CHRONIC KIDNEY DISEASE DETECTION USING FUZZY NEURAL NETWORK

A THESIS SUBMITTED TO THE GRADUATE SCHOOL OF APPLIED SCIENCES OF NEAR EAST UNIVERSITY

By

REBAR DARA MOHAMMED

In Partial Fulfillment of the Requirements for the Degree of Master of Science

In

Computer Engineering

NICOSIA, 2021

CHRONIC KIDNEY DISEASE DETECTION USING FUZZY NEURAL NETWORK

A THESIS SUBMITTED TO THE
GRADUATE SCHOOL OF APPLIED SCIENCES
OF
NEAR EAST UNIVERSITY

By
REBAR DARA MOHAMMED

In Partial Fulfillment of the Requirements for
the Degree of Master of Science
in
Computer Engineering

NICOSIA, 2021

REBAR DARA MOHAMMED: CHRONIC KIDNEY DISEASE DETECTION USING FUZZY NEURAL NETWORKS

Approval of Director of Graduate School of Applied Sciences

Prof Dr K. Hüsnü Can Başer

We certify that this thesis is satisfactory for the award of the degree of Master of Science in Computer Engineering

Examining Committee in Charge:

Assoc. Prof. Dr. Boran Şekeroglu	Committee Member, Department of Information Systems Engineering, NEU
Assist. Prof. Dr. Elbrus Imanov	Committee Member, Department of Computer Engineering, NEU
Prof. Dr. Rahib Abiyev	Supervisor, Committee Member, Department of Computer Engineering, NEU

I hereby declare that all information in this document has been obtained and presented in accordance with academic rules and ethical conduct. I also declare that, as required by these rules and conduct, I have fully cited and referenced all material and results that are not original to this work.

Name, Last name: Rebar Mohammed

Signature: 

Date:

To my parents...

ACKNOWLEDGEMENTS

I'm appreciative and obliged wholeheartedly to Prof. Dr. Rahib Abiyev for his incredible oversight, help, resilience, and determination all through my postulation at Near East University. His recommendation, what's more, the direction was the way to progress and aided me scholastically as well as I took in a great deal about social science also. The oversight of the boss assisted me with the yearning route since I have first begun. He has not just propelled me to depict the examination abilities through the proposal yet additionally been a good example for me. This chance to create and compose a proposal isn't without a doubt, truly noteworthy for me yet additionally their quality in the center, it is consistently fundamental to complete such autonomous studies to move past progress and flourishing in their life.

My sincere gratitude to all of my instructors and teachers throughout my academic career, particularly those who helped to enlighten me on my path to this academic stage of master's degree study: Assist. Prof. Dr. Elbrus Imanov, Prof.Dr.Fahreddin Sadikoglu Assist. Prof. Dr. Fadi Al-Turjman, Assist. Prof. Dr. Melike Sah Direkoglu, Assist. Prof. Dr. Kaan Uyar.

I would like to express my deepest appreciation to my cousin Mr.Kamaran Tahir that he helped me in the duration of my master study.

I would also like to extend my deepest gratitude to my dear friend Mr.Arsalan Afrasyaw that he was supported me and helped me in some part of my master.

Last but not least, I am grateful to all of my family members for their unwavering support and unwavering faith.

ABSTRACT

Chronic kidney disease (CKD) is a loss of the important blood-filtering functions of kidneys. Kidney disease remains a global health issue due to numerous patients. Due to the rapid and precise recognition capabilities of machine learning models, they successfully assist doctors in achieving this goal. Artificial intelligence (AI) has been widely used in medical disciplines to enhance medical advancement, with applications mostly in early detection, disease diagnosis, and management. We present a machine learning framework for diagnosing CKD in this study. The fuzzy neural network is proposed for the detection of kidney diseases. The CKD data set was obtained from the AI vault at the University of California, Irvine (UCI), which has a substantial number of missing values. Effective prediction approaches should be considered as the percentage of individuals affected by CKD continues to rise. The dataset includes the statistics of 400 patients that have 24 input attributes characterizing chronic kidney diseases. The Fuzzy Neural Network (FNN) was utilized for the classification of the diseases in this study. The architecture of FNN has been designed for CKD classification. The design of the FNN is performed using a different number of rules (hidden neurons). Learning of the system has been performed using a cross-validation approach by applying a gradient descent algorithm. To conduct the learning, the mean of the respective attributes was used to replace all missing values in the CKD dataset. After learning, the designed FNN system is applied for the detection of diseases. For the classification of chronic kidney disease, fuzzy neural networks with various numbers of rules were used. The recognition rate with 16 rules was obtained as 99.75%. The obtained experimental and comparative results demonstrate the efficiency of using the FNN system in diagnosing chronic kidney diseases.

Keywords: Chronic kidney disease, neural networks, fuzzy neural networks, AI, machine learning.

ÖZET

Kronik böbrek hastalığı (KBH), böbreklerin önemli kan filtreleme işlevlerinin kaybıdır. Böbrek hastalığı, çok sayıda hasta nedeniyle küresel bir sağlık sorunu olmaya devam etmektedir. Makine öğrenimi modellerinin hızlı ve kesin tanıma yetenekleri sayesinde, doktorlara bu amaca ulaşmada başarılı bir şekilde yardımcı olurlar. Yapay zeka (AI), çoğunlukla erken teşhis, hastalık teşhisi ve yönetimindeki uygulamalarla tıbbi ilerlemeyi geliştirmek için tıp disiplinlerinde yaygın olarak kullanılmaktadır. Bu çalışmada KBH teşhisi için bir makine öğrenimi çerçevesi sunuyoruz. Bulanık sinir ağı böbrek hastalıklarının tespiti için önerilmiştir. CKD veri seti, önemli sayıda eksik değere sahip olan California Üniversitesi, Irvine'deki (UCI) AI kasasından elde edildi. KBH'den etkilenen bireylerin yüzdesi artmaya devam ettikçe etkili tahmin yaklaşımları düşünülmelidir. Veri seti, kronik böbrek hastalıklarını karakterize eden 24 girdi özelliğine sahip 400 hastanın istatistiklerini içerir. Bu çalışmada hastalıkların sınıflandırılması için Fuzzy Neural Network (FNN) kullanılmıştır. FNN mimarisi, CKD sınıflandırması için tasarlanmıştır. FNN'nin tasarımı, farklı sayıda kural (gizli nöronlar) kullanılarak gerçekleştirilir. Sistemin öğrenilmesi, bir gradyan iniş algoritması uygulanarak bir çapraz doğrulama yaklaşımı kullanılarak gerçekleştirilmiştir. Öğrenmeyi yürütmek için, CKD veri setindeki tüm eksik değerleri değiştirmek için ilgili özniteliklerin ortalaması kullanıldı. Öğrendikten sonra hastalıkların tespiti için tasarlanan FNN sistemi uygulanır. Kronik böbrek hastalığının sınıflandırılması için çeşitli sayıda kurallara sahip bulanık sinir ağları kullanılmıştır. 16 kural ile tanınma oranı %99.75 olarak elde edilmiştir. Elde edilen deneysel ve karşılaştırmalı sonuçlar, FNN sisteminin kronik böbrek hastalıklarının teşhisinde kullanılmasının etkinliğini göstermektedir

Anahtar Kelimeler: Kronik böbrek hastalığı, sinir ağları, bulanık sinir ağları, AI, makine öğrenimi.

TABLE OF CONTENTS

ACKNOWLEDGEMENTS ..... **i**

ABSTRACT ..... **ii**

ÖZET ..... **iii**

TABLE OF CONTENT ..... **iv**

LIST OF FIGURES ..... **vii**

LIST OF TABLES ..... **viii**

LIST OF ABBREVIATIONS ..... **ix**

CHAPTER 1 ..... **1**

INTRODUCTION 1

1.1 Motivation 1

1.2 Contribution 2

1.3 Purpose of the study 2

1.4 Importance of the Research 3

1.5 Study area and data 3

1.6 Study Limitation 4

1.7 Overview of the Research 4

CHAPTER 2 ..... **6**

LITERATURE REVIEW 6

CHAPTER 3 ..... **10**

BACKGROUND STUDIES **10**

3.1 Kidneys 10

3.2 Chronic Kidney Disease (CKD) 11
3.3 Chronic Kidney Disease (CKD) Statistics 12
3.4 Chronic Kidney Disease (CKD) Risk Factor 13
3.5 Artificial Intelligence 15
3.6 Machine Learning 16
3.6.1 Machine Learning (ML) Algorithms 17
3.7 Neural Network 20
3.7.1 Overview 20
3.7.2 Technical description 22
3.8 Support Vector Machine (SVM) 22
3.9 Deep Learning Method 24
3.10 Cross-Validation Method 25
CHAPTER 4 27
4.1 FNN Model 27
4.2 Parameter Learning of Fuzzy Neural Network (FNN) 31
4.3 CKD Dataset 34
4.4 Data Pre-processing 36
CHAPTER 5 39
SIMULATION AND RESULTS 39
5.1 Simulation 39
5.2 Performance Evaluation Measurements 39
5.2.1 Accuracy 40
5.2.2 Precision 40

5.2.3 Recall 40
5.2.4 F1-Score 40
5.3 Simulation Results 41
5.4 Comparisons and related works using CKD dataset 44
CHAPTER 6 46
CONCLUSIONS 46
REFERENCES 47
APPENDICES 55
APPENDIX 1 56
APPENDIX 2 62
APPENDIX 3 63

LIST OF FIGURES

Figure 3.1: Kidneys 10
Figure 3.2: Stages in Progression of CKD.....13
Figure 3.3: Machine Learning Algorithms Diagram. 17
Figure 3.4: Training data nature for machine learning algorithm types. 19
Figure 3.5: Reinforcement Learning Process. 20
Figure 3.6: Neural Network Architecture. 21
Figure 3.7: SVM for Multi-Class Classification.....23
Figure 3.8: AI, ML, and DL Introduction. 25
Figure 3.9: N-Fold Croass Validation Graphical. 25
Figure 4.1: Flowchart for CKD Detection. 27
Figure 4.2: Fuzzy neural network (FNN) architecture. 30
Figure 4.3: The process of converting the numerical values 36
Figure 5.1: The plot of root means square error (RMSE) obtained from fuzzy neural network (FNN) classifier by using 5 rules (hidden neuron).. 43
Figure 5.2: The plot of root mean square error (RMSE) obtained from fuzzy neural network (FNN) classifier by using 8 rules (hidden neuron). 43
Figure 5.3: The plot of root mean square error (RMSE) obtained from fuzzy neural network (FNN) classifier by using 16 rules (hidden neuron). 44

LIST OF TABLES

Table 4.1: CKD Dataset description. 35
Table 4.2: Dataset for 400 patients with Missing Values. 36
Table 4.3: Missing value representation. 37
Table 4.4: CKD Datasets after finding Missing values. 38
Table 5.1: The FNN model's experimental results. 44
Table 5.2: Comparisons and related works using CKD dataset. 45

LIST OF ABBREVIATIONS

AI: Artificial Intelligence

ML: Machine Learning

DL: Deep Learning

FNN: Fuzzy Neural Network

NN: Neural Network

CKD: Chronic Kidney Disease

ANN: Artificial Neural Network

SVM: Support Vector Machine

KNN: K- Nearest Neighbor

NB: Naïve Bayes

FCM: Fuzzy C-Means

QOF: Quality Outcomes Framework

BP: Blood Pressure

SG: Specific Gravity

RBC: Red Blood Cells

PC: Pus Cell

PCC: Pus Cell Clumps

BA: Bacteria

BGR: Blood Glucose Random

BU: Blood Urea

SC: Serum Creatinine

SOD: Sodium

POT: Potassium

HEMO: Hemoglobin

PCV: Packed Cell Volume

WC: White Blood Cell Count

RC: Red Blood Cell Count

HTN: Hypertension

DM: Diabetes Mellitus

CAD: Coronary Artery Disease

APPET: Appetite

PE: Pedal Edema

ANE: Anemia

RBF: Radial Basis Function

TSK: Takagi-Sugeno-Kang

RMSE: Root Mean Square Error

GFR: Glomerular Filtration Rate

TP: True Positive

FP: False Positive

TN: True Negative

FN: False Negative

CHAPTER 1

INTRODUCTION

1.1 Motivation

Chronic kidney disease (CKD) is serious global medical problem raised by a losing kidney functions of the patient. Even though it is widely recognized that CKD is associated with increased risks of end-stage excretory organ disease, vessel accidents, and all-cause mortality, there are currently no reliable records on individual patients. The term "discharge organ damage" refers to a condition in which the kidney's capacity is reduced due to a significant drop in the vessel filtration rate (GFR) (A. Salekin and J. Stankovic, 2016). The kidneys act as filters, removing waste from the blood through a network of small veins. It degrades in some situations, and the kidneys lose their ability to recognize supplements, resulting in nephropathy. CKD has no known cause, but it is usually irreversible and can lead to serious medical problems. According to the Asian country excretory organ establishment, there are more than ten million people in the country who have nephropathy.

According to the Asian nation excretory organ establishment, more than ten million people in the country are currently suffering from nephropathy. Almost 160,000 excretory organ patients, according to the UN office, are in good health and must be forced to undergo regular subjective examinations week after week. About 195 million young girls are affected by CKD around the world, and it is the eighth leading cause of death among women, with about 600.000 people dying each year. Between 2011 and 2012, there were approximately 1.9 million adults in the United Kingdom with CKD who were enrolled in the Quality Outcomes Framework (QOF). Regardless, the total yearly expense of CKD to the United Kingdom's National Health Service is quantifiable.

In respect of £795 per personate with CKD and recorded in the QOF; that equates to £1.45 billion each year (Kunwar et al., 2016). Distinguishing individuals with CKD in the beginning

phases will assist with diminishing the danger of end-stage excretory organ sickness. As a result, this thesis aims to provide this list and to focus on specific approaches that may be used to resolve the issues that are projected in the context of CKD or in other settings. The systems that want to carry out the writing survey are outlined in the next section. within the future outcomes section, we will first show the basic CKD findings that are researched in conjunction with selected documents, then show and discuss the relapse procedures used with each set of results. We prefer to bless some possible different logical methodologies that are ne'er or only rarely used with regards to CKD but may, in any case, be of interest where applicable.

1.2 Contribution

In this thesis, FNN model is proposed for detection of kidney diseases. After discussing feature reduction, learning algorithms, and comparing several ML techniques on the CKD dataset, we found that the FNN model performed well. The most critical features were ranked and correspondingly classified using the FNN model. After that, we compared the accuracy of various machine learning approaches such as Logistic regression, Gradient Boosting, Decision Tree, AdaBoost, Support Vector Regression, and AdaBoost to prove efficiency of the model. The comparative results ML models that are scheduled in chapter 5 indicate the efficiency of proposed approach.

1.3 Purpose of the study

The goal of this thesis is to create automated approaches to assist clinicians in the diagnosis kidney diseases to avoid misdiagnosis and cut down on patient waiting times. This research work accomplishes the goal using the dataset having CKD or not-CKD classes. The fuzzy neural network (FNN) model, which can be designed and used in treatment dynamic, is perhaps the main stage in automation of diseases diagnose. For finding of CKD, we examined a novel strategy in an incorporated fuzzy neural structure (multi-input and multi-output) in

light of the Takagi-Sugeno-Kang (TSK)- type rule. We used selected FNN to analyze clinical data of different kidney diseases in real-world medical diagnosis. However, due to the restrictions of the settings, the available data samples for establishing the model are quite little, with just 400 samples. As a result, the model's generalization performance may be limited. Furthermore, because the data set contains only two types of data samples (CKD and not-CKD), the model is unable to determine the severity of CKD. We anticipate that as the size and quality of the data grow, our model will become even more perfect.

