{ij}^{MUk} - a_j^{ML0}\right), v\left(a_{ij}^{Uk} - a_j^{L0}\right)\right]$$

where $v(x)$ denotes the value function on the original prospect theory formula. While the interval number represents the *k*th status of probability weight function under the *j*th attribute $w_j^k = \left[w\left(\bar{p}_j^{Lk}\right), w\left(\bar{p}_j^{Uk}\right)\right]$, fuzzy number denotes the prospect value function $\left[w\left(\bar{p}_j^{Lk}\right), w\left(\bar{p}_j^{Lk}\right), w\left(\bar{p}_j^{Uk}\right), w\left(\bar{p}_j^{Uk}\right)\right]$. Subsequently, *j*th attribute under the *i*th alternative of the prospect function z_{ij} can be formulated as;

(58)

$$\begin{aligned} z_{ij} &= \sum_{k=1}^{l_i}\left(w_j^k z_{ij}^k\right) = \left[z_{ij}^L, z_{ij}^{ML}, z_{ij}^{MU}, z_{ij}^U\right] \\ &= \left[\sum_{k=1}^{l_i}\left(w\left(\bar{p}_j^{Lk}\right)v\left(a_{ij}^{Lk} - a_j^{U0}\right)\right),\right. \\ &\sum_{k=1}^{l_i}\left(w\left(\bar{p}_j^{Lk}\right)v\left(a_{ij}^{MLk} - a_j^{MU0}\right)\right), \\ &\sum_{k=1}^{l_i}\left(w\left(\bar{p}_j^{Uk}\right)v\left(a_{ij}^{MUk} - a_j^{ML0}\right)\right), \\ &\left.\sum_{k=1}^{l_i}\left(w\left(\bar{p}_j^{Uk}\right)v\left(a_{ij}^{Uk} - a_j^{L0}\right)\right)\right] \end{aligned}$$

while the ith alternative of the weighted prospect function value can be formulated as;

(59)

$$z_i = \sum_{j=1}^{n} (\omega_j \times z_{ij}) = [z_i^L, z_i^{ML}, z_i^{MU}, z_i^U]$$

$$= \left[\sum_{j=1}^{n} \left(\omega_j \times \sum_{k=1}^{l_i} \left(w(\bar{p}_j^{Lk}) v(a_{ij}^{Lk} - a_j^{U0}) \right) \right), \right.$$

$$\sum_{j=1}^{n} \left(\omega_j \times \sum_{k=1}^{l_i} \left(w(\bar{p}_j^{Lk}) v(a_{ij}^{MLk} - a_j^{MU0}) \right) \right),$$

$$\sum_{j=1}^{n} \left(\omega_j \times \sum_{k=1}^{l_i} \left(w(\bar{p}_j^{Uk}) v(a_{ij}^{MUk} - a_j^{ML0}) \right) \right),$$

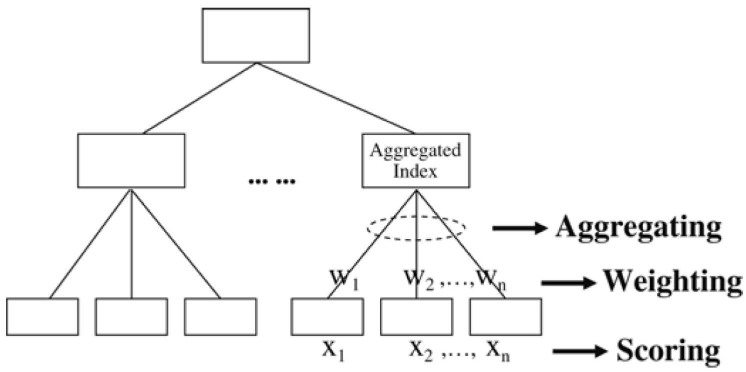
$$\left. \sum_{j=1}^{n} \left(\omega_j \times \sum_{k=1}^{l_i} \left(w(\bar{p}_j^{Uk}) v(a_{ij}^{Uk} - a_j^{L0}) \right) \right) \right]$$

where ω and ω_j denotes weight vector and weight of attribute respectively.

4.7.2 Weighted Fuzzy Logic

A hierarchical structure has to be composed to identify the evaluation problem in a situation of multi-criteria decision-making. Assessing elements with a hierarchical structure proven to perform better on classifying the alterations between a clusters of assessed objects (Stillwell et al, 1987). Evaluating the importance of a criteria to form a hierarchical structure is completely varies on the decision makers' relative and subjective assessment. For that reason, a situation of multi-criteria decision-making cannot be accepted as a deterministic and a crisp value, and should be considered as fuzzy (Guh, et al., 2008). In order to evaluate the hierarchy of multi-criteria in a vague environment, the operation that has been designed by 3 main operators has to be followed in order to achieve aggregate index. These operators are; scoring, weighting, and aggregating the criteria, scoring indicates the attained performance for each criteria, weighting indicates the relative significance of individual criteria based on cluster of analogous, and aggregating indicates the connective operator in based on the decision makers' sensitivity. The structure has been visualized in Figure 25.

Fig. 25 Operation Structure for Aggregation Index



Fuzzy Weighted Average has composed by integrating the fuzzy variables of hierarchical evaluation that forms the operators into fuzzy set theory. Fuzzy weight average can be formulized as;

(60)

$$y = f(x_1, x_2, \dots, x_n, w_1, w_2, \dots, w_n) = \left(\frac{w_1 x_1 + w_2 x_2, \dots, +w_n x_n}{w_1 + w_2, \dots, +w_n}\right)$$
$$= (w_1' x_1 + w_2' x_2, \dots, +w_n' x_n)$$

where w_i and x_i denotes relative weighting and scoring criteria respectively. Function f denotes mapping of $X_1 \times X_2 \times \dots X_n \times Z_1 \times Z_2 \times \dots \times Z_n$ where the fuzzy weighted average numbers defined as $A_1,\ A_2, \dots\ A_n$ and weights of fuzzy numbers $W_1,\ W_2, \dots\ W_n$, on the universes $X_1,\ X_2, \dots\ X_n\ and\ Z_1,\ Z_2, \dots, Z_n$, while normalized weight is denoted as w_i'. Conversely, as the aforementioned indications of scoring, weighting, and aggregating signifies, it is obvious that they have a significant influence on understanding decision makers' assessment, thus aggregation factor. In this instance, aggregation should also be considered as fuzzy variable alongside score and weight operators, as aggregation indicates the subjective approach of the decision makers. While aggregation index function towards maximum on scenarios with an optimistic behavior of decision maker, function towards minimum on scenarios with a pessimistic behavior of decision maker is implemented. Therefore, domination effect of operators has to be bared in mind, as higher score with lower weight still need to be considered with a low weight score.

4.7.2.1 Fuzzy Weighted Average Within a Generalized Mean Operator

Dujmovic was introduced weighted generalized mean operator in 1974. Weighted generalized mean operator is monotonously incessant function, and represented as; $\min(x_1, x_2 \ldots, x_n) \leq y \leq \max(x_1, x_2 \ldots, x_n)$. It is possible to derive the membership function of fuzzy weighted average by Dujmovic's (1974) generalized mean operator. Dyckoff and Pedrycz (1984) contributed to Dujmovic's (1974) weighted generalized mean operator by fuzzifying the aggregation operator, ⊕. Fuzzy weighted average is a hierarchical evaluation method that addresses the approximate membership values. Dyckoff and Pedrycz's (1984) fuzzy weighted average function integrating the generalized mean operator is formulated as;

(61)

$$y = (w_1' x_1 \oplus w_2' x_2, \ldots, \oplus w_n' x_n)$$

$$y = f(x_1, x_2, \ldots, x_n, w_1', w_2', \ldots, w_n', p) = w_1' x_1^p + w_2' x_2^p, \ldots, + w_n' x_n^p)^{1/p}$$

$$y = f(x_1, x_2, \ldots, x_n, w_1, w_2, \ldots, w_n, p) = \left(\frac{w_1 x_1^p + w_2 x_2^p, \ldots, + w_n x_n^p}{w_1 + w_2, \ldots, + w_n} \right)^{1/p}$$

$$\alpha_i \leq x_i \leq b_i, c_i \leq w_i \leq d_i$$

where w_i', and x_i denotes normalized weight within a hierarchy, and score for all fuzzy numbers respectively, while their corresponding intervals denoted as c_i and d_i for w_i, and a_i and b_i for x_i. Varieties of aggregation operators are obtainable by the weighted generalized means operator, by differentiating the parameter value *p. Primary* operators obtained by the variations of valuating the parameter value is presented below;

$$p = -\infty, \text{the minimum operator}$$
$$p = -1, \text{the harmonic mean operator}$$
$$p = 0, \text{the geometric mean operator}$$
$$p = +1, \text{the arithmetic mean operator}$$
$$p = +\infty, \text{the maximum operator}$$

Altered values of *p* identify the decisive behavior of the decision maker. While higher value of *p* for weight generalized mean function indicates the openness on decision maker's behavior of evaluation, lower value of *p* for weight generalized mean function indicates that decision maker's evaluation is moderate. Weight of *p* on generalized mean function can be measured by calculating the summation on the inequality, and dividing it by $\sum_{i=1}^{n} w_i$, which can be formulated as;

(62)

$$\frac{\sum_{i=1}^{n} w_i \alpha_i^p}{\sum_{i=1}^{n} w_i} \leq \frac{\sum_{i=1}^{n} w_i x_i^p}{\sum_{i=1}^{n} w_i} \leq \frac{\sum_{i=1}^{n} w_i b_i^p}{\sum_{i=1}^{n} w_i}$$
$$\Rightarrow \left(\frac{\sum_{i=1}^{n} w_i \alpha_i^p}{\sum_{i=1}^{n} w_i}\right)^{1/p} \leq \left(\frac{\sum_{i=1}^{n} w_i x_i^p}{\sum_{i=1}^{n} w_i}\right)^{1/p} \leq \left(\frac{\sum_{i=1}^{n} w_i b_i^p}{\sum_{i=1}^{n} w_i}\right)^{1/p}$$

Lower and upper bounds for each α-cut interval is expressed as $L_{p\geq 0}, U_{p\geq 0}$ while $p \geqq 0$, where $L_{p\geq 0} = Min_{p\geq 0} f_L$, and $U_{p\geq 0} = Max_{p\geq 0} f_U$, and formulated as;

(63)

$$L_{p\geq 0} = Min_{p\geq 0} f_L = Min_{p\geq 0}\left(\frac{w_1}{\sum_{i=1}^{n} w_i}\alpha_1^p + \frac{w_2}{\sum_{i=1}^{n} w_i}\alpha_2^p, \dots + \frac{w_n}{\sum_{i=1}^{n} w_i}\alpha_n^p\right)^{1/p}$$
$$c_i \leq w_i \leq d_i, i = 1, 2, \dots, n$$

$$U_{p\geq 0} = Max_{p\geq 0} f_U = Max_{p\geq 0}\left(\frac{w_1}{\sum_{i=1}^{n} w_i}b_1^p + \frac{w_2}{\sum_{i=1}^{n} w_i}b_2^p, \dots + \frac{w_n}{\sum_{i=1}^{n} w_i}b_n^p\right)^{1/p}$$
$$c_i \leq w_i \leq d_i, i = 1, 2, \dots, n$$

In order to take the mean of the natural logarithm, Napierian logarithm is implemented for $L_{p\geq 0}$ and $U_{p\geq 0}$, which is formulated as;

(64)

$$\text{In } L_{p\geq 0} = Min_{p\geq 0} f_L = Min_{p\geq 0} \frac{1}{p} \text{In} \left(\frac{w_1}{\sum_{i=1}^{n} w_i} \alpha_1^p + \frac{w_2}{\sum_{i=1}^{n} w_i} \alpha_2^p, \dots + \frac{w_n}{\sum_{i=1}^{n} w_i} \alpha_n^p \right)$$
$$c_i \leq w_i \leq d_i, i = 1, 2, \dots, n$$
$$= \frac{1}{p} \text{In} Min_{p\geq 0} \left(\frac{w_1}{\sum_{i=1}^{n} w_i} \alpha_1^p + \frac{w_2}{\sum_{i=1}^{n} w_i} \alpha_2^p, \dots + \frac{w_n}{\sum_{i=1}^{n} w_i} \alpha_n^p \right) = \frac{1}{p} \text{In} L'_{p\geq 0}$$
$$c_i \leq w_i \leq d_i, i = 1, 2, \dots, n$$

$$\text{In } U_{p\geq 0} = Max_{p\geq 0} f_U = Max_{p\geq 0} \frac{1}{p} \text{In} \left(\frac{w_1}{\sum_{i=1}^{n} w_i} b_1^p + \frac{w_2}{\sum_{i=1}^{n} w_i} b_2^p, \dots + \frac{w_n}{\sum_{i=1}^{n} w_i} b_n^p \right)$$
$$c_i \leq w_i \leq d_i, i = 1, 2, \dots, n$$
$$= \frac{1}{p} \text{In} Max_{p\geq 0} \left(\frac{w_1}{\sum_{i=1}^{n} w_i} b_1^p + \frac{w_2}{\sum_{i=1}^{n} w_i} b_2^p, \dots + \frac{w_n}{\sum_{i=1}^{n} w_i} b_n^p \right) = \frac{1}{p} \text{In} U'_{p\geq 0}$$
$$c_i \leq w_i \leq d_i, i = 1, 2, \dots, n$$

4.7.3 Feature-Weighted Support Vector Regression

Feature-Weighted Support Vector Regression emphasized discernibility matrix to approximate relative significance/weights of each under multifarious weighting differentiations in order to diminish the impact of relatively non-acute samples and noise. So as to weight the inputs on the feature space generated by using the kernel function, primarily the distance between sample points has to be obtained. The distance between inputs may be measured by;

(65)

$$d_{ij} = \left\| \phi(x_i) - \phi(x_j) \right\|^2$$
$$= \langle \phi(x_i) - \phi(x_j), \phi(x_i) - \phi(x_j) \rangle$$
$$= \langle \phi(x_i), \phi(x_i) \rangle - 2\langle \phi(x_i), \phi(x_j) \rangle + \langle \phi(x_j), \phi(x_j) \rangle$$
$$= K(x_i, x_i) - 2K(x_i, x_j) + K(x_j, x_j)$$

where $\phi(x_i)$, $\phi(x_j)$ maps the sample points x_i and x_j in hyperspace, and d_{ij}, denotes the distance between sample points in hyperspace. When RBF kernel is implemented, d_{ij} can be formulated as;

(66)

$$d_{ij} = 2 - 2K(x_i, x_j) = 2 - 2\exp\left(-\gamma\|x_i - x_j\|^2\right)$$

The distance between sample points indicates the rate of similarity between sample points, where smaller distance indicates higher rate of similarity, and $d_{ij} = 0$ indicates that x_i and x_j are equal. Therefore, if the weight value $w_k \in [0,1]$, kernel function can be formulated as;

(67)

$$K_w(x_i, x_j) = exp\left(-\gamma\left(\sum_{k=1}^{n}\left(w_k(x_{ik} - x_{jk})\right)^2\right)\right)$$

Based on the aforementioned implementations of weight values, quadratic programming of optimization problem can be revised as;

(68)

$$min_{\alpha,\alpha*}\frac{1}{2}\sum_{i=1,j=1}^{l}(\alpha_i - a_i^*)(\alpha_j - \alpha_j^*)K_w\left(x_i, x_j\right) + \varepsilon\sum_{i=1}^{l}(\alpha_i + \alpha_i^*) - \sum_{i=1}^{l}(\alpha_i - \alpha_i^*)y_i$$

subject to:

$$\begin{cases}\sum_{i=1}^{l}(\alpha_i - \alpha_i^*) = 0\\ 0 \leq \alpha_i, a_i^* \leq C\end{cases}$$

In this case, formula for support vector regression should also be revised into featured weight form. Featured weight support vector regression formula can be formulated as;

(69)

$$f(x) = \sum_{i=1}^{l}(\bar{\alpha}_i^* - \bar{\alpha}_i)K_w(x_i, x) + \bar{b}$$

CHAPTER 5

The Instantaneous Impact of Recently Published Macroeconomic News on Exchange Rate

5.1 Introduction

A market survey conducted by Cheung and Chinn (2001) proposed the significance of macroeconomic indicators in forecasting exchange rates in changing economic environments. Continuous fluctuation of the focal point instigates the precarious interaction between the exchange rate and macroeconomic indicators. Moreover, Rossi (2006) presented empirical evidence for the causality of macroeconomic fundamentals on devastating volatility of exchange rates, while rejecting the hypothesis of random walks. Studies pioneered the theory that macroeconomic fundamentals provide an appreciably sufficient source of information that suggests frequent changes in the exchange rates (Sarno & Valente, 2009). Bacchetta and Wincoop (2009) claimed that the largest portion of the interaction had naturally been established. The macroeconomic fundamentals and exchange rate are the misty subjects of the economics that are likely to change, which leads to the interaction of structural parameters (Beckmann et al., 2010). Resultantly, the consideration of macroeconomic fundamentals as a significant parameter during specific periods is adequate (Bacchetta & Wincoop, 2004). Goldberg and Frydman (1996a, 2007) correspondingly supported the imperfect knowledge approach, and clarified the instability of parameters. Market participants were insensible of the principles of the model; however, they preferred to use the macroeconomic fundamentals in order to forecast the exchange rates effectively. As the deeper understanding of the cause and effect connection was developed and digested by the market participants, stronger bonds between fundamentals and exchange rates started to occur. The majority of the recent literature also shares the parallel hypothesis that supports the existence of interaction and causality in both short and long run. Besides, while examining historical data of macroeconomic variables may be useful to predict exchange rate in long run, behavioral analysis should be adopting exchange rates in daily basis to run prediction (Klitgaard and Weir, 2004). The motive behind using exchange rate in daily basis is; as the exchange rate is considered to be an essential endogenous variable that responds to both exogenous and policy-driven distortions, it is a variable that is either directly or indirectly targeted by policymakers (Montiel and Ostry, 1991). Macroeconomic indicators and financial prices reflect the thoughts, understandings, and expectation of others on relatively endogenous information structure (Angeletos and Werning, 2006). Any kind of event-based information that may influence the investors' sentiments has a direct impact on exchange rate prediction and expectation (Yasir et al., 2019). Moreover, economic condition and policy of a country is highly dependent on exchange rates (Kandil et al., 2007; Cehreli et al., 2017) due to its significant role on international trade, investment determination, risk management, and balance of payments (Sharma et al, 2016; Korol, 2014). Therefore, volatile exchange rate may influence sentiments of investors, and cause macroeconomic indicators negatively (De Grauwe, 1988). Angeletos and La'O (2013) defined sentiment shocks as the primary beliefs on endogenous economic outcomes, which is adequate to cause a change in related equilibrium. Accuracy of their argument on primary beliefs has also been proven by the survey-based research. Furthermore, recently announced macroeconomic information orients investors to reconsider and reshapes their thoughts, understanding, and expectations about the state of the economy, and respond accordingly (Bahloul and Gupta, 2018). Conversely, the significance of exchange rate on macroeconomic management is crucial (Saraç and Karagöz, 2016). Unstable exchange rates have adverse impact on macroeconomic variables in short term (Bakhshi and Ebrahimi, 2016). Additionally, fluctuating exchange rate may cause domino effect on macroeconomic variables (Ali et al., 2014). Consequently, literature supports and verifies the existence of a robust and significant interaction between macroeconomic indicators and exchange rates, while authenticating that exchange rates are highly volatile and substantially information dependent. Hence, consideration of subjective and qualitative information that has been introduced above while predicting the volatility of exchange rates is critical for the minimization of the margin for error.

This chapter aims to bring a novel approach to exchange rate prediction by investigating the impact of official macroeconomic statements and recently published news by examining the reaction of the investors and understanding their psychological response on particular variations of news contents. A hybrid machine-learning algorithm is going to be generated by the author to unveil the significance of macroeconomic news on exchange rates fluctuations, and predict the price of the exchange rates on a daily basis. Macroeconomic news have been used to predict macroeconomic fundamentals, and exchange rate in existing literature (Caporale et al., 2018; Bahloul and Gupta, 2018; Hachula and Nautz, 2018), however this chapter is pioneering on using hybrid machine learning algorithm, and examining both and qualitative and quantitative data to predict exchange rate. News regarding to government officials' statements, expectations on the attitude on economic policies, and publication of macroeconomic data plays a crucial role on informing the community about recent economic situation that country is currently in. The information provided by the news plays a significant role on the behavior of investors' in short run. In order to comprehend the correlation between news and exchange rate volatility, recently published macroeconomic are going to be examined alongside to historical exchange rate and macroeconomic data.

Emerging economies are perfect fit to study due to their exceptionally responsive nature on both domestic and foreign economic decisions, expectations, and events (Mehdian et al., 2008). In recent years, Turkish economy is an obvious proof to clarify this hypothesis. Turkish Lira depreciated by 152% against US Dollar, and hit record braking lows while becoming the worst performing currency of emerging markets between the dates of 29 December 2017 to 01 November 2019. Furthermore, macroeconomic indicators conspicuously fluctuated significantly between aforementioned date interval, which may clearly be observed as the lowest inflation rate was 8.55%, while highest was 25.52%, lowest unemployment rate was 9.60%, while highest was 14.7%, and lowest interest rate was 8.0%, while highest was 24%. Meanwhile, trade gap decreased down to -0.46 Billion USD, and increased up to -9.21 Billion USD. Moreover, most significant rating agencies, namely; Standard & Poor's (S&P), Moody's, and Fitch downgrade Turkey's rating from highest band of Speculative Grade Rating down to mid range ratings. S&P downgraded Turkey's credit rating from BB to B+, Moody's downgraded Turkey's credit rating from Ba1 to B1, and Fitch downgraded Turkey's credit rating from BB+ to BB-. Hence, consecutive macroeconomic fluctuations that have been occurred in a limited date interval, which been aforementioned above provided a unique opportunity to study instantaneous effects of official macroeconomic statements, and publicly published news on exchange rate fluctuations.

This chapter is structured as follows: Section 2 presents and details the selected macroeconomic indicators and collected data and fundamental variables; Section 3 describes applied methodologies; Section 4 presents empirical workflow and results; and Section 5 presents concludes the chapter.

5.2 Data and Fundamental Variables

5.2.1 Macroeconomic Indicators

To perceive the impact of expected and existing fluctuations on macroeconomic indicators by examining the related news, inflation rate, unemployment rate, interest rate, balance of trade, and credit rating have selected to be examined, as the causality between exchange rate and macroeconomic indicators have broadly be verified by the literature. Volatility of aforementioned indicators during selected date intervals have also presented on the previous section of this chapter. What literature suggests regarding to the correlation between aforementioned macroeconomic indicators and exchange rate may be found below;

5.2.1.1 Inflation Rate

Inflation rate is one of the key macroeconomic indicators, and has a crucial role in monetary policy, in addition to guiding both occupants' and enterprises' financial planning by indicating strong signals on purchasing power of residents (Zhang and Li, 2012). Theoretically, correlation between exchange rate and inflation rate had proposed by Barro and Gordon (1983). Their argument has been empirically supported by recent studies that introduced both direct and indirect impacts of exchange rate fluctuations on inflation rate. Because of its direct impact on purchasing power, as well as demand and supply of imported and exporting goods, exchange rate valuation plays a significant role on inflation rate (Şen et al., 2019). The recent researches, Parsley and Wei (2007) and Qui et al. (2011) shares parallel findings that indicate the significance of the country's currency demand and supply on the inflation rate under the sway of change in the level of imports and exports.

Şen, Kaya, Kaptan, and Cömert (2019) investigated the interrelationship between inflation rate, interest rate, and exchange rate in emerging economies. To do so, authors have selected Brazil, India, Indonesia, South Africa, and Turkey. The main similarity of selected countries is being highly import-dependent. Results indicate a strong correlation between inflation rate and exchange rates. Findings specify that, appreciation of the foreign currency value cause inflation through import channel.

5.2.1.2 Interest Rate

Interest rate is one of the key macroeconomic variables, and constitutes a vital part of policy variables in managing unpremeditated exchange rate fluctuations (Saraç and Karagöz, 2016). Unpremeditated exchange rate fluctuations may force central banks to take monetary action by intervening interest rate in order to take fluctuating exchange rate under control. The effectiveness of exchange rate on interest rate, vice versa have been a subject to many theoretical and empirical researches throughout the literature. Cheung et al. (2002) presented significant impacts of exchange rate uncertainty on interest rates. Karahan and Çolak (2012) have also investigated the causality of exchange rate uncertainty on interest rate in Turkey. Their results indicate that, rising exchange rate uncertainty causes interest rates to increase.

As aforementioned studies suggests, the correlation between exchange rate and interest race cannot be ignored. Moreover, interest rate is a qualitatively essential variable to perform an effective exchange rate prediction (Greun and Wilkinson, 1994; Chen, 2007).

5.2.1.3 Balance of Trade

Balance of trade is another variable, which its linkage with exchange rate became a subject for numerous theoretical and empirical studies. Recent literature approves the impact of exchange rate fluctuations on balance of trade (Ozturk, 2006).

Asteriou et al. (2016) investigated the causality of exchange rate volatility on balance of trade for four different countries including Turkey in both short and long term. Research results indicate a substantial short term, but only a moderate long term causality between exchange rate volatility and trade balance in Turkey.

Unlike relevant studies, Serenis and Tsounis (2016) developed an alternative method to measure exchange rate fluctuations and its affect on balance of trade. Instead of evaluating the standard deviation of the moving average of the logarithm, authors take high and low values of exchange rate into account. Results specified a significant association between fluctuating exchange rates and balance of trade.

5.2.1.4 Unemployment Rate

The correlation between exchange rate and unemployment rate in both emerging and developed economies has been widely examined and approved in existing literature (Frenkel and Ros, 2006; He, 2013; Milas and Legrenzi, 2006).

As an alternative perspective, factors, which have an impact on unemployment rate have been investigated by Ranjbar and Moazen (2009) on eight different countries including Turkey. Their study advocates an adverse correlation between unemployment rate and exchange rate.

Uncertain fluctuations of exchange rate also play a crucial role on unemployment rate. Chang (2011) investigated the relationship between uncertain exchange rate and unemployment. Correlation between uncertain exchange rate and unemployment detected in both short and long term.

Furthermore, exchange rate volatility is another significant factor that has an impact on unemployment rate. The relationship between exchange rate volatility and unemployment rate examined by Nyahokwe and Ncwadi (2013), and substantial correlation between variables has been observed.

5.2.1.5 Credit Ratings

Credit rating has been defined by Standard & Poor's as "a present view on the creditworthiness of an obligor relating to particular financial obligation, a special class of obligations, or a particular financial program" (Martell, 2005). Furthermore, a sovereign credit rating suggests the capability, likelihood, and readiness of the governments to operate their financial obligations, according to rating agencies (Mateev, 2012).

The effects of sovereign credit ratings on exchange volatility have been examined by Subaşi (2008). Results of the study indicate a correlation between sovereign credit ratings and USD/TRY and EUR/TRY exchange rate volatility. Findings confirms an increase in volatility, which causes a depreciation of the exchange rate when the credit rating downgrades, while upgrading credit rating causes mixed effects on both the price range and volatility.

The behavior of exchange rates after sovereign credit rating announcements have been examined by several studies. Nelson (1991) identified that exchange rates response immediately and heavily to news and become volatile. Moreover, Jansen and Haan (2005) found an asymmetric behavior of exchange rate to positive or negative credit rating announcements. Credit Rating measures of the top three rating agencies, namely; Moody's, Standard & Poor's, and Fitch, known as the 'Big Three' is presented in Table 1.

5.2.2 Selected News Sources

For sentiment analysis, HTTP REST API has extracted daily news articles regarding to official macroeconomic statements and published news that reflects the expectations and opinions of economists from leading news websites Reuters, Bloomberg, The Guardian, and The New York Times. Over eighty news articles have been analyzed, and textual content transformed into numeric sentiment inputs.

5.2.3 Exchange Rate Data, and Reasoning Behind Variations

To perceive the degree of responsiveness of exchange rate to recently published macroeconomic news; daily price, daily change in percentage, opening price, highest price, lowest price, and difference between highest and lowest prices of USD/TRY be investigated for each day. Especially opening price, close price, highest price, lowest price, and the difference between these prices plays a crucial role on overcoming any possible accuracy error on the ML algorithm caused by the coincidence of news content on opposite polarities that may take place on the same day. The accommodation of news on opposite polarities on the same day may neutralize the significant effect of vital news, and cause a dramatic error on further calculations of the algorithm. To cope with this issue, the philosophy behind the candlestick chart has been adopted. In this case, even if neutralization happened in any day during the selected date interval, algorithm will detect it by observing the volatility of the exchange rate price throughout the day, and weight related news accordingly to match the volatility and the price by considering previous/further news on both subjects.

5.2.4 Date Interval for Data Collection

Historical data for macroeconomic indicators and USD/TRY exchange rate have been collected. Year to year, and month-to-month changes in percentage of inflation rate, interest rate in percentage and its difference in percentage compared to last month, unemployment rate in percentage, and its difference in percentage compared to last month, exports of Turkey, imports of Turkey, balance of trade of Turkey, and change compared to last month, and announced credit rating of Turkey, and the difference of the rating compared to previous rating is collected as data for macroeconomic indicators. For exchange rate data; USD/TRY exchange rate prices; daily price, daily change in percentage, opening price, highest price, lowest price, and difference between highest and lowest prices have been collected. Mentioned macroeconomic and exchange rate data have been collected over the period between 29 December 2017 and 01 November 2019. Collected macroeconomic and USD/TRY exchange rate data is equivalent to 4239 variables.

5.3 Methodology

This chapter adopted an advanced methodology to comprehend the correlation between recently published macroeconomic news and exchange rate volatility, pursuing to predict exchange rates by evaluating the sentiment of recently published news articles. To do so, methodology is constructed as three phases, which are “Sentiment Analysis”, “Fuzzy Logic”, and “Support Vector Regression” respectively. In depth methodology for aforementioned phases has been presented in Chapter 4. Related sections and subsections for relevant phases are going to be referenced below. Throughout the first two phases, methodology is going to address each macroeconomic indicator individually. The main purpose is to evaluate the significance and the correlation of recently published news regarding to each indicator independently, alongside to enhance understanding of the influence of recently published news on particular macroeconomic indicator on decision makers’, and behavior of decision makers under the specific environment that causes fluctuation on exchange rate. After achieving the outcome of the second phase individually for each macroeconomic indicator, second phase will be repeated, following by the third phase for all indicators collectively.

First phase “Sentiment Analysis” (Chapter 4.2), practices components of Natural Language Processing, namely; Lexicon (Chapter 4.2.1), and Learning Automata (Chapter 4.2.2) for textual analysis, which reinforced by Prospect Theory (Chapter 4.2.3). Outcome of Sentiment Analysis is going to be examined by the second phase, “Fuzzy Logic” that adopts fuzzy logic (Chapter 4.3) algorithm from machine learning, and going to achieve a correlation value between each political event and exchange rate. Final phase, “Support Vector Regression” (Chapter 4.4.1) is going to run support vector regression. Monte Carlo Cross Validation (Chapter 4.5.1.1) is going to be used to split training and test data. The linearity of the provided dataset is

going to be examined by Principal Component Analysis (PCA), Principal Component Regression (PCR), and Partial Least Square (PLS) methods, and kernel function for the model is going to be assigned according to the linearity of the dataset. The performances of Polynomial, and Gaussian Radial Basis kernel functions are going to be compared, and the one that provides the most accurate hyperparameter selection is going to be adopted for the model. Hyperparameters for selected kernel function; Gaussian Radial Basis Function is going to be obtained by Grid Search with Cross Validation method (Chapter 4.5.2.1). Grid Search with Cross Validation method plays a crucial role on fine-tuning the hyperparameters for the kernel function cautiously and suitably to provide most accurate hyperparameter selection, while eliminating the heteroscedasticity and multicollinearity in regression analysis. The accuracy of the obtained parameters is going to be validated by Accuracy, Precision, Recall, and F1 Score metrics (Chapter 4.5.2.2). The performance of the constructed model is going to be evaluated by the statistical tests, V Measure Score (VMS), Explained Variance Score (EVS), Mean Squared Error (MSE), Root Mean Squared Error (RMSE), Mean Absolute Error (MAE), Mean Absolute Percentage Error (MAPE), and Coefficient of Determination (R^2) (Chapter 4.6), and future exchange rates are going to be predicted.
The workflow mechanism of the presented methodology is going to be detailed in the following subsection.

5.3.1 Workflow Mechanism of Methodology

Methodology has three main phases, namely; sentiment analysis, fuzzy logic, and support vector regression. First phase, sentiment analysis starts with laying the foundation of textual analysis. To do so, every published news articles related to macroeconomic indicators and decisions on economic policies are going to be collected from the leading news publishers alongside with authorities' publishing services that have been mentioned in Section 2. Extracted titles and annotated text of articles constitutes the database for Sentiment Analysis. Annotated text of the extracted news are going to be feed into Learning Automata to compose a Lexicon dedicated predominantly for macroeconomic indicators, which is going to be authorized by the word polarity functions of existing lexicons dedicated to economics. Constructed Lexicon extracts sentiments that annotated texts encompasses by investigating each word involved in the text, and produce an overall polarity vector value for every annotated text feed into algorithm. Attained polarity vector values are going to be evaluated by the historical macroeconomic data alongside to daily price, daily change in percentage, opening price, highest price, lowest price, and difference between highest and lowest prices of USD/TRY to obtain a Prospect value that defines the subjective feelings and valuation of the decision maker on published news for each macroeconomic indicator. Especially opening price, close price, highest price, lowest price, and the difference between these prices plays a crucial role on overcoming any possible accuracy error on the machine learning algorithm caused by the coincidence of news content on opposite polarities that may take place on the same day. The outcome of the sentiment analysis is going to characterize the fuzzy antecedents of second phase, fuzzy logic. Obtained prospect values for each macroeconomic indicator are also designated as fuzzy antecedents under title of every macroeconomic indicator, namely; interest rate, inflation rate, balance of trade, unemployment rate, and credit ratings. There are three universal outcomes on macroeconomic news and policy announcements directly perceived by decision maker, and also have a direct psychological impact on the subjective feelings and decision making process, which are; increase, steady, and decrease. These three universal outcomes assigned as membership functions for fuzzy antecedents. The transition between the memberships of the determined sets for each fuzzy antecedent, aside from the degrees of truth for each defined set that have been mentioned above are going to be defined by the upper and lower bounds, width, concave, convex, and the steepness of the associated membership, which will be presented on the future graphs. The fuzzy relations are going to be obtained by the presence or absence of association, interaction, or interconnectedness between the elements of two or more sets. As interpretations of news are personal and subjective, decision makers may also interpret outcomes of macroeconomic announcements subjectively. For that reason, 245 fuzzy rules have designated by the author and feed into fuzzy engine to generate fuzzy consequent. Fuzzy consequent is principally a dependency rate of USD/TRY on recently published macroeconomic news. In other words, fuzzy consequent signifies the correlation between USD/TRY and news. Obtained correlation value is used as input alongside to historical USD/TRY and macroeconomic data to run support vector regression. The linearity of the provided dataset is going to be examined by Principal Component Analysis (PCA), Principal Component Regression (PCR), and Partial Least Square (PLS) methods. PCA/PCR is a mono-dimensional, observed variable approach that aims to find a linear correlation by reducing dimensionality. PLS is a multidimensional, latent variable approach that suits best for when there is a multicollinearity. Both PCA/PCR and PLS methods can be used to tests outliers and sensitivity analysis to provide diagnostic tools for the model. Gaussian Radial Basis Function utilized as kernel function to project nonlinear historical data into high-dimensional space. Constructed hyperplane provides adjustability to converge optimal set of weights. Gaussian Radial Basis Function utilized as kernel function due to its good fit on non-linear inputs (Grigoryan, 2016). Gaussian Radial Basis Function tuned by hyperparameters that have been attained by grid search with cross validation method, which also eliminated the possibility of heteroscedasticity and multicollinearity. Accuracy of the parameters attained by the grid search with cross validation is assessed by accuracy, precision, recall, and f1-score metrics. Monte-Carlo cross-validation method has been adopted to split presented data into training and test groups for prediction. Accuracy of the prediction investigated by the statistical tests. Prediction consistency of the model is evaluated against seven widely used statistical tests for machine learning algorithms, namely, V Measure Score (VMS), Explained Variance Score (EVS), Mean Squared Error (MSE), Root Mean Squared Error (RMSE), Mean Absolute Error (MAE), Mean Absolute Percentage Error (MAPE), and Coefficient of Determination (R^2). VMS is a decisive measure to begin with, as it indicates clustering performance of the model. To perform an accurate prediction with machine learning algorithms, superior performance on homogeneity and completeness on clustering task is a necessity to achieve an accurate regression without multicollinearity issue. Algorithm is appropriate to run further statistical tests and accurate prediction only when satisfying VMS value is achieved, as it is an indicator of a reliable model. Grade of achieved VMS value provides insight on the prejudices of measured deviation of errors. MSE, RMSE, MAE, and MAPE are calculated in order to measure the deviation between predicted and actual values. RMSE and MAE are calculated simultaneously to identify possible deviation in the errors. While equivalent RMSE and MAE values indicates similarity of magnitude on all errors, the difference between the results specifies the variance associated with the individual errors. Greater difference indicates higher variance, and occurrence dissemination of the specific errors. The diversity of RMSE and MAE results denotes of underfitting and/or overfitting issues. Furthermore, MAPE is also considered and calculated to normalize the absolute error and obtain altered perspective on association with actual values. The homoscedasticty can be evaluated by EVS result, as it evaluates the explainability of the variance by the factors presented in the actual data. Additionally, R^2 and EVS targets inconsistency between predicted and actual value. However, while R^2 uses raw sum of squares, Explained Variance Score uses the biased variance. Both variables can only be similar if the mean of the residuals equals to zero, and prediction is unbiased.

Conclusively, if the statistical tests' results are convenient, exchange rates are going to be predicted in daily basis by considering the impact of published news. The empirical workflow that described above has presented as a flowchart in Figure 26, which has been located in appendix section.

5.4 Empirical Findings of the Hybrid Machine Learning Algorithms

The main objective of the sentiment analysis is to derive subjective perspective of decision makers on particular events that took place on news. To do so, unique lexicon for macroeconomics is built by Learning automata, which supplemented by the polarity functions of existing lexicons on economics. Probability of sentiment outcome is going to be extracted for each published news article. Prospect theory enhanced the outcome by observing the historical exchange rate prices. Subjective feelings of the decision makers on variety of scenarios on macroeconomic indicators have aimed to be reached by this process to understand the valuation behavior of decision makers on particular scenarios. This procedure provided an enhanced overlook on the content of the news, perspective of decision makers on each scenario, and fuzzy antecedents for the next step.

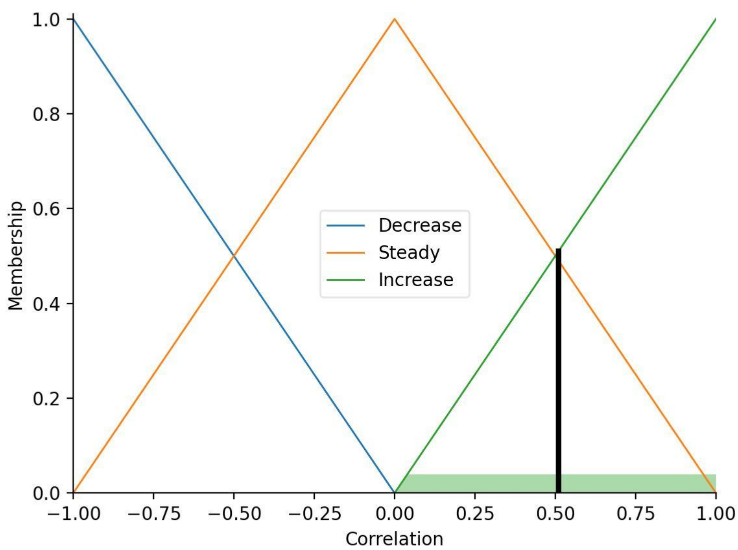
Sentiment analysis has produced sentiment scores from 1 to 5 for every analyzed news; where the value 5 defines substantial appreciation of USD/TRY, values between 5 and 3 represent diminishing appreciation of USD/TRY, value 3 defines steady state, values between 3 and 1 represents gradual depreciation of USD/TRY, and 1 defines very strong depreciation of USD/TRY. Obtained dataset as the output of sentiment analysis have not feed into fuzzy logic algorithm directly, but separated and allocated under the title of related macroeconomic variable to form fuzzy antecedents.

Due to acute effects of change in macroeconomic variables, trapezoidal membership function has selected to be used. For each fuzzy antecedent, membership functions, namely; ‘increase’, ‘steady’, and ‘decrease’ that represent appreciation, steady state, and depreciation of Macroeconomic variables respectively, which have been assigned to determine the degree of belonging to a fuzzy set based on outcome of the sentiment analysis.

Fuzzy engine determined that inflation rate, interest rate, and unemployment rate are positively correlated with USD/TRY, while balance of trade, and credit ratings are inversely correlated with USD/TRY. For that reason, values that denote the degree of belonging alter. For inflation rate, interest rate, and unemployment rate; values from 1 to 3 denotes the degree of belonging to fuzzy set ‘increase’, value 3 defines the belonging to fuzzy set ‘steady’, and values from 3 to 5 denotes the degree of belonging to fuzzy set ‘decrease’. Meanwhile, for balance of trade, and credit ratings; values from 1 to 3 denotes the degree of belonging to fuzzy set ‘decrease’, value 3 defines the belonging to fuzzy set ‘steady’, and values from 3 to 5 denotes the degree of belonging to fuzzy set ‘increase’. Degree of membership to a fuzzy set denoted between 0 and 1, While 0 and 1 represents 0% and 100% membership degree respectively, anything in between considered as a partial membership to belonging set. As each macroeconomic indicator has unique dynamics, significance and influence on exchange rate, the membership function of fuzzy logic algorithm is diversified and constructed specifically for each indicator, while fuzzy rules comprises macroeconomic indicators collectively, and function regarding to their weights. Dependency of USD/TRY on inflation rate, interest rate, balance of trade, unemployment rate, and credit ratings have presented in Figures 27, 28, 29, 30, and 31 respectively in the appendix section.

Fuzzy consequents for each macroeconomic indicator alongside the consequent for combination of macroeconomic fundamentals have obtained. Fuzzy consequents provided the correlation value, which is necessary to run prediction by using support vector regression. Obtained correlation values for each macroeconomic indicator specify that; published news regarding the increase in inflation rate of Turkey, and downgrading credit rating of Turkey have significant impact on the appreciation of USD/TRY. Besides, published news on increase in unemployment rate of Turkey has insignificant impact on the appreciation of Turkey. On the other hand, the impact of published news regarding the decrease on interest rate in Turkey on depreciation of USD/TRY is inevitably significant. Meanwhile, published news on increase in balance of trade of Turkey has insignificant impact on the depreciation of USD/TRY. In the light of the correlation values of macroeconomic news on each indicator, the impulsive dependency of exchange rate on macroeconomic news has been simulated, and presented in Figure 32.

Fig. 32. Correlation between Macroeconomic News and USD/TRY

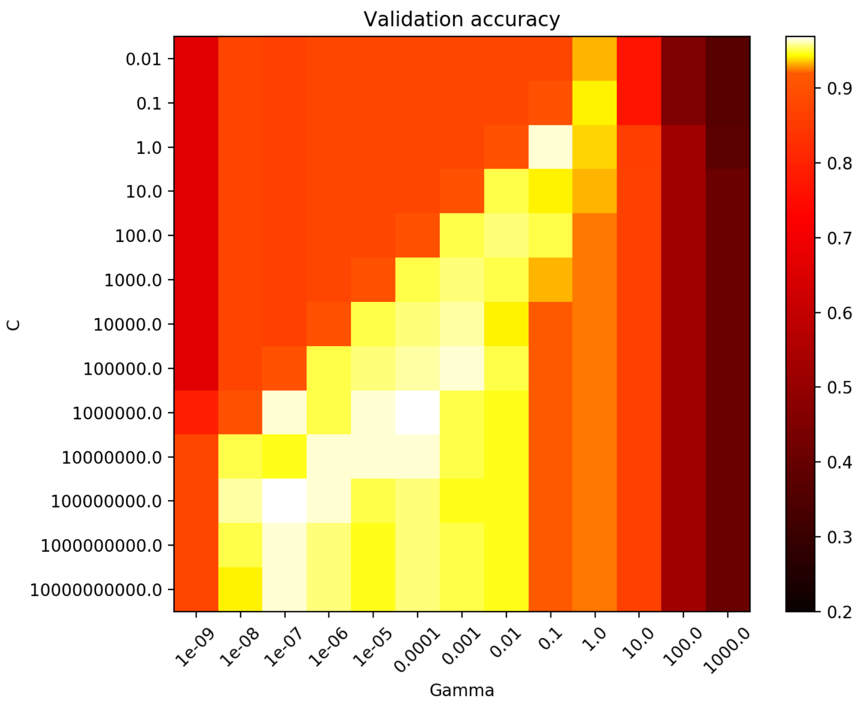


Simulation result presents the significant impulsive dependency of USD/TRY on recently published macroeconomic news. The simulation result verifies and emphasizes the importance and significance of consideration of published macroeconomic news on understanding the fluctuations of exchange rate.

Correlation values that have been obtained as fuzzy consequents used as input in addition to historical USD/TRY data and historical data of each macroeconomic indicator to run support vector regression. As a final step, USD/TRY prices are going to be predicted in the light of consequents achieved during sentiment analysis, and fuzzy logic processes regarding to the correlation between macroeconomic news and exchange rate prices. The linearity of the provided dataset is evaluated by PCA/PCR and PLS methods. Figure 33, which presented in the appendix section illustrates PCA/PCR and PLS results. First principal component in PCA/PCR and PLS indicates the axis in the K-dimensional variable space that accommodates the largest variance of the samples in a direction, while second principal component is orthogonal axis to the first principal component in the K-dimensional variable space, and identifies the second largest source of samples in a direction that can be denoted as the residual variance of the samples that improves the approximation. Moreover, the direction that provides the lowest variance is aimed to be captured by the PLS. The nonlinearity of the provided dataset has confirmed by the tests applied. For that reason, kernel function for the SVR analysis is going to be assigned accordingly in order to prevent heteroscedasticity and multicollinearity. The application of Gaussian Radial Basis Function on PCA and the projection difference is exemplified as Figure 35 in the appendix section.

Nonlinear historical data have been projected into high-dimensional space by using the kernel function. Constructed hyperplane provides adjustability to converge optimal set of weights (Oliveira et al., 2017). Gaussian Radial Basis Function utilized as kernel function due to its good fit on non-linear inputs (Grigoryan, 2016). Selected kernel function tuned by $C, \varepsilon, \text{and}\, \gamma$ hyperparameters. In order to set the hyperparametes, a grid search with cross validation has been adopted to eliminate the possibility of multicollinearity and achieve highest accuracy possible. Grid search ranges of $C, \varepsilon, \text{and}\, \gamma$ hyperparameters assigned as suggested by Hsu, Chang, and Lin (2003). *C* ranges from 1 to 100000, and γ from 2^{-10} to 2^{11}. Grid search with cross validation for hyperparameter $C, \text{and}\ \gamma$ value selection is presented as Table 2, and layers view for visualization of grid search cross validation accuracy is presented as Figure 36 in the appendix section. Grid search cross validation accuracy heat map is also presented as Figure 37.

Fig. 37. Grid Search Cross Validation Accuracy Heat Map for Hyperparameter Selection



Hyperparameter ε is assigned by considering the highest change in the historical data, and determined as high as possible in order to eliminate any possible overfitting issue. Accuracy of the parameter selection using grid search with cross-validation is evaluated by consideration of three metrics, namely; precision, recall, and f1-score. Results have indicated the accuracy of 0.96 for parameter selection, and detailed in Table 3, which has been presented below.

Tab. 3 Accuracy of the Parameter Estimation Using Grid Search with Cross-Validation for Macroeconomic News

The model is trained on the full development set.
The scores are computed on the full evaluation set.

	precision	recall	f1-score	support
0	1.00	1.00	1.00	21
1	0.93	0.97	0.94	30
2	0.97	0.96	0.96	24
Accuracy			0.96	75
Macro Avg	0.95	0.98	0.95	75
Weighted Avg	0.95	0.97	0.95	75

The critical values of Accuracy, Precision, Recall, and F1 Score are 0.20, 0.40, 0.60, and 0.80 denotes slight, fair, moderate, substantial, and almost perfect agreement respectively.

Figure 38, 39, and 40 in the appendix section present the historical USD/TRY data associated to macroeconomic news mirrored on to the constructed hyperplane. In the mentioned historical USD/TRY figures, the colors of data points are fading from red to yellow. Color red indicates the low price range, and yellow indicates the high price range. By observing the displacements of color-coded data points, it is possible to comprehend how nonlinear data points forms and/or separated linearly in the constructed hyperplane. Figures captured from different angles once achieving the optimal parameters of selected kernel function. Macroeconomic indicators provided restricted historical data due to their nature. Therefore, splitting presented data into training and test groups is essential to appraise the performance of the regression model. To do so, Monte-Carlo cross-validation (MCCV) method adopted. As the provided historical macroeconomics data is relatively restricted, 85% of provided data used as train data, while 15% reserved as test data. MCCV randomly splits reserved portion of data into sub-samples and assign them as test sets. The process of selecting random independent partitions repeated for multiple times. Prediction consistency of the model is evaluated against seven key statistical metrics, which are widely be accepted and using to validate machine learning prediction performance, namely; V Measure Score (VMS), Explained Variance Score (EVS), Mean Squared Error (MSE), Root Mean Squared Error (RMSE), Mean Absolute Error (MAE), Mean Absolute Percentage Error (MAPE), and Coefficient of Determination (R^2).

Primarily, VMS is a crucial measure to check, and attain an overview on clustering performance of the model. Homogeneity and completeness of clustering task plays a significant role on achieving higher accuracy of prediction on machine learning models. If VMS value satisfies the expectations, model is suitable to run further statistical methods to measure the deviation of errors, and inconsistency of predicted values. To measure the deviation between predicted and actual values, MSE, RMSE, MAE, and MAPE are calculated. RMSE and MAE are calculated simultaneously to identify possible deviation in the errors. While equivalent RMSE and MAE values indicates similarity of magnitude on all errors, the difference between the results specifies the variance associated with the individual errors. Greater difference indicates higher variance, and occurrence dissemination of the specific errors. Additionally, MAPE is also considered and calculated to normalize the absolute error and obtain altered perspective on association with actual values. While MSE, RMSE, MAE, and MAPE measure the deviation between predicted and actual value, R^2 and EVS targets inconsistency between predicted and actual value. However, while R^2 uses raw sum of squares, Explained Variance Score uses the biased variance. Both variables can only be similar if the mean of the residuals equals to zero, and prediction is unbiased.

Monte-Carlo cross-validation run numerous times, and statistical test results attained for each run. Achieved results for statistical test are presented exhaustively in Table 4 in the appendix section. Then, obtained statistical test results averaged. Averaged results for MCCV statistical test results are presented in Table 5.

Table 5. Averaged Monte Carlo Cross Validation Statistical Test Results.

VMS	0.919392***
EVS	0.928443***
MSE	0.050812
RMSE	0.225206**
MAE	0.192386
MAPE	0.039060***
$\boldsymbol{R^2}$	0.924597***

† Exhaustive Monte-Carlo cross validation statistical test results are presented in the appendix section, as Table 4.

The critical values of VMS are 0.70, 0.80, and 0.90. Therefore, *, **, and *** denotes threshold, significant, and highly significant at 0.70, 0.80, and 0.90 respectively.

The critical values of EVS are 0.60, 0.80, and 0.90. Therefore, *, **, and *** denotes threshold, significant, and highly significant at 0.60, 0.80, and 0.90 respectively.

The critical values of RMSE are set by SI at $\leq 10\%$, and $\leq 5\%$, which indicate RMSE value of 0.520908, and 0.260454. Therefore, *, and ** indicates the threshold and significance at respective critical points.

The critical values of the difference between RMSE and MAE at $\leq 5\%$, and ≤1% denotes acceptable and perfect fit of the model respectively.

The critical values of MAPE are 0.25, 0.10, and 0.05. Therefore, *, **, and *** denotes threshold, low but acceptable, and highly acceptable accuracy at 0.25, 0.10, and 0.05 respectively.

The critical values of R^2 are 0.50, 0.75, and 0.90. Therefore, *, **, and *** denotes weak, moderate, and substantial prediction at 0.50, 0.75, and 0.90 respectively.

The critical value of the difference between EVS and R^2 is $\leq 2\%$. Therefore * denotes acceptable bias at $\leq 2\%$.

In purpose of validating the results, alongside of measuring the sustainability of the results, the Monte Carlo Cross Validation has repeated 40 times, the mean of achieved statistical test results has been taken and presented above. The mean of MCCV statistical test results indicates that;

VMS result is 0.919392, which is higher than 0.90 criteria. Therefore, it is possible to denote VMS value as highly significant thus indicates a harmonic mean between homogeneity and completeness and indicates an accomplished clustering task for the future weight support vector regression. Additionally, significant VMS result, 0.919392 indicates no multicollinearity issue, and authorizes the model for further tests and regression analysis.

EVS result is 0.928443, which is higher than 0.90 criteria. Therefore, it is possible to signify EVS value as highly significant. The results indicate that the variance can be explained by the factors presented by the actual data. Moreover, EVS is another measure that is crucial for regression analysis, as it scores homoscedasticity. Highly significant EVS result, 0.928443 denotes that the variance can be explained by the factors presented by the actual data, and not heteroscedastic.

MSE result is 0.050812. As it has been emphasized on the Chapter 4, MSE closer to 0 indicates better fit of regression line, as it indicates the variance of residuals. However, as there isn’t acceptable range set for MSE as it does not shares the same unit as the original values, MSE has been used to achieve RMSE values.

RMSE value is in the same unit with the original data, and indicates the standard deviation of errors emerged during prediction, thus signifies the accuracy of the model. RMSE result is 0.225206, which denotes that the standard deviation of error is 0.225206 when compared to actual USD/TRY value. In order to evaluate RMSE result, the value transformed into interpretable value, SI. Transformed RSME, SI value is 0.043233, which is lower than 0.05 criteria. Therefore it is possible to denote RMSE value as significant.

MAE is relatively insensitive to outliers compared to RMSE, while RMSE magnifies the bigger errors and ignoring the smaller errors. For that reason, MAE may accommodate bias, while RMSE is related to variance. When MAE and RMSE compared, it is possible to observe underfitting and overfitting issues that have been targeted by the application of MCCV method. Underfitting and overfitting is the situation of high bias, low variance, and low bias, high variance respectively. The difference between RMSE and MAE results, 0.225206 and 0.192386 respectively is 17.05%, which is higher than the 5% criteria. Underfitting issue is detected, as the difference is higher than the criteria. This indicates that the model was deficient on detecting the relationship between actual and predicted values accurately.

The accuracy of the prediction has also been measured by the MAPE results. MAPE results, 0.039060 that is equivalent to 3.90%, which is lower than the 0.05 criteria, equivalent to 5%. Therefore, it is possible to say that the model has highly acceptable accuracy.

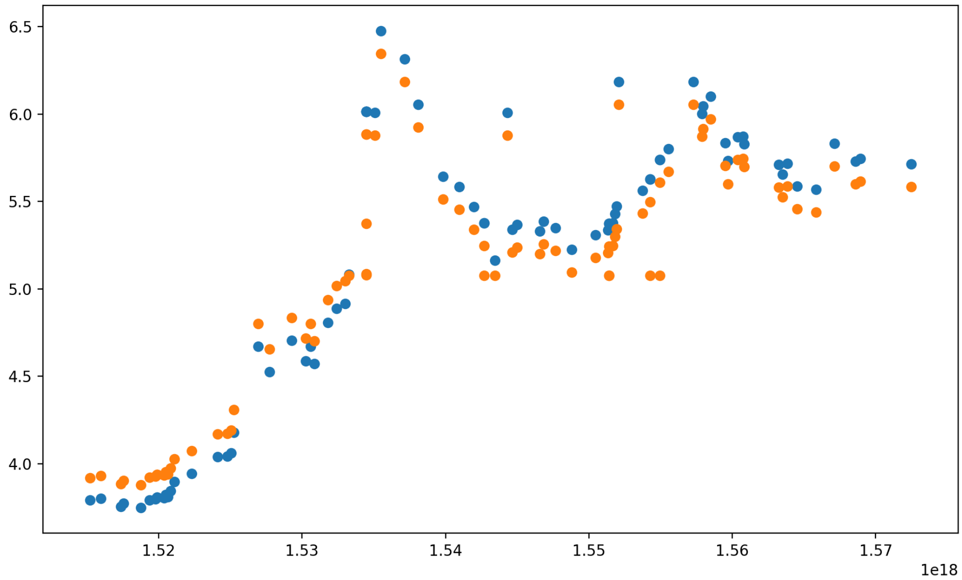
R^2 result is 0.924597, which is higher than the 0.90 criteria. Highly significant results have verified that the prediction is substantial with a superior accuracy, as the highly significant majority of the predicted data points are on the regression line.

Moreover, the deviation between EVS and R^2 results, 0.928443 and 0.924597 respectively is only 0.42%, which is lower than 2% criteria. Thus the deviation between EVS and R^2 results signifies that the prediction is unbiased.

When evaluated in daily manners, macroeconomic news may react deficiently on comprehending the relationship between actual and predicted data, due to lack of continuity of macroeconomic news taking place in the media. In order to understand the reasoning behind the detected deficiency on comprehending the relationship between actual and predicted data, and figuring out a solution for the attained reason, a detailed observation has been made. After detailed observation on the historical data and events, author anticipates the possibility of variance of error to, and underfitting model to be caused by ongoing political tension and uncertainty between United States and Turkey that triggers the daily fluctuations massively, as the affected date interval for deviation intersects with ongoing political events between mentioned countries. As the model is concentrated on evaluating and predicting exchange rate in daily manners, the possibility of failing to notice important fluctuations caused by the political tension alongside to macroeconomic news is highly possible. This allegation is going to be investigated in the further chapter.

Consequently, reputable EVS, MSE, RMSE, MAE, MAPE and R^2 results revealed the importance and usability of recently published macroeconomic news on understanding and predicting the fluctuations on exchange rate prices. Prediction results for USD/TRY prices visually presented in Figure 41. Actual USD/TRY prices presented as blue dots, while predicted prices presented as orange dots.

Fig. 41. Actual versus predicted USD/TRY prices obtained by analyzing recently published news and statements regarding on macroeconomic indicators.



**Blue Dots denote Actual USD/TRY Prices

**Orange Dots denote Predicted USD/TRY Prices

5.5 Conclusion

This chapter developed a robust methodology and a novel approach that allows evaluating the significance of recently published macroeconomic news, their causality on fluctuating exchange rates in daily basis, and usefulness to predict exchange rates. To perceive the causality of recently published macroeconomic news on exchange rate fluctuations, published news on inflation rate, unemployment rate, interest rate, balance of trade, and credit rating have been evaluated alongside the historical data. For this purpose; related news and historical data have been collected to form the necessary database that is going to be examined throughout the interconnected phases of constructed hybrid machine learning methodology. The hybrid machine learning model encapsulate learning automata for textual analysis, prospect theory for enhancing textual analysis scores by associating with the human behavior by examining the change in exchange rate prices and generate fuzzy inputs, fuzzy logic for understanding the degrees of dependency of exchange rate fluctuations on macroeconomic news regarding to each macroeconomic indicator, and support vector regression for predicting exchange rate by using variables obtained by previous phases of methodology alongside to historical data. Simulations on news articles, subjective feelings of the decision makers, and historical prices indicated a significant correlation between exchange rate and the news regarding to inflation rate, interest rate, and credit ratings, while insignificant correlation detected for balance of trade, and unemployment rate. By considering the correlation values obtained on fuzzy logic phase, support vector regression analysis has taken place as a final step. A successful and a strong prediction on USD/TRY exchange rate have obtained on the final phase.

Detailed fuzzy logic analysis indicates that exchange rate is highly responsive to breaking news on inflation rate, interest rate, and credit ratings. On the other hand, similar responsiveness of exchange rate has not been observed to breaking news on unemployment rate and balance of trade.

The result of this chapter specifically verifies that; recently published macroeconomic news based on stories such as sights and thoughts on current macroeconomic conditions, expectations on upcoming rates, changes on policies, and official announcements regarding to macroeconomic indicators plays a vital role and have a significant immediate impact on exchange rates right after the publication. Consequently, this chapter evidently revealed that; not only the objective quantitative variables, but also subjective qualitative information can be used to obtain enhanced prediction performance by examining decision makers' behavior on particular scenarios.

Furthermore, even if EVS, R^2, MSE, and MAPE scores indicated a valid and accurate prediction, anomalies observed on the variance between RMSE and MAE. The variance specifies the necessity of further investigation to uncover the reasoning underneath. In the light of detailed examination on the dataset and extended research, author anticipates that; variance occurred between RMSE and MAE caused neither by the macroeconomic fluctuations, biased measurements, poor fit of the model, nor the related qualitative or quantitative data. In fact, it believed to be caused by the external factors other than economics, such as political tension and uncertainty that took place during particular date intervals. Matching date interval of ongoing political tension and uncertainty between United States and Turkey with the deviated variance errors reinforces this hypothesis. Political tension and uncertainty should deeply be investigated for better understanding, and possibly achieving superior statistical results, and achieving a model that can take prediction accuracy and methodology even further. For that reason, the impact of recently published news regarding to political tension and uncertainty on exchange rate is going to be examined in the next chapter.

CHAPTER 6
Impact of Particular Political Event Categories on Exchange Rate

6.1 Introduction

In the era of technology, accessibility of information widen the research areas while augmenting the understanding of cause and effects of various circumstances corresponding to political instability, uncertainty triggered by unexpected political and economic events that may mainly be categorized such as; elections, referendum, national and international political tensions, embargos, protests, blasts, etc. Attributable to its substantial role on the risk management, international trade, and investment determination (Sharma et al, 2016; Korol, 2014), economic condition and policy of a country is highly dependent on exchange rates (Kandil et al., 2007; Cehreli et al., 2017). Unlike other economic indicators, exchange rate is highly volatile and substantially information dependent. Any kind of event-based information that may influence the investors' sentiments has a direct impact on exchange rate prediction and expectation (Yasir et al., 2019).

The speeches that provides opinions or official statements regarding to a occurred political event that accommodates validation of political tension, which presented by political figures has a direct impact on public's political sentiment (He et al., 2017). Average person may possibly be emotionally biased while evaluating political tension, or may completely be out of his/her interest, but financial markets adopt unbiased approach and highly responsive to any variability on political tension. Besides, due to possible arise of potential wars, economic sanctions, tariffs, and boycotts, firms are also heterogeneous in their attitudes (He, et al., 2017). Studies support that fading political harmony has a direct adverse impact on economic activities (Besley and Mueller, 2012), which distresses corporate investment cycles (Julio and Yook, 2012), cross-border trading and capital flows (Julio and Yook, 2016), and valuation of publicly listed firms (Pastor and Veronesi, 2013). Even when there isn't any political instability and uncertainty, financial markets still reacts on negative news by interpreting it as a trigger point of a forthcoming instability (Veronesi, 1999). Moreover, instability and uncertainty on financial markets triggers economic policy uncertainty, which influences macroeconomic variables (Baker et al., 2012).

Adverse impact of the military conflict on economics and the stock market of the involved countries is broadly studied and accepted by the literature as it instigates deterioration of expected profitability which has a direct adverse impact on both firms' profitability, valuation of investors', and fundamental macroeconomic variables as a result of a chain reaction (Abadie and Gardeazabal, 2003; Amihud and Wohl, 2004; Rigobon and Sack, 2005). Although, literature indicates that uncertain political environment that has been triggered by various political events, which have a potential of affecting political harmony and international relations, causing ambiguous policies, and increasing political tension causes a negative impact on economic activities, even if it would not cause any hypothetical military conflict (He et al., 2017). Earlier researches indicate that economic volatilities are highly derived by political uncertainty (Bloom, 2014). Furthermore, political uncertainty and instability such as geopolitical conflicts, terrorism and immigrant crisis as the main reason for slow global economic growth over past decade (Gholipour, 2019).

Comprehensive previous studies efficiently introduced the strong negative impacts of collapsing government, regime instability, political polarization, and government repression on the tendency of macroeconomic fundamentals (Alesina et al., 1996; Chen and Feng, 1996; Jong-a-Pin, 2009). Additionally, the degree of instability and uncertainty on political regime plays is a significantly influential factor on macroeconomic variables in a short run (Aisen and Veiga, 2013; Campos and Nugent, 2002), as socio-political instability instigates politic-economic uncertainty, which causes a decline on the volume of investments due to surging risks (Alesina and Perotti, 1996; Darby et al., 2004). For that reason, it is possible to advocate that; political crisis in any variety has a direct negative impact on financial markets and economics (Berkman et al., 2011).

Consequently, consideration of subjective and qualitative information that has been introduced above while predicting the volatility of exchange rates is critical for the minimization of the margin for error.

This chapter concentrates on the conceiving the potential impact of particular political event categories comprises and/or arose in Turkey on USD/TRY exchange rate between the dates 29 December 2017 and 01 November 2019. During stated date interval, plentiful political event categories regarding to Turkey's national and international policy have taken place in frequent manners. Diversity of those political event categories provides a unique opportunity to examine the impact of each political event category involving domestic and foreign policy individually on the volatility of exchange rate in short run. This chapter predominantly concentrated on four main political events, namely; Pastor Andrew Brunson case, Parliamentary and Presidential elections, S-400 crisis, and eventful Istanbul Mayoral elections. Selection of political events have meticulously done and homogenously distributed as national and international. Additionally, unlike traditional uniform categorization of political events, selected events are members of various political categories, such as; judicial system, international politics, international relations, political tension, political sanctions on government officials, financial sanctions, warfare, defense industry, NATO, human rights, elections, cabinet changes, and democratic uncertainty. Multiple memberships of each political event provide a unique opportunity to study, and observe multiple categories individually under the name of single event. Besides, both of aforementioned international political events that triggered political tension, instability and uncertainty had occurred between United States of America and Turkey, which results in abrupt dramatic fluctuations in

USD/TRY exchange rate. That is why; USD/TRY exchange rate is going to be examined to comprehend the effects of political events on exchange rate.

In this chapter, the effectiveness of officials' statements regarding to politics and international relations, publicly published news on aforementioned political cases and events, and their degree of certainty and consensus on exchange rate are going to be investigated. The chapter will especially concentrate on the aforementioned political events. Each event is going to be investigated individually in daily manners. Investigating in daily manners will provide an opportunity to categorize the official statements and published news under the title of each political category while comprehending how exchange rate behaves under the influence of each political category, as well as political event bodily, and the degree of certainty and consensus exclusively. The impact of each political case and event on exchange rate is going to be observed by examining the related news articles, while combination all varieties of political events will be examined to propound the significance of political news on exchange rate.

The main purpose of this chapter is to bring a novel approach to exchange rate prediction by investigating the effects particular political event categories by examining recently published news, officials' statements, in addition to historical exchange rates. While doing so, the chapter also serves the purpose of creating a behavioral database for each political event categories derived from daily-published news and exchange rate fluctuations to prepare a foundation for complex algorithms that are aiming to deal with much boarder spectrum of national and international political events under the extensive degree of certainty and consensus. To achieve this purpose, a hybrid machine-learning algorithm has generated by amalgamating natural language processing (NLP), fuzzy logic (FL), and supports vector machines (SVM) to create a database for further analysis, and to unveil the significance of political events on exchange rates fluctuations, furthermore predict the price of the exchange rates on a daily basis under the light of generated algorithms. The chapter is structured as follows: Section 2 introduces main political events that had been examined, Section 3 briefly describes data collection and fundamental variables; Section 4 describes application of methodologies, Section 5 presents empirical results, and prediction of USD/TRY exchange rate, and Section 6 concludes the chapter.

6.2 Political Cases

This section is dedicated to provide detailed information on four main political events that the chapter is predominantly concentrated on, namely; Pastor Andrew Brunson Case, 2018 Parliamentary and Presidential Elections, S-400 Crisis, and 2019 Istanbul Mayoral Elections. Selection of political events has meticulously done in order to distribute national and international events homogenously so as to validate the significance of national and international events comprehensively.

6.2.1 Pastor Andrew Brunson Case

Andrew Craig Brunson is an American-Hungarian pastor who had lived at Izmir, Turkey since the mid 1990s with his wife and three children. He was teaching at the Izmir Resurrection Church (BBC News, 2018; Time, 2018).

Pastor Brunson was accused for spying, and having links with the outlawed Kurdistan Workers' Party (PKK) and Gulenist movement that were blaming by the authorities for failed military coup in July 2016. He was arrested and sent to a detention facility by the Turkish authorities in October 2016 (BBC News, 2018; Time, 2018). Pastor Brunson has been sent to prison in December 2016, without any evidence of accusations against him (EPC, 2018).

While Pastor Brunson's case reaches back to 2016, case continued passively under the shadows of failed military coup in 2016, and constitutional referendum in 2017 until the successive court hearings in 2018.

Four court hearings occurred for Pastor Brunson's case on 16 April, 7 May, 18 July 2018, and 12 October 2018 respectively (USCIRF, 2018).

Meantime, political disagreement and uncertainty embitter between U.S. and Turkey over the case. U.S. was forcing Turkey to release Pastor Brunson, while Turkey was demanding the to exchange Pastor Brunson for the leader of Gulenist movement, Fetullah Gülen that resides in the U.S..

Debates between Turkish president Recep Tayyip Erdoğan, and United States president Donald Trump regarding to Pastor Brunson's case increased the tension between United States and Turkey, which were already overwrought. In August 2018, U.S. imposed political and financial sanctions on Turkey, which includes sanctions on top two Turkish government officials, and doubling the tariffs on Turkish steel to 20% and aluminum to 50% (The New York Times, 2018), while Turkey reacted back by doubling the tariffs on American cars to 120%, and alcoholic drinks to 140% (CBS News, 2018).

Brunson released from prison to house arrest in 25 July 2018 until 12 October hearing. Then, released from Turkish custody after his last hearing.

Political disagreements and uncertainty instigated during Pastor Brunson's case cause Turkish Lira to drop record low against the US Dollar by depreciating more than 40% since the beginning of the 2018 till the release date of Pastor Brunson, while Turkey's inflation rate dramatically increased due to uncertain economic and financial situation.

6.2.2 2018 Parliamentary and Presidential Elections

Turkish parliamentary elections and presidential elections held on the same day, 24 June 2018. Substantially, presidential election was scheduled for 3 November 2019. Following 2017 constitutional referendum that abolished the office of Prime Minister as well as the parliamentary system that replace it with presidential system, the leader of Nationalist Movement Party (MHP), Kemal Kılıçdaroğlu suggested in October 2017 for early presidential elections to accelerate the implementation process of presidential system, however officials did not respond to his suggestion (Yenicag Gazetesi, 2017). Subsequently, the leader of Nationalist Movement Party (MHP), Devlet Bahçeli suggested in April 2018 for early elections (Turkiye Gazetesi, 2018). Incumbent president Recep Tayyip Erdoğan accepted the suggestion of for early elections and announced early elections to be held on 24 June 2018, pursuing the purpose of disentangling economic and political uncertainty

(Hürriyet, 2018). Political parties formed official alliances for both Parliamentary and Presidential elections. While coalition of Justice and Development Party (AKP) and MHP formed People's Alliance, coalition of Republican People's Party (CHP), Felicity Party (SP), and Good Party (IYI) formed Nation Alliance. Alliance of Free Cause Party (HUDAPAR), and Peoples' Democratic Party (HDP) was also speculated; however alliance was rejected by HDP. People's Alliance nominated Recep Tayyip Erdoğan (Reuters, 2018), CHP nominated Muharrem Ince (NTV, 2018), HDP nominated Selehattin Demirtaş, and IYI nominated Meral Akşener (CNN Türk, 2017) as their candidates, besides to two more candidates of other parties for presidential election. Recep Tayyip Erdoğan and People's Alliance won presidential and parliamentary elections by taking 52.59% and 53.66% of the votes respectively, while their main rival Muharrem Ince and Nation's Alliance had 30.64%, and 33.94% of the votes respectively.

6.2.3 S-400 Crisis

The second largest army amongst the 29 member military alliance NATO belongs to Turkey. Moreover, Turkey's geographically strategic borders with Iran, Iraq, and Syria matter for U.S. considerably. Conversely, President Erdogan's rising authority retrograde relations with European Union (EU) and some of the NATO members since Erdogan's latest move on missile defense system in November 2016 (BBC News, 2019a). In order to empower its independent defense policy and enhancing airspace and counter threats protection against wars raged borders, Turkey signed a deal with Russia to buy Russia's cutting-edge long-range anti-aircraft defense system S-400 Triumf surface to air missile that costs $2.5 billion, instead of more expensive NATO version of the similar missile which is called SA-21 Growler. According to initial agreements, two battery shipments will be made from Russia to Turkey, and next two batteries will be produced in Turkey (TRT World, 2017; BBC News, 2019a).

Turkey's bold purchasing decision in a pursuit of independent defense policy, developing indigenous defense technology, and purpose of defending larger area deteriorate already tense relations with the EU and the U.S., furthermore cause accusations between sides. While accusations regarding Turkey to be violating human rights and acting undemocratic made by the U.S. and EU, the U.S. and EU accused by Turkey for supporting terror organizations Kurdistan Workers' Party (PKK) and its extension in Syria that is called YPG, and for supporting Fethullahist Terrorist Organization (FETO) (TRT World, 2017). Moreover, the U.S. has warned purchasing S-400 may result with imposing economic sanctions and excluding Turkey from F-35 program, as the U.S. will not prefer their F-35 Fighter Jets to be close to S-400 missile to not to provide any possibility for Russian technicians to observe any susceptibilities of F-35 (BBC News, 2019a).