1.4 Importance of the Research

Significant advances in medical science have been made in recent years as a result of Machine Learning and Deep Learning application in the medical image processing. These techniques allow diagnosing diseases earlier and more quickly. Before the invention of these techniques, diagnosing diseases was tedious and time-consuming. Because the medical industry requires precise and efficient methods to identify life-threatening diseases such as cancer, which is the world's leading cause of patient mortality, computer-aided technology is critical in overcoming such limits. Furthermore, with the use of CKD datasets, we develop a model for the detection of chronic kidney disease utilizing the fuzzy neural network (FNN) model in our study.

1.5 Study area and data

This study will be conducted by taking advantage of the CKD dataset of Dr. P. Soundarapandian's open-source research library at the University of California Irvine (UCI); thus, the area of this study will be the same area of the CKD dataset. A total of 400 individuals were included in the data collection, 250 of whom had CKD and 150 of whom were healthy.

1.6 Study Limitation

Chronic kidney disease is unquestionably one of the most difficult deadly diagnoses to diagnose with great accuracy and precision. In a nutshell, developing an application for detecting chronic sickness will benefit not only medical experts in treating crucial situations but also people who have difficulty contacting a doctor. One of the reasons for the difficulty in diagnosing CKD is that it is difficult to anticipate and detect the diseases, because it is not dependent on a single parameter. Furthermore, common CKD symptoms do not play a substantial role in diagnosing the diseases. The necessity for robust validation in real-world investigations is probably the most significant drawback of the AI/ML methodology. We recognize that AI software capabilities outnumber AI excitement, owing to a lack of clinical validation and everyday care deployment.

1.7 Overview of the Research

To create a chronic kidney disease system, the thesis contains the following chapters.

Chapter 1 presents the importance of research study, motivation, purpose, importance, study area and data, study constraints, contribution, and a brief overview of the thesis's topic.

Chapter 2 presents the review of machine learning techniques used for the detection of kidney diseases. Different AI techniques have been analyzed. Especially, the discussions provide some relevant research works presented to solve the chronic kidney disease detection issues by using machine learning algorithms.

Chapter 3 presents the description of AI methods, specifically, neural networks (NN), Support Vector Machines (SVM), and Fuzzy Neural Networks. This chapter also comprehensively and in detail talks about deep learning methods and background studies.

The section also explains the basics of the machine learning techniques, the preprocessing, feature extraction, and classification stages used for the detection of chronic kidney disease.

Chapter 4 presents the proposed multi-input multi-output FNN model used for the detection of kidney diseases. The structure and mathematical background of FNN model are presented, the learning algorithm used is described.

Chapter 5 presents the modeling FNN based kidney diseases detection model. An simulation results of the FNN model is given. Comparative results of different models are presented to demonstrate the efficiency of the proposed FNN model.

Conclusions present the important simulation results obtained from the thesis.

CHAPTER 2

LITERATURE REVIEW

Chronic kidney disease (CKD) is a worldwide public health problem that claims an incalculable number of lives. CKD is the eleventh leading cause of death worldwide, accounting for 1.2 million deaths each year, and according to the Bangladeshi Kidney Foundation, approximately 40,000 CKD patients experience kidney failure each year, with a few thousand dying prematurely due to CKD. Kidney disease remains a global health issue due to a large number of patients.

Early detection of kidney diseases allows treatment and help to the people suffering from the diseases. Different techniques were applied for the detection of kidney diseases. Recently machine learning techniques have been widely applied for the detection of CKD.

Radial Basis Function (RBF) was used by Xun L. et al. (A. Salekin and J. Stankovic, 2016) to calculate the Glomerular Filtration Rate (GFR) in kidney infection patients. Three hundred and twenty-seven instances were included in the dataset. The exactness of their RBF model was higher than established equations such as Jelliffe-1973-condition and Ruijinequation for a Standard GFR with less than 30% variance, according to their findings. In diagnosing CKD, the RBF model has a precision of 82.1 percent. To diagnose CKD, Salekin A. (D. Gupta et al., 2016) compared the performance of three models dependent on K-NN, Random Forest, and Neural Network. They used the IBK algorithm to replace missing data in the dataset for the K-NN model and the Neural Network in their paper, whilst the C4.5 Random Forest Model was utilized to build execution about missing data in the dataset. The Random Forest strategy outperformed the other two approaches,

The F1 score was 0.993, with the Root Mean Square Error (RMSE) being 0.0184. Both writers in (D. Gupta et al., 2016) went on to reduce the model's attributes to pick the most relevant ones.

These characteristics were used to create a more basic and efficient CKD model. The association between eleven chronic diseases, including kidney disease, was determined using the ML technique by Gupta D. et al. (A. Y. Al-Hyari et al.,2013). Their dataset included 4384 examples, and they used many machine learning techniques to diagnose CKD.

The AdaBoost technique, according to the authors, provided the best results, with classification accuracy of 98.87 percent and 88.66 percent for traineeship and research, sequentially. A. Y. Al-Hyari et al. (Q. Zheng et al.,2018) developed a CKD model using Decision Tree (DT), Artificial Neural Network (ANN), and Naive Bayes methods (NB).

For each of their models, a total of 102 were decided. The DT, NB, and ANN have an accuracy of 92.2 percent, 88.2 percent, and 82.4 percent, respectively. To enhance the diagnosis of congenital defects of the kidney and urinary tract (CAKUT) in infants, Zheng et al. (S. P. Deng et al.,2017) used the transfer-learning approach to derive imaging features from ultrasound kidney videos.

The Support Vector Machine has been used to identify the features, which included a mix of transfer learning and traditional imaging features. The model's accuracy and area under the ROC curves are 0.87 and 0.92, respectively.

Deng et al. (Vijayarani et al.,2015) used genetic code and DNA methylation datum to create a fused network to identify the phases of kidney Renal Cell Carcinoma (KIRC). Each type of data was used to create a patient's network, which was then fused using network fusion methods. Their model was 0.852 percent accurate.

Ravindra et al. (Vijayarani et al.,2015) classified CKD and NCKD patients utilizing a Support Vector Machine (SVM) with a radial foundation kernel action. They used the equivalent dataset from the UCI ML repository and divided it into four categories, including titular and numeral qualities. Cases 1, 2, 3, and 4 each had 3, 4, 4, and 7 qualities, respectively. The classifier was tested in all 4 cases, and the sensitivity, specificity, and classification accuracy scores were reported.

Case 4 had the most elevated in general precision of 93.75 percent, according to their findings. Case 1 received the greatest Sensitivity score of 100 percent, while Case 4 had the best Specificity score of 98% (R. B.V, N. Sriraam, and M. Geetha,2017).

The rendering of the Support vector machine (SVM) and the K-Nearest Neighbour (KNN) allocations in categorizing CDK was compared by Parul et al. (Sinha et al.,2015). The representation of the classifiers was assessed utilizing accuracy. Precision, rendering, and the F-measure are all important factors to consider. KNN outperformed SVM regarding accuracy, precision, rendering, and f-measure scores, with 78 percent accuracy, 85 percent precision, and 80 percent f-measure. SVM, on the other hand, has the maximum recall rating of 100 percent.

Dr. Uma et al. (U. N. Dulhare and M. Ayesha,2016) utilized the Naive Bayes classification with the OneR (One Rule) accredit elect to categorize CKD and develop act rule to restraint the disease from progressing further. In comparison to previous systems, the OneR attribute annotator with the Naive Bayes classifier chose only 5 traits out of 25 and decreased the attributes by 80%, improving prediction accuracy by 12.5 percent. This approach also extracted six action rules for each stage of chronic kidney disease, allowing the essential treatments to be administered following the action rules to prevent CKD progression to the next stage. (Boukenze et al.,2017). They deciphered segregation between ANNs, SVM, KNN, and NB calculations to foresee ongoing kidney sickness. The comparison was made based on the WEKA tool's forecast accuracy. The best method, according to their research, was ANN with SVM, which had a 62.5 percent accuracy.

(Panwong et al.,2016) used the WEKA tool to compare the KNN, ANN, RF, J48, and NB algorithms. With an accuracy of 86.6 percent, they achieved the highest precision with Random Forest Classifier.

(U. Tulare et al.,2016) used the WEKA tool to examine the prediction of CKD using the NB classifier and found an accuracy of 97.5 percent.

(Abeer et. al.) utilized Classifiers like SVM and LR are analyzed dependent on their presentation. They saw that the presentation of the SVM has shown better results when contrasted with the other calculation with an exactness of 93.1%.

(R. Dhruvi et al.,2016) used J48 and Naive Bayes methods in Java to evaluate CKD prediction. According to their findings, the best algorithm was J48, which had a 62.5 percent accuracy rate.

To predict chronic kidney failure, (Boukenze et al.,2016) conducted a comparison study of support vector machines, decision trees, and Bayesian networks. In terms of forecasting chronic kidney failure, decision trees outperform support vector machines and Bayesian networks. The decision tree's accuracy is 97 percent. Hadoop and the map/reduce tool are used to implement intelligent approaches. From the perspective of the researchers, a decision tree might accurately predict chronic renal failure.

(Padmanaban et al.,2016) published a study that compared the use of Nave Bayes and decision trees to predict chronic renal failure. In terms of forecasting chronic kidney failure, decision trees surpass Nave Bayes. The decision tree's accuracy is 91 percent. On the WEKA mining tool, clever strategies are used. From the viewpoint of the researchers, a decision tree can accurately predict chronic renal failure.

CHAPTER 3

BACKGROUND STUDIES

3.1 Kidneys

The kidney is a pair of fist-sized bean-shaped organs (NIDDK, 2018) (Kidney.org, 2015). They are located below the ribs, and there is one on each side of the spine. Every day, the kidney filters 120 to 150 liters of blood to produce 1 to 2 liters of urine. The main function of the kidneys is to expel waste and excess fluid from the body through urine. Urine is composed of a series of complex stages of excretion and reabsorption. This process is essential to keep the body's chemical balance stable. The kidneys are responsible for controlling the levels of salt, potassium, and acid in the body and producing hormones that affect the function of other organs. For example, the kidneys produce a hormone that can promote red blood cell production, regulate blood pressure, and regulate calcium metabolism (Kidney.org, 2014).

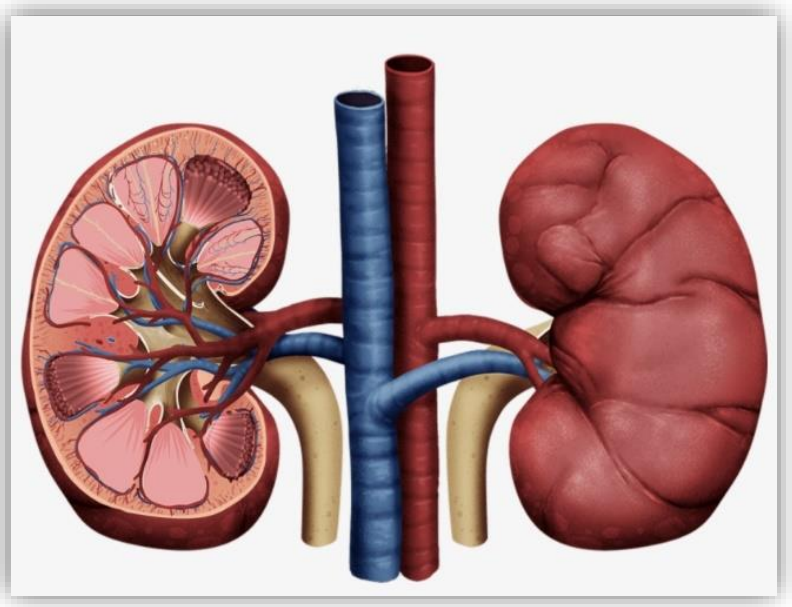


FIGURE 3.1: Kidneys

3.2 Chronic Kidney Disease (CKD)

Chronic kidney disease (CKD) is a global public health problem that affects about 10% of the world's population (Z. Chen et al., A. Subasi et al.,2016). In China, 10.8% of people have chronic kidney disease (L. Zhang et al.,2012), and in the United States, the prevalence ranges from 10% to 15% (A. Singh et al.,2015). According to another study, among the general adult population in Mexico, this number has now reached 14.7% (A. M. Cueto-Manzano et al.,2014). Renal function gradually deteriorates with this condition, eventually leading to total renal failure. In the early stages of CKD, there are no visible symptoms. As a result, the disease may not be recognized until roughly 25% of the kidney's function has been lost (H. Polat et al.,2017). Furthermore, CKD is associated with a high rate of morbidity and death, as well as a global influence on the human body (C. Barbieri et al.,2015).

Furthermore, chronic kidney disease (CKD) is a long-term kidney impairment that affects 5-10% of the world population. Moreover, millions of individuals die each year as a result of cardiovascular disorders linked to CKD. Furthermore, because early-stage CKD has no visible symptoms, approximately 90% of CKD patients are unaware of their condition. Taiwan has the highest estimated total CKD prevalence (15.46 percent) in the world, according to a recent study (M.-H. Tsai et al.,2018).

Chronic Kidney Disease (CKD) is a disorder that causes the kidneys to lose their normal function and, in the long term, causes permanent kidney damage. Wastes and excessive fluids begin to wreak havoc on our bodies in the end stage of renal failure, causing other organs and hormones to malfunction (V. Kunwar et al.,2016). As a result, it is necessary to address concerns about CKD through research focused on the early prediction of diseases. Early prediction leads to early detection and, as a result, early remediation for chronic illness patients. In the healthcare industry, data mining has become an essential specialist subject for accurate and early identification of chronic diseases in patients who have been afflicted for a long time. Both the physician and the patient will benefit from the disease prediction results.

Early detection and treatment can prevent or delay the progression of CKD to renal failure or other negative effects (Zewei Chena et al.,2016) (Singh N et al.,2017).

3.3 Chronic Kidney Disease (CKD) Statistics

Using the PRISMA guidelines, a methodical analysis of pertinent existing literature from South Asian countries was conducted. Two experts independently searched Pub Med, Google Scholar, and POPLINE for potential literary works. In addition, national online journals in India, Pakistan, Bangladesh, Nepal, and Sri Lanka were searched. Bhutan, the Maldives, and Afghanistan, on the other hand, did not have access to any domestic online journals. The following keywords were used in the search: Epidemiology, Prevalence, Chronic Renal Failure, Chronic Kidney Disease, India, Bangladesh, Sri Lanka, Nepal, Bhutan, Maldives, Pakistan, and Afghanistan.

A worldwide search term for potential literature searches has been developed using these key terms and Boolean operators. To identify additional papers, the bibliography of all chosen researches (snow bowling) was manually scanned.

In India, eight trials have really been reported, all of which used a cross-sectional research designing. The plurality of the study was conducted entirely in urban settings, with only one sample including respondents from urban, semi-urban, and rural settings. Three of these studies used random sampling to select respondents. The number of people who took part in this study ranged from 1104 to 12,271, with the major part of them being grown males. The clinical trials looked at spot quantitative urine protein and/or eGFR as biomarkers for CKD diagnosis. Three researches utilized the NIDRD formula, one used the CKD-EPI formula, and two used the eGFR calculation. The CG-BSA and NIDRD formulas were used in the rest of the study (Tahmid Abrar et al.,2019).

3.4 Chronic Kidney Disease (CKD) Risk Factor

The gradual and generally irreparable ailment is chronic kidney disease (CKD). Different types of results are concerned with CKD, such as time to the dialog, transplantation, or GFR. Statistical analyzes to examine the relationship between these results and risk factors boosted several methodological questions. The aim of this research was for these challenges to be overviewed and certain statistical methods identified which could address them.

FIGURE 3.2: Stages in Progression of CKD.

Chronic kidney disease (CKD) is a collective term for various diseases that affect the structure and function of the kidney. It is rarely reversible and usually follows a gradual direction. To carry out possible therapeutic interventions, it is important to identify risk factors for the development of CKD. Due to increased mortality and treatment costs, it is necessary to prevent progression to renal failure, the glomerular filtration rate (GFR) is less than 15 ml / min / 1.73 m2, or the need for dialysis or transplantation. In statistical analysis, several types of outcome variables can be used to study risk factors associated with CKD progression. For example, the time to reach a certain GFR value, the start of dialysis or transplantation, cardiovascular events, or all-cause death may all be outcome variables. The slope of the GFR decline or its overall trend over time can also be used as a vector of results. If the outcome variable is the time of the event of particular interest, then survival regression models (such as the Cox model) seem to be able to examine the risk factors associated with that event. Normal

survival analysis requires knowing the exact time of the event in all patients who have experienced the event. For example, for events such as death or the start of renal replacement therapy, the exact date can usually be found. Even so, the time-to-event for several other events of concern in the development of kidney disease may not be precisely understood. It is understood, for example, that the progression to a specific GFR value occurred only between two consecutive GFR readings. The time between these two measurements is said to be "interval-censored" in such a situation. Interval censorship should ideally be included in the evaluation, particularly if the time interval between successive measurements is long. Furthermore, competing hazards may have to be taken into account when determining survival. A competing occurrence, by definition, prevents the interesting case from being observed. For example, in all cases of progression, death is a competitive event because patients who die during follow-up will no longer progress. If the decline in the number of GFRs over time is of concern, linear regression models can be used to investigate the relationship between risk variables and aggregate statistics (such as mountains) of individual GFRs over time. On the other hand, such methods cannot account for all the data available for repeat measurements of kidney function and cannot draw certain conclusions that are not always accurate. If enough patients take at least three measurements, it may be preferable to use a mixed linear model of all repeated quantitative measurements of kidney function. The regression model reflects the correlation between repeated measurements from the same patient and can be adapted to multiple measurement numbers from each patient acquired at unequal intervals over time. In addition, the non-linear time trajectory, therefore, each smoothness of the ERC result variable raises a series of methodological questions in the statistical analysis of the risk variable. Although some of these statistical problems have been identified in the nephrology literature, no article provides an overview of these statistical problems for our understanding. Therefore, this article aims to provide an overview and focus on some of the suggested methods in the context of CKD or other contexts to help solve these problems. To do this, we conducted a literature review of the methodology used to investigate CKD risk factors in the last ten years. Second, the findings of this analysis of the literature were used to identify major methodological issues and to highlight some methods for

addressing them. The visible CKD findings analyzed in the selected papers are listed in the section on the following outcomes, and the methods of stagnation used in each type of outcome are defined and discussed. Where it is appropriate, we present some future research methods (Tahmid Abrar et al.,2019).

3.5 Artificial Intelligence

Artificial intelligence (AI) solutions may now be found in almost every medical and non-medical industry. In difficult medical conditions where the medical community has reached a stalemate, new algorithms have emerged (M.Hueso et al.,2018). Machine learning (ML) solutions were provided to medical registries for enhanced event prediction that outperformed human accuracy (H. M. Krumhozl et al.,2014). Because AI and machine learning "...have the potential to adapt and improve device performance in real-time to constantly enhance health care for patients..." the US Food and Drug Administration this year announced AI / ML-based software as a regulatory framework for medical devices (Administration Fad et al.,2019). Several challenges and complications caused by end-stage renal illness necessitating dialysis (T. A. Mayrakanas et al.,2016) (chronic kidney disease (CKD) stage G5D), the dialysis technique itself, or kidney transplantation (CKD stage G5T) (F. Reyna-Sepulyeda et al.,2017) have received incipient contributions from AI algorithms in the last 15 years.

A computer software that calculates and deduces task-related information and derives the features of the relevant pattern is referred to as machine learning (M. Alloghani et al.,2018). This technique can provide reliable and cost-effective illness diagnosis; therefore, it could be a potential way for detecting CKD. With the advancement of information technology (D. Gupta et al.,2016), it has evolved into a new type of medical instrument with a wide range of applications due to the rapid growth of electronic health records (L. Du et al., 2018). Machine learning has already been utilized in the medical industry to detect human body status (R. Abbas et al.,2018), assess disease-related parameters (M. Mahyoub et al.,2018).

3.6 Machine Learning

Machine learning (ML) is a kind of artificial intelligence that allows computers to employ intelligent software to do tasks more efficiently. The core of smart software that is utilized to generate machine intelligence is statistical learning approaches. Because data is required for machine learning algorithms to learn, the discipline must be linked to database training. Pattern recognition, data mining, and information discovery from data are all well-known concepts. One could be perplexed as to how to comprehend the massive image that depicts such a link (Mohammed et al., 2016).

Machine learning is one of the most promising and rapidly expanding areas of computer science. Different algorithms and functions help computers work efficiently. Machine learning is the process of teaching a computer to learn new algorithms to experience the machine in autonomous smart data processing. Machine learning is used in a variety of fields to improve the efficiency and accuracy of data processing.

Machine learning is based on efficient algorithms that employ a specific collection of tools and functions to address complicated and large data sets. Machine learning helps in a variety of sectors; these are often artificial intelligence applications that are used for recognition and prediction, such as in computer engineering and medicine. Machine learning is very common in today's computer technology, and it has a lot of advantages. As explained in the hybrid model, machine learning generates some rules for the input data that enable the machining process the similar cases each tie efficiently. It is very vital to understand how the variables in the inputs are moved into vectors because it operates on prediction. Machine learning has reduced the number of manual occupations available to people, reducing the potential for errors and inaccuracy (Smola & Vishwanathan, 2008).

An efficient algorithm offers two distinct advantages. For starters, technology can take the place of hard and repetitive human tasks. Second, and more importantly, it can extract more complicated and nuanced patterns from the input data than a human observer could. Radiation

therapy relies on both of these advantages. The regular contouring of organs and tumors in danger, for example, is a time-consuming pattern recognition procedure that is based on the knowledge and observer's intimacy with the appearance of anatomy in picture diagnosis. This familiarity, however, has its limits, resulting in ambiguity and inter-observer heterogeneity in the contours that arise. It is feasible that a contouring algorithm can combine the experience of many observers while simultaneously integrating input from different shapes or sources in one image, raising textural subtleties and therefore reducing contour uncertainty (Naqa & Murphy, 2015).

3.6.1 Machine Learning (ML) Algorithms

As illustrated in Figure 3-3, ML can be classified as supervised, unsupervised, semi-supervised, and reinforcement learning, based on the nature of the data.

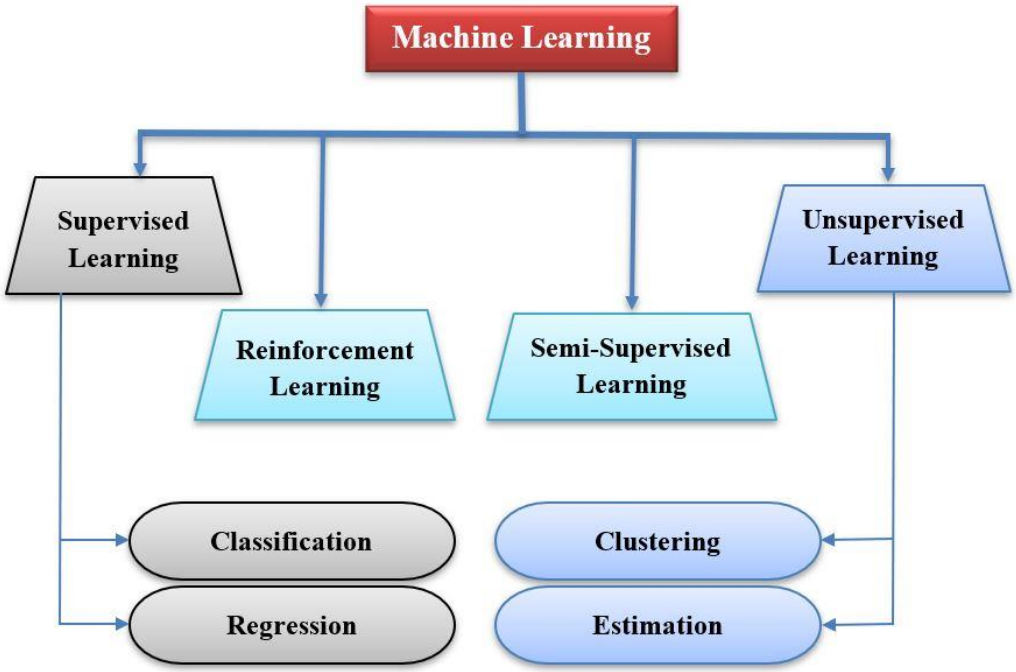


FIGURE 3.3: Machine Learning Algorithms Diagram.

3.6.1.1 Supervised Learning:

Pair of inputs and outputs is used for supervised learning. The goal of the learning process is to create the desired result for each entry. In this case, the cost function is related to the elimination of incorrect deductions. (Ojha et al., 2017). The root mean square error is a common cost; try to reduce the mean square error between the network output and the expected output. Pattern recognition (as well-referred to as classification) and regression are two tasks that lend themselves to supervised learning (also known as function approximation). Sequential data can benefit from supervised learning (e.g., for handwriting, speech, and gesture recognition).

3.6.1.2 Unsupervised Learning:

It's only natural to be concerned about classification issues. Provide only those examples that the learning model is permitted to use. For example, evaluation of the probability distribution function and clustering.

3.6.1.3 Semi-supervised Learning:

This type of learning combines both supervised and uncontrolled learning. It is used to create a suitable model for data classification by labeling a subset of the data and leaving the rest unlabeled. The labeled component of this design can be used to aid learning of the unlabeled portion. This type of approach applies to almost all-natural activities and more closely resembles how humans gain their skills. The goal of this learning type is to comprehend a system that can better predict classes of incoming test data than a system that just uses labeled data (Mohammed et al., 2016); as illustrated in Figure 3-4.

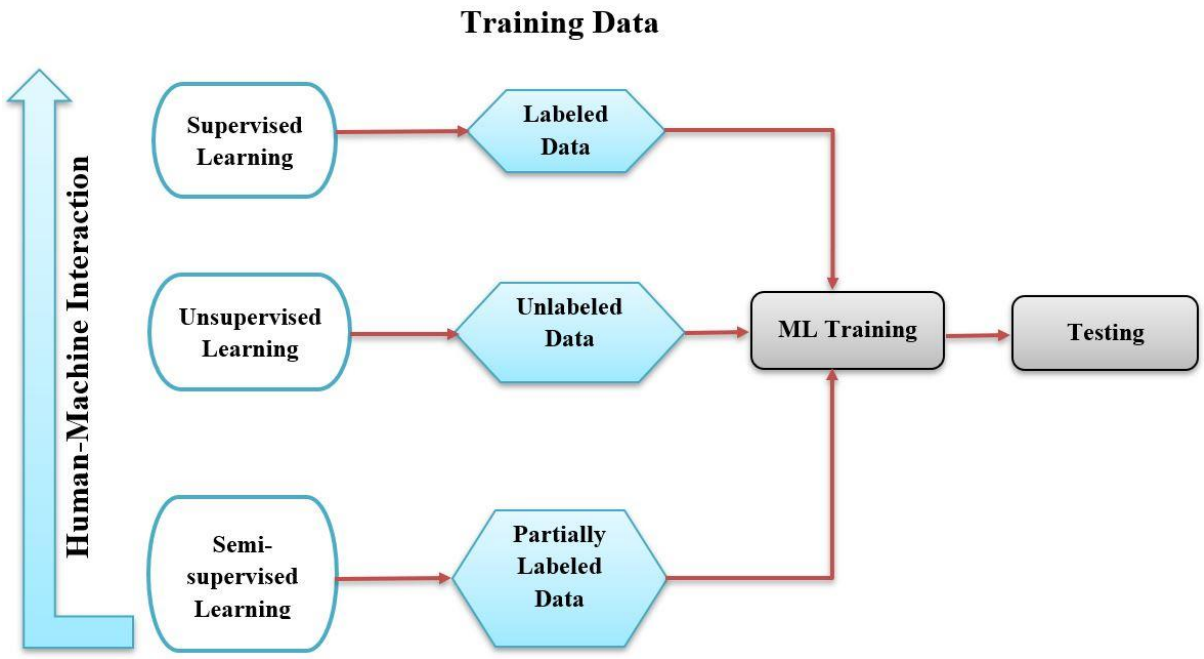


FIGURE 3.4: Training data nature for machine learning algorithm types.

3.6.1.4 The Reinforcement Learning

is a procedure that aims to employ observations gathered during the encounter, as well as the environment, to take actions that reduce risks or maximize benefits. Agents are used to creating intelligent systems. Reinforcement learning is used to carry out the following necessary process: The agent observes the input state in the first process. In the second process, the decision-making function is used to create an agent that executes an action. After the action is completed, the agent receives reinforcement from the environment or a reward in the third process. Finally, the state-action pair's reward knowledge is preserved. (Mohammed and colleagues, 2016). Figure 3-5 depicts this.

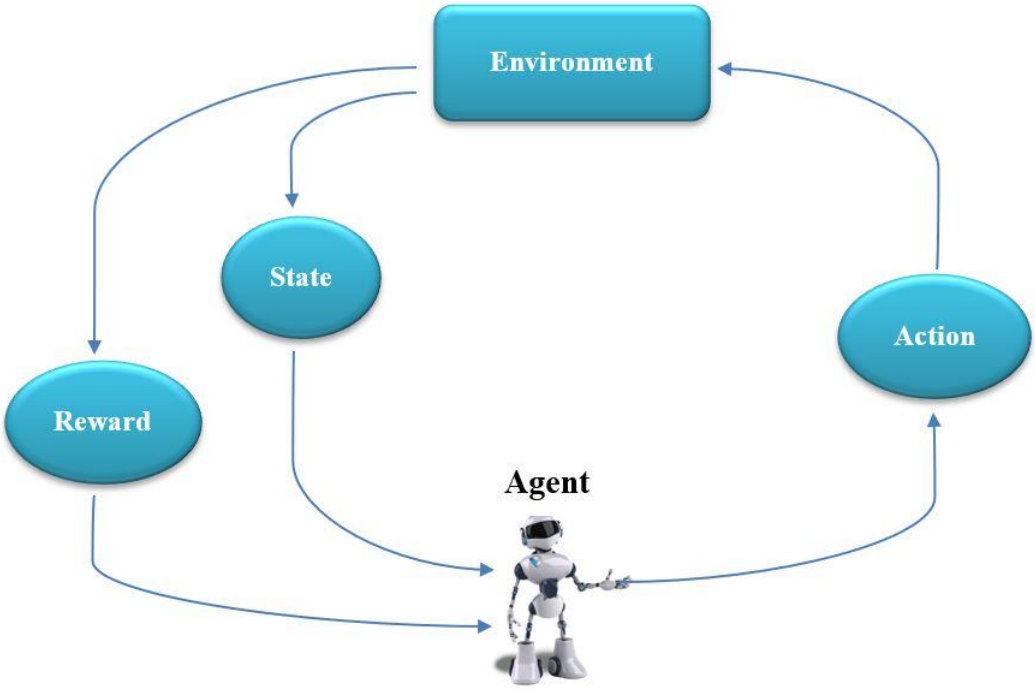


FIGURE 3.5: Reinforcement Learning Process.