Turkey previously became a partner country by investing comprehensively in U.S. F-35 warplane program. Turkey's aviation industry is currently producing 937 parts of the F-35 Fighter Jet (BBC News, 2019b), and signed up to buy 100 of U.S. F-35 Fighter Jets (BBC News, 2019a).

Despite to all political tension, accusations, and threats between sides, Turkey did not step back from the decision to purchase S-400 air missile system from Russia, and the first battery shipment has arrived to Turkey in July 2019. The actualization of receiving the parts for S-400 enraged U.S., which resulted with the exclusion of Turkey from the F-35 program. Turkey's aviation industry will also lose its production participation of 937 parts for the F-35 Fighter Jet. However, according to top officials of Pentagon, Turkey may regain supplier status and F-35 Fighter Jets if S-400 purchasing decision be reversed (Defensenews, 2019). Ongoing debates between sides continued, and S-400 deal could not be finalized for delivering the second shipment till August 2020.

6.2.4 2019 Istanbul Mayoral Election

Istanbul mayoral election was held on 31 March 2019. While there were thirteen candidates for the position of Mayor of Istanbul, there were two favorites who are; Nation Alliance's candidate Ekrem Imamoglu and People's Alliance's candidate Binali Yıldırım. Election had resulted in favor of Ekrem Imamoglu by a narrow margin of only 0.2% (Haberturk, 2019). Although, governing Justice and Development Party (AKP) petitioned Supreme Electoral Council for recounting the invalid votes and correction of ballot records in nine districts of Istanbul two days after the elections, which followed by another petition for recounting the votes in thirty-one districts. Supreme Electoral Council denied AKP's petition on recounting in thirty-one districts. Supreme Electoral Council conclude their fourteen hours long meeting with a result of recounting votes in 63 districts of Istanbul (Sozcu, 2019), and votes were recounted two weeks later (Hürriyet, 2019). As a result of recounting process, Ekrem Imamoglu received certificate of election on 17 April 2019 (Cumhuriyet, 2019). Nevertheless, AKP objected to latest result by presenting 44-page long report to Supreme Electoral Council. On 6 May 2019, Supreme Electoral Council accepted AKP's objection and decided to cancel the Istanbul mayoral election that was held in 31 March 2019, and to repeat on 23 June 2019 (Indipendent, 2019). New election calendar announced by Supreme Electoral Council and election process has started again in 15 May, while election campaigns have started again in 13 June till 22 June 2019. Meanwhile, accusations by the members of AKP targeting Ekrem Imamoglu became more and more violent each passing day (T24, 2019a; T24, 2019b). Supreme Electoral Council announced official results of 23 June elections on 11 July 2019. Official results indicates Ekrem Imamoglu as a new Mayor of Istanbul, with a record breaking margin of votes, 54.2% against Binali Yıldırım's 45.0% (BBC News, 2019c). Whole election process had followed closely by international press and interpreted as very politically instable and uncertain situation. The results of Istanbul mayoral election evaluated as the 'beginning of the end' for AKP and President Erdogan by international commentators, which obviously indicates expectations for political instability and uncertainty to continue (Euronews, 2019; The National News, 2019; BBC News, 2019d).

6.3 Data and Fundamental Variables

For sentiment analysis, HTTP REST API has extracted daily news articles regarding to each aforementioned political cases, and also any further miniscule cases that might not by substantial as others, yet still a conceivable factor that may influence the decision makers' decision from leading news websites BBC, Reuters, Bloomberg, The Guardian, and The New York Times. 613 news articles regarding to aforementioned political events have been analyzed, and textual content transformed into numeric sentiment inputs.

To perceive the degree of responsiveness of exchange rate to recently published government officials' political and international relations related official verbal or written statements; daily price, daily change in percentage, opening price, highest price, lowest price, and difference between highest and lowest prices of USD/TRY be investigated for each day.

Especially opening price, close price, highest price, lowest price, and the difference between these prices plays a crucial role on overcoming any possible accuracy error on the ML algorithm caused by the coincidence of news content on opposite polarities that may take place on the same day. The accommodation of news on opposite polarities on the same day may neutralize the significant effect of vital news, and cause a dramatic error on further calculations of the algorithm. To cope with this issue, the philosophy behind the candlestick chart has been adopted. In this case, even if neutralization happened in any day during the selected date interval, algorithm will detect it by observing the volatility of the exchange rate price throughout the day, and weight related news accordingly to match the volatility and the price by considering previous/further news on both subjects.

The abovementioned varieties of historical USD/TRY price data have been collected, and split accordingly regarding to the date intervals of the political events.

The degree of sensitivity of exchange rate to recently published government officials' political and international relations related official verbal or written statements regarding to Pastor Andrew Brunson has been perceived by examining the aforementioned historical exchange data, which is equivalent to

1746 data points has been collected between the dates of 29 December 2017 and 16 October 2018.

In the best interests of distinguished understanding on the possessions of 2018 Parliamentary and Presidential elections, 648 historical data points that has been collected between the dates of 15 April 2018 and 01 August 2018 so as to perceive the grade of sensitivity of exchange rate to recently published parliamentary and presidential elections related official verbal or written statements with higher machine learning algorithm accuracy.

On behalf of sophisticated understanding of the consequences of the S-400 Crisis, above-mentioned exchange rate data have been collected over the period between 29 December 2018 and 15 October 2019, which is equivalent to 1740 variables, and investigated daily in order to perceive the degree of responsiveness of exchange rate to recently published government officials' political and international relations related official verbal or written statements with advanced machine learning algorithm precision.

To perceive the degree of responsiveness of exchange rate to recently published government officials' and mayor candidates' mayoral election related official verbal or written statements regarding to 2019 Istanbul Mayoral elections, which causes uncertainty continuum, aforesaid varieties of historical exchange rate data, which is equivalent to 648 variables have been collected over the period between 15 March 2019 and 01 July 2019.

6.4 Methodology

The methodology for this chapter is constructed to comprehend the impact of particular political events on exchange rate fluctuations, and predict exchange rates accordingly by evaluating news articles. This chapter is addressing each political event individually. For that reason, presented methodology is going to be applied repeatedly for each political event in order to enhance understanding of the influence of event on decision makers', and behavior of decision makers under the specific environments that causes fluctuation on exchange rate. To do so, methodology is constructed as three phases, which are "Sentiment Analysis", "Fuzzy Logic", and "Support Vector

Regression” respectively. In depth methodology for aforementioned phases has been presented in Chapter 4. Related sections and subsections for relevant phases are going to be referenced below. First phase “Sentiment Analysis” (Chapter 4.2), practices components of Natural Language Processing, namely; Lexicon (Chapter 4.2.1), and Learning Automata (Chapter 4.2.2) for textual analysis, which reinforced by Prospect Theory (Chapter 4.2.3). Outcome of Sentiment Analysis is going to be examined by the second phase, “Fuzzy Logic” that adopts fuzzy logic (Chapter 4.3) algorithm from machine learning, and going to achieve a correlation value between each political event and exchange rate. Final phase, “Support Vector Regression” (Chapter 4.4.1) is going to run support vector regression. Monte Carlo Cross Validation (Chapter 4.5.1.1) is going to be used to split training and test data. The linearity of the provided dataset is going to be examined by Principal Component Analysis (PCA), Principal Component Regression (PCR), and Partial Least Square (PLS) methods, and kernel function for the model is going to be assigned according to the linearity of the dataset. The performances of Polynomial, and Gaussian Radial Basis kernel functions are going to be compared, and the one that provides the most accurate hyperparameter selection that eliminates heteroscedasticity in regression is going to be adopted for the model. Hyperparameters for selected kernel function; Gaussian Radial Basis Function is going to be obtained by Grid Search with Cross Validation (Chapter 4.5.2.1). Grid Search with Cross Validation method plays a crucial role on fine-tuning the hyperparameters for the kernel function cautiously and suitably to provide most accurate hyperparameter selection, while eliminating the heteroscedasticity and multicollinearity in regression analysis. The accuracy of the obtained parameters is going to be validated by Accuracy, Precision, Recall, and F1 Score metrics (Chapter 4.5.2.2). The performance of the constructed model is going to be evaluated by the statistical tests, V Measure Score (VMS), Explained Variance Score (EVS), Mean Squared Error (MSE), Root Mean Squared Error (RMSE), Mean Absolute Error (MAE), Mean Absolute Percentage Error (MAPE), and Coefficient of Determination (R^2) (Chapter 4.6), and future exchange rates are going to be predicted.

The workflow mechanism of the presented methodology is going to be detailed in the following subsection.

6.4.1 Workflow Mechanism of Methodology

As this chapter principally concentrates on four main political events, namely; Pastor Andrew Brunson case, 2018 Parliamentary and Presidential elections, S-400 crisis, and eventful 2019 Istanbul Mayoral elections, every single news articles that has been published by the leading reliable news sources that has been mentioned in Section 3, have been collected between the identified date intervals for each political event, which was also mentioned in Section 2, and extracted based on the country/countries involved, and related case/events.

Collected and extracted articles form the base for Sentiment Analysis. Annotated text of every news article is going to be feed into Learning Automata to construct a unique Lexicon dedicated to the particular case. Word polarity function forms the vertebral column, and is going to be authorized by extracting the word polarity functions of existing associative Lexicons to ensure the perception, and precise consideration of multidisciplinary words provided in the annotated text. Generated Lexicons for particular cases extracts sentiments that annotated text comprises by examining the text word by word, and produce an overall polarity vector value for every text feed into algorithm. Positive expressions of polarity vector values are normalized into positive and negative values by normalized polarity vector.

Besides, a hidden layer of algorithm that has been embedded in the python will dissociate and classify the news articles into political categories, based on the wording structure of the annotated text. The hidden layer will classify each news article regardless of the political event it articulates, but concentrates on the political categories, such as; judicial system, international politics, international relations, financial sanctions, warfare, defense industry, elections, cabinet changes, and democratic uncertainty. Besides, political news that accommodates possible direct or indirect discourse on macroeconomic complexion is also going to be noted and categorized. Singular dissociation of articles into political and macroeconomic categories plays a significant role on understanding the complex behavioral structure of human reasoning, yet the response that causes fluctuations on exchange rate prices. Attained categorization is going to be hardcoded into the ID of the particular news article, and stored to be used on complex behavioral analysis on the impact of politics on exchange rate fluctuations, which is going to be examined on the next chapters.

Obtained normalized polarity vector values provide a probability of the sentiment outcome, which is going to be evaluated by observing the effects of decision made on the historical USD/TRY data concurrent within the date intervals that includes particular cases, which have been mentioned in Section 3. Moreover, the outcome of observation on the decision making process can shape a database to provide algorithm the ability to bypass the process of observing historical data to derive the subjective feelings of the decision maker regarding to the breaking news, and predict the exchange rate even before the effects of the flash news transpired. Subjective feelings of the decision maker are going to be extracted, which is going to be used as an input for prospect theory to generate a prospect value for every news article examined as an outcome of sentiment analysis. The outcome of the sentiment analysis is going to shape the fuzzy antecedents for fuzzy logic phase. When the content of the political news considered, there are two main themes that psychologically affects the subjective feelings of the decision maker, which were also proven by the existing literature. These are the degree of certainty and the degree of consensus that encountered relatively as the main viewpoints of the published news. For that reason; fuzzy antecedents have been distributed into two main classes for political cases, which are; Certainty and Consensus. The membership function for Certainty and Consensus divided into three main sets. The membership function for Certainty defined as Certain, Neutral, and Uncertain, while defined as Agreement, Neutral, and Disagreement for Consensus. The transition between the memberships of the determined sets for each fuzzy antecedent, aside from the degrees of truth for each defined set that have been mentioned above are going to be defined by the upper and lower bounds, width, concave, convex, and the steepness of the associated membership, which will be presented on the future graphs. The fuzzy relations are going to be obtained by the presence or absence of association, interaction, or interconnectedness between the elements of two or more sets. Interpretation of the resulting fuzzy relations may be obtained by the fuzzy rules designated by the author, which are going to be feed into fuzzy engine as well as membership functions. Fuzzy engine is going to generate Fuzzy Consequent, which is essentially a dependency of USD/TRY behavior on recently published news articles. In other words, fuzzy consequent signifies the correlation between USD/TRY and news articles regarding to political tension. Correlation value that obtained as fuzzy consequent used as input as well as historical USD/TRY data to run support vector regression. The linearity of the provided dataset is going to be examined by Principal Component Analysis (PCA), Principal Component Regression (PCR), and Partial Least Square (PLS) methods. PCA/PCR is a mono-dimensional, observed variable approach that aims to find a linear correlation by reducing dimensionality. PLS is a multidimensional, latent variable approach that suits best for when there is a multicollinearity. Both PCA/PCR and PLS methods can be used to tests outliers and sensitivity analysis to provide diagnostic tools for the model. By using the kernel functions, nonlinear USD/TRY data have been projected into high-dimensional space. Constructed hyperplane provides adjustability to converge optimal set of weights (Oliveira et al., 2017). Gaussian Radial Basis Function utilized as kernel function due to its good fit on non-linear inputs (Grigoryan, 2016). Gaussian Radial Basis Function tuned by hyperparameters that have been attained by grid search with cross validation method, which also eliminated the possibility of heteroscedasticity and multicollinearity. Accuracy of the parameters attained by the grid search with cross validation is assessed by accuracy, precision, recall, and f1-score metrics. The data will randomly be divided into training and test groups for multiple times in order to appraise the performance of the regression model. To do so, Monte-Carlo cross validation (MCCV) method has been used. Splitting percentages of training and test samples have been adjusted according to the data size of the political event, and the portion of the data is reserved to run prediction, and observe the accuracy. Accuracy of the prediction investigated by the statistical tests. Prediction consistency of the model is evaluated against seven widely used statistical tests for machine learning algorithms, namely, V Measure Score (VMS), Explained Variance Score (EVS), Mean Squared Error (MSE), Root Mean Squared Error (RMSE), Mean Absolute Error (MAE), Mean Absolute Percentage Error (MAPE), and Coefficient of Determination (R^2). VMS is a decisive measure to begin with, as it indicates clustering performance of the model. To perform an accurate prediction with machine learning algorithms, superior performance on homogeneity and completeness on clustering is an essential in order to achieve a precise prediction without multicollinearity issue. For mentioned reasons, achieving significant VMS results is required for successful regression analysis. Grade of achieved VMS value provides insight on the prejudices of measured deviation of errors. MSE, RMSE, MAE, and MAPE are calculated in order to measure the deviation between predicted and actual values. RMSE and MAE are calculated simultaneously to identify possible deviation in the errors. While equivalent RMSE and MAE values indicates similarity of magnitude on all errors, the difference between the results specifies the variance associated with the individual errors. Greater difference indicates higher variance, and occurrence dissemination of the specific errors. The presence of underfitting and overfitting issues on the model are going to be evaluated by examining the difference between RMSE and MAE results. Additionally, MAPE is also considered and calculated to normalize the absolute error and obtain altered perspective on association with actual values. While MSE, RMSE, MAE, and MAPE measure the deviation between predicted and actual value, R^2 and EVS targets inconsistency between predicted and actual value. However, while R^2 uses raw sum of squares, Explained Variance Score uses the biased variance. Both variables can only be similar if the mean of the residuals equals to zero, and prediction is unbiased. Moreover, the explainability of the variance by the factors presented in the actual data is evaluated by the EVS results to check the homoscedasticity.

Finally, if the statistical tests' results are convenient, future exchange rates are going to be predicted in daily basis. The empirical workflow that described above has presented as a flowchart in Figure 42.

Presented workflow mechanism of methodology is going to be applied for every political event individually and collectively in purpose of unraveling the impact of the variety of tension and uncertainty grades on exchange rate.

6.5 Empirical Findings of the Hybrid Machine Learning Algorithms

The sentiment analysis primarily aims to conceive subjective perspective of decision makers under diversified events and situations. In order to acquire, unique lexicon for the various kinds of scenarios and cases built by Learning automata, which supplemented by the polarity functions of existing lexicons. Especially two main inferences aimed to reach by this process, which are; degree of political certainty, and political consensus. Prospect theory enhanced the outcome with the historical USD/TRY data to conceive decision-making patterns of decision makers on degrees of certainty and consensus that obtained as normalized polarity vector. This procedure provided an enhanced overlook on the content of the news, perspective of decision makers on each scenario, and fuzzy antecedents for the upcoming step.

Analyzed news has been scored regarding to the outcome of sentiment analysis. Sentiment scores have been assigned between 1 and 5, where 1 denotes highest likelihood for definitive sharp depreciation of USD/TRY, progressive increase in the score from 1 to 3 denotes fading depreciation on USD/TRY, while 3 denotes the steady state of the exchange rate. Scores from 3 to 5 denotes a gradual appreciation in USD/TRY, while 5 denotes substantial absolute sharp appreciation in USD/TRY. Attained sentiment scores are going to be feed into fuzzy logic algorithm as an antecedent under the title of each political case and event that has been previously mentioned above.

Fuzzy sets defined with membership functions for the antecedents ‘consensus’ and ‘certainty’, which indicates the subjective level of political consensus between sides, and subjective level of political certainty respectively. Political consensus and certainty have acute effects on international relations and economics. For that reason, trapezoidal membership function has selected to be used for fuzzy antecedents, while fuzzy consequent is defined with sigmoidal function.

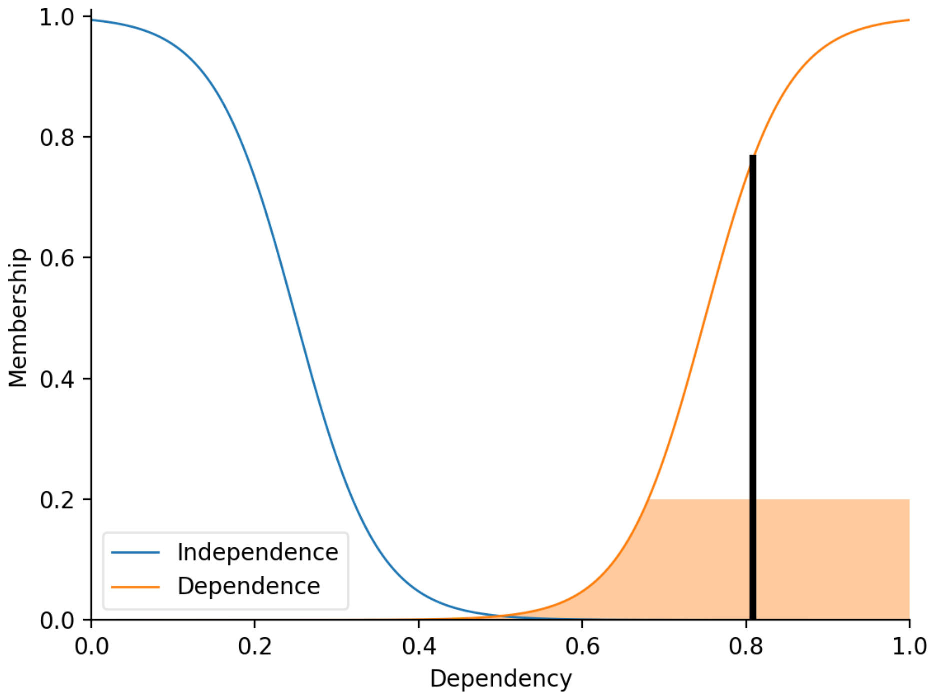
Membership functions defined for each fuzzy antecedent. For the antecedent ‘consensus’, assigned membership functions are: ‘agreement’, ‘neutral’, and ‘disagreement’, while ‘certain’, ‘neutral’, and ‘uncertain’ are assigned as membership functions for antecedent ‘certainty’. Sets for membership functions assigned by the fuzzy engine, in order to determine the degree of belonging to a fuzzy set according to the consequence of the sentiment analysis. ‘Certainty’ and ‘consensus’ values between 1 and 3 denotes the degree of belonging to fuzzy subsets ‘certain’ and ‘agreement’, 3 defined as ‘neutral’, and values between 3 and 5 denotes the degree of belonging to fuzzy subsets ‘uncertain’ and ‘disagreement’ respectively. Degree of membership to a fuzzy set denoted between 0 and 1, While 0 and 1 represents 0% and 100% membership degree respectively, anything in between considered as a partial membership to belonging set. Correlation value for each case is going to be generated by the combining the results derived from certainty and consensus values within the fuzzy engine based on the preset fuzzy rules. Correlation between published news regarding to each case and USD/TRY exchange rate is defined as dependency value, where value 0 defines that case related news and fluctuations in exchange rate is extremely independent, and value 1 defines that case related news and fluctuations in exchange rate is extremely dependent.

As each political case and event has its particularly unique dynamics, fuzzy logic algorithm is diversified and constructed specifically for each individual political case. Each fuzzy logic algorithm that has been constructed purpose specific for each political case and event compose their own membership function, questioned under event specific fuzzy rules, examined with unique fuzzy engine, and provided a fuzzy consequent specifically for that case. Attained fuzzy logic phase results, and further phases are going to be detailed and presented under the title of each case;

6.5.1 Fuzzy Logic and Support Vector Regression Results for Pastor Andrew Brunson Case

Fuzzy consequents regarding to certainty, consensus, and dependency of exchange rate on recently published news concerning the Pastor Andrew Brunson case provided the correlation values, which are necessary to comprehend significance and impact of the varieties of political tension, market's reaction on gradual grade of tension, and to run accurate prediction by using support vector regression accordingly. Obtained averaged consequent values for Pastor Andrew Brunson case indicates significant membership value of disagreement for consensus, while indicating insignificant membership value of certain for certainty when each antecedent analyzed individually. Averaged correlation values indicate a significant adverse impact of disagreement between Turkey and United States on USD/TRY exchange rate. Furthermore, while certain news are impacting USD/TRY exchange rate positive but insignificantly, uncertain news observed to impact exchange rate negative and significantly. Fuzzy consequent for degree of political certainty and consensus of the news articles designates a certain disagreement between U.S. and Turkey. Simulation results for 'certainty' and 'consensus' are visually presented in Figure 43 and 44 in the appendix section. Within the scope of certainty and consensus values regarding particularly to this case, dependency between Pastor Andrew Brunson case and USD/TRY exchange rate has been simulated. Simulation results indicate a highly significant dependency value, 0.822 between the published news regarding to the case, and fluctuations on USD/TRY exchange rate. Dependency rate is visually presented below, as Figure 45.

Fig. 45 Dependency Rate between Published News Related to Pastor Andrew Brunson Case and USD/TRY Exchange Rate Fluctuations



Attained dependency rate from the fuzzy logic phase is going to be used as an input alongside to historical exchange rate data to run SVR. The linearity of the provided dataset is evaluated by PCA/PCR and PLS methods. Figure 46, which presented in the appendix section illustrates PCA/PCR and PLS results. First principal component in PCA/PCR and PLS indicates the axis in the K-dimensional variable space that accommodates the largest variance of the samples in a direction, while second principal component is orthogonal axis to the first principal component in the K-dimensional variable space, and identifies the second largest source of samples in a direction that can be denoted as the residual variance of the samples that improves the approximation. Besides, the direction that provides the lowest variance is aimed to be captured by the PLS. The nonlinearity of the provided dataset has confirmed by the tests applied. For that reason, kernel function for the SVR analysis is going to be assigned accordingly in order to prevent heteroscedasticity and multicollinearity. The application of Gaussian Radial

Basis Function on PCA and the projection difference is exemplified as Figure 35 in the appendix section.

By using the kernel functions, nonlinear USD/TRY data have been projected into high-dimensional space. Constructed hyperplane provides adjustability to converge optimal set of weights (Oliveira et al., 2017). Gaussian Radial Basis Function utilized as kernel function due to its good fit on non-linear inputs (Grigoryan, 2016). Selected kernel function tuned by $C, \varepsilon, \text{and}\ \gamma$ hyperparameters. In order to set the hyperparametes, a grid search has been adopted to eliminate the possibility of multicollinearity and achieve highest accuracy possible. Grid search ranges of $C, \varepsilon, \text{and}\ \gamma$ hyperparameters assigned between the range suggested by Hsu, Chang, and Lin (2003). *C* ranges from 1 to 10^5, and γ from 2^{-10} to 2^3. Grid search with cross validation for hyperparameter $C, \text{and}\ \gamma$ value selection is presented as Table 6 in the appendix section. Grid search cross validation accuracy heat map is also presented as Figure 47 in the appendix section. Hyperparameter ε is assigned by considering the highest change in the historical data, and determined as high as possible in order to eliminate any possible overfitting issue. Accuracy of the parameter selection using grid search with cross-validation is evaluated by consideration of three metrics, namely; precision, recall, and f1-score. Results have indicated the accuracy of 0.97 for parameter selection, and detailed in Table 7, which has been presented below.

Tab. 7 Accuracy of the Parameter Estimation Using Grid Search with Cross-Validation for Pastor Andrew Brunson Case

The model is trained on the full development set.
The scores are computed on the full evaluation set.

	precision	recall	f1-score	support
0	0.94	1.00	0.97	20
1	1.00	0.96	0.97	30
Accuracy			0.97	75
Macro Avg	0.97	0.97	0.97	75
Weighted Avg	0.97	0.97	0.97	75

The critical values of Accuracy, Precision, Recall, and F1 Score are 0.20, 0.40, 0.60, and 0.80 denotes slight, fair, moderate, substantial, and almost perfect agreement respectively.

Pastor Andrew Brunson case has provided restricted historical data due to its relatively short existence. Therefore, splitting presented data into training and test groups is essential to appraise the performance of the regression model. To do so, Monte-Carlo cross-validation (MCCV) method adopted. 80% of provided data used as train data, while 20% reserved as test data. MCCV randomly splits reserved portion of data into sub-samples and assign them as test sets. The process of selecting random independent partitions repeated for multiple times. Prediction consistency of the model is evaluated against seven key statistical metrics, which are widely be accepted and using to validate machine learning prediction performance, namely; V Measure Score (VMS), Explained Variance Score (EVS), Mean Squared Error (MSE), Root Mean Squared Error (RMSE), Mean Absolute Error (MAE), Mean Absolute Percentage Error (MAPE), and Coefficient of Determination (R^2). Monte-Carlo cross-validation run numerous times, and statistical test results attained for each run. Achieved results for statistical test are presented exhaustively in Table 8 in the appendix section. Then, obtained statistical test results averaged. Averaged results for MCCV statistical test results are presented in Table 9.

Table 9. Averaged Monte Carlo Cross Validation Statistical Test Results for Pastor Andrew Brunson Case.

VMS	0.928399***
EVS	0.935027***
MSE	0.057164
RMSE	0.239008*
MAE	0.230802
MAPE	0.050095**
$\boldsymbol{R^2}$	0.931198***

† Exhaustive Monte-Carlo cross validation statistical test results are presented in the appendix section, as Table 8.

The critical values of VMS are 0.70, 0.80, and 0.90. Therefore, *, **, and *** denotes threshold, significant, and highly significant at 0.70, 0.80, and 0.90 respectively.

The critical values of EVS are 0.60, 0.80, and 0.90. Therefore, *, **, and *** denotes threshold, significant, and highly significant at 0.60, 0.80, and 0.90 respectively.

The critical values of RMSE are set by SI at $\leq 10\%$, and $\leq 5\%$, which indicate RMSE value of 0.469161 and 0.234580 respectively. Therefore, *, and ** indicates the threshold and significance at respective critical points.

The critical values for the difference between RMSE and MAE at $\leq 5\%$, and ≤1% denote acceptable and perfect fit of the model respectively.

The critical values of MAPE are 0.25, 0.10, and 0.05. Therefore, *, **, and *** denotes threshold, low but acceptable, and highly acceptable accuracy at 0.25, 0.10, and 0.05 respectively.

The critical values of R^2 are 0.50, 0.75, and 0.90. Therefore, *, **, and *** denotes weak, moderate, and substantial prediction at 0.50, 0.75, and 0.90 respectively.

The critical value of the difference between EVS and R^2 is $\leq 2\%$. Therefore * denotes acceptable bias at $\leq 2\%$.

In purpose of validating the results, alongside of measuring the sustainability of the results, the Monte Carlo Cross Validation has repeated 40 times, the mean of achieved statistical test results has been taken and presented above. The mean of MCCV statistical test results indicates that;

VMS result is 0.928399, which is higher than 0.90 criteria. Therefore, it is possible to denote VMS value as highly significant, thus indicates a superior harmonic mean between homogeneity and completeness and indicates a perfectly accomplished clustering task for the future weight support vector regression. Additionally, highly significant VMS result, 0.928399 indicates no multicollinearity issue, and authorizes the model for further tests and regression analysis.

EVS result is 0.935027, which is higher than 0.90 criteria. Therefore, it is possible to signify EVS value as highly significant. The results indicate that the variance can be explained by the factors presented by the actual data. Moreover, EVS is another measure that is crucial for regression analysis, as it scores homoscedasticity. Highly significant EVS result, 0.935027 denotes that the variance can be explained by the factors presented by the actual data, and not heteroscedastic.

MSE result is 0.057164. As it has been emphasized on the Chapter 4, MSE closer to 0 indicates better fit of regression line, as it indicates the variance of residuals. However, as there isn't acceptable range set for MSE as it does not shares the same unit as the original values, MSE has been used to achieve RMSE values.

RMSE value is in the same unit with the original data, and indicates the standard deviation of errors emerged during prediction, thus signifies the accuracy of the model. RMSE result is 0.239008, which denotes that the standard deviation of error is 0.239008 when compared to actual USD/TRY value. In order to evaluate RMSE result, the value transformed into interpretable value, SI. Transformed RSME, SI value is 0.050944, which is lower than 0.10 criteria. Therefore it is possible to denote RMSE value is at acceptable range.

MAE is relatively insensitive to outliers compared to RMSE, while RMSE magnifies the bigger errors and ignoring the smaller errors. For that reason, MAE may accommodate bias, while RMSE is related to variance. When MAE

and RMSE compared, it is possible to observe underfitting and overfitting issues that have been targeted by the application of MCCV method. Underfitting and overfitting is the situation of high bias, low variance, and low bias, high variance respectively. The difference between RMSE and MAE results, 0.239008 and 0.230802 respectively is 3.55%, which is lower than the 5% criteria. Therefore, the difference is in acceptable range, and denotes that model has fitted without any underfitting or overfitting issues.

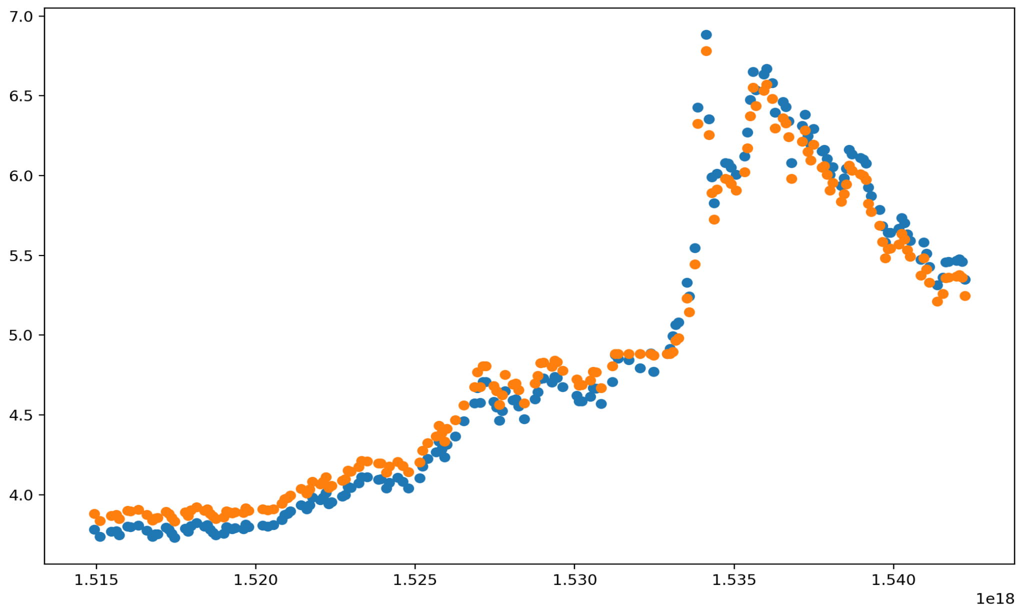
The accuracy of the prediction has also been measured by the MAPE results. MAPE results, 0.050095 that is equivalent to 5%, which is equal to 0.05 criteria, equivalent to 5%. Therefore, it is possible to say that the model has highly acceptable accuracy.

R^2 result is 0.931198, which is higher than the 0.90 criteria. Highly significant results have verified that the prediction is substantial with a superior accuracy, as the highly significant majority of the predicted data points are on the regression line.

Furthermore, the deviation between EVS and R^2 results, 0.935027 and 0.931198 respectively is only 0.41%, which is lower than 2% criteria. Thus the deviation between EVS and R^2 results signifies that the prediction is unbiased.

Comprehensive examination has been made methodically throughout the fuzzy logic and support vector regression phases to understand the reasoning behind the behaviour of RMSE, MAE, and especially MAPE results. It has been observed that, relatively high, yet acceptable RMSE, MAE, and MAPE results have been occurred due to the critical, uncertain, and multidirectionally shifting tension nature of the case. During the fuzzy logic phase, a certain disagreement between United States and Turkey has been observed. However, phase examined in detail, it also be seen that uncertain disagreement have also be occurred betweenwhiles. Strict behaviour of both sides before and after the court hearings, expected court orders, political and financial sanctions against each other, and releasing Pastor Brunson are some of the featured reasons for uncertainty and political tension between United States and Turkey. Uncertainty, uncertain disagreement, and uncertain agreements throughout the case have elevated the possibility of sudden change in direction on the trend of the news. For that reason, margin for error should have been kept wide in order to eliminate any fitting problem, and also biased calculations.

Fig. 49. Actual versus predicted USD/TRY prices obtained by analyzing recently published news and statements regarding on Pastor Andrew Brunson Case.



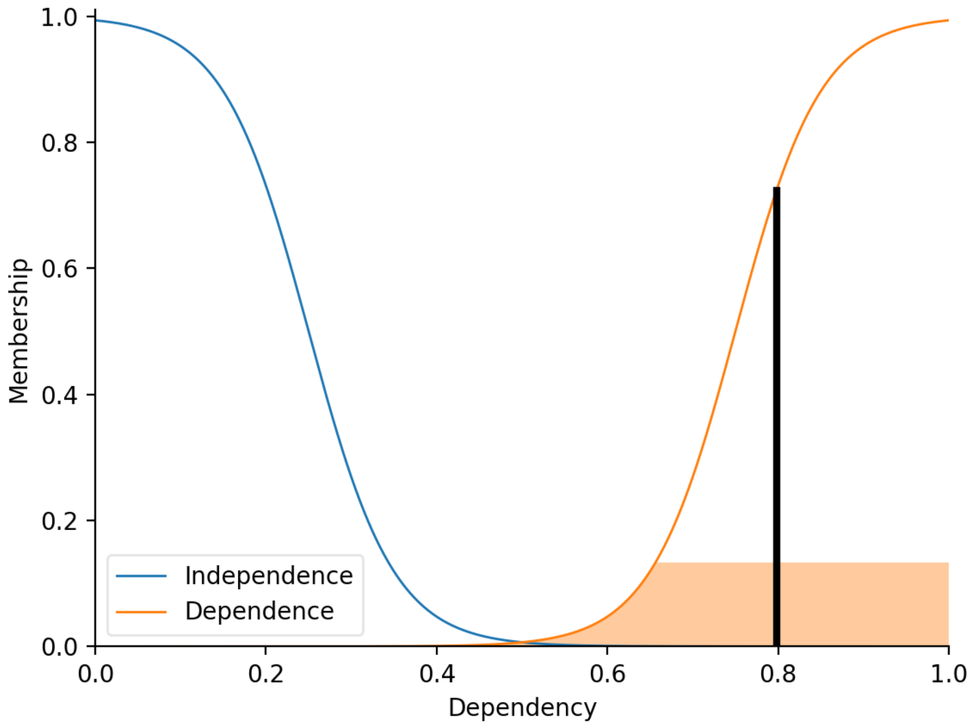
**Blue Points denote Actual USD/TRY Prices

**Orange Points denote Predicted USD/TRY Prices

6.5.2 Fuzzy Logic and Support Vector Regression Results for 2018 Turkish Parliamentary and Presidential Elections

In order to comprehend weight and impact of the encountered political situations like electoral competition, governmental uncertainty, and policy insulation on exchange rate, fuzzy consequents concerning certainty, consensus, and dependency rate of recently published news concerning the 2018 Parliamentary and Presidential elections plays a vital role. Fuzzy logic phase is targeting to comprehend investors' interpretation on breaking news regarding on electoral competition, and expectancy on cabinet changes to be able to run accurate prediction by using support vector regression. Obtained averaged consequent values for 2018 Parliamentary and Presidential Elections signifies insignificant membership value of disagreement for consensus, and insignificant membership value of certain for certainty when each antecedent analyzed individually. Averaged correlation values indicate an inconsequential negative impact of incongruity during the electoral competition on USD/TRY exchange rate. Furthermore, while certain news articles are impacting USD/TRY exchange rate positive but insignificantly, news causing uncertainty on the course of electoral competition observed to cause Turkish Lira to depreciate. Fuzzy consequent for degree of political certainty and consensus of the news articles defines certain disagreement between sides, as the nature of the election required. Simulation results for 'certainty' and 'consensus' are visually presented in Figure 50 and 51 individually. Within the scope of separated certainty and consensus values regarding particularly on ongoing parliamentary and presidential election campaigns and related published news articles, dependency between 2018 Parliamentary and Presidential Elections and USD/TRY exchange rate fluctuations has been simulated. Simulation results denote a highly substantial dependency value of 0.791 between the published news regarding to the aforementioned elections, and fluctuations on USD/TRY exchange rate throughout the election process. Dependency rate is visually presented below, as Figure 52.

Fig. 52 Dependency Rate between Published News Related to 2018 Turkish Parliamentary and Presidential Elections and USD/TRY Exchange Rate Fluctuations



Dependency rate that has been achieved on fuzzy logic phase is going to be used as an input together with historical USD/TRY exchange rate data for the last phase, support vector regression. The linearity of the provided dataset is evaluated by PCA/PCR and PLS methods. Figure 53, which presented in the appendix section illustrates PCA/PCR and PLS results. First principal component in PCA/PCR and PLS indicates the axis in the K-dimensional variable space that accommodates the largest variance of the samples in a direction, while second principal component is orthogonal axis to the first principal component in the K-dimensional variable space, and identifies the second largest source of samples in a direction that can be denoted as the residual variance of the samples that improves the approximation. Besides, the direction that provides the lowest variance is aimed to be captured by the PLS. The nonlinearity of the provided dataset has confirmed by the tests applied. For that reason, kernel function for the SVR analysis is going to be assigned accordingly in order to prevent heteroscedasticity and multicollinearity. The application of Gaussian Radial Basis Function on PCA and the projection difference is exemplified as Figure 35 in the appendix section.

Kernel functions are going to be featured to project nonlinear exchange rate data into high-dimensional space. Constructed hyperplane provides adjustability to converge optimal set of weights (Oliveira et al., 2017). Gaussian Radial Basis Function utilized as kernel function due to its good fit on non-linear inputs (Grigoryan, 2016). Kernel function selection a criterion has been selected Gaussian Radial Basis Function, and hyperparameters $C, \varepsilon, \text{and } \gamma$ have been tuned. In order to set the hyperparametes, a grid search technique has been adopted to eliminate the probability of multicollinearity and attain utmost accuracy achievable. Grid search ranges of $C, \varepsilon, \text{and } \gamma$ hyperparameters assigned as; *C* ranges from $2^{-3.3}$ to 10^5, and γ from $2^{-3.3}$ to 2000. Grid search with cross validation has run to obtain harmonious hyperparameter $C, \text{and } \gamma$ that fits perfectly particular to the input data of this case is accessible as Table 10 in the appendix section. Grid search cross validation accuracy heat map is also presented as Figure 54 in the appendix section. Hyperparameter ε is assigned by considering the highest change in the historical data, and determined as high as possible in order to eliminate any possible overfitting issue. Accuracy of the parameter selection using grid search with cross-validation is evaluated by consideration of three metrics, namely; precision, recall, and f1-score. Results have indicated the accuracy of 0.94 for parameter selection, and detailed in Table 11, as presented and detailed below.

Tab. 11 Accuracy of the Parameter Estimation Using Grid Search with Cross-Validation for 2018 Turkish Parliamentary and Presidential Elections

The model is trained on the full development set.
The scores are computed on the full evaluation set.

	precision	recall	f1-score	support
0	0.92	1.00	0.95	30
1	1.00	0.88	0.93	20
Accuracy			0.94	50
Macro Avg	0.95	0.93	0.94	50
Weighted Avg	0.94	0.94	0.94	50

The critical values of Accuracy, Precision, Recall, and F1 Score are 0.20, 0.40, 0.60, and 0.80 denotes slight, fair, moderate, substantial, and almost perfect agreement respectively.

Due to the nature of presidential and parliamentary elections, time interval for election campaign constrained. For that reason, number of published news regarding to topic, and also the historical data are restricted. Hence, splitting obtainable data into training and test groups is obligatory to recuperate the performance of the regression model. To do so, Monte-Carlo cross-validation (MCCV) approach is embraced. In order to enhance the performance of the algorithm, 90% of provided data used as train data to provide maximum data for training possible, while 10% reserved as test data, which is low, yet sufficient for a test run. MCCV randomly splits reserved portion of data into sub-samples and assign them as test sets. The process of selecting random independent partitions repeated for multiple times. Prediction consistency of the model is evaluated against seven key statistical metrics, which are widely be accepted and using to validate machine learning prediction performance, namely; V Measure Score (VMS), Explained Variance Score (EVS), Mean Squared Error (MSE), Root Mean Squared Error (RMSE), Mean Absolute Error (MAE), Mean Absolute Percentage Error (MAPE), and Coefficient of Determination (R^2). Monte-Carlo cross-validation run numerous times, and statistical test results attained for each run. Achieved results for statistical test are presented exhaustively in Table 12 in the appendix section. Then, obtained statistical test results averaged. Averaged results for MCCV statistical test results are presented in Table 13.

Table 13. Averaged Monte Carlo Cross Validation Statistical Test Results for 2018 Turkish Parliamentary and Presidential Elections.

VMS	0.905317***
EVS	0.870601**
MSE	0.011431
RMSE	0.106894**
MAE	0.103208
MAPE	0.022936***
$\boldsymbol{R^2}$	0.862592**

† Exhaustive Monte-Carlo cross validation statistical test results are presented in the appendix section, as Table 12.

The critical values of VMS are 0.70, 0.80, and 0.90. Therefore, *, **, and *** denotes threshold, significant, and highly significant at 0.70, 0.80, and 0.90 respectively.

The critical values of EVS are 0.60, 0.80, and 0.90. Therefore, *, **, and *** denotes threshold, significant, and highly significant at 0.60, 0.80, and 0.90 respectively.

The critical values of RMSE are set by SI at $\leq 10\%$, and $\leq 5\%$, which indicate RMSE value of 0.453034, and 0.226517 respectively. Therefore, *, and ** indicates the threshold and significance at respective critical points.

The critical values for the difference between RMSE and MAE at $\leq 5\%$, and ≤1% denotes acceptable and perfect fit of the model respectively.

The critical values of MAPE are 0.25, 0.10, and 0.05. Therefore, *, **, and *** denotes threshold, low but acceptable, and highly acceptable accuracy at 0.25, 0.10, and 0.05 respectively.

The critical values of R^2 are 0.50, 0.75, and 0.90. Therefore, *, **, and *** denotes weak, moderate, and substantial prediction at 0.50, 0.75, and 0.90 respectively.

The critical value of the difference between EVS and R^2 is $\leq 2\%$. Therefore * denotes acceptable bias at $\leq 2\%$.

In purpose of validating the results, alongside of measuring the sustainability of the results, the Monte Carlo Cross Validation has repeated 40 times, the mean of achieved statistical test results has been taken and presented above. The mean of MCCV statistical test results indicates that;

VMS result is 0.905317, which is higher than 0.90 criteria. Therefore, it is possible to denote VMS value as highly significant thus indicates a harmonic mean between homogeneity and completeness and indicates an accomplished clustering task for the future weight support vector regression. Additionally, significant VMS result, 0.905317 indicates no multicollinearity issue, and authorizes the model for further tests and regression analysis.

EVS result is 0.870601, which is higher than 0.80 criteria. Therefore, it is possible to signify EVS value as significant. The results indicate that the variance can be explained by the factors presented by the actual data. Moreover, EVS is another measure that is crucial for regression analysis, as it scores homoscedasticity. Significant EVS result, 0.870601 denotes that the variance can be explained by the factors presented by the actual data, and not heteroscedastic.

MSE result is 0.011431. As it has been emphasized on the Chapter 4, MSE closer to 0 indicates better fit of regression line, as it indicates the variance of residuals. However, as there isn't acceptable range set for MSE as it does not shares the same unit as the original values, MSE has been used to achieve RMSE values.

RMSE value is in the same unit with the original data, and indicates the standard deviation of errors emerged during prediction, thus signifies the accuracy of the model. RMSE result is 0.106894, which denotes that the standard deviation of error is 0.106894 when compared to actual USD/TRY value. In order to evaluate RMSE result, the value transformed into interpretable value, SI. Transformed RSME, SI value is 0.023595, which is lower than 0.05 criteria. Therefore it is possible to denote RMSE value as significant.

MAE is relatively insensitive to outliers compared to RMSE, while RMSE magnifies the bigger errors and ignoring the smaller errors. For that reason, MAE may accommodate bias, while RMSE is related to variance. When MAE and RMSE compared, it is possible to observe underfitting and overfitting issues that have been targeted by the application of MCCV method. Underfitting and overfitting is the situation of high bias, low variance, and low bias, high variance respectively. The difference between RMSE and MAE results, 0.106894 and 0.103208 respectively is 3.57%, which is lower than the 5% criteria. As the difference is in acceptable range, it is possible to say that the model has fitted without any underfitting or overfitting issues.

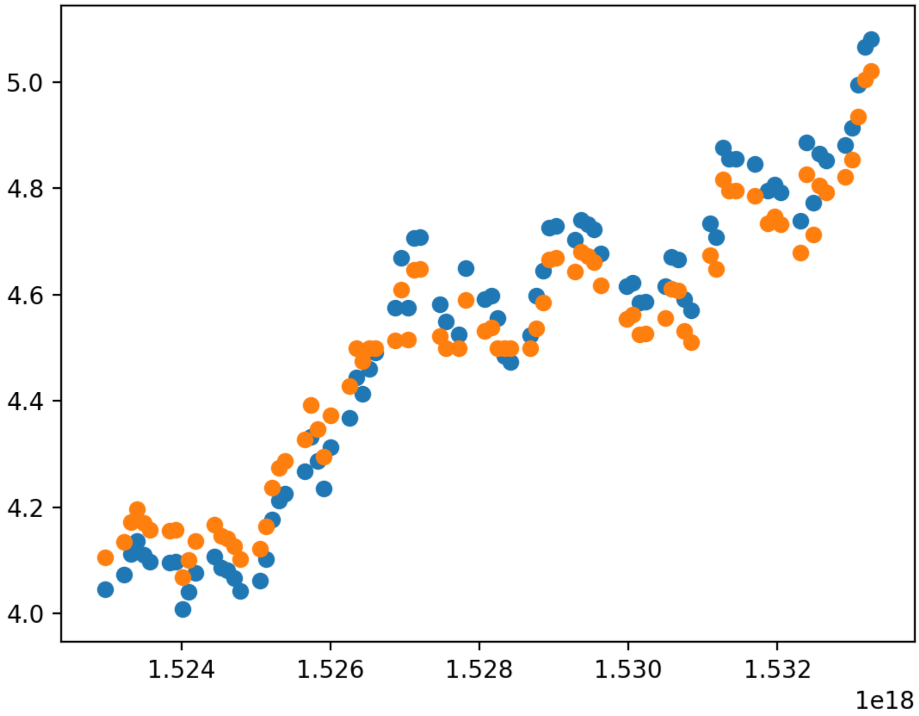
The accuracy of the prediction has also been measured by the MAPE results. MAPE results, 0.022936 that is equivalent to 2.29%, which is lower than the 0.05 criteria, equivalent to 5%. Therefore, it is possible to say that the model has highly acceptable accuracy.

R^2 result is 0.862592, which is higher than the 0.75 criteria. Moderate results have verified that the prediction is acceptable and accurate, as the majority of the predicted data points are on the regression line.

Furthermore, the deviation between EVS and R^2 results, 0.870601 and 0.862592 respectively is only 0.928%, which is lower than 2% criteria. Thus the deviation between EVS and R^2 results signifies that the prediction is unbiased.

Relatively lower rates for MAPE, RMSE, and MAE have been achieved by the relatively nonviolent and steady state nature of the elections process compared to cases that accommodates international political tension. This hypothesis was also empirically supported by the specific fuzzy consensus and certainty levels (Figure 50 and 51 in appendix) that has been mentioned on the previous phase, fuzzy logic. On deeper analysis, it has been observed that, election process only considered by the investors' as an uncertain environment, when the Election Day gets closer. On the other hand, as the day arrives, consensus turns into favour of agreement, while previous days were considering as a member of disagreement set. While these changes do not cause any dramatic fluctuations on exchange rate, the fact that elections' guidance on exchange rate cannot be ignored, as any possible cabinet change can directly impact government authoritarianism, policy, regime, and democracy which may directly affect investor's attitude. Prediction results for USD/TRY prices obtained by examining recently published news and statements regarding on 2018 Turkish Parliamentary and Presidential elections is visually presented in Figure 56.

Fig. 56. Actual versus predicted USD/TRY prices obtained by analyzing recently published news and statements regarding on 2018 Turkish Parliamentary and Presidential Elections.



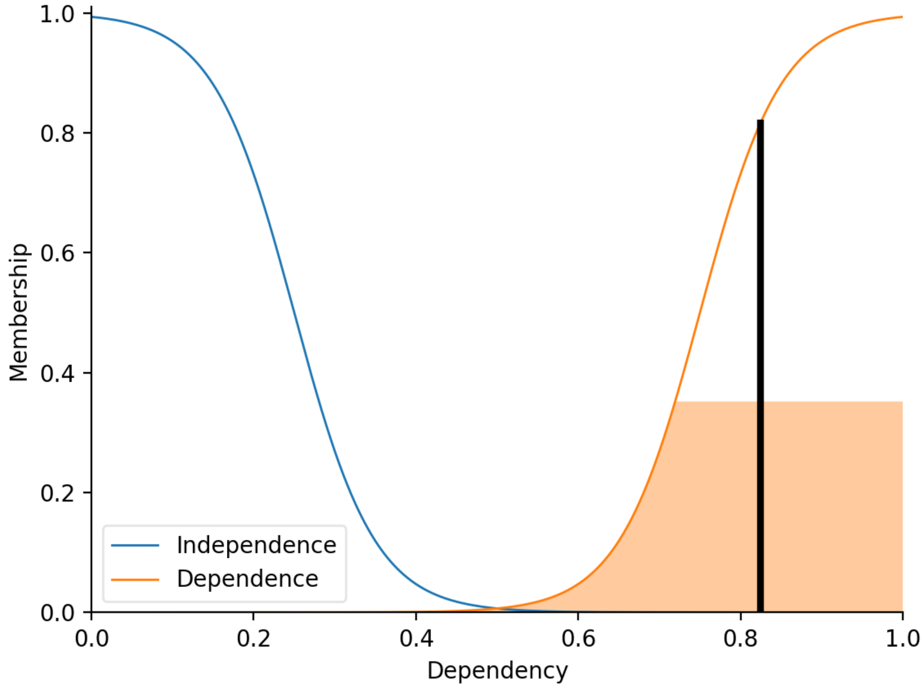
**Blue Points denote Actual USD/TRY Prices

**Orange Points denote Predicted USD/TRY Prices

6.5.3 Fuzzy Logic and Support Vector Regression Results for S-400 Crisis

Correlation values for S-400 crises regarding to certainty, consensus, and dependency of exchange rate on published news articles on the case has been obtained as fuzzy consequents, which are noteworthy indicators to understand how decision makers' react to fluctuating political tension between countries. Comprehending subjective perspective of decision makers on varieties of political tension, and being able to describe it quantitatively plays a significant role on running accurate predictions. S-400 crisis has obtained highest averaged membership values compared to other political cases that have been examined throughout this chapter. Once each fuzzy antecedent analyzed individually, membership values that have been derived by examining the related news articles evidently indicate a certain disagreement between sides thru the case. Averaged correlation values denote a substantial negative impact of disagreement between Turkey and United States on the price of Turkish Lira, which causes an appreciation on the price of USD/TRY exchange rate. Furthermore, certainty of the political disagreement between aforementioned countries amplifies the impact of disagreement on the exchange rate under the subjective judgment of decision makers on a certain situation that is less feasible to interpret. Fuzzy consequent for degree of political certainty and consensus of the news articles designates a certain disagreement between U.S. and Turkey. Simulation results for 'certainty' and 'consensus' are visually presented in Figure 57 and 58. Dependency rate of USD/TRY exchange rate on recently published news regarding to S-400 crises has also been simulated within the obtained certainty and consensus values. Simulation results indicate an exceptionally substantial dependency value of 0.846 between the published news and statements regarding to the case, and fluctuations on USD/TRY exchange rate. Dependency rate is visually presented below, as Figure 59.

Fig. 59 Dependency Rate between Published News Related to S-400 Crisis and USD/TRY Exchange Rate Fluctuations



Dependency rate that has been achieved on fuzzy logic phase is going to be used as an input together with historical USD/TRY exchange rate data for the last phase, support vector regression. The linearity of the provided dataset is evaluated by PCA/PCR and PLS methods. Figure 60, which presented in the appendix section illustrates PCA/PCR and PLS results. First principal component in PCA/PCR and PLS indicates the axis in the K-dimensional variable space that accommodates the largest variance of the samples in a direction, while second principal component is orthogonal axis to the first principal component in the K-dimensional variable space, and identifies the second largest source of samples in a direction that can be denoted as the residual variance of the samples that improves the approximation. Besides, the direction that provides the lowest variance is aimed to be captured by the PLS. The nonlinearity of the provided dataset has confirmed by the tests applied. For that reason, kernel function for the SVR analysis is going to be assigned accordingly in order to prevent heteroscedasticity and multicollinearity. The application of Gaussian Radial Basis Function on PCA and the projection difference is exemplified as Figure 35 in the appendix section.

Nonlinear historical data is projected into high-dimensional space by using kernel functions. Constructed hyperplane provides adjustability to converge optimal set of weights (Oliveira et al., 2017). Gaussian Radial Basis Function employed as kernel function for this case particularly due to its respectable fit on non-linear data (Grigoryan, 2016). Gaussian Radial Basis kernel function is tuned by $C, \varepsilon, \text{and}\, \gamma$ hyperparameters. In order to eliminate the possibility of multicollinearity and achieve highest accuracy possible, grid search has been adopted to set the hyperparametes. Grid search ranges of $C, \varepsilon, \text{and}\, \gamma$ hyperparameters assigned as suggested by Hsu, Chang, and Lin (2003). *C* ranges from $2^{-6.6}$ to 10^4, and γ from 2^{-10} to 2000. Grid search with cross validation for hyperparameter $C, \text{and}\, \gamma$ value selection is obtainable on Table 14 in the appendix section. Grid search cross validation accuracy heat map is also presented as Figure 61 in the appendix section. Hyperparameter ε is assigned by bearing in mind the uppermost difference in the historical data, and determined as high as possible in order to reject any potential overfitting issue. Accuracy of the parameter selection using grid search with cross-validation is evaluated by consideration of three metrics, namely; precision, recall, and f1-score. Results have indicated the accuracy of 0.98 for parameter selection, and detailed in Table 15, which is specified on the table below.

Tab. 15 Accuracy of the Parameter Estimation Using Grid Search with Cross-Validation for S-400 Crises

The model is trained on the full development set.
The scores are computed on the full evaluation set.

	precision	recall	f1-score	support
0	0.95	1.00	0.98	20
1	1.00	0.97	0.98	30
Accuracy			0.98	50
Macro Avg	0.98	0.98	0.98	50
Weighted Avg	0.98	0.98	0.98	50

The critical values of Accuracy, Precision, Recall, and F1 Score are 0.20, 0.40, 0.60, and 0.80 denotes slight, fair, moderate, substantial, and almost perfect agreement respectively.

Presented data is going to be split into training and test groups are essential to appraise the performance of the regression model. To do so, Monte-Carlo cross-validation (MCCV) method adopted. S-400 crisis is a long running case. That is why, its influence has been observed in a wide range of quantitative and qualitative data that has been collected. 80% of provided data used as train data, while 20% reserved as test data, as the case is generous in terms of data provided. MCCV randomly splits reserved portion of data into sub-samples and assign them as test sets. The process of selecting random independent partitions repeated for multiple times.

Prediction consistency of the algorithm is assessed in contradiction of seven fundamental statistical metrics, which are broadly be accepted and operating to authenticate machine learning algorithm's prediction performance, namely; V Measure Score (VMS), Explained Variance Score (EVS), Mean Squared Error (MSE), Root Mean Squared Error (RMSE), Mean Absolute Error (MAE), Mean Absolute Percentage Error (MAPE), and Coefficient of Determination (R^2). Monte-Carlo cross-validation run several times, and statistical test results achieved for each run. Achieved results for statistical test are accessible exhaustively in Table 16 in the appendix section. Then, obtained statistical test results averaged. Averaged results for MCCV statistical test results are presented in Table 17.

Table 17. Averaged Monte Carlo Cross Validation Statistical Test Results for S-400 Crisis.

VMS	0.885410**
EVS	0.905757***
MSE	0.005274
RMSE	0.072623**
MAE	0.069744
MAPE	0.012321***
$\boldsymbol{R^2}$	0.900212***

† Exhaustive Monte-Carlo cross validation statistical test results are presented in the appendix section, as Table 16.

The critical values of VMS are 0.70, 0.80, and 0.90. Therefore, *, **, and *** denotes threshold, significant, and highly significant at 0.70, 0.80, and 0.90 respectively.

The critical values of EVS are 0.60, 0.80, and 0.90. Therefore, *, **, and *** denotes threshold, significant, and highly significant at 0.60, 0.80, and 0.90 respectively.

The critical values of RMSE are set by SI at $\leq 10\%$, and $\leq 5\%$, which indicate RMSE value of 0.564463, and 0.282231 respectively. Therefore, *, and ** indicates the threshold and significance at respective critical points.

The critical values for the difference between RMSE and MAE at $\leq 5\%$, and ≤1% denote acceptable and perfect fit of the model respectively.

The critical values of MAPE are 0.25, 0.10, and 0.05. Therefore, *, **, and *** denotes threshold, low but acceptable, and highly acceptable accuracy at 0.25, 0.10, and 0.05 respectively.

The critical values of R^2 are 0.50, 0.75, and 0.90. Therefore, *, **, and *** denotes weak, moderate, and substantial prediction at 0.50, 0.75, and 0.90 respectively.

The critical value of the difference between EVS and R^2 is $\leq 2\%$. Therefore * denotes acceptable bias at $\leq 2\%$.

In purpose of validating the results, alongside of measuring the sustainability of the results, the Monte Carlo Cross Validation has repeated 40 times, the mean of achieved statistical test results has been taken and presented above. The mean of MCCV statistical test results indicates that;

VMS result is 0.885410, which is higher than 0.80 criteria. Therefore, it is possible to denote VMS value as significant, thus indicates a harmonic mean between homogeneity and completeness and indicates a well accomplished clustering task for the future weight support vector regression. Additionally, significant VMS result, 0.885410 indicates no multicollinearity issue, and authorizes the model for further tests and regression analysis.

EVS result is 0.905757, which is higher than 0.90 criteria. Therefore, it is possible to signify EVS value as highly significant. The results indicate that the variance can be explained by the factors presented by the actual data. Moreover, EVS is another measure that is crucial for regression analysis, as it scores homoscedasticity. Highly significant EVS result, 0.905757 denotes that the variance can be explained by the factors presented by the actual data, and not heteroscedastic.

MSE result is 0.005274. As it has been emphasized on the Chapter 4, MSE closer to 0 indicates better fit of regression line, as it indicates the variance of residuals. However, as there isn’t acceptable range set for MSE as it does not shares the same unit as the original values, MSE has been used to achieve RMSE values.

RMSE value is in the same unit with the original data, and indicates the standard deviation of errors emerged during prediction, thus signifies the accuracy of the model. RMSE result is 0.072623, which denotes that the standard deviation of error is 0.072623 when compared to actual USD/TRY value. In order to evaluate RMSE result, the value transformed into interpretable value, SI. Transformed RSME, SI value is 0.012866, which is lower than 0.05 criteria. Therefore it is possible to denote RMSE value as significant.

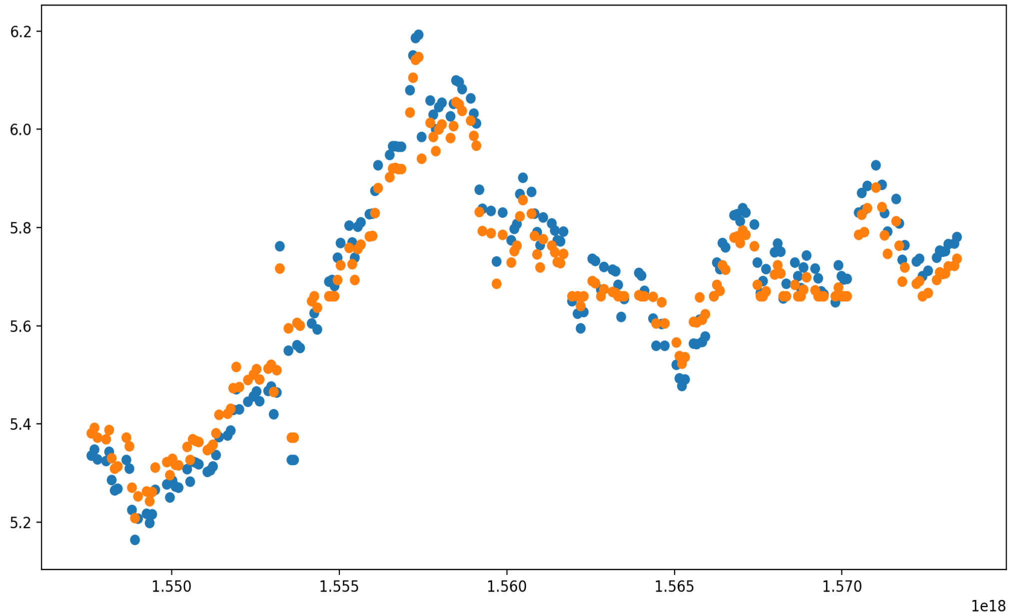
MAE is relatively insensitive to outliers compared to RMSE, while RMSE magnifies the bigger errors and ignoring the smaller errors. For that reason, MAE may accommodate bias, while RMSE is related to variance. When MAE and RMSE compared, it is possible to observe underfitting and overfitting issues that have been targeted by the application of MCCV method. Underfitting and overfitting is the situation of high bias, low variance, and low bias, high variance respectively. The difference between RMSE and MAE results, 0.072623 and 0.069744 respectively is 4.13%, which is lower than the 5% criteria. Therefore, the difference is in acceptable range, and denotes that model has fitted without any underfitting or overfitting issues.

The accuracy of the prediction has also been measured by the MAPE results. MAPE results, 0.012321 that is equivalent to 1.23%, which is lower than the 0.05 criteria, equivalent to 5%. Therefore, it is possible to say that the model has highly acceptable accuracy.

R^2 result is 0.900212, which is higher than the 0.90 criteria. Highly significant results have verified that the prediction is substantial with a superior accuracy, as the highly significant majority of the predicted data points are on the regression line.

Furthermore, the deviation between EVS and R^2 results, 0.935027 and 0.900212 respectively is only 0.62%, which is lower than 2% criteria. Thus the deviation between EVS and R^2 results signifies that the prediction is unbiased.

Fig. 63. Actual versus predicted USD/TRY prices obtained by analyzing recently published news and statements regarding on S-400 Crisis.



**Blue Points denote Actual USD/TRY Prices

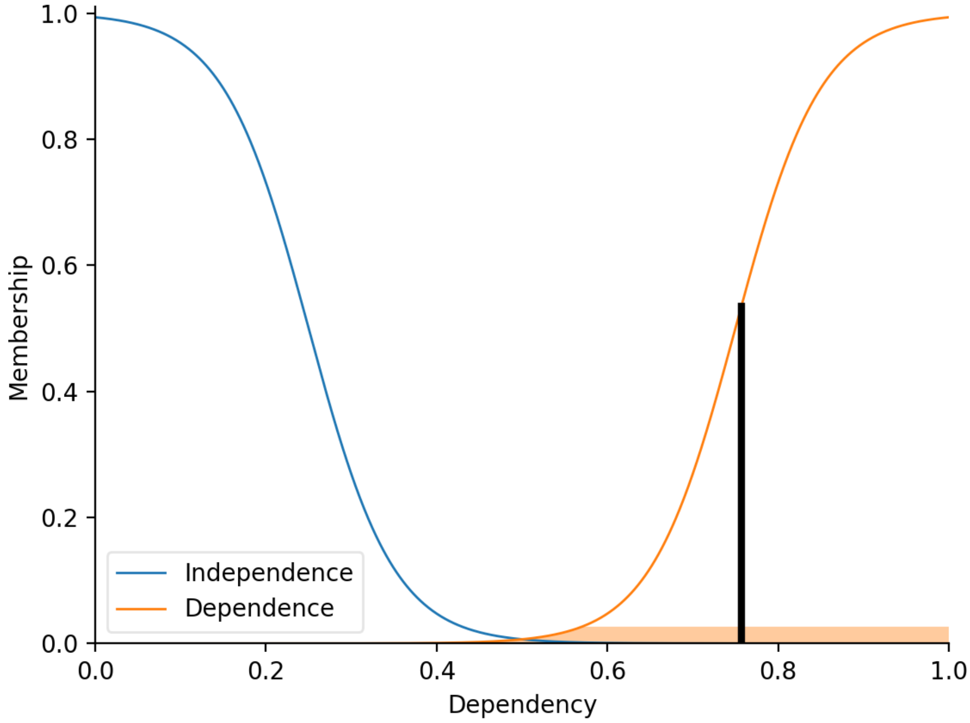
**Orange Points denote Predicted USD/TRY Prices

A certain disagreement between United States and Turkey has been observed by a highest margin during the previous phase, fuzzy logic. The clarity of certainty and disagreement have significant effect on achieving superior dependency rate, which has a direct impact on the minimization of MAPE, MAE, and MSE, alongside to augmenting prediction accuracy and excellent EVS and R^2 values. On the other hand, sudden shocks of uncertainty have also occurred occasionally, and caused sudden fluctuations and deviations, which were decisive during the measurements. Nevertheless, any fitting problems or biased calculations were not detected while dialing with sudden fluctuations on correlating values.

6.5.4 Fuzzy Logic and Support Vector Regression Results for 2019 Istanbul Mayoral Election

2019 Istanbul Mayoral elections were the most complicated, uncertain, and challenging electoral case to comprehend. This case is of great importance to comprehend decision makers' perception and interpretation on eventfully extended mayoral election process, and its causality on USD/TRY exchange rate. Fuzzy consequents regarding certainty, consensus, and dependency rate are targeting to uncover decision makers' perspective of valuation on breaking news regarding on eventful mayoral electoral process. Obtained averaged consequent values for 2019 Istanbul Mayoral elections signifies that the mentioned case is almost neither a member of agreement nor disagreement on consensus, and neither a member of certain or uncertain sets on certainty regarding to the perspective of decision makers'. Closeness of both certainty and consensus values to 3 represents that the correlation between USD/TRY exchange rate and recently published news on 2019 Istanbul Mayoral elections is almost diminished into the set Neutral, which indicates a steady state of exchange rate. However, when fuzzy antecedents examined comprehensively, it is observed that there are specific days that cause the disharmony on averaged correlation to be a perfect member of neutral set. These particular days directly addresses days right after the election day, which election results officially be announced when the vote counting ends. On those days, significant correlation between exchange rate and published news has detected, while gradually increasing, and diminishing causality has detected pre and post days of announcement respectively. Simulation results for 'certainty' and 'consensus' are visually presented in Figure 64 and 65. Dependency rate between 2019 Istanbul Mayoral elections and USD/TRY exchange rate has been simulated within the scope of separately obtained certainty and consensus values concerning eventful mayoral elections, and related news articles that were transferring the process. Simulation results indicated a dependency value of 0.718 with a minimal membership for dependence set for the causality between published news regarding to the mayoral elections, and fluctuations on USD/TRY exchange rate throughout the election process. Dependency rate is visually presented below, as Figure 66.

Fig. 66 Dependency Rate between Published News Related to 2019 Istanbul Mayoral Election and USD/TRY Exchange Rate Fluctuations



Attained dependency rate from the fuzzy logic phase is going to be used as an input alongside to historical exchange rate data to run SVR. The linearity of the provided dataset is evaluated by PCA/PCR and PLS methods. Figure 67, which presented in the appendix section demonstrates PCA/PCR and PLS results. First principal component in PCA/PCR and PLS indicates the axis in the K-dimensional variable space that accommodates the largest variance of the samples in a direction, while second principal component is orthogonal axis to the first principal component in the K-dimensional variable space, and identifies the second largest source of samples in a direction that can be denoted as the residual variance of the samples that improves the approximation. Besides, the direction that provides the lowest variance is aimed to be captured by the PLS. The nonlinearity of the provided dataset has confirmed by the tests applied. For that reason, kernel function for the SVR analysis is going to be assigned accordingly in order to prevent heteroscedasticity and multicollinearity. The application of Gaussian Radial Basis Function on PCA and the projection difference is exemplified as Figure 35 in the appendix section. Provided USD/TRY data has been projected into high-dimensional space by the kernel functions, and hyperplane has been constructed to provide adjustability to converge optimal set of weights (Oliveira et al., 2017). Gaussian Radial Basis Function utilized as kernel function due to its good fit on non-linear inputs (Grigoryan, 2016). Selected kernel function tuned by $C, \varepsilon, \text{and } \gamma$ hyperparameters, and grid search has been adopted to set the best-suited set of hyperparameters, and eliminate the probability of multicollinearity, besides attain the uppermost accuracy rate possible. Grid search ranges of $C, \varepsilon, \text{and } \gamma$ hyperparameters assigned as suggested by Hsu, Chang, and Lin (2003). *C* ranges from 1 to 10^5, and γ from 2^{-10} to 2000.

Grid search with cross validation for hyperparameter $C, \text{and } \gamma$ value that performed best specifically for this case is accessible as Table 18 in the appendix section. Grid search cross validation accuracy heat map is also presented as Figure 68 in the appendix section. Hyperparameter ε is assigned by considering the highest change in the historical data, and determined as high as possible in order to eliminate any possible overfitting issue. Accuracy of the parameter selection using grid search with cross-validation is evaluated by consideration of three metrics, namely; precision, recall, and f1-score. Results have indicated the accuracy of 0.96 for parameter selection, and detailed in Table 19, which has been presented below.

Tab. 19 Accuracy of the Parameter Estimation Using Grid Search with Cross-Validation for 2019 Istanbul Mayoral Elections

The model is trained on the full development set.
The scores are computed on the full evaluation set.

	precision	recall	f1-score	support
0	0.94	1.00	0.97	30
1	1.00	0.90	0.95	20
Accuracy			0.96	50
Macro Avg	0.97	0.95	0.96	50
Weighted Avg	0.96	0.96	0.96	50

The critical values of Accuracy, Precision, Recall, and F1 Score are 0.20, 0.40, 0.60, and 0.80 denotes slight, fair, moderate, substantial, and almost perfect agreement respectively.

Splitting provided data into training and test groups is essential to appraise the performance of the regression model. To do so, Monte-Carlo cross-validation (MCCV) method adopted. 2019 Istanbul Mayoral elections provided delimited historical data during its relatively short, yet substantial continuation. For that reason, 85% of provided data used as train data, while 15% reserved as test data. MCCV randomly splits reserved portion of data into sub-samples and assign them as test sets. The process of selecting random independent partitions repeated for multiple times. Monte-Carlo cross-validation run numerous times, and statistical test results attained for each run. Achieved results for statistical test are presented exhaustively in Table 20 in the appendix section. Prediction consistency of the model is evaluated against seven key statistical metrics, which are widely be accepted and using to validate machine learning prediction performance, namely; V Measure Score (VMS), Explained Variance Score (EVS), Mean Squared Error (MSE), Root Mean Squared Error (RMSE), Mean Absolute Error (MAE), Mean Absolute Percentage Error (MAPE), and Coefficient of Determination (R^2) . Subsequently, obtained statistical test results averaged. Averaged results for MCCV statistical test results are presented in Table 21.

Table 21. Averaged Monte Carlo Cross Validation Statistical Test Results for 2019 Istanbul Mayoral Elections.