3.7 Neural Network

3.7.1 Overview

With its amazing ability to learn, the human brain performs surprisingly better than digital computers when dealing with complex tasks such as recognition and prediction. The human brain is made up of hundreds of billions of nerve cells called neurons. The axons that connect these neurons to other cells are called axons. The human brain uses a network of interconnected neurons to communicate via electrical impulses to complete calculations. Dendrites receive stimuli from the outside world and information from the sense organs. These inputs generate electrical impulses that pass through the neural network. A neuron can send a message to another neuron to solve the problem, or it can ignore the message and solve the problem itself. The concept of ANN is inspired by the brain's ability to make precise

connections (DACS et al., 2014). Warren McCulloch and the neurophysiologist and mathematician Walter Pitts simulated a neuron in the 1940s. A basic switch with two inputs and one binary output serves as the basis for the model. The neuron's job is to receive information from the outside world. It receives information from other neurons and produces zero or one output, indicating that it is active. Based on the overall weighted input (A. Krogh et al., 2008), you can choose to be active or inactive. Rosenblatt later revealed in the 1960s that this patterned neural network can recognize and function, and some of the destroyed neurons behave similarly to the human brain (A. Krogh et al., 2008). The artificial neural network (ANN) is a mathematical model that simulates the function of neurons in the human brain. In real life, artificial neural networks are used for various tasks, including stock market forecasting (M. Kamuda et al., 2017). The ability and generalizability to learn to perform its function after adequate training, that is, the ability to provide satisfactory solutions to invisible data, are two basic attributes of the ANN. Classification, noise reduction, and extrapolation are the three most common problems ANN uses to address (A. Survey et al., 2017).

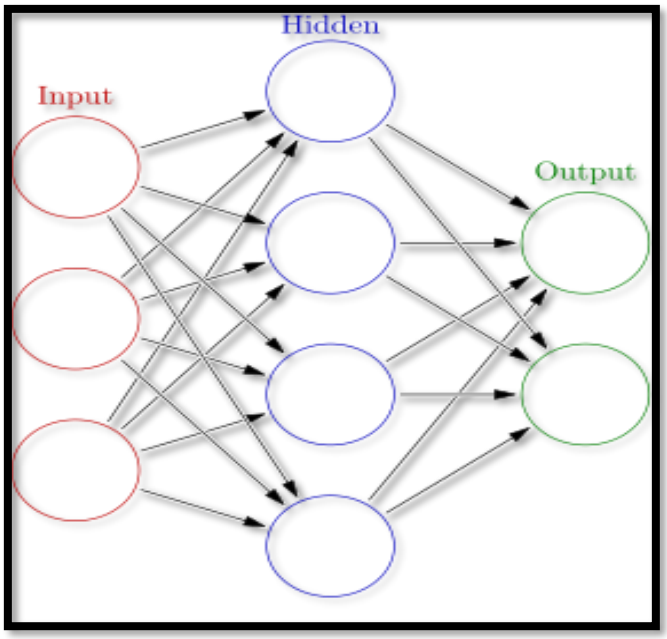


FIGURE 3.6: Neural Network Architecture.

3.7.2 Technical description

As indicated in Fig. 1, an Artificial Neural Network (ANN) can be separated into three layers:

1. **Input layer:** This one is where the data/features are received for the first time. These inputs are subjected to some normalization techniques to confine them to a specific range. The neural network's task is made easier by the normalized inputs, which results in increased precision.

2. **Hidden (intermediate or invisible) layer:** Depending on the network's application, this layer may be made up of several layers. These layers are in charge of identifying a process or system's pattern. In these layers, the majority of neural network operations are carried out.

3. **Output layer:** This one is also including neurons to represent the final network output generated by the neuron processing layer above. The principles of ANN composition include how neurons interact and how to generate layers; furthermore, its arrangement can be divided into four categories. Single-layer, mesh, loop, and multi-layer feedforward networks belong to these types. (I.N. da Silva et al.,2017).

3.8 Support Vector Machine (SVM)

Vapnik, Cortes, and Boser proposed a classifier based on statistical learning theory in the late 1990s [3]. Support Vector Machine (SVM) is the name of the classifier, which became widely known in the machine learning community later that year. SVM is a supervised machine learning technique used to solve classification and regression problems. For binary classification problems, SVM is the most widely used method. SVM is becoming more and more popular and is becoming one of the most widely used machine learning techniques. SVM has been used in various applications, including bioinformatics and handwriting

recognition (R.G. Brereton et al., 2010). Medical diagnosis, weather forecast, stock market analysis, and image processing are applications that use SVM. SVM is a calculation method, like all other machine learning algorithms, it learns from experience and examples to assign labels to objects. For example, SVM needs to filter a large number of real and fake credit card photos to distinguish them. The main function of SVM is to separate binary label data according to the line that reaches the maximum distance between the labels (T.S. Furey et al., 2000). The so-called "curse of dimensionality" affects most machine learning algorithms. When the model checks a small number of samples and has little experience with the existence of various features, it suffers from the curse of dimensionality. Due to such limitations, the performance of the model may be affected. SVM models are susceptible to the curse of dimensionality (A.A.A.A. Adewumi et al., 2016) (K.O. Akande et al., 2015) (C. Cortes et al., 1995). Even with fewer cases, SVM can work normally and has good accuracy. SVM is different from other machine learning methods (A.A.A.A. Adewumi et al., 2016) (K.O. Akande et al., 2015) (C. Cortes et al., 1995); Because of these advantageous advantages. To help separate tag data, SVM uses kernel functions. One of the benefits of using kernels in SVM is that the kernel definitions are applicable to non-vector inputs, that is, inputs that have no size or direction, which is very relevant in the medical field and vital in biological applications. Due to this benefit, SVM can mark DNA and protein sequences. Additionally, a mixture of multiple data types can be used to define the kernel (W.S. Noble et al., 2006).

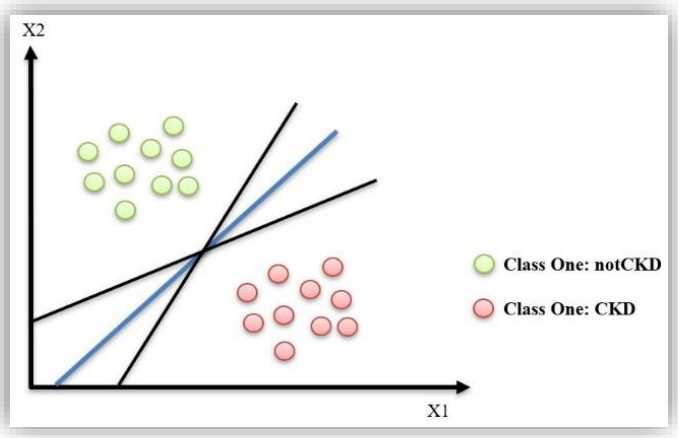


FIGURE 3.7: SVM for Multi-Class Classification.

3.9 Deep Learning Method

Deep Learning (DL) is a subtype of machine learning that focuses on discovering data patterns rather than solving specific problems, and it first gained popularity in 2012. According to (Goodfellow et al., 2016), DL is primarily an ML branch and a computer systems booster. In a dynamic, real-world setting, in most cases, deep learning is just a developer approach for AI systems content creation (text, picture, and music production) is a burgeoning field of deep learning applications that follow the two traditional machine learning tasks of classification and prediction, such as picture and speech translation and recognition.

The ability of deep learning architectures and training approaches has been utilized to autonomously learn musical genres from arbitrary musical corporations without the need for human user involvement, and then produce samples from the projected distribution. (Briot and Pachet, 2017).

The motivation for using pre-trained DL is that it saves time because it does not require a large data set to get results. These systems were also able to extract random characteristics from picture categorization. Lower-level features such as edges, texture, and color were retrieved by the higher layers. High-level features such as contours and objects were retrieved by the algorithms' bottom layers. Writings are often extracted from bottom levels by using pre-trained systems because the features on top layers of pre-trained systems are nearly the same in medical and natural images. Medical photos to natural photographs would be included in the bottom layers. The main purpose is to extract features from different layers of pre-trained systems that were trained on our dataset and are assembled to extract multi-scale information from input photos to improve the feature ability of the classifier model. Figure 3-8 shows AI, ML, and DL Introduction.

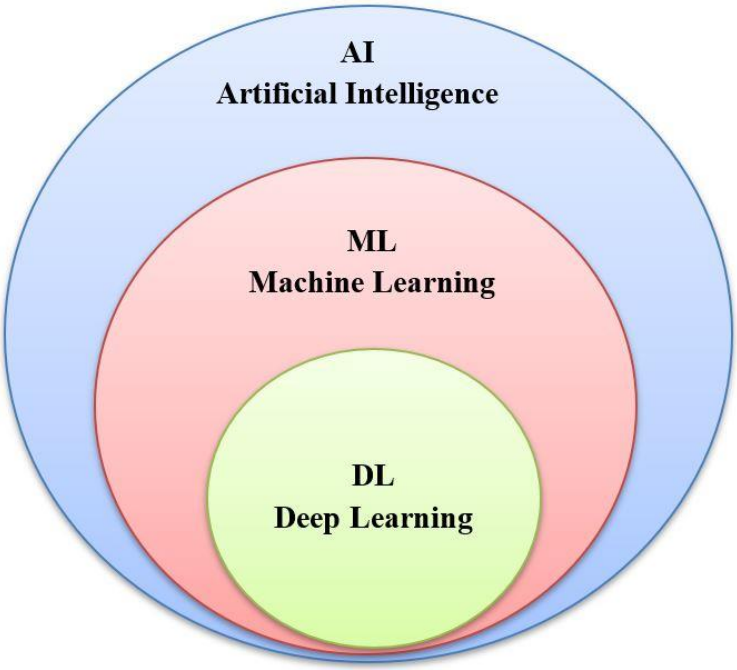


FIGURE 3.8: AI, ML, and DL Introduction.

3.10 Cross-Validation Method

Rotation estimation is the most common name for the cross-validation approach. It's a model validation approach for determining how well a statistical analytical result would generalize to a different dataset. In general, cross-validation is used in situations where the goal is to evaluate how well a predictive model will perform, i.e., where the goal is to forecast how well a model will perform. For a specific predictive problem, known data (training dataset) is usually supplied into the model for training, while unknown data (testing dataset) is utilized to validate the model.

Cross-validation also averages / combines fit measurements (prediction error) to more accurately estimate model prediction performance. In this study, we use cross-validation methods to describe the data set (validation data set) to test the model in the training phase.

We used a 10-time cross-validation in this survey, which involves randomly dividing the original data set into 10 equal subsamples. A single subsample from each of the ten subsamples is used as validation data to test the model, and the remaining nine subsamples are used as training data.

This method was then performed ten times, with each of the ten subsamples serving as the validation dataset exactly once each. The average of these ten possibilities was then used to create a single prediction. The main advantage of this method over arbitrarily iterated sub-sampling is that all perceptions are used for both training and validation, and each perception is used exactly once for validation (John Bush Idoko et al.,2018).

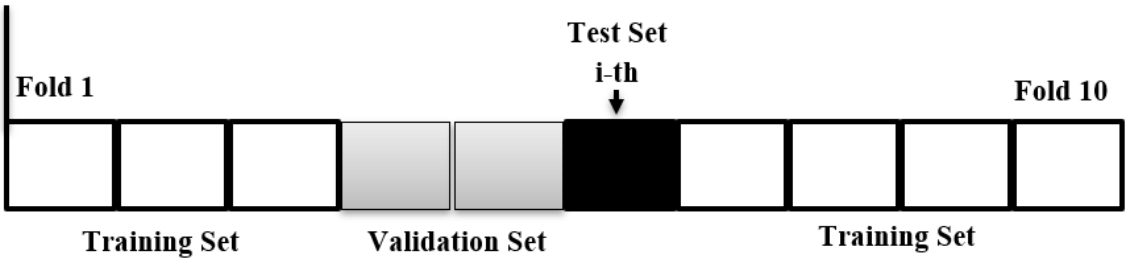


FIGURE 3.9: N-Fold Cross Validation Graphical

CHAPTER 4

CHRONIC KIDNEY DISEASE DETECTION USING FUZZY NEURAL NETWORK

4.1 FNN Model

The following is a summary of this thesis, we suggested a fuzzy neural network for the categorization of chronic kidney disease using the CKD data set. A fuzzy logic resembles human reasoning process. They're commonly utilized to tackle a variety of issues, including control, categorization, prediction, and identification. One of Zadeh's studies shows a typical explanation of fuzzy logic, in which he advocated fuzzy logic as a way to emulate the human brain's reasoning processes (Zadeh.,1996).

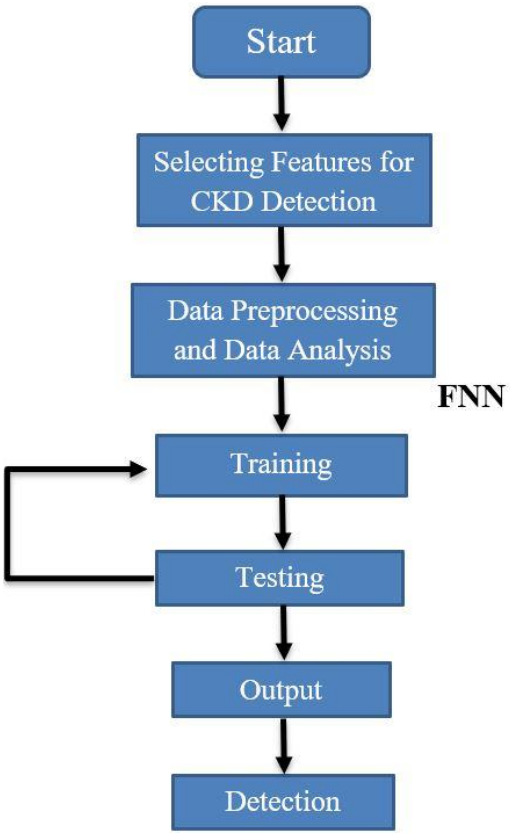


FIGURE 4.1: Flowchart for CKD Detection.

Fuzzy logic architectures are commonly used to tackle problems such as categorization, monitoring, consistency, and prognosis. Fuzzy logic is one of the leading simple and straightforward approaches for setting input area to its matching output area. Fuzzy logic is a simple and practical way to set an input area to an output area. The rules with a previous and eventual portion are used to do this. The input variables are included in the antecedent section. The system's output variables are included in the next section. A pattern recognition system's primary goal is to map inputs to their matching outputs, Patterns can be used as inputs, and classes can be used as outputs. Fuzzy values or linguistic phrases are used to describe the values of variables on a fuzzy rule basis. A membership function characterizes each fuzzy value.