VMS	0.765247*
EVS	0.759593*
MSE	0.009110
RMSE	0.095443**
MAE	0.084592
MAPE	0.014609***
$\boldsymbol{R^2}$	0.759524**

† Exhaustive Monte-Carlo cross validation statistical test results are presented in the appendix section, as Table 20.
The critical values of VMS are 0.70, 0.80, and 0.90. Therefore, *, **, and *** denotes threshold, significant, and highly significant at 0.70, 0.80, and 0.90 respectively.
The critical values of EVS are 0.60, 0.80, and 0.90. Therefore, *, **, and *** denotes threshold, significant, and highly significant at 0.60, 0.80, and 0.90 respectively.
The critical values of RMSE are set by SI at $\leq 10\%$, and $\leq 5\%$, which indicate RMSE value of 0.581499, and 0.290749 respectively. Therefore, *, and ** indicates the threshold and significance at respective critical points.
The critical values for the difference between RMSE and MAE at ≤5%, and ≤1% denote acceptable and perfect fit of the model respectively.
The critical values of MAPE are 0.25, 0.10, and 0.05. Therefore, *, **, and *** denotes threshold, low but acceptable, and highly acceptable accuracy at 0.25, 0.10, and 0.05 respectively.
The critical values of R^2 are 0.50, 0.75, and 0.90. Therefore, *, **, and *** denotes weak, moderate, and substantial prediction at 0.50, 0.75, and 0.90 respectively.
The critical value of the difference between EVS and R^2 is $\leq 2\%$. Therefore * denotes acceptable bias at $\leq 2\%$.

In purpose of validating the results, alongside of measuring the sustainability of the results, the Monte Carlo Cross Validation has repeated 40 times, the mean of achieved statistical test results has been taken and presented above. The mean of MCCV statistical test results indicates that;

Complexity of the 2019 Istanbul mayoral elections, and the uncertainty of its possible consequences on further economic policies and exchange rate have restrained clustering task to perform at its peak levels. VMS result is 0.765247, which is higher than 0.70 criteria. Therefore, it is possible to denote VMS value as acceptable thus indicates a harmonic mean between homogeneity and completeness and indicates an accomplished clustering task for the future weight support vector regression. Additionally, significant VMS result, 0.765247 indicates no multicollinearity issue, and authorizes the model for further tests and regression analysis.

EVS result is 0.759593, which is higher than 0.60 criteria. Therefore, it is possible to signify EVS value as acceptable. The results indicate that the variance can be explained by the factors presented by the actual data. Moreover, EVS is another measure that is crucial for regression analysis, as it scores homoscedasticity. Acceptable EVS result, 0.759593 denotes that the variance can be explained by the factors presented by the actual data, and not heteroscedastic.

MSE result is 0.009110. As it has been emphasized on the Chapter 4, MSE closer to 0 indicates better fit of regression line, as it indicates the variance of residuals. However, as there isn't acceptable range set for MSE as it does not shares the same unit as the original values, MSE has been used to achieve RMSE values.

RMSE value is in the same unit with the original data, and indicates the standard deviation of errors emerged during prediction, thus signifies the accuracy of the model. RMSE result is 0.095443, which denotes that the standard deviation of error is 0.072623 when compared to actual USD/TRY value. In order to evaluate RMSE result, the value transformed into interpretable value, SI. Transformed RSME, SI value is 0.016413, which is lower than 0.05 criteria. Therefore it is possible to denote RMSE value as significant.

MAE is relatively insensitive to outliers compared to RMSE, while RMSE magnifies the bigger errors and ignoring the smaller errors. For that reason, MAE may accommodate bias, while RMSE is related to variance. When MAE

and RMSE compared, it is possible to observe underfitting and overfitting issues that have been targeted by the application of MCCV method. Underfitting and overfitting is the situation of high bias, low variance, and low bias, high variance respectively. The difference between RMSE and MAE results, 0.095443 and 0.084592 respectively is 12.8%, which is higher than the 5% criteria. Underfitting issue is detected, as the difference is higher than the criteria. This indicates that the model was unable to detect the relationship between actual and predicted values accurately.

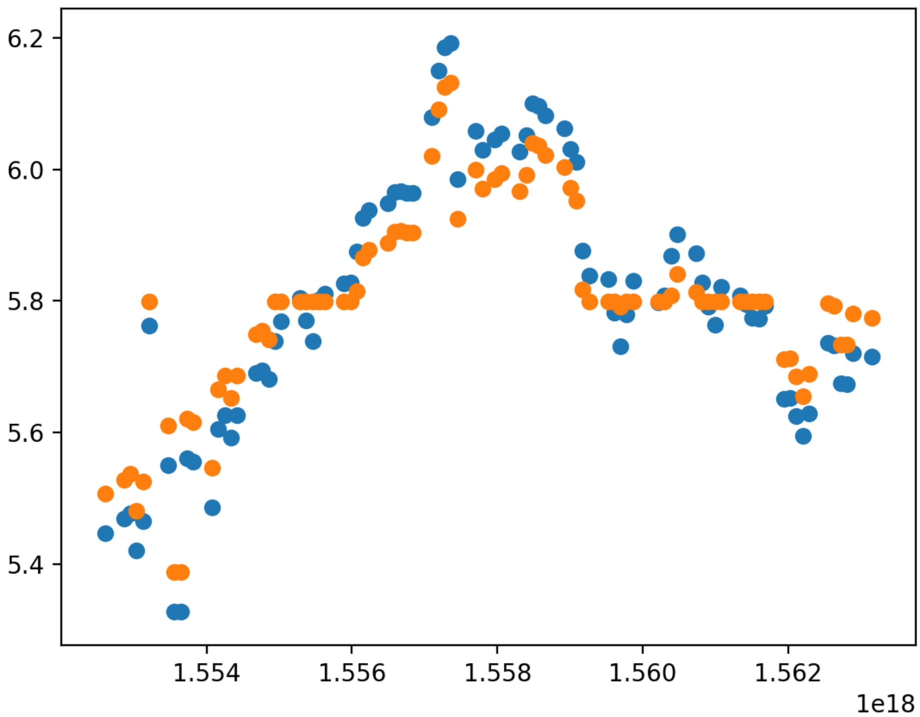
The accuracy of the prediction has also been measured by the MAPE results. MAPE results, 0.014609 that is equivalent to 1.46%, which is lower than the 0.05 criteria, equivalent to 5%. Therefore, it is possible to say that the model has highly acceptable accuracy.

R^2 result is 0.759524, which is higher than the 0.75 criteria. Moderate results have verified that the prediction is acceptable and accurate, as the majority of the predicted data points are on the regression line.

Furthermore, the deviation between EVS and R^2 results, 0.759593 and 0.759524 respectively is only 0.009%, which is lower than 2% criteria. Thus the deviation between EVS and R^2 results signifies that the prediction is unbiased.

Successful statistical test results have been attained in an election period of high political tension and uncertainty, which supports the hypothesis of exchange rate fluctuations under politically uncertain circumstances. On the other hand, EVS and R^2 results contrary to expectations, and indicating moderate prediction performance. Findings from further investigations indicates that no matter how chaotic and democratically questionable the election process is, the election process only considered by the investors' as an uncertain and risky environment, when the election day gets closer. It may also clearly be observed that, as the election day approaches, consensus turns into favour of agreement, while previous days were considering as a member of disagreement set. While these changes do not cause any dramatic fluctuations on exchange rate, the fact that elections' influence on exchange rate cannot be overlooked, as any potential cabinet change can straightforwardly influence government authoritarianism, policy, regime, and democracy which may directly distress investor's assertiveness. In other words, not the election process as a whole, but only the expectations on election results, and election results themselves matters for the decision makers, which may also explain the reason how the model underfit, but still manage to achieve acceptable test results. Prediction performance is also supporting this argument, as the prediction accuracy increase as the election day approaches, and hit up to superior rates, while prediction performed poorly during the days between prior and next election days.

Fig. 70. Actual versus predicted USD/TRY prices obtained by analyzing recently published news and statements regarding on 2019 Istanbul Mayoral Elections.



**Blue Points denote Actual USD/TRY Prices

**Orange Points denote Predicted USD/TRY Prices

6.6 Conclusion

This chapter emphasized the significance of particular political events on exchange rate fluctuations. Principally, four political events that have been homogenously distributed as national and international have been examined throughout the chapter, namely; Pastor Andrew Brunson case, 2018 Parliamentary and Presidential elections, S-400 crisis, and 2019 Istanbul Mayoral elections. The core intention is to conceive the weight of various political event categories on investors' decision-making process, and how it impacts the volatility of exchange rate. To accomplish this intention, officials' statements, recently published news articles regarding to aforementioned political cases, and historical exchange rates have been examined under the light of advanced hybrid machine learning algorithm that consists from natural language processing (NLP), fuzzy logic (FL), and supports vector machines (SVM). The methodology encapsulate learning automata for textual analysis, prospect theory for associating the textual analysis scores with change in exchange rate prices, fuzzy logic for simulating degree of certainty and consensus for every published political news, and deliver the degree of dependency between political event and exchange rate fluctuation. Support vector regression is practiced to predict exchange rate by obtained correlation value, alongside the historical exchange rate data in order to evaluate the accuracy of the algorithm, as well as predictability of exchange rate by analyzing news articles.

As selected political events have a partial membership for multiple categories, such as; judicial system, international politics, international relations, financial sanctions, sanctions on government officials, warfare, defense industry, political tension, human rights, NATO, elections, cabinet changes, and democratic uncertainty, findings provided a broad spectrum on political events under the extensive degree of certainty and consensus.

Findings derived from examined political events indicate that; recently published news that concerns certain disagreement on international politics that composed under uncertain environment far from compromise plays a vital role and has a direct rapid impact on exchange rates. While Pastor Andrew Brunson case enlightened the algorithm concerning the categories of international judicial process, court hearings, expected court orders, political, sanctions on government officials, and financial sanctions under every possible degrees of certainty and consensus, S-400 crises has contributed to algorithm by inspecting the categories of warfare, defence industry, political tension, human rights, and NATO under certainty and disagreement. Besides, 2018 Parliamentary and Presidential elections, and 2019 Istanbul Mayoral elections have enhanced the algorithm concerning the categories of political instability, elections, Supreme Court, cabinet changes, and democratic uncertainty.

The lucidity of certainty and disagreement have substantial impact on attaining superior dependency rate that directly originates minimization of MAPE, MAE, and MSE, alongside to augmenting prediction accuracy and excellent EVS and R^2 values. Instead, unexpected shockwaves of uncertainty transpired sporadically have triggered abrupt fluctuations and deviations, which were influential throughout the measurements. Moreover, statistical test results indicated a valid and accurate prediction without the issues of multicollinearity, heteroscedasticity, biasness, and underfitting or overfitting.

The results of this chapter unambiguously authenticate that varieties of political tension plays a vital role and has a direct rapid impact on exchange rate fluctuations. The correlation between obtained dependency rates, and prediction accuracy of each predicted event have also validated the superiority of the algorithm and significance of political events and regarding news. Furthermore, hypothesis of which supports that; not only the objective quantitative variables, but also subjective qualitative information can be used to obtain enhanced prediction performance by examining decision makers' behavior on particular scenarios has been validated.

Discoveries of the chapter also serves the purpose of creating a behavioral database for each political event categories derived from daily-published news and exchange rate fluctuations to prepare a foundation for complex algorithms that are aiming to deal with much boarder spectrum of national and international political events under the extensive degree of certainty and consensus.

CHAPTER 7

The Correlation Between Recently Published Political News and Exchange Rate Fluctuations

7.1 Introduction

Previous chapter have concentrated on the influence of specific types of political events, and political tension and uncertainty caused by them on the exchange rate fluctuations. Chapter predominantly concentrated on four main political events that occurred successively in a very short date interval, namely; Pastor Andrew Brunson case, Parliamentary and Presidential elections, S-400 crisis, and eventful Istanbul Mayoral elections. Selection of political events have meticulously done and homogenously distributed as national and international in order to comprehend the impact of national and international events, under the influence of various kinds of political cases. Once the content of selected political events examined in depth, it is possible to observe interconnection of political events with the political categories such as; judicial system, international politics, international relations, financial sanctions, warfare, defense industry, elections, cabinet changes, and democracy under the influence of political tension, disagreement, and uncertainty on involved countries economics, domestic currency, and exchange rate. Likewise, the significance of homologous political events on economics and exchange rate has also supported by the existing literature, which has been presented throughout the previous chapter. The chapter actively revealed dramatic impact of political tension and uncertainty on exchange rate fluctuations under the influence of specific political events. Besides, chapter also propounds the weights for particular political categories through attained statistical test results.

Fundamentally, an exquisite database have composed and shaped in the process of each of three phases that has shaped the methodological workflow of previous chapter, which provides a boarder view on the impact of politics and political events on economics, and can be subsidiary for further sophisticated investigations. The first phase of the methodology, sentiment analysis has contributed the database by constructing unique lexicons derived from each political event whereas an additional hidden layer of algorithm has classified the news articles into political categories. Matching the polarity vectors with historical data comprehended the subjective valuation of decision makers', which enhanced the lexicons' polarity function, and make them applicable not only for particular political event, but political categories under the influence of any political tension variants, even if only single annotated text have been categorized under that variant of particular category. Second phase fuzzy logic has subsidized by providing dependency rate of exchange rate on political news under various levels of certainty and consensus. Final phase, support vector regression has contributed to database by evaluating the accuracy of the clustering, errors, and prediction that has been attained under the supervision of the outcomes obtained by the aforementioned phases.

This chapter is constructed on the foundation that has been laid by the previous chapter concerning specific political events, cases, and categories, and emphasizes on the correlation between officials' statements, speeches addressing the nation, speeches addressing government agencies and officials, recently published political news on media regardless of the news' political category, event, or case involving Turkey and/or predominantly concerning Turkey and United States, but also Russia, Middle East, and Europe subsequently, and instantaneous USD/TRY exchange rate fluctuations concerning the dates 29 December 2017 and 01 November 2019. Tremendous depreciation of TRY against USD has occurred during the aforementioned date interval by enormous 152% that is quite extraordinary to be happened in a very short time interval compared to mentioned depreciation percentage, and has to be scrutinized thoroughly.

In this chapter, responsiveness of exchange rate to government officials' statements, speeches addressing the nation, speeches addressing government agencies and officials, and recently published political news are going to be assessed under the degrees of certainty and consensus considered from the perspective of investors. The database, which has been attained by the hidden layer of algorithm embedded into the hybrid machine learning code structured for previous chapter, as well as the outcome achieved in the previous chapter is going to enforce the algorithm of this chapter, and guide for further behavioral analysis on investors' evaluations on various degrees of certainty and consensus on aforementioned kinds of political news, and measure their instinctive response, thus exchange rate volatility comprehensively.

The main purpose of this chapter is to deeply investigate the instinctive attitude of investors when the political breaking news took place on the media regardless of political category and presentation form of the news, and use the findings to evolve the hybrid machine-learning algorithm to compose an exchange rate prediction technique that predominantly involves diverse disciplines, and practices both qualitative and quantitative data. To achieve this purpose, an advanced hybrid machine-learning algorithm compared to previous chapter's algorithm has been generated by implementing the findings of the previous chapter and using adapted forms of machine learning techniques, namely; natural language processing (NLP), weighted prospect theory (WPT), weighted fuzzy logic (WFL), and feature weighted supports vector machines (FWSVM) to unveil the significance of any kind of political tension on exchange rates fluctuations, and run daily exchange rate predictions consequently.

The chapter is structured as follows: Section 2 briefly describes data collection and fundamental variables; Section 3 describes application of methodologies, Section 4 presents empirical results, and prediction of USD/TRY exchange rate, and Section 5 concludes the chapter.

7.2 Data and Fundamental Variables

For sentiment analysis, HTTP REST API has extracted daily news articles regarding to each aforementioned political cases, and also any further miniscule cases that might not by substantial as others, yet still a conceivable factor that may influence the decision makers' decision from leading news websites BBC, Reuters, Bloomberg, The Guardian, and The New York Times. 821 news articles have been analyzed, and textual content transformed into numeric sentiment inputs.

In the preeminent curiosities to refine understanding on the effects of political tension, instability, uncertainty, and consensus triggered by aggregate factors, USD/TRY exchange rate is going to be investigated daily. In order to perceive the degree of responsiveness of exchange rate to recently published government officials' political and international relations related official verbal or written statements; daily price, daily change in percentage, opening price, highest price, lowest price, and difference between highest and lowest prices of USD/TRY be investigated for each day.

Especially opening price, close price, highest price, lowest price, and the difference between these prices plays a crucial role on overcoming any possible accuracy error on the ML algorithm caused by the coincidence of news content on opposite polarities that may take place on the same day. The accommodation of news on opposite polarities on the same day may neutralize the significant effect of vital news, and cause a dramatic error on further calculations of the algorithm. To cope with this issue, the philosophy behind the candlestick chart has been adopted. In this case, even if neutralization happened in any day during the selected date interval, algorithm will detect it by observing the volatility of the exchange rate price throughout the day, and weight related news accordingly to match the volatility and the price by considering previous/further news on both subjects. Abovementioned exchange rate data have been collected over the period between 29 December 2017 and 01 November 2019, which is equivalent to 4032 variables.

7.3 Methodology

When compared to previous chapters, the methodology of this chapter has evolved to meet the superior requirements arising by the complex structure of this chapter, and designed to be able to compute the influence of varying grades of political tension regardless of the political category the tension involved on exchange rate fluctuations, and predict exchange rates accordingly.

The methodology for this chapter is involves three stages. The inaugurating stage is the “Sentiment Analysis”, which adopted Natural Language Processing for textual analysis, which is reinforced by Weighted Prospect Theory. Outcome of Sentiment Analysis is going to be examined by the second stage, which is “Weighted Fuzzy Logic” in order to attain a weighted correlation value. Last stage “Feature-Weighted Support Vector Regression” is going to run regression, apply statistical tests to perceive validation and accuracy of the model, and predict future exchange rates.

Cited stages that constituents the methodology, and accommodated under the workflow of this chapter has clarified thoroughly in Chapter 4. Sections and subsections of Chapter 4 regarding to relevant stages of the methodology are going to be indicated below.

Primary stage, “Sentiment Analysis” (Chapter 4.2), practices the components of Natural Language Processing. Lexicons (Chapter 4.2.1) for this chapter have been provided by the output obtained on previous chapter, Chapter 6.5. Learning Automata (Chapter 4.2.2) is adopted for textual analysis, which reinforced by Weighted Prospect Theory (Chapter 4.7.1), which targets correct application of weights. Outcome of Sentiment Analysis is going to be providing input for second stage, “ Weighted Fuzzy Logic” that adopts fuzzy logic (Chapter 4.7.2), and fuzzy weighted average within a generalized mean operator (Chapter 4.7.2.1) algorithms of machine learning, and going to achieve weighted correlation value between recently published political news and exchange rate. Final stage, “Weighted Support Vector Regression” is composed by enhancing the SVR to WSVR with a blend of methodologies (Chapter 4.4.1, and 4.7.3) by using the presented weighted correlation values is going to run support vector regression. In order to split data into training and test sets, Monte Carlo Cross Validation (Chapter 4.5.1.1) is going to be adopted. Hyperparameters for WSVR is going to be acquired by Grid Search with Cross Validation (Chapter 4.5.2.1). Grid Search with Cross Validation method plays a crucial role on fine-tuning the hyperparameters for the kernel function cautiously and suitably to provide most accurate hyperparameter selection, while eliminating the heteroscedasticity and multicollinearity in regression analysis. The accuracy of the obtained parameters is going to be validated by the metrics, namely; Accuracy, Precision, Recall, and F1 Score (Chapter 4.5.2.2). The linearity of the provided dataset is going to be evaluated by Principal Component Analysis (PCA), Principal Component Regression (PCR), and Partial Least Square (PLS) methods, and kernel function for the model is going to be assigned according to the linearity of the dataset. The performances of Polynomial, and Gaussian Radial Basis kernel functions are going to be compared, and the one that provides the most accurate hyperparameter selection that eliminates heteroscedasticity in regression is going to be adopted for the model. The performance of the fabricated model is going to be assessed by the statistical tests, namely; V Measure Score (VMS), Explained Variance Score (EVS), Mean Squared Error (MSE), Root Mean Squared Error (RMSE), Mean Absolute Error (MAE), Mean Absolute Percentage Error (MAPE), and Coefficient of Determination (R^2) (Chapter 4.6), and future exchange rates are going to be predicted. The workflow mechanism of the presented methodology is going to be detailed in the following subsection.

7.3.1 Workflow Mechanism of Methodology

Findings of the previous chapter have played a vital role on understanding the multidirectional impact of various political environments on exchange rate, and constructing enhanced methodology for further investigations accordingly. In order to perform the purpose of this chapter, officials' statements, speeches addressing the nation, speeches addressing government agencies and officials, and recently published political news on media regardless of the news' political category, event, or case in the aforementioned date intervals by a feverish work of complex algorithms.

The main determination of the first stage, Sentiment Analysis is to extract and separate collected annotated news regarding to their political categories, and weight them accordingly. To do so, annotated text collected news are going to be feed into Learning Automata. Learning Automata is going to use unique Lexicons that have been constructed in Chapter 6. These Lexicons are constructed uniquely to perform under the complex structure of diversified causality for political tension that involves Turkey. Unique Lexicons have been developed through the embedded hidden layer of algorithm into Chapter 6's python code, political categories associable with Turkey's current political tension, like; judicial system, international politics, international relations, financial sanctions, warfare, defense industry, elections, cabinet changes, and democracy has been classified under the degree of certainty and consensus. This application during the construction phase of Lexicons, which has taken place in Chapter 6, strengthened the performance of Sentiment Analysis stage of this chapter. The word polarity function for Lexicons concentrated on particular categories has attained harmoniously by considering findings of Chapter 6, which unlocked the power of enrichment and integration of Lexicons under the influence of diversified political environments. Sentiment extraction is going to take place by inspecting the annotated text supplied previously, and produces an overall polarity vector value for every text feed into algorithm. Positive expressions of polarity vector values are normalized into positive and negative values by normalized polarity vector. Normalized polarity vector values that have been attained will provide probability of the sentiment outcome. As Chapter 6 suggested, the process of involving historical data to derive the subjective feelings of the decision makers have been bypassed completely in this chapter, as the hidden layer embedded into code of previous chapter have already composed required database to supervise the algorithm to behave accordingly under the influence of various political environments. In this chapter, in order to comprehend the decision making reference point of each attribute of decision makers', attained probability values are going to be fuzzified and weighted accordingly, as each breaking news and/or official statement's hierarchical order of importance that determines the decision makers' attitude is fuzzy in itself. This will provide an insight of weighting and valuation approach of decision makers to the algorithm. Obtained weighted fuzzy values are going to be used as an input for prospect theory, and weighted prospect values are going to be generated. Generated weighted prospect values will be the output of sentiment analysis, which is indicating the subjective assessment of decision makers. The output of sentiment analysis is going to be used as an input on the second phase, weighted fuzzy logic. As the political news have been evaluated collectively, clustering the fuzzy antecedents is necessary. Clustering aims to cluster the fuzzy antecedents, which has been obtain on previous stage sentiment analysis depending on the subjective weighting and valuation of decision makers. Performing clustering task will distinguish the antecedents into cluster of similar weights and importance, and these clusters will be arranged hierarchically. A hierarchical structure is necessary to identify the evaluation in a multi-criteria decision-making scenario, as the significance of criteria entirely depends on the subjective weighting and valuation of decision maker. Each cluster that forms a step in hierarchy will be evaluated separately under the methodology of fuzzy logic. As it has previously be suggested, the content of the political news are going to be considered in two main themes that psychologically affects the subjective feelings of the decision maker, which were also proven by the existing literature. These are the degree of certainty and the degree of consensus that encountered relatively as the main viewpoints of the published news. For that reason; fuzzy antecedents have been distributed into two main classes for political cases, which are; Certainty and Consensus. The membership function for Certainty and Consensus divided into three main sets. The membership function for

Certainty defined as Certain, Neutral, and Uncertain, while defined as Agreement, Neutral, and Disagreement for Consensus. The transition between the memberships of the determined sets for each fuzzy antecedent, aside from the degrees of truth for each defined set that have been mentioned above are going to be defined by the upper and lower bounds, width, concave, convex, and the steepness of the associated membership, which will be presented on the future graphs.

After assessing the membership function for the fuzzy antecedents in each cluster, fuzzy weighted average within a generalized mean operator is going to approximate membership values by hierarchical evaluation method. By weighting the parameters within a hierarchy, altered valuation of the decision maker is going to be obtained by weighted generalized mean operator, which indicates the decision makers' behavior of evaluation. The fuzzy relations are going to be obtained by the presence or absence of association, interaction, or interconnectedness between the elements of two or more sets. Interpretation of the resulting fuzzy relations may be obtained by the fuzzy rules designated by the author, which are going to be feed into fuzzy engine as well as membership functions. Fuzzy engine is going to generate Fuzzy Consequent, which is essentially a dependency of USD/TRY behavior on recently published news, and official statements. In other words, fuzzy consequent signifies the correlation between USD/TRY and news and officials' statements that are about degrees of political consensus and certainty, and/or have possibility to cause political tension.

Correlation value that has been achieved, as fuzzy consequents will be used as input in addition to historical USD/TRY data to run feature weighted support vector regression. The linearity of the provided dataset is going to be examined by Principal Component Analysis (PCA), Principal Component Regression (PCR), and Partial Least Square (PLS) methods. PCA/PCR is a mono-dimensional, observed variable approach that aims to find a linear correlation by reducing dimensionality. PLS is a multidimensional, latent variable approach that suits best for when there is a multicollinearity. Both PCA/PCR and PLS methods can be used to tests outliers and sensitivity analysis to provide diagnostic tools for the model. Nonlinear USD/TRY data have been projected into high-dimensional space by using the kernel functions. Constructed hyperplane delivers adjustability to congregate optimal set of weights (Oliveira et al., 2017). Gaussian Radial Basis Function is selected as kernel function due to its good fit on non-linear inputs (Grigoryan, 2016). Gaussian Radial Basis Function tuned by hyperparameters that have been achieved by grid search with cross validation method, which also eliminated the possibility of heteroscedasticity and multicollinearity. Accuracy of the parameters attained by the grid search with cross validation is evaluated by accuracy, precision, recall, and f1-score metrics. The data will randomly be divided into training and test groups for multiple times in order to consider the performance of the regression model. To do so, Monte-Carlo cross validation (MCCV) method has been implemented to split presented data into training and test groups for prediction. Accuracy of the prediction has been investigated by the statistical tests. Prediction reliability of the model is measured against seven broadly adopted statistical tests for machine learning algorithms, namely; V Measure Score (VMS), Explained Variance Score (EVS), Mean Squared Error (MSE), Root Mean Squared Error (RMSE), Mean Absolute Error (MAE), Mean Absolute Percentage Error (MAPE), and Coefficient of Determination (R^2). As VMS signifies the clustering precision of the model, it is a pivotal measure to check primarily. Grander performance on homogeneity and completeness on clustering task is an essential in order to accomplish an accurate regression without multicollinearity issue. It is only appropriate to run further statistical tests and accurate prediction when satisfying VMS value is achieved, as it indicates the reliability of the algorithm. Attained VMS value specifies perception on the prejudices of measured deviation of errors. MSE, RMSE, MAE, and MAPE are calculated in order to measure the deviation between predicted and actual values. RMSE and MAE are considered simultaneously to identify potential deviation in the errors. While equivalent RMSE and MAE values specifies correspondence of magnitude on all errors, the alteration between the results specifies the variance associated with the individual errors. Larger the dissimilarity indicates higher variance, and occurrence dissemination of the specific errors. The differentiation between RMSE and MAE is also an indication of underfitting and/or overfitting issues. Additionally, MAPE is correspondingly reflected and intended to normalize the absolute error and obtain altered perspective on association with real values. While MSE, RMSE, MAE, and MAPE measure the deviation between predicted and actual value, R^2 and EVS targets inconsistency between predicted and actual value. Conversely, while R^2 uses raw sum of squares, EVS practices the biased variance. Moreover, the homoscedasticty can be evaluated by EVS result, as it evaluates the explainability of the variance by the factors presented in the actual data. Both variables can only be parallel if the mean of the residuals equals to zero, and prediction is unbiased. Conclusively, if the statistical tests' results are convenient, future exchange rates are going to be predicted in daily basis. The empirical workflow, which has been explained above is accessible as a flowchart in Figure 71.

7.4 Empirical Findings of the Hybrid Machine Learning Algorithms

Subjective perception of decision maker under diversified political events and situations has targeted to be considered through the sentiment analysis. To do so, the NLP algorithm has adopted unique Lexicons that have been constructed through Chapter 6, specifically to perform under the complex structure of diversified causality for political tension that involves Turkey. Attained NPV through the sentiment extraction have fuzzified and weighted accordingly, and reference points for each attribute of decision makers obtained. Obtained weighted fuzzy values have feed into prospect theory, to generate weighted prospect values. Generated weighted prospect values provided an enhanced overlook on the subjective assessment of decision makers, and also form the output of sentiment analysis, which have been used as fuzzy antecedents for the weighted fuzzy logic phase.

Decision makers' subjective assessments have been scored for every published news article through the sentiment analysis, which provided the output of sentiment scores that are fuzzified. Fuzzified sentiment scores are going to be clustered, hierarchically evaluated, and used as fuzzy antecedents on the second phase, weighted fuzzy logic.

Clustering aims to cluster the fuzzy antecedents based on the weights given through the previous phase, which evaluated the subjective weighting and valuation of decision makers. This process is going to be followed by arranging the clusters hierarchically in order to identify the evaluation in a multi-criteria decision-making scenario. When the hierarchical order of clusters have been set, the sentiment scores that has been achieved on previous phase is going to be feed into the algorithm. Sentiment scores have been assigned between the range of 1 and 5, where 1 designates highest likelihood for definitive severe depreciation of USD/TRY, gradual increase in the range between 1 and 3 signifies decline in depreciation of USD/TRY, while 3 denotes the steady state of the exchange rate. Scores from 3 to 5 on the scale symbolizes a gradual appreciation in USD/TRY, while 5 denotes significantly absolute sharp appreciation in USD/TRY. Fuzzy sets defined with membership functions for the antecedents ‘consensus’ and ‘certainty’, which indicates the subjective level of political consensus between sides, and subjective level of political certainty respectively. Political consensus and certainty have acute effects on international relations and economics. For that reason, trapezoidal membership function has selected to be used for fuzzy antecedents, while fuzzy consequent is defined with sigmoidal function.

Membership functions defined for each fuzzy antecedent. For the antecedent ‘consensus’, assigned membership functions are: ‘agreement’, ‘neutral’, and ‘disagreement’, while ‘certain’, ‘neutral’, and ‘uncertain’ are assigned as membership functions for antecedent ‘certainty’. Sets for membership functions assigned by the fuzzy engine, in order to determine the degree of belonging to a fuzzy set according to the consequence of the sentiment analysis. ‘Certainty’ and ‘consensus’ values between 1 and 3 denotes the degree of belonging to fuzzy subsets ‘certain’ and ‘agreement’, 3 defined as ‘neutral’, and values between 3 and 5 denotes the degree of belonging to fuzzy subsets ‘uncertain’ and ‘disagreement’ respectively. Degree of membership to a fuzzy set denoted between 0 and 1, While 0 and 1 represents 0% and 100% membership degree respectively, anything in between considered as a partial membership to belonging set.

The relationship between political tension and exchange rate has been indicated by the correlation values. Political tension, and the correlation between political tension and exchange rate have been evaluated on the basis of certainty, consensus, and dependency rates respectively. While certainty and consensus rates are providing deeper understanding of the content of the published political news, dependency rate is the fuzzy consequent, which is a noteworthy indicator to understand how decision makers' react to fluctuating political tension. Understanding the subjective standpoint of decision makers on diversities of political tension, and being able to designate it quantitatively plays a substantial role on running precise projections. Once weighing the membership function for the fuzzy antecedents in each cluster, fuzzy weighted average within a generalized mean operator is going to approximate membership values by hierarchical evaluation method. The fuzzy relations are going to be obtained by the presence or absence of association, interaction, or interconnectedness between the elements of two or more sets.

Attained fuzzy weighted averaged values for the degree of belonging signifies insignificant impact for both certainty and consensus, yet indicates a certain disagreement when the contents of the political news have been generalized with a mean operator. Besides, the membership value for consensus indicates a significance of disagreement, yet insignificance of certainty for the published political news that has been examined. It is possible to say that, certainty of the political disagreement between Turkey and U.S. have amplified the impact of disagreement on the exchange rate under the subjective judgment of decision makers on a certain situation that is less feasible to interpret. Decisively, Fuzzy consequent for the degree of political certainty and consensus of the news articles designates a certain disagreement between U.S. and Turkey. Simulations that has been run to comprehend the degree of political certainty and consensus for recently published political news has visually presented in Figure 72 and 73 respectively.

Fig. 72. Degree of Political Certainty for Recently Published Political News

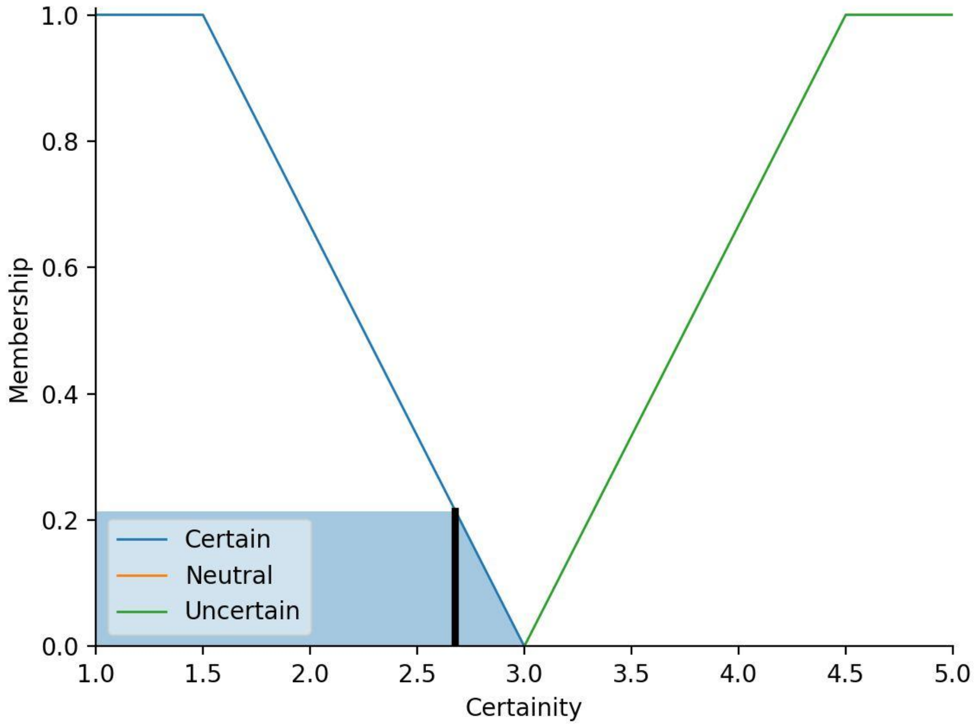
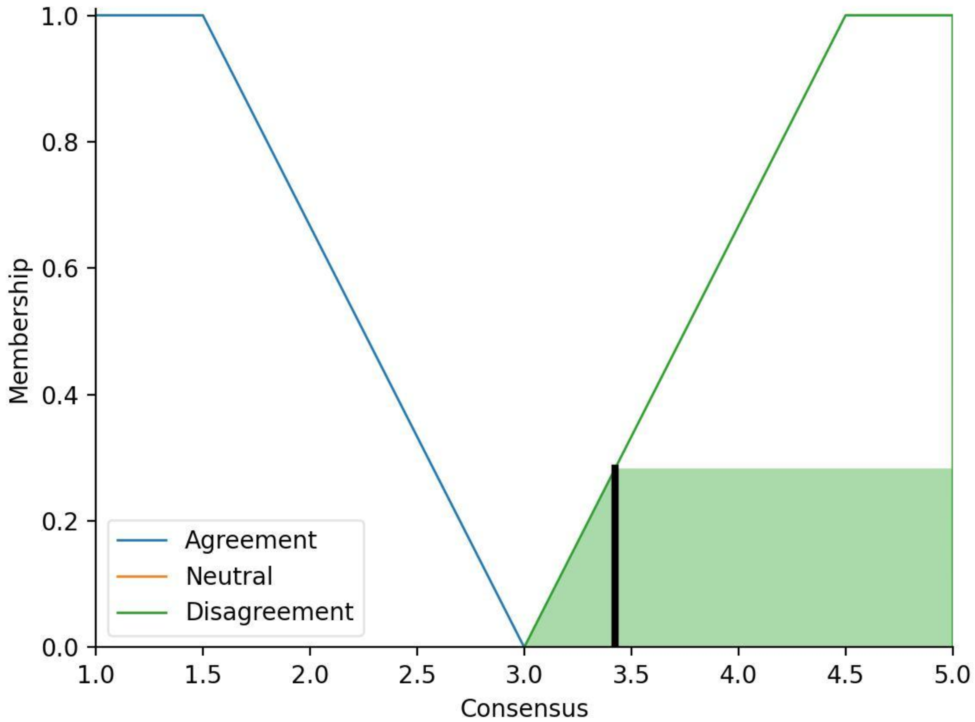


Fig. 73. Degree of Political Consensus for Recently Published Political News

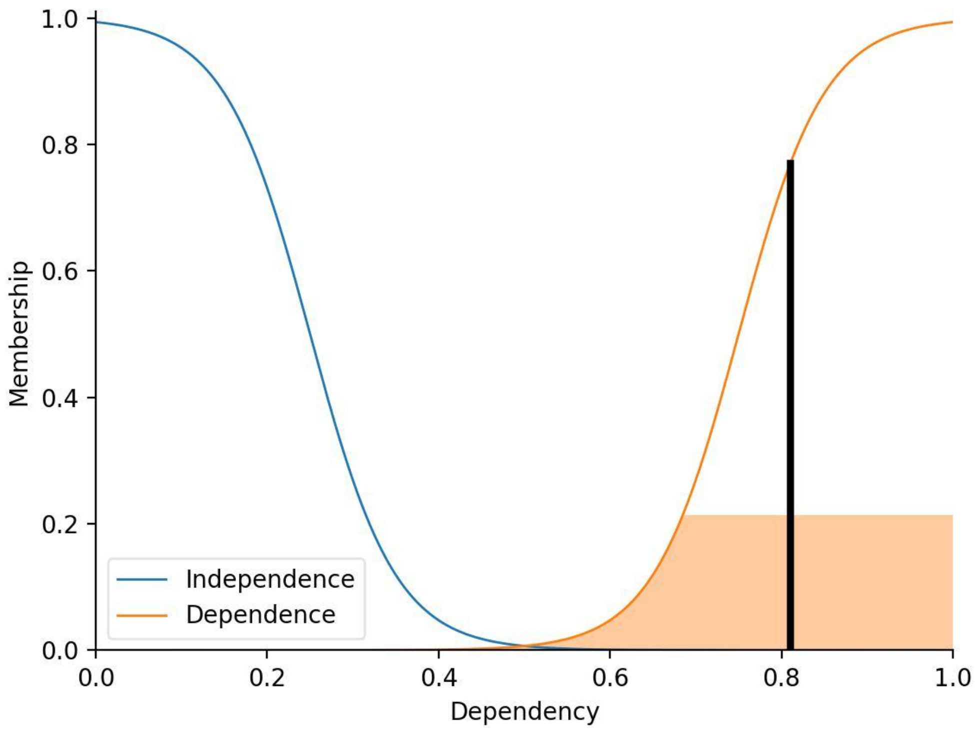


Simulation results of fuzzy weighted average within a generalized mean operator, which have been presented above exposed the inconsistency of political consensus and certainty on political news, thus verifies the political tension between Turkey and U.S.. When compared to previous chapter that has been evaluated the political cases individually, which indicates the uniqueness of each case, obtained generalized mean operator results are obviously expected, which are reflecting the nature of politically tense environment.

Correlation value political news is going to be generated by the combining the results derived from certainty and consensus values within the fuzzy engine based on the preset fuzzy rules. Correlation between political news and USD/TRY exchange rate is defined as dependency value, where value 0 defines that case related news and fluctuations in exchange rate is extremely independent, and value 1 defines that case related news and fluctuations in exchange rate is extremely dependent.

Averaged correlation values indicate a negative impact of political tension between Turkey and United States on the price of Turkish Lira, which causes an appreciation on the price of USD/TRY exchange rate. Dependency rate of USD/TRY exchange rate on recently published political news has also been simulated within the obtained certainty and consensus values. Simulation results indicate an exceptionally substantial dependency value of 0.846, and verify the hypotheses that exchange rate fluctuations are highly dependent on the published news and statements regarding to the political tension. Dependency rate of USD/TRY exchange rate fluctuations on the recently published political news is visually presented below, as Figure 74.

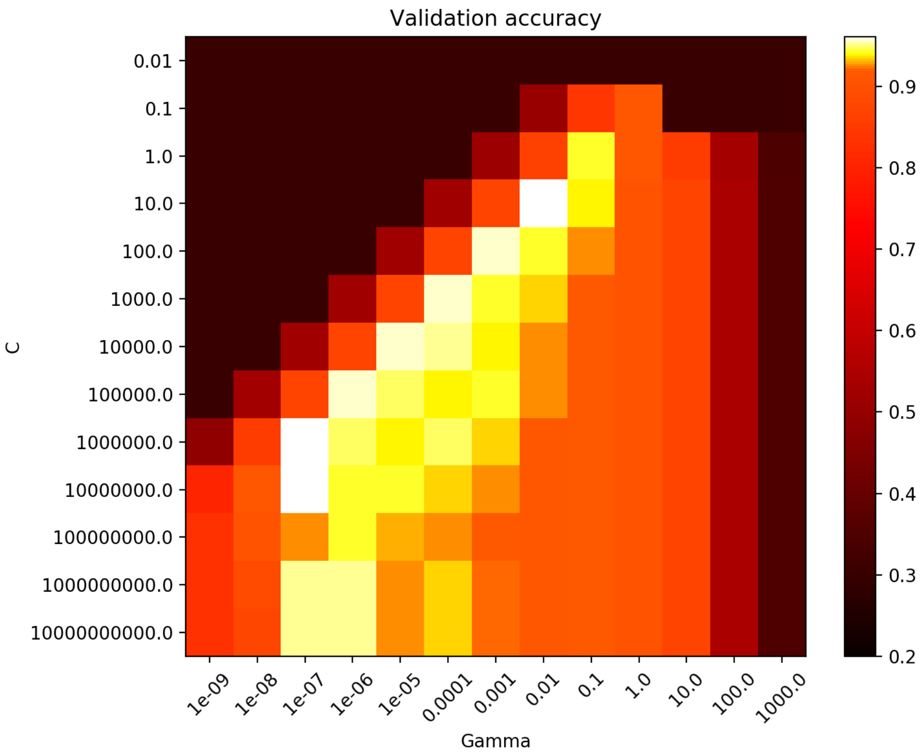
Fig. 74 Dependency Rate of USD/TRY Exchange Rate Fluctuations on Recently Published Political News.



On the final phase, previously achieved fuzzy consequent, weighted correlation value is going to be used as an input for feature-weighted support vector regression, as well as historical USD/TRY data. The linearity of the provided dataset is examined by the PCA/PCR and PLS methods. Figure 75 that visualized in the appendix section has presented the PCA/PCR and PLS results. First principal component in PCA/PCR and PLS indicates the axis in the K-dimensional variable space that accommodates the largest variance of the samples in a direction, while second principal component is orthogonal axis to the first principal component in the K-dimensional variable space, and identifies the second largest source of samples in a direction that can be denoted as the residual variance of the samples that improves the approximation. Besides, the direction that provides the lowest variance is aimed to be captured by the PLS. The nonlinearity of the provided dataset has confirmed by the tests applied. For that reason, kernel function for the SVR analysis is going to be assigned accordingly in order to prevent heteroscedasticity and multicollinearity. The application of Gaussian Radial Basis Function on PCA and the projection difference is exemplified as Figure 35 in the appendix section.

As a first step of FWSVR, nonlinear USD/TRY data has to be projected into high-dimensional space by using the kernel function. Constructed Hyperplane has been captured from various angles to enhance consideration of hyperspace, and comparability with traditional 2-D space. In purpose of enhanced comparability and understanding, USD/TRY data has visually been presented in the appendix section in both 2-D space and high-dimensional space once the hyperspace is constructed after achieving the optimal hyperparameters that has been selected by grid search with cross validation method. Figure 76 that takes place in the appendix section embodies the USD/TRY historical data as traditional 2 dimensional vectors in 2-D input space, while Figure 77, 78, and 79 illustrates the high-dimensional space. In order to converge optimal set of weights for constructed hyperplane, kernel function has to be used to deliver adjustability. Gaussian Radial Basis Function utilized as kernel function due to its good fit on non-linear inputs (Grigoryan, 2016). Gaussian Radial Basis function is formed by $C, \varepsilon, \text{and}\ \gamma$ hyperparameters that are need to be tuned. For superior tuning, grid search with cross validation method has been adopted to set the best-suited set of hyperparameters targeting to eliminate the possibility of multicollinearity and achieve highest accuracy rate for the hyperparameters possible. Grid search ranges of $C, \varepsilon, \text{and}\ \gamma$ hyperparameters assigned as suggested by Hsu, Chang, and Lin (2003). *C* ranges from 1 to 10^4, and γ from 2^{-10} to 2000. Grid search with cross validation for hyperparameter $C, \text{and}\ \gamma$ value selection is presented as Table 22 in the appendix section, and layers view for visualization of grid search cross validation accuracy is presented as Figure 80. Grid search cross validation accuracy heat map is also presented below as Figure 80. Hyperparameter ε is assigned by considering the highest change in the historical data, and determined as high as possible in order to eliminate any possible overfitting issue.

Fig. 80 Grid Search Cross Validation Accuracy Heat Map for Political News



Accuracy of the parameter selection using grid search with cross-validation is evaluated by consideration of three metrics, namely; precision, recall, and f1-score. Results have indicated the accuracy of 0.95 for parameter selection, and detailed in Table 23, which has been presented below.

Tab. 23 Accuracy of the Parameter Estimation Using Grid Search with Cross-Validation for Political News

The model is trained on the full development set.
The scores are computed on the full evaluation set.

	precision	recall	f1-score	support
0	1	1	1	21
1	0.91	0.97	0.94	30
2	0.95	0.96	0.95	24
Accuracy			0.95	75
Macro Avg	0.98	0.98	0.98	75
Weighted Avg	0.97	0.97	0.97	75

The critical values of Accuracy, Precision, Recall, and F1 Score are 0.20, 0.40, 0.60, and 0.80 denotes slight, fair, moderate, substantial, and almost perfect agreement respectively.

When the hyperplane is constructed, input data for future weighted support vector regression has needed to be separated into train and test sets before running regression analysis. Splitting feed data into training and test groups is essential to enhance the performance of the regression model. Monte-Carlo cross-validation (MCCV) method adopted to split the input accordingly. 80% of provided data used as train data, while 20% reserved as test data. MCCV randomly splits reserved portion of data into sub-samples and assign them as test sets. The process of selecting random independent partitions repeated for multiple times sufficient enough compared to the volume of the data, and consistency of the output.

Prediction consistency of the model is evaluated against seven key statistical metrics, which are widely be accepted and using to validate machine learning prediction performance, namely; V Measure Score (VMS), Explained Variance Score (EVS), Mean Squared Error (MSE), Root Mean Squared Error (RMSE), Mean Absolute Error (MAE), Mean Absolute Percentage Error (MAPE), and Coefficient of Determination (R^2). MCCV has run numerous times sufficient enough regarding to the volume of the data, and evaluated consistency of the output. Statistical test results attained for each run. Achieved results of statistical test for each run are presented exhaustively in Table 24 in the appendix section. Subsequently, obtained statistical test results averaged. Averaged results for MCCV statistical test results are presented in Table 25.

Table 25. Averaged Monte Carlo Cross Validation Statistical Test Results for Political News.

VMS	0.947610***
EVS	0.945922***
MSE	0.036425
RMSE	0.190830**
MAE	0.188578
MAPE	0.037097***
$\boldsymbol{R^2}$	0.939205***

† Exhaustive Monte-Carlo cross validation statistical test results are presented in the appendix section, as Table 24.

The critical values of VMS are 0.70, 0.80, and 0.90. Therefore, *, **, and *** denotes threshold, significant, and highly significant at 0.70, 0.80, and 0.90 respectively.

The critical values of EVS are 0.60, 0.80, and 0.90. Therefore, *, **, and *** denotes threshold, significant, and highly significant at 0.60, 0.80, and 0.90 respectively.

The critical values of RMSE are set by SI at $\leq 10\%$, and $\leq 5\%$, which indicate RMSE value of 0.520908, and 0.260454 respectively. Therefore, *, and ** indicates the threshold and significance at respective critical points.

The critical values of the difference between RMSE and MAE at $\leq 5\%$, and ≤1% denotes acceptable and perfect fit of the model respectively.

The critical values of MAPE are 0.25, 0.10, and 0.05. Therefore, *, **, and *** denotes threshold, low but acceptable, and highly acceptable accuracy at 0.25, 0.10, and 0.05 respectively.

The critical values of R^2 are 0.50, 0.75, and 0.90. Therefore, *, **, and *** denotes weak, moderate, and substantial prediction at 0.50, 0.75, and 0.90 respectively.

The critical value of the difference between EVS and R^2 is $\leq 2\%$. Therefore * denotes acceptable bias at $\leq 2\%$.

In purpose of validating the results, alongside of measuring the sustainability of the results, the Monte Carlo Cross Validation has repeated 40 times, the mean of achieved statistical test results has been taken and presented above. The mean of MCCV statistical test results indicates that;

VMS result is 0.947610, which is higher than 0.90 criteria. Therefore, it is possible to denote VMS value as highly significant, thus indicates a superior harmonic mean between homogeneity and completeness and indicates a perfectly accomplished clustering task for the future weight support vector regression. Additionally, highly significant VMS result, 0.947610 indicates no multicollinearity issue, and authorizes the model for further tests and regression analysis.

EVS result is 0.945922, which is higher than 0.90 criteria. Therefore, it is possible to signify EVS value as highly significant. The results indicate that the variance can be explained by the factors presented by the actual data. Moreover, EVS is another measure that is crucial for regression analysis, as it scores homoscedasticity. Highly significant EVS result, 0.945922 denotes that the variance can be explained by the factors presented by the actual data, and not heteroscedastic.

MSE result is 0.036425. As it has been emphasized on the Chapter 4, MSE closer to 0 indicates better fit of regression line, as it indicates the variance of residuals. However, as there isn't acceptable range set for MSE as it does not shares the same unit as the original values, MSE has been used to achieve RMSE values.

RMSE value is in the same unit with the original data, and indicates the standard deviation of errors emerged during prediction, thus signifies the accuracy of the model. RMSE result is 0.190830, which denotes that the standard deviation of error is only 0.190830 when compared to actual USD/TRY value. In order to evaluate RMSE result, the value transformed into interpretable value, SI. Transformed RSME, SI value is 0.036634, which is lower than 0.05 criteria. Therefore it is possible to denote RMSE value as significant.

MAE is relatively insensitive to outliers compared to RMSE, while RMSE magnifies the bigger errors and ignoring the smaller errors. For that reason, MAE may accommodate bias, while RMSE is related to variance. When MAE

and RMSE compared, it is possible to observe underfitting and overfitting issues that have been targeted by the application of MCCV method. Underfitting and overfitting is the situation of high bias, low variance, and low bias, high variance respectively. The difference between RMSE and MAE results, 0.190830 and 0.188578 respectively is 1.18%, which is lower than the 5% criteria. Therefore, the difference is in acceptable range, and denotes that model has fitted without any underfitting or overfitting issues.

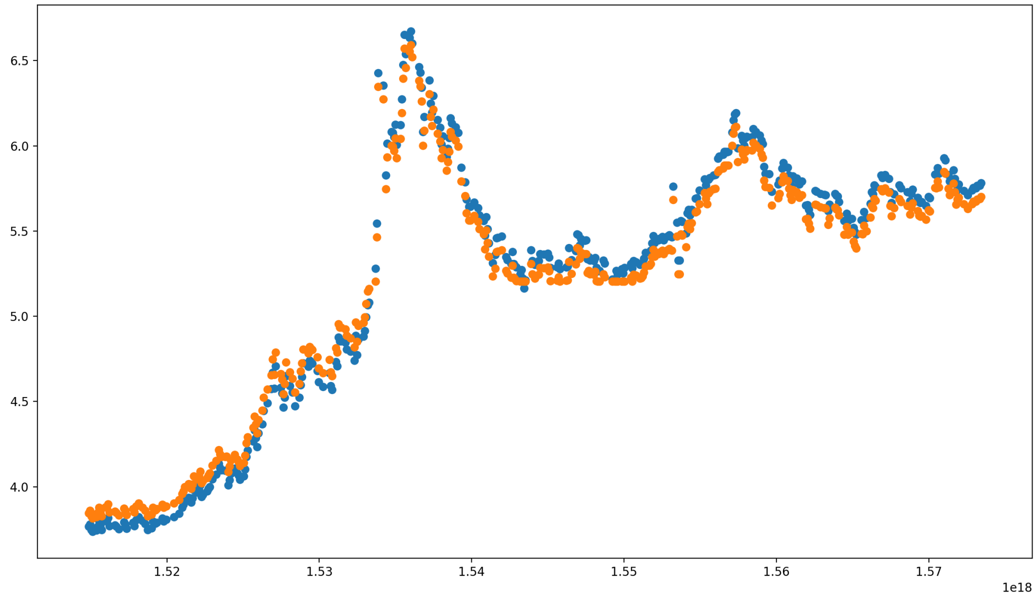
The accuracy of the prediction has also been measured by the MAPE results. MAPE results, 0.037097 that is equivalent to 3.7%, which is higher than 0.05 criteria, equivalent to 5%. Therefore, it is possible to say that the model has highly acceptable accuracy.

R^2 result is 0.939205, which is higher than the 0.90 criteria. Highly significant results have verified that the prediction is substantial with a superior accuracy, as the highly significant majority of the predicted data points are on the regression line.

Furthermore, the deviation between EVS and R^2 results, 0.945922 and 0.939205 respectively is only 0.72%, which is lower than 2% criteria. Thus the deviation between EVS and R^2 results signifies that the prediction is unbiased.

Achieved prediction results for USD/TRY prices between examined time intervals by using FWSVR is visually presented below as Figure 82.

Fig. 82. Actual versus predicted USD/TRY prices obtained by analyzing recently published news and statements regarding on recently published political news.



**Blue Points denote Actual USD/TRY Prices

**Orange Points denote Predicted USD/TRY Prices

Decisively, VMS, EVS, MSE, RMSE, MAE, MAPE and R^2 addressed the most common issues of regression analysis, namely; multicollinearity, heteroscedasticity, biasedness. Test results have verified the accuracy of the model by confirming the harmony of clustering, fit of the model, and unbiased prediction with minimal percentage of error, while clarifying that the regression analysis is purified from possible issues, then still achieved a successful prediction performance for daily exchange rates for a date interval with a highly fluctuating exchange rate prices. Furthermore, statistical test results had clarified the significance of recently published political news on exchange rate, and their verified their impact on the fluctuating exchange rates.

7.5 Conclusion

The impacts of political news that accommodate political tension in a varied degree of certainty on exchange rate fluctuations have been examined through this chapter. The chapter has run a complex algorithm to observe the instinctive attitude of investors on breaking news regarding to recent political environment, which triggers the act of decision-making, and causes the impulsive response of exchange rate. Predominantly, this chapter intended to measure the responsiveness of exchange rate to government officials' statements, speeches addressing the nation, speeches addressing government agencies and officials, and recently published political news beneath the wide spectrum of certainty and consensus considered through the relative perspective of investors. In order to evaluate the dependency of exchange rate fluctuations on any political environment and predict future exchange rates accordingly, machine learning techniques that has been used previously have necessarily been enhanced and reformed based on the database that has been conquered by the hidden layers of algorithm embedded and assessed through the previous chapter. Further investigation to evaluate investors' initial response on the political news have obligated the algorithm to be redesigned in order to meet the requirements of the modified techniques that have heavily relied on the subjective and qualitative data. The enhanced and reformed methods to meet the requirements of complex algorithms that have been used through this chapter are; natural language processing, weighted prospect theory, weighted fuzzy logic, and feature-weighted supports vector regression. The aforementioned methods are interconnected, and associated with the tasks of; calculating the probability of the sentiment outcome for every collected news article, assessing decision makers' subjective valuation of the news article, investigating the degree of certainty and consensus of the political environment to attain the correlation between political news and exchange rate fluctuations, and regression analysis for daily exchange rate prediction respectively.

Simulations observing recently published political news have verified the inconsistency of political consensus and certainty, thus the political tension between Turkey and U.S.. Simulation outcomes indicates that, certain disagreement and uncertain disagreement have a strong negative impact, uncertain agreement has mild positive/negative impact, and certain agreement has strong positive impact on exchange rates. Furthermore, dependency rate that has been attained as an outcome of the simulations has verified the strong correlation between political tension and exchange rate volatility. The regression analysis that has been guided by the attained correlation values for every possible scenario of political tension such as; judicial system, international politics, international relations, financial sanctions, sanctions on government officials, warfare, defense industry, political tension, human rights, NATO, elections, cabinet changes, and democratic uncertainty. Provided input has split into train and test portions to run MCCV. MCCV has repeated for numerous times to test the consistency of the output, and have been evaluated against seven key statistical metrics in order to clarify the performance of the regression analysis. Statistical test results indicated a valid and accurate prediction without the issues of multicollinearity, heteroscedasticity, biasness, and underfitting or overfitting. The findings of this chapter explicitly verified the correlation between government officials' statements, speeches addressing the nation, speeches addressing government agencies and officials, and recently published political news that accommodates political tension in a varied spectrum of certainty and exchange rate fluctuations through the significance of achieved dependency rates and prediction accuracy. The substantial role of cognitive approach and analysis has also been emphasized throughout the chapter. Chapter 3 have previously mentioned about Turkey's geopolitical and strategic importance for U.S., Russia, Europe, and Middle East, and how political tension on the region is influential on investors' decision making process. This chapter has scientifically verified the hypothesis of Chapter 3 on the causality between political tension on the region and the highly responsive exchange rate fluctuations. Besides, the novel collaboration of diverse disciplines; soft computing, politics, and psychology in purpose of enhancing effectiveness of economic analysis, and the usability of subjective qualitative information alongside the objective quantitative variables has been firmly confirmed by the dramatics accuracy of the composed model, thus, findings of the chapter clarified that model performed unambiguously marvelous on observing decision makers' initial responses and behaviors during political events, extracting events' weights relatively, and predict future exchange rates accordingly.

CHAPTER 8

Effective Forecasting of Exchange Rate in Daily Basis by Examining Macroeconomic and Political News

8.1 Introduction

The influence of breaking news and officials' announcements on macroeconomic indicators, political tension, and political uncertainty on exchange rate fluctuations have been introduced throughout the previous three chapters. The significance of breaking news accommodating the expectations and the analysis on macroeconomic indicators, and officials' reports and announcements on daily exchange rate fluctuations has been emphasized in Chapter 5. Following chapter, Chapter 6 intended to conceive the weight distribution of various political categories, which involve judicial system, international politics, international relations, financial sanctions, warfare, defense industry, elections, cabinet changes, and democracy by examining four main political events that took place during the selected date interval for this study. Then, Chapter 7 disseminate the political analysis into a broader spectrum by benefitting from findings of the previous chapter on how investors' response on altering grades of certainty and consensus on differing political categories, then weighting the breaking political news based on their hierarchical order of importance on investors' subjective perspective.

Although, understanding, uncovering, and weighting the instantaneous impact of breaking news regarding macroeconomic indicators and political tension separately enhances the prediction accuracy of further exchange rates, yet it is not sufficient enough to fully reflect the achievable prediction performance of the algorithm in real world scenarios, as both macroeconomic fluctuations and political tension may appear simultaneously at the same time. Moreover, political tension is a triggering factor for fluctuations of macroeconomic indicators, which was emphasized through the Chapter 3 that concentrated on the geopolitical importance and the influence of Turkey for United States, Europe, and Middle East, and point out the economical consequences of political tension on the region for Turkey. Due to aforementioned reasons, in order to comprehend the actual impact of breaking news on instantaneous exchange rate fluctuations, not only the political or macroeconomic news be examined separately, but both should be examined simultaneously to be able to simulate the timeline of the news streaming of the real world situations, and predict accordingly.

Previous empirical chapters, Chapter 5, 6, and 7 serve the purpose of weighting diversified macroeconomic and political breaking news and announcements regarding to their hierarchical order of importance on investors' subjective perspective. This chapter is going to examine macroeconomic and political breaking news simultaneously in an exact timeline that the news have been streamed in the media so as to simulate real world environment and the actual impact emerged. The algorithm is going to be benefitting from the diversified algorithm clusters that have been generated throughout previous empirical chapters. Complex integration of algorithms that concentrates on diverse focuses aims to enhance the prediction performance and accuracy by consolidating the hierarchical evaluation of macroeconomic and political news, by especially targeting the dates that macroeconomic and political impact overlaps.

The main purposes of this chapter is simulating the actual frequency and contents of news feed, impersonate the subjective viewpoint and instinct behaviors of the investors to the breaking news via robust AI, and achieve a highly accurate exchange rate prediction. Previous empirical chapters played a significant role on understanding the subjective perspective of the investors on the macroeconomic news concerning the indicators; interest rate, inflation rate, unemployment rate, balance of trade, and credit ratings, alongside to the political news considering the categories of judicial system, international politics, international relations, financial sanctions, warfare, defense industry, elections, cabinet changes, and democracy under the influence of political tension, disagreement, and uncertainty. Although, the emergence of both macroeconomic and political news back to back in the same day is a highly possible situation in real world scenarios. In this case, the impact of the particular news on the valuation mechanism of the investors may differ compared to the separate evaluations that had examined throughout the previous empirical chapters. As this chapter is aiming to simulate the investors' evaluation process of real world scenarios to achieve highly accurate exchange rate prediction, the hierarchical evaluation is crucial to set the subjective order of importance, and assigning altering weights required to adapt varying complicated situations and environments.

To achieve this purpose, generated hybrid machine-learning algorithms throughout the previous empirical chapters have been combined to form an enhanced form of AI that is capable enough to adapt varying complicated situations and environments simultaneously by making the right decisions on weight distribution under complex situations, and run the prediction accordingly. To do so, adapted variations of machine learning algorithms; natural language processing (NLP), weighted prospect theory (WPT), weighted fuzzy logic (WFL), and feature weighted supports vector machines (FWSVM) have been adopted.

The chapter is structured as follows: Section 2 briefly describes data collection and fundamental variables; Section 3 describes application of methodologies, Section 4 presents empirical results, and prediction of USD/TRY exchange rate, and Section 5 concludes the chapter.

8.2 Data and Fundamental Variables

In the preeminent curiosities to refine understanding on the initial impact of macroeconomic and political breaking news that accommodates officials' announcements, tension, instability, uncertainty, and consensus triggered by aggregate factors, USD/TRY exchange rate is going to be investigated in daily manners as objective and quantitative data. In order to perceive the degree of responsiveness of exchange rate to recently published government officials' political and international relations related official verbal or written statements; daily price, daily change in percentage, opening price, highest price, lowest price, and difference between highest and lowest prices of USD/TRY be investigated for each day. To cope with any overlapping news on opposite polarities on the same day, opening price, close price, highest price, lowest price, and the difference between highest and lowest prices plays a crucial role on overcoming any possible accuracy error on the ML algorithm, as the appearance of news on opposite polarities on the same day may neutralize the significant effect of vital news, and cause a dramatic error on further calculations of the algorithm. The philosophy behind the candlestick chart has been adopted in order to overcome the possibility of mentioned issue. In this case, even if neutralization happened in any day during the selected date interval, algorithm will be able to detect it by observing the volatility of the exchange rate price throughout the day, and weight related news accordingly to match the volatility and the price by considering previous/further news on both subjects. Abovementioned exchange rate data have been collected over the period between 29 December 2017 and 01 November 2019, which is equivalent to 4032 variables.

Alongside to historical USD/TRY exchange rate, historical inflation rate, unemployment rate, interest rate, balance of trade, and credit rating are going to be observed to perceive the initial impact of expected and existing macroeconomic fluctuations, and guide the algorithm on investigating and valuating the related news articles accordingly.

For subjective qualitative analysis, HTTP REST API has extracted daily news articles regarding officials' macroeconomic and political statements, announcements, alongside to published news articles that reflects tension, instability, uncertainty, consensus, and expectations and opinions of experts that may influence the decision makers' decision from the websites of leading news publishers', such as; BBC, Reuters, Bloomberg, The Guardian, and The New York Times. Almost 1000 news articles have been collected analyzed, and subjective and qualitative textual content transformed into numeric sentiment inputs.

8.3 Methodology

An augmented methodology is going to be adopted by this chapter to meet the requirements of the complex integration of algorithms that involve the contribution of diversified algorithm clusters generated throughout previous empirical chapters on diverse focuses to conceive the hierarchical evaluation and weight distribution, which is essential to accomplish enhanced prediction performance and accuracy under the influence of real world environments involving macroeconomic fluctuations and political tension simultaneously.

The methodology of this chapter is constructed as three phases. The initial phase is the "Sentiment Analysis", which adopted Natural Language Processing for textual analysis, and reinforced the analysis with Weighted Prospect Theory. Aftermath of Sentiment Analysis is going to be studied by the following phase "Weighted Fuzzy Logic" in order to attain a weighted correlation value. Final phase, "Feature-Weighted Support Vector Regression" is going to run regression, apply statistical tests to perceive validation and accuracy of the model, and predict future exchange rates.

Mentioned phases that constituents the methodology, which is also going to be accommodated through the workflow of this chapter has clarified thoroughly in Chapter 4. Sections and subsections of Chapter 4 that involves relevant phases of the methodology are going to be indicated below.

Initial phase, "Sentiment Analysis" (Chapter 4.2), practices the components of Natural Language Processing. Lexicons (Chapter 4.2.1) for this chapter have been provided by the output obtained on previous chapters, Chapter 5.4 and Chapter 6.5. Learning Automata (Chapter 4.2.2) is adopted for textual analysis, which reinforced by Weighted Prospect Theory (Chapter 4.7.1), which targets correct application of weights. Outcome of Sentiment Analysis is going to be providing input for second phase, " Weighted Fuzzy Logic" that adopts fuzzy logic (Chapter 4.7.2), and fuzzy weighted average within a generalized mean operator (Chapter 4.7.2.1) algorithms of machine learning, and going to achieve weighted correlation value between recently published macroeconomic and political news and exchange rate. Final phase, "Weighted Support Vector Regression" is composed by enhancing the SVR to WSVR with a combination of methodologies (Chapter 4.4.1, and 4.7.3) by using the presented weighted correlation values is going to run support vector regression. In order to split data into training and test sets, Monte Carlo Cross Validation (Chapter 4.5.1.1) is going to be adopted. Hyperparameters for WSVR is going to be acquired by Grid Search with Cross Validation (Chapter 4.5.2.1). Grid Search with Cross Validation method plays a crucial role on fine-tuning the hyperparameters for the kernel function cautiously and suitably to provide most accurate hyperparameter selection, while eliminating the heteroscedasticity and multicollinearity in regression analysis. The accuracy of the obtained parameters is going to be validated by the metrics, namely; Accuracy, Precision, Recall, and F1 Score (Chapter 4.5.2.2). The linearity of the provided dataset is going to be evaluated by Principal Component Analysis (PCA), Principal Component Regression (PCR), and Partial Least Square (PLS) methods, and kernel function for the model is going to be assigned according to the linearity of the dataset. The performances of Polynomial, and Gaussian Radial Basis kernel functions are going to be compared, and the one that provides the most accurate hyperparameter selection that eliminates heteroscedasticity in regression is going to be adopted for the model. The performance of the fabricated model is going to be assessed by the statistical tests, namely; V Measure Score (VMS), Explained Variance Score (EVS), Mean Squared Error (MSE), Root Mean Squared Error (RMSE), Mean Absolute Error (MAE), Mean Absolute Percentage Error (MAPE), and Coefficient of Determination (R^2) (Chapter 4.6), and future exchange rates are going to be predicted. The workflow mechanism of the introduced methodology is going to be elaborated through the upcoming subsection.

8.3.1 Workflow Mechanism of Methodology

Throughout the previous empirical chapters, Chapter 5, 6, and 7 a substantial qualitative database has been collected. Besides, fuzzy quantitative database has been composed by processing collected qualitative database. Obtained qualitative and fuzzy quantitative database aims to contribute to the main purpose of this study by examining the usability of qualitative datasets on particular disciplines, and revealing their correlation and multidirectional impact on quantitative time series. Obtained results indicated the significance of examining qualitative dataset on understanding their impact on quantitative time series individually on the examined discipline, while going to be contributing to this thesis' final chapter by providing a augmented hierarchical order of importance and precise weight distribution to any potential qualitative input, which is essential in order to compose a accurate and superior machine learning algorithm.

The foundations of the first phase, Sentiment Analysis had been shaped through the previous empirical chapters, Chapter 5, 6, and 7. Officials' statements, announcements, speeches addressing the nation, speeches addressing government agencies and officials, and recently published news regarding macroeconomic indicators and politics had been extracted from the news sources, and analyzed through the unique lexicons that had been generated in associated chapter in order to extract the sentiments and provide polarity vectors as an output for every article feed into the algorithm throughout previously mentioned empirical chapters, which transformed into normalized polarity vectors that provides the probability of the sentiment outcome.

Empirical chapters, Chapter 5 and 6 derived the subjective feelings of the decision makers by examining the fluctuations on the historical data, while the capabilities of the algorithm is augmented in Chapter 7 to derive the subjective standpoint of the decision makers without needing to observe the fluctuations on the historical data. The augmented capabilities of the algorithm is modified and enhanced accordingly for this chapter in order to be able to examine the macroeconomic and political news simultaneously by benefitting from the provided weights and the hierarchical importance throughout the previous empirical chapters. To comprehend the decision making reference point of each attribute of decision makers', attained probability values are going to be fuzzified and weighted accordingly, as each breaking news and/or official statement's hierarchical order of importance that determines the decision makers' attitude is fuzzy in itself.

This will provide an insight of weighting and valuation approach of decision makers to the algorithm. Obtained weighted fuzzy values are going to be used as an input for prospect theory, and weighted prospect values are going to be generated. Generated weighted prospect values will be the output of sentiment analysis, which is indicating the subjective assessment of decision makers.

Achieved outcome of sentiment analysis is appointed as an input for the second phase, weighted fuzzy logic. As the macroeconomic and political articles have been evaluated collectively, clustering task is essential for the fuzzy antecedents. Clustering intends to cluster the fuzzy antecedents, which has been obtaining on earlier phase, sentiment analysis depending on the subjective weighting, besides the valuation of decision makers. Performing clustering task will distinguish the antecedents into the clusters of similar weights and significance, which are going to be arranged hierarchically. A hierarchical structure is required to identify the evaluation in a multi-criteria decision-making scenario, as the significance of criteria entirely depends on the subjective weighting and valuation of decision maker. Each cluster that forms a step in hierarchy will be evaluated separately under the methodology of fuzzy logic.

As it has been suggested through empirical chapters 5 and 7, the announcements and articles on macroeconomic indicators, namely; interest rate, inflation rate, unemployment rate, credit ratings, and balance of trade are going to be considered, while the content of the political news are going to be considered under two main interpretations that encountered relatively as the main viewpoints of the published articles, which are the degree of certainty, and the degree of consensus. The psychological impacts of both mentioned macroeconomic indicators and political interpretations' on the subjective approaches of the decision makers has been proven by the existing literature. Thus, fuzzy antecedents have been distributed accordingly as 'Interest Rate', 'Inflation Rate', 'Unemployment Rate', 'Balance of Trade', and Credit Ratings' for macroeconomic indicators, while distributed into two main classes for political news; 'Certainty' and 'Consensus'.

The membership function for Macroeconomic Indicators are considered under three main universal and subjective interpretations of the decision makers, which are; Increase, Steady, and Decrease. For the fuzzy antecedents of political articles, Certainty and Consensus the membership function divided into three main sets. The membership function for Certainty defined as Certain, Neutral, and Uncertain, while defined as Agreement, Neutral, and Disagreement for Consensus. The transition between the memberships of the determined sets for each fuzzy antecedent, aside from the degrees of truth for each defined set that have been mentioned above are going to be defined by the upper and lower bounds, width, concave, convex, and the steepness of the associated membership, which will be presented on the future graphs.