We can use associateship functions to quantify linguistic terms. The fuzzy system design includes the accurate development of the antecedent and consequent aspects of the rules. The usage of neural networks is one of the most successful technologies for constructing if-then rules. Neural Networks have self-learning and generalization abilities, nonlinear mapping, parallel processing, and vitality.

The neural network-based model's accuracy can be improved because of its self-learning properties. Fuzzy logic provides for the reduction of data complexity as well as the handling of uncertainty and imprecision.

We can create an order with quick education capabilities that can explain nonlinearly orders with doubts using a mixture of fuzzy logic and neural networks.

These paths are used in this paper to create fuzzy neural networks and solve the pattern categorization problem. The input signals for the FNN base classifier come from a dataset obtained from chronic kidney disease diagnosis. This FNN classifier divides the domain's input signals into six distinct classes. The NN structure enabled our proposed algorithm (FNN) to achieve the process of fuzzy reasoning. The generation of an appropriate rule base using the "IF-THEN" structure was used to create this FNN. For the classification system utilizing learning capability, it is important to set the appropriate depiction of the fuzzy IF-THEN base

premise and the consequents' part (Abiyev RH et al.,2016) (Mamdani EH et al.,1975). This was accomplished by evaluating the created framework's error response. The frameworks were also designed using Mamdani and Takagi-Sugeno-Kang (TSK) fuzzy rules, sequentially (Takagi T et al.,1985) (Abiyev RH et al.,2008). The design of the system in this study is based on the latter, namely the TSK-type fuzzy rule. The fuzzy former and fragile eventual components of a TSK-type fuzzy rule are fuzzy and crisp, respectively. Fuzzy systems with the following structure are used to approximate nonlinearity and linearity:

$$\text{If } x_1 \text{ is } A_{1j} \text{ and } x_2 \text{ is } A_{2j} \text{ and} \ldots x_m \text{ is } A_{mj}. \quad \text{Then} \quad y_j = b_j + \sum_{i=1}^{m} a_{ij} x_i \tag{4.1}$$

The system's input and output signals are denoted by x_i and y_j respectively. i=1,..., m represents the number of input signals, and j = 1, ..., r represents the number of output signals. The fuzzy set's input is represented by A_{ij}, and the b_j and a_{ij} is used to represent the coefficients.

As illustrated in Figure 4.2, our FNN structure was used to distinguish the six layers of chronic kidney disease. Six layers make up the FNN suggested. The first layer uniformly distributes

$$\mu 1_j(x_i) = e^{-\frac{(x_i - c_{ij})^2}{\sigma_{ij}^2}}, \; i=1..m, \; j=1..r \tag{4.2}$$

The input signal xi of the system (i = 1 ..., m). Linguistic terms are described using the membership function in the second layer. The membership degree of the fuzzy set to which the input value belongs is estimated. taking into account each of the model's input signals. We use Gaussian membership functions to represent the language concepts proposed in (Abiyev RH, 2011), as follows:

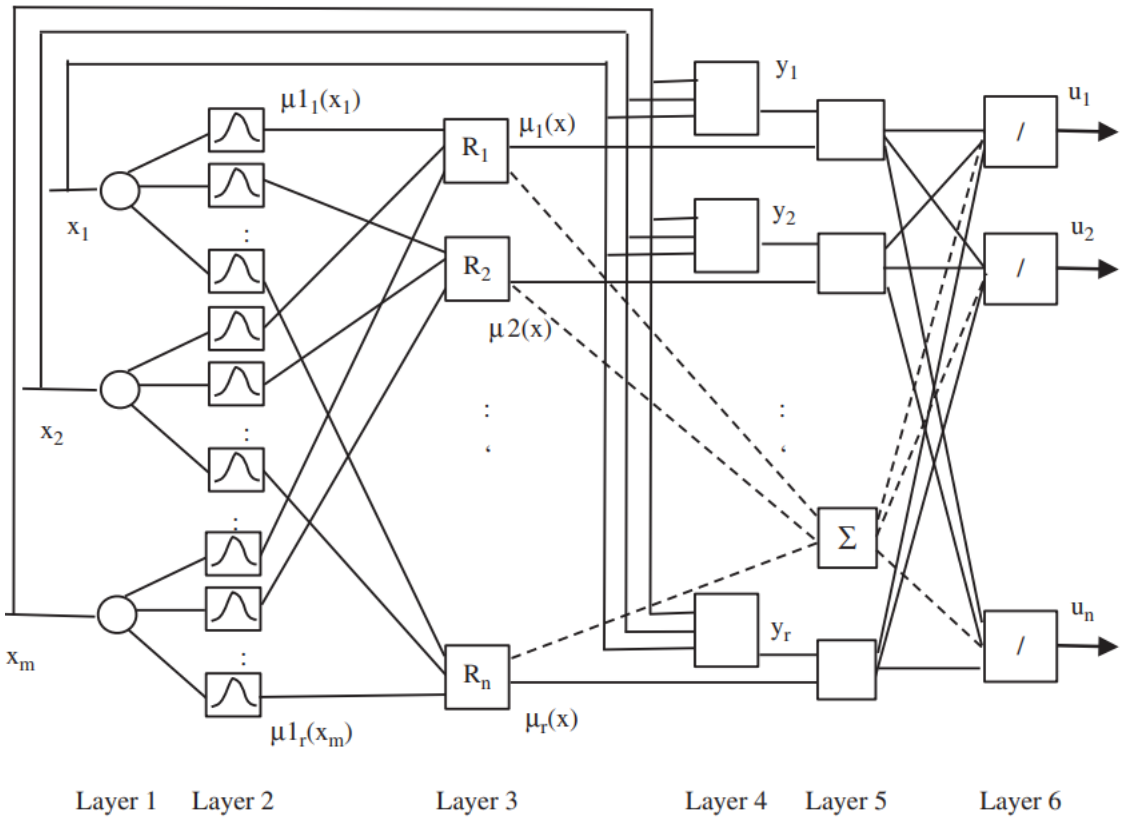


FIGURE 4.2: Fuzzy neural network (FNN) architecture.

Here, 'm' represents the number of input signals and 'r' represents the number of hidden neurons in the third layer, commonly known as the number of fuzzy rules. Represent respectively the center and the width of the Gaussian membership function. The membership function of the ith input variable of the jth term is 1j (xi) (Mamdani EH et al., 1975) (Takagi T, 1985) (Abiyev RH et al., 2008).

In our design, the rules layer is installed on the third layer, and the total number of nodes equals the total number of rules. The letters "R1, R2..., Rr" stand for the many rules. The t-norm min procedure is used to determine the output signal of this layer.

$$\mu_j(x) = \prod_i \mu 1_j(x_i), \quad i = 1, \ldots, m, \; j = 1, \ldots, r \tag{4.3}$$

In this case, the min operation is represented by the Π.

The input signals in the fifth layer were denoted by these j(x) signals. The subsequent layer was attached to the fourth layer. The 'n' linear systems that made up this layer. During this step, we calculated the following values for the rule output:

$$y1_j = b_j + \sum_{i=1}^{m} a_{ij} x_i \tag{4.4}$$

The fifth layer was induced by multiplying the third layer output signals μj (x) by output signals from the fourth layer. After then, the output of the "j-th" node was identified.

$$y_j = \mu_j(x) y1_j \tag{4.5}$$

The following is how the sixth layer determines FNN output signals:

$$u_k = \frac{\sum_{j=1}^{r} w_{jk} y_j}{\sum_{j=1}^{r} \mu_j(x)} \tag{4.6}$$

The network output signals (k=1,...,n) are denoted by u_k. As soon as the output signal of the fuzzy rule's subsequent component is determined, the network parameters are trained.

4.2 Parameter Learning of Fuzzy Neural Network (FNN)

The FNN parameters are initially created randomly. Fig. 6 The membership function (4.2) of the second-level fuzzy rule (4.1) and the parameters of the fourth and fifth-level linear functions (4.4) are these parameters. Parameter training of membership functions "$c_{ij}(t)$ and $\sigma_{ij}(t)$, $w_{jk}(t)$, $a_{ij}(t)$, $b_j(t)$, ($i = 1, ..., m$, $j = 1, ..., r$, $k = 1, ..., n$)". The FNN model is constructed in the latter part. The parameters of FNN were updated in this work through fuzzy grouping and gradient techniques [R. Abiyev et al., 2011] [R. H. Abiyev et al. 2016]. Use the fuzzy clustering of c-means to find the leading component parameter as the membership function parameter. After clustering, the parameters are trained using a gradient descent

learning algorithm with an adaptive learning rate. Using an adaptive learning rate can accelerate learning and ensure convergence. Calculate the value of the error cost function when learning the network output.

$$E = \frac{1}{2}\sum_{k=1}^{n}(u_k^d - u_k)^2 \tag{4.7}$$

Consequently, n denotes the network's number of output signals, and u_k^d and u_k denote the network's desired and current output values (k = 1,...,n). The following formulas are used to alter the network parameters "w_{jk}, a_{ij}, b_j, I = 1,...,m, j = 1,...,r, k = 1,...,n" and parameters of the membership function "c_{ij} and σ_{ij} I = 1,...,m, j = 1,...,r" of the structure of the FNN (R. Abiyev et al., 2011) (R. H. Abiyev et al.,2016).

$$\begin{aligned}
w_{jk}(t+1) &= w_{jk}(t) - \gamma\frac{\partial E}{\partial w_{jk}} + \lambda(w_{jk}(t) - w_{jk}(t-1)); \\
a_{ij}(t+1) &= a_{ij}(t) - \gamma\frac{\partial E}{\partial a_{ij}} + \lambda(a_{ij}(t) - a_{ij}(t-1)); \qquad (4.8) \\
b_j(t+1) &= b_j(t) - \gamma\frac{\partial E}{\partial b_j} + \lambda(b_j(t) - b_j(t-1)); \\
c_{ij}(t+1) &= c_{ij}(t) - \gamma\frac{\partial E}{\partial c_{ij}} + \lambda(c_{ij}(t) - c_{ij}(t-1)); \qquad (4.9) \\
\sigma_{ij}(t+1) &= \sigma_{ij}(t) - \gamma\frac{\partial E}{\partial \sigma_{ij}} + \lambda(\sigma_{ij}(t) - \sigma_{ij}(t-1)); \\
i &= 1,\ldots,m;\ j = 1,\ldots,r;\ k = 1,\ldots,n
\end{aligned}$$

Here, m represents the number of input signals (input neurons) and "r" represents the number of fuzzy rules (hidden neurons), γ represents the learning rate, λ and represents

momentum. In (4.8), the derivatives are computed as follows:

$$\begin{aligned}
\frac{\partial E}{\partial w_{jk}} &= \frac{\partial E}{\partial u_k}\frac{\partial u_k}{\partial w_{jk}} = (u_k(t) - u_k^d(t)) \cdot y1_j \Big/ \sum_{j=1}^{n} \mu_j \\
\frac{\partial E}{\partial a_{ij}} &= \frac{\partial E}{\partial u_k}\frac{\partial u_k}{\partial y1_j}\frac{\partial y1_j}{\partial y_j}\frac{\partial y_j}{\partial a_{ij}} \\
&= \sum_k (u_k(t) - u_k^d(t)) \cdot w_{kj}\mu_j x_i \Big/ \sum_{j=1}^{n} \mu_j \\
\frac{\partial E}{\partial b_j} &= \frac{\partial E}{\partial u_k}\frac{\partial u_k}{\partial y1_j}\frac{\partial y1_j}{\partial y_j}\frac{\partial y_j}{\partial b_j} \\
&= \sum_k (u_k(t) - u_k^d(t)) \cdot w_{kj}\mu_j \Big/ \sum_{j=1}^{n} \mu_j
\end{aligned} \tag{4.10}$$

In (4.9), the derivatives are calculated as follows:

$$\begin{aligned}
\frac{\partial E}{\partial c_{ij}} &= \sum_k \frac{\partial \mathrm{E}}{\partial u_k}\frac{\partial u_k}{\partial \mu_j}\frac{\partial \mu_j}{\partial c_{ij}} \\
\frac{\partial E}{\partial \sigma_{ij}} &= \sum_k \frac{\partial \mathrm{E}}{\partial u_k}\frac{\partial u_k}{\partial \mu_j}\frac{\partial \mu_j}{\partial \sigma_{ij}}
\end{aligned}$$

Here, $i = 1, \ldots, m,\ j = 1, \ldots, r,\ k = 1, \ldots, n.$

$$\begin{aligned}
&\frac{\partial E}{\partial u_k} = u_k(t) - u_k^d(t); \quad \frac{\partial u_k}{\partial \mu_j} = \frac{y_j - u_k}{\sum_{j=1}^{n} \mu_j} \\
&\frac{\partial \mu_j(x_i)}{\partial c_{ij}} = \mu_j(x_i)\frac{2(x_i - c_{ij})}{\sigma_{ij}^2}; \\
&\frac{\partial \mu_j(x_i)}{\partial \sigma_{ij}} = \mu_j(x_i)\frac{2(x_i - c_{ij})^2}{\sigma_{ij}^3}
\end{aligned} \tag{4.11}$$

The derivatives in (4.8) and (4.9) are computed using Eqs. (4.10)– (4.11), and the FNN parameters are corrected using Eqs.

4.3 CKD Dataset

The CKD dataset, in our thesis, was collected from Dr. P. Soundarapandian's open-source research library at the University of California Irvine (UCI) (Tahmid Abrar et al.,2019). The data collection contains 25 attributes from a total of 400 patients, 250 of whom have CKD and 150 of whom are healthy. Individual patients in the dataset are between the ages of 60 and 90. The dataset's average density fluctuates between one.005 and 1.025, while each glucose and albumen value ranges between “0-5”. “BP” Ranges of Patients ‘Hemoglobin’ (HEMO) is measured between “5.6 and 17.7 gms”, “PCV” is 16 53, “WBCC” is a pair of otI 26400 cell/cumm, “RBCC” is two point one ‘8.0’ million/cumm, blood sugar random (BGR) is “seventy 490 mgs/dl”, and “BU” is ‘fifteen 424 mgs/cum’, the dataset includes measurements of body fluid creatinine (SC) ranging from “0 to 48.1 mgs/dl”, metallic element “SOD 4.5 150 mEq,'L”, and potassium (POT) a pair of.5 7.6 mEqL. (outliers). The dataset's remaining alternatives were painted in binary: gift or absence, yes or no, smalt or harmful. Table 4 displays the dataset's attribute and abbreviation, while Table 4.1 shows the dataset.

TABLE 4.1: CKD Dataset description.

No.	Attribute	Abbreviation	Type	Value
1	Age	Age	Numerical	Years
2	Blood Pressure	Bp	Numerical	mm/Hg
3	Specific Gravity	Sg	Nominal	1.005, 1.010, 1.015, 1.020, 1.025
4	Albumin	Al	Nominal	0, 1, 2, 3, 4, 5
5	Sugar	Su	Nominal	0, 1, 2, 3, 4, 5
6	Red Blood Cells	Rbc	Nominal	Normal, Abnormal
7	Pus Cell	Pc	Nominal	Normal, Abnormal
8	Pus Cell Clumps	Pcc	Nominal	Present, Not present
9	Bacteria	Ba	Nominal	Present, Not present
10	Blood Glucose Random	Bgr	Numerical	mgs/dl
11	Blood Urea	Bu	Numerical	mgs/dl
12	Serum Creatinine	Sc	Numerical	mgs/dl
13	Sodium	Sod	Numerical	mEq/L
14	Potassium	Pot	Numerical	mEq/L
15	Hemoglobin	Hemo	Numerical	Gms
16	Packed Cell Volume	Pcv	Numerical	_____
17	White Blood Cell Count	Wc	Numerical	cells/cumm
18	Red Blood Cell Count	Rc	Numerical	millions/cmm
19	Hypertension	Htn	Nominal	Yes, No
20	Diabetes Mellitus	Dm	Nominal	Yes, No
21	Coronary Artery Disease	Cad	Nominal	Yes, No
22	Appetite	Appet	Nominal	Good, Poor
23	Pedal Edema	Pe	Nominal	Yes, No
24	Anemia	Ane	Nominal	Yes, No
25	Class	Class	Nominal	Ckd, notckd

TABLE 4.2: Dataset for 400 patients with Missing Values.

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y
1	age	bp	sg	al	su	rbc	pc	pcc	ba	bgr	bu	sc	sod	pot	hemo	pcv	wc	rc	htn	dm	cad	appet	pe	ane	class
2	48	80	1.02	1	0	?	normal	notpresent	notpresent	121	36	1.2	?	?	15.4	44	7800	5.2	yes	yes	no	good	no	no	ckd
3	7	50	1.02	4	0	?	normal	notpresent	notpresent	?	18	0.8	?	?	11.3	38	6000	?	no	no	no	good	no	no	ckd
4	62	80	1.01	2	3	normal	normal	notpresent	notpresent	423	53	1.8	?	?	9.6	31	7500	?	no	yes	no	poor	no	yes	ckd
5	48	70	1.005	4	0	normal	abnormal	present	notpresent	117	56	3.8	111	2.5	11.2	32	6700	3.9	yes	no	no	poor	yes	yes	ckd
6	51	80	1.01	2	0	normal	normal	notpresent	notpresent	106	26	1.4	?	?	11.6	35	7300	4.6	no	no	no	good	no	no	ckd
7	60	90	1.015	3	0	?	?	notpresent	notpresent	74	25	1.1	142	3.2	12.2	39	7800	4.4	yes	yes	no	good	yes	no	ckd
8	68	70	1.01	0	0	?	normal	notpresent	notpresent	100	54	24	104	4	12.4	36	?	?	no	no	no	good	no	no	ckd
9	24	?	1.015	2	4	normal	abnormal	notpresent	notpresent	410	31	1.1	?	?	12.4	44	6900	5	no	yes	no	good	yes	no	ckd
10	52	100	1.015	3	0	normal	abnormal	present	notpresent	138	60	1.9	?	?	10.8	33	9600	4	yes	yes	no	good	no	yes	ckd
11	53	90	1.02	2	0	abnormal	abnormal	present	notpresent	70	107	7.2	114	3.7	9.5	29	12100	3.7	yes	yes	no	poor	no	yes	ckd
12	50	60	1.01	2	4	?	abnormal	present	notpresent	490	55	4	?	?	9.4	28	?	?	yes	yes	no	good	no	yes	ckd
13	63	70	1.01	3	0	abnormal	abnormal	present	notpresent	380	60	2.7	131	4.2	10.8	32	4500	3.8	yes	yes	no	poor	yes	no	ckd
14	68	70	1.015	3	1	?	normal	present	notpresent	208	72	2.1	138	5.8	9.7	28	12200	3.4	yes	yes	yes	poor	yes	no	ckd
15	68	70	?	?	?	?	?	notpresent	notpresent	98	86	4.6	135	3.4	9.8	?	?	?	yes	yes	yes	poor	yes	no	ckd
16	68	80	1.01	3	2	normal	abnormal	present	present	157	90	4.1	130	6.4	5.6	16	11000	2.6	yes	yes	yes	poor	yes	no	ckd
17	40	80	1.015	3	0	?	normal	notpresent	notpresent	76	162	9.6	141	4.9	7.6	24	3800	2.8	yes	no	no	good	no	yes	ckd
18	47	70	1.015	2	0	?	normal	notpresent	notpresent	99	46	2.2	138	4.1	12.6	?	?	?	no	no	no	good	no	no	ckd
19	47	80	?	?	?	?	?	notpresent	notpresent	114	87	5.2	139	3.7	12.1	?	?	?	yes	no	no	poor	no	no	ckd
20	60	100	1.025	0	3	?	normal	notpresent	notpresent	263	27	1.3	135	4.3	12.7	37	11400	4.3	yes	yes	yes	good	no	no	ckd
21	62	60	1.015	1	0	?	abnormal	present	notpresent	100	31	1.6	?	?	10.3	30	5300	3.7	yes	no	yes	good	no	no	ckd
22	61	80	1.015	2	0	abnormal	abnormal	notpresent	notpresent	173	148	3.9	135	5.2	7.7	24	9200	3.2	yes	yes	yes	poor	yes	yes	ckd

CKD_Dataset_With_Missing_Values

4.4 Data Pre-processing

The dataset we used from UCI in this thesis has a lot of blank or null data that needs to be solved in this pre-processing stage, as seen in Table 2. Furthermore, non-numerical values must be translated to numerical values. We transformed all non-numerical values to numerical using the following procedure:

```
data['class'] = data['class'].map({'ckd':1, 'notckd':0})
data['htn'] = data['htn'].map({'yes':1, 'no':0})
data['dm'] = data['dm'].map({'yes':1, 'no':0})
data['cad'] = data['cad'].map({'yes':1, 'no':0})
data['appet'] = data['appet'].map({'good':1, 'poor':0})
data['ane'] = data['ane'].map({'yes':1, 'no':0})
data['pe'] = data['pe'].map({'yes':1, 'no':0})
data['ba'] = data['ba'].map({'present':1, 'notpresent':0})
data['pcc'] = data['pcc'].map({'present':1, 'notpresent':0})
data['pc'] = data['pc'].map({'abnormal':1, 'normal':0})
data['rbc'] = data['rbc'].map({'abnormal':1, 'normal':0})
```

FIGURE 4.3: The process of converting the numerical values.

TABLE 4.3: Missing value representation.

Attribute	Missing Values	% Of Total Values
Rbc	152	38.00
Rc	131	32.75
Wc	106	26.50
Pot	88	22.00
Sod	87	21.75
Pcv	71	17.75
Pc	65	16.25
Hemo	52	13.00
Su	49	12.25
Sg	47	11.75
Al	46	11.50
Bgr	44	11.00
Bu	19	4.75
Sc	17	4.25
Bp	12	3.00
Age	9	2.25
Ba	4	1.00
Pcc	4	1.00
Htn	2	0.50
Dm	2	0.50
Cad	2	0.50
Appet	1	0.25
Pe	1	0.25
Ane	1	0.25

Table 4.3 shows the rate of the missing value before preprocessing. Now, to preprocessing the dataset, we propose four cases,

1. Remove every missing value.
2. Substitute the minimum values for all missing values.
3. Use max to fill in any missing values.
4. Substitute mean values for all missing values.

To begin, when the dataset is transformed to a data frame with pandas, all blank values are replaced with NaN. The aforesaid methods are then used for constructing cases, as the medical data is unbalanced and only 158 patients are available, using the “dropna( )” and “fillna(parameters)” functions. We'll weigh all of the possibilities to come up with the best-case scenario.

TABLE 4.4: CKD Datasets after finding Missing values.

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y
1	age	bp	sg	al	su	rbc	pc	pcc	ba	bgr	bu	sc	sod	pot	hemo	pcv	wc	rc	htn	dm	cad	appet	pe	ane	class
2	48	80	1.02	1	0	0.8097	1	0	0	121	36	1.2	137.5176	4.630868	15.4	44	7800	5.2	1	1	0	1	0	0	1
3	7	50	1.02	4	0	0.8097	1	0	0	148.085	18	0.8	137.5176	4.630868	11.3	38	6000	4.7022	0	0	0	1	0	0	1
4	62	80	1.01	2	3	1	1	0	0	423	53	1.8	137.5176	4.630868	9.6	31	7500	4.7022	0	1	0	0	0	1	1
5	48	70	1.005	4	0	1	0	1	0	117	56	3.8	111	2.5	11.2	32	6700	3.9	1	0	0	0	1	1	1
6	51	80	1.01	2	0	1	1	0	0	106	26	1.4	137.5176	4.630868	11.6	35	7300	4.6	0	0	0	1	0	0	1
7	60	90	1.015	3	0	0.8097	0.77	0	0	74	25	1.1	142	3.2	12.2	39	7800	4.4	1	1	0	1	1	0	1
8	68	70	1.01	0	0	0.8097	1	0	0	100	54	24	104	4	12.4	36	8411.6041	4.7022	0	0	0	1	0	0	1
9	24	76.46	1.015	2	4	1	0	0	0	410	31	1.1	137.5176	4.630868	12.4	44	6900	5	0	1	0	1	1	0	1
10	52	100	1.015	3	0	1	0	1	0	138	60	1.9	137.5176	4.630868	10.8	33	9600	4	1	1	0	1	0	1	1
11	53	90	1.02	2	0	0	0	1	0	70	107	7.2	114	3.7	9.5	29	12100	3.7	1	1	0	0	0	1	1
12	50	60	1.01	2	4	0.8097	0	1	0	490	55	4	137.5176	4.630868	9.4	28	8411.6041	4.7022	1	1	0	1	0	1	1
13	63	70	1.01	3	0	0	0	1	0	380	60	2.7	131	4.2	10.8	32	4500	3.8	1	1	0	0	1	0	1
14	68	70	1.015	3	1	0.8097	1	1	0	208	72	2.1	138	5.8	9.7	28	12200	3.4	1	1	1	0	1	0	1
15	68	70	1.017386	1.02	0.451	0.8097	0.77	0	0	98	86	4.6	135	3.4	9.8	38.8	8411.6041	4.7022	1	1	1	0	1	0	1
16	68	80	1.01	3	2	1	0	1	1	157	90	4.1	130	6.4	5.6	16	11000	2.6	1	1	1	0	1	0	1
17	40	80	1.015	3	0	0.8097	1	0	0	76	162	9.6	141	4.9	7.6	24	3800	2.8	1	0	0	1	0	1	1
18	47	70	1.015	2	0	0.8097	1	0	0	99	46	2.2	138	4.1	12.6	38.8	8411.6041	4.7022	0	0	0	1	0	0	1
19	47	80	1.017386	1.02	0.451	0.8097	0.77	0	0	114	87	5.2	139	3.7	12.1	38.8	8411.6041	4.7022	1	0	0	0	0	0	1
20	60	100	1.025	0	3	0.8097	1	0	0	263	27	1.3	135	4.3	12.7	37	11400	4.3	1	1	1	1	0	0	1
21	62	60	1.015	1	0	0.8097	0	1	0	100	31	1.6	137.5176	4.630868	10.3	30	5300	3.7	1	0	1	1	0	0	1
22	61	80	1.015	2	0	0	0	0	0	173	148	3.9	135	5.2	7.7	24	9200	3.2	1	1	1	0	1	1	1

CKD_Dataset

CHAPTER 5

SIMULATION AND RESULTS

5.1 Simulation

The detection of chronic kidney disease is the main problem considered in this thesis. The study's major goal and purpose are to use machine learning in particular FNN model to detect chronic renal disease with accuracy and precision. The simulation results have been provided in order to demonstrate the suitability of using FNN model in diagnosing chronic renal disease. For training and testing validation, the model's accuracy, sensitivity, specificity, and F-1 score are examined. As a result, in the pre-processing phase, the study contains and evaluates the source data, and then uses this data in the stimulation or processing phase to discover the output that is predicting effectively and efficiently.

5.2 Performance Evaluation Measurements

The models, as well as their accuracy and standard deviation, can be derived by comparing different ML algorithms and various parameter values. Additional criteria should be considered as well. We adhere to the traditional definitions of all performance measurements. The fundamental terms are as follows:

- TP: the figure of CKD cases that were rightly detected as CKD.
- FP: this refers to the number of healthy people who are mistakenly diagnosed with CKD.
- TN: the figure of healthy instances that are rightly forecasted as healthy.
- FN: this is the figure of CKD cases that are incorrectly categorized as healthy.

5.2.1 Accuracy

Accuracy is a statistic for determining how well a certain class predicts a given amount over a given sample size. The calculation can be done in the following way:

$$\text{Accuracy} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{FP} + \text{TN} + \text{FN}} \tag{5.1}$$

5.2.2. Precision

Precision is defined as the ratio of accurately predicted CKD cases to the total number of anticipated CKD cases, and is computed as follows:

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}} \tag{5.2}$$

5.2.3. Recall

Recall or sensitivity has been clarified as the rate of correctly anticipated figure of CKD cases to the synoptic figure of CKD cases, and is computed as follows:

$$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}} \tag{5.3}$$

5.2.4. F1-Score

Combining precision and recall into a single statistic is another important metric. It is the precision and recall mean of harmony. The mathematical equation for F1-Score is as follows:

(5.4)

$$\mathrm{F1}_{\mathrm{Score}} = 2 * \frac{\mathrm{P} * \mathrm{R}}{\mathrm{P} + \mathrm{R}} = \frac{TP}{TP + \frac{1}{2}(FP + FN)}$$

5.3 Simulation Results

The classification system was designed using a dataset of 24 input and 1 output signal. Training, evaluation, and testing was done with the input/output data pairs. A varying number of rules were employed to test the classifier during simulation. The membership functions of the preceding part and subsequent part parameters were first constructed at random. The output signals were computed using input signals and the FNN structure; the deviation of current outputs from goal outputs was then determined using the FNN output. For calculating the root mean square error, this deviation was used (RMSE). The FNN system performance was calculated using the RMSE and recognition rate. The following formula was used to calculate the RMSE:

$$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(y_i^d - y_i)^2} \tag{5.5}$$

N represents the number of samples, and y_i^d and y_i are the destination and current exit sign, respectively. The learning of FNN was done using RMSE values.

We used the following formula to calculate the recognition rate:

$$RecofnitionRate = \frac{Number\ ofitems\ correctly\ classified}{Total\ number\ of\ items} * 100\% \tag{5.6}$$

The FNN's parameters were learned using fuzzy c-means rating and slope descent techniques. Through choose the cores of the membership methods of the input and cryptic layers, the input was first fed to a fuzzy c-means classifier (John Bush Idoko et al.,2018).

The simulations have been done using different number of rules (hidden neurons). We use 5, 8 and 16 rules and fixed performance characteristics of the system. Figure 5.1, 5.2 and 5.3 depicts convergence graphics of the FNN system obtained during learning using 5, 8 and 16 rules. Table 5.1 depicts training, validation RMSE obtained during training, test errors and accuracies of the system obtained using test data. As shown by increasing number of rules the values of RMSE errors are decreased and accuracies are increases. Using 16 rules, we have obtained the following performance characteristics for the FNN model. The accuracy was 99.75%, sensitivity was 100%, specificity was 99.34%, precision was 99.6%. True Positive=249, True Negative=150, False Positive=1, and False Negative=0. F1 score is obtained as F1=249/ (249+0.5(1+0)) = 0.998.