After assessing the membership function for the fuzzy antecedents in each cluster, fuzzy weighted average within a generalized mean operator is going to approximate membership values by hierarchical evaluation method for both categories of fuzzy antecedents, macroeconomic and political. By weighting the parameters within a hierarchy, altered valuation of the decision maker is going to be obtained by weighted generalized mean operator, which indicates the decision makers' behavior of evaluation. The fuzzy relations are going to be obtained by the presence or absence of association, interaction, or interconnectedness between the elements of two or more sets. Interpretation of the resulting fuzzy relations may be obtained by the fuzzy rules designated by the author, which are going to be feed into fuzzy engine as well as membership functions. Over 400 fuzzy rules have been designated from both macroeconomic and political fuzzy categories of the algorithm to feed into fuzzy engine. Fuzzy engine is going to generate Fuzzy Consequent, which is essentially a dependency of USD/TRY behavior on recently published news, and officials' statements. In other words, fuzzy consequent

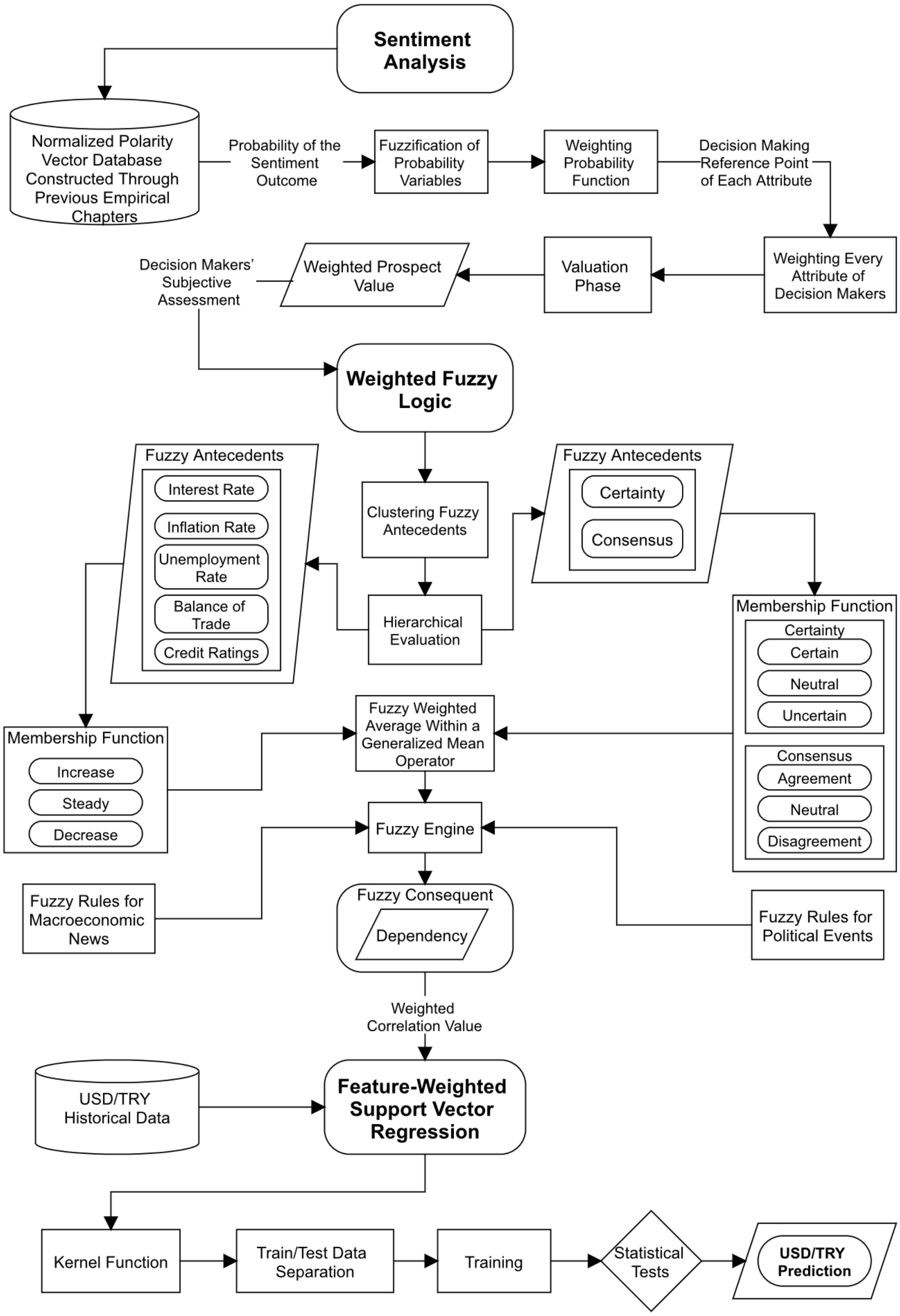
signifies the weighted correlation between USD/TRY and news and officials' statements.
Correlation value that has been achieved, as fuzzy consequents will be used as input in addition to historical USD/TRY data to run feature weighted support vector regression. The linearity of the provided dataset is going to be examined by Principal Component Analysis (PCA), Principal Component Regression (PCR), and Partial Least Square (PLS) methods.

PCA/PCR is a mono-dimensional, observed variable approach that aims to find a linear correlation by reducing dimensionality.
PLS is a multidimensional, latent variable approach that suits best for when there is a multicollinearity. Both PCA/PCR and PLS methods can be used to tests outliers and sensitivity analysis to provide diagnostic tools for the model.

Nonlinear USD/TRY data have been projected into high-dimensional space by using the kernel functions. Constructed hyperplane delivers adjustability to congregate optimal set of weights (Oliveira et al., 2017). The performance of Polynomial and Gaussian Radial Basis Function kernel functions are going to be compared by the grid search with cross validation results to be selected as the kernel function for the Future Weighted Support Vector Regression (FWSVR) analysis. As Gaussian Radial Basis Function scored better on grid search with cross validation for hyperparameter selection, it has been selected to utilize as kernel function for the FWSVR analysis.
Gaussian Radial Basis Function tuned by hyperparameters that have been achieved by grid search with cross validation method, which also eliminated the possibility of heteroscedasticity and multicollinearity. Accuracy of the parameters attained by the grid search with cross validation is evaluated by accuracy, precision, recall, and f1-score metrics. The data will randomly be divided into training and test groups for multiple times in order to consider the performance of the regression model. To do so, Monte-Carlo cross validation (MCCV) method has been implemented to split presented data into training and test groups for prediction. Accuracy of the prediction has been investigated by the statistical tests. Prediction reliability of the model is measured against seven broadly adopted statistical tests for machine learning

algorithms, namely; V Measure Score (VMS), Explained Variance Score (EVS), Mean Squared Error (MSE), Root Mean Squared Error (RMSE), Mean Absolute Error (MAE), Mean Absolute Percentage Error (MAPE), and Coefficient of Determination (R^2). As VMS signifies the clustering precision of the model, it is a pivotal measure to check primarily. Grander performance on homogeneity and completeness on clustering task is an essential in order to accomplish an accurate prediction deprived of multicollinearity issue. For mentioned reasons, achieving significant VMS results is necessary for successful regression analysis. Attained VMS value specifies perception on the prejudices of measured deviation of errors. MSE, RMSE, MAE, and MAPE are calculated in order to measure the deviation between predicted and actual values. RMSE and MAE are considered simultaneously to identify potential deviation in the errors. While equivalent RMSE and MAE values specifies correspondence of magnitude on all errors, the alteration between the results specifies the variance associated with the individual errors. Larger the dissimilarity indicates higher variance, and occurrence dissemination of the specific errors. Underfitting and overfitting issues of the model can also be observed by examining the difference between RMSE and MAE results. Furthermore, MAPE is correspondingly reflected and intended to normalize the absolute error and obtain altered perspective on association with real values. While MSE, RMSE, MAE, and MAPE measure the deviation between predicted and actual value, R^2 and EVS targets inconsistency between predicted and actual value. Conversely, while R^2 uses raw sum of squares, EVS practices the biased variance. Both variables can only be parallel if the mean of the residuals equals to zero, and prediction is unbiased. Additionally, as EVS evaluates the explainability of the variance by the factors presented in the actual data, it is also possible to check the homoscedasticity by EVS results. Conclusively, if the statistical tests' results are convenient, future exchange rates are going to be predicted in daily basis. Abovementioned empirical workflow is presented as a flowchart below as Figure 83.

Fig. 83 Flowchart of the proposed methodology of Chapter 8



8.4 Empirical Findings of the Hybrid Machine Learning Algorithms

Subjective perception of decision makers' under diversified macroeconomic and political announcement, statement, expectations, events, and situations have targeted to be evaluated through the sentiment analysis. Lexicons addressing aforementioned macroeconomic and political environments have been constructed uniquely through the empirical chapters Chapter 4, 5, and 6 specifically to perform under the complex structure of diversified causality for macroeconomic fluctuations and political tension that involves Turkey. Attained normalized polarity vectors through the sentiment extraction have been fuzzified and weighted correspondingly, targeting to attain the reference points for each attribute of decision makers. Attained weighted fuzzy values have feed into prospect theory, to generate weighted prospect values. Generated weighted prospect values provided an enhanced overlook on the subjective assessment of decision makers, and also form the output of sentiment analysis, which have been used as fuzzy antecedents for the weighted fuzzy logic phase.

Decision makers' subjective assessments have been scored for every published news article through the sentiment analysis, which provided the output of sentiment scores that are fuzzified. Fuzzified sentiment scores are going to be clustered, hierarchically evaluated, and used as fuzzy antecedents on the second phase, weighted fuzzy logic.

Clustering aims to cluster the fuzzy antecedents based on the weights given through the previous phase, which evaluated the subjective weighting and valuation of decision makers. This process is going to be followed by arranging the clusters hierarchically in order to identify the evaluation in a multi-criteria decision-making scenario. When the hierarchical order of clusters have been set, the sentiment scores that has been achieved on previous phase is going to be feed into the algorithm. Sentiment scores have been assigned between the range of 1 and 5, where 1 designates highest likelihood for definitive severe depreciation of USD/TRY, gradual increase in the range between 1 and 3 signifies decline in depreciation of USD/TRY, while 3 denotes the steady state of the exchange rate. Scores from 3 to 5 on the scale symbolizes a gradual appreciation in USD/TRY, while 5 denotes significantly absolute sharp appreciation in USD/TRY.

Membership functions for given fuzzy antecedents of macroeconomic fundamentals and political tension do not match due to their diverse nature. For that reason, macroeconomic and political fuzzy antecedents should be evaluated separately based on the attained weights through the empirical chapters, Chapter 5, 6, and 7.

The outcome of sentiment analysis for macroeconomic indicators have been clustered and arranged hierarchically under the title of related macroeconomic variable to form fuzzy antecedents for macroeconomic indicators. Decision makers' identify the grade of change in macroeconomic indicators in simple, yet subjective manners. Fluctuating rates of macroeconomic indicators expressed as 'increase', 'steady', and 'decrease', while the meaning of each aforementioned expression subjectively differs from person to person. Due to the particular reason for circumstance, membership functions for macroeconomic antecedents have been defined as 'increase', 'steady', and 'decrease'. Trapezoidal structure for membership function has been adopted to adapt the acute effects of change in macroeconomic variables.

Fuzzy engine determined that inflation rate, interest rate, and unemployment rate are positively correlated with USD/TRY, while balance of trade, and credit ratings are inversely correlated with USD/TRY. For that reason, values that denote the degree of belonging alter. For inflation rate, interest rate, and unemployment rate; values from 1 to 3 denotes the degree of belonging to fuzzy set 'increase', value 3 defines the belonging to fuzzy set 'steady', and values from 3 to 5 denotes the degree of belonging to fuzzy set 'decrease'. Meanwhile, for balance of trade, and credit ratings; values from 1 to 3 denotes the degree of belonging to fuzzy set 'decrease', value 3 defines the belonging to fuzzy set 'steady', and values from 3 to 5 denotes the degree of belonging to fuzzy set 'increase'.

On the other hand, fuzzy sets for political news have defined with membership functions for the antecedents 'consensus' and 'certainty', which indicates the subjective level of political consensus between sides, and subjective level of political certainty respectively. Political consensus and certainty have acute effects on international relations and economics. For that reason, trapezoidal membership function has selected to be used for fuzzy antecedents, while fuzzy consequent is defined with sigmoidal function.

Membership functions defined for each fuzzy antecedent. For the antecedent 'consensus', assigned membership functions are: 'agreement', 'neutral', and 'disagreement', while 'certain', 'neutral', and 'uncertain' are assigned as membership functions for antecedent 'certainty'. Sets for membership functions assigned by the fuzzy engine, in order to determine the degree of belonging to a fuzzy set according to the consequence of the sentiment analysis. 'Certainty' and 'consensus' values between 1 and 3 denotes the degree of belonging to fuzzy subsets 'certain' and 'agreement', 3 defined as 'neutral', and values between 3 and 5 denotes the degree of belonging to fuzzy subsets 'uncertain' and 'disagreement' respectively.

For both macroeconomic and political fuzzy antecedents, degree of membership to a fuzzy set denoted between 0 and 1, While 0 and 1 represents 0% and 100% membership degree respectively, anything in between considered as a partial membership to belonging set.

Once weighing the membership function for the fuzzy antecedents in each cluster, fuzzy weighted average within a generalized mean operator is going to approximate membership values by hierarchical evaluation method. The fuzzy relations are going to be obtained by the presence or absence of association, interaction, or interconnectedness between the elements of two or more sets.

The impacts of recently published news regarding macroeconomic indicators and political tension on exchange rate are going to be indicated by the correlation values. While macroeconomic news regarding to selected macroeconomic indicators deliver a deeper understanding on the direction of country's economy, attained fuzzy rates of certainty and consensus are going to provide deeper understanding on the political environment and situation by examining the political news in order to achieve dependency rate as a fuzzy consequent that is a noteworthy indicator to understand how decision makers' react to fluctuating macroeconomic indicators and political tension. The ability of describing qualitative information and subjective standpoint and valuation of decision makers on macroeconomic fluctuations and political tension in a quantitative form plays a crucial role on running precise predictions.

In order to examine the impact of macroeconomic and political news on exchange rate simultaneously, the hierarchy of importance, and the weights for each variable and scenario should be specified accurately. The hierarchical evaluation and the weight distribution for macroeconomic and political news have been achieved according to the unique dynamics that has been obtained through previous empirical chapters. The weights for the impact on exchange rate fluctuations, and the membership function for each macroeconomic indicator has attained through empirical chapters 5, while the fuzzy weighted average values for political tension have been generalized with a mean operator through the empirical chapter 7.

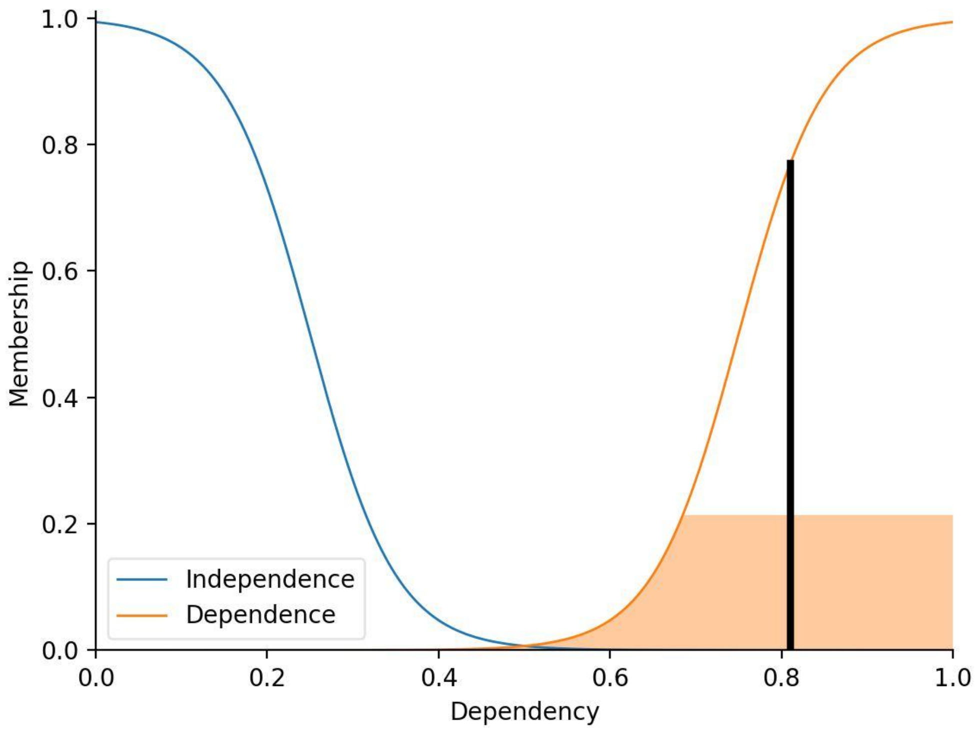
Simulation results for the dependency of USD/TRY on inflation rate, interest rate, balance of trade, unemployment rate, and credit ratings have presented in Chapter 5 and illustrated as Figures 27, 28, 29, 30, and 31 respectively, while the dependency of USD/TRY on generalized macroeconomic news has presented as Fig 32 in the appendix section. Obtained correlation values for each macroeconomic indicator had specified that; published news regarding the increase in inflation rate of Turkey, and downgrading credit rating of Turkey have significant impact on the appreciation of USD/TRY. Besides, published news on increase in unemployment rate of Turkey has insignificant impact on the appreciation of Turkey. On the other hand, the impact of published news regarding the decrease on interest rate in Turkey on depreciation of USD/TRY is inevitably significant. Meanwhile, published news on increase in balance of trade of Turkey has insignificant impact on the depreciation of USD/TRY.

Chapter 7 had emphasized the degree of political certainty and consensus for recently published political news, and presented the results visually as Figure 72 and 73, while the dependency rate of USD/TRY exchange rate fluctuations on the recently published political news has been introduced throughout the empirical chapter, and illustrated as Figure 74 in the appendix section. Fuzzy consequent for the degree of political certainty and consensus of the news articles designates a certain disagreement between U.S. and Turkey, while the averaged correlation values indicate a negative impact of political tension between Turkey and United States on the price of Turkish Lira, which causes an appreciation on the price of USD/TRY exchange rate.

As this chapter has focused on to scrutinize the instantaneous impact of simultaneous appearance of macroeconomic and political news on USD/TRY, both macroeconomic and political variables have been feed into fuzzy weighted average within a generalized mean operator simultaneously, and trained by the fuzzy engine that is using over 400 rules designated from both macroeconomic and political fuzzy categories of the algorithm. Correlation value for merged macroeconomic and political news are going to be generated by combining the attained results derived within the fuzzy engine based on the preset fuzzy rules. The fuzzy consequent, dependency value that represents the correlation between recently published news and fluctuating USD/TRY prices have been generated by the fuzzy engine, where value 0 defines that case related news and fluctuations in exchange rate is extremely independent, and value 1 defines that case related news and fluctuations in exchange rate is extremely dependent.

Dependency rate of USD/TRY exchange rate on recently published macroeconomic and political news has been simulated. Simulation results indicate a significant dependency value of 0.817, and verify the hypotheses that exchange rate fluctuations are highly dependent on the published news and statements regarding to the macroeconomic fluctuations and political tension. Dependency rate of USD/TRY exchange rate fluctuations on the recently published news is visually presented below, as Figure 84.

Fig. 84 Dependency for Organic News Feed Simulation**.**



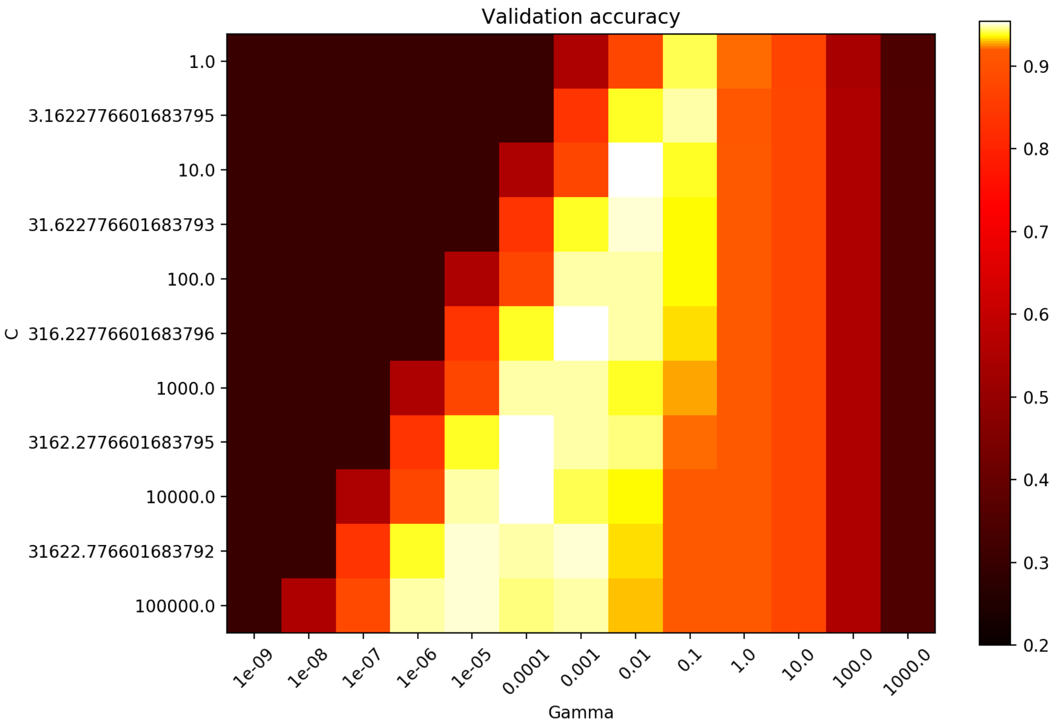
On the final phase, previously achieved fuzzy consequent, weighted correlation value is going to be used as an input for feature-weighted support vector regression, as well as historical USD/TRY data.

The linearity of the provided dataset is evaluated by PCA/PCR and PLS methods. Figure 85, which is presented in the appendix section illustrates PCA/PCR and PLS results. First principal component in PCA/PCR and PLS indicates the axis in the K-dimensional variable space that accommodates the largest variance of the samples in a direction, while second principal component is orthogonal axis to the first principal component in the K-dimensional variable space, and identifies the second largest source of samples in a direction that can be denoted as the residual variance of the samples that improves the approximation. Besides, the direction that provides the lowest variance is aimed to be captured by the PLS. The nonlinearity of the provided dataset has confirmed by the tests applied. For that reason, kernel function for the SVR analysis is going to be assigned accordingly in order to prevent heteroscedasticity and multicollinearity. The application of Gaussian Radial Basis Function on PCA and the projection difference is exemplified as Figure 35 in the appendix section.

Kernel functions have provided an ability of projecting nonlinear USD/TRY data into high-dimensional space. Projection of USD/TRY in high-dimensional space is presented in Figure 86 below, and 87 and 88 in the appendix section. Figure 89 represents the USD/TRY historical data as traditional 2 dimensional vectors in 2-D input space. By using kernel function, Constructed hyperplane provides adjustability to converge optimal set of weights (Oliveira et al., 2017). It has been observed that, Polynomial and Gaussian Radial Basis Function kernel functions are the ones that may suit best for provided data points. For that reason, aforementioned kernel functions have been compared, according to grid scores on development set attained by grid search with cross validation for hyperparameter selection. As Gaussian Radial Basis Function scored better on grid scores, it has been selected to utilize as kernel function for the FWSVR analysis. Grid scores on development sets for both kernel functions have presented as Table 26 exhaustively in the appendix section.

Selected Gaussian Radial Basis function has tuned by $C, \varepsilon, \text{and}\, \gamma$ hyperparameters, and grid search has been adopted to set the best-suited set of hyperparameters. Grid search ranges of $C, \varepsilon, \text{and}\, \gamma$ hyperparameters assigned as suggested by Hsu, Chang, and Lin (2003). *C* ranges from 1 to 10^5, and γ from 2^{-8} to 2000. In depth grid search with cross validation for hyperparameter $C, \text{and}\, \gamma$ value selection is accessible as Table 26 in the appendix section. Grid search cross validation accuracy heat map is also presented as Figure 90 below, and layers view for 3D Visualization of Grid Search Cross Validation Accuracy Heat Map for news has ben presented as Figure 91 in the appendix section.

Fig. 90. Grid Search Cross Validation Accuracy Heat Map for Organic News Feed Simulation.



Hyperparameter ε is assigned by considering the highest change in the historical data, and determined as high as possible in order to eliminate any possible overfitting issue. Accuracy of the parameter selection using grid search with cross-validation is evaluated by consideration of three metrics, namely; precision, recall, and f1-score. Results have indicated the accuracy of 0.95 for parameter selection, and detailed in Table 27.

Tab. 27 Accuracy of the Parameter Estimation Using Grid Search with Cross-Validation for Organic News Feed Simulation.

The model is trained on the full development set.
The scores are computed on the full evaluation set.

	precision	recall	f1-score	support
0	1	1	1	21
1	0.91	0.97	0.94	30
2	0.95	0.88	0.91	24
Accuracy			0.95	75
Macro Avg	0.95	0.95	0.95	75
Weighted Avg	0.95	0.95	0.95	75

The critical values of Accuracy, Precision, Recall, and F1 Score are 0.20, 0.40, 0.60, and 0.80 denotes slight, fair, moderate, substantial, and almost perfect agreement respectively.

In order to enhance the performance of the regression model, splitting presented data into training and test groups is essential to appraise the performance of the regression model. To do so, Monte-Carlo cross-validation (MCCV) method adopted. 80% of provided data used as train data, while 20% reserved as test data. MCCV randomly splits reserved portion of data into sub-samples and assign them as test sets. The MCCV procedure repeated for 65 times in order to observe the consistency of the results.

Prediction consistency of the model is evaluated against seven key statistical metrics, which are widely be accepted and using to validate machine learning prediction performance, namely; V Measure Score (VMS), Explained Variance Score (EVS), Mean Squared Error (MSE), Root Mean Squared Error (RMSE), Mean Absolute Error (MAE), Mean Absolute Percentage Error (MAPE), and Coefficient of Determination (R^2). Statistical test results attained every time MCCV procedure repeated. Achieved results for statistical test are presented exhaustively in Table 28 in the appendix section. Obtained statistical test results averaged. Averaged results for MCCV statistical test results are presented in Table 29.

Table 29. Averaged Monte Carlo Cross Validation Statistical Test Results for Organic News Feed Simulation.

.

VMS	0.959443***
EVS	0.963683***
MSE	0.024505
RMSE	0.156535**
MAE	0.155023
MAPE	0.030480***
$\boldsymbol{R^2}$	0.958967***

† Exhaustive Monte-Carlo cross validation statistical test results are accessible in the appendix section, as Table 28.

The critical values of VMS are 0.70, 0.80, and 0.90. Therefore, *, **, and *** denotes threshold, significant, and highly significant at 0.70, 0.80, and 0.90 respectively.

The critical values of EVS are 0.60, 0.80, and 0.90. Therefore, *, **, and *** denotes threshold, significant, and highly significant at 0.60, 0.80, and 0.90 respectively.

The critical values of RMSE are set by SI at $\leq 10\%$, and $\leq 5\%$, which indicate RMSE value of 0.520908, and 0.260454 respectively. Therefore, *, and ** indicates the threshold and significance at respective critical points.

The critical values of the difference between RMSE and MAE at $\leq 5\%$, and ≤1% denotes acceptable and perfect fit of the model respectively.

The critical values of MAPE are 0.25, 0.10, and 0.05. Therefore, *, **, and *** denotes threshold, low but acceptable, and highly acceptable accuracy at 0.25, 0.10, and 0.05 respectively.

The critical values of R^2 are 0.50, 0.75, and 0.90. Therefore, *, **, and *** denotes weak, moderate, and substantial prediction at 0.50, 0.75, and 0.90 respectively.

The critical value of the difference between EVS and R^2 is $\leq 2\%$. Therefore * denotes acceptable bias at $\leq 2\%$.

The Monte Carlo Cross Validation has repeated for 65 times in order to validate the results alongside of measuring sustainability, and the mean of attained statistical test results has been presented above. The mean of MCCV statistical test results indicates that;

VMS result is 0.959443, which is higher than 0.90 criteria. Therefore, it is possible to denote VMS value as highly significant, thus indicates a superior harmonic mean between homogeneity and completeness and indicates a perfectly accomplished clustering task for the future weight support vector regression. Moreover, highly significant VMS result, 0.959443 indicates no multicollinearity issue, and authorizes the model for further tests and regression analysis.

EVS result is 0.963683, which is higher than 0.90 criteria. Therefore, it is possible to signify EVS value as highly significant. The results indicate that the variance can be explained by the factors presented by the actual data. Moreover, EVS is another measure that is crucial for regression analysis, as it scores homoscedasticity. Highly significant EVS result, 0.963683 denotes that the variance can be explained by the factors presented by the actual data, and not heteroscedastic.

MSE result is 0.024505. As it has been emphasized on the Chapter 4, MSE closer to 0 indicates better fit of regression line, as it indicates the variance of residuals. However, as there isn't acceptable range set for MSE as it does not shares the same unit as the original values, MSE has been used to achieve RMSE values.

RMSE value is in the same unit with the original data, and indicates the standard deviation of errors emerged during prediction, thus signifies the accuracy of the model. RMSE result is 0.156535, which denotes that the standard deviation of error is only 0.156535 when compared to actual USD/TRY value. In order to evaluate RMSE result, the value transformed into interpretable value, SI. Transformed RSME, SI value is 0.030050, which is lower than 0.05 criteria. Therefore it is possible to denote RMSE value as significant.

MAE is relatively insensitive to outliers compared to RMSE, while RMSE magnifies the bigger errors and ignoring the smaller errors. For that reason, MAE may accommodate bias, while RMSE is related to variance. When MAE

and RMSE compared, it is possible to observe underfitting and overfitting issues that have been targeted by the application of MCCV method. Underfitting and overfitting is the situation of high bias, low variance, and low bias, high variance respectively. The difference between RMSE and MAE results, 0.156535 and 0.155023 respectively is 0.97%, which is lower than the 1% criteria. Therefore, the difference denotes that model has fitted perfectly without any underfitting or overfitting issues.

The accuracy of the prediction has also been measured by the MAPE results. MAPE results, 0.030480 that is equivalent to 3.048%, which is higher than 0.05 criteria, equivalent to 5%. Therefore, it is possible to say that the model has highly acceptable accuracy.

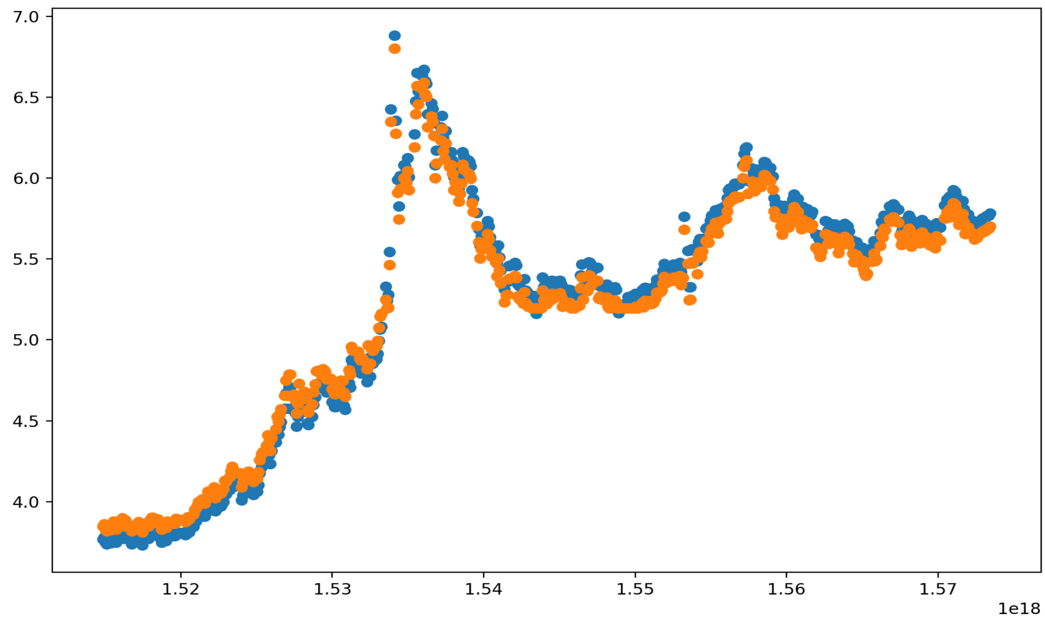
R^2 result is 0.958967, which is higher than the 0.90 criteria. Highly significant results have verified that the prediction is substantial with a superior accuracy, as the highly significant majority of the predicted data points are on the regression line.

Furthermore, the deviation between EVS and R^2 results, 0.963683 and 0.958967 respectively is only 0.49%, which is lower than 2% criteria. Thus the deviation between EVS and R^2 results signifies that the prediction is unbiased.

Conclusively, statistical tests addressed the most common issues of regression analysis, namely; multicollinearity, heteroscedasticity, biasedness. Test results clarified that the regression analysis is purified from possible issues, then still achieved a superior prediction performance for daily exchange rates for a date interval with a highly fluctuating exchange rate prices.

Achieved prediction results for USD/TRY prices between examined time intervals by using FWSVR is visually presented below as Figure 92.

Fig. 92. Actual versus predicted USD/TRY prices obtained by analyzing recently published news and statements regarding Organic News Feed Simulation.



**Blue Points denote Actual USD/TRY Prices

**Orange Points denote Predicted USD/TRY Prices

8.5 Conclusion

This chapter focused on simulating the responsiveness of USD/TRY to organic news streaming by impersonating the subjective perception and behaviors of the investors. As the organic news feed scenario encapsulates both publications regarding macroeconomic and political news simultaneously, comprehending the subjective viewpoint and instincts of investors on particular scenarios are crucial on hierarchical evaluation and assigning altering weights required to adapt varying complicated situations and environments. This chapter has benefitted from the findings of the previous empirical chapters on weighting diversified macroeconomic and political breaking news and announcements regarding to their hierarchical order of importance on investors' subjective perspective. The occurrence of both macroeconomic and political news back to back in the same day is a highly possible situation in real world scenarios. Complex integration of algorithms has concentrated on diverse focuses to enhance the prediction performance and accuracy by consolidating the hierarchical evaluation of macroeconomic and political news, by especially targeting the dates that macroeconomic and political impact overlaps.

The advanced artificial intelligence algorithm that has been modulated by the findings of the previous empirical chapters has evaluated the dependency of the exchange rate fluctuations on the initial release of macroeconomic and political announcements. The requirements of the complicated artificial intelligence algorithm that relied heavily on the subjective and qualitative data have met by the enhanced and reformed versions of previously adopted machine learning algorithms, which are; natural language processing, weighted prospect theory, weighted fuzzy logic, and feature-weighted supports vector regression. The aforementioned methods are interconnected, and associated with the tasks of; calculating the probability of the sentiment outcome for each of almost a thousand news article, assessing decision makers' subjective valuation of those articles, investigating the perspective and expectations on macroeconomic variables, and the degree of certainty and consensus of the political environment to attain the correlation between macroeconomic news, political news and exchange rate fluctuations, and regression analysis for daily exchange rate prediction respectively.

Simulation results indicated that exchange rate is highly responsive to initial publication of announcements, statements, and breaking news regarding to inflation rate, interest rate, and credit ratings, while similar responsiveness of exchange rate did not observed for unemployment rate, and balance of trade. Moreover, it has been determined that inflation rate, interest rate, and unemployment rate are positively correlated with USD/TRY, while balance of trade, and credit ratings are inversely correlated with USD/TRY. Furthermore, simulations on political tension have confirmed the inconsistency of political consensus and certainty, thus the political tension between Turkey and U.S., while the results indicate that certain disagreement and uncertain disagreement have a strong negative impact, uncertain agreement has mild positive/negative impact, and certain agreement has strong positive impact on exchange rates. Comprehending the dependency of exchange rate fluctuations on the news that accommodate macroeconomic announcements and political tension provided a valuable insight for possible combinations of real world scenarios to be guided through the regression analysis. Provided input has split into train and test portions to run MCCV. MCCV has repeated for 65 times to test the consistency of the output, and have been evaluated against seven key statistical metrics in order to clarify the performance of the regression analysis. Statistical test results indicated a valid and accurate prediction without the issues of multicollinearity, heteroscedasticity, biasness, and underfitting or overfitting.

The findings of this chapter explicitly verified the initial impact of government officials' statements, speeches addressing the nation, speeches addressing government agencies and officials, policy implications, and breaking news on impulsive exchange rate fluctuations, thus the substantial role of cognitive approach and analysis. The endnotes have indicated the significance of stance of the government on both internal and foreign policies on investors' decisions, thus the course of the economy, and emphasized the influence of investors' sentiments on fluctuating exchange rates on uncertain environments. Besides, the novel collaboration of diverse disciplines; soft computing, politics, and psychology in purpose of enhancing effectiveness and performance of the economic analysis, and the usability of subjective qualitative information alongside the objective quantitative variables has been firmly confirmed by the superior accuracy of the composed model, thus, findings of the chapter clarified that model performed unambiguously marvelous on observing decision makers' initial responses and behaviors to macroeconomic indicators and political tension, extracting their weights successfully, and predict future exchange rates accordingly.

General Conclusion and Interpretations

This study has implemented qualitative and subjective data alongside the quantitative and objective data to enhance the accuracy of exchange rate prediction, while emphasizing the significance of political sciences in economic fluctuations by investigating the initial impact of recently published news regarding to macroeconomic indicators and political tension to derive the sentiments of investors. To do so, the author has generated a hybrid machine-learning algorithm. The basis of the study has been introduced in Chapter 1, which includes epistemological approach of the thesis, statement of the research problem, objective of the study, research motivation, methodological approach, and the overview of the thesis. General review of theoretical and empirical literature has been introduced in Chapter 2, which comprises studies that mainly investigates the impact of political tension and uncertainty on economics, prediction performance of machine learning algorithms, and the comparisons on the prediction accuracy of machine learning and econometrics models. This chapter had bluntly revealed the significance of political tension and uncertainty on economics in both short and long term, usability of soft computing techniques on economics and finance fields, the superior prediction accuracy of machine learning algorithms compared to traditional econometrics models, and verified the multidisciplinary approach of the author while addressing economic issues.

Chapter 3 emphasized the reasoning behind selecting Turkey, Turkish economy, and USD/TRY exchange rate to study. Turkey and Turkish economy provides an immense opportunity to study the impact of political tension and uncertainty on economics. As Turkey is currently undertaking the role of transit country for natural gas and oil trade between Russia, OPEC and EU, Turkey has a geopolitical influence on the region. Moreover, there is a continuously ongoing war scene in the Middle East. As a NATO member, Turkey has an obligation to make its military bases accessible and usable to US, which makes the combat more accessible and cost effective for the US military. Due to aforementioned reasons, EU and US are continuously applying political and economic pressure on Turkey to weaken Turkey's empowerment, growth, and independence, to restrict Turkey's foreign policies and make them match their interests. For that reason, applied political pressure causes political tension, instability, and uncertainty, which impacts investors' rational expectations and prevent foreign and domestic investments to Turkey, which impacts Turkey's macroeconomic indicators, and currency negatively, which was also supported by the existing literature that has been mentioned throughout the chapter.

The thesis has adopted a multidisciplinary methodology to meet the predetermined objectives of the study, which has been introduced in Chapter 4. Arising expectations from sciences unveil the soft computing techniques that also study qualitative, informal, and approximate information besides the traditional quantitative, formal, and precise theories. The principal elements of soft computing are experimental data, which also known as statistical learning, and fuzzy logic methods. Neural Networks and Support Vector Machines uses statistical learning to learn from feed data, while Fuzzy Logic methods entrenches human knowledge into an analytical model. A sophisticated combination of soft computing methods, namely; natural language processing, fuzzy logic, and supports vector machines in addition to prospect theory of behavioral economics are going to be used in order to accomplish a realistic prediction by overcoming the real life scenarios, through simulating the real life decision maker's behaviors, which is the fundamental goal of the thesis. The methodology is constructed as three phases, namely; Sentiment Analysis, Fuzzy Logic, and Support Vector Regression respectively. Following three sections of this chapter are dedicated to in depth introduction of basic principles of the methods that constructs the three phases of the methodology of empirical Chapters 5 and 6, while weighted variations that are going to construct the phases of the methodology of empirical Chapters 7 and 8.

The causal relationship between macroeconomic indicators and exchange rate fluctuations has been proven by existing literature. While existing literature has concentrated on the impact of the changes in macroeconomic indicators that have spread over a month, which formed as a quantitative research that completely relies on objective historical data, the initial impact of macroeconomic announcements and breaking news on exchange rate that reflects the subjective valuation of qualitative information of investors' has completely be ignored.

The first empirical chapter, Chapter 5 had especially concentrated on the aforementioned research gap. Chapter had emphasized the instantaneous responsiveness of exchange rates to the announcements of macroeconomic variables, government officials' statements and expectations on economic policies, and the breaking news in the media that scrutinize macroeconomic indicators to comment on the direction of the economy.

Existing literature has accepted the correlation between exchange rate and inflation rate (Barro and Gordon, 1983); (Parsley and Wei, 2007); (Qui et al., 2011); (Şen et al., 2019), exchange rate and interest rate (Saraç and Karagöz, 2016); (Cheung et al., 2002); (Karahan and Çolak, 2012); (Greun and Wilkinson, 1994); (Chen, 2007), exchange rate and unemployment rate (Frenkel and Ros, 2006); (He et al., 2013); (Milas and Legrenzi, 2006); (Ranjbar and Moazen, 2009); (Chang, 2011); (Nyahokwe and Ncwadi, 2013), exchange rate and balance of trade (Ozturk, 2006); (Asteriou et al., 2016); (Serenis and Tsounis, 2016), existing rate and credit ratings (Mateev, 2012); Subaşi, 2008); (Nelson, 1991); (Jansen and Haan, 2005).

Parallel to the existing literature, the key findings of the chapter 5 of this thesis have also indicate that inflation rate, interest rate, and unemployment rate are positively correlated with USD/TRY, while balance of trade, and credit ratings are inversely correlated. However, initial responsiveness of exchange rate to the news regarding those macroeconomic indicators was never measured before through the existing literature. This chapter has analyzed macroeconomic announcements, government officials' statements and expectations on economic policies, and the breaking news appeared on the news sources to derive the initial responsiveness of exchange rate, and fill the aforementioned research gap. After detailed fuzzy logic analysis, it has been found that exchange rate is highly responsive to breaking news on inflation rate, interest rate, and credit ratings. On the other hand, similar responsiveness of exchange rate has not been observed to breaking news on unemployment rate and balance of trade.

The result of this chapter specifically verifies that; recently published macroeconomic news based on stories such as sights and thoughts on current macroeconomic conditions, expectations on upcoming rates, changes on policies, and official announcements regarding to macroeconomic indicators plays a vital role and have a significant immediate impact on exchange rates right after the publication.

Turkish Lira depreciated by 152% against US Dollar, and hit record braking lows while becoming the worst performing currency of emerging markets in between the relatively short date interval that this study has concentrated on. If the time interval is restricted into a single week for further investigation, the difference between daily highest and lowest price of USD/TRY exchange rate was 23.00%, 12.48%, 10.93%, 12.55%, 6.59%, and 10.91% for 6 consecutive days in between the dates of 10 August 2018 and 17 August 2018 respectively. The overlapping dates of exaggerated fluctuations on USD/TRY exchange rate and heated political tension between United States and Turkey in consequence of successive decisions of Turkish lower and high courts' to reject Pastor Andrew Brunson's appeal to end his house arrest and lift his travel ban cannot be justified as just a coincidence that has no impact on Turkish economy. In fact, Pastor Andrew Brunson case is just one of many political cases that accommodates political tension, uncertainty, and disagreement that has a clearly visible excessive influence on Turkish economy. However, the political tension on Turkey doesn't look like it will settle down soon, due to Turkey's influential geographical location, thus geopolitical significance on global energy trade, besides the international and regional theatre of war, as clarified through Chapter 3. Due to mentioned reasons; political tension is not a temporary, but ongoing situation for Turkey, thus cannot be excluded from the equation, yet the ways of eliminating/lessen its impact on economy should be considered. To do so, the weight distribution, dependency, and causality of volatile political situations have to be comprehended primarily.

Empirical chapters, Chapter 6 and 7 have especially concentrated on understanding the impact of various degrees of certainty and agreement on exchange rate fluctuations by examining national and international political events that accommodates political tension under the political subcategories of; judicial system, international politics, international relations, sanctions on government officials, warfare, defense industry, political tension, human rights, NATO, elections, cabinet changes, and democracy. To do so, government officials' statements, speeches addressing the nation, speeches addressing government agencies and officials, and breaking political news have been derived from the media. Empirical chapter 6 concentrated on specific national and international political events to examine the impact of aforementioned political subcategories in particular, while chapter 7 benefitted from the findings of the previous chapter to scrutinize the political tension in broader spectrum by not limiting the event count and political subcategories, but concentrating on political tension parallel to the real world timeline and diversity.

Current literature indicates that the speeches that provides opinions or official statements regarding to a occurred political event that accommodates validation of political tension, which presented by political figures has a direct impact on public's political sentiment (He et al., 2017), and fading political harmony has a direct adverse impact on economic activities (Besley and Mueller, 2012). Moreover, military conflict has a direct adverse impact on both firms' profitability, valuation of investors', and fundamental macroeconomic variables as a result of a chain reaction (Abadie and Gardeazabal, 2003; Amihud and Wohl, 2004; Rigobon and Sack, 2005). Although, literature indicates that uncertain political environment that has been triggered by various political events, which have a potential of affecting political harmony and international relations, causing ambiguous policies, and increasing political tension causes a negative impact on economic activities, even if it would not cause any hypothetical military conflict (He et al., 2017); (Gholipour, 2019); (Bloom, 2014); (Besley and Mueller, 2012); (Julio and Yook, 2016); (Pastor and Veronesi, 2013). As a result of previously mentioned studies, it is possible to say that literature had efficiently introduced the strong negative impacts of collapsing government, regime instability, political polarization, and government repression on the tendency of macroeconomic fundamentals (Alesina et al., 1996); (Chen and Feng, 1996); (Jong-a-Pin, 2009); (Berkman et al., 2011). Furthermore, the degree of instability and uncertainty on political regime plays is a significantly influential factor on macroeconomic variables in a short run (Aisen and Veiga, 2013; Campos and Nugent, 2002), as socio-political instability instigates politic-economic uncertainty, which causes a decline on the volume of investments due to surging risks (Alesina and Perotti, 1996; Darby et al., 2004).

Correspondingly, Chapter 6 investigated four political events, namely; Pastor Andrew Brunson case, 2018 Parliamentary and Presidential elections, S-400 crisis, and 2019 Istanbul Mayoral elections to conceive the weight of various political event categories on investors' decision-making process, and how related news on particular political events impact the volatility of exchange rate on instant manners. As a result of deep analysis, it has been found out that, exchange rate fluctuations are highly dependent on the news regarding to Pastor Andrew Brunson Case and S-400 Crisis, and have substantial initial response. On the other hand, the news on Parliamentary and Presidential elections have a moderate initial impact, while Istanbul Mayoral elections have weak initial impact on exchange rate. When the causality of moderate and weak initial impacts of elections has intensely be investigated, it has been found out that, both elections have very high but gradually increasing impact on exchange rate. As the Election Day approaches, news on the elections starts to gain importance. On last couple of days before and after the elections, every election related news becomes very effective, and has a substantial impact on exchange rate fluctuations as they indicates and boosts uncertainty, while that impact may not be observed weeks before elections, which lowers the average impact of news regarding to elections.

Findings derived from examined political events indicate that; recently published news that concerns certain disagreement on international politics that composed under uncertain environment far from compromise plays a vital role and has a direct rapid impact on exchange rates. While Pastor Andrew Brunson case enlightened the algorithm concerning the categories of international judicial process, court hearings, expected court orders, political, sanctions on government officials, and financial sanctions under every possible degrees of certainty and consensus, S-400 crises has contributed to algorithm by inspecting the categories of warfare, defence industry, political tension, human rights, and NATO under certainty and disagreement. Besides, 2018 Parliamentary and Presidential elections, and 2019 Istanbul Mayoral elections have enhanced the algorithm concerning the categories of political instability, elections, Supreme Court, cabinet changes, and democratic uncertainty.

Findings of these chapters revealed the significance of political tension on exchange rate fluctuations, substantial role of cognitive approach and analysis, and achievability of performance enhancement on exchange rate prediction when political analysis implemented. Additionally, it has been observed that; when political tension, disagreement, and uncertainty is caused by the political subcategories like judicial system, international politics, political sanctions, warfare, defense industry, human rights, and NATO, the impact on nation's economy is inevitable, while the impact of political subcategories like human rights, democracy, elections, cabinet changes, and internal politics are only crucial only days prior and after the main event.

In line with the findings of empirical chapters 5, 6, and 7, the final empirical chapter, Chapter 8 has simulated the actual frequency and contents of news feed, impersonate the subjective viewpoint and instinct behaviors of the investors to the breaking news via robust AI by understanding the investors' evaluation process of real world scenarios through setting the subjective order of importance, and assigning altering weights required to adapt varying complicated situations and environments to achieve a highly accurate exchange rate prediction.

Empirical chapters 5 and 6 derived the subjective feelings of the decision makers by examining the fluctuations on the historical data, while the capabilities of the algorithm is augmented in Chapter 7 to derive the subjective standpoint of the decision makers without needing to observe the fluctuations on the historical data. The augmented capabilities of the algorithm is modified and enhanced accordingly for chapter 8 to be able to examine the macroeconomic and political news simultaneously by benefitting from the provided weights and the hierarchical importance throughout the previous empirical chapters. To understand the decision making reference point of each attribute of decision makers', achieved probability values have fuzzified and weighted accordingly, as each breaking news and/or official statement's hierarchical order of importance that determines the decision makers' attitude is fuzzy in itself. By weighting the parameters within a hierarchy, altered valuation of the decision maker have obtained by weighted generalized mean operator, which indicates the decision makers' behavior of evaluation. The fuzzy relations have been obtained by the presence or absence of association, interaction, or interconnectedness between the elements of two or more sets. Interpretation of the resulting fuzzy relations has obtained by the fuzzy rules designated by the author, which feed into fuzzy engine as well as membership functions. Over 400 fuzzy rules have been designated from both macroeconomic and political fuzzy categories of the algorithm and feed into fuzzy engine. Fuzzy engine has generated Fuzzy Consequent, which signifies the weighted correlation between USD/TRY and news and officials' statements. The performance of Polynomial and Gaussian Radial Basis Function kernel functions have compared by the grid search with cross validation results to be selected as the kernel function for the Future Weighted Support Vector Regression analysis. Monte-Carlo cross validation method has implemented to split presented data into training and test groups for prediction. Accuracy of the prediction has been investigated by the statistical tests.

Findings have indicated the significance of stance of the government on both internal and foreign policies on investors' decisions, thus the course of the economy, and emphasized the influence of investors' sentiments on fluctuating exchange rates on uncertain environments.

The main consequent derived from the outcomes of empirical chapters 5, 6, 7, and 8 can be interpreted as;

The impact of political tension and disagreement on nation’s economy, alongside to macroeconomic indicators has been proven. However, it is impracticable to be anticipating easing of political tension and disagreement. Political tension and disagreement between government officials’ and countries will always be a part of politics as long as there is a conflict of interest. However, the economic impact of political tension and disagreement may vary for every country. The determining factor that differentiates the influence on economy is uncertainty. This study has revealed that certainty, whether it’s an agreement or disagreements situation causes exchange rate fluctuations in expected levels. However, there is a serious variability of given weights to uncertain situations in order to predict exchange rates successfully. Not only the uncertainty of the steps taken due to political tension, but also the uncertain attitude of the government causes lack of trust in government, which has a direct influences on domestic and foreign investors’ investment plans, as trust is the foundation of increasing confidence of consumers and investors. In order to obtain public trust, the phenomenon of uncertainty is needed to be broken. Even if the decision made may worsen the political tension, government should exhibit a solid stand behind its decision and attitude as a representation of being faithful to its values, integrity, fairness, and transparency, which are the building blocks of public trust.

In addition to political tension and uncertainty, institutions and organizations that are subject to the government but have the autonomy of policy implementation and decision-making, such as central bank should not be pressured or tried to be directed by the higher authorities. Moreover, authorized persons in the assigned positions should be given the necessary independence and time to implement and run their policies that they believe is the most beneficial solution for the issue encountered. Interpreting the policies of institutions and organizations like central bank, and/or frequently appointing a replacement on authority are the primary factors that are causing uncertainty, lack of trust, thus ease or eliminate the effectiveness of policies applied. For investors, formulation and implementation of economic policies,

and commitment on coping with policy instability and tension under uncertain environment plays a vital role on their investment decisions.
On the other hand, if public trust cannot be obtained, investors' become more inclined to evaluate the speculations, which cause the responsiveness of the market thus influence of speculations to become unnecessarily huge, even if the allegations are baseless.
Consequently, if the perception of integrity, fairness, transparency, and democracy are imposed to public, thus the public trust has gained; the absurd fluctuations on nation's economy can be prevented in case of political tension.

In search of pioneering a paradigm shift on the prediction methodology on the fields of economics and finance, author has generated a sophisticated artificial intelligence algorithm to implement qualitative and subjective data alongside the quantitative and objective data to excel a prediction model. This thesis has underlined the significance of multidisciplinary approach by merging the disciplines of economics, finance, political sciences, psychology, and soft computing on achieving enhanced prediction performance, and suggests considering the sentiments of investors' while valuating and predicting variables that's valuation mechanism related to humans' perception under vague environments. By doing so, thesis has achieved its key objectives.

REFERENCES

Agathangelou, P., Trihinas, D., Katakis, I., (2020). *A Multi-Factor Analysis of Forecasting Methods: A Study on the M4 Competition*. Data 2020, Vol 5 (2), 41-65.

Aisen, A., Veiga, F,. J. (2013). *How does political instability affect economic growth?*. European Journal of Political Economy, 29, 151–167.

Akdemir, I., O., (2011). *Global energy circulation, Turkey's geographical location and petropolitics*. Procedia Social and Behavioral Sciences 19, 71-80.

Angeletos, G. M., Werning, I., (2006). *Crisis and Prices: Information Aggregation, Multiplicity, and Volatility.* American Economic Review, 1720-1736.

Antonakakis, N., Chatziantoniou, I., Filis, G. (2013). *Dynamic co-movements of stock market returns, implied volatility and policy uncertainty*. Economic Letters 120, 87-92.

Bagci H., Kardas S., (2003). *Post-11 September Impact: The Strategic Importance of Turkey Revisited*. CEPS/IISS European Security Forum, Brussels, 12 May 2003, No.13.

Baker, S., R., Bloom, N., Davis, S., J. (2012). *Has Economic Policy Uncertainty Hampered the Recovery?* In: Ohanian, L., E., Taylor, J., B., Wright, I. J. (Eds.), Government Policies and the Delayed Economic Recovery. Hoover Institution, Stanford University. Chapter 3.

Barro, R. J. and Gordon, D. B. (1983). *A Positive Theory of Monetary Policy in Natural Rate Model*. Journal of Political Economy, Vol: 91(4), pp. 589-610.

Berkman, Henk, Jacobsen, Ben, Lee, John B. (2011). *Time-varying rare disaster risk and stock returns*. Journal of Financial Economics. 101, 313–332.

Bernanke, B.S. (1983). *Irreversibility, uncertainty and cyclical investment*. The Quarterly Journal of Economics Vol 98(1), 85-106.

Besley, T., Mueller, H. (2012). *Estimating the peace dividend: The impact of violence on house prices in Northern Ireland*. American Economic Review 102 (2), 810–833.

Bezerra, P. C. S. , & Albuquerque, P. H. M. (2017). *Volatility forecasting via SVR—GARCH with mixture of Gaussian kernels*. Computational Management Science, 14 (2), 179–196.

BBC News (2019). *Trump threatens Turkey sanctions over pastor Andrew Brunson.* Available at: https://www.bbc.com/news/world-us-canada-44974722

BBC News, (2019a). Turkey defies US as Russian S-400 missile defence arrives. Available at: https://www.bbc.com/news/world-europe-48962885

BBC News, (2019b). Where does Turkey's S-400 missile deal with Russia leave the US?. Available at: https://www.bbc.com/news/world-europe-48962886

BBC News, (2019c). Erdogan's party suffers blow after Istanbul re-run poll defeat. Available at: https://www.bbc.com/news/world-europe-48739256

BBC News, (2019d). Istanbul mayoral vote: Is ‘disastrous’ loss beginning of Erdogan’s end?. Available at: https://www.bbc.com/news/world-europe-48744733

Bollen, J., Mao, H., Zeng, X., (2011). *Twitter mood predicts the stock market*. Journal of Computational Science 2 (1), 1–8.

Boostan, E., Tahernia, N., Shafiee, A., (2015). *Fuzzy – Probabilistic seismic hazard assessment, case study: Tehran region, Iran*. Natural Hazards, vol 77(2), 525-541.

CBS News (2018). *Turkey aims new tariffs at U.S. as court rejects pastor's appeal.* Available at: https://www.cbsnews.com/news/turkey-us-tariffs-court-rejects-pastor-andrew-brunson-appeal-turkish-lira-crisis/

Cehreli, C., Dursun, I., Barlas, Y. (2017). *Speculative Dynamics of Exchange Rates in Turkey: A System Dynamics Approach.* Yildiz Social Science Review, 103-120.

Chen, H. , Xiao, K. , Sun, J. , & Wu, S. (2017). *A double-layer neural network frame- work for high-frequency forecasting*. ACM Transactions on Management Information Systems (TMIS), 7 (4), 11:2–11:17.

Chen, S., (2007). *A note on interest rate defense policy and exchange rate volatility*. Economic Modeling, 24 (5), 768–777.

Chen, Y. S. , Cheng, C.-H. , & Tsai, W. L. (2014). *Modeling fitting-function-based fuzzy time series patterns for evolving stock index forecasting*. Applied Intelligence, 41 (2), 327–347.

Cheng, M-Y., Roy, A. F. V. (2011). *Evolutionary fuzzy decision model for cash flow prediction using time-dependent support vector machines.* International Journal of Project Management, 29, 56-65.

Cheung, Y., Chinn, M.D., Pascual, A.,G., (2002). *What do We Know about Recent Exchange Rate Models? In-Sample Fit and Out-of Sample Performance Evaluated*. CESIFO Working Paper 902

Chiang, W.C., Enke, D., Wu, T., Wang, R., (2016). *An adaptive stock index trading decision support system*. Expert Systems with Applications, 59 (1), 195–207.

Christodoulou, S., Agathokleous, A., Kranioti, S., (2012). *Analytical and Numerical Models for the Risk-of-Failure Analysis of Urban Water Distribution Network Components.* NIREAS International Water Research Center. NIREAS-IWRC/D5.13.2 Version 1.1.

Chuku, C., Simpasa, A., Oduor, J. (2019). *Intelligent forecasting of economic growth for developing economies.* International Economics, 159, 74-93.

CNN Türk, (2017). *Anayasa değişikliği maddeleri tam metni | Yeni anayasa maddeleri nelerdir?*. Available at: https://www.cnnturk.com/video/turkiye/anayasa-degisikligi-maddeleri-tam-metni-yeni-anayasa-maddeleri-nelerdir

Cohen, S., B., (2003). *Geopolitics of the World System*. Political Geography Vol 23(6), 801-804.

Cumhuriyet, (2019). İstanbul Büyükşehir Belediye Başkanı Ekrem İmamoğlu. Available at: http://www.cumhuriyet.com.tr/haber/siyaset/1350201/istanbul_Buyuksehir_Belediye_Baskani_Ekrem_imamoglu.html

Çakar, N., (1998). *A Strategic Overview of Turkey*. Journal of International Affairs, Vol3(2).

Defensenews, (2019). Turkey officially kicked out of F-35 program, costing US half a billion dollars. Available at: https://www.defensenews.com/air/2019/07/17/turkey-officially-kicked-out-of-f-35-program/

Drucker, H., Burgers, C., J., C., Kaufman, L., Smola, A., J., Vapnik, V. (1996). *Support Vector Regression Machines.* Advances in Neural Information Processing Systems 9, Massachusetts Institute of Technology Press.

Dubitzky, W., Granzow, M., Berrar, D., P., (2007). *Fundamentals of Data*

Mining in Genomics ad Proteomics. Springer Science & Business Media. ISBN 978-0-387-47509-7.

EPC (2018). *Andrew Brunson Summary/Timeline.* Available at: https://epc.org/wp-content/uploads/Files/4-Resources/2-News-And-Information/AndrewBrunsonSummaryTimeline.pdf

Euronews, (2019). Could Imamoglu victory in Istanbul be 'beginning of the end' for Erdogan?. Available at: https://www.euronews.com/2019/06/24/could-imamoglu-victory-in-istanbul-be-beginning-of-the-end-for-erdogan

Geisser, S. (1975). *The predictive sample reuse method with applications*. Journal of the American Statistical Association, 70(350):320 328.

Gholipour, H. F. (2019). *The effects of economic policy and political uncertainties on economic acrivities*. Research in International Business and Finance, 48, 210-218.

Giuso, L., Parigi, G. (1999). *Investment and demand uncertainty*. The Quarterly Journal of Economics 114, 185-227.

Grigoryan, H. (2016). *A Stock Market Prediction Method Based on Support Vector Machines (SVM) and Independent Component Analysis (ICA).* Database Systems Journal, 7 (1), 12-21.

Grubbs, F., E., (1973). *Errors of Measurement, Precision, Accuracy, and the Statistical Comparison of Measuring Instruments*. Technometrics, Vol 15(1), 53-66.

Gu, S., Kelly, B., Xiu, D. (2020). *Empirical Asset Pricing via Machine Learning.* The Review of Financial Studies, 33 (5), 2223-2273.

Guh, Y-Y., Po, R-W., and Lee, E., S., (2008). *The fuzzy weighted average within a generalized means function*. Computers and Mathematics with Applications, 55 (2008) 2699-2706.

Gunay., S. (2016). *Is Political Risk still an Issue for Turkish Stock Market?*. Borsa Istanbul Review, Research and Business Development Department, Borsa Istanbul, 16(1), 21-31.

Habertürk, (2019). YSK Başkanı açıkladı! İmamoğlu önde. Available at: https://www.haberturk.com/ysk-baskani-acikladi-imamoglu-onde-2420335

Hair, J., F., Jr., Black, W., C., Babin, B., J., Anderson, R., E., (2010). *Multivariate Data Analysis: A Global Perspective*. 7th Edition. Pearson Prentice Hall: Upper Saddle River, NJ, USA. ISBN: 0135153093.

Hair, J., F., Ringle, C., M., Sarstedt, M., (2013). *Partial Least Squares Structural Equation Modeling: Rigorous Applications, Better Results and Higher Acceptance*. Long Range Planning 46(1-2), 1-12.

Hair, J., F., Sarstedt M., Hopkins, L., Kuppelweiser, V., (2014). *Partial Least Squares Structural Equation Modeling (PLS-SEM): An Emerging Tool for Business Research*. European Business Review 26(2): 106-121.

He, Y., Nielsson, U., Wang, Y. (2017). *Hurting without hitting: The economic cost of political tension.* Journal of International Financial Markets, Institutions & Money, 51, 106-124.

Henrique, B. M. , Sobreiro, V. A. , & Kimura, H. (2018). *Building direct citation networks*. Scientometrics, 115 (2), 817–832.

Henseler, J., Ringle, C., M., Sinkovics, R., R., (2009). *The use of partial least squares path modeling in international marketing*. New Challenges to International Marketing Advances in International Marketing, Vol 20, 277-319.

Hillier, D., Loncan, T. (2019). *Political uncertainty and Stock returns: Evidence from the Brazilian Political Crisis.* Pasific-Basin Finance Journal, 54, 1-12.

Hsu, C-W., Chang C.-C., Lin, C.-J., (2003) *A Practical Guide to Support Vector Classification*. Technical Report, Department of Computer Science, National Taiwan University.

Huang, C., F., (2012). *A hybrid stock selection model using genetic algorithms and support vector regression*. Applied Soft Computing,12, 807–818.

Hürriyet, (2018). *Erdoğan açıkladı... Erken seçim tarihi belli oldu*. Available at: https://www.hurriyet.com.tr/gundem/erdogandan-erken-secim-icin-son-dakika-aciklamasi-40809175

Hürriyet, (2019). AK Parti'den flaş karar! İstanbul'un tüm ilçelerinde seçime itiraz edilecek... Available at: https://www.hurriyet.com.tr/gundem/ak-partiden-flas-karar-istanbulun-tum-ilcelerinde-secime-itiraz-edilecek-41169769

Independent, (2019). YSK İstanbul seçimini iptal etti, 23 Haziran'da yeniden sandık kurulacak. Available at: https://www.indyturk.com/node/28476/haber/ysk-istanbul-seçimini-iptal-etti-23-haziranda-yeniden-sandık-kurulacak

James G., Witten D., Hastie T., Tibshirani R., (2017). *An Introduction to Statistical Learning: with Applications in R*. 7th printing 2017 edition, ISBN-13: 978-1461471370

Jansen, D., J., Haan, J., (2005). *Talking heads: the effects of ECB statements on the euro-dollar exchange rate*. Journal of International Money and Finance, Elsevier, vol. 24(2), pages 343-361.

Julio, B., Yook, Y. (2012). *Political uncertainty and corporate investment cycles*. Journal of Finance 67 (1), 45–83.

Julio, B., Yook, Y. (2016). *Policy uncertainty, irreversibility, and cross-border flows of capital*. Journal of International Economics, 103(C), 13-26.

Kahneman, D., Tuersky, A., (1979). *Prospect Theory: An Analysis of Decision under Risk*. Econometrica, 47 (2), 263-292.

Kahneman, D., Tuersky, A., (1992). *Advances in prospect theory: Cumulative representation of uncertainty.* Journal of Risk and Uncertainty, 5, 297-323.

Kandil, M., Berument, H., Dincer, N. N. (2007). *The effects of exchange rate fluctuations on economic activity in Turkey*. Journal of Asian Economics, 18, 466-489.

Kara, Y. , Boyacioglu, M. A. , & Baykan, Ö. K. (2011). *Predicting direction of stock price index movement using artificial neural networks and support vector machines: The sample of the Istanbul Stock Exchange*. Expert Systems with Applications, 38 (5), 5311–5319.

Karatas, S., (2010). *Delays inn Turkish-Azeri Gas Deal Raises Uncertainty Over Nabucco*. Eurasia Daily Monitor, Vol 7(39).

Kecman, V. (2001). *Learing and Soft Computing: Support Vector Machines, Neural Networks, and Fuzzy Logic Models*. The MIT Press. ISBN 978-0-262-11255-0.

Kim, K. (2003). *Financial time series forecasting using support vector machines*. Neurocomputing, 55 (1–2), 307–319.

Kleynhans, T., Montanaro, M., Gerace, A., Kanan, C., (2017). *Predicting Top-of-Atmosphere Thermal Radiance Using MERRA-2 Atmospheric Data with Deep Learning*. Remote Sensing, Anaheim, USA, Vol 10178.

Klir, G., J., Folger T., A. (1988). *Fuzzy Sets, Uncertainty, and Information.* Prentice-Hall International, ISBN: 0133456382.

Koch, G. G., Landis, J. R., Freeman,J. L., Freeman,D. H., Jr., Lehnen,R. G. (1977). A General Methodology for the Analysis of Experiments with Repeated Measurement of Categorical Data. Biometrics Vol 33, 133-158.

Korol, T. (2014). *A fuzzy logic model for forecasting exchange rates*. Knowledge-Based Systems, 67,49–60.

Kumar, D. , Meghwani, S. S. , & Thakur, M. (2016). *Proximal support vector machine based hybrid prediction models for trend forecasting in financial markets*. Journal of Computational Science, 17 (1), 1–13.

Kumar, M. , & Thenmozhi, M. (2014). *Forecasting stock index returns using ARIMA-SVM, ARIMA-ANN, and ARIMA-random forest hybrid models*. International Journal of Banking, Accounting and Finance, 5 (3), 284–308.

Landis J., R., Koch, G., G., (1977). *The Measurement of Observer Agreement for Categorical Data*. Biometrics, Vol 33(1), 159-174.

Liu, S., Lin, Y., (2010). *Grey Systems: Theory and Application*. Berlin: Springer-Verlag. ISBN: 978-3-642-16157-5.

Liu, P., Jin, F., Zhang, X., Su, Y., Wang, M., (2011). *Research on the multi-attribute decision-making under risk with interval probability based on prospect theory and the uncertain linguistic variables*. Knowledge-Based Systems, 24, 554-561.

Martell, R., (2005). *The Effect of Sovereign Credit Rating Changes on Emerging Stock Markets*. Working Paper. Purdue University (2005).

Mateev, M., (2012). *The Effect of Sovereign Credit Rating Announcements on Emerging Bond and Stock Markets: New Evidences*. Oxford Journal: An International Journal of Business & Economics, 7(1), 28-41.

Mehdian, S., Nas, T., Perry, M., J. (2008). *An examination of investor reaction to unexpected political and economic events in Turkey*. Global Finance Journal, 18, 337–350.

Mentaschi, L., Besio, G., Cassola, F., Mazzino, A., (2013). *Problems in RMSE-based wave model validations*. Ocean Modeling 72, 53-58.

Nasehi, S., Olia, S., S., B., Karimi, S., Heydari, S., (2017). *Modeling site selection for solar power establishment by fuzzy logic and ordered weighted averaging methods in arid and semi-arid regions (Case study Yazd province-Iran)*. Journal of Biodiversity and Environmental Sciences, Vol 10(5), p 177-192.

Nelson, D. (1991). *Conditional Heteroscedasticity in Asset Returns: A New Approach*. Econometrica , 59 (2), pp. 347-370.

Nelson, M., Illingworth, W.T., (1991). *A practical guide to neural nets*. Addison-Wesley Longman Publishing Co., Inc., ISBN:978-0-201-52376-8.

Nilashi, M., Bagherifard, K., Othman, I., Janahmadi, N., Barisami, M., (2011). *An Application Expert System for Evaluating Effective Factors on Trust in B2C Websites.* Engineering, Vol 3(11), ID: 8656.

NTV, (2018). *İYİ Parti Genel Başkanı Akşener'den 24 Haziran erken seçiminde adaylık açıklaması: 100 bin imzayla Cumhurbaşkanı adayı olacağım*. Available at: https://www.ntv.com.tr/turkiye/iyi-parti-genel-baskani-aksenerden-24-haziran-erken-seciminde-adaylik-aciklamasi,bDEX_kvnB065HK_Y2E_tnw

Oliveira, N. , Cortez, P. , & Areal, N. (2017). *The impact of microblogging data for stock market prediction: Using Twitter to predict returns, volatility, trading volume and survey sentiment indices*. Expert Systems with Applications, 73 (1), 125–144.

Parsley, D., Wei, S., (2007). *A Prism into the PPP Puzzles: The Micro-Foundations of Big Mac Real Exchange Rates*. The Economic Journal, Volume 117, Issue 523, P1336–1356.

Pastor, L., Veronesi, P. (2013). *Political uncertainty and risk premia*. Journal of Financial Economics 110, 520–545.

Pradigis, I., Karapistoli, E., Tsintzos, P., Sarigiannidis, A., Ravenko, A., Panou, G. (2020). *Market Sentiment Models and Indicators.* Horizon 2020 Information and Communication Technologies, 12-64.

Picard, R., R., Cook, R., D. (1984). *Cross-validation of regression models*. Journal of the American Statistical Association, 79(387):575 583, 1984.

Qui, M., Pinfold, J., Rose, L., (2011). *Predicting foreign exchange movements using historic deviations from PPP*. International Review of Economics and Finance, 20, 485-497.

Remesan R., Mathew J., (2014). *Hydrological Data Driven Modelling: A Case Study Approach***.** Vol. 1. Springer, 2014. ISBN 978-3-3-319-09235-5 (eBook)

Reuters, (2018). *Turkey's main opposition nominates combative former teacher to challenge Erdogan*. Available at: https://www.reuters.com/article/us-turkey-election/turkeys-main-opposition-nominates-combative-former-teacher-to-challenge-erdogan-idUSKBN1I50V4

Robinson, V., B. (2003). *A Perspective onn the Fundamentals of Fuzzy Sets and their Use in Geographic Information Systems.* Transactions in GIS, 7(1), 3-30.

Rodrik, D. (1991). *Policy uncertainty and private investment in developing countries*. Journal of Development Economics 36, 229-242.

Rosenberg, A., Hirschberg, J., (2007). *V-Measure: A conditional entropy-based external cluster evaluation measure.* Joint Conference on Empirical Methods in Natural Language Processing and Computational Natural Language Learning, June 2007, 410-420.

Saraç, T., B., Karagöz, K. (2016). *Impact of Short-term Interest Rate on Exchange Rate: The Case of Turkey*. Procedia Economics and Finance, 38, 195-202.

Sarigiannidis, A., Karypidis P-A., and Pragidis I-C., (2018). *A Novel Lexicon-Based Approach in Determining Sentiment in Financial Data Using Learning Automata*. Internet Science. Springer International Publishing. DOI: 10.1007/978-3-319-77547-0_4.

Sarstedt, M., Mooi, E., (2014). *A Concise Guide to Market Research: The Process, Data, and Methods Using IBM SPSS Statistics*. Second Edition, Springer Berlin Heidelberg. ISBN: 978-3-642-53965-7.

SAS Institute (2017). *JMP 13 Fitting Linear Models*. Second Edition. Publisher(s): SAS Institute ISBN: 9781629609522

Saul, B., C., (2004). *The Geopolitics of Turkey's Accession to the European Union*. Eurasian Geography and Economics, 45(8), 575-582.

Schumaker, R. P. , & Chen, H. (2009). *Textual analysis of stock market prediction using breaking financial news: The AZFin text system*. ACM Transactions on Information Systems (TOIS), 27 (2), 12.

Sharma, H., Sharma, D., K., Hota, H., S. (2016). *A Hybrid Neuro-Fuzzy Model For Foreign Exchange Rate Prediction.* Academy of Accounting and Financial Studies Journal, 20 (3), 1-13.

Smyl, S., Ranganathan, J., Pasqua, A., (2018). *M4 Forecasting Competition: Introducing a New Hybrid ES-RNN Mode*l. Uber Engineering. Published in 25 June 2018. Available at: https://eng.uber.com/m4-forecasting-competition/

Sözcü, (2019). SEÇİM SONUÇLARI 2019: İstanbul'da oylar yeniden sayılacak mı? İşte İstanbul'da son durum… Available at: https://www.sozcu.com.tr/2019/gundem/istanbul-secim-sonuclari-2019-secim-yeniden-mi-yapilacak-iste-istanbulda-son-durum-4365689/

Specht D. F., (1991). *A General Regression Neural Network*. IEEE Transactions on Neural Networks. Vol 2. No. 6. November 1991.

Subaşi, F.Ö. (2008). *The effect of sovereign rating changes on stock returns and exchange rates*. International Research Journal of Finance and Economics, 20, 46-54.

Swanson, D., A., (2015). *On the Relationship among Values of the Same Summary Measure of Error when it is used across Multiple Characteristics at the Same Point in Time: An Examination of MALPE and MAPE*. Review of Economics & Finance, Better Advances Press, Vol 5(3), 1-14.