TABLE 5.1: The FNN model's experimental results.

Number of rules (Hidden neurons)	**Training**		**Testing**	
	Training Error	**Evaluation Error**	**Test Error**	**Accuracy**
5	0.356755	0.356648	0.356078	92%
8	0.255578	0.338906	0.344146	98%
16	0.163081	0.169469	0.167446	**99.75%**

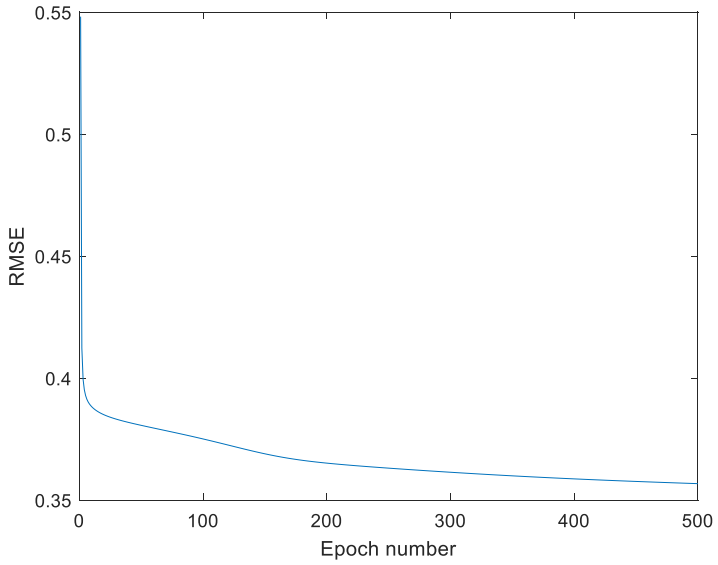


FIGURE 5.1: The plot of root means square error (RMSE) obtained from fuzzy neural network (FNN) classifier by using 5 rules (hidden neuron).

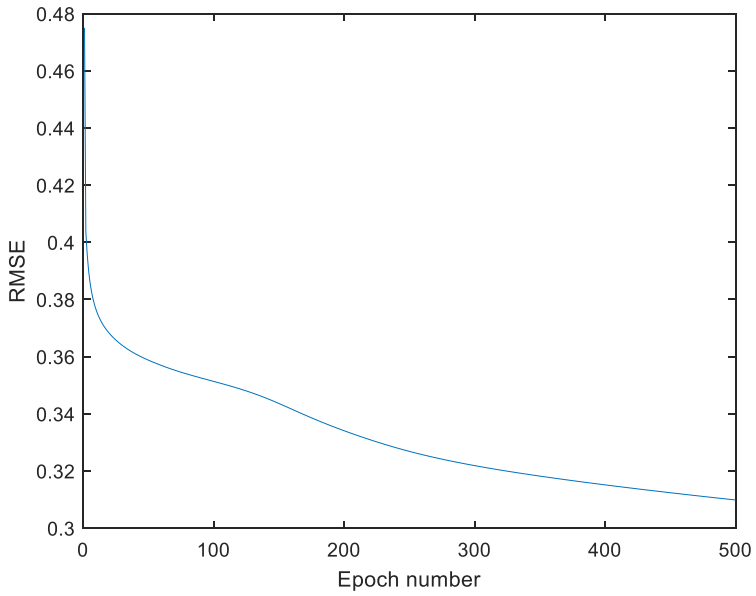


FIGURE 5.2: The plot of root means square error (RMSE) obtained from fuzzy neural network (FNN) classifier by using 8 rules (hidden neuron).

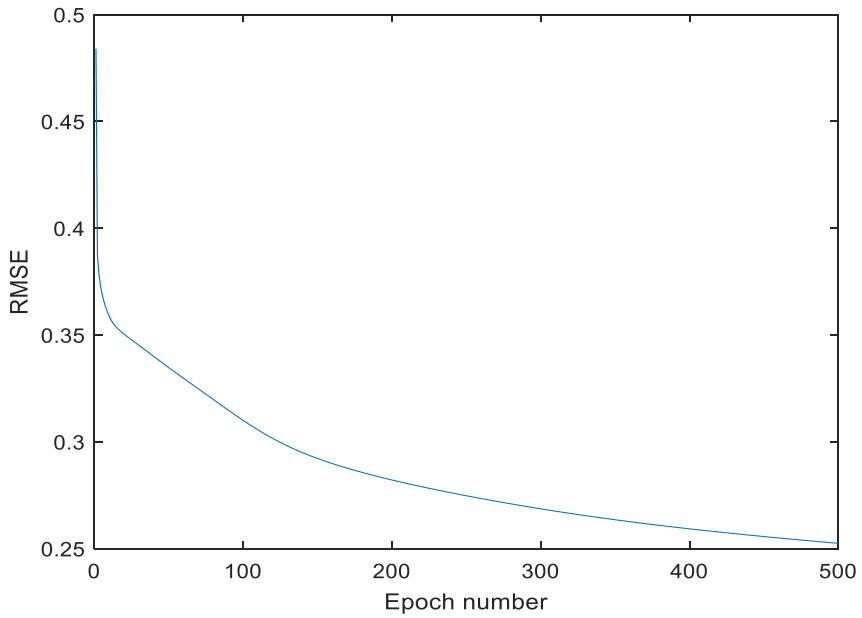


FIGURE 5.3: The plot of root means square error (RMSE) obtained from fuzzy neural network (FNN) classifier by using 16 rules (hidden neuron).

5.4 Comparisons and related works using CKD dataset

In this section, we compared our model result with some other related works that using the same dataset (CKD dataset). as clear as that our model is a fuzzy neural network (FNN) that contains neuron fuzzy and neural network, we make these comparisons by the model's accuracy rating.The simulation of FNN based CKD detection system have been done using different number of rules: 5, 8 and 16. The model accuracy results with the 5 rules was obtained as 92%, with the 8 rules- 98%. We have obtained better results using 16 rules. The model accuracy result was obtained as 99.75%, sensitivity- 100%, specificity- 99.34%, precision was 99.6%, F1 score- 99.8%. Table 5.2 depicts comparative results of different models used for detection of kidney diseases. As shown FNN model has obtained better result than other models. Comparative results demonstrate the efficiency of FNN model in kidney diseases detection.

TABLE 5.2: Comparisons and related works using CKD dataset.

Authors	**Model**	**Dataset**	**Accuracy (%)**
Xun L.et al	RBF	CKD dataset	82.1 %
Ravindra et al	SVM & RBK function	CKD dataset	93.75 %
B.Boukenze et al	ANN & SVM	CKD dataset	62.5 %
Panwong P. et al	Random forest	CKD dataset	86.6 %
U.Dulhare et al	NB	CKD dataset	97.5 %
Abeer et al	SVM	CKD dataset	93.1 %
R.Dhruvi et al	Naïve Bayes	CKD dataset	62.5 %
Padmanaban et al	Decision Tree	CKD dataset	91 %
Proposed Method	**FNN with 5 rules (hidden neurons)**	**CKD dataset**	**92 %**
Proposed Method	**FNN with 8 rules (hidden neurons)**	**CKD dataset**	**98 %**
Proposed Method	**FNN with 16 rules (hidden neurons)**	**CKD dataset**	**99.75 %**

CHAPTER 6

CONCLUSIONS

Chronic kidney disease is unquestionably one of the most difficult deadly diagnoses to diagnose with great accuracy and precision. In a nutshell, developing an application for detecting chronic sickness will benefit not only medical experts in treating crucial situations but also people who have difficulty contacting a doctor. One of the reasons for the difficulty in diagnosing CKD is that it is difficult to anticipate and detect because it is not dependent on a single parameter. Furthermore, common CKD symptoms do not play a substantial role in diagnosing the condition. The purpose of this study is to see how well the FNN classifier performs when used on a real-world dataset to diagnose chronic kidney disease. We combined the neuron fuzzy and neural network learning skills to simplify the uncertainties observed in the dataset. The design of fuzzy neural networks for chronic kidney disease detection is given in this research. Fuzzy neural networks are proposed for the detection of kidney diseases. The CKD data set was taken from the machine learning repository at the University of California, Irvine (UCI), The dataset includes the statistics of 400 patients that have 24 input attributes characterizing the related chronic kidney disease. The Fuzzy Neural Network (FNN) utilized as a classification approached in this study. The design of the FNN is performed using a different number of rules (hidden neurons). Learning of the system has been performed using a cross-validation approach by applying a gradient descent algorithm. To conduct the learning, the mean of the respective attributes was used to replace all missing values in the CKD dataset. After learning, the designed FNN system is applied for the detection of diseases. The simulation of FNN based CKD detection system have been done using different number of rules: 5, 8 and 16. We have obtained better results using 16 rules. The model accuracy was obtained as 99.75%, sensitivity- 100%, specificity- 99.34%, precision was 99.6%, F1 score- 99.8%. Comparative results were provided in order to demonstrate the efficiency of FNN model in kidney diseases detection. The obtained experimental and comparative results demonstrate the efficiency of using the FNN system in diagnosing chronic kidney diseases.

REFERENCES

A.Salekin and J. Stankovic, (2016) “Detection of chronic kidney disease and selecting important predictive attributes,” IEEE International Conference on Healthcare Informatics.

Kunwar, V., Chandel, K., Sabitha, A. S., & Bansal, A. (2016). Chronic Kidney Disease analysis using data mining classification techniques. In 2016 6th International Conference- Cloud System and Big Data Engineering (Confluence) (pp. 300-305).

Chronic_Kidney_Disease Data Set [Online].Available: https://archive.ics.uci.edu/ml/datasets/chronic_kidney_disease.

D. Gupta, S. Khare, and A. Aggarwal, (2016)“A method to predict diagnostic codes for chronic diseases using machine learning techniques,” 2016 International Conference on Computing, Communication and Automation (ICCCA), pp. 281–287.

A. Y. Al-Hyari, A. M. Al-Taee and M. A. Al-Taee, (2013) "Clinical decision support system for diagnosis and management of Chronic Renal Failure," 2013 IEEE Jordan Conference on Applied Electrical Engineering and Computing Technologies (AEECT), Amman, pp. 1-6.

Q. Zheng, G. Tasian, and Y. Fan, “Transfer learning for diagnosis of congenital abnormalities of the kidney and urinary tract in children based on Ultrasound imaging data,” International Symposium on Biomedical Imaging, vol. abs/1801.0, no. Isbi, pp. 1487–1490, 2018.

S. P. Deng, S. Cao, D.-S. Huang, and Y.-P. Wang, (2017) “Identifying Stages of Kidney Renal Cell Carcinoma by Combining Gene Expression and DNA Methylation Data,” IEEE/ACM Transactions on Computational Biology and Bioinformatics, vol. 14, no. 5, pp. 1147–1153.

Vijayarani, S., Dhayanand, S., & Phil, M. (2015). Kidney disease prediction using SVM and ANN algorithms. International Journal of Computing and Business Research (IJCBR), 6(2).

R. B.V, N. Sriraam, and M. Geetha, “Classification of non-chronic and chronic kidney disease using SVM neural networks,” International Journal of Engineering & Technology, vol. 7, no. 1.3, p. 191, Dec. 2017.

Sinha, P., & Sinha, P. (2015). Comparative study of chronic kidney disease prediction using KNN and SVM. International Journal of Engineering Research and Technology, 4(12), 608-12.

U. N. Dulhare and M. Ayesha, “Extraction of action rules for chronic kidney disease using Na¨ıve bayes classifier,” in 2016 IEEE International Conference on Computational Intelligence and Computing Research (ICCIC), 2016.

Boukenze, Basma, Abdelkrim Haqiq, and Hajar Mousannif. "Predicting Chronic Kidney Failure Disease Using Data Mining Techniques." Advances in Ubiquitous Networking 2. Springer, Singapore, 2017. 701-712.

Panwong, Patcharaporn, and Natthakan Iam-On. "Predicting transitional interval of kidney disease stages 3 to 5 using data mining method." 2016 second Asian conference on defence technology (ACDT). IEEE, 2016.

Dulhare, Uma N., and Mohammad Ayesha. "Extraction of action rules for chronic kidney disease using Naïve bayes classifier." 2016 IEEE International and Computing Research (ICCIC). IEEE, 2016.

Abeer, Ahmad, “Diagnosis and Classification of Chronic Renal failure Utilizing Intelligent Data Mining Classification”.

Dhruvi. R, Yavnika P, Nutan R, " Prediction of Probability of Chronic Diseases and Providing Relative Real-Time Statistical Report using data mining and machine learning

techniques". International Journal of Science, Engineering and Technology Research (IJSETR) vol. 5, no. 4. 2016.

Boukenze B, Mousannif H, Haqiq A (2016) Performance of data mining techniques to predict in healthcare case study: chronic kidney failure disease. IJDMS 8(3):1–9.

Padmanaban KR, Parthiban G (2016) Applying machine learning techniques for predicting the risk of chronic kidney disease. IJST 9(29):1–5.

"Your Kidneys & How They Work | NIDDK." [Online]. Available: https://www.niddk.nih.gov/health-information/kidneydisease/kidneys-how-they-work, June 2018.

"Kidney disease: The Basics," Aug. 2014. [Online]. Available: https://www.kidney.org/news/newsroom/factsheets/KidneyDiseaseBasics.(NVS 2021 report of 2018 data)

"Global Facts: About Kidney Disease," [Online]. Available: https://www.kidney.org/kidneydisease/global-facts-about-kidneydisease/. [Accessed: 20-Feb-2020].

Z. Chen et al., "Diagnosis of patients with chronic kidney disease by using two fuzzy classifiers," Chemometr. Intell. Lab., vol. 153, pp. 140-145, Apr. 2016.

A. Subasi, E. Alickovic, J. Kevric, "Diagnosis of chronic kidney disease by using random forest," in Proc. Int. Conf. Medical and Biological Engineering, Mar. 2017, pp. 589-594.

L. Zhang et al., "Prevalence of chronic kidney disease in china: a crosssectional survey," Lancet, vol. 379, pp. 815-822, Aug. 2012.

A. Singh et al., "Incorporating temporal EHR data in predictive models for risk stratification of renal function deterioration," J. Biomed. Inform., vol. 53, pp. 220-228, Feb. 2015.

A. M. Cueto-Manzano et al., "Prevalence of chronic kidney disease in an adult population," Arch. Med. Res., vol. 45, no. 6, pp. 507-513, Aug. 2014.

H. Polat, H.D. Mehr, A. Cetin, “Diagnosis of chronic kidney disease based on support vector machine by feature selection methods,” J. Med. Syst., vol. 41, no. 4, Apr. 2017.

C. Barbieri et al., “A new machine learning approach for predicting the response to anemia treatment in a large cohort of end stage renal disease patients undergoing dialysis,” Comput. Biol. Med., vol. 61, pp. 56-61, Jun. 2015.

M.-H. Tsai, et al., “Incidence, Prevalence, and Duration of Chronic Kidney Disease in Taiwan : Results from a Community-Based Screening Program of 106094 Individuals,” Nephron, vol. 140, pp. 175–184, 2018.

V. Kunwar, K. Chandel, A. S. Sabitha and A. Bansal. (2016). “Chronic Kidney Disease analysis using data mining classification techniques." In 6th International Conference - Cloud System and Big Data Engineering (Confluence),Noida,2016,pp.300-305. doi: 10.1109/CONFLUENCE.2016.7508132.