Şen, H., Kaya, A., Kaptan, S., Cömert, M., (2019). *Interest rates, inflation, and exchange rates in fragile EMEs: A fresh look at the long-run interrelationships*. 2019. hal-02124985.

Taboada, M., Brooke, J., Tofiloski, M., Voll, K., Stede, M. (2011). *Lexicon-based methods for sentiment analysis.* Computational Linguistics, 37(2), 267-307.

Tekin, A., Walterova, I., (2007). *Turkey's Geopolitical Role: The Energy Angle*. Midde East Policy Vol 14(1), 84-95.

The National News, (2019). The beginning of the end for Erdogan?. Available at: https://www.thenationalnews.com/opinion/editorial/the-beginning-of-the-end-for-erdogan-1.878567

The New York Times (2018). *Trump Hits Turkey When It's Down, Doubling Tarrifs.* Available at: https://www.nytimes.com/2018/08/10/us/politics/trump-turkey-tariffs-currency.html

Time (2018). *Who is Andrew Brunson, the Evangelical Pastor Freed in Turkey?* Available at: https://time.com/5351025/andrew-brunson-trump-turkey/

TRT World, (2017). Why is Turkey buying the Russian S-400 missile defence system?. Available at: https://www.trtworld.com/turkey/what-does-the-purchase-of-russian-s-400s-mean-to-turkey-409563

Turkish Probe, (2002). *Turkey Overhauling Foreign Investment Procedures*. Turkish Probe, No.487, 26 May 2002.

Turney, P. D. (2002). *Thumbs Up or Thumbs Down?: Semantic Orientation Applied to Unsupervised Classification of Reviews.* Proceedings of the 40th Annual Meeting of the Association for Computational Linquistics (ACL), Philadelphia, 2002, 417-424.

Türkiye Gazetesi, (2018). *Bahçeli: Erdoğan'ı destekliyoruz*. Available at: https://www.turkiyegazetesi.com.tr/gundem/534712.aspx

T24, (2019a). AK Parti, 'Çünkü çaldılar' söylemini yerleşik hale getirmeye çalışıyor. Available at: https://t24.com.tr/haber/cunku-caldilar-soylemini-yerlesik-hale-getirmeye-calisiyorlar,821604

T24, (2019b). Erdoğan: İBB Başkanlığı seçimlerinde demokrasi ve sandığa gölge düştü, oyları çaldılar. Available at: https://t24.com.tr/haber/erdogan-ibb-baskanligi-secimlerinde-demokrasi-ve-sandiga-golge-dustu-oylari-caldilar,823964

United States Commission on International Religious Freedom (2018). *Pastor Andrew Brunson.* Available at: https://www.uscirf.gov/pastor-andrew-brunson

Vapnik, V. (1995). *The nature of statistical learning theory*. New York, New York: Springer Heidelberg.

Vapnik, V. (1998). *Statistical Learning Theory*. Wiley-Interscience, ISBN 0-471-03003-1.

Veronesi, P. (1999). *Stock market overreactions to bad news in good times: a rational expectations equilibrium model*. Review of Financial Studies 12 (5), 975–1007.

Von Neumann, J., Morgenstern, O. (1947). *Theory of games and economic behavior.* Princeton University Press. Second Edition.

Wang, J. J. , Wang, J. Z. , Zhang, Z. G. , & Guo, S. P. (2012). *Stock index forecasting based on a hybrid model*. Omega, 40 (6), 758–766.

Weng, B., Ahmed, M. A., & Megahed, F. M. (2017). *Stock market one-day ahead movement prediction using disparate data sources*. Expert Systems with Applications, 79 (1), 153–163.

Yasir, M., Durrani, M., Afzal, S., Maqsood, M., Aadil, F., Mehmood, I., Rho, S. (2019). *An Intelligent Event-Sentiment-Based Daily Foreign Exchange Rate Forecasting System*. Journal of Applied Sciences, 9, 2980.

Yenicaggazetesi, (2017). *Kılıçdaroğlu erken seçim dedi*. Available at: https://www.yenicaggazetesi.com.tr/kilicdaroglu-erken-secim-dedi-175644h.htm

Zahed, L. (1965). *Fuzzy sets*. Information and Control, 8 (3), 338-353.

Zahed, L. (1973). *Outline of a new approach to the analysis of complex systems and decision processes*. IEEE Trans. Systems, Man and Cybernetics, (3), 28-44.

Zanotti, J., Thomas, C., (2020). *Turkey: Background and U.S. Relations In Brief*. Congressinal Research Service, CRS Report R44000.

Zhong, X., Enke, D. (2017). *Forecasting daily stock market return using dimensionality reduction*. Expert Systems with Applications, 67 (1), 126–139.

APPENDIX

TABLES

Tab. 1 Credit Rating Measures of the Big Three Rating Agencies

Interpretation	Moody's	Standard&Poor's	Fitch
Investment-Grade Ratings			
Highest Quality	Aaa	AAA	AAA
High Quality	Aa1	AA+	AA+
	Aa2	AA	AA
	Aa3	AA-	AA-
Strong Payment Capacity	A1	A+	A+
	A2	A	A
	A3	A-	A-
Adequate Payment Capacity	Baa1	BBB+	BBB+
	Baa2	BBB	BBB
	Baa3	BBB-	BBB-
Speculative-Grade Ratings			
Likely to fulfill obligations, but with ongoing uncertainty	Ba1	BB+	BB+
	Ba2	BB	BB
	Ba3	BB-	BB-
High-Risk Obligation	B1	B+	B+
	B2	B	B
	B3	B-	B-
Highly vulnerable to nonpayment	Caa1	CCC+	CCC
	Caa2	CCC	CC
	Caa3	CCC-	C
Payment Default	Ca	CC	DDD
	C	SD	DD
		D	D

Tab. 2 Parameter Estimation Using Grid Search with Cross-Validation for Macroeconomic News

Grid scores on development set:		C	Gamma	Kernel
0.333	(+/-0.000)	1	2.00E-10	RBF
0.333	(+/-0.000)	1	2.00E-09	RBF
0.333	(+/-0.000)	1	2.00E-08	RBF
0.333	(+/-0.000)	1	2.00E-07	RBF
0.333	(+/-0.000)	1	2.00E-06	RBF
0.667	(+/-0.000)	1	0.002	RBF
0.333	(+/-0.000)	1	0.001	RBF
0.870	(+/-0.193)	1	0.01	RBF
0.942	(+/-0.106)	1	0.1	RBF
0.717	(+/-0.191)	1	20	RBF
0.347	(+/-0.053)	1	100	RBF
0.333	(+/-0.000)	1	200	RBF
0.333	(+/-0.000)	1	2000	RBF
0.333	(+/-0.000)	10	2.00E-10	RBF
0.333	(+/-0.000)	10	2.00E-09	RBF
0.333	(+/-0.000)	10	2.00E-08	RBF
0.333	(+/-0.000)	10	2.00E-07	RBF
0.333	(+/-0.000)	10	2.00E-06	RBF
0.909	(+/-0.106)	10	0.002	RBF
0.903	(+/-0.118)	10	0.001	RBF
0.953	(+/-0.077)	10	0.01	RBF
0.967	(+/-0.082)	10	0.1	RBF
0.744	(+/-0.223)	10	20	RBF
0.360	(+/-0.065)	10	100	RBF
0.333	(+/-0.000)	10	200	RBF
0.333	(+/-0.000)	10	2000	RBF
0.333	(+/-0.000)	100	2.00E-10	RBF
0.333	(+/-0.000)	100	2.00E-09	RBF
0.333	(+/-0.000)	100	2.00E-08	RBF
0.333	(+/-0.000)	100	2.00E-07	RBF
0.333	(+/-0.000)	100	2.00E-06	RBF
0.967	(+/-0.082)	100	0.002	RBF
0.953	(+/-0.077)	100	0.001	RBF
0.967	(+/-0.082)	100	0.01	RBF
0.923	(+/-0.107)	100	0.1	RBF
0.744	(+/-0.223)	100	20	RBF
0.360	(+/-0.065)	100	100	RBF
0.333	(+/-0.000)	100	200	RBF
0.333	(+/-0.000)	100	2000	RBF
0.333	(+/-0.000)	1000	2.00E-10	RBF
0.333	(+/-0.000)	1000	2.00E-09	RBF

0.333	(+/-0.000)	1000	2.00E-08	RBF
0.333	(+/-0.000)	1000	2.00E-07	RBF
0.667	(+/-0.000)	1000	2.00E-06	RBF
0.967	(+/-0.082)	1000	0.002	RBF
0.967	(+/-0.082)	1000	0.001	RBF
0.927	(+/-0.096)	1000	0.01	RBF
0.923	(+/-0.107)	1000	0.1	RBF
0.744	(+/-0.223)	1000	20	RBF
0.360	(+/-0.065)	1000	100	RBF
0.333	(+/-0.000)	1000	200	RBF
0.333	(+/-0.000)	1000	2000	RBF
0.333	(+/-0.000)	10000	2.00E-10	RBF
0.333	(+/-0.000)	10000	2.00E-09	RBF
0.333	(+/-0.000)	10000	2.00E-08	RBF
0.667	(+/-0.000)	10000	2.00E-07	RBF
0.909	(+/-0.106)	10000	2.00E-06	RBF
0.940	(+/-0.113)	10000	0.002	RBF
0.927	(+/-0.096)	10000	0.001	RBF
0.923	(+/-0.171)	10000	0.01	RBF
0.923	(+/-0.107)	10000	0.1	RBF
0.744	(+/-0.223)	10000	20	RBF
0.360	(+/-0.065)	10000	100	RBF
0.333	(+/-0.000)	10000	200	RBF
0.333	(+/-0.000)	10000	2000	RBF
0.333	(+/-0.000)	10000	2.00E-10	RBF
0.333	(+/-0.000)	10000	2.00E-09	RBF
0.333	(+/-0.000)	10000	2.00E-08	RBF
0.667	(+/-0.000)	10000	2.00E-07	RBF
0.909	(+/-0.106)	10000	2.00E-06	RBF
0.940	(+/-0.113)	10000	0.002	RBF
0.927	(+/-0.096)	10000	0.001	RBF
0.923	(+/-0.171)	10000	0.01	RBF
0.923	(+/-0.107)	10000	0.1	RBF
0.744	(+/-0.223)	10000	20	RBF
0.360	(+/-0.065)	10000	100	RBF
0.333	(+/-0.000)	10000	200	RBF
0.333	(+/-0.000)	10000	2000	RBF
0.333	(+/-0.000)	100000	2.00E-10	RBF
0.333	(+/-0.000)	100000	2.00E-09	RBF
0.667	(+/-0.000)	100000	2.00E-08	RBF
0.909	(+/-0.106)	100000	2.00E-07	RBF
0.967	(+/-0.082)	100000	2.00E-06	RBF
0.940	(+/-0.113)	100000	0.002	RBF
0.957	(+/-0.072)	100000	0.001	RBF
0.923	(+/-0.171)	100000	0.01	RBF
0.923	(+/-0.107)	100000	0.1	RBF

0.744	(+/-0.223)	100000	20	RBF
0.360	(+/-0.065)	100000	100	RBF
0.333	(+/-0.000)	100000	200	RBF
0.333	(+/-0.000)	100000	2000	RBF

Tab. 3 Accuracy of the Parameter Estimation Using Grid Search with Cross-Validation for Macroeconomic News

The model is trained on the full development set.
The scores are computed on the full evaluation set.

	precision	recall	f1-score	support
0	1.00	1.00	1.00	21
1	0.93	0.97	0.94	30
2	0.97	0.96	0.96	24
Accuracy			0.96	75
Macro Avg	0.95	0.98	0.95	75
Weighted Avg	0.95	0.97	0.95	75

Table 4. Averaged Monte Carlo Cross Validation Statistical Test Results for Macroeconomic News

VMS	0.919392
EVS	0.928443
MSE	0.050812
RMSE	0.225206
MAE	0.192386
MAPE	0.039060
R^2	0.924597

Table 5 Exhaustive Monte-Carlo Cross Validation Statistical Test Results for Macroeconomic News

VMS	EVS	MSE	RMSE	MAE	MAPE	R^2
0.917043	0.925387	0.054812	0.234120	0.196124	0.039657	0.920749
0.927700	0.931250	0.046563	0.215784	0.187587	0.038305	0.928321
0.913102	0.921740	0.053304	0.230876	0.197203	0.040284	0.918441
0.905373	0.932245	0.050255	0.224177	0.192872	0.038489	0.927420
0.922020	0.921689	0.053494	0.231288	0.192281	0.038785	0.916168
0.918056	0.930977	0.049036	0.221441	0.191699	0.039250	0.928093
0.921476	0.929501	0.051350	0.226604	0.194039	0.039429	0.925886
0.931755	0.937034	0.043030	0.207436	0.184206	0.037053	0.933583
0.917140	0.928337	0.050804	0.225397	0.191501	0.039224	0.925121
0.928150	0.930048	0.049391	0.222242	0.192233	0.039151	0.926531
0.927822	0.930826	0.048096	0.219308	0.191509	0.039170	0.927617
0.919871	0.930318	0.051292	0.226478	0.189875	0.038525	0.926935
0.915559	0.933407	0.049670	0.222867	0.188541	0.038120	0.929943
0.912885	0.923607	0.053707	0.231748	0.194833	0.039721	0.920545
0.921878	0.929890	0.048518	0.220269	0.189951	0.038797	0.926912
0.927250	0.929170	0.052956	0.230122	0.190004	0.038345	0.925184
0.925316	0.932335	0.048599	0.220453	0.190450	0.038882	0.929030
0.921321	0.930038	0.048779	0.220859	0.191653	0.038423	0.925389
0.920277	0.917187	0.057483	0.239756	0.198595	0.039989	0.911595
0.906313	0.924245	0.055804	0.236228	0.198692	0.040067	0.918329
0.928415	0.935346	0.045093	0.212352	0.186670	0.038470	0.933414
0.922916	0.945243	0.038882	0.197186	0.179696	0.036542	0.942830
0.926905	0.928613	0.051583	0.227119	0.186654	0.037947	0.924574
0.923057	0.924832	0.053471	0.231237	0.199243	0.040578	0.920730
0.917681	0.922560	0.052550	0.229237	0.195605	0.039869	0.919071
0.915810	0.923124	0.054884	0.234274	0.193677	0.039093	0.917299
0.921914	0.922460	0.054603	0.233674	0.198356	0.040080	0.918118
0.917134	0.928621	0.050906	0.225623	0.194879	0.039809	0.925993
0.911042	0.924324	0.056433	0.237557	0.196566	0.039589	0.919334
0.924989	0.940784	0.042012	0.204969	0.181652	0.037174	0.938455
0.907759	0.920534	0.056922	0.238584	0.198146	0.040004	0.914902
0.923764	0.922209	0.054709	0.233900	0.198087	0.040112	0.917676
0.904044	0.926725	0.055767	0.236151	0.199034	0.040353	0.921534
0.928594	0.929905	0.047618	0.218216	0.189624	0.038638	0.926673
0.916978	0.937709	0.041793	0.204433	0.183751	0.037097	0.934891
0.918914	0.930901	0.048844	0.221008	0.190968	0.038848	0.927469
0.916599	0.930783	0.048747	0.220786	0.191353	0.038970	0.927057
0.914654	0.922916	0.053187	0.230623	0.196904	0.040114	0.919371
0.918114	0.931046	0.049621	0.222757	0.192316	0.039144	0.927385
0.921163	0.925086	0.053272	0.230808	0.195504	0.039791	0.921799
0.914338	0.923214	0.055463	0.235506	0.195309	0.039559	0.918120

Tab. 6 Parameter Estimation Using Grid Search with Cross-Validation for Pastor Andrew Brunson Case

Grid scores on development set:		C	Gamma	Kernel
0.250	(+/-0.000)	1	2.00E-10	RBF
0.250	(+/-0.000)	1	2.00E-09	RBF
0.821	(+/-0.113)	1	2.00E-08	RBF
0.720	(+/-0.320)	1	2.00E-07	RBF
0.838	(+/-0.173)	1	2.00E-06	RBF
0.838	(+/-0.173)	1	2.00E-05	RBF
0.838	(+/-0.173)	1	0.0002	RBF
0.838	(+/-0.173)	1	0.002	RBF
0.838	(+/-0.173)	1	0.001	RBF
0.838	(+/-0.173)	1	0.01	RBF
0.852	(+/-0.156)	1	0.1	RBF
1.000	(+/-0.000)	1	20	RBF
0.938	(+/-0.110)	1	100	RBF
0.912	(+/-0.119)	1	200	RBF
0.679	(+/-0.430)	1	2000	RBF
0.250	(+/-0.000)	10	2.00E-10	RBF
0.250	(+/-0.000)	10	2.00E-09	RBF
0.700	(+/-0.253)	10	2.00E-08	RBF
0.720	(+/-0.320)	10	2.00E-07	RBF
0.838	(+/-0.173)	10	2.00E-06	RBF
0.838	(+/-0.173)	10	2.00E-05	RBF
0.838	(+/-0.173)	10	0.0002	RBF
0.838	(+/-0.173)	10	0.002	RBF
0.838	(+/-0.173)	10	0.001	RBF
0.838	(+/-0.173)	10	0.01	RBF
0.983	(+/-0.067)	10	0.1	RBF
1.000	(+/-0.000)	10	20	RBF
0.938	(+/-0.110)	10	100	RBF
0.912	(+/-0.119)	10	200	RBF
0.801	(+/-0.063)	10	2000	RBF
0.250	(+/-0.000)	100	2.00E-10	RBF
0.250	(+/-0.000)	100	2.00E-09	RBF
0.700	(+/-0.253)	100	2.00E-08	RBF
0.720	(+/-0.320)	100	2.00E-07	RBF
0.838	(+/-0.173)	100	2.00E-06	RBF
0.838	(+/-0.173)	100	2.00E-05	RBF
0.838	(+/-0.173)	100	0.0002	RBF
0.838	(+/-0.173)	100	0.002	RBF
0.838	(+/-0.173)	100	0.001	RBF
0.852	(+/-0.156)	100	0.01	RBF
1.000	(+/-0.000)	100	0.1	RBF

1.000	(+/-0.000)	100	20	RBF
0.938	(+/-0.110)	100	100	RBF
0.912	(+/-0.119)	100	200	RBF
0.801	(+/-0.063)	100	2000	RBF
0.250	(+/-0.000)	1000	2.00E-10	RBF
0.250	(+/-0.000)	1000	2.00E-09	RBF
0.700	(+/-0.253)	1000	2.00E-08	RBF
0.720	(+/-0.320)	1000	2.00E-07	RBF
0.838	(+/-0.173)	1000	2.00E-06	RBF
0.838	(+/-0.173)	1000	2.00E-05	RBF
0.838	(+/-0.173)	1000	0.0002	RBF
0.845	(+/-0.166)	1000	0.002	RBF
0.838	(+/-0.173)	1000	0.001	RBF
0.983	(+/-0.067)	1000	0.01	RBF
1.000	(+/-0.000)	1000	0.1	RBF
1.000	(+/-0.000)	1000	20	RBF
0.938	(+/-0.110)	1000	100	RBF
0.912	(+/-0.119)	1000	200	RBF
0.801	(+/-0.063)	1000	2000	RBF
0.250	(+/-0.000)	10000	2.00E-10	RBF
0.250	(+/-0.000)	10000	2.00E-09	RBF
0.700	(+/-0.253)	10000	2.00E-08	RBF
0.720	(+/-0.320)	10000	2.00E-07	RBF
0.838	(+/-0.173)	10000	2.00E-06	RBF
0.838	(+/-0.173)	10000	2.00E-05	RBF
0.838	(+/-0.173)	10000	0.0002	RBF
0.861	(+/-0.151)	10000	0.002	RBF
0.852	(+/-0.156)	10000	0.001	RBF
1.000	(+/-0.000)	10000	0.01	RBF
1.000	(+/-0.000)	10000	0.1	RBF
1.000	(+/-0.000)	10000	20	RBF
0.938	(+/-0.110)	10000	100	RBF
0.912	(+/-0.119)	10000	200	RBF
0.801	(+/-0.063)	10000	2000	RBF
0.250	(+/-0.000)	10000	2.00E-10	RBF
0.250	(+/-0.000)	10000	2.00E-09	RBF
0.700	(+/-0.253)	10000	2.00E-08	RBF
0.720	(+/-0.320)	10000	2.00E-07	RBF
0.838	(+/-0.173)	10000	2.00E-06	RBF
0.838	(+/-0.173)	10000	2.00E-05	RBF
0.838	(+/-0.173)	10000	0.0002	RBF
0.861	(+/-0.151)	10000	0.002	RBF
0.852	(+/-0.156)	10000	0.001	RBF
1.000	(+/-0.000)	10000	0.01	RBF
1.000	(+/-0.000)	10000	0.1	RBF
1.000	(+/-0.000)	10000	20	RBF

0.938	(+/-0.110)	10000	100	RBF
0.912	(+/-0.119)	10000	200	RBF
0.801	(+/-0.063)	10000	2000	RBF

Tab. 7 Accuracy of the Parameter Estimation Using Grid Search with Cross-Validation for Pastor Andrew Brunson Case

The model is trained on the full development set.
The scores are computed on the full evaluation set.

	precision	recall	f1-score	support
0	0.94	1.00	0.97	20
1	1.00	0.96	0.97	30
Accuracy			0.97	75
Macro Avg	0.97	0.97	0.97	75
Weighted Avg	0.97	0.97	0.97	75

Tab. 9. Averaged Monte Carlo Cross Validation Statistical Test Results for Pastor Andrew Brunson Case.

VMS	0.928399
EVS	0.935027
MSE	0.057164
RMSE	0.239008
MAE	0.230802
MAPE	0.050095
$\boldsymbol{R^2}$	0.931198

Tab. 8 Exhaustive Monte-Carlo Cross Validation Statistical Test Results for Pastor Andrew Brunson Case

VMS	EVS	MSE	RMSE	MAE	MAPE	R^2
0.932435	0.938555	0.054969	0.234454	0.226857	0.049327	0.934439
0.942626	0.937360	0.054675	0.233826	0.224992	0.048676	0.934045
0.951302	0.934990	0.058973	0.242843	0.237216	0.050949	0.931560
0.911351	0.933829	0.053070	0.230370	0.221495	0.048031	0.930601
0.928652	0.935948	0.059291	0.243498	0.236147	0.051380	0.932240
0.927821	0.938905	0.054790	0.234072	0.226507	0.049290	0.935322
0.928541	0.937389	0.054457	0.233360	0.225388	0.048804	0.934322
0.915010	0.928845	0.062932	0.250862	0.242023	0.052763	0.924887
0.918016	0.926500	0.062116	0.249232	0.239457	0.051846	0.922367
0.926497	0.937550	0.058488	0.241843	0.232941	0.050658	0.933068
0.922980	0.936697	0.053969	0.232313	0.224200	0.048355	0.933448
0.922260	0.939227	0.052710	0.229587	0.220150	0.048503	0.934840
0.928430	0.937382	0.053983	0.232342	0.223870	0.048965	0.934040
0.910631	0.936811	0.052369	0.228843	0.219315	0.047891	0.933104
0.918589	0.927675	0.062036	0.249070	0.239406	0.052478	0.923041
0.920278	0.939940	0.053747	0.231833	0.223491	0.048660	0.935428
0.927821	0.933280	0.059326	0.243570	0.236143	0.051477	0.929003
0.936554	0.934071	0.055492	0.235567	0.229180	0.048853	0.930605
0.933979	0.932295	0.059548	0.244025	0.236491	0.050995	0.928266
0.918016	0.940222	0.053637	0.231597	0.222763	0.048543	0.935168
0.928652	0.927804	0.063305	0.251605	0.242278	0.052544	0.924171
0.931276	0.939388	0.054727	0.233938	0.226683	0.049394	0.935485
0.923501	0.929319	0.063346	0.251686	0.243281	0.052629	0.926332
0.941091	0.937164	0.055448	0.235474	0.228776	0.049071	0.934265
0.942714	0.938590	0.055444	0.235466	0.228444	0.049289	0.934862
0.918016	0.932974	0.058543	0.241958	0.233761	0.050945	0.928437
0.919434	0.938671	0.053874	0.232109	0.223674	0.048796	0.934266
0.942714	0.933030	0.059988	0.244924	0.236652	0.051283	0.929793
0.949542	0.940562	0.055468	0.235517	0.228253	0.049588	0.937042
0.931169	0.932257	0.058918	0.242731	0.234544	0.050630	0.927349
0.942952	0.935522	0.057879	0.240580	0.233282	0.050179	0.931457
0.936652	0.938251	0.056883	0.238501	0.230443	0.050417	0.934472
0.925660	0.936008	0.056286	0.237247	0.228265	0.049465	0.932794
0.933259	0.933166	0.059280	0.243476	0.235388	0.051015	0.929674
0.935931	0.934071	0.059443	0.243809	0.236262	0.050944	0.930356
0.918016	0.933660	0.057259	0.239289	0.229815	0.050164	0.929685
0.917166	0.929750	0.056404	0.237495	0.227178	0.049815	0.925660
0.920154	0.932103	0.058585	0.242043	0.233080	0.050544	0.928088
0.914409	0.934758	0.055794	0.236208	0.227453	0.049443	0.931357
0.917296	0.936153	0.055629	0.235858	0.227100	0.049516	0.932036
0.952958	0.935445	0.060661	0.246295	0.240243	0.051770	0.931731

Tab. 10 Parameter Estimation Using Grid Search with Cross-Validation for 2018 Turkish Parliamentary and Presidential Elections

Grid scores on development set:		C	Gamma	Kernel
0.280	(+/-0.049)	0.1	0.1	RBF
0.645	(+/-0.369)	0.1	20	RBF
0.500	(+/-0.000)	0.1	100	RBF
0.500	(+/-0.000)	0.1	200	RBF
0.500	(+/-0.000)	0.1	2000	RBF
0.525	(+/-0.100)	1	0.1	RBF
1.000	(+/-0.000)	1	20	RBF
0.947	(+/-0.137)	1	100	RBF
0.833	(+/-0.284)	1	200	RBF
0.578	(+/-0.135)	1	2000	RBF
0.975	(+/-0.100)	10	0.1	RBF
1.000	(+/-0.000)	10	20	RBF
0.947	(+/-0.137)	10	100	RBF
0.850	(+/-0.310)	10	200	RBF
0.578	(+/-0.135)	10	2000	RBF
1.000	(+/-0.000)	100	0.1	RBF
1.000	(+/-0.000)	100	20	RBF
0.950	(+/-0.133)	100	100	RBF
0.896	(+/-0.196)	100	200	RBF
0.572	(+/-0.529)	100	2000	RBF
1.000	(+/-0.000)	1000	0.1	RBF
1.000	(+/-0.000)	1000	20	RBF
0.947	(+/-0.137)	1000	100	RBF
0.850	(+/-0.310)	1000	200	RBF
0.578	(+/-0.135)	1000	2000	RBF
1.000	(+/-0.000)	10000	0.1	RBF
1.000	(+/-0.000)	10000	20	RBF
0.947	(+/-0.137)	10000	100	RBF
0.850	(+/-0.310)	10000	200	RBF
0.578	(+/-0.135)	10000	2000	RBF
1.000	(+/-0.000)	10000	0.1	RBF
1.000	(+/-0.000)	10000	20	RBF
0.950	(+/-0.133)	10000	100	RBF
0.896	(+/-0.196)	10000	200	RBF
0.572	(+/-0.529)	10000	2000	RBF
1.000	(+/-0.000)	100000	0.1	RBF
1.000	(+/-0.000)	100000	20	RBF
0.947	(+/-0.137)	100000	100	RBF
0.850	(+/-0.310)	100000	200	RBF
0.578	(+/-0.135)	100000	2000	RBF

Tab. 11 Accuracy of the Parameter Estimation Using Grid Search with Cross-Validation for 2018 Turkish Parliamentary and Presidential Elections

The model is trained on the full development set.
The scores are computed on the full evaluation set.

	precision	recall	f1-score	support
0	0.92	1.00	0.95	30
1	1.00	0.88	0.93	20
Accuracy			0.94	50
Macro Avg	0.95	0.93	0.94	50
Weighted Avg	0.94	0.94	0.94	50

Tab. 13. Averaged Monte Carlo Cross Validation Statistical Test Results for 2018 Turkish Parliamentary and Presidential Elections.

VMS	0.905317
EVS	0.870601
MSE	0.011431
RMSE	0.106894
MAE	0.103208
MAPE	0.022936
R^2	0.862592

Tab. 12 Exhaustive Monte-Carlo Cross Validation Statistical Test Results for 2018 Turkish Parliamentary and Presidential Elections

VMS	EVS	MSE	RMSE	MAE	MAPE	R^2
0.915286	0.875708	0.011591	0.107661	0.104579	0.023305	0.869743
0.884266	0.869771	0.011114	0.105421	0.101826	0.022549	0.860421
0.906264	0.870173	0.011259	0.106110	0.102093	0.022684	0.863504
0.884266	0.871611	0.011503	0.107252	0.103715	0.022964	0.861095
0.876558	0.859328	0.011567	0.107552	0.103169	0.022840	0.850672
0.899117	0.851899	0.012275	0.110792	0.106406	0.023735	0.843539
0.899117	0.868049	0.011950	0.109314	0.105276	0.023354	0.859092
0.906264	0.869590	0.012092	0.109964	0.105800	0.023483	0.860834
0.908737	0.859879	0.012167	0.110302	0.106012	0.023602	0.853393
0.901623	0.861361	0.012296	0.110886	0.107151	0.023857	0.851703
0.906264	0.874975	0.011226	0.105951	0.102224	0.022727	0.867700
0.899117	0.868488	0.011186	0.105765	0.102153	0.022692	0.858220
0.913225	0.886545	0.010338	0.101678	0.098443	0.021854	0.878462
0.903779	0.872481	0.011281	0.106213	0.102966	0.022843	0.865713
0.894324	0.874887	0.010884	0.104324	0.100755	0.022319	0.865998
0.932997	0.874514	0.011442	0.106968	0.103561	0.023008	0.868502
0.906264	0.887794	0.010289	0.101433	0.098532	0.021863	0.879895
0.913225	0.874121	0.011499	0.107234	0.104616	0.023188	0.862299
0.935351	0.860164	0.012312	0.110961	0.107046	0.023912	0.853712
0.920001	0.867726	0.012154	0.110244	0.107388	0.023892	0.859035
0.906264	0.868379	0.011916	0.109162	0.105696	0.023522	0.861501
0.939213	0.861648	0.011749	0.108394	0.104315	0.023335	0.856228
0.899117	0.870211	0.011572	0.107572	0.103424	0.023059	0.860734
0.906264	0.875434	0.011288	0.106245	0.102800	0.022890	0.869215
0.913225	0.877908	0.011364	0.106603	0.103017	0.022903	0.870277
0.920001	0.871119	0.011543	0.107437	0.103891	0.023123	0.863367
0.913225	0.872982	0.011310	0.106348	0.103117	0.022912	0.865183
0.884266	0.871145	0.010901	0.104410	0.100371	0.022179	0.861156
0.913225	0.871584	0.011538	0.107413	0.104519	0.023190	0.864101
0.928975	0.871661	0.011642	0.107897	0.104126	0.023170	0.865184
0.886840	0.869346	0.011126	0.105481	0.101879	0.022597	0.861122
0.920001	0.871493	0.011881	0.108999	0.105099	0.023443	0.866240
0.899117	0.877480	0.010935	0.104570	0.100600	0.022315	0.867946
0.876558	0.871349	0.010776	0.103805	0.099814	0.022108	0.862917
0.906264	0.873352	0.011163	0.105655	0.101623	0.022607	0.866773
0.901623	0.873548	0.011160	0.105640	0.101694	0.022599	0.864224
0.906264	0.861282	0.011504	0.107256	0.103732	0.023072	0.852170
0.894324	0.872057	0.011155	0.105616	0.101590	0.022542	0.864343
0.891785	0.875527	0.011072	0.105222	0.101589	0.022509	0.866504
0.899117	0.866718	0.011480	0.107147	0.103271	0.022980	0.858126
0.906264	0.871363	0.011182	0.105746	0.101641	0.022639	0.865432

Tab. 14 Parameter Estimation Using Grid Search with Cross-Validation for S-400 Crises

Grid scores on development set:		C	Gamma	Kernel
0.500	(+/-0.000)	0.01	2.00E-10	RBF
0.500	(+/-0.000)	0.01	2.00E-09	RBF
0.567	(+/-0.267)	0.01	2.00E-08	RBF
0.567	(+/-0.267)	0.01	2.00E-07	RBF
0.567	(+/-0.267)	0.01	2.00E-06	RBF
0.567	(+/-0.267)	0.01	0.002	RBF
0.567	(+/-0.267)	0.01	0.001	RBF
0.567	(+/-0.267)	0.01	0.01	RBF
0.567	(+/-0.267)	0.01	0.1	RBF
0.600	(+/-0.400)	0.01	20	RBF
0.550	(+/-0.200)	0.01	100	RBF
0.550	(+/-0.200)	0.01	200	RBF
0.500	(+/-0.000)	0.01	2000	RBF
0.500	(+/-0.000)	0.1	2.00E-10	RBF
0.500	(+/-0.000)	0.1	2.00E-09	RBF
0.567	(+/-0.267)	0.1	2.00E-08	RBF
0.567	(+/-0.267)	0.1	2.00E-07	RBF
0.567	(+/-0.267)	0.1	2.00E-06	RBF
0.567	(+/-0.267)	0.1	0.002	RBF
0.567	(+/-0.267)	0.1	0.001	RBF
0.567	(+/-0.267)	0.1	0.01	RBF
0.567	(+/-0.267)	0.1	0.1	RBF
0.920	(+/-0.150)	0.1	20	RBF
0.550	(+/-0.200)	0.1	100	RBF
0.550	(+/-0.200)	0.1	200	RBF
0.500	(+/-0.000)	0.1	2000	RBF
0.500	(+/-0.000)	1	2.00E-10	RBF
0.500	(+/-0.000)	1	2.00E-09	RBF
0.567	(+/-0.267)	1	2.00E-08	RBF
0.567	(+/-0.267)	1	2.00E-07	RBF
0.567	(+/-0.267)	1	2.00E-06	RBF
0.567	(+/-0.267)	1	0.002	RBF
0.567	(+/-0.267)	1	0.001	RBF
0.567	(+/-0.267)	1	0.01	RBF
0.607	(+/-0.302)	1	0.1	RBF
1.000	(+/-0.000)	1	20	RBF
0.935	(+/-0.108)	1	100	RBF
0.915	(+/-0.087)	1	200	RBF
0.600	(+/-0.219)	1	2000	RBF
0.500	(+/-0.000)	10	2.00E-10	RBF
0.500	(+/-0.000)	10	2.00E-09	RBF

0.567	(+/-0.267)	10	2.00E-08	RBF
0.567	(+/-0.267)	10	2.00E-07	RBF
0.567	(+/-0.267)	10	2.00E-06	RBF
0.567	(+/-0.267)	10	0.002	RBF
0.567	(+/-0.267)	10	0.001	RBF
0.607	(+/-0.302)	10	0.01	RBF
1.000	(+/-0.000)	10	0.1	RBF
1.000	(+/-0.000)	10	20	RBF
0.960	(+/-0.098)	10	100	RBF
0.915	(+/-0.087)	10	200	RBF
0.625	(+/-0.195)	10	2000	RBF
0.500	(+/-0.000)	100	2.00E-10	RBF
0.500	(+/-0.000)	100	2.00E-09	RBF
0.567	(+/-0.267)	100	2.00E-08	RBF
0.567	(+/-0.267)	100	2.00E-07	RBF
0.567	(+/-0.267)	100	2.00E-06	RBF
0.767	(+/-0.211)	100	0.002	RBF
0.587	(+/-0.288)	100	0.001	RBF
0.727	(+/-0.165)	100	0.01	RBF
1.000	(+/-0.000)	100	0.1	RBF
1.000	(+/-0.000)	100	20	RBF
0.960	(+/-0.098)	100	100	RBF
0.915	(+/-0.087)	100	200	RBF
0.625	(+/-0.195)	100	2000	RBF
0.500	(+/-0.000)	1000	2.00E-10	RBF
0.500	(+/-0.000)	1000	2.00E-09	RBF
0.567	(+/-0.267)	1000	2.00E-08	RBF
0.567	(+/-0.267)	1000	2.00E-07	RBF
0.567	(+/-0.267)	1000	2.00E-06	RBF
0.727	(+/-0.165)	1000	0.002	RBF
0.727	(+/-0.165)	1000	0.001	RBF
0.980	(+/-0.080)	1000	0.01	RBF
1.000	(+/-0.000)	1000	0.1	RBF
1.000	(+/-0.000)	1000	20	RBF
0.960	(+/-0.098)	1000	100	RBF
0.915	(+/-0.087)	1000	200	RBF
0.625	(+/-0.195)	1000	2000	RBF
0.500	(+/-0.000)	10000	2.00E-10	RBF
0.500	(+/-0.000)	10000	2.00E-09	RBF
0.567	(+/-0.267)	10000	2.00E-08	RBF
0.567	(+/-0.267)	10000	2.00E-07	RBF
0.567	(+/-0.267)	10000	2.00E-06	RBF
0.803	(+/-0.187)	10000	0.002	RBF
0.727	(+/-0.165)	10000	0.001	RBF
1.000	(+/-0.000)	10000	0.01	RBF
1.000	(+/-0.000)	10000	0.1	RBF

1.000	(+/-0.000)	10000	20	RBF
0.960	(+/-0.098)	10000	100	RBF
0.915	(+/-0.087)	10000	200	RBF
0.625	(+/-0.195)	10000	2000	RBF
0.500	(+/-0.000)	10000	2.00E-10	RBF
0.500	(+/-0.000)	10000	2.00E-09	RBF
0.567	(+/-0.267)	10000	2.00E-08	RBF
0.567	(+/-0.267)	10000	2.00E-07	RBF
0.567	(+/-0.267)	10000	2.00E-06	RBF
0.803	(+/-0.187)	10000	0.002	RBF
0.727	(+/-0.165)	10000	0.001	RBF
1.000	(+/-0.000)	10000	0.01	RBF
1.000	(+/-0.000)	10000	0.1	RBF
1.000	(+/-0.000)	10000	20	RBF
0.960	(+/-0.098)	10000	100	RBF
0.915	(+/-0.087)	10000	200	RBF
0.625	(+/-0.195)	10000	2000	RBF
0.500	(+/-0.000)	100000	2.00E-10	RBF
0.500	(+/-0.000)	100000	2.00E-09	RBF
0.567	(+/-0.267)	100000	2.00E-08	RBF
0.567	(+/-0.267)	100000	2.00E-07	RBF
0.767	(+/-0.211)	100000	2.00E-06	RBF
1.000	(+/-0.000)	100000	0.002	RBF
0.980	(+/-0.080)	100000	0.001	RBF
1.000	(+/-0.000)	100000	0.01	RBF
1.000	(+/-0.000)	100000	0.1	RBF
1.000	(+/-0.000)	100000	20	RBF
0.960	(+/-0.098)	100000	100	RBF
0.915	(+/-0.087)	100000	200	RBF
0.625	(+/-0.195)	100000	2000	RBF

Tab. 15 Accuracy of the Parameter Estimation Using Grid Search with Cross-Validation for S-400 Crises

The model is trained on the full development set.
The scores are computed on the full evaluation set.

	precision	recall	f1-score	support
0	0.95	1.00	0.98	20
1	1.00	0.97	0.98	30
Accuracy			0.98	50
Macro Avg	0.98	0.98	0.98	50
Weighted Avg	0.98	0.98	0.98	50

Tab. 17. Averaged Monte Carlo Cross Validation Statistical Test Results for S-400 Crisis.

VMS	0.885410
EVS	0.905757
MSE	0.005274
RMSE	0.072623
MAE	0.069744
MAPE	0.012321
$\boldsymbol{R^2}$	0.900212

Tab. 16 Exhaustive Monte-Carlo Cross Validation Statistical Test Results for S-400 Crisis

VMS	EVS	MSE	RMSE	MAE	MAPE	R^2
0.886283	0.905519	0.005301	0.072810	0.069844	0.012333	0.900960
0.883238	0.905032	0.005248	0.072440	0.069296	0.012241	0.899034
0.872316	0.904065	0.005256	0.072502	0.069586	0.012293	0.896677
0.882563	0.906755	0.005262	0.072540	0.069579	0.012287	0.901641
0.875471	0.904151	0.005193	0.072066	0.068857	0.012156	0.898779
0.897507	0.903103	0.005403	0.073503	0.071077	0.012570	0.896915
0.883913	0.900521	0.005271	0.072601	0.069556	0.012285	0.895562
0.886959	0.904716	0.005303	0.072820	0.069979	0.012363	0.899094
0.888456	0.906842	0.005186	0.072013	0.069310	0.012257	0.901089
0.881021	0.910151	0.005235	0.072357	0.069457	0.012260	0.904154
0.885782	0.907881	0.005283	0.072683	0.069801	0.012333	0.903307
0.885608	0.909293	0.005297	0.072783	0.069861	0.012334	0.903855
0.880164	0.905213	0.005255	0.072492	0.069502	0.012286	0.899738
0.894415	0.901620	0.005359	0.073202	0.070483	0.012463	0.896174
0.890774	0.901538	0.005326	0.072982	0.070250	0.012419	0.894935
0.890144	0.904804	0.005328	0.072994	0.070140	0.012385	0.898473
0.892290	0.905728	0.005350	0.073145	0.070358	0.012434	0.899360
0.887131	0.908652	0.005323	0.072961	0.070455	0.012438	0.903045
0.881021	0.904424	0.005225	0.072283	0.069306	0.012249	0.898537
0.891614	0.906941	0.005309	0.072859	0.070199	0.012404	0.902048
0.880839	0.907399	0.005208	0.072164	0.069134	0.012215	0.901839
0.890939	0.903461	0.005283	0.072687	0.069933	0.012370	0.898291
0.874604	0.905423	0.005218	0.072233	0.069259	0.012237	0.899155
0.895928	0.907740	0.005354	0.073169	0.070393	0.012442	0.902353
0.878784	0.906035	0.005221	0.072259	0.069419	0.012245	0.900884
0.863561	0.904964	0.005106	0.071458	0.068113	0.012017	0.899033
0.870969	0.905468	0.005165	0.071868	0.068555	0.012094	0.899002
0.894575	0.906722	0.005315	0.072903	0.070014	0.012371	0.901304
0.883238	0.906298	0.005237	0.072365	0.069190	0.012212	0.900383
0.872316	0.905977	0.005176	0.071941	0.068853	0.012148	0.901077
0.876337	0.905785	0.005196	0.072083	0.068915	0.012167	0.901384
0.889977	0.904101	0.005274	0.072619	0.069692	0.012318	0.898903
0.893967	0.907986	0.005308	0.072854	0.070173	0.012410	0.903744
0.880164	0.905836	0.005252	0.072471	0.069627	0.012305	0.900522
0.893129	0.905334	0.005334	0.073032	0.070409	0.012460	0.899705
0.896832	0.908340	0.005342	0.073090	0.070432	0.012461	0.903739
0.894575	0.903864	0.005387	0.073398	0.070899	0.012530	0.898103
0.894415	0.903435	0.005330	0.073009	0.070474	0.012466	0.897246
0.881877	0.907662	0.005226	0.072289	0.069207	0.012211	0.902454
0.887131	0.908482	0.005281	0.072670	0.069778	0.012326	0.902437
0.890986	0.908764	0.005323	0.072957	0.070149	0.012386	0.903763

Tab. 18 Parameter Estimation Using Grid Search with Cross-Validation for 2019 Istanbul Mayoral Elections

Grid scores on development set:		C	Gamma	Kernel
0.270	(+/-0.049)	1	2.00E-10	RBF
0.270	(+/-0.049)	1	2.00E-09	RBF
0.377	(+/-0.458)	1	2.00E-08	RBF
0.500	(+/-0.000)	1	2.00E-07	RBF
0.500	(+/-0.000)	1	2.00E-06	RBF
0.572	(+/-0.529)	1	0.002	RBF
0.500	(+/-0.000)	1	0.001	RBF
0.876	(+/-0.168)	1	0.01	RBF
0.935	(+/-0.108)	1	0.1	RBF
1.000	(+/-0.000)	1	20	RBF
0.986	(+/-0.057)	1	100	RBF
0.935	(+/-0.108)	1	200	RBF
0.585	(+/-0.087)	1	2000	RBF
0.270	(+/-0.049)	10	2.00E-10	RBF
0.270	(+/-0.049)	10	2.00E-09	RBF
0.377	(+/-0.458)	10	2.00E-08	RBF
0.500	(+/-0.000)	10	2.00E-07	RBF
0.500	(+/-0.000)	10	2.00E-06	RBF
0.572	(+/-0.529)	10	0.002	RBF
0.500	(+/-0.000)	10	0.001	RBF
0.876	(+/-0.168)	10	0.01	RBF
0.935	(+/-0.108)	10	0.1	RBF
1.000	(+/-0.000)	10	20	RBF
0.986	(+/-0.057)	10	100	RBF
0.935	(+/-0.108)	10	200	RBF
0.585	(+/-0.087)	10	2000	RBF
0.270	(+/-0.049)	100	2.00E-10	RBF
0.377	(+/-0.458)	100	2.00E-09	RBF
0.500	(+/-0.000)	100	2.00E-08	RBF
0.500	(+/-0.000)	100	2.00E-07	RBF
0.500	(+/-0.000)	100	2.00E-06	RBF
0.517	(+/-0.267)	100	0.002	RBF
0.500	(+/-0.000)	100	0.001	RBF
0.717	(+/-0.235)	100	0.01	RBF
1.000	(+/-0.000)	100	0.1	RBF
1.000	(+/-0.000)	100	20	RBF
0.986	(+/-0.057)	100	100	RBF
0.935	(+/-0.108)	100	200	RBF
0.585	(+/-0.087)	100	2000	RBF
0.500	(+/-0.000)	1000	2.00E-10	RBF
0.500	(+/-0.000)	1000	2.00E-09	RBF

0.572	(+/-0.529)	1000	2.00E-08	RBF
0.500	(+/-0.000)	1000	2.00E-07	RBF
0.500	(+/-0.000)	1000	2.00E-06	RBF
0.717	(+/-0.235)	1000	0.002	RBF
0.720	(+/-0.258)	1000	0.001	RBF
1.000	(+/-0.000)	1000	0.01	RBF
1.000	(+/-0.000)	1000	0.1	RBF
0.986	(+/-0.057)	1000	20	RBF
0.935	(+/-0.108)	1000	100	RBF
0.585	(+/-0.087)	1000	200	RBF
0.630	(+/-0.174)	1000	2000	RBF
0.270	(+/-0.049)	10000	2.00E-10	RBF
0.270	(+/-0.049)	10000	2.00E-09	RBF
0.377	(+/-0.458)	10000	2.00E-08	RBF
0.500	(+/-0.000)	10000	2.00E-07	RBF
0.500	(+/-0.000)	10000	2.00E-06	RBF
0.873	(+/-0.142)	10000	0.002	RBF
0.717	(+/-0.235)	10000	0.001	RBF
1.000	(+/-0.000)	10000	0.01	RBF
1.000	(+/-0.000)	10000	0.1	RBF
0.986	(+/-0.057)	10000	20	RBF
0.935	(+/-0.108)	10000	100	RBF
0.935	(+/-0.108)	10000	200	RBF
0.630	(+/-0.174)	10000	2000	RBF
0.270	(+/-0.049)	10000	2.00E-10	RBF
0.270	(+/-0.049)	10000	2.00E-09	RBF
0.377	(+/-0.458)	10000	2.00E-08	RBF
0.500	(+/-0.000)	10000	2.00E-07	RBF
0.500	(+/-0.000)	10000	2.00E-06	RBF
0.873	(+/-0.142)	10000	0.002	RBF
0.717	(+/-0.235)	10000	0.001	RBF
1.000	(+/-0.000)	10000	0.01	RBF
1.000	(+/-0.000)	10000	0.1	RBF
0.986	(+/-0.057)	10000	20	RBF
0.935	(+/-0.108)	10000	100	RBF
0.935	(+/-0.108)	10000	200	RBF
0.630	(+/-0.174)	10000	2000	RBF
0.270	(+/-0.049)	100000	2.00E-10	RBF
0.377	(+/-0.458)	100000	2.00E-09	RBF
0.500	(+/-0.000)	100000	2.00E-08	RBF
0.500	(+/-0.000)	100000	2.00E-07	RBF
0.517	(+/-0.267)	100000	2.00E-06	RBF
0.572	(+/-0.529)	100000	0.002	RBF
0.717	(+/-0.235)	100000	0.001	RBF
0.876	(+/-0.168)	100000	0.01	RBF
1.000	(+/-0.000)	100000	0.1	RBF

0.986	(+/-0.057)	100000	20	RBF
0.935	(+/-0.108)	100000	100	RBF
0.935	(+/-0.108)	100000	200	RBF
0.630	(+/-0.174)	100000	2000	RBF

Tab. 19 Accuracy of the Parameter Estimation Using Grid Search with Cross-Validation for 2019 Istanbul Mayoral Elections

The model is trained on the full development set.
The scores are computed on the full evaluation set.

	precision	recall	f1-score	support
0	0.94	1.00	0.97	30
1	1.00	0.90	0.95	20
Accuracy			0.96	50
Macro Avg	0.97	0.95	0.96	50
Weighted Avg	0.96	0.96	0.96	50

Tab. 21 Averaged Monte Carlo Cross Validation Statistical Test Results for 2019 Istanbul Mayoral Elections.

VMS	0.765247
EVS	0.759593
MSE	0.009110
RMSE	0.095443
MAE	0.084592
MAPE	0.014609
$\boldsymbol{R^2}$	0.759524

Tab. 20 Exhaustive Monte-Carlo Cross Validation Statistical Test Results for 2019 Istanbul Mayoral Elections

VMS	EVS	MSE	RMSE	MAE	MAPE	R^2
0.744692	0.777927	0.008778	0.093690	0.082043	0.014187	0.777775
0.771957	0.764627	0.009138	0.095594	0.084913	0.014654	0.764577
0.773206	0.756277	0.009195	0.095891	0.084847	0.014654	0.756175
0.752765	0.754362	0.008965	0.094683	0.083613	0.014430	0.754351
0.771957	0.759113	0.009219	0.096018	0.085290	0.014736	0.759104
0.752765	0.762289	0.008918	0.094436	0.083283	0.014391	0.762267
0.752765	0.759611	0.008961	0.094663	0.083579	0.014426	0.759609
0.771957	0.764696	0.009110	0.095448	0.084333	0.014559	0.764692
0.762462	0.761183	0.009007	0.094906	0.083733	0.014473	0.761161
0.762462	0.752976	0.009119	0.095493	0.084845	0.014650	0.752923
0.762462	0.763287	0.009122	0.095508	0.084856	0.014652	0.763225
0.762462	0.761350	0.009105	0.095420	0.084816	0.014648	0.761342
0.762462	0.761475	0.009109	0.095440	0.084742	0.014635	0.761463
0.781254	0.760082	0.009378	0.096839	0.086195	0.014886	0.759977
0.752765	0.759140	0.008907	0.094377	0.083205	0.014379	0.759139
0.752765	0.758992	0.009070	0.095236	0.084415	0.014594	0.758920
0.771957	0.755385	0.009202	0.095925	0.085109	0.014706	0.755370
0.771957	0.737672	0.009266	0.096258	0.086004	0.014822	0.737187
0.762462	0.762699	0.009022	0.094984	0.083819	0.014479	0.762678
0.762462	0.763062	0.009083	0.095304	0.084341	0.014566	0.763037
0.771957	0.762574	0.009231	0.096077	0.085493	0.014767	0.762509
0.771957	0.754911	0.009225	0.096049	0.085429	0.014764	0.754894
0.762462	0.762196	0.009081	0.095296	0.084539	0.014613	0.762028
0.771957	0.757241	0.009243	0.096141	0.085575	0.014786	0.757240
0.771957	0.761165	0.009137	0.095586	0.084743	0.014619	0.761149
0.752765	0.752371	0.009231	0.096079	0.085583	0.014791	0.752370
0.771957	0.765010	0.009105	0.095419	0.084497	0.014592	0.765001
0.771957	0.762604	0.009213	0.095984	0.085110	0.014693	0.762595
0.771957	0.758938	0.009251	0.096183	0.085791	0.014823	0.758872
0.771957	0.758749	0.009205	0.095942	0.085213	0.014723	0.758749
0.752765	0.766010	0.008919	0.094442	0.083292	0.014402	0.765610
0.771957	0.762808	0.009188	0.095855	0.084902	0.014665	0.762793
0.752765	0.763753	0.008917	0.094429	0.083288	0.014391	0.763710
0.771957	0.757415	0.009102	0.095406	0.084415	0.014584	0.757361
0.762462	0.761132	0.009068	0.095226	0.084129	0.014527	0.760932
0.771957	0.753722	0.009166	0.095737	0.085059	0.014680	0.753435
0.752765	0.752795	0.008927	0.094481	0.083338	0.014398	0.752734
0.762462	0.759041	0.008995	0.094840	0.083827	0.014464	0.758964
0.781254	0.754381	0.009329	0.096589	0.086387	0.014906	0.754277
0.771957	0.759709	0.009151	0.095662	0.084875	0.014640	0.759707
0.771957	0.760577	0.009146	0.095635	0.084807	0.014631	0.760576

Tab. 22 Parameter Estimation Using Grid Search with Cross-Validation for Political News

Grid scores on development set:		C	Gamma	Kernel
0.129	(+/-0.018)	1	2.00E-10	RBF
0.129	(+/-0.018)	1	2.00E-09	RBF
0.129	(+/-0.018)	1	2.00E-08	RBF
0.129	(+/-0.018)	1	2.00E-07	RBF
0.129	(+/-0.018)	1	2.00E-06	RBF
0.129	(+/-0.018)	1	2.00E-05	RBF
0.129	(+/-0.018)	1	0.0002	RBF
0.494	(+/-0.011)	1	0.002	RBF
0.129	(+/-0.018)	1	0.001	RBF
0.910	(+/-0.154)	1	0.01	RBF
0.948	(+/-0.103)	1	0.1	RBF
0.859	(+/-0.069)	1	20	RBF
0.197	(+/-0.279)	1	100	RBF
0.129	(+/-0.018)	1	200	RBF
0.129	(+/-0.018)	1	2000	RBF
0.129	(+/-0.018)	10	2.00E-10	RBF
0.129	(+/-0.018)	10	2.00E-09	RBF
0.129	(+/-0.018)	10	2.00E-08	RBF
0.129	(+/-0.018)	10	2.00E-07	RBF
0.129	(+/-0.018)	10	2.00E-06	RBF
0.129	(+/-0.018)	10	2.00E-05	RBF
0.494	(+/-0.011)	10	0.0002	RBF
0.920	(+/-0.113)	10	0.002	RBF
0.930	(+/-0.101)	10	0.001	RBF
0.966	(+/-0.057)	10	0.01	RBF
0.979	(+/-0.051)	10	0.1	RBF
0.870	(+/-0.077)	10	20	RBF
0.266	(+/-0.344)	10	100	RBF
0.129	(+/-0.018)	10	200	RBF
0.129	(+/-0.018)	10	2000	RBF
0.129	(+/-0.018)	100	2.00E-10	RBF
0.129	(+/-0.018)	100	2.00E-09	RBF
0.129	(+/-0.018)	100	2.00E-08	RBF
0.129	(+/-0.018)	100	2.00E-07	RBF
0.129	(+/-0.018)	100	2.00E-06	RBF
0.494	(+/-0.011)	100	2.00E-05	RBF
0.920	(+/-0.113)	100	0.0002	RBF
0.979	(+/-0.051)	100	0.002	RBF
0.966	(+/-0.057)	100	0.001	RBF
0.979	(+/-0.051)	100	0.01	RBF
0.945	(+/-0.063)	100	0.1	RBF

0.870	(+/-0.077)	100	20	RBF
0.266	(+/-0.344)	100	100	RBF
0.129	(+/-0.018)	100	200	RBF
0.129	(+/-0.018)	100	2000	RBF
0.129	(+/-0.018)	1000	2.00E-10	RBF
0.129	(+/-0.018)	1000	2.00E-09	RBF
0.129	(+/-0.018)	1000	2.00E-08	RBF
0.129	(+/-0.018)	1000	2.00E-07	RBF
0.494	(+/-0.011)	1000	2.00E-06	RBF
0.920	(+/-0.113)	1000	2.00E-05	RBF
0.979	(+/-0.051)	1000	0.0002	RBF
0.979	(+/-0.051)	1000	0.002	RBF
0.979	(+/-0.051)	1000	0.001	RBF
0.934	(+/-0.097)	1000	0.01	RBF
0.945	(+/-0.063)	1000	0.1	RBF
0.870	(+/-0.077)	1000	20	RBF
0.266	(+/-0.344)	1000	100	RBF
0.129	(+/-0.018)	1000	200	RBF
0.129	(+/-0.018)	1000	2000	RBF
0.129	(+/-0.018)	10000	2.00E-10	RBF
0.129	(+/-0.018)	10000	2.00E-09	RBF
0.129	(+/-0.018)	10000	2.00E-08	RBF
0.494	(+/-0.011)	10000	2.00E-07	RBF
0.920	(+/-0.113)	10000	2.00E-06	RBF
0.979	(+/-0.051)	10000	2.00E-05	RBF
0.979	(+/-0.051)	10000	0.0002	RBF
0.947	(+/-0.110)	10000	0.002	RBF
0.934	(+/-0.097)	10000	0.001	RBF
0.933	(+/-0.164)	10000	0.01	RBF
0.945	(+/-0.063)	10000	0.1	RBF
0.870	(+/-0.077)	10000	20	RBF
0.266	(+/-0.344)	10000	100	RBF
0.129	(+/-0.018)	10000	200	RBF
0.129	(+/-0.018)	10000	2000	RBF

Tab. 23 Accuracy of the Parameter Estimation Using Grid Search with Cross-Validation for Political News

The model is trained on the full development set.
The scores are computed on the full evaluation set.

	precision	recall	f1-score	support
0	1	1	1	21
1	0.91	0.97	0.94	30
2	0.95	0.96	0.95	24
Accuracy			0.95	75
Macro Avg	0.98	0.98	0.98	75
Weighted Avg	0.97	0.97	0.97	75

Tab. 25 Averaged Monte Carlo Cross Validation Statistical Test Results for Political News.

VMS	0.947610
EVS	0.945922
MSE	0.036425
RMSE	0.190830
MAE	0.188578
MAPE	0.037097
$\boldsymbol{R}^2$	0.939205

Tab. 24 Exhaustive Monte-Carlo Cross Validation Statistical Test Results for Political News

VMS	EVS	MSE	RMSE	MAE	MAPE	R^2
0.944990	0.944191	0.036894	0.192079	0.189555	0.037208	0.937024
0.949575	0.944610	0.037416	0.193431	0.191622	0.037837	0.937939
0.939724	0.943863	0.036394	0.190772	0.187833	0.036810	0.936696
0.954753	0.944732	0.037718	0.194211	0.192698	0.038097	0.938820
0.955493	0.945572	0.037470	0.193572	0.191811	0.037877	0.939465
0.939448	0.947069	0.036295	0.190513	0.187552	0.036799	0.939738
0.942539	0.944726	0.036446	0.190908	0.187810	0.036789	0.937325
0.951324	0.945746	0.037449	0.193517	0.191758	0.037897	0.939657
0.943803	0.944499	0.036635	0.191402	0.188818	0.037097	0.938174
0.944031	0.944925	0.037137	0.192709	0.190695	0.037553	0.937810
0.949276	0.944624	0.037203	0.192882	0.190713	0.037527	0.937634
0.936550	0.946825	0.036137	0.190098	0.186782	0.036593	0.939516
0.948707	0.945323	0.037035	0.192445	0.190357	0.037427	0.938301
0.945622	0.944437	0.037210	0.192899	0.190825	0.037601	0.937745
0.943732	0.945298	0.037101	0.192615	0.190634	0.037532	0.938446
0.955940	0.945113	0.037790	0.194396	0.192735	0.038060	0.938306
0.950963	0.950700	0.033731	0.183661	0.181676	0.035819	0.945243
0.950974	0.948738	0.035458	0.188303	0.186625	0.036668	0.942116
0.944436	0.947420	0.035151	0.187485	0.185349	0.036459	0.940717
0.950468	0.947157	0.036231	0.190345	0.188131	0.036916	0.940244
0.950767	0.945639	0.036982	0.192307	0.189992	0.037352	0.938991
0.959746	0.946529	0.036066	0.189911	0.188599	0.037308	0.940544
0.943768	0.947684	0.034961	0.186979	0.184638	0.036296	0.941400
0.944735	0.950067	0.034579	0.185955	0.183020	0.035882	0.943290
0.946571	0.946254	0.036115	0.190041	0.187497	0.036891	0.940483
0.940023	0.943765	0.036561	0.191208	0.188453	0.037009	0.936669
0.943472	0.947446	0.036295	0.190513	0.188090	0.036984	0.940829
0.945871	0.944893	0.036677	0.191514	0.189229	0.037197	0.937862
0.949775	0.943271	0.037185	0.192835	0.190840	0.037509	0.936371
0.955995	0.944104	0.037603	0.193914	0.192252	0.037934	0.936980
0.952831	0.943964	0.037571	0.193832	0.192140	0.037952	0.937571
0.947983	0.944730	0.037029	0.192429	0.190113	0.037346	0.936866
0.947412	0.946051	0.036803	0.191842	0.189663	0.037292	0.938739
0.957797	0.946903	0.037321	0.193187	0.191510	0.037770	0.941144
0.936930	0.945475	0.035765	0.189116	0.185912	0.036445	0.937633
0.949839	0.945215	0.036667	0.191487	0.189138	0.037151	0.938304
0.954866	0.945052	0.037138	0.192712	0.190806	0.037708	0.939567
0.948581	0.944955	0.036526	0.191118	0.188665	0.036997	0.938358
0.943768	0.945422	0.036304	0.190537	0.187851	0.036888	0.938227
0.947379	0.949628	0.033558	0.183188	0.181089	0.035628	0.943084
0.941565	0.950179	0.032818	0.181158	0.178203	0.034890	0.943574

Tab. 26 Parameter Estimation Using Grid Search with Cross-Validation for Organic News Feed Simulation.

Grid scores on development set:		C	Gamma	Kernel
0.129	(+/-0.018)	1	2.00E-08	RBF
0.129	(+/-0.018)	1	2.00E-07	RBF
0.129	(+/-0.018)	1	2.00E-06	RBF
0.129	(+/-0.018)	1	2.00E-05	RBF
0.129	(+/-0.018)	1	0.0002	RBF
0.494	(+/-0.011)	1	0.002	RBF
0.129	(+/-0.018)	1	0.001	RBF
0.91	(+/-0.154)	1	0.01	RBF
0.948	(+/-0.103)	1	0.1	RBF
0.968	(+/-0.052)	1	1	RBF
0.916	(+/-0.127)	1	10	RBF
0.859	(+/-0.069)	1	20	RBF
0.197	(+/-0.279)	1	100	RBF
0.129	(+/-0.018)	1	200	RBF
0.129	(+/-0.018)	1	1000	RBF
0.129	(+/-0.018)	1	2000	RBF
0.129	(+/-0.018)	10	2.00E-08	RBF
0.129	(+/-0.018)	10	2.00E-07	RBF
0.129	(+/-0.018)	10	2.00E-06	RBF
0.129	(+/-0.018)	10	2.00E-05	RBF
0.494	(+/-0.011)	10	0.0002	RBF
0.92	(+/-0.113)	10	0.002	RBF
0.93	(+/-0.101)	10	0.001	RBF
0.966	(+/-0.057)	10	0.01	RBF
0.979	(+/-0.051)	10	0.1	RBF
0.955	(+/-0.047)	10	1	RBF
0.921	(+/-0.119)	10	10	RBF
0.87	(+/-0.077)	10	20	RBF
0.266	(+/-0.344)	10	100	RBF
0.129	(+/-0.018)	10	200	RBF
0.129	(+/-0.018)	10	1000	RBF
0.129	(+/-0.018)	10	2000	RBF
0.129	(+/-0.018)	100	2.00E-08	RBF
0.129	(+/-0.018)	100	2.00E-07	RBF
0.129	(+/-0.018)	100	2.00E-06	RBF
0.494	(+/-0.011)	100	2.00E-05	RBF
0.92	(+/-0.113)	100	0.0002	RBF
0.979	(+/-0.051)	100	0.002	RBF
0.966	(+/-0.057)	100	0.001	RBF
0.979	(+/-0.051)	100	0.01	RBF
0.945	(+/-0.063)	100	0.1	RBF

0.955	(+/-0.047)	100	1	RBF
0.921	(+/-0.119)	100	10	RBF
0.87	(+/-0.077)	100	20	RBF
0.266	(+/-0.344)	100	100	RBF
0.129	(+/-0.018)	100	200	RBF
0.129	(+/-0.018)	100	1000	RBF
0.129	(+/-0.018)	100	2000	RBF
0.129	(+/-0.018)	1000	2.00E-08	RBF
0.129	(+/-0.018)	1000	2.00E-07	RBF
0.494	(+/-0.011)	1000	2.00E-06	RBF
0.92	(+/-0.113)	1000	2.00E-05	RBF
0.979	(+/-0.051)	1000	0.0002	RBF
0.979	(+/-0.051)	1000	0.002	RBF
0.979	(+/-0.051)	1000	0.001	RBF
0.934	(+/-0.097)	1000	0.01	RBF
0.945	(+/-0.063)	1000	0.1	RBF
0.955	(+/-0.047)	1000	1	RBF
0.921	(+/-0.119)	1000	10	RBF
0.87	(+/-0.077)	1000	20	RBF
0.266	(+/-0.344)	1000	100	RBF
0.129	(+/-0.018)	1000	200	RBF
0.129	(+/-0.018)	1000	1000	RBF
0.129	(+/-0.018)	1000	2000	RBF
0.129	(+/-0.018)	10000	2.00E-08	RBF
0.494	(+/-0.011)	10000	2.00E-07	RBF
0.92	(+/-0.113)	10000	2.00E-06	RBF
0.979	(+/-0.051)	10000	2.00E-05	RBF
0.979	(+/-0.051)	10000	0.0002	RBF
0.947	(+/-0.110)	10000	0.002	RBF
0.934	(+/-0.097)	10000	0.001	RBF
0.933	(+/-0.164)	10000	0.01	RBF
0.945	(+/-0.063)	10000	0.1	RBF
0.955	(+/-0.047)	10000	1	RBF
0.921	(+/-0.119)	10000	10	RBF
0.87	(+/-0.077)	10000	20	RBF
0.266	(+/-0.344)	10000	100	RBF
0.129	(+/-0.018)	10000	200	RBF
0.129	(+/-0.018)	10000	1000	RBF
0.129	(+/-0.018)	10000	2000	RBF
0.494	(+/-0.011)	100000	2.00E-08	RBF
0.92	(+/-0.113)	100000	2.00E-07	RBF
0.979	(+/-0.051)	100000	2.00E-06	RBF
0.979	(+/-0.051)	100000	2.00E-05	RBF
0.964	(+/-0.061)	100000	0.0002	RBF
0.947	(+/-0.110)	100000	0.002	RBF
0.964	(+/-0.061)	100000	0.001	RBF

0.933	(+/-0.164)	100000	0.01	RBF
0.945	(+/-0.063)	100000	0.1	RBF
0.955	(+/-0.047)	100000	1	RBF
0.921	(+/-0.119)	100000	10	RBF
0.87	(+/-0.077)	100000	20	RBF
0.266	(+/-0.344)	100000	100	RBF
0.129	(+/-0.018)	100000	200	RBF
0.129	(+/-0.018)	100000	1000	RBF
0.129	(+/-0.018)	100000	2000	RBF
0.129	(+/-0.018)	1	2.00E-08	Polynomial
0.129	(+/-0.018)	1	2.00E-07	Polynomial
0.129	(+/-0.018)	1	2.00E-06	Polynomial
0.129	(+/-0.018)	1	2.00E-05	Polynomial
0.129	(+/-0.018)	1	0.0002	Polynomial
0.129	(+/-0.018)	1	0.002	Polynomial
0.129	(+/-0.018)	1	0.001	Polynomial
0.129	(+/-0.018)	1	0.01	Polynomial
0.734	(+/-0.389)	1	0.1	Polynomial
0.934	(+/-0.097)	1	1	Polynomial
0.945	(+/-0.063)	1	10	Polynomial
0.945	(+/-0.063)	1	20	Polynomial
0.945	(+/-0.063)	1	100	Polynomial
0.945	(+/-0.063)	1	200	Polynomial
0.945	(+/-0.063)	1	1000	Polynomial
0.945	(+/-0.063)	1	2000	Polynomial
0.129	(+/-0.018)	10	2.00E-08	Polynomial
0.129	(+/-0.018)	10	2.00E-07	Polynomial
0.129	(+/-0.018)	10	2.00E-06	Polynomial
0.129	(+/-0.018)	10	2.00E-05	Polynomial
0.129	(+/-0.018)	10	0.0002	Polynomial
0.129	(+/-0.018)	10	0.002	Polynomial
0.129	(+/-0.018)	10	0.001	Polynomial
0.129	(+/-0.018)	10	0.01	Polynomial
0.916	(+/-0.044)	10	0.1	Polynomial
0.945	(+/-0.063)	10	1	Polynomial
0.945	(+/-0.063)	10	10	Polynomial
0.945	(+/-0.063)	10	20	Polynomial
0.945	(+/-0.063)	10	100	Polynomial
0.945	(+/-0.063)	10	200	Polynomial
0.945	(+/-0.063)	10	1000	Polynomial
0.945	(+/-0.063)	10	2000	Polynomial
0.129	(+/-0.018)	100	2.00E-08	Polynomial
0.129	(+/-0.018)	100	2.00E-07	Polynomial
0.129	(+/-0.018)	100	2.00E-06	Polynomial
0.129	(+/-0.018)	100	2.00E-05	Polynomial
0.129	(+/-0.018)	100	0.0002	Polynomial

0.129	(+/-0.018)	100	0.002	Polynomial
0.129	(+/-0.018)	100	0.001	Polynomial
0.42	(+/-0.290)	100	0.01	Polynomial
0.957	(+/-0.119)	100	0.1	Polynomial
0.945	(+/-0.063)	100	1	Polynomial
0.945	(+/-0.063)	100	10	Polynomial
0.945	(+/-0.063)	100	20	Polynomial
0.945	(+/-0.063)	100	100	Polynomial
0.945	(+/-0.063)	100	200	Polynomial
0.945	(+/-0.063)	100	1000	Polynomial
0.945	(+/-0.063)	100	2000	Polynomial
0.129	(+/-0.018)	1000	2.00E-08	Polynomial
0.129	(+/-0.018)	1000	2.00E-07	Polynomial
0.129	(+/-0.018)	1000	2.00E-06	Polynomial
0.129	(+/-0.018)	1000	2.00E-05	Polynomial
0.129	(+/-0.018)	1000	0.0002	Polynomial
0.129	(+/-0.018)	1000	0.002	Polynomial
0.129	(+/-0.018)	1000	0.001	Polynomial
0.734	(+/-0.389)	1000	0.01	Polynomial
0.934	(+/-0.097)	1000	0.1	Polynomial
0.945	(+/-0.063)	1000	1	Polynomial
0.945	(+/-0.063)	1000	10	Polynomial
0.945	(+/-0.063)	1000	20	Polynomial
0.945	(+/-0.063)	1000	100	Polynomial
0.945	(+/-0.063)	1000	200	Polynomial
0.945	(+/-0.063)	1000	1000	Polynomial
0.945	(+/-0.063)	1000	2000	Polynomial
0.129	(+/-0.018)	10000	2.00E-08	Polynomial
0.129	(+/-0.018)	10000	2.00E-07	Polynomial
0.129	(+/-0.018)	10000	2.00E-06	Polynomial
0.129	(+/-0.018)	10000	2.00E-05	Polynomial
0.129	(+/-0.018)	10000	0.0002	Polynomial
0.407	(+/-0.275)	10000	0.002	Polynomial
0.129	(+/-0.018)	10000	0.001	Polynomial
0.916	(+/-0.044)	10000	0.01	Polynomial
0.945	(+/-0.063)	10000	0.1	Polynomial
0.945	(+/-0.063)	10000	1	Polynomial
0.945	(+/-0.063)	10000	10	Polynomial
0.945	(+/-0.063)	10000	20	Polynomial
0.945	(+/-0.063)	10000	100	Polynomial
0.945	(+/-0.063)	10000	200	Polynomial
0.945	(+/-0.063)	10000	1000	Polynomial
0.945	(+/-0.063)	10000	2000	Polynomial
0.129	(+/-0.018)	100000	2.00E-08	Polynomial
0.129	(+/-0.018)	100000	2.00E-07	Polynomial
0.129	(+/-0.018)	100000	2.00E-06	Polynomial

0.129	(+/-0.018)	100000	2.00E-05	Polynomial
0.129	(+/-0.018)	100000	0.0002	Polynomial
0.653	(+/-0.369)	100000	0.002	Polynomial
0.42	(+/-0.290)	100000	0.001	Polynomial
0.957	(+/-0.119)	100000	0.01	Polynomial
0.945	(+/-0.063)	100000	0.1	Polynomial
0.945	(+/-0.063)	100000	1	Polynomial
0.945	(+/-0.063)	100000	10	Polynomial
0.945	(+/-0.063)	100000	20	Polynomial
0.945	(+/-0.063)	100000	100	Polynomial
0.945	(+/-0.063)	100000	200	Polynomial
0.945	(+/-0.063)	100000	1000	Polynomial
0.945	(+/-0.063)	100000	2000	Polynomial

Tab. 27 Accuracy of the Parameter Estimation Using Grid Search with Cross-Validation for Organic News Feed Simulation.

The model is trained on the full development set.
The scores are computed on the full evaluation set.

	precision	recall	f1-score	support
0	1	1	1	21
1	0.91	0.97	0.94	30
2	0.95	0.88	0.91	24
Accuracy			0.95	75
Macro Avg	0.95	0.95	0.95	75
Weighted Avg	0.95	0.95	0.95	75

Tab. 29 Averaged Monte Carlo Cross Validation Statistical Test Results for Organic News Feed Simulation.

VMS	0.959443
EVS	0.963683
MSE	0.024505
RMSE	0.156535
MAE	0.155023
MAPE	0.030480
R^2	0.958967

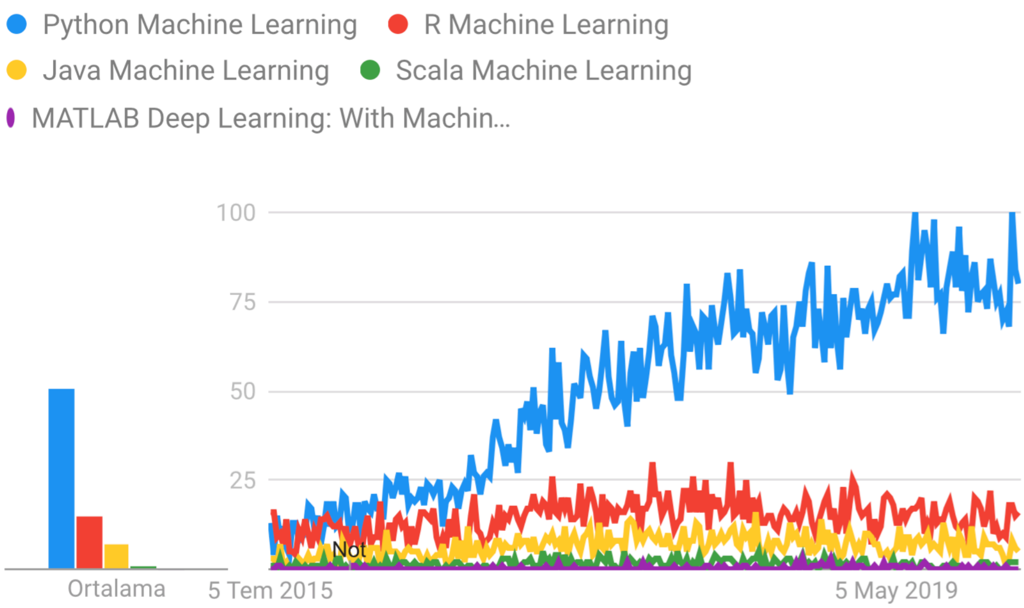
Tab. 28 Exhaustive Monte-Carlo Cross Validation Statistical Test Results for Organic News Feed Simulation.

VMS	EVS	MSE	RMSE	MAE	MAPE	R2
0.954511	0.964728	0.023996	0.154907	0.153306	0.030116	0.959540
0.945871	0.965212	0.023648	0.153781	0.151796	0.029723	0.960176
0.970435	0.964137	0.024494	0.156507	0.155461	0.030636	0.960017
0.972109	0.964373	0.024675	0.157083	0.156325	0.030928	0.960427
0.955138	0.964276	0.023952	0.154765	0.153104	0.030028	0.958699
0.960549	0.964558	0.024174	0.155479	0.154065	0.030313	0.960237
0.951623	0.964004	0.023857	0.154458	0.152706	0.029947	0.959302
0.948282	0.964035	0.023798	0.154266	0.152401	0.029891	0.959038
0.959396	0.963547	0.024793	0.157458	0.156085	0.030677	0.958570
0.971352	0.963299	0.025150	0.158588	0.157663	0.031097	0.958788
0.962135	0.963690	0.024795	0.157466	0.156096	0.030702	0.958827
0.944401	0.964761	0.024006	0.154939	0.152549	0.029886	0.960013
0.952173	0.962074	0.024353	0.156055	0.153987	0.030162	0.957240
0.950138	0.964204	0.024324	0.155962	0.154010	0.030197	0.959341
0.965106	0.964154	0.024296	0.155873	0.154487	0.030419	0.959638
0.949807	0.964240	0.023582	0.153565	0.151265	0.029605	0.959152
0.956295	0.964923	0.024065	0.155128	0.153509	0.030173	0.960190
0.954810	0.962741	0.024131	0.155342	0.153775	0.030182	0.957389
0.969405	0.963505	0.024556	0.156703	0.155711	0.030732	0.959418
0.960046	0.964074	0.024249	0.155720	0.154489	0.030368	0.959391
0.964659	0.964794	0.024362	0.156082	0.154841	0.030561	0.960597
0.943544	0.964200	0.023506	0.153316	0.151040	0.029548	0.958964
0.958886	0.962259	0.025193	0.158722	0.157425	0.030942	0.957675
0.949276	0.963388	0.024547	0.156675	0.154737	0.030267	0.958256
0.942802	0.964155	0.024265	0.155772	0.153632	0.030029	0.959109
0.963862	0.962603	0.025059	0.158301	0.156740	0.030834	0.957735
0.967263	0.961676	0.025355	0.159234	0.158012	0.031168	0.957220
0.959422	0.963306	0.025082	0.158372	0.156876	0.030867	0.958763
0.956295	0.963948	0.025026	0.158195	0.156804	0.030815	0.958457
0.950468	0.962298	0.024661	0.157038	0.155150	0.030408	0.957251
0.961022	0.963664	0.025092	0.158405	0.156863	0.030890	0.958914
0.955465	0.962953	0.024849	0.157635	0.155932	0.030627	0.958530
0.956492	0.963428	0.024988	0.158077	0.156567	0.030760	0.958717
0.952203	0.962962	0.024807	0.157501	0.155796	0.030609	0.958444
0.952532	0.964965	0.023880	0.154531	0.152665	0.029969	0.960401
0.969285	0.962774	0.024452	0.156372	0.155232	0.030623	0.958588
0.978464	0.962520	0.024803	0.157489	0.156743	0.031059	0.958744
0.959097	0.963363	0.024138	0.155366	0.153887	0.030298	0.958914
0.944955	0.964772	0.023412	0.153009	0.150544	0.029427	0.959698
0.961692	0.964573	0.024802	0.157486	0.155956	0.030653	0.959853
0.950767	0.963251	0.024127	0.155329	0.153021	0.029954	0.958914

0.967149	0.962772	0.025002	0.158119	0.156911	0.030919	0.958155
0.978200	0.963631	0.025372	0.159285	0.158459	0.031385	0.959438
0.963057	0.962262	0.024848	0.157633	0.156233	0.030731	0.957344
0.950767	0.963800	0.024318	0.155941	0.153854	0.030147	0.958694
0.956846	0.963541	0.024636	0.156959	0.155296	0.030507	0.958894
0.971492	0.963712	0.025214	0.158789	0.157757	0.031102	0.959569
0.966829	0.963300	0.024752	0.157328	0.156124	0.030780	0.958885
0.978812	0.964629	0.025148	0.158583	0.157982	0.031287	0.960601
0.963562	0.964085	0.024570	0.156747	0.155407	0.030618	0.958871
0.960524	0.962675	0.024621	0.156910	0.155687	0.030577	0.957607
0.963885	0.964543	0.024625	0.156925	0.155824	0.030667	0.959855
0.948613	0.963861	0.023995	0.154902	0.152842	0.029952	0.958765
0.959396	0.962837	0.024411	0.156239	0.154704	0.030325	0.957432
0.961645	0.965022	0.024568	0.156743	0.155580	0.030634	0.960591
0.942874	0.963889	0.023835	0.154386	0.152133	0.029780	0.958733
0.972092	0.963012	0.024922	0.157866	0.156910	0.031009	0.958728
0.964659	0.962596	0.024541	0.156655	0.155207	0.030528	0.957577
0.971370	0.964368	0.024899	0.157793	0.156893	0.030990	0.960627
0.958313	0.962756	0.024310	0.155916	0.154216	0.030234	0.957312
0.952831	0.964915	0.024145	0.155388	0.153536	0.030091	0.959117
0.966829	0.963714	0.024790	0.157450	0.156388	0.030810	0.959075
0.973106	0.963745	0.025020	0.158177	0.157278	0.031070	0.959899

FIGURES

Fig. 1 Worldwide Popularity of Programming Languages for Machine Learning

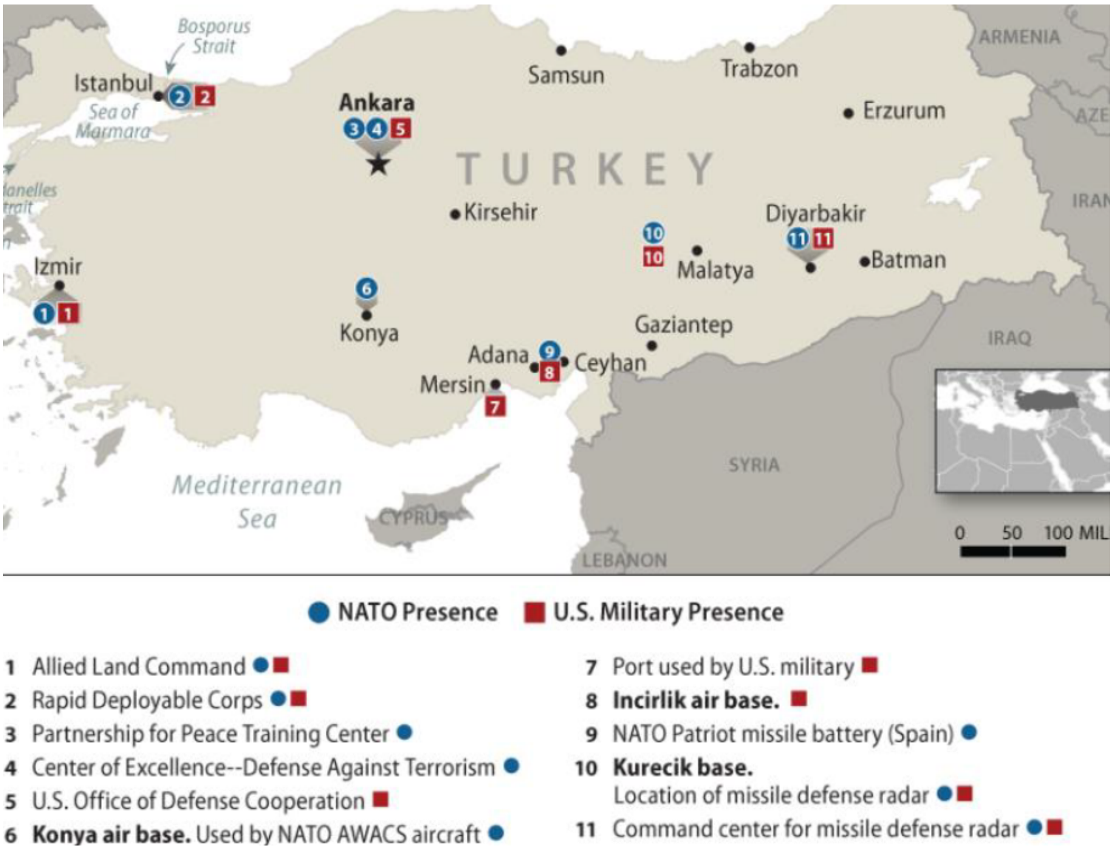


(Source: Google Trends)

Fig. 2 Oil and Natural Gas Production, and Gas Pipeline Structure between Europe and West Asia

(Source: Akdemir, 2010)

Fig. 3 Occurrence of United States and NATO Military in Turkey

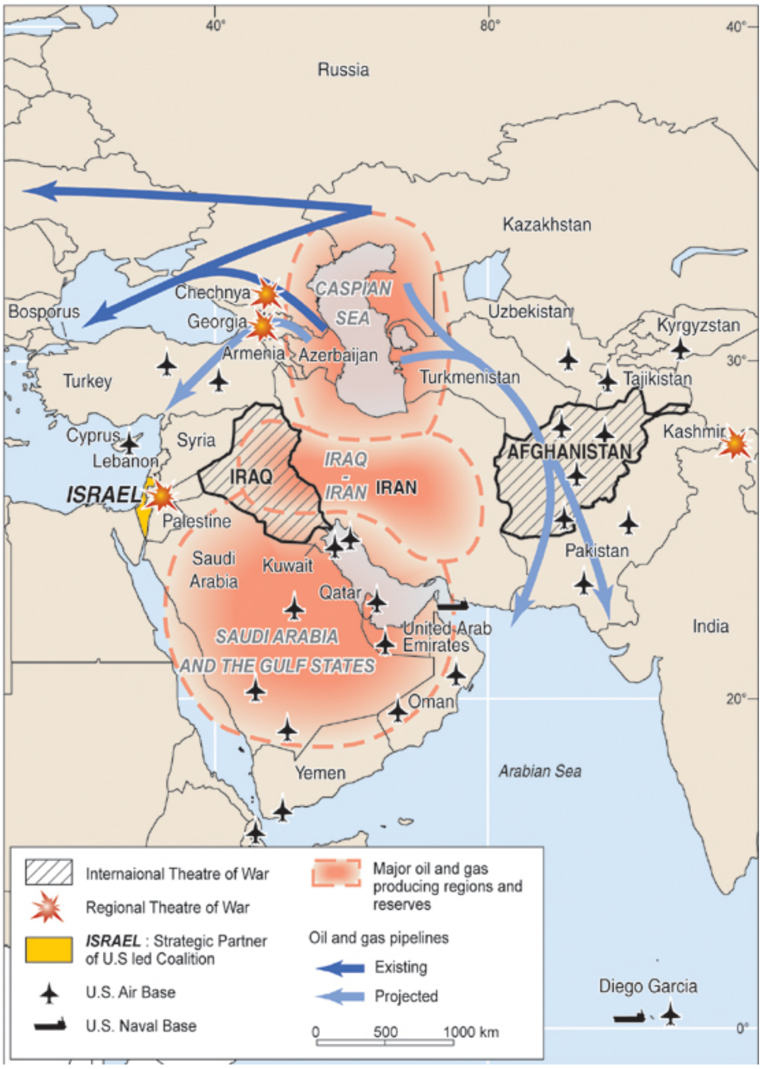


(Source: Congressional Research Service, 2020)

Fig. 5 Geopolitical Vectors of Turkey

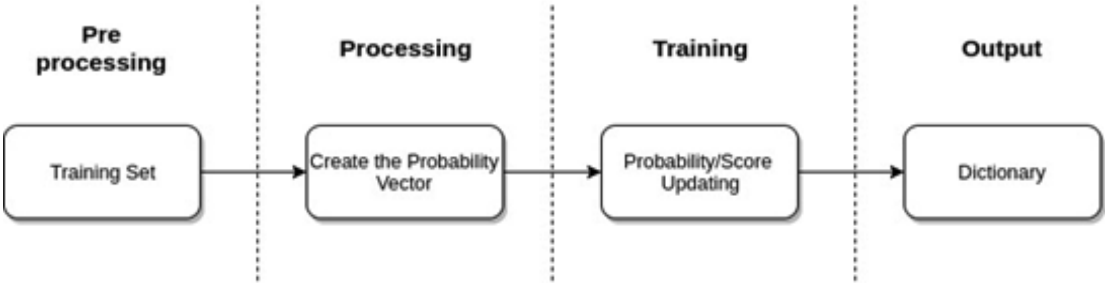
(Source: Geopolitical Intelligence Services, 2018)

Fig. 4 United States' Middle East Policy, and Oil Wars in Recent History



(Source: American Foreign Policy, 2008)

Fig. 6 Workflow of the Learning Automata



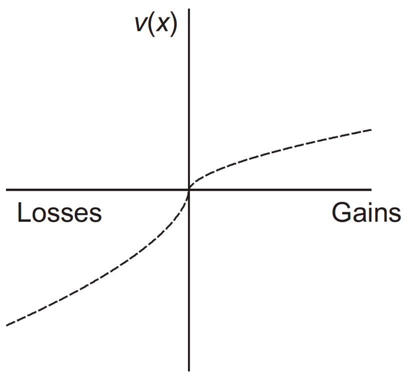
(Source: Sarigiannidis et al., 2018)

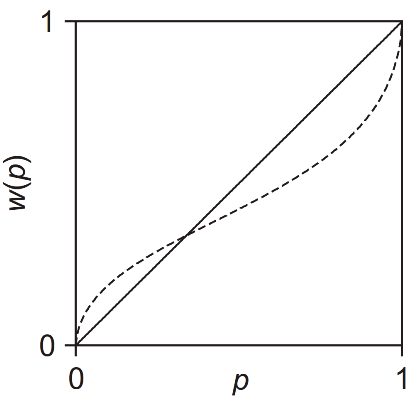
Fig. 7 Numerical exemplification Learning Automata operating mechanism

(Source: Sarigiannidis et al., (2018) modified by the author)

Fig. 8 Risk aversion of utility function

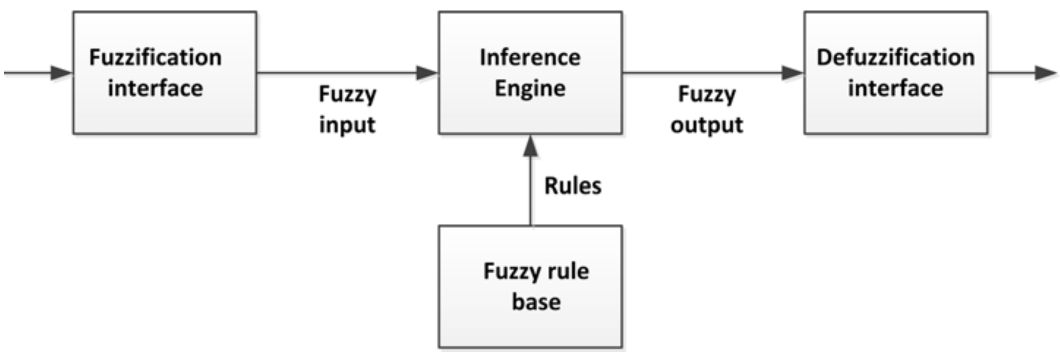
Fig. 9 Loss aversion of utility function





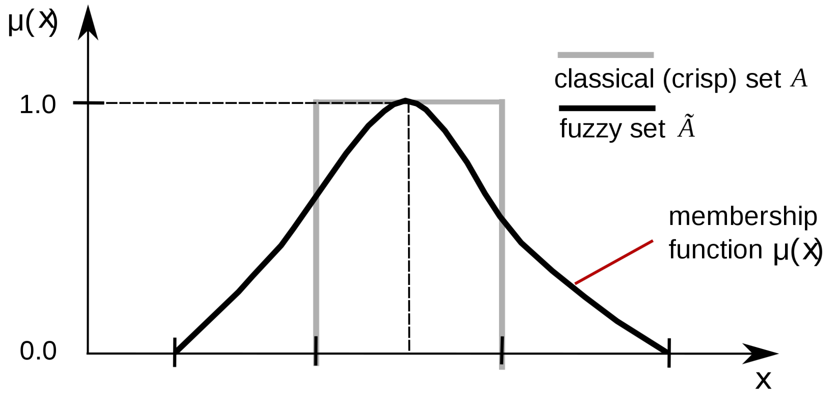
(Source: Kahneman and Tversky, 1991)

Fig. 10 Workflow of Fuzzy Logic



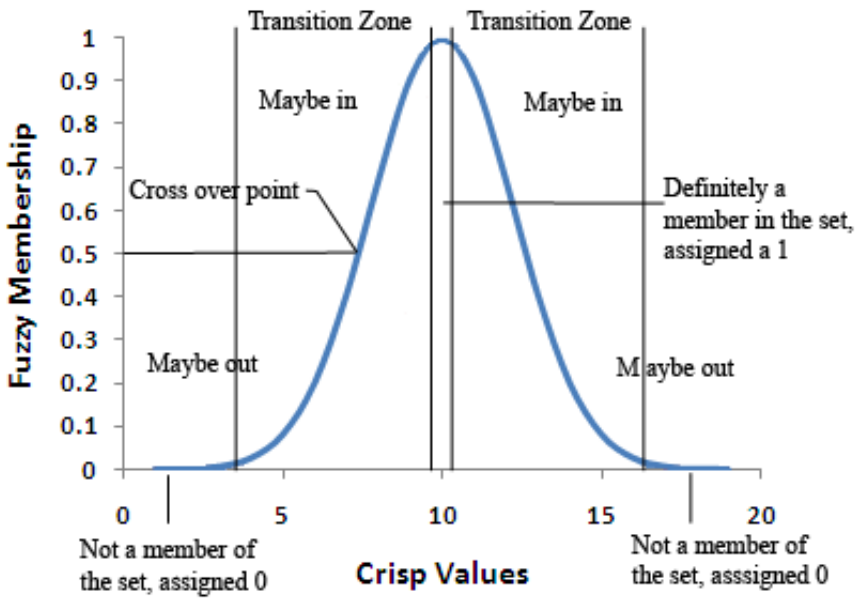
(Source: Nilashi M., et al., 2011)

Fig. 11 Classical (crisp) set vs Fuzzy set



(Source: Christodoulou S., et al., 2012)

Fig. 12 Illustration of Fuzzy Membership on Vague States



(Source: Nasehi S., et al., 2017)

Fig.13 Right Shoulder of Sigmoidal Function

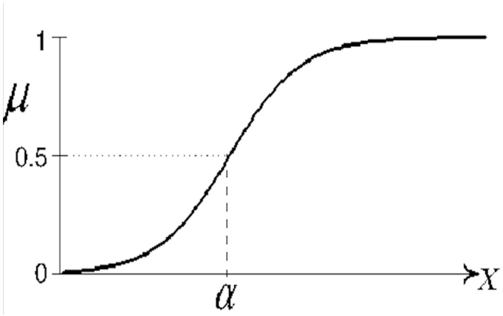
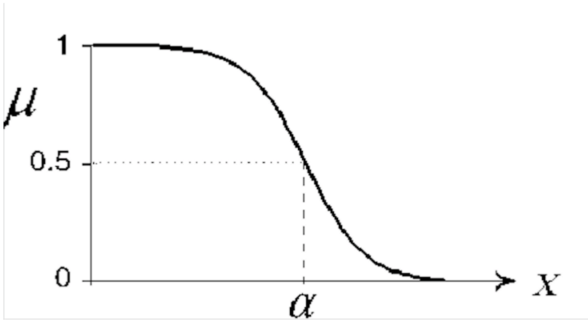
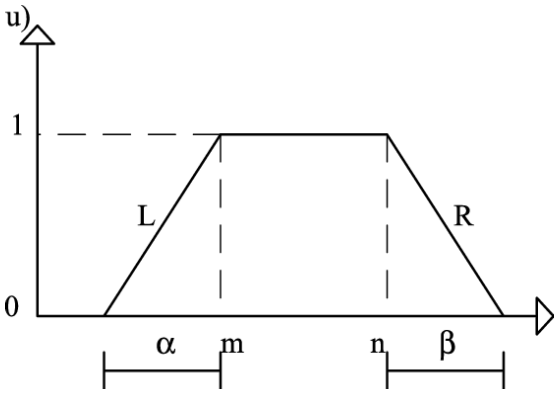


Fig.14 Left Shoulder of Sigmoidal Function.



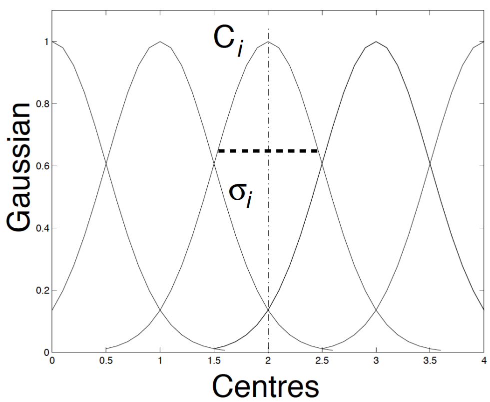
(Source: Robinson V., 2003)

Fig. 15 Illustration of Trapezoidal Membership Function



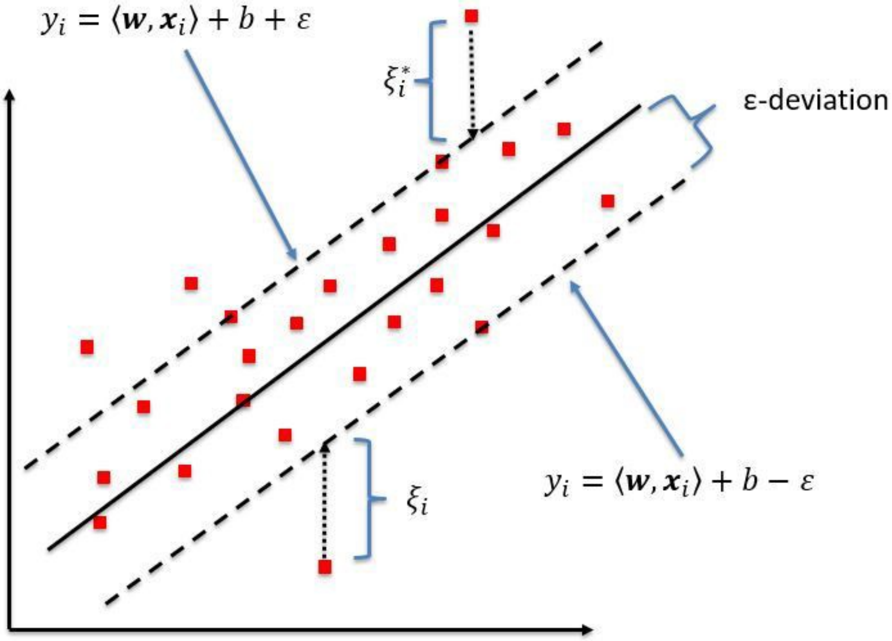
(Source: Kecman, 2001)

Fig. 16 Illustration of Gaussian Membership Function



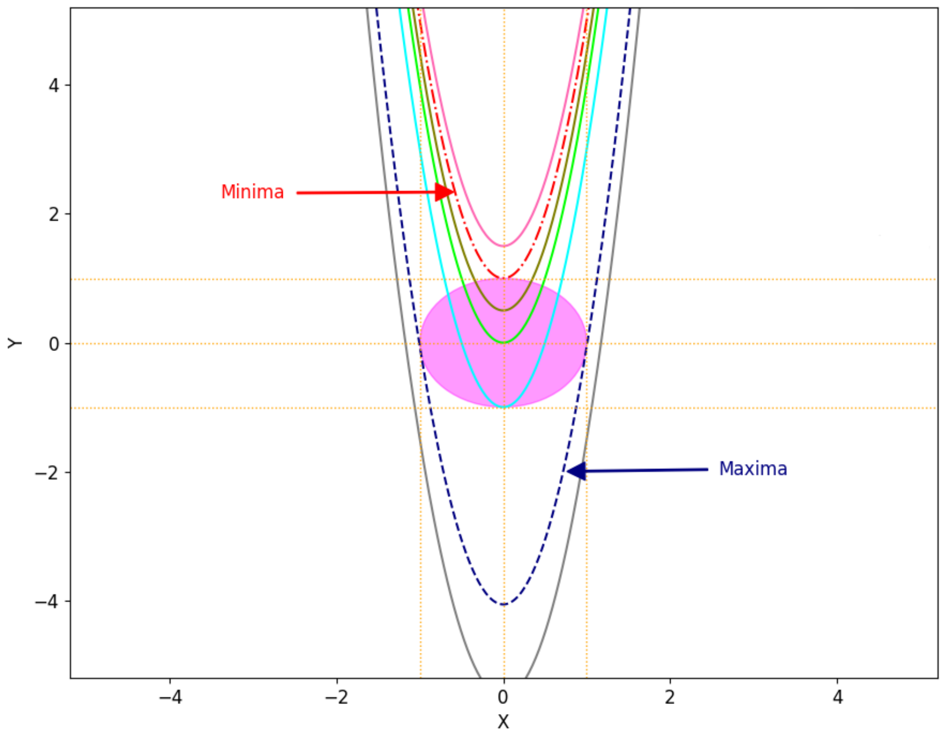
(Source: Boostan E., et al., 2015)

Fig. 17 Visual Demonstration of Support Vector Regression and ε-Tubes



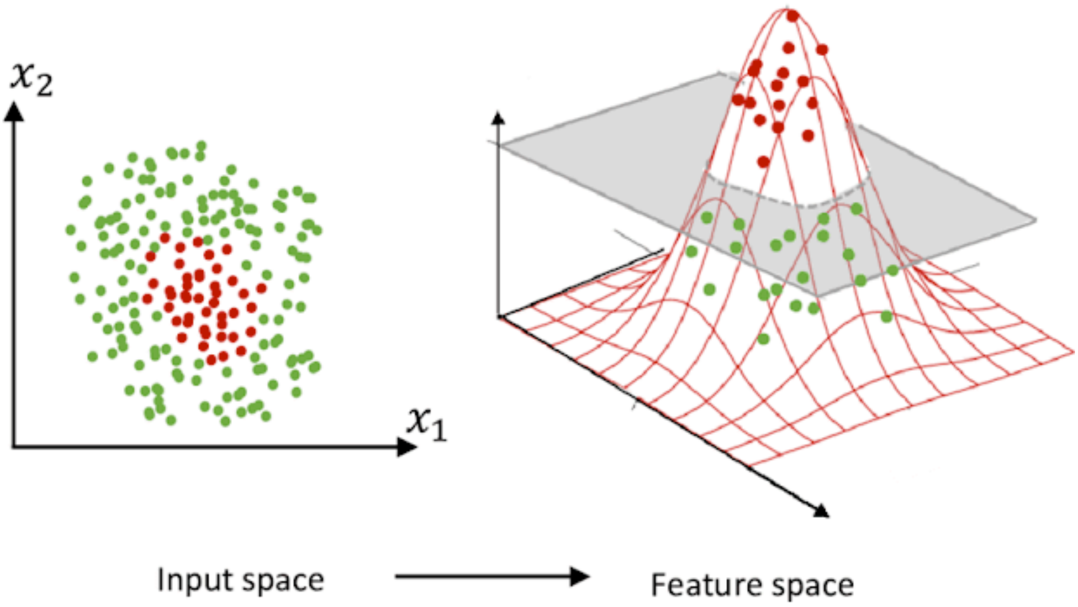
(Source: Kleynhans et al., 2017)

Fig. 18 Visual Demonstration of Lagrange Multiplier Method



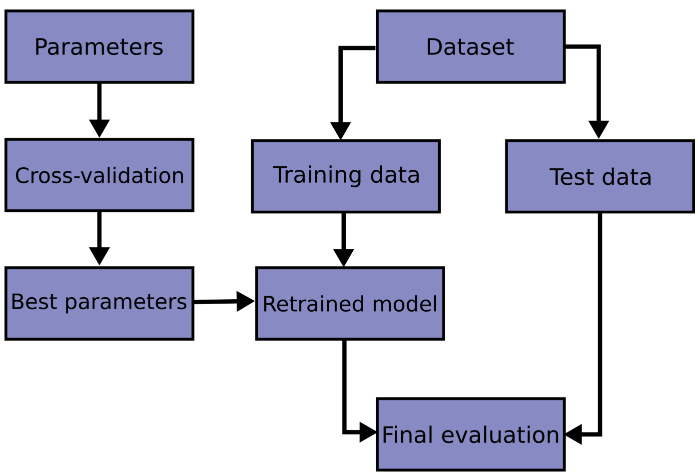
(Source: Bhattacharyya, S., (2018) Support Vector Machine: Complete Theory. Towardsdatascience)

Fig. 19 Mapping Inputs into High-Dimensional Feature Space by Using Kernel Function



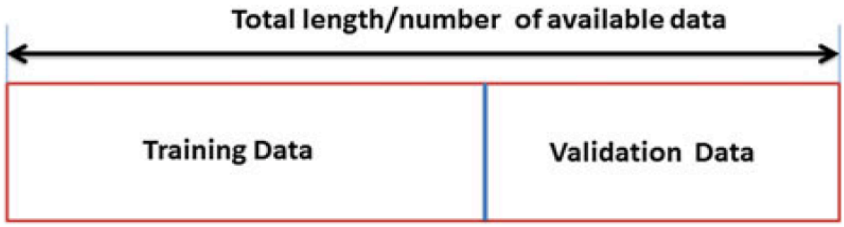
(Source: Jenis Algoritma Pada Machine Learning, 2019)

Fig. 20 Accommodation of Cross Validation and Data Separation on Methodological Workflow



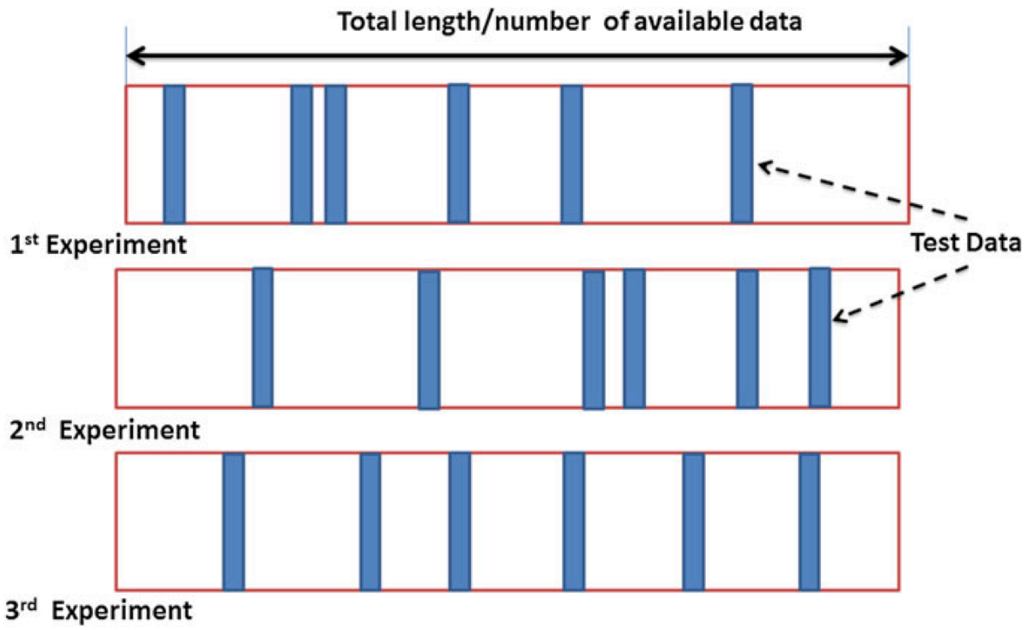
(Source: Scikit-learn, Cross-Validation: Evaluating Estimator Performance)

Fig. 21 Traditional Way of Separating Training and Validation Groups



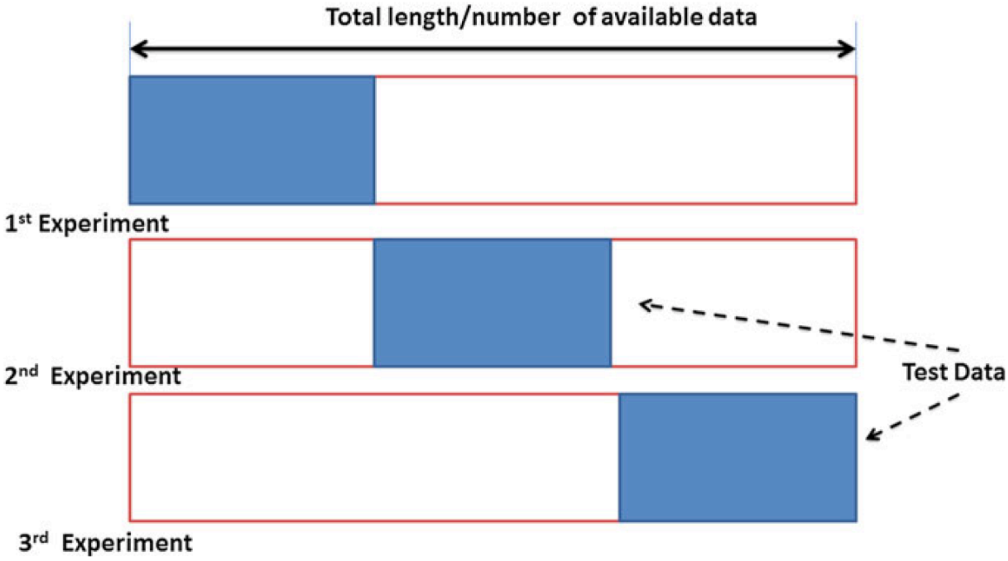
(Source: Dubitzky et al., 2007)

Fig. 22 Monte Carlo Cross Validation



(Source: Dubitzky et al., 2007)

Fig. 23 K-Fold Cross Validation

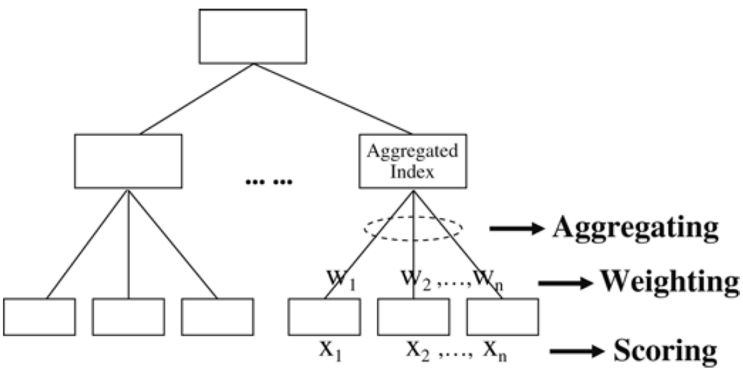


(Source: Dubitzky et al., 2007)

Fig. 24. Confusion Matrix

Confusion Matrix		Predicted Value	
		Positive	Negative
Actual Value	Positive	True Positive	False Negative
	Negative	False Positive	True Negative

Fig. 25 Operation Structure for Aggregation Index



(Source: Guh et al., 2008)

Fig. 26 Flowchart of the proposed methodology for Chapter 5

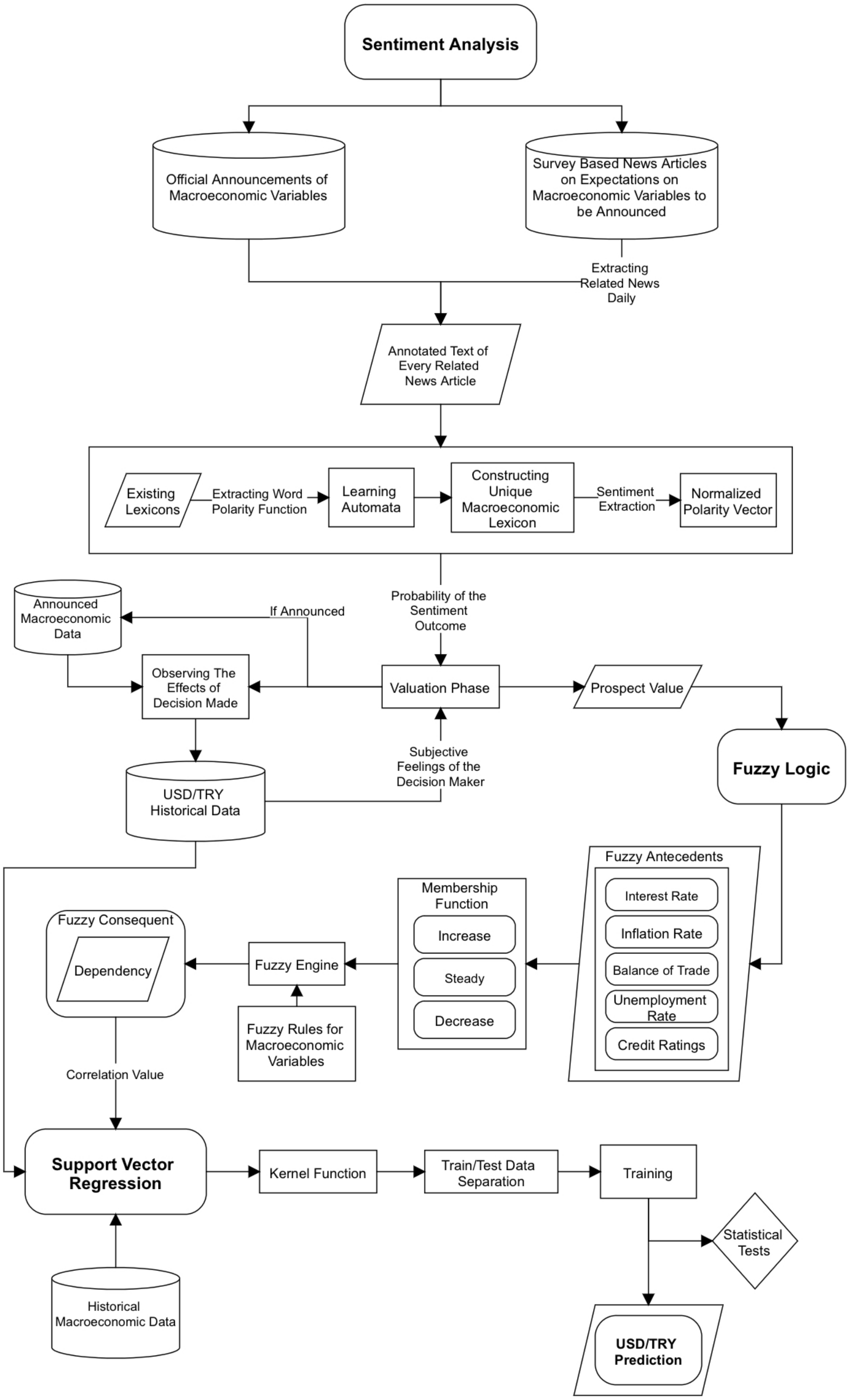


Fig. 27. Dependency of USD/TRY on News Regarding to Inflation Rate

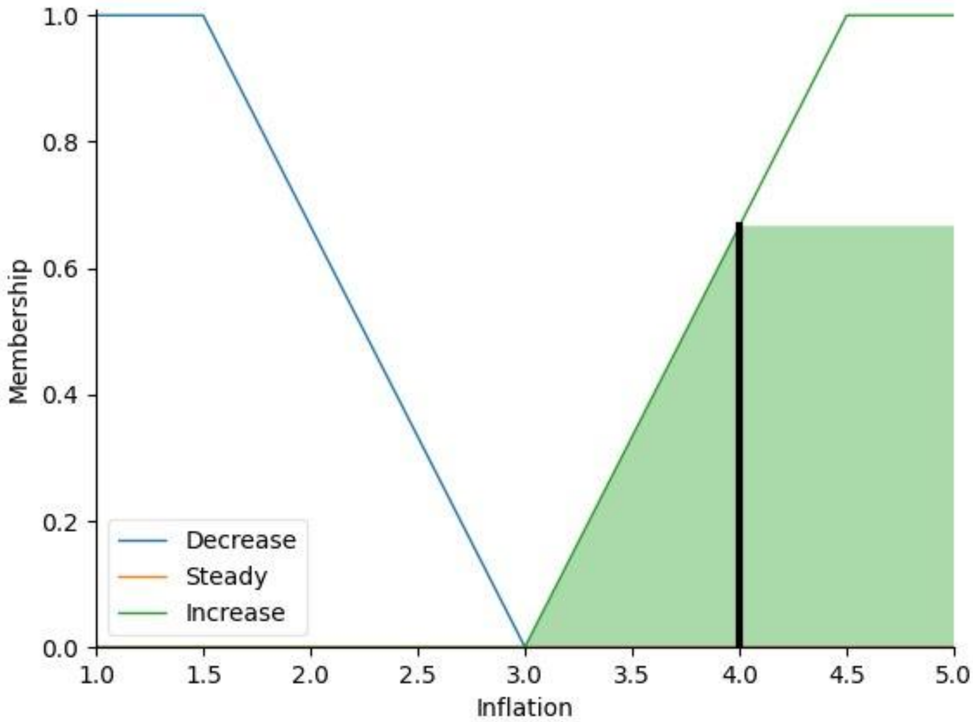


Fig. 28. Dependency of USD/TRY on News Regarding to Interest Rate

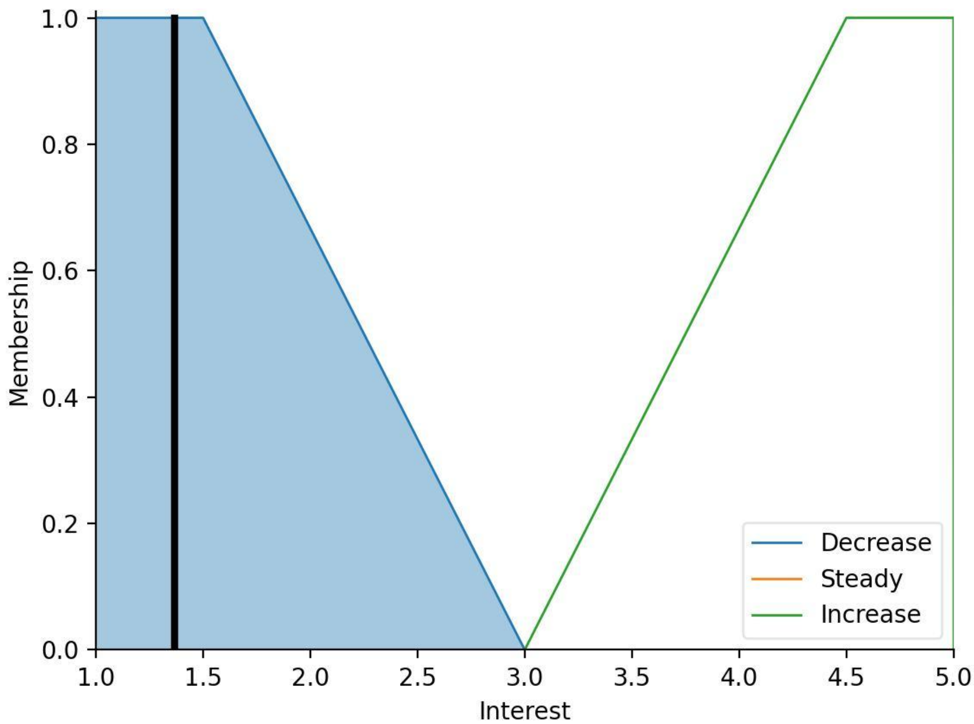


Fig. 29. Dependency of USD/TRY on News Regarding to Balance of Trade

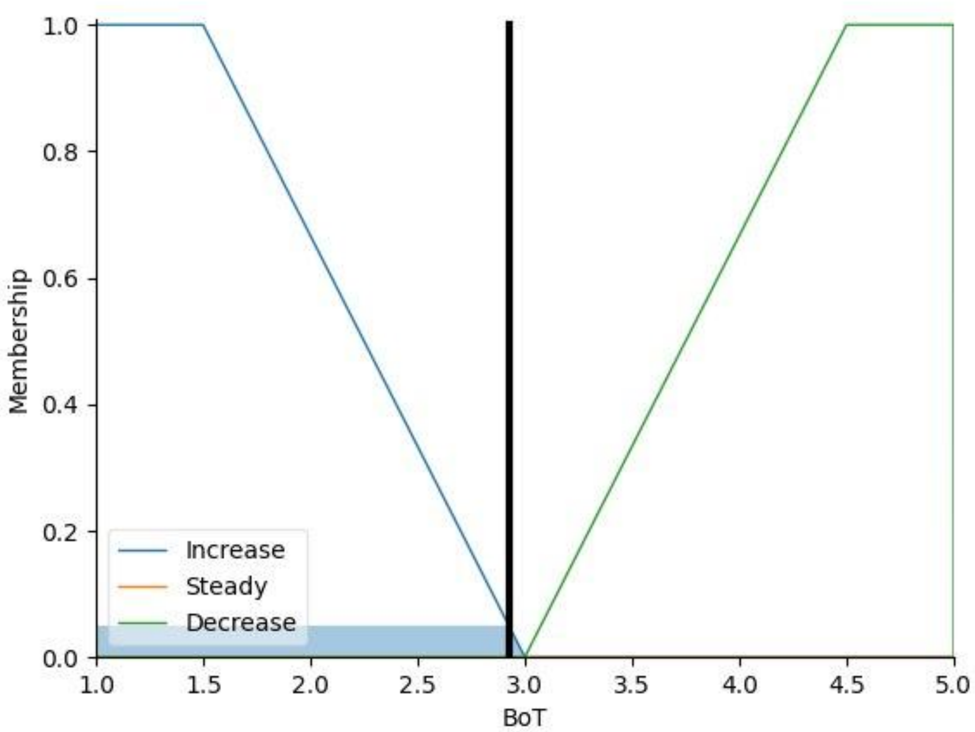


Fig. 30. Dependency of USD/TRY on News Regarding to Unemployment Rate

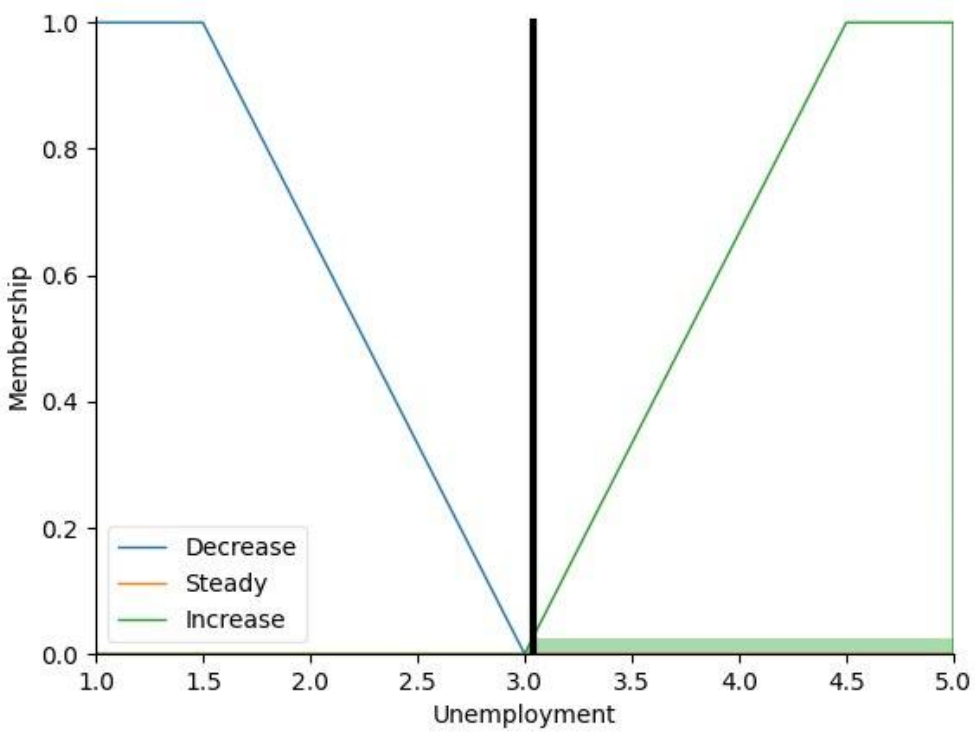


Fig. 31. Dependency of USD/TRY on News Regarding to Credit Ratings

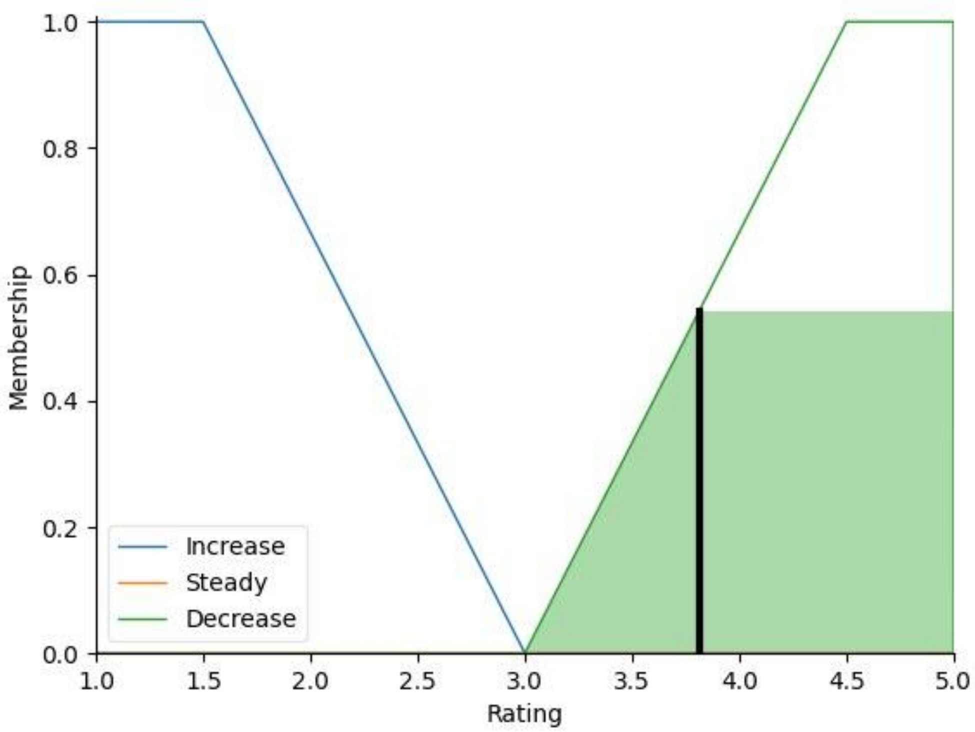


Fig. 32. Correlation between Macroeconomic News and USD/TRY

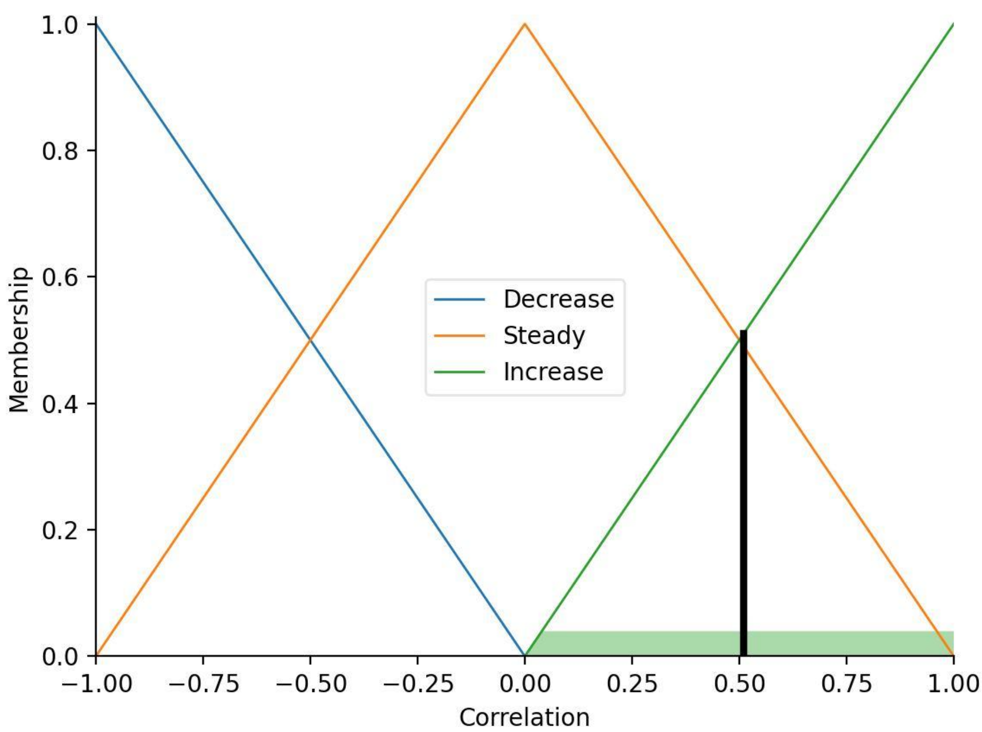


Fig. 33 Illustration of PCR/PCA and PLS Results for Macroeconomic News

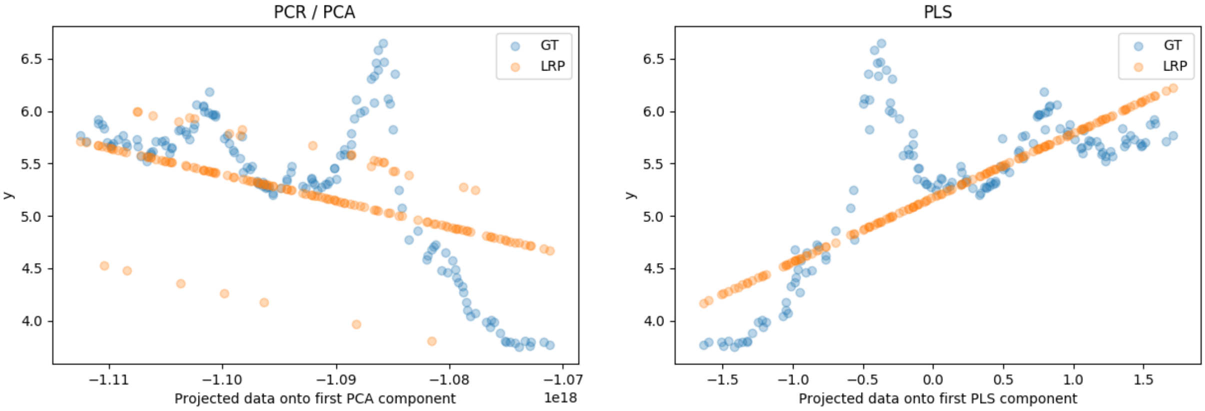


Fig. 34 Kernel Selection for PCA: Linear Kernel

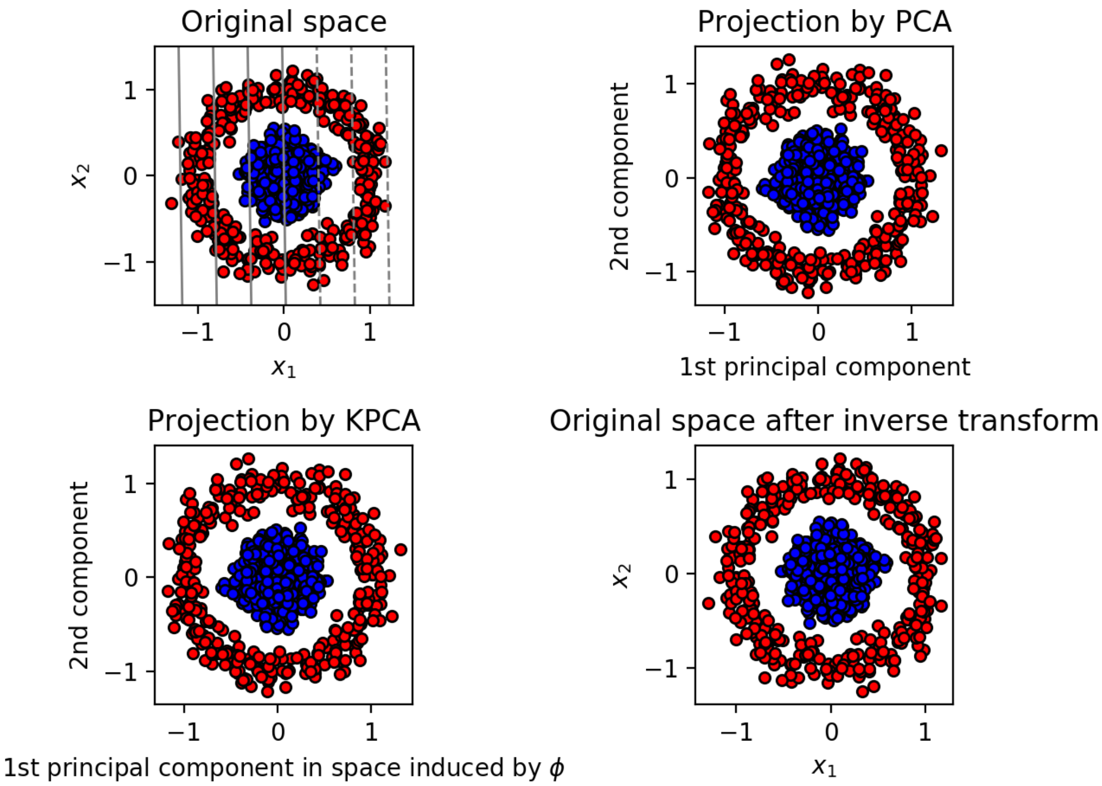


Fig. 35 Kernel Selection for PCA: Gaussian RBF Kernel RBF

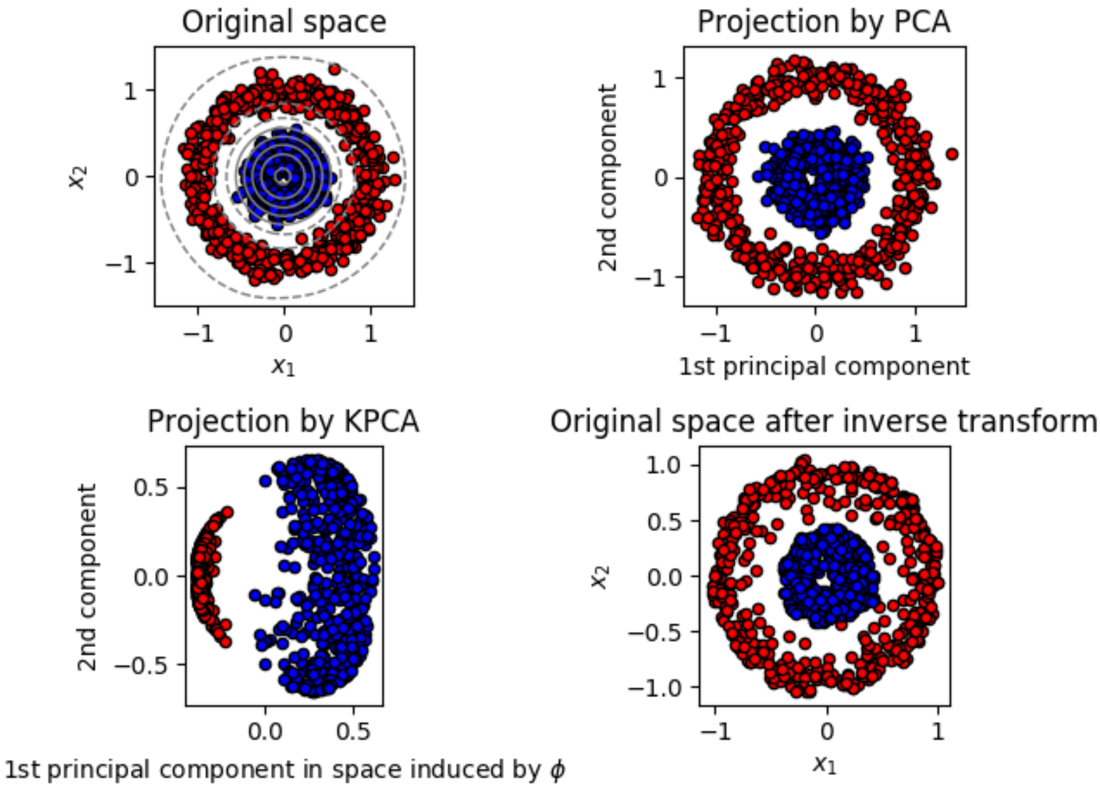


Fig. 36 Layers View for 3D Visualisation of Grid Seach Cross Validation Accuracy of Macroeconomic News

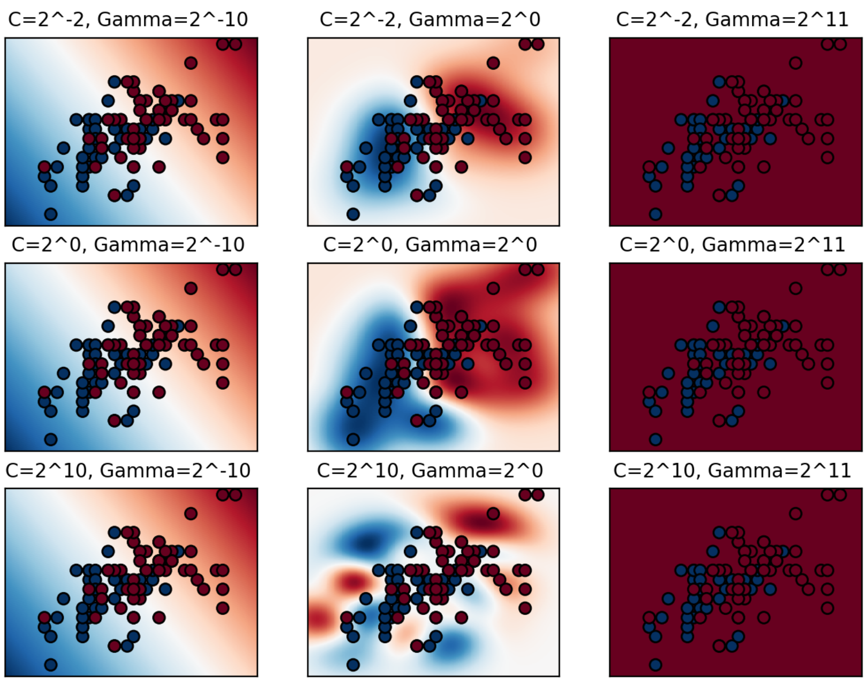


Fig. 37. Grid Search Cross Validation Accuracy Heat Map for the Hyperparameter Selection of Macroeconomic News

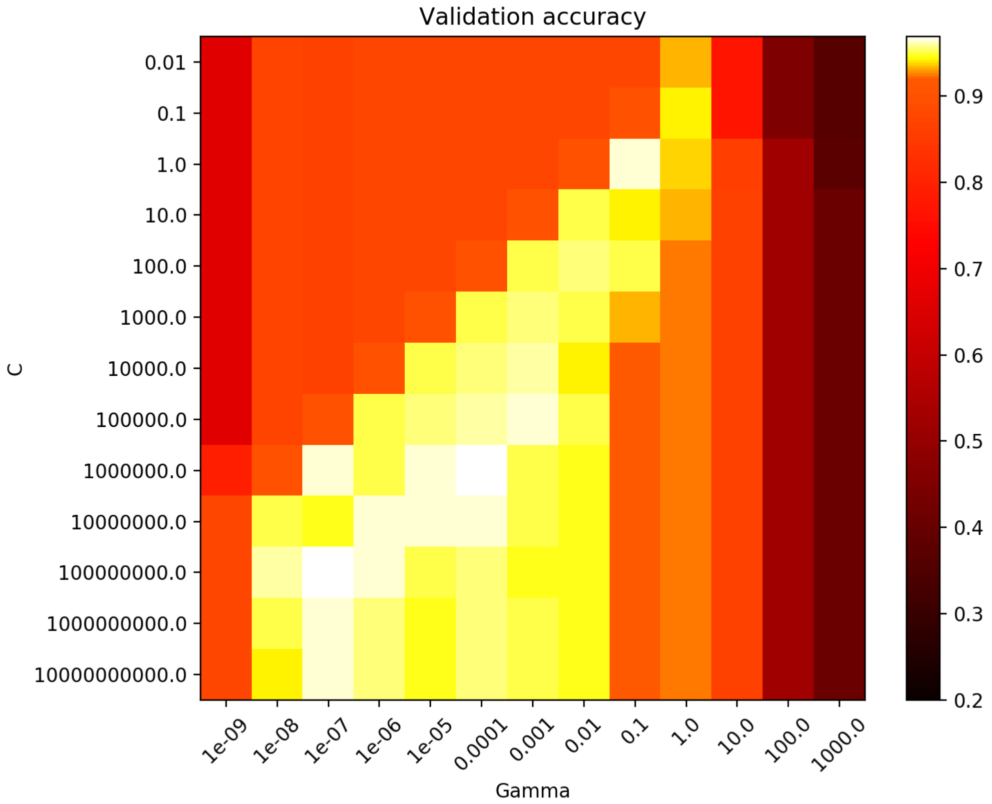


Fig. 38. USD/TRY Data Parallel to Macroeconomic Events mirrored in Constructed Hyperspace (1)

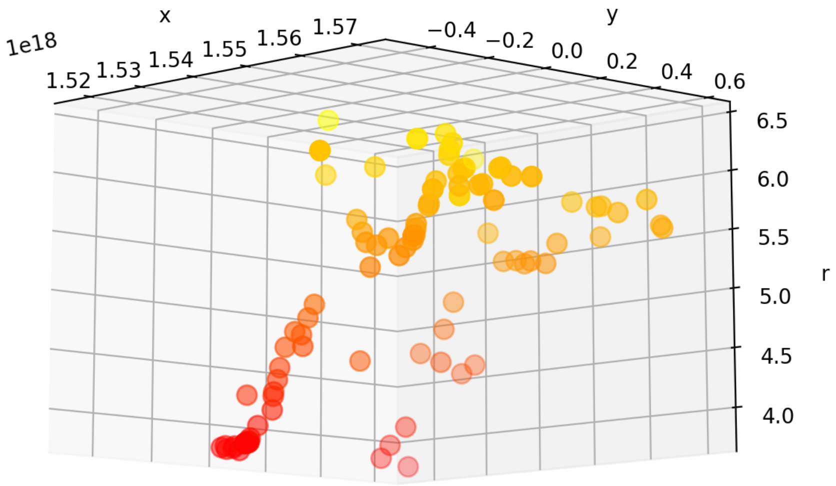


Fig. 39. USD/TRY Data Parallel to Macroeconomic Events mirrored in Constructed Hyperspace (2)

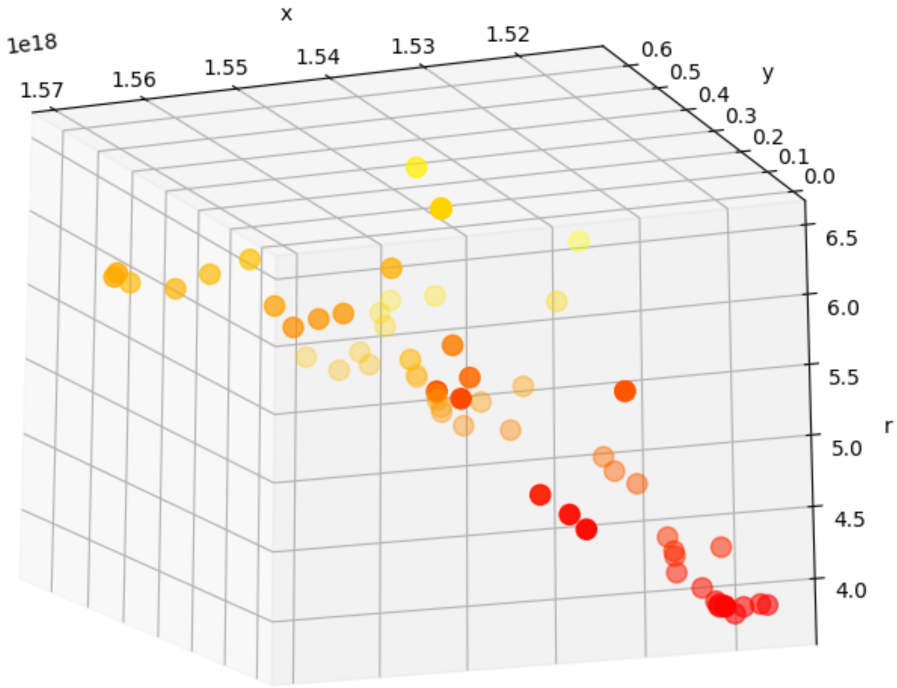


Fig. 40. USD/TRY Data Parallel to Macroeconomic Events mirrored in Constructed Hyperspace (3)

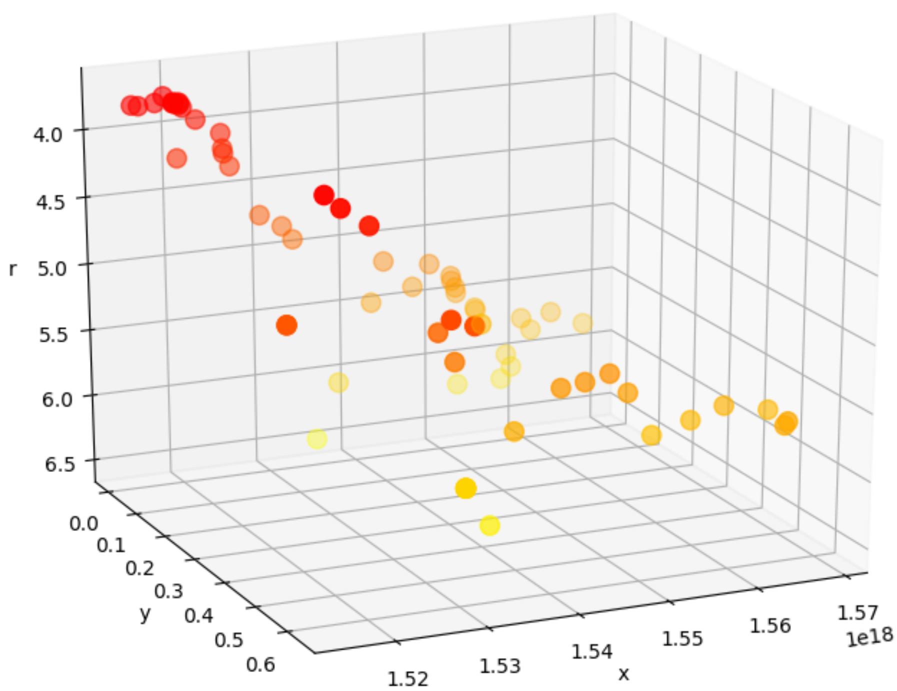
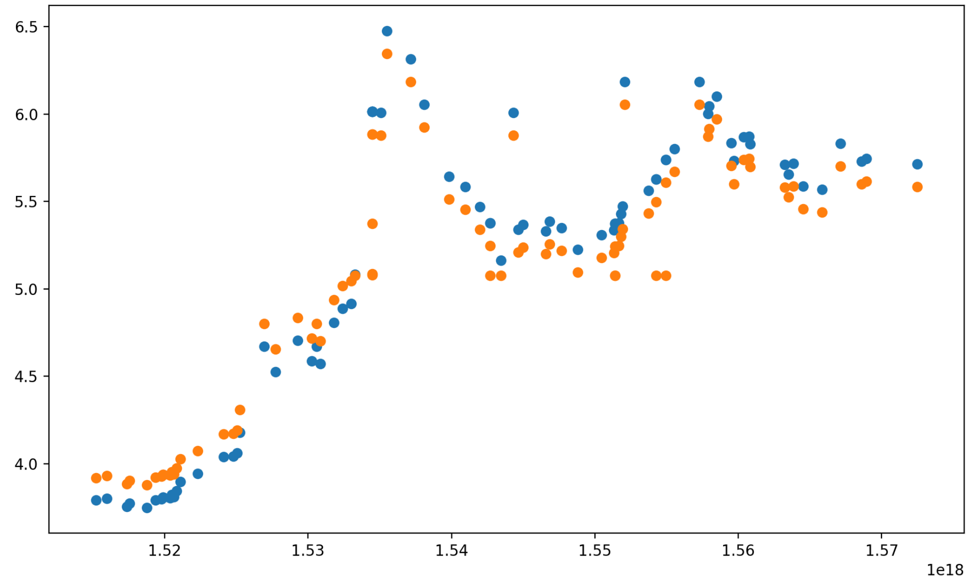


Fig. 41. Actual versus predicted USD/TRY prices obtained by analyzing recently published news and statements regarding on macroeconomic indicators.