Zewei Chena et al. (2016). “Diagnosis of patients with chronic kidney disease by using two fuzzy classifiers.” Chemometrics and Intelligent Laboratory Systems, Volume 153, 15 April 2016, Pages 140-145.

Singh N, Singh P. Rule Based Approach for Prediction of Chronic Kidney Disease: A Comparative Study. Biomed Pharmacol J 2017;10(2).

Tahmid Abrar,Samiha Tasnim,Md. Mehrab Hossain , “EARLY DETECTION OF CHRONIC KIDNEY DISEASE USING MACHINE LEARNING”,Department of Computer Science and Engineering BRAC University (2019).

M. Hueso, A. Vellido, N. Montero et al., “Artificial intelligence for the artificial kidney: pointers to the future of a personalized hemodialysis therapy,” Kidney Diseases, vol. 4, no. 1, pp. 1–9, 2018.

H. M. Krumholz, “Big data and new knowledge in medicine: the thinking, training, and tools needed for a learning health system,” Health Affairs, vol. 33, no. 7, pp. 1163–1170, 2014.

Administration FaD, “Proposed regulatory framework for modifications to artificial intelligence/machine learning (AI/ML)-based software as a medical device (SaMD)-discussion paper and request for feedback,” 2019, https://www.fda .gov/media/122535/download.

T. A. Mavrakanas and D. M. Charytan, “Cardiovascular complications in chronic dialysis patients,” Current Opinion in Nephrology and Hypertension, vol. 25, no. 6, pp. 536–544, 2016.

F. Reyna-Sepúlveda, A. Ponce-Escobedo, A. Guevara-Charles et al., “Outcomes and surgical complications in kidney transplantation,” International Journal Organ Transplantation Medicine, vol. 8, no. 2, pp. 78–84, 2017.

M. Alloghani et al., “Applications of machine learning techniques for software engineering learning and early prediction of students’ performance,” in Proc. Int. Conf. Soft Computing in Data Science, Dec. 2018, pp. 246- 258.

D. Gupta, S. Khare, A. Aggarwal, “A method to predict diagnostic codes for chronic diseases using machine learning techniques,” in Proc. Int. Conf. Computing, Communication and Automation, Apr. 2016, pp. 281-287.

L. Du et al., “A machine learning based approach to identify protected health information in Chinese clinical text,” Int. J. Med. Inform., vol. 116, pp. 24-32, Aug. 2018.

R. Abbas et al., “Classification of foetal distress and hypoxia using machine learning approaches,” in Proc. Int. Conf. Intelligent Computing, Jul. 2018, pp. 767-776.

M. Mahyoub, M. Randles, T. Baker and P. Yang, “Comparison analysis of machine learning algorithms to rank alzheimer’s disease risk factors by importance,” in Proc. 11th Int. Conf. Developments in eSystems Engineering, Sep. 2018.

Mohammed, M., Khan, M. B., & Bashier, E. B. M. (2016). Machine Learning Algorithms and Applications. In CRC Press.

Smola, A., & Vishwanathan, S. (2008). Introduction to Machine Learning. Cambridge: The Press Syndicate of the University of Cambridge.

Naqa, I. El, & Murphy, M. J. (2015). Machine Learning in Radiation Oncology. 3–11. https://doi.org/10.1007/978-3-319-18305-3

Ojha, Varun Kumar; Abraham, Ajith; Snášel, Václav (1 April 2017). "Metaheuristic design of feedforward neural networks: A review of two decades of research". Engineering Applications of Artificial Intelligence. 60: 97–116.

Briot, J., & Pachet, F. (2017). Music Generation by Deep Learning - Challenges and Directions. Computer Science, Engineering. https://arxiv.org/abs/1712.04371.A. Survey, Importance of artificial neural networks for location of faults in transmission systems, Surveyor (2017) 357-362.

DACS, Artificial Neural Networks Technology, (2014).

A. Krogh, What are artificial neural networks? Nat. Biotechnol. 26 (2) (2008) 195–197 http://doi.org/10.1038/nbt1386.

M. Kamuda, J. Stinnett, C. Sullivan, Automated isotope identification algorithm using artificial neural networks, IEEE Trans. Nucl. Sci. 9499 (c) (2017) 1–1 http:// doi.org/10.1109/TNS.2017.2693152.

I.N. da Silva, D. Hernane Spatti, R. Andrade Flauzino, L.H.B. Liboni, S.F. dos Reis Alves, Artificial Neural Networks vols. 21–29, (2017) http://doi.org/10.1007/978- 3-319-43162-8.

R.G. Brereton, G.R. Lloyd, Support vector machines for classification and regression, Analyst 135 (2) (2010) 230–267 http://doi.org/10.1039/B918972F.

T.S. Furey, N. Cristianini, N. Duffy, D.W. Bednarski, M. Schummer, D. Haussler, Support vector machine classification and validation of cancer tissue samples using microarray expression data, 16 (10) (2000) 906–914.

A.A.A.A. Adewumi, T.O. Owolabi, I.O.I.O. Alade, S.O.S.O. Olatunji, Estimation of physical, mechanical and hydrological properties of permeable concrete using computational intelligence approach, Appl. Soft Comput. 42 (2016) 342–350 http://doi.org/10.1016/j.asoc.2016.02.009.

K.O. Akande, S.O. Olatunji, T.O. Owolabi, A. AbdulRaheem, Feature SelectionBased ANN for Improved Characterization of Carbonate Reservoir, CPAPER, Society of Petroleum Engineers, 2015, http://doi.org/10.2118/178029-MS.

C. Cortes, V. Vapnik, Support-vector networks, Mach. Learn. 297 (1995) 273–297.

W.S. Noble, What is a support vector machine? Nat. Biotechnol. 24 (12) (2006) 1565–1568 http://doi.org/10.1038/nbt1206-1565.

Idoko JB, Arslan M, Abiyev R. Fuzzy Neural System Application to Differential Diagnosis of Erythemato-Squamous Diseases. Cyprus J Med Sci 2018; 3: 90-7.

L. A. Zadeh, Fuzzy logic = computing with words. IEEE Trans. Fuzzy Syst. 4, 103 (1996).

Abiyev RH, Abizade S. Diagnosing Parkinson's diseases using fuzzy neural system. Comput Math Methods Med 2016; 1267919.

Mamdani EH, Assilian S. An experiment in linguistic synthesis with a fuzzy logic controller, Int J Man Mach Stud 1975; 7: 1–3.

Takagi T, Sugeno M. Fuzzy identification of systems and its applications to modeling and control. IEEE Syst, Man, Cybern 1985; 1: 116- 132.

Abiyev RH, Kaynak O. Fuzzy wavelet neural networks for identification and control of dynamic plants – a novel structure and a comparative study. IEEE T Ind Electron 2008; 55: 3133-3140.

Abiyev RH. Fuzzy wavelet neural network based on fuzzy clustereing and gradient techniques for time series prediction. Neural Comput Appl 2011; 20: 249-25.

Mamdani EH, Assilian S. An experiment in linguistic synthesis with a fuzzy logic controller, Int J Man Mach Stud 1975; 7: 1–3.

Takagi T, Sugeno M. Fuzzy identification of systems and its applications to modeling and control. IEEE Syst, Man, Cybern 1985; 1: 116- 132.

Abiyev RH, Kaynak O. Fuzzy wavelet neural networks for identification and control of dynamic plants – a novel structure and a comparative study. IEEE T Ind Electron 2008; 55: 3133-3140.

R. Abiyev, Fuzzy wavelet neural network based on fuzzy clustering and gradient techniques for time series prediction. Neural. Comput. Appl. 20, 249 (2011).

R. H. Abiyev, N. Akkaya, E. Aytac, I. Günsel, and A. Ça ˘gman, Brain-computer interface for control of wheelchair using fuzzy neural networks. Biomed Research International 2016 (2016).

Chronic_Kidney_Disease Data Set [Online]. Available: https://archive.ics.uci.edu/ml/datasets/chronic_kidney_disease.

Tahmid Abrar, Samiha Tasnim, Md. Mehrab Hossain. EARLY DETECTION OF CHRONIC KIDNEY DISEASE USING MACHINE LEARNING, Department of Computer Science and Engineering BRAC University September, 2019.

Abiyev, R., Abizada, S. Type-2 fuzzy wavelet neural network for estimation energy performance of residential buildings. *Soft Comput* (2021), https://rdcu.be/ckGS8. https://doi.org/10.1007/s00500-021-05873-4

Abiyev R.H, Idoko J.B., Dara R. Fuzzy Neural Networks for Detection Kidney Diseases. In: Kahraman C., Cevik Onar S., Oztaysi B., Sari I., Cebi S., Tolga A. (eds) Intelligent and Fuzzy Techniques: Smart and Innovative Solutions. INFUS 2021. Advances in Intelligent Systems and Computing, Springer,Cham (accepted).

APPENDICES

APPENDIX 1

(1): Class_nfuztsk File content:

```
function [c1,o1,w1,w2]=class_nfuztsk(c1,o1,w1,w2)

% Fuzzy Neural Networks (Muti-input multi-output networks) MIN T-norm

%   clustering Multioutput bsisic program, using linear function in consequent part => TSK type

% option1 - train=1,  test=2

% N1 - number of neurons of input layer

% N2 - number of neurons of rule (hidden) layer

% M - number of neurons of output layer

% a1 - learning rate

% a2 - Momentum

% maxc - output classes

% clustering   1-with clustering,   0- without clustering

menu1='K'

if(menu1=='K') load kidneydp3.txt; xt=kidneydp3(:,3:21); end    %kidney diseases

[row,col]=size(xt);  minc=min(xt(1:row,col));

for i=1:col-1

   max1=max(xt(:,i));    xt(:,i)=xt(:,i)/max1;

end

tbeg=0;  tend=row;
```

```
for t=tbeg+1:tend
   if(minc==1)  xt(t,col)=xt(t,col)-1;   end
      Data(t-tbeg,:)=[xt(t,1:col)];
end
maxc=max(xt(1:row,col))+1;
N2=16; N1=col-1;   M=maxc;
[N1   N2   M]
pause
clc
disp('0- Initialization of parameters ')
disp('1- Learning                    ')
disp('2- Testing                     ')
disp('3- Exit                        ')
num=input('Enter number : ');
switch num
   case 0, k1=1;k2=1; k3=1; k4=1;
         [c1,o1,w1,w2]=nfuzTSKgen(N1,N2,M,k1,k2,k3,k4);
         [c1,o1,w1,w2]=class_nfuztsk(c1,o1,w1,w2)
   case 1,  epoch =200;    option1=1; clustering=0; a1=0.01;      a2=0;
```

[c1,o1,w1,w2]=nfuzTSKtrain_crosv(menu1,option1,N1,N2,M,a1,a2,Data,epoch,maxc,clusteri
ng,c1,o1,w1,w2);

[c1,o1,w1,w2]=class_nfuztsk(c1,o1,w1,w2);

case 2, option1=2;epoch =1; clustering=0; a1=0.0; a2=0;

[c1,o1,w1,w2]=nfuzTSKtrain_crosv(menu1,option1,N1,N2,M,a1,a2,Data,epoch,maxc,clusteri
ng,c1,o1,w1,w2);

[c1,o1,w1,w2]=class_nfuztsk(c1,o1,w1,w2);

case 3, quit cancel;

end

end

(2): nfuzTSK File content:

function [ys,net,m,minm,summin]=nfuzTSK(N1, N2,M, c,o,w1,w2,x);

% Fuzzy Neural Networks (Muti-input multi-output networks) MIN T-norm operation

% The NEFUZM is used to create network realising Mamdany type fuzzy system.

% ys - output array ys(M)

% net - outputs of linear function net(N2)

% minm - outputs of rule layer minm(N2)

% m - membership functions beween input and rule layers m(N1,N2)

% summin - sum of membership functions

% N1 - number of neurons of input layer

% N2 - number of neurons of rule (hidden) layer

% M - number of neurons of output layer

% c - cerntres of membersip functions c(N1,N2)

% o - widths of membersip functions o(N1,N2)

% w1 - weight coefficients used in linear functionx w1(N1,N2)

% w2 - weight coefficients between hidden and output layers w2(N2,M)

% x - input signals x(N1)

%

%inference using min operation in premise

(3): nfuzTSKgen File content:

function [c,o,w1,w2]=...

nfuzTSKgen(N1,N2,M,k1,k2,k3,k4)

% Fuzzy Neural Networks (Muti-input multi-output networks) MIN T-norm operation

% The NEFUZM is used to create network realising Mamdany type fuzzy system.

% c1 - cerntres of membersip functions

% o1 - widths of membersip functions

% w1 - weight coefficients between hidden and output layers

% k1, k2, k3 are the coefficients for scaling the c1, o1, w1

% N1 - number of neurons of input layer

% N2 - number of neurons of rule (hidden) layer

% M - number of neurons of output layer

(4): nfuzTSKtrain File content:

function
[c,o,w1,w2,co,oo,w1o,w2o]=nfuzTSKtrain(N1,N2,M,a1,a2,x,er,minm,summin,m,ys,net,c,o,w1,w2,co,oo,w1o,w2o)

% Fuzzy Neural Networks (Muti-input multi-output networks) MIN T-norm operation

% The NEFUZM is used to create network realising Mamdany type fuzzy system.

% c - cerntres of membersip functions c(N1,N2)

% o - widths of membersip functions o(N1,N2)

% w1 - weight coefficients used in linear functionx w1(N1,N2)

% w2 - weight coefficients between hidden and output layers w2(N2,M)

% co,oo,w1o,w2o are previous vales of c,o,w1,w2

% N1 - number of neurons of input layer

% N2 - number of neurons of rule (hidden) layer

% M - number of neurons of output layer

% x - input array(signals) x(N1)

% er - error array

% minm - minimum values of mf in rule layer

% summin - sum of membership functions

% m - membership funcions

% ys - output array

% net - outputs of linear functions

(5): nfuzTSKtrain_crosv File content:

function [c1find,o1find,w1find,w2find]=nfuzTSKtrain_crosv(menu1,option1,N1,N2,M,a1,a2,Data,epoch,maxc,clustering,c1,o1,w1,w2)

% Fuzzy Neural Networks (Muti-input multi-output networks) MIN T-norm operation

% The NEFUZM is used to create network realising Mamdany type fuzzy system.

% option1 - train=1, test=2

% N1 - number of neurons of input layer

% N2 - number of neurons of rule (hidden) layer

% M - number of neurons of output layer

% a1 - learning rate

% a2 - Momentum

% maxc - output classes

% clustering   1-with clustering, 0- without clustering

APPENDIX 2: Ethical Approval letter

NEAR EAST UNIVERSITY

ETHICAL APPROVAL DOCUMENT

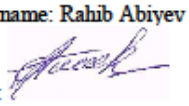
Date:01/06/2021

To the Graduate School of Applied Sciences

For the thesis project entitled as "Chronic kidney disease detection using fuzzy neural network", the researchers declare that they did not collect any data from human/animal or any other subjects. Therefore, this project does not need to go through the ethics committee evaluation.

Title: Prof. Dr

Name Surname: Rahib Abiyev

Signature: 

Role in the Research Project: Supervisor

APPENDIX 3. Similarity