**Blue Points denote Actual USD/TRY Prices

**Orange Points denote Predicted USD/TRY Prices

Fig. 42 Flowchart of the proposed methodology of Chapter 6

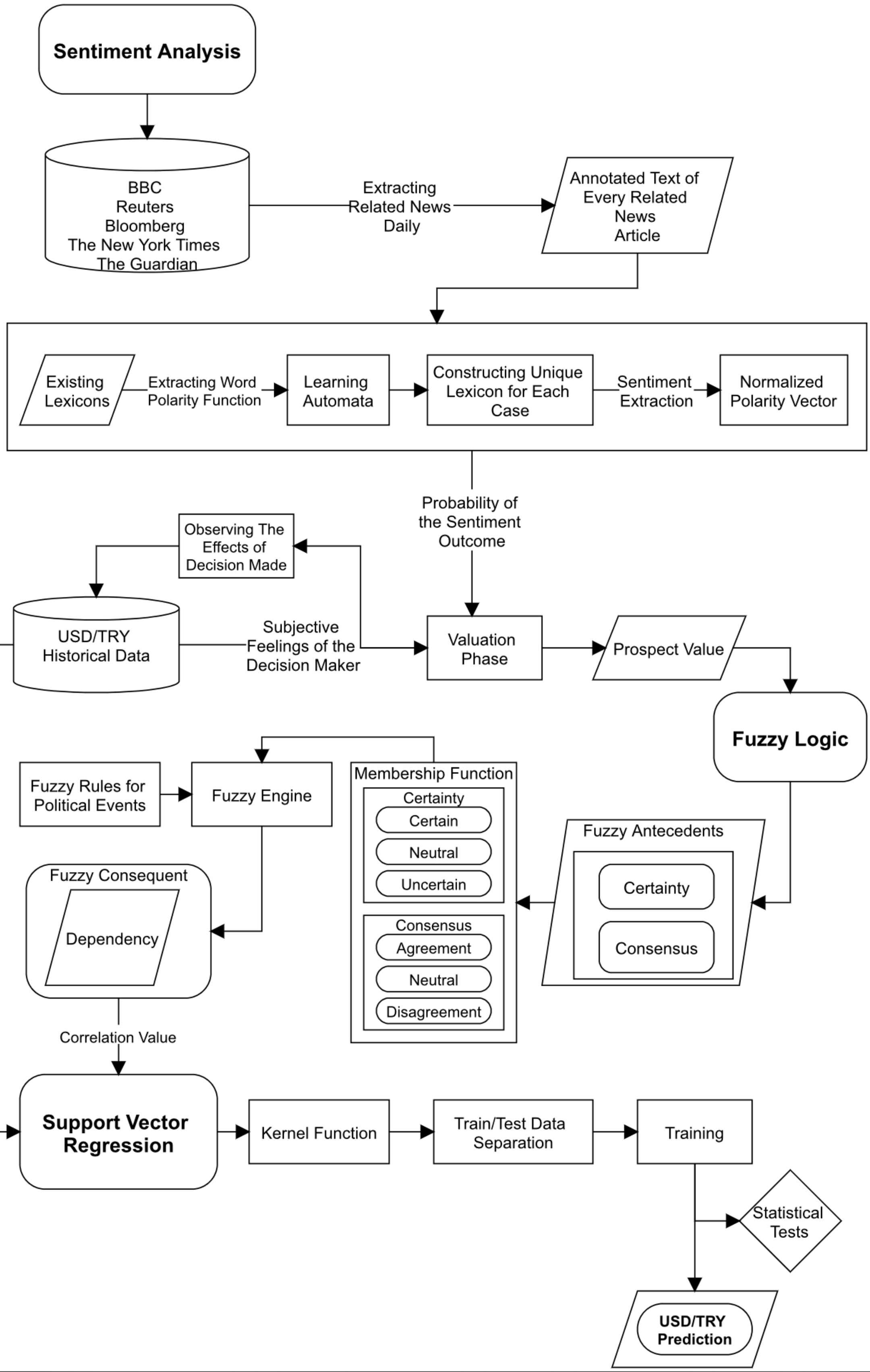


Fig. 43 Degree of Political Certainty Pastor Andrew Brunson Case

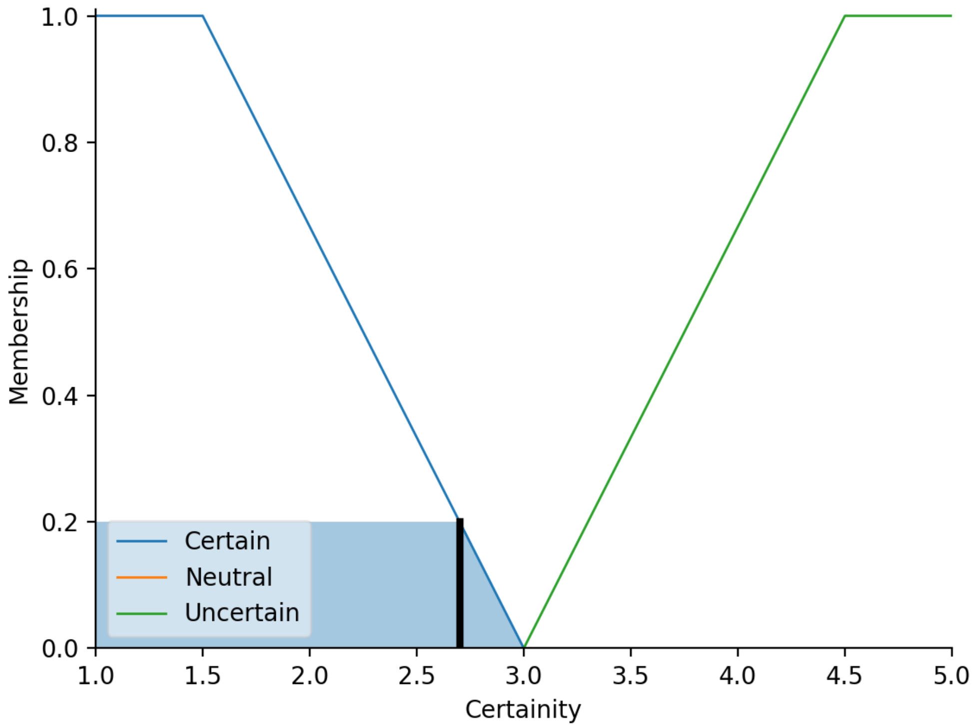


Fig. 44 Degree of Political Consensus Pastor Andrew Brunson Case

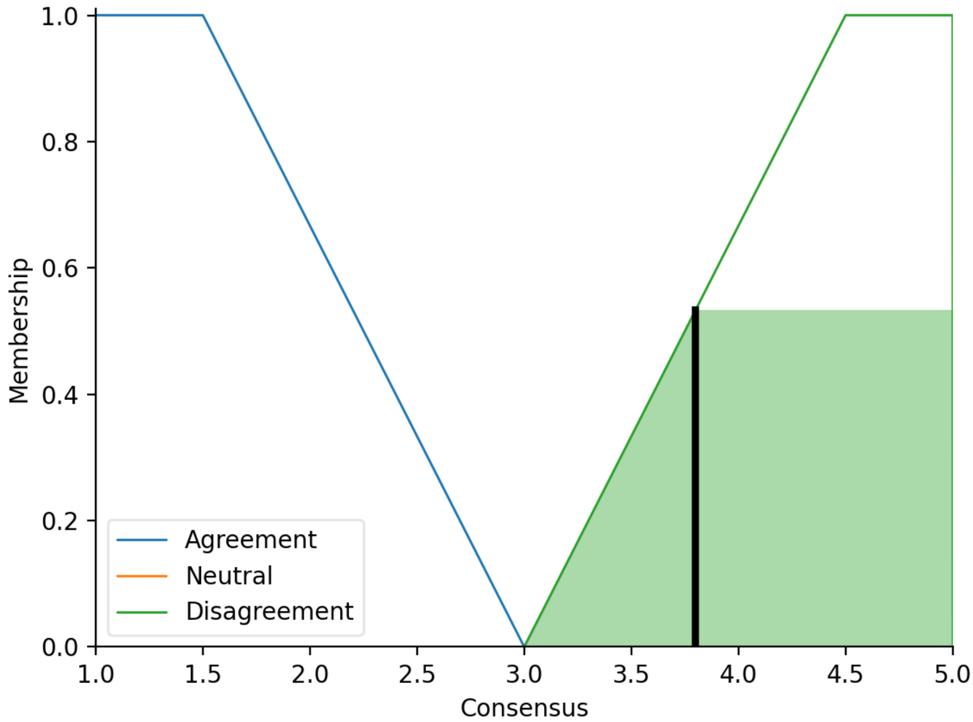


Fig. 45 Dependency Rate between Published News Related to Pastor Andrew Brunson Case and USD/TRY Exchange Rate Fluctuations

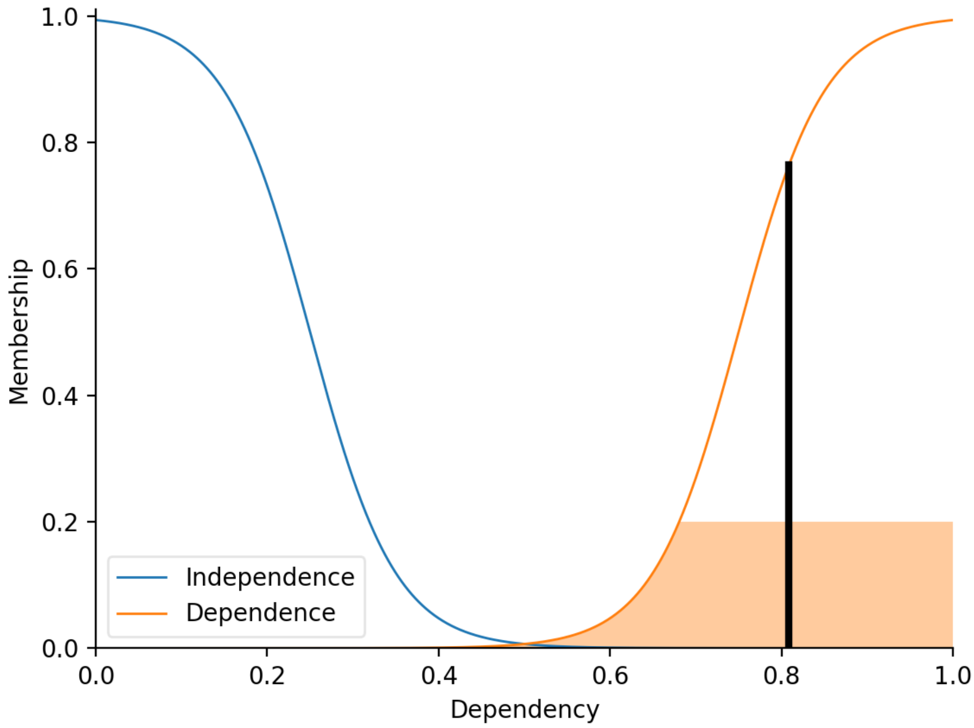


Fig. 46 Illustration of PCR/PCA and PLS Results for Pastor Andrew Brunson Case

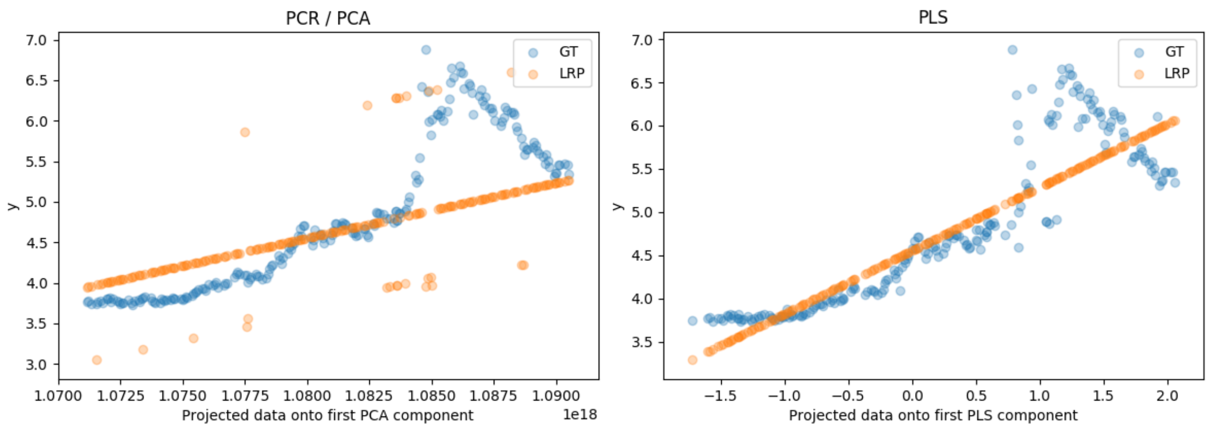


Fig. 47. Grid Search Cross Validation Accuracy Heat Map of Hyperparameter Selection for the Pastor Andrew Brunson Case

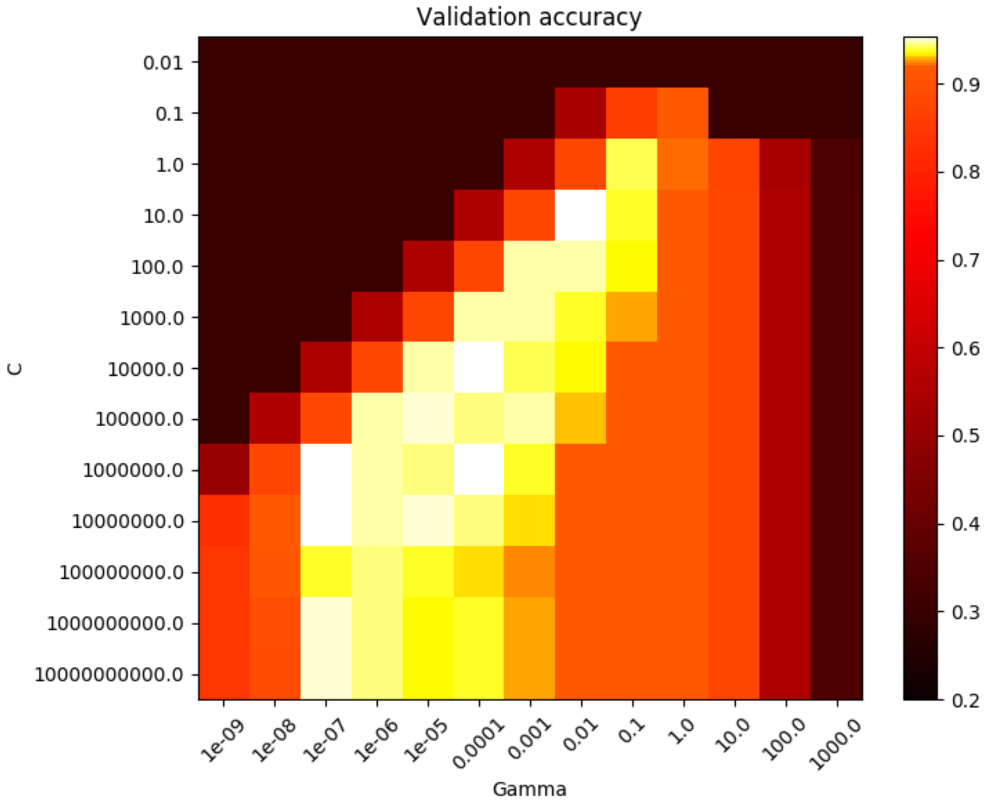


Fig. 48 Layers View for 3D Visualisation of Grid Seach Cross Validation Accuracy of Pastor Andrew Brunson Case

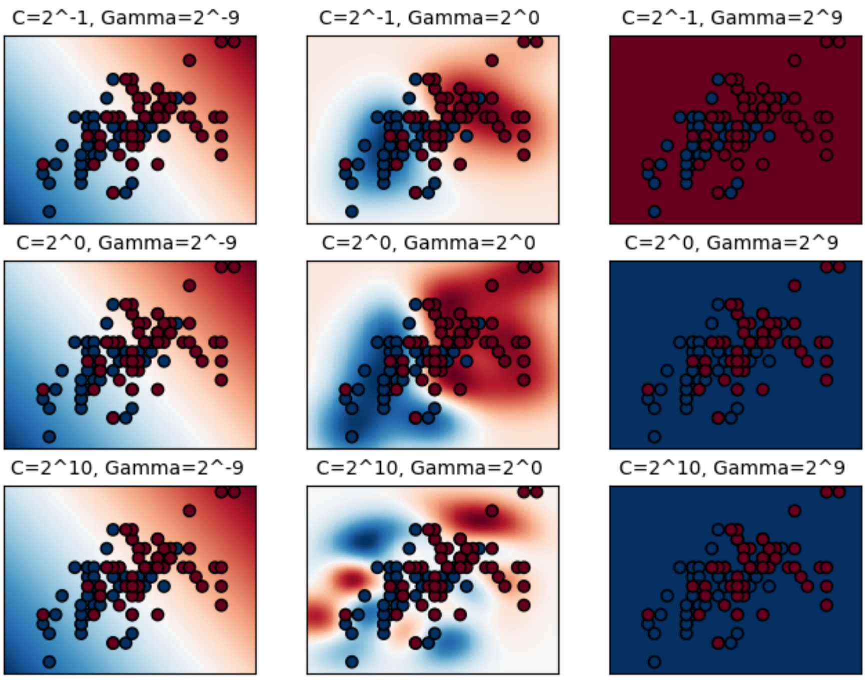
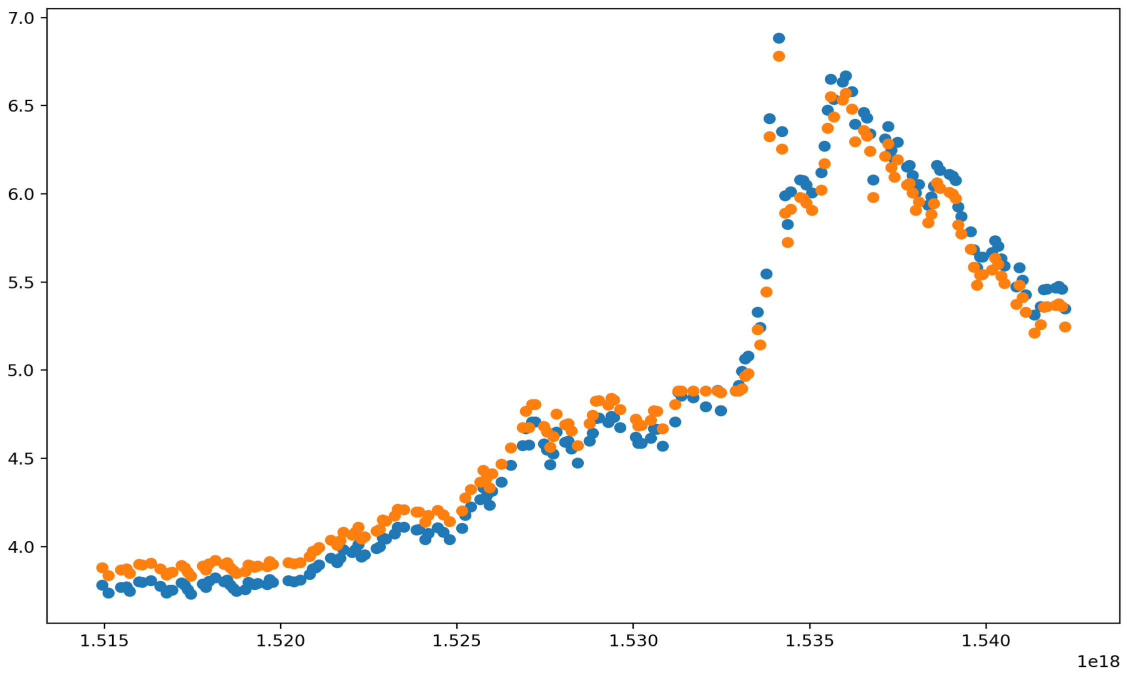


Fig. 49. Actual versus predicted USD/TRY prices obtained by analyzing recently published news and statements regarding on Pastor Andrew Brunson Case.



**Blue Points denote Actual USD/TRY Prices

**Orange Points denote Predicted USD/TRY Prices

Fig. 50 Degree of Political Certainty for 2018 Turkish Parliamentary and Presidential Elections

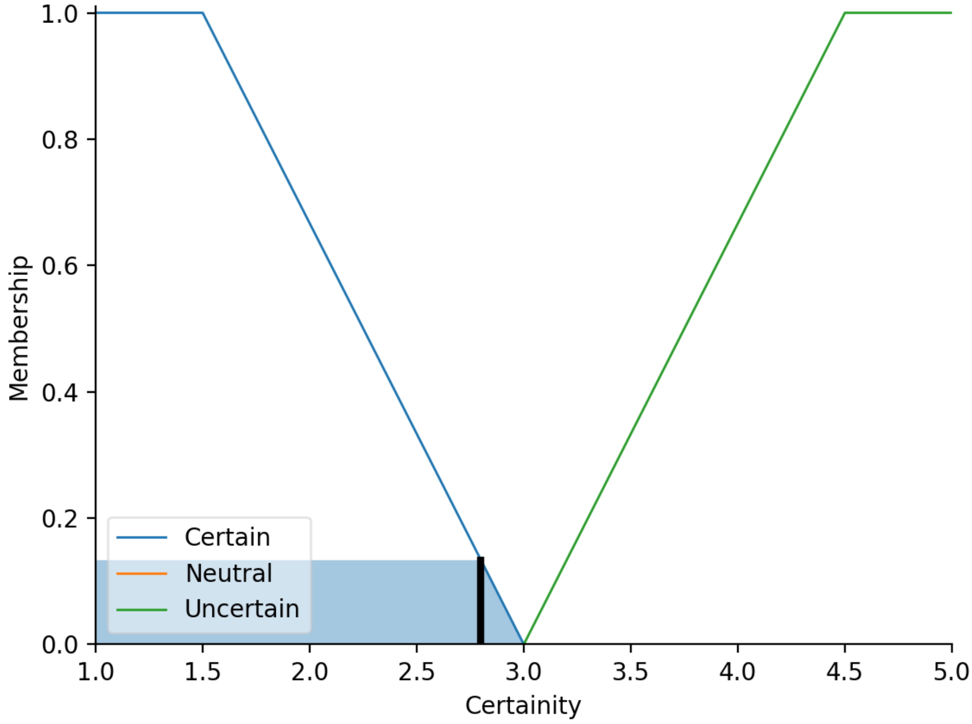


Fig. 51 Degree of Political Consensus for 2018 Turkish Parliamentary and Presidential Elections

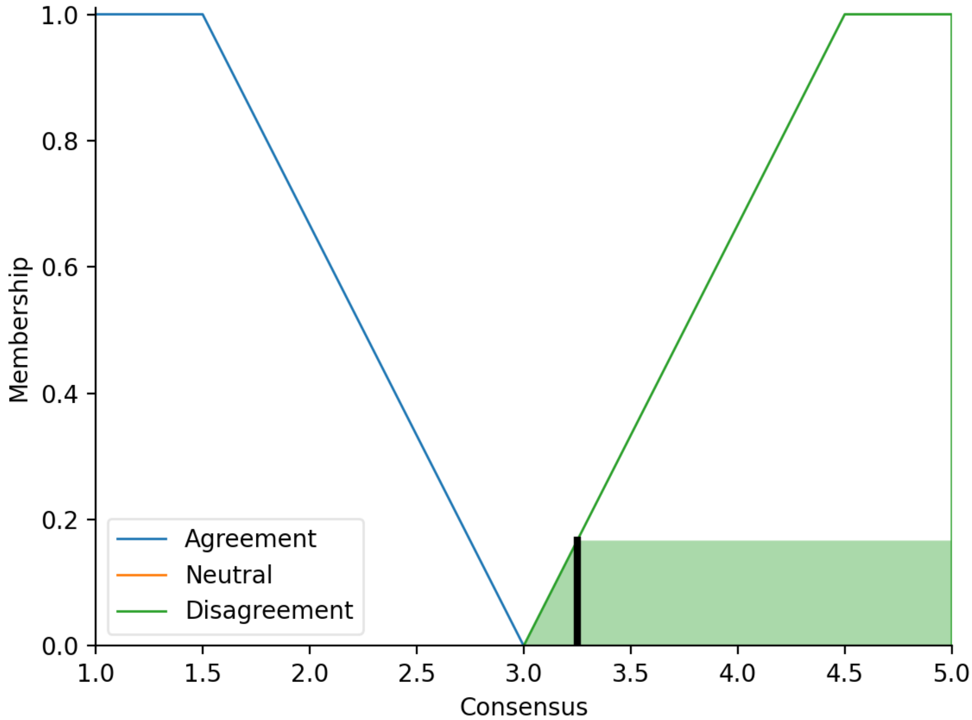


Fig. 52 Dependency Rate between Published News Related to 2018 Turkish Parliamentary and Presidential Elections and USD/TRY Exchange Rate Fluctuations

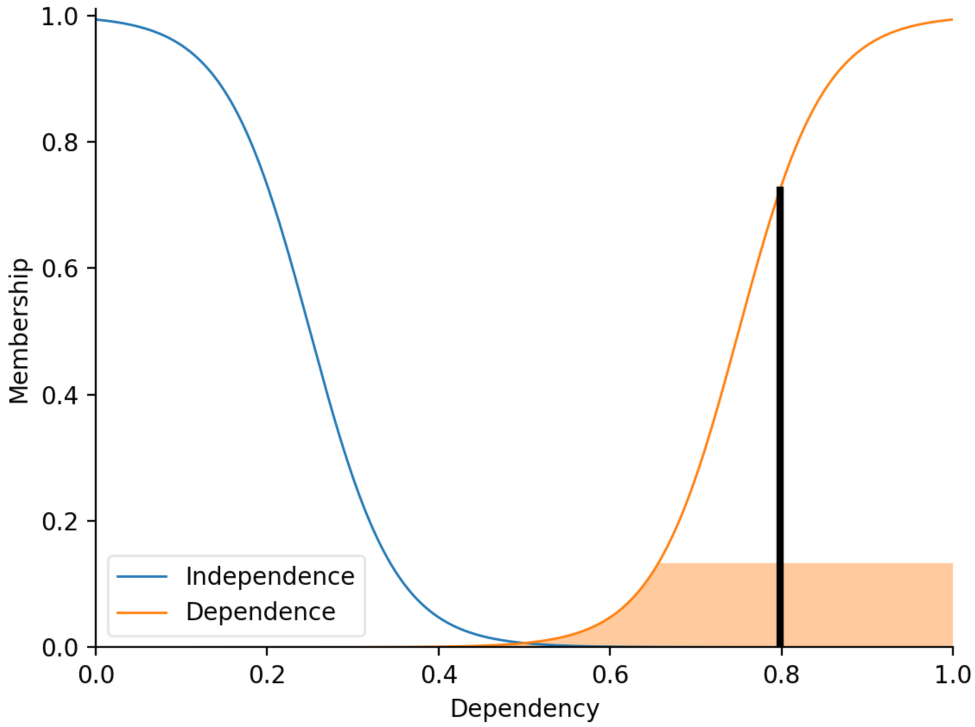


Fig. 53 Illustration of PCR/PCA and PLS Results for 2018 Turkish Parliamentary and Presidential Elections

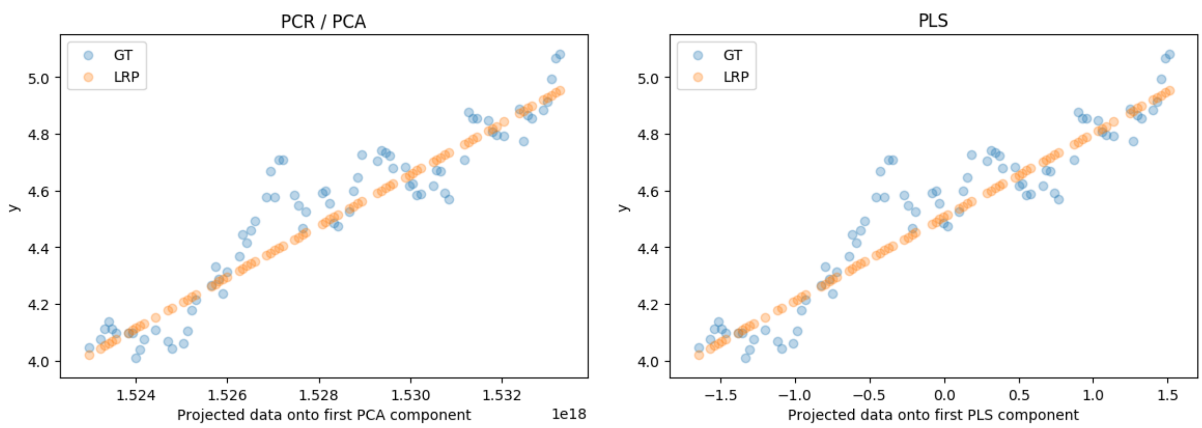


Fig. 54 Grid Search Cross Validation Accuracy Heat Map of Hyperparameter Selection for 2018 Turkish Parliamentary and Presidential Elections

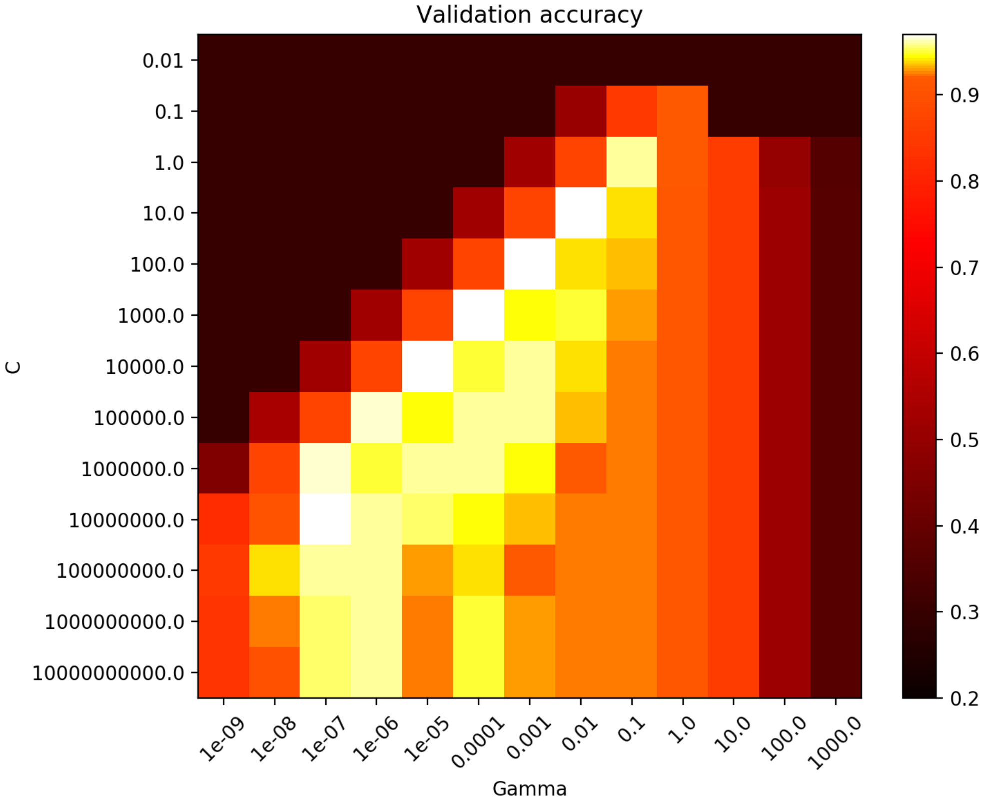


Fig. 55 Layers View for 3D Visualisation of Grid Seach Cross Validation Accuracy of 2018 Turkish Parliamentary and Presidential Elections.

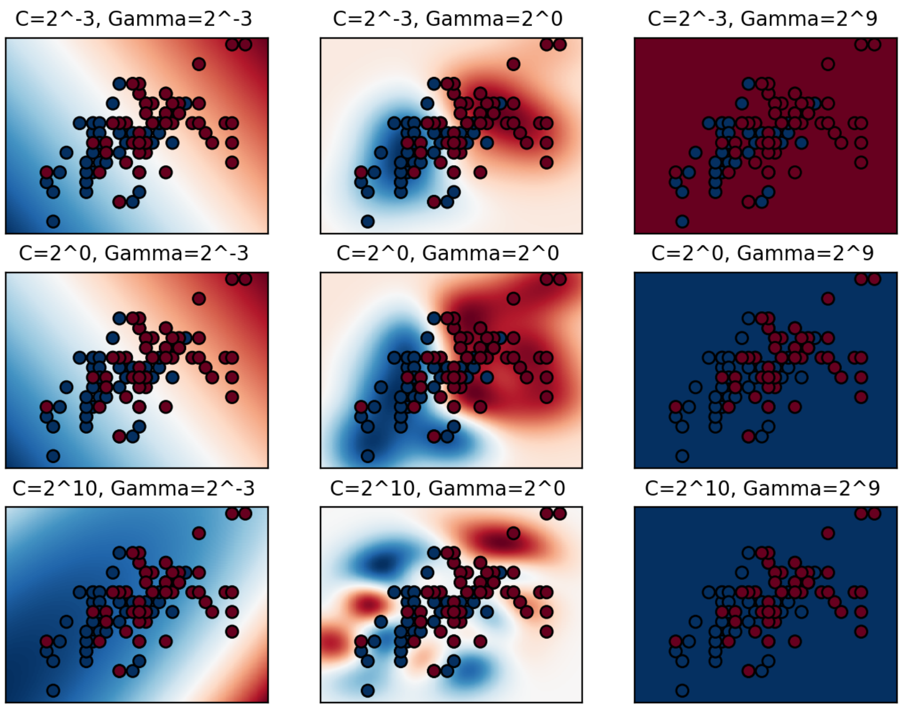
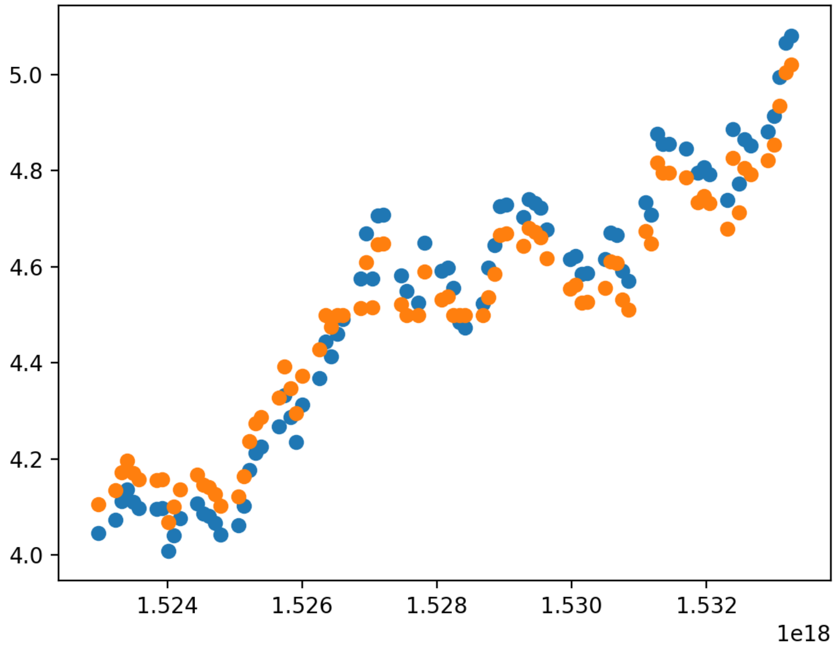


Fig. 56. Actual versus predicted USD/TRY prices obtained by analyzing recently published news and statements regarding on 2018 Turkish Parliamentary and Presidential Elections.



**Blue Points denote Actual USD/TRY Prices

**Orange Points denote Predicted USD/TRY Prices

Fig. 57. Degree of Political Certainty for S-400 Crisis

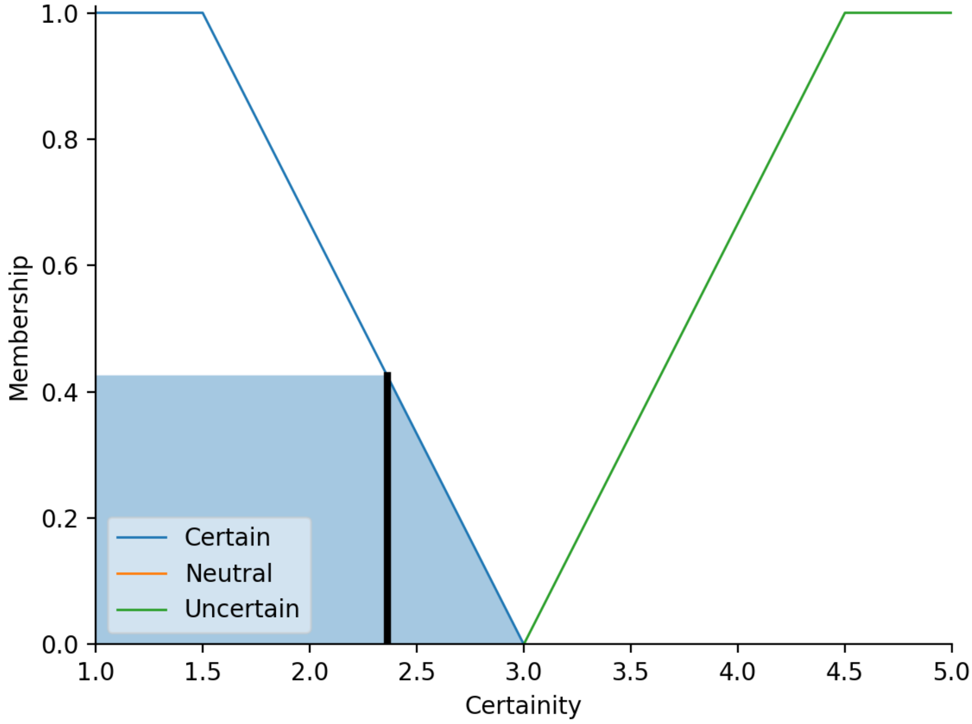


Fig. 58. Degree of Political Consensus for S-400 Crisis

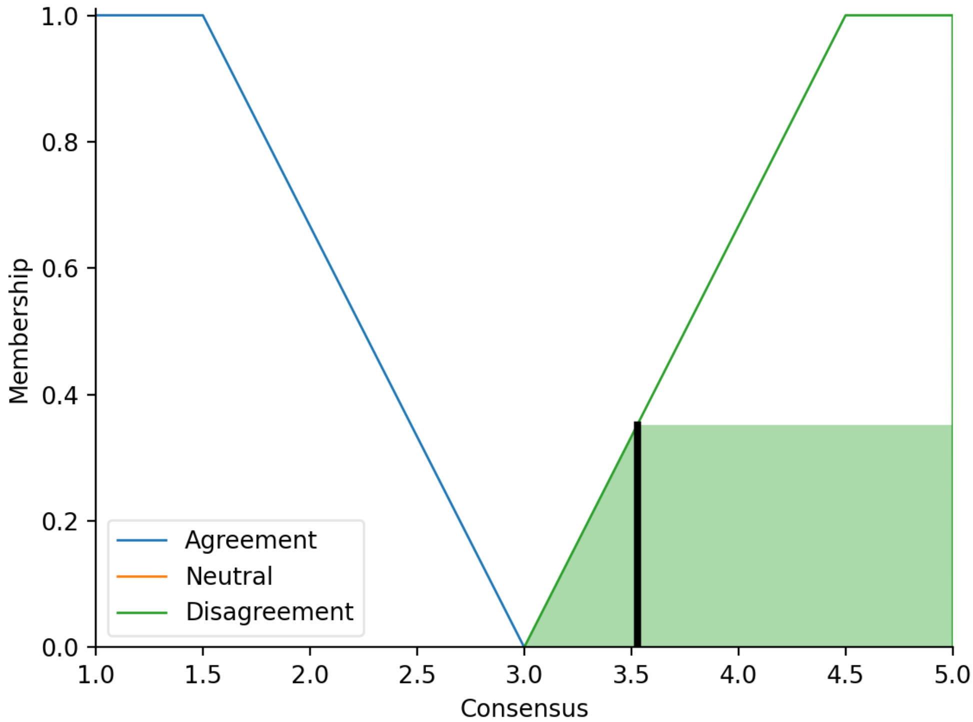


Fig. 59 Dependency Rate between Published News Related to S-400 Crisis and USD/TRY Exchange Rate Fluctuations

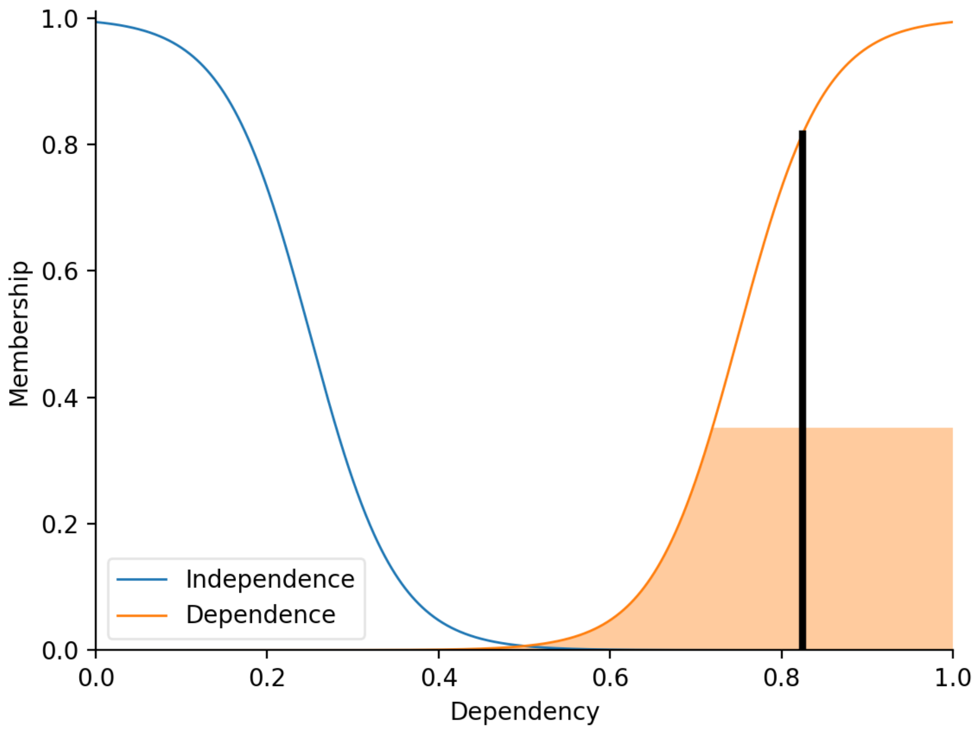


Fig. 60 Illustration of PCR/PCA and PLS Results for S-400 Crisis

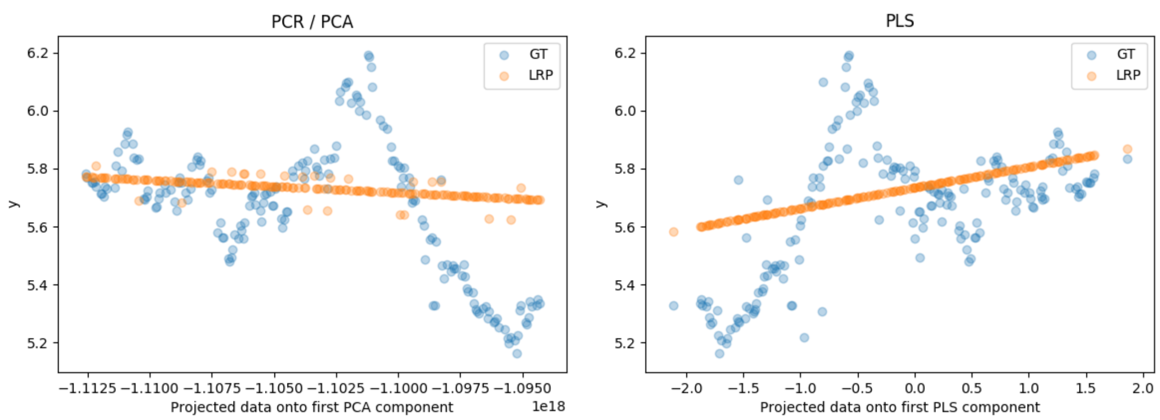


Fig. 61. Grid Search Cross Validation Accuracy Heat Map of Hyperparameter Selection for S-400 Crisis

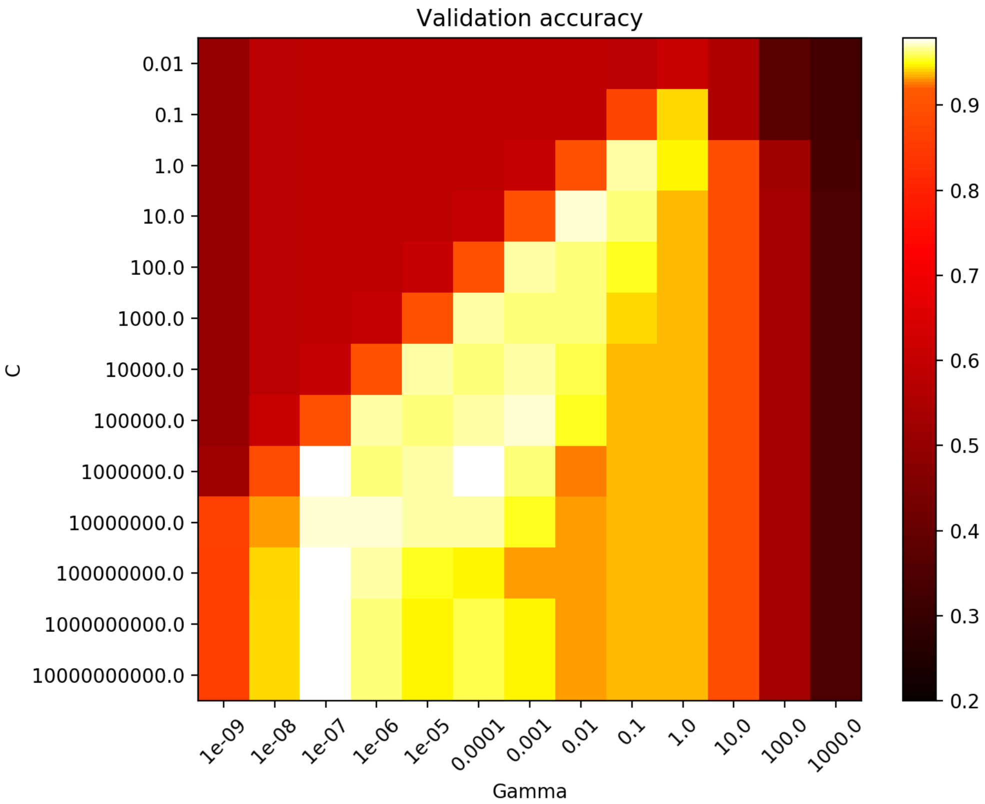


Fig. 62 Layers View for 3D Visualisation of Grid Seach Cross Validation Accuracy of S400 Crisis

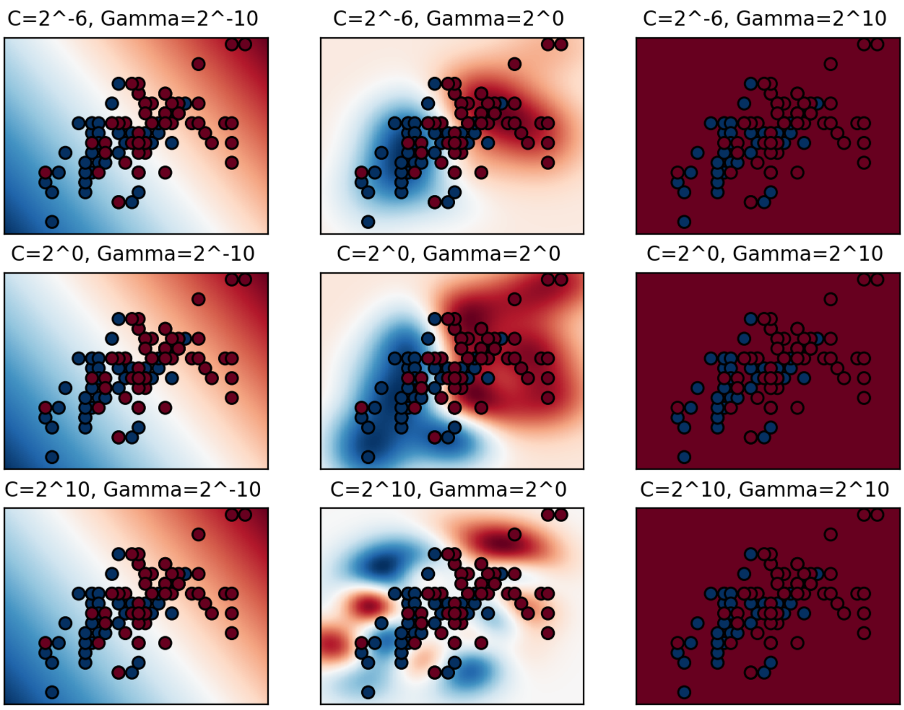
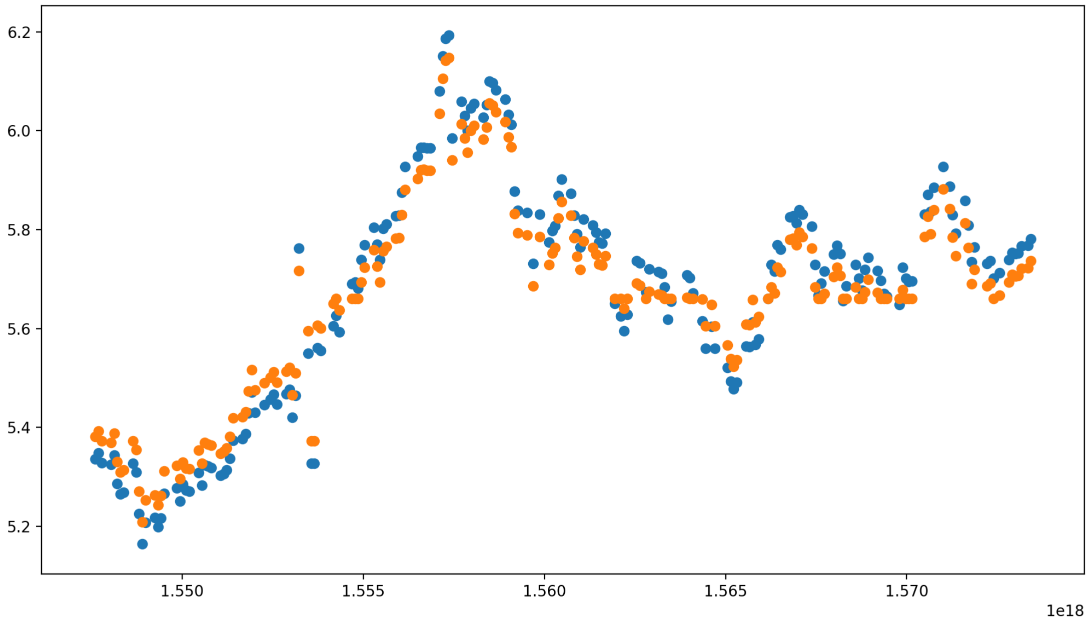


Fig. 63. Actual versus predicted USD/TRY prices obtained by analyzing recently published news and statements regarding on S-400 Crisis.



**Blue Points denote Actual USD/TRY Prices

**Orange Points denote Predicted USD/TRY Prices

Fig. 64. Degree of Political Certainty for 2019 Istanbul Mayoral Election

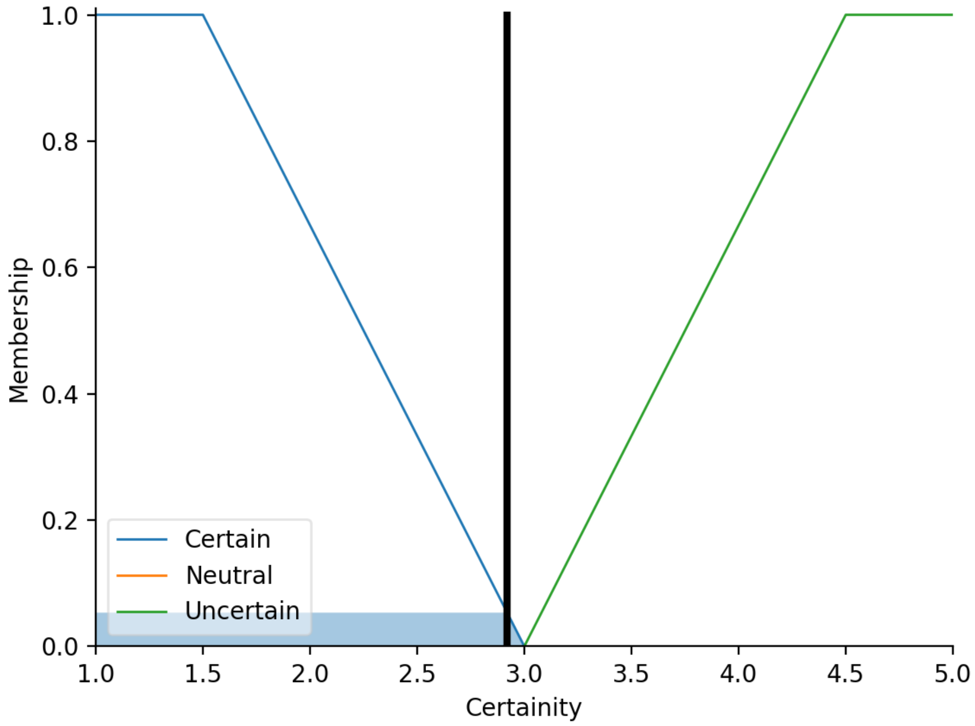


Fig. 65. Degree of Political Consensus for 2019 Istanbul Mayoral Election

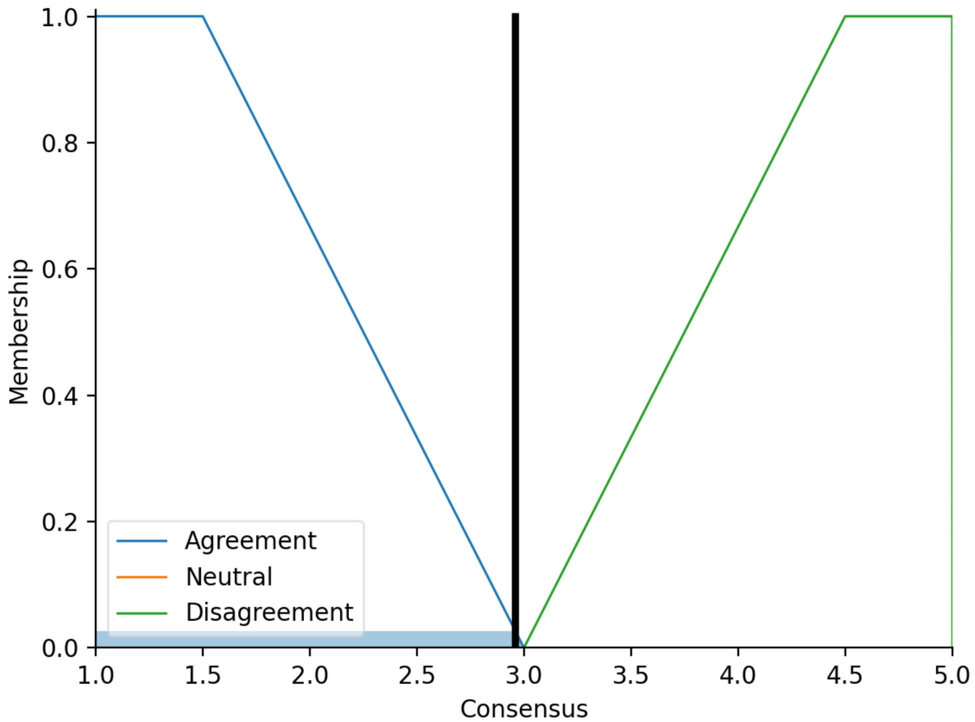


Fig. 66 Dependency Rate between Published News Related to 2019 Istanbul Mayoral Election and USD/TRY Exchange Rate Fluctuations

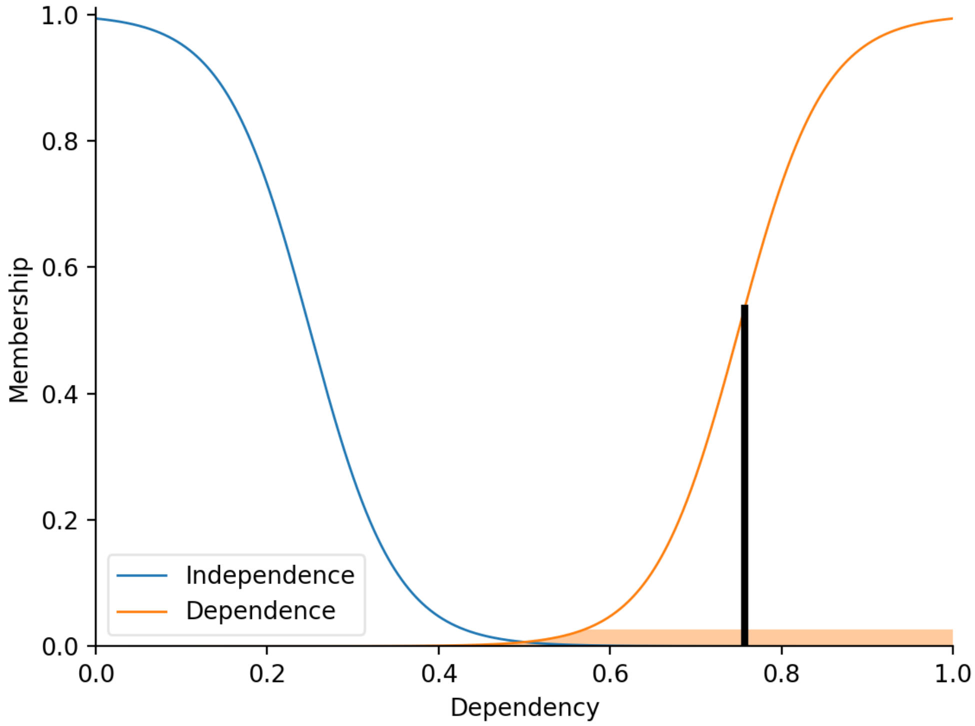


Fig. 67 Illustration of PCR/PCA and PLS Results for 2019 Istanbul Mayoral Election

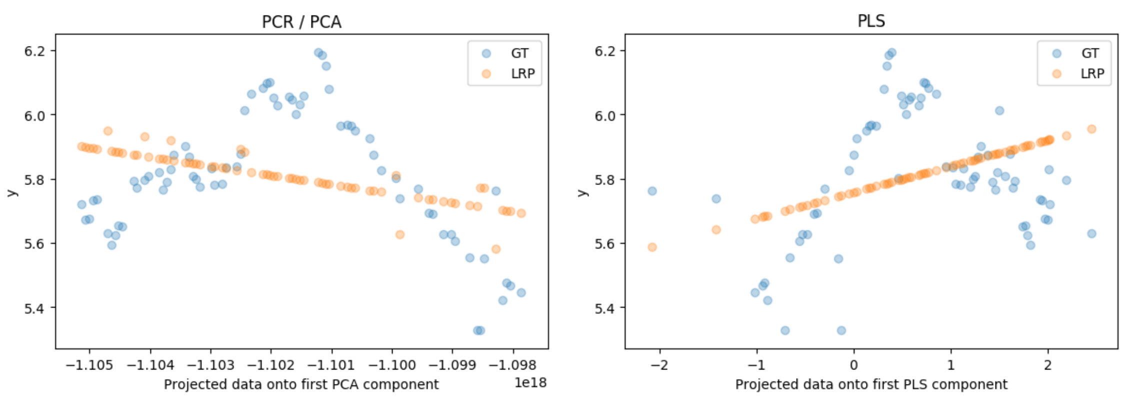


Fig. 68. Grid Search Cross Validation Accuracy Heat Map of Hyperparameter Selection for 2019 Istanbul Mayoral Election

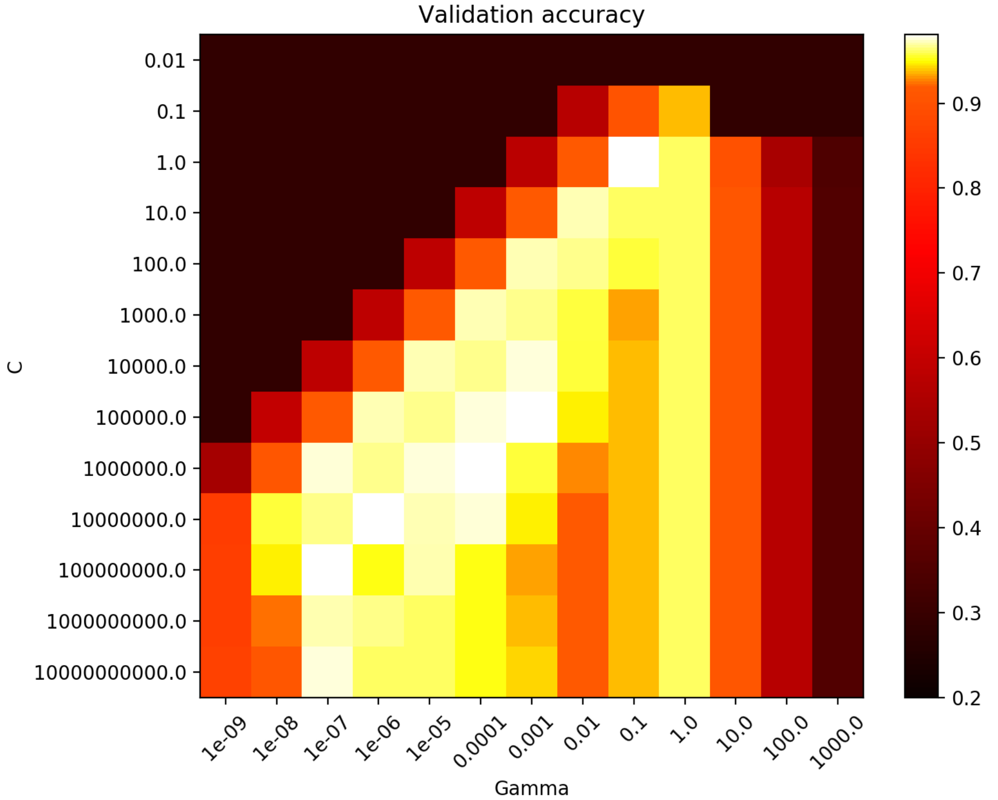


Fig. 69 Layers View for 3D Visualisation of Grid Seach Cross Validation Accuracy of 2019 Istanbul Mayoral Election

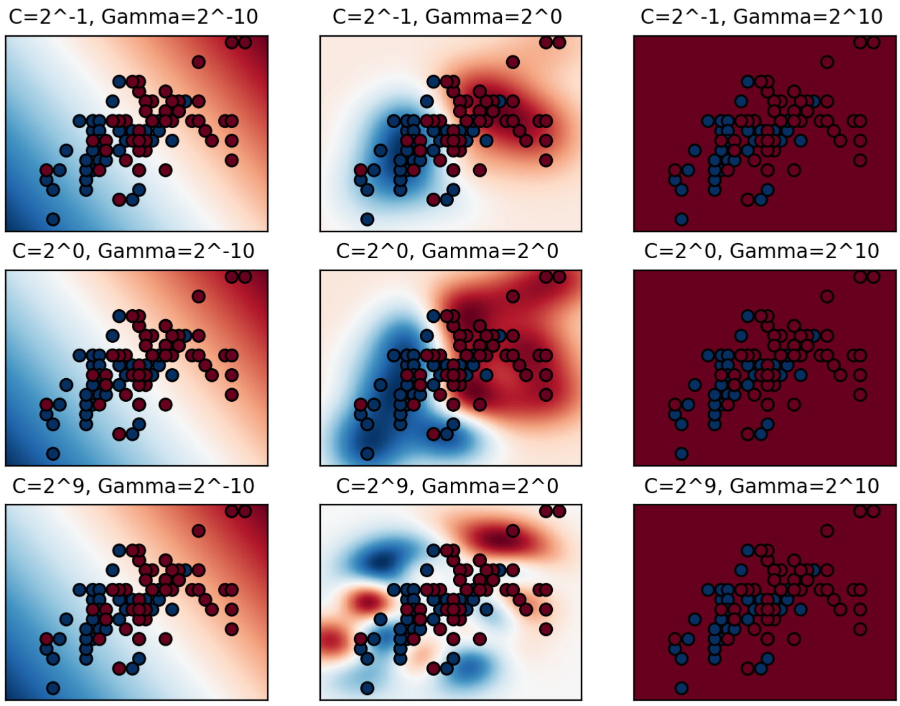
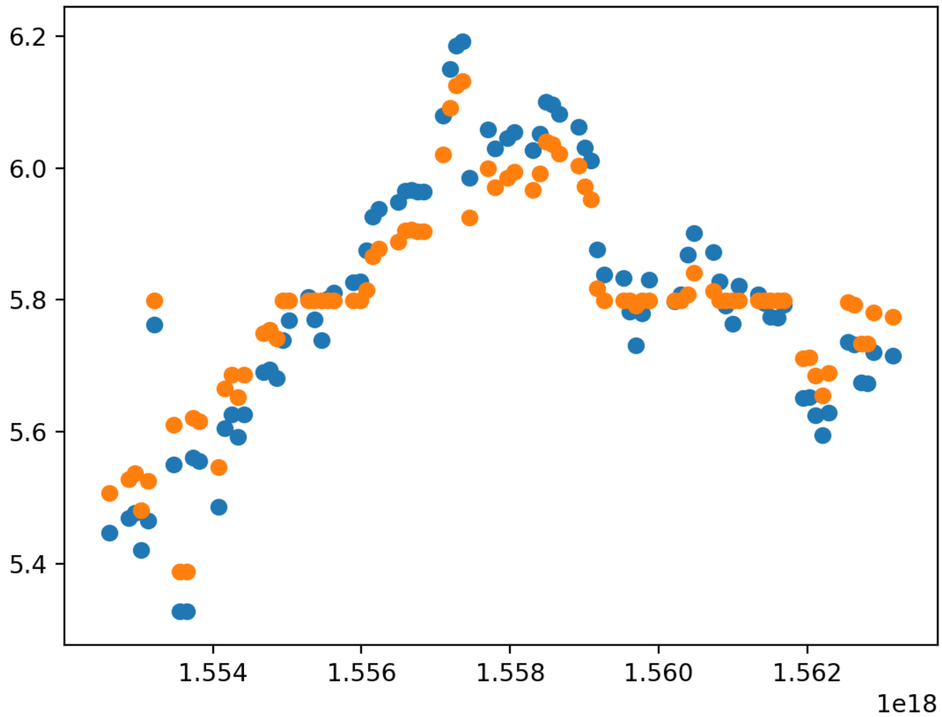


Fig. 70. Actual versus predicted USD/TRY prices obtained by analyzing recently published news and statements regarding on 2019 Istanbul Mayoral Elections.



**Blue Points denote Actual USD/TRY Prices

**Orange Points denote Predicted USD/TRY Prices

Fig. 71. Flowchart of the proposed methodology of Chapter 7

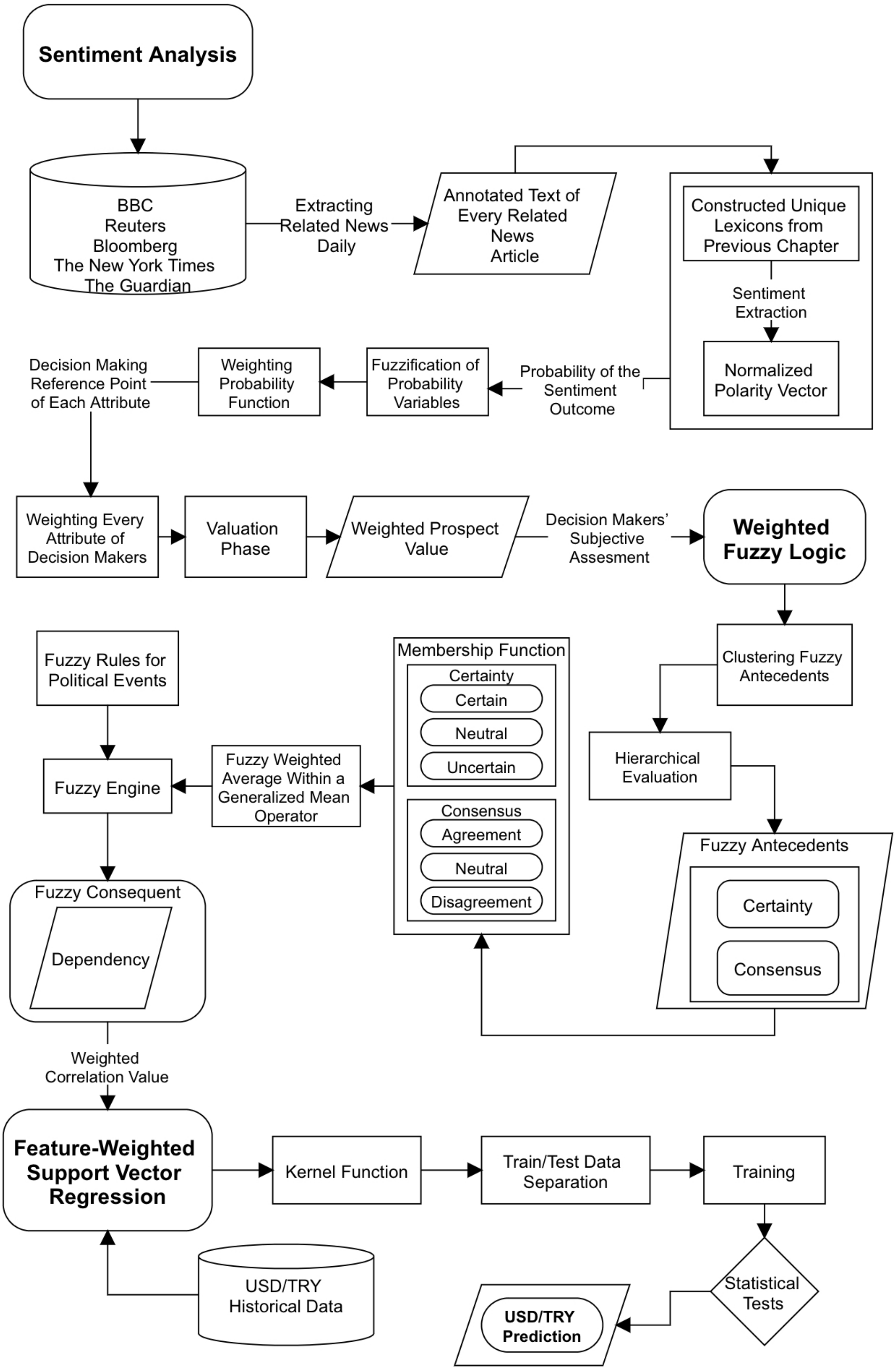


Fig. 72. Degree of Political Certainty for Recently Published Political News

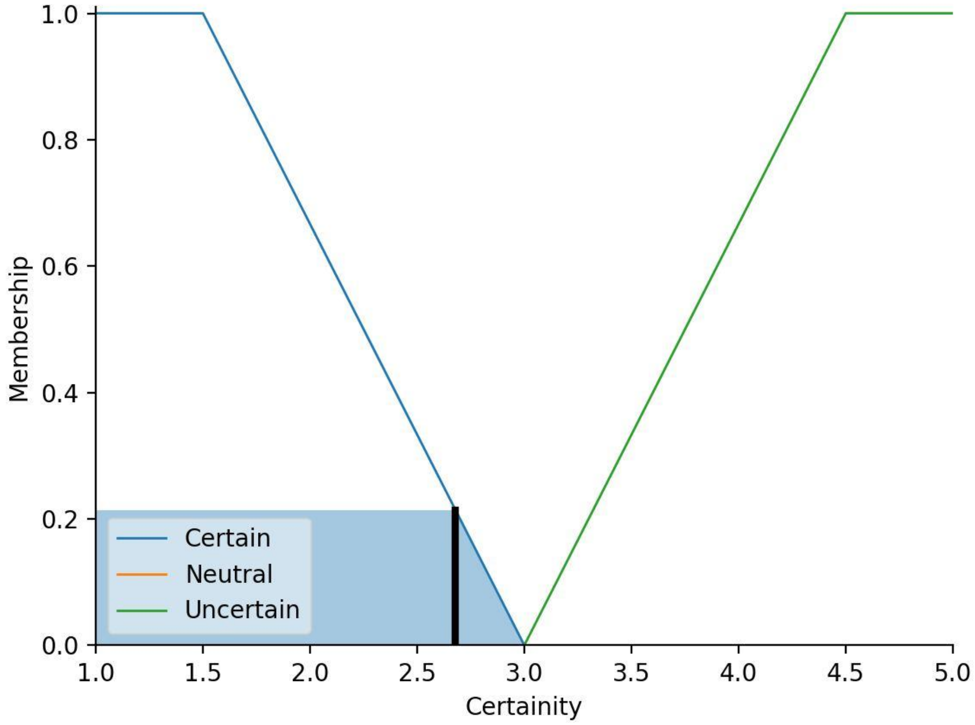


Fig. 73. Degree of Political Consensus for Recently Published Political News

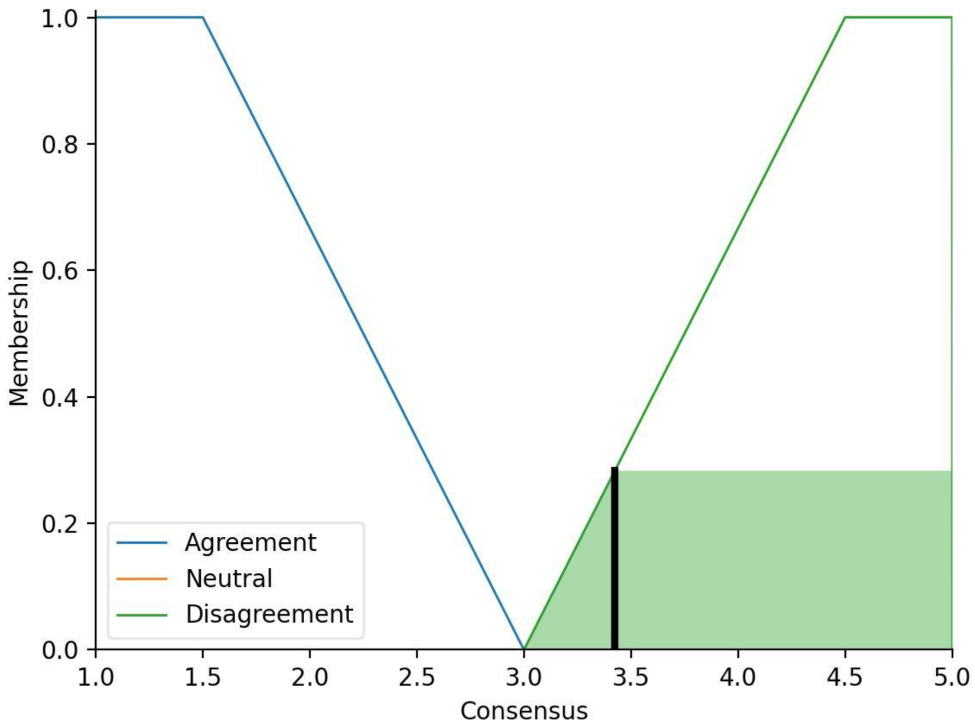


Fig. 74 Dependency Rate between Political News and USD/TRY Exchange Rate Fluctuations

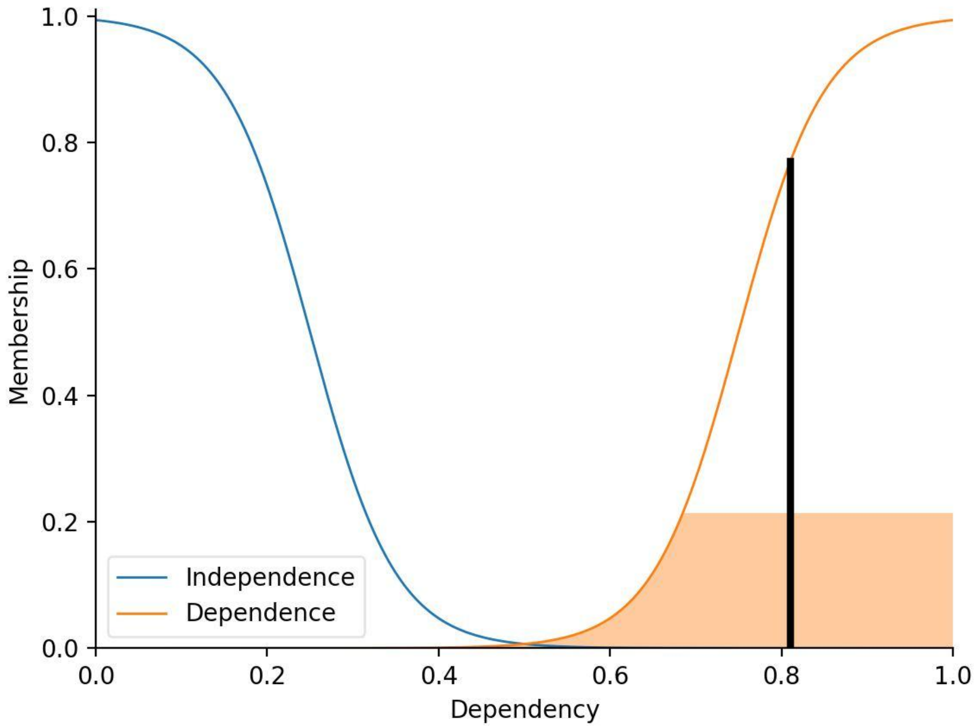


Fig. 75 Illustration of PCR/PCA and PLS Results for Political News

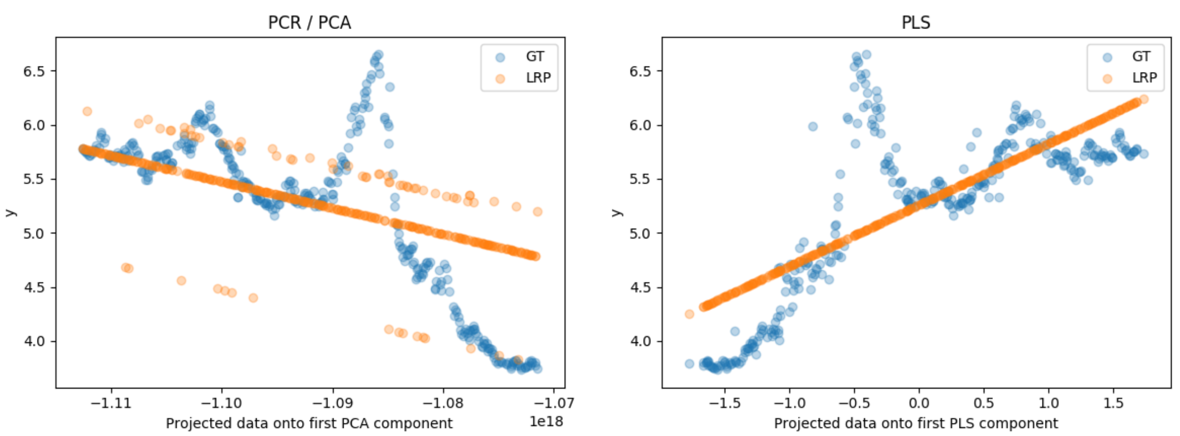


Fig. 76. 2D visualization of USD/TRY Data Parallel to Political Events

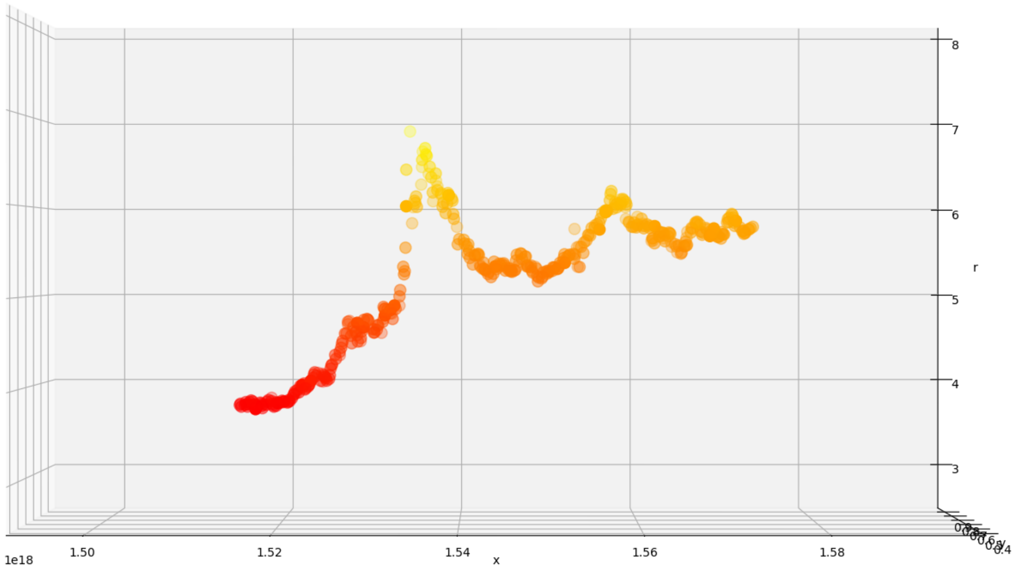


Fig. 77. USD/TRY Data Parallel to Political Events mirrored in Constructed Hyperspace (1)

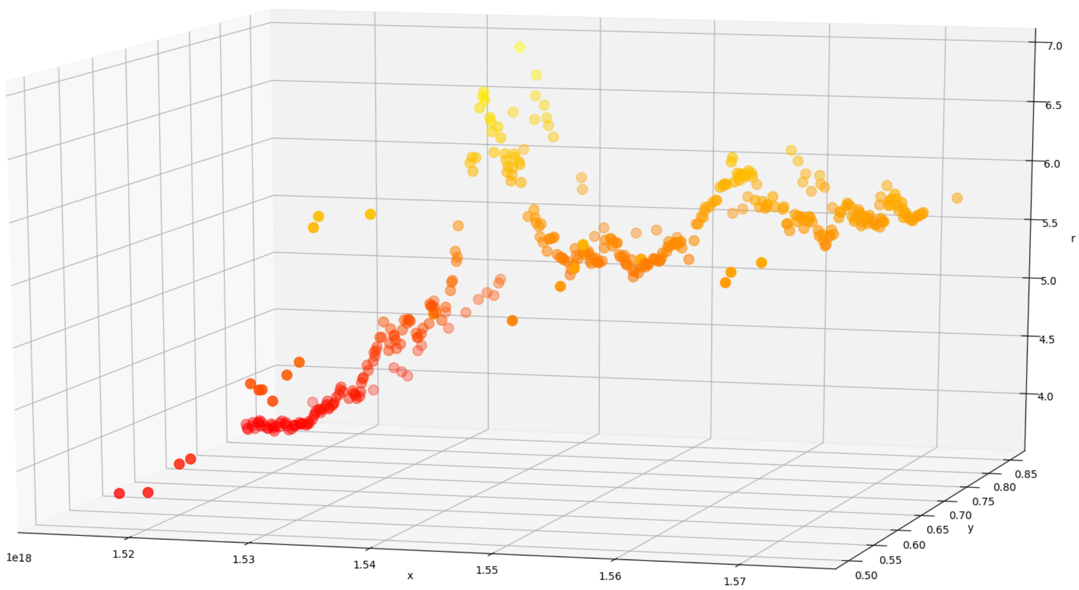


Fig. 78. USD/TRY Data Parallel to Political Events mirrored in Constructed Hyperspace (2)

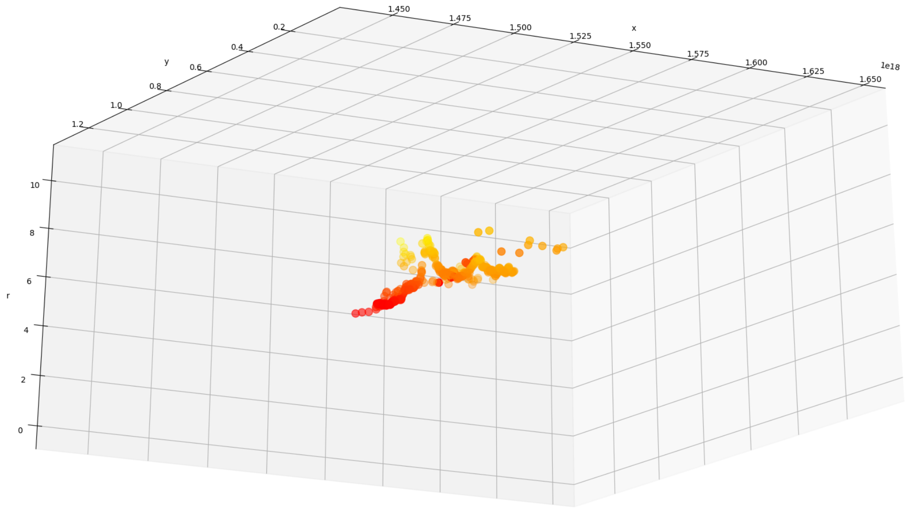


Fig. 79. USD/TRY Data Parallel to Political Events mirrored in Constructed Hyperspace (3)

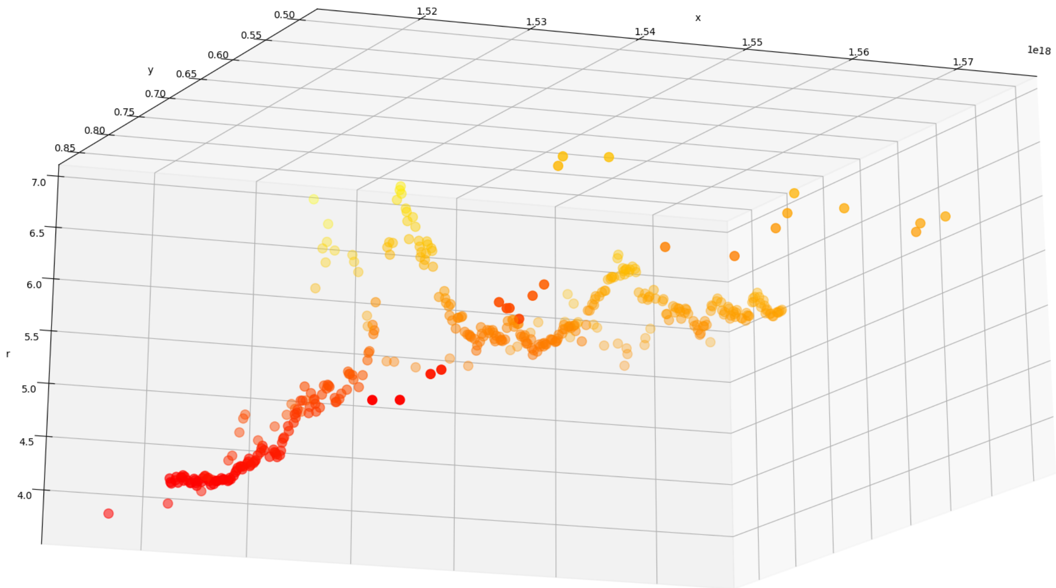


Fig. 80. Grid Search Cross Validation Accuracy Heat Map of Hyperparameter Selection for Political News

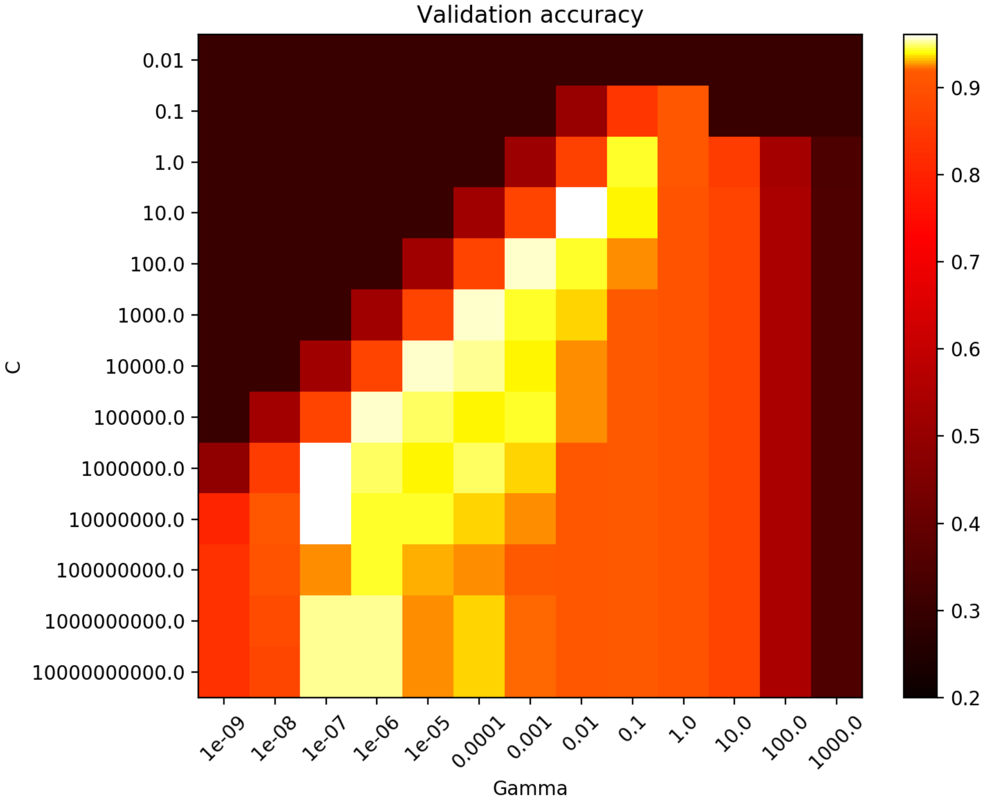


Fig. 81 Layers View for 3D Visualisation of Grid Seach Cross Validation Accuracy of Political News

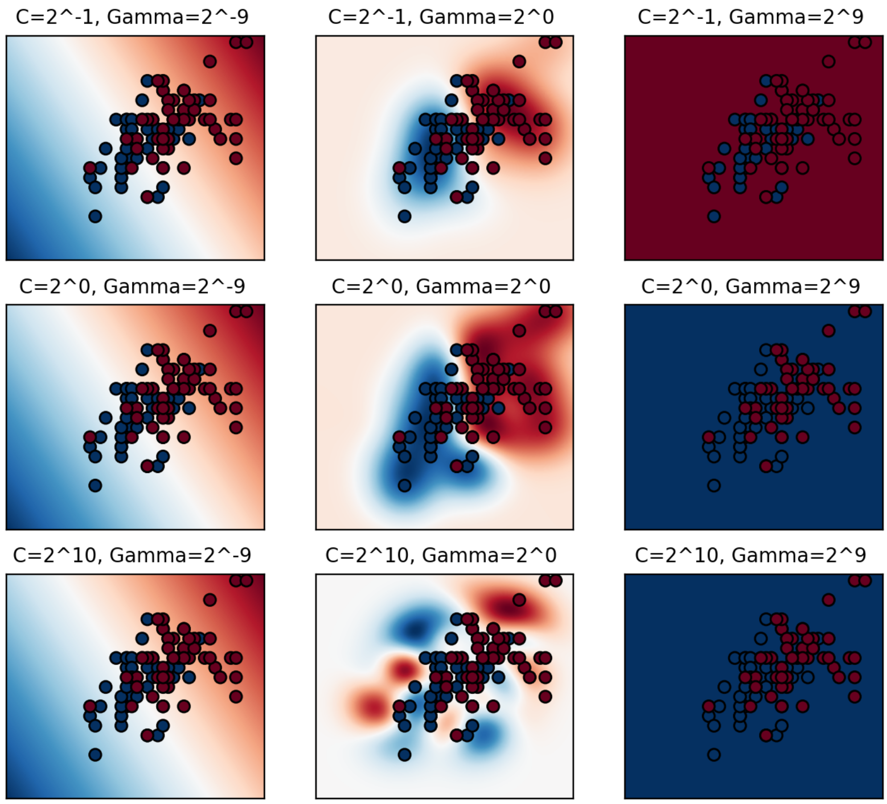
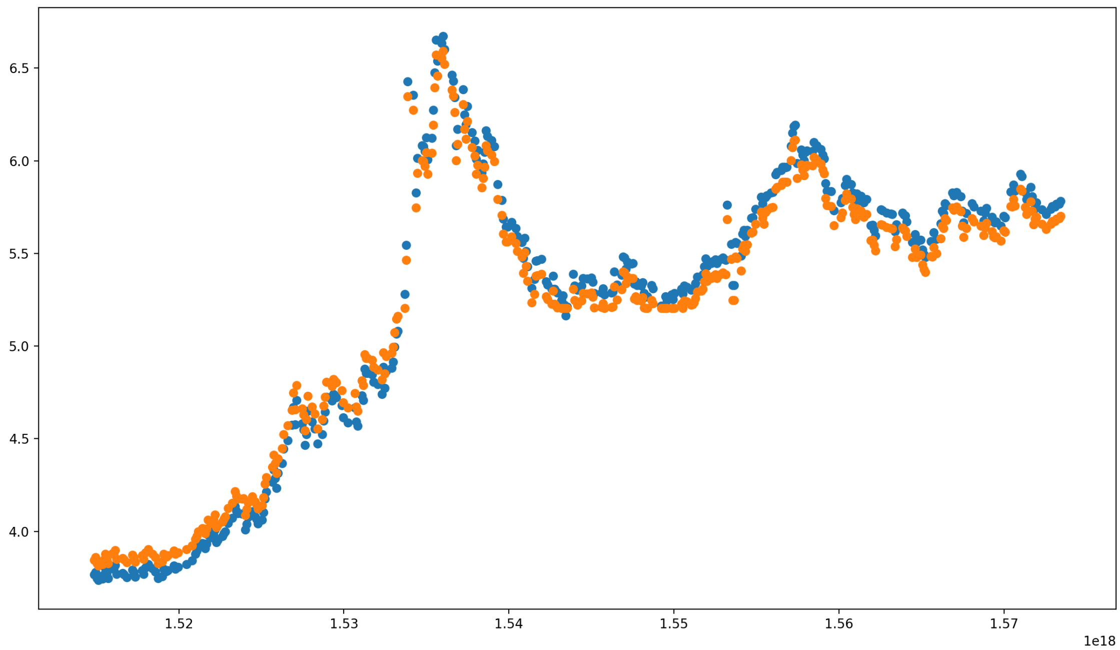


Fig. 82. Actual versus predicted USD/TRY prices obtained by analyzing recently published news and statements regarding on recently published political news.



**Blue Points denote Actual USD/TRY Prices

**Orange Points denote Predicted USD/TRY Prices

Fig. 83. Flowchart of the proposed methodology of Chapter 8

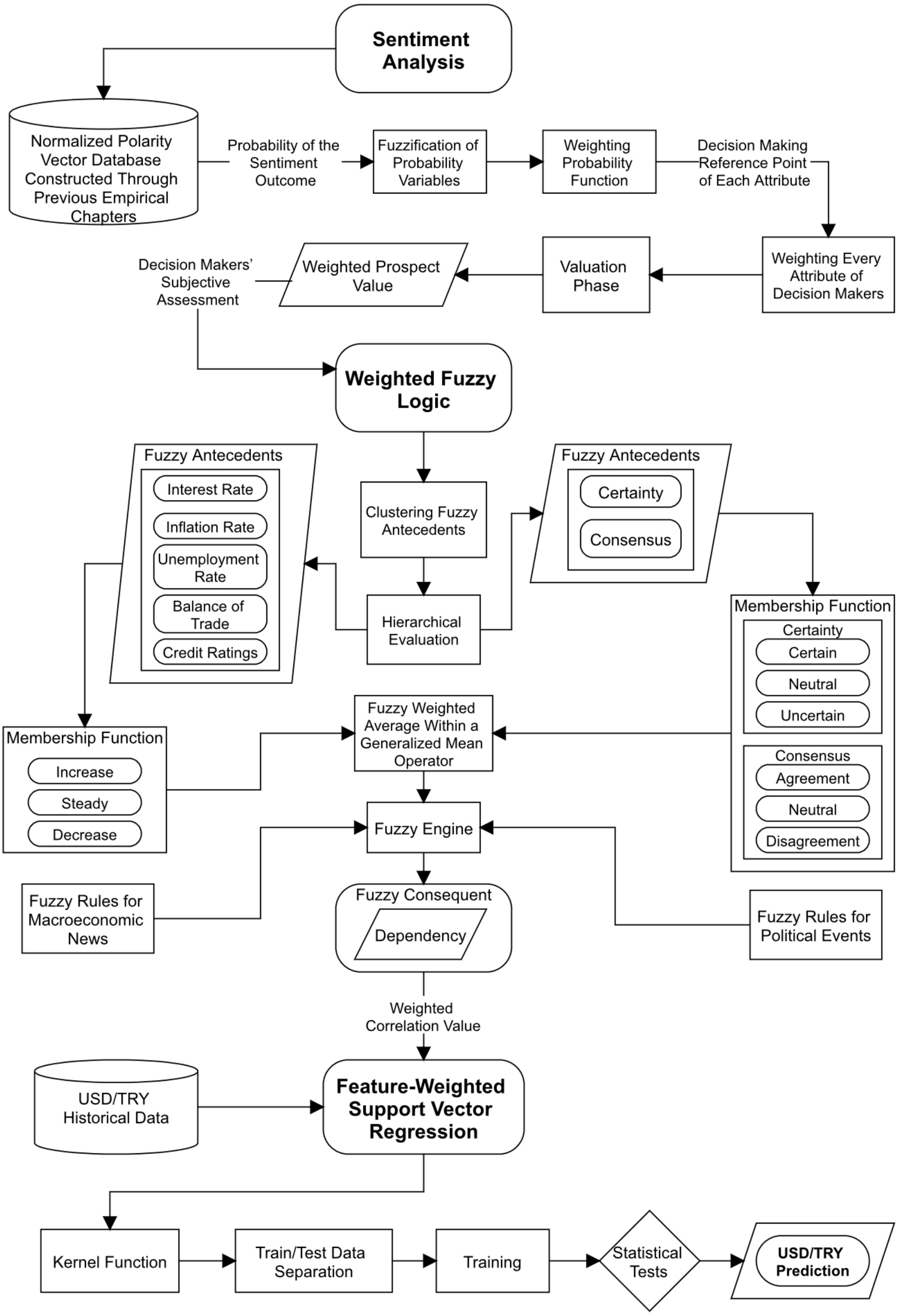


Fig. 84 Dependency Rate between Organic News Feed Simulation and USD/TRY Exchange Rate Fluctuations

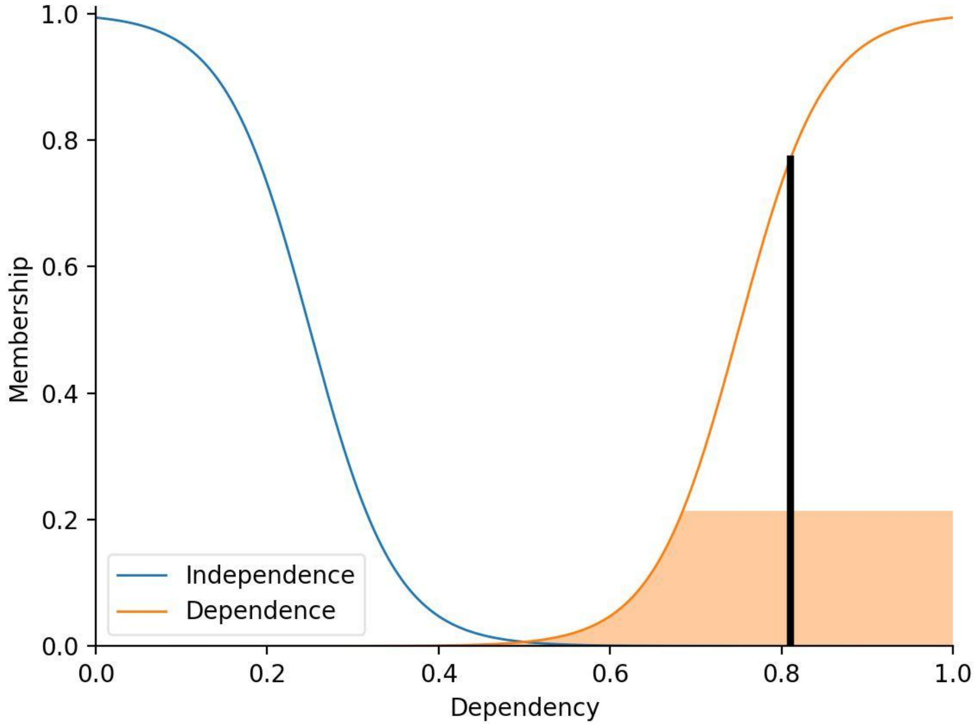


Fig. 85 Illustration of PCR/PCA and PLS Results for Organic News Feed Simulation

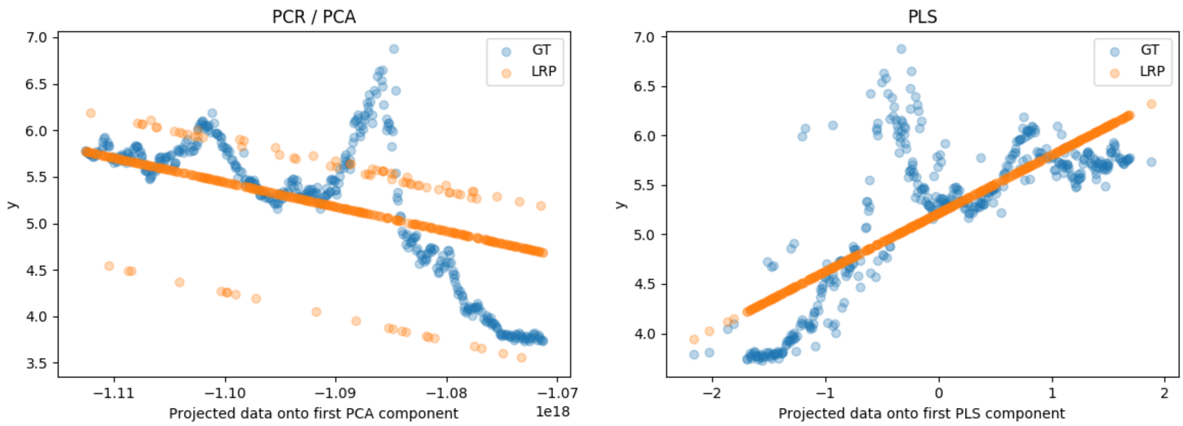


Fig. 86. Daily USD/TRY Data mirrored in Constructed Hyperspace (1)

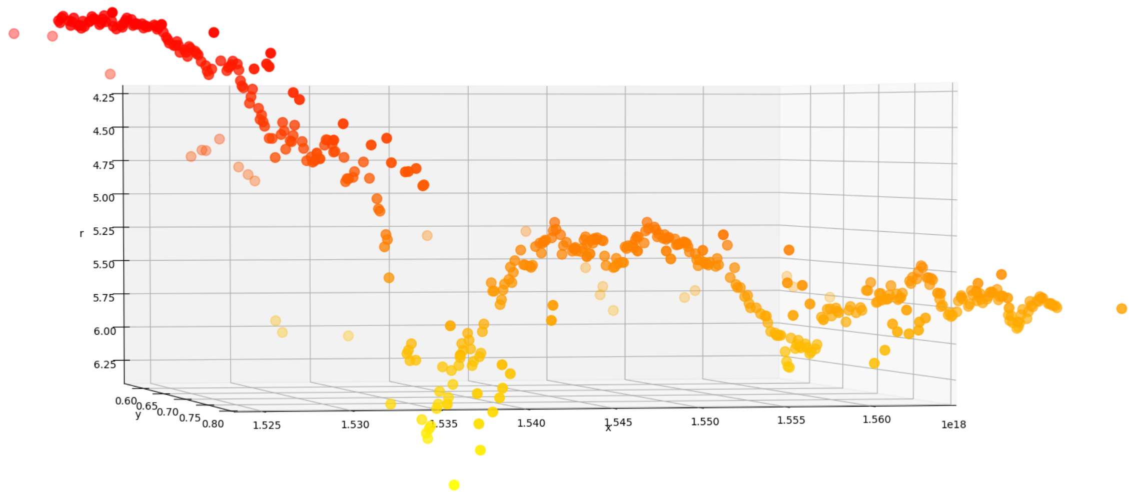


Fig. 87. Daily USD/TRY Data mirrored in Constructed Hyperspace (2)

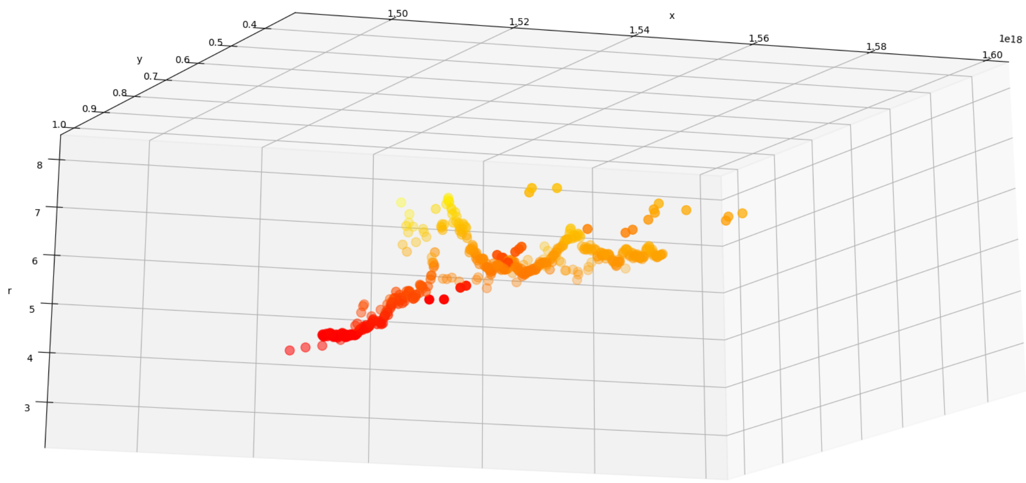


Fig. 88. Daily USD/TRY Data mirrored in Constructed Hyperspace (3)

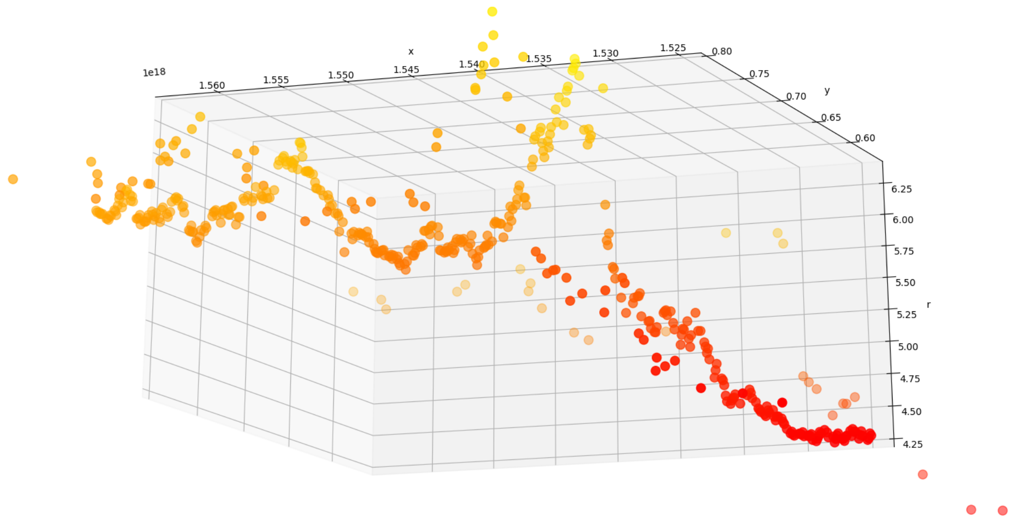


Fig. 90. Grid Search Cross Validation Accuracy Heat Map of Hyperparameter Selection for Organic News Feed Simulation

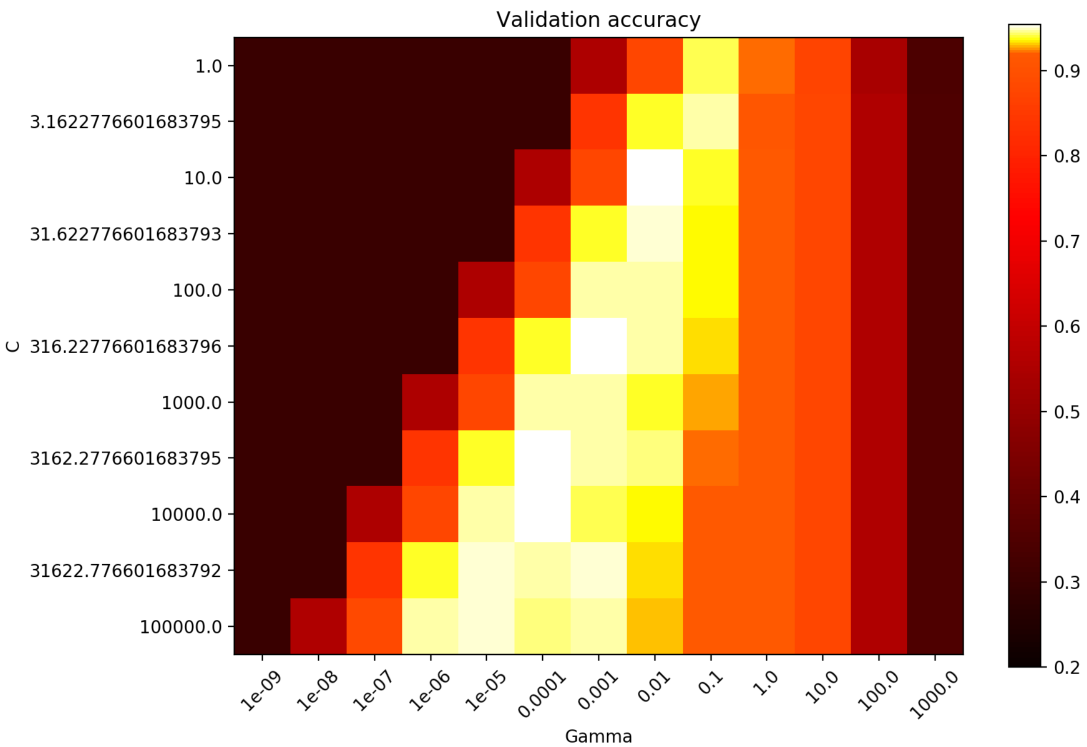


Fig. 91. Layers View for 3D Visualization of Grid Search Cross Validation Accuracy Heat Map for Organic News Feed Simulation

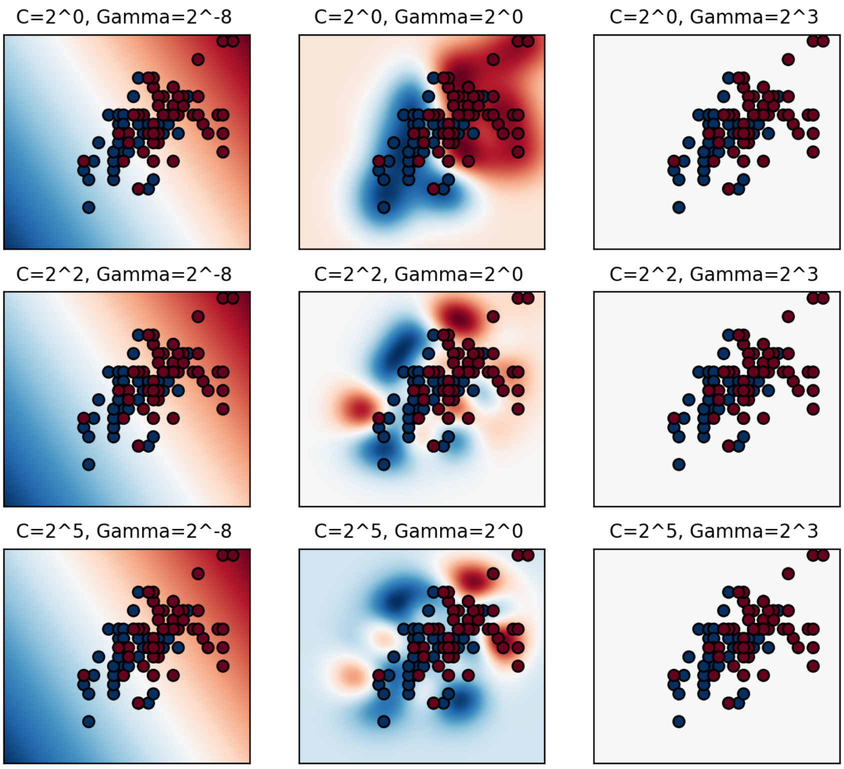
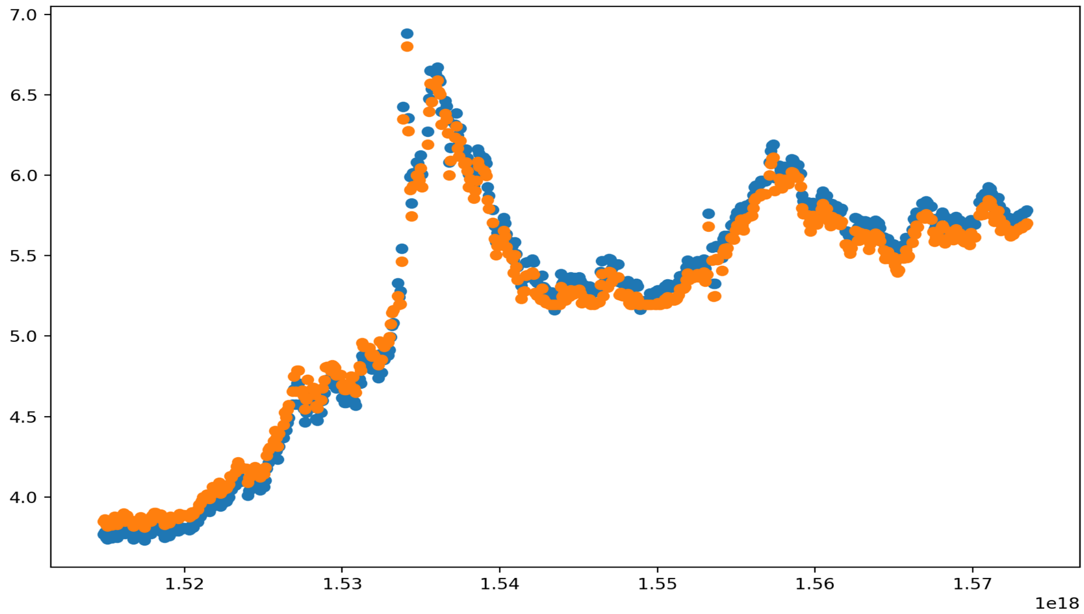


Fig. 92. Actual versus predicted USD/TRY prices obtained by analyzing Organic News Feed regarding on macroeconomics and political tension



**Blue Points denote Actual USD/TRY Prices

**Orange Points denote Predicted USD/TRY Prices

Definitions of Variables and Derivations of Equations

Defining variables of Lexicon

g_i – Positive annotation

h_i – Negative annotation

a_i – Nominated action

$\beta(n)$ - Received environmental response

n – Cycle

$p(i)$ - Score of each existing word in the training set

$p(x)$ - Word polarity factor

m – Total number of words appeared in the training set

R – Score of the examined sentence

L – Coverage speed of the automaton

S – Produced overall value of the sentence based on the polarity vector

w – Number of the words found in the particular sentence

min – Most negative polarity value

max – Most positive polarity value

Derivation of Equations for Lexicon

Polarity Function

$$p_i(n+1) = p_i(n) - \big(1-\beta(n)\big)g_i\big(p(n)\big) + \beta(n)h_i\big(p(n)\big) \;\; if\; a(n) \neq a_i$$

$$p_i(n+1) = p_i(n) + (1-\beta(n))\sum_{j\neq i} g_i\big(p(n)\big) - \beta(n)\sum_{j\neq i} h_j\big(p(n)\big) \;\; if\; a(n) = a_i$$

$$p(i) = \frac{1}{m} \text{ for each word } i, 1 \leq i \leq m$$

$p(x) = p(x) + R/L \cdot p(x)$ for every word found in the sentence

$p(x) = p(x) + S/(m-w) \cdot p(x)$ for every other word, not in the sentence

$S = S_1 + S_2 + \cdots + S_w$ Overall value received by the words appeared in the training set.

Normalized Polarity Vector

$P_N(x) = \frac{p(x)-min}{max-min}$ As polarity vector can only be expressed in positive values, Normalization of polarity vector is necessary to achieve sentiment score in the range of [-1, 1].

Defining variables of Lexicon Prospect Theory

V – Prospect Value

w(p) – Probability weight function

$v(\Delta\chi)$ – Derived value of the subjective feelings of decision maker

χ_0 – Reference point for the subjective feelings of decision maker

$\Delta\chi_i$ - Deviation of the subjective feelings of decision maker. Positive value indicates gains, while negative value indicates losses.

$v(\chi)$ - Power function of the value

α – Concave degree of gains regarding to the power function value of χ. Tendency of decision maker's risk taking is indicated by this value. Higher value indicated higher risk taking tendency.

β – Convex degree of losses regarding to the power function value of χ. Tendency of decision maker's risk taking is indicated by this value. Higher value indicated higher risk taking tendency.

θ – Possibility of losses rather than gains. Loss aversion can be indicated by $\theta > 1$.

p – Probability of the event outcome. Overvaluation of the event outcome can be denoted as $w\,(p) > p$, while undervaluation of the event outcome can be denoted as $w\,(p) < p$.

γ – Coefficient of the risk gain attitude/Curvature parameter. Degree of curvature can be measured as $\gamma > 0$.

δ – Coefficient of the risk loss attitude/ Elevation parameter. Degree of elevation can be measured as $\delta > 0$.

ω - Weight vector

ω_j - Weight of attribute

Derivation of Equations for Prospect Theory

$$V = \sum_{i=1}^{k} (w(p_i)v(\Delta\chi_i))$$

$$v(\chi) = \begin{cases} \chi^{\alpha}, & \chi \geq 0 \\ -\theta(-\chi)^{\beta}, & \chi < 0 \end{cases}$$

Nonlinear weight function of the gains

$$w^{+}(p) = \frac{p^{\gamma}}{(p^{\gamma} + (1-p)^{\gamma})^{1/\gamma}}$$

Nonlinear weight function of the losses

$$w^{-}(p) = \frac{p^{\delta}}{(p^{\delta} + (1-p)^{\delta})^{1/\delta}}$$

Weighting interval probability:

$$\left[w(\bar{p}_j^{L1}), w(\bar{p}_j^{U1})\right]\left[w(\bar{p}_j^{L2}), w(\bar{p}_j^{U2})\right] \cdots \left[w\left(\bar{p}_j^{Ll_i}\right), w\left(\bar{p}_j^{Ul_i}\right)\right]$$
$$(j = 1, 2, \dots, n)$$

Uncertain linguistic variables

$$\left[x_{ij}^{L0}, x_{ij}^{U0}\right]$$

Transforming uncertain linguistic variables into fuzzy numbers

$$\left[a_j^{L0}, a_j^{ML0}, a_j^{MU0}, a_j^{U0}\right]$$

Uncertain linguistic variables on kth status of jth attribute under ith alternative

$$\left[x_{ij}^{Lk}, x_{ij}^{Uk}\right]$$

Fuzzy numbers on kth status of jth attribute under ith alternative

$$\left[a_{ij}^{Lk}, a_{ij}^{MLk}, a_{ij}^{MUk}, a_{ij}^{Uk}\right]$$

Prospect value function for fuzzy numbers

$$z_{ij}^{k} = \left[v\left(a_{ij}^{Lk} - a_j^{U0}\right), v\left(a_{ij}^{MLk} - a_j^{MU0}\right), v\left(a_{ij}^{MUk} - a_j^{ML0}\right), v\left(a_{ij}^{Uk} - a_j^{L0}\right)\right]$$

Prospect Function on jth attribute under the ith alternative:

$$\begin{aligned} z_{ij} &= \sum_{k=1}^{l_i}\left(w_j^k z_{ij}^k\right) = \left[z_{ij}^L, z_{ij}^{ML}, z_{ij}^{MU}, z_{ij}^U\right] \\ &= \left[\sum_{k=1}^{l_i}\left(w\left(\bar{p}_j^{Lk}\right)v\left(a_{ij}^{Lk} - a_j^{U0}\right)\right),\right. \\ &\quad \sum_{k=1}^{l_i}\left(w\left(\bar{p}_j^{Lk}\right)v\left(a_{ij}^{MLk} - a_j^{MU0}\right)\right), \\ &\quad \sum_{k=1}^{l_i}\left(w\left(\bar{p}_j^{Uk}\right)v\left(a_{ij}^{MUk} - a_j^{ML0}\right)\right), \\ &\quad \left.\sum_{k=1}^{l_i}\left(w\left(\bar{p}_j^{Uk}\right)v\left(a_{ij}^{Uk} - a_j^{L0}\right)\right)\right] \end{aligned}$$

Weighted Prospect Function Value

$$z_i = \sum_{j=1}^{n} \left(\omega_j \times z_{ij}\right) = [z_i^L, z_i^{ML}, z_i^{MU}, z_i^U]$$

$$= \left[\sum_{j=1}^{n}\left(\omega_j \times \sum_{k=1}^{l_i}\left(w\left(\bar{p}_j^{Lk}\right)v\left(a_{ij}^{Lk} - a_j^{U0}\right)\right)\right),\right.$$

$$\sum_{j=1}^{n}\left(\omega_j \times \sum_{k=1}^{l_i}\left(w\left(\bar{p}_j^{Lk}\right)v\left(a_{ij}^{MLk} - a_j^{MU0}\right)\right)\right),$$

$$\sum_{j=1}^{n}\left(\omega_j \times \sum_{k=1}^{l_i}\left(w\left(\bar{p}_j^{Uk}\right)v\left(a_{ij}^{MUk} - a_j^{ML0}\right)\right)\right),$$

$$\left.\sum_{j=1}^{n}\left(\omega_j \times \sum_{k=1}^{l_i}\left(w\left(\bar{p}_j^{Uk}\right)v\left(a_{ij}^{Uk} - a_j^{L0}\right)\right)\right)\right]$$

Defining variables, and Derivation of Equations for Fuzzy Logic

$\mu(x) \rightarrow [0,1]$ - Membership function

$\mu_A(x): X \longrightarrow [0, 1]$ – Partial membership function for a fuzzy set.

μ - Degree of membership

$\mu_A(x)$ - Membership function for x in fuzzy set *A*.

X – Universal set.

Membership Functions;

Sigmoidal membership function

Right shoulder sigmoidal function

$$\mu = \frac{1}{1 + e^{-\beta(x-\alpha)}}$$

Left shoulder sigmoidal function

$$\mu = \frac{1}{1 + e^{\beta(x-\alpha)}}$$

β – Steepness of the sigmoid

α – Crossover point.

Trapezoid membership function

$$\mu(x) = \begin{cases} 0 & if\ x < a \\ \frac{x-a}{m-a} & if\ x \in [a,m] \\ 1 & if\ x \in [m,n] \\ \frac{b-x}{b-n} & if\ x \in [n,b] \\ 0 & if\ x > b \end{cases}$$

a - Upper bound

b - Lower Bound

m and n – Coordinates of tolerance

Gaussian membership function

$$\mu_{S^i}(x) = \exp\left(-\frac{(c_i - x)^2}{2\sigma_i^2}\right)$$

c_i – center of the ith fuzzy set S^i.

σ – width of the ith fuzzy set S^i.

$\mu_B(x) = 1 - \mu_A(x)$ - Complement operator of the fuzzy set, which is theoretically regarded as operator NOT.

$MIN\ (\mu_A(x), \mu_B(x))$ - Intersection operator of the fuzzy set, which is theoretically regarded as operator MIN that construed as logical AND.

$MAX\ (\mu_A(x), \mu_B(x))$ – Union operator of the fuzzy set, which is theoretically regarded as operator MAX that construed as logical OR.

$\mu_{A\wedge B} = MIN\ (\mu_A, \mu_B)$ - Lower intersection point of the membership function.

$\mu_{A\vee B} = MAX\ (\mu_A, \mu_B)$ – Upper intersection point of the membership function.

$\mu_R\ :\ X_1 \times \cdots \times X_m \rightarrow [0, 1]$ - Simplification of *n*-tuples (relationship matrix).

μ_R - Membership function of a multidimensional fuzzy set

X_i - Universes of discourse

$X_1 \times \cdots \times X_m$ – Product space

$R\ (X, Y)$ / R (in simple form) – Fuzzy relation between two sets

Membership function of the MAX-MIN composition of two fuzzy relations

$$R_1 \circ R_2(x, z) = \left\{\left[(x, z), \underset{y}{max} \left\{min\left\{\mu_{R_1}(x, y), \mu_{R_2}(y, z)\right\}\right\}\right] \middle| \ x \in X, y \in Y, z \in Z\right\}$$

Fuzzy inference

$$H = MIN(\mu_L(x_1'), \mu_M(x_2'))$$

H – Fuzzy inference

x' - Fuzzified crisp value

$\mu_{L'}$ - Transformation of fuzzified crisp value into a membership vector.

Defuzzification formula

$$y' = \frac{\sum_{i=1}^{N} y_i\, H_i}{\sum_{i=1}^{N} H_i}$$

N – Number of membership functions

Defining variables, and Derivation of Equations for Fuzzy Weighted Average

Integrating the fuzzy variables of hierarchical evaluation that forms the operators into fuzzy set theory:

w_i – Relative weighting criteria

x_i – Scoring criteria

f - mapping of $X_1 \times X_2 \times \ldots X_n \times Z_1 \times Z_2 \times \ldots \times Z_n$

$A_1, A_2, \ldots A_n$ - fuzzy weighted average numbers

$W_1, W_2, \ldots W_n$ - weights of fuzzy numbers

$X_1, X_2, \ldots X_n$ - the universe of fuzzy numbers

w_i'- normalized weight

$$y = f(x_1, x_2, \ldots, x_n, w_1, w_2, \ldots, w_n) = \left(\frac{w_1x_1 + w_2x_2, \ldots, +w_nx_n}{w_1 + w_2, \ldots, +w_n}\right)$$
$$= (w_1'x_1 + w_2'x_2, \ldots, +w_n'x_n)$$

$\min(x_1, x_2 \ldots, x_n) \leq y \leq \max(x_1, x_2 \ldots, x_n)$ - Weighted generalized mean operator.

$\oplus$ - aggregation operator

w_i'- normalized weight within a hierarchy

x_i - score for all fuzzy numbers

fuzzy weighted average function integrating the generalized mean operator;

$$= (w_1'x_1 \oplus w_2'x_2, \ldots, \oplus w_n'x_n)$$
$$y = f(x_1, x_2, \ldots, x_n, w_1', w_2', \ldots, w_n', p) = w_1'x_1^p + w_2'x_2^p, \ldots, +w_n'x_n^p)^{1/p}$$
$$y = f(x_1, x_2, \ldots, x_n, w_1, w_2, \ldots, w_n, p) = \left(\frac{w_1x_1^p + w_2x_2^p, \ldots, +w_nx_n^p}{w_1 + w_2, \ldots, +w_n}\right)^{1/p}$$
$$\alpha_i \leq x_i \leq b_i, c_i \leq w_i \leq d_i$$

$$p = -\infty, \text{the minimum operator}$$
$$p = -1, \text{the harmonic mean operator}$$
$$p = 0, \text{the geometric mean operator}$$
$$p = +1, \text{the arithmetic mean operator}$$
$$p = +\infty, \text{the maximum operator}$$

Weight of *p* on generalized mean function can be measured by calculating the summation on the inequality, and dividing it by $\sum_{i=1}^{n} w_i$;

$$\frac{\sum_{i=1}^{n} w_i \alpha_i^p}{\sum_{i=1}^{n} w_i} \leq \frac{\sum_{i=1}^{n} w_i x_i^p}{\sum_{i=1}^{n} w_i} \leq \frac{\sum_{i=1}^{n} w_i b_i^p}{\sum_{i=1}^{n} w_i}$$
$$\Rightarrow \left(\frac{\sum_{i=1}^{n} w_i \alpha_i^p}{\sum_{i=1}^{n} w_i}\right)^{1/p} \leq \left(\frac{\sum_{i=1}^{n} w_i x_i^p}{\sum_{i=1}^{n} w_i}\right)^{1/p} \leq \left(\frac{\sum_{i=1}^{n} w_i b_i^p}{\sum_{i=1}^{n} w_i}\right)^{1/p}$$

Lower and upper bounds for each α-cut interval is expressed as $L_{p\geq 0}, U_{p\geq 0}$ while $p \geqq 0$, where $L_{p\geq 0} = Min_{p\geq 0} f_L$, and $U_{p\geq 0} = Max_{p\geq 0} f_U$,

$$L_{p\geq 0} = Min_{p\geq 0} f_L = Min_{p\geq 0} \left(\frac{w_1}{\sum_{i=1}^{n} w_i} \alpha_1^p + \frac{w_2}{\sum_{i=1}^{n} w_i} \alpha_2^p, \dots + \frac{w_n}{\sum_{i=1}^{n} w_i} \alpha_n^p\right)^{1/p}$$
$$c_i \leq w_i \leq d_i, i = 1, 2, \dots, n$$

$$U_{p\geq 0} = Max_{p\geq 0} f_U = Max_{p\geq 0} \left(\frac{w_1}{\sum_{i=1}^{n} w_i} b_1^p + \frac{w_2}{\sum_{i=1}^{n} w_i} b_2^p, \dots + \frac{w_n}{\sum_{i=1}^{n} w_i} b_n^p\right)^{1/p}$$
$$c_i \leq w_i \leq d_i, i = 1, 2, \dots, n$$

In order to take the mean of the natural logarithm, Napierian logarithm is implemented for $L_{p\geq 0}$ and $U_{p\geq 0}$;

$$\text{In } L_{p\geq 0} = Min_{p\geq 0} f_L = Min_{p\geq 0} \frac{1}{p} \text{In}\left(\frac{w_1}{\sum_{i=1}^{n} w_i} \alpha_1^p + \frac{w_2}{\sum_{i=1}^{n} w_i} \alpha_2^p, \ldots + \frac{w_n}{\sum_{i=1}^{n} w_i} \alpha_n^p\right)$$

$$c_i \leq w_i \leq d_i, i = 1, 2, \ldots, n$$

$$= \frac{1}{p} \text{In} Min_{p\geq 0} \left(\frac{w_1}{\sum_{i=1}^{n} w_i} \alpha_1^p + \frac{w_2}{\sum_{i=1}^{n} w_i} \alpha_2^p, \ldots + \frac{w_n}{\sum_{i=1}^{n} w_i} \alpha_n^p\right) = \frac{1}{p} \text{In} L'_{p\geq 0}$$

$$c_i \leq w_i \leq d_i, i = 1, 2, \ldots, n$$

$$\text{In } U_{p\geq 0} = Max_{p\geq 0} f_U = Max_{p\geq 0} \frac{1}{p} \text{In}\left(\frac{w_1}{\sum_{i=1}^{n} w_i} b_1^p + \frac{w_2}{\sum_{i=1}^{n} w_i} b_2^p, \ldots + \frac{w_n}{\sum_{i=1}^{n} w_i} b_n^p\right)$$

$$c_i \leq w_i \leq d_i, i = 1, 2, \ldots, n$$

$$= \frac{1}{p} \text{In} Max_{p\geq 0} \left(\frac{w_1}{\sum_{i=1}^{n} w_i} b_1^p + \frac{w_2}{\sum_{i=1}^{n} w_i} b_2^p, \ldots + \frac{w_n}{\sum_{i=1}^{n} w_i} b_n^p\right) = \frac{1}{p} \text{In} U'_{p\geq 0}$$

$$c_i \leq w_i \leq d_i, i = 1, 2, \ldots, n$$

Defining variables of Support Vector Regression

$f(x)$ - Function of support vectors
w – Weight parameter that determine the hyperplane
$\phi_i(x)$ - Mapping function. Nonlinear input data can be separated linearly by mapping the input into high-dimensional feature space by using Kernel function.
$K(x_i, x_j)$ - Kernel function. The value of kernel function is equal to the inner products x_i and x_j in the feature space $\phi(x_i)$ and (x_j).
b – Bias term
ξ – Slack variable that measures above of the ε-tube.
ξ^* - Slack variable that measures below of the ε-tube.
l – Number of training data

C - Unit of regularization that controls the trade-off between the empirical risk and misclassification.

α_i – Unique minimum values that are above the ε-tube.

α_i^* - Unique maximum values that are below the ε-tube.

d : degree of the kernel

σ^2 : bandwidth of the kernel

$\phi(x_i), \phi(x_j)$ - maps the sample points x_i and x_j in hyperspace.

d_{ij} - the distance between sample points in hyperspace

Derivation of Equations for Support Vector Regression

Support Vector Regression;

$$f(x) = w^T\phi(x) + b$$

Nonlinear regression hyperplane;

$$f(x) = \sum_{i=1}^{l}(\alpha_i^* - \alpha_i)\, K(x_i, x) + b$$

Desired weights can be obtained by;

$$w = \sum_{i=1}^{l}(\alpha_i^* - \alpha_i)x_i$$

Optimal bias of the regression can be found by;

$$b = \frac{1}{l}\left(\sum_{i=1}^{l} y_i - x_i^T w\right)$$

Mapping the input data in the high dimensional feature space as linear function;

$$K(x,y) = \sum_{i=1}^{\infty} \lambda_i \phi_i(x)\phi_i(y)$$

Loss function with ε-intensivity zones;

$$|y - f(x,w)|_\varepsilon = \begin{cases} 0 & if\ |y - f(x,w)| \le \varepsilon \\ |y - f(x,w)| - \varepsilon & otherwise \end{cases}$$

$f(x,w)$ - Function of the weights *w* that are the subjects of learning
ε - Margin of tolerance
Empirical Risk;

$$R_{emp}^{\varepsilon}(w,b) = \frac{1}{l}\sum_{i=1}^{l} |y_i - w^T x_i - b|_\varepsilon$$

To minimize the empirical risk;

$$R = \frac{1}{2}\|w\|^2 + C\left(\sum_{i=1}^{l} |y_i - f(x_i,w)|_\varepsilon\right)$$

$\|w\|^2$ - Flatness of ε-tube.
Optimization of risk minimization for nonlinear data separation;

$$R_{w,\xi,\xi^*} = \left[\frac{1}{2}\|w\|^2 + C\left(\sum_{i=1}^{l}\xi_i + \sum_{i=1}^{l}\xi_i^*\right)\right]$$

$$Subject\ to \begin{cases} y_i - w^T x_i - b \le \varepsilon + \xi_i \\ w^T x_i + b - y_i \le \varepsilon + \xi_i^* \\ \xi_i \ge 0 \\ \xi_i^* \ge 0 \end{cases}$$

Implementation of the unique minimum and maximum values of the function that are above and below of the ε-tube by Lagrange multipliers;

$L_p(w, \xi, \xi^*, \alpha_i, \alpha_i^*, \beta_i, \beta_i^*)$:

$$L = \frac{1}{2}w^T w + C\left(\sum_{i=1}^{l} \xi_i + \sum_{i=1}^{l} \xi_i^*\right) - \sum_{i=1}^{l} \alpha_i^*[y_i - w^T x_i - b + \varepsilon + \xi_i^*]$$
$$- \sum_{i=1}^{l} \alpha_i[w^T x_i + b - y_i + \varepsilon + \xi_i^*] - \sum_{i=1}^{l} (\beta_i^* \xi_i^* + \beta_i \xi_i)$$

For regression, applying Karush-Kuhn-Tucker conditions can maximize Lagrange multipliers;

$$L_d(\alpha, \alpha^*) = -\varepsilon \sum_{i=1}^{l} (\alpha_i^* + \alpha_i) + \sum_{i=1}^{l} (\alpha_i^* - \alpha_i) y_i - \frac{1}{2} \sum_{i,j=1}^{l} (\alpha_i^* - \alpha_i)\left(\alpha_j^* + \alpha_j\right) x_i^T x_j$$

$$Subject\ to \begin{cases} \sum_{i=1}^{l} \alpha_i^* = \sum_{i=1}^{l} \alpha_i \\ 0 \le \alpha_i^* \le C \\ 0 \le \alpha_i \le C \end{cases}$$

Optimal desired weights vector *w* of the regression hyperplane;

$$w = \sum_{i=1}^{l} (\alpha_i^* - \alpha_i) x_i$$

Optimal bias *b* of the regression hyperplane;

$$b = \frac{1}{l}\left(\sum_{i=1}^{l} y_i - x_i^T w\right)$$

Best linear regression hyperplane;

$$z = f(x, w) = w^T x + b$$

Kernel function to theoretically map the nonlinear input data in the high dimensional feature space as linear data

$$K(x, y) = \sum_{i=1}^{\infty} \lambda_i \phi_i(x) \phi_i(y)$$

Gaussian Kernel Function: $K_{gaussian} = e^{-1/_{\sigma^2}[(x-x_i)^T \Sigma^{-1}(x-x_i)]}$

Polynomial Kernel Function: $K_{polynomial} = [(x^T x_i) + 1]^d$

Nonlinear regression formula after kernel function selection:

$$f(x) = \sum_{i=1}^{l} (\alpha_i^* - \alpha_i) K(x_i, x) + b$$

Feature-Weighted Support Vector Regression

$$d_{ij} = 2 - 2K(x_i, x_j) = 2 - 2\exp\left(-\gamma \|x_i - x_j\|^2\right)$$

w_k - Weight value

$$K_w(x_i, x_j) = exp\left(-\gamma \left(\sum_{k=1}^{n} \left(w_k (x_{ik} - x_{jk})\right)^2\right)\right)$$

Implementation of weight values by quadratic programming of optimization

$$min_{\alpha,\alpha*}\frac{1}{2}\sum_{i=1,j=1}^{l}(\alpha_i-a_i^*)(\alpha_j-\alpha_j^*)K_w\left(x_i,x_j\right)+\varepsilon\sum_{i=1}^{l}(\alpha_i+\alpha_i^*)-\sum_{i=1}^{l}(\alpha_i-\alpha_i^*)y_i$$

subject to:

$$\begin{cases}\sum_{i=1}^{l}(\alpha_i-\alpha_i^*)=0\\ 0\leq\alpha_i,a_i^*\leq C\end{cases}$$

Feature weight support vector regression:

$$f(x)=\sum_{i=1}^{l}(\bar{\alpha}_i^*-\bar{\alpha}_i)K_w(x_i,x)+\bar{b}$$

Defining variables of Grid Search with Cross Validation

TP: Correctly predicted positive values

TN: Correctly predicted negative values

FP: Prediction error when actual values and predicted value contradicts

FN: Prediction error when predicted value and actual values contradicts

Derivation of Equations for Grid Search with Cross Validation

There are four main indicators to evaluate the performance of the Grid Search with Cross Validation, which are; Accuracy, Precision, Recall, and F1 Score.

$$Accuracy = \frac{True\ Positives + True\ Negatives}{True\ Positives + False\ Positives + True\ Negatives + False\ Negatives}$$

$$Precision = \frac{True\ Positives}{True\ Positives + False\ Positives}$$

$$Recall = \frac{True\ Positives}{True\ Positives + False\ Negatives}$$

$$F1\ Score = \frac{2 \times (Precision \times Recall)}{Precision + Recall}$$

Defining variables, and Derivation of Equations for Statistical Tests

Key metrics to appraise machine learning model's accuracy; V Measure Score (VMS), Explained Variance Score (EVS), Mean Square Error (MSE), Root Mean Square Error (RMSE), Mean Absolute Error (MAE), Mean Absolute Percentage Error (MAPE), and Coefficient of Determination (R^2).

V Measure Score

K - set of clusters

C - set of classes

N – data points

n – number of classes

a_{ij} *– number of data points*

c_i *– member of class*

k_j *– element of cluster*

Homogeniety of clustering task

$$h = \begin{cases} 1 & \text{if } H(C,K) = 0 \\ 1 - \dfrac{H(C|K)}{H(C)} & else \end{cases}$$

where;

$$H(C|K) = -\sum_{k=1}^{|K|}\sum_{c=1}^{|C|}\frac{a_{ck}}{N}\log\frac{a_{ck}}{\sum_{c=1}^{|C|}a_{ck}}$$

$$H(C) = -\sum_{c=1}^{|C|}\frac{\sum_{k=1}^{|K|}a_{ck}}{n}\log\frac{\sum_{k=1}^{|K|}a_{ck}}{n}$$

Completeness of clustering task

$$c = \begin{cases} 1 & \text{if } H(K,C) = 0 \\ 1 - \dfrac{H(K|C)}{H(K)} & else \end{cases}$$

where;

$$H(K|C) = -\sum_{c=1}^{|C|}\sum_{k=1}^{|K|}\frac{a_{ck}}{N}\log\frac{a_{ck}}{\sum_{k=1}^{|K|}a_{ck}}$$

$$H(K) = -\sum_{k=1}^{|K|}\frac{\sum_{c=1}^{|C|}a_{ck}}{n}\log\frac{\sum_{c=1}^{|C|}a_{ck}}{n}$$

V measure score – Harmonic mean between homogeneity and completeness;

β - Weight

h - homogeniety

c – completeness

$$V_{\beta} = \frac{(1+\beta) * h * c}{(\beta * h) + c}$$

Explained Variance Score (EVS)

y – True value

$\hat{y}$ – Predicted value

Var – Variance (Squared standard deviation)

$$explained\ variance(y, \hat{y}) = 1 - \frac{Var\{y - \hat{y}\}}{Var\{y\}}$$

Mean Squared Error (MSE)

n – Number of data points

y_i – Value observed

$\hat{y}_i$ - Predicted value of the i-th sample

$$MSE = \frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2$$

Root Mean Square Error (RMSE)

N - Number of data points

x_i *– Actual values*

$\hat{x}_i$ - Predicted value of the i-th sample

$$RMSE = \sqrt{\frac{\frac{1}{n}\sum_{i=1}^{N}(x_i - \hat{x}_i)^2}{N}}$$

Scatter Index (SI)

$\bar{O}$ - values of average observation

$\bar{S}$- values of average simulation

$$SI = \sqrt{\frac{\sum_{i=1}^{N}[(S_i - \bar{S}) - (O_i - \bar{O})]^2}{\sum_{i=1}^{N} O_i^2}}$$

Mean Absolute Error (MAE)

n – Number of data points

y_i – Value observed

$\hat{y}_i$ - Predicted value of the i-th sample

$$MAE = \frac{1}{n}\sum_{i=1}^{n}|y_i - \hat{y}_i|$$

Mean Absolute Percentage Error (MAPE)

n - number of times the summation iteration happens

A_t - actual value

F_t - forecast value

$$MAPE = \frac{1}{n}\sum_{t=1}^{n}\left|\frac{A_t - F_t}{A_t}\right|$$

Coefficient of Determination

y - result variable

$\tilde{y}$ - meta-model result variable

$\bar{\tilde{y}}$ - mean variable

$$CoD = \left(\frac{\sum_{i=1}^{n} [(\tilde{y}_i - \bar{\tilde{y}}) \cdot (y_i - \bar{y})]}{\sqrt{\sum_{i=1}^{n} (\tilde{y}_i - \tilde{y})^2 \cdot \sum_{i=1}^{n} (y_i - \bar{y})^2}} \right)^2$$

BIOGRAPHY

Hüseyin İlker Erçen had received his Bachelor of Arts Degree in Business Management and Economics (Dual Honours) from Keele University, United Kingdom in 2013. He carried on his education at Near East University, where he received his Master of Science Degree in Banking and Finance in 2015, with the CGPA of 4.0. During his Masters and Doctorate programs, he also attended to online certificate programs of world's leading universities, such as; Harvard University, Massachusetts Institute of Technology, Princeton University, and Ecole Polytechnique Fédérale de Lausanne in various subjects like; Economics, Finance, Business Management, Philosophy, Neuroscience, and Quantum Mechanics. In 2015, he took his first step to his academic career as a part-time lecturer at Faculty of Economics & Administrative Sciences in Near East University, while undertaking Senior Researcher role in NEU Robotics Laboratory, and Co-Founder of NEU 3D Laboratories.

Hüsyin İlker

ORIGINALITY REPORT

6%	4%	4%	2%
SIMILARITY INDEX	INTERNET SOURCES	PUBLICATIONS	STUDENT PAPERS

PRIMARY SOURCES

1	dokumen.pub Internet Source	1 %
2	doctorpenguin.com Internet Source	<1 %
3	Guh, Y.Y.. "The fuzzy weighted average within a generalized means function", Computers and Mathematics with Applications, 200806 Publication	<1 %
4	www.scribd.com Internet Source	<1 %
5	Sariyar, B.. "Monte Carlo sampling and principal component analysis of flux distributions yield topological and modular information on metabolic networks", Journal of Theoretical Biology, 20060921 Publication	<1 %
6	"Internet Science", Springer Science and Business Media LLC, 2018 Publication	<1 %

Lisansüstü Eğitim Enstitüsü Müdürlüğü'ne;

Tezin yazılıp hazırlanmasında etik kurallarına aykırı hiçbir unsurun yer almadığını tez danışmanları olarak beyan ederiz.

Prof.Dr.Hüseyin ÖZDEŞER (Supervisor)

Doç.Dr.Turgut TÜRSOY (Co-supervisor)