COMPARISON OF INTELLIGENT LEARNING ALGORITHMS

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

By

SHAKAR SHERWAN HASAN HASAN

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

NICOSIA, 2018

COMPARISON OF INTELLIGENT LEARNING ALGORITHMS

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

By

SHAKAR SHERWAN HASAN HASAN

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

NICOSIA, 2018

SHAKAR SHERWAN HASAN: COMPARISON OF INTELLIGENT LEARNING ALGORITHMS

Approval of Director of Graduate School of

Applied Sciences

Prof. Dr. Nadire ÇAVUŞ

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

Software Engineering

Examining Committee in Charge:

Assist. Prof. Dr. Yöney Kırsal EVER Department of Software Engineering, NEU

Assoc. Prof. Dr. Kamil DİMİLİLER Department of Automotive Engineering, NEU

Assist. Prof. Dr. Boran ŞEKEROĞLU Department of Information Systems Engineering, NEU

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

Name, Last name;

Signature:

Date:

To my country and my family …

ACKNOWLEDGEMENT

First of all, I want thanks God for this fantastic chance which gave me to start a master study in software engineering.

I want send a very thank full to my mom who are beside me in everything and my dad, he is also great power to me and I promise I will try to remain your names massively, mom and dad I can't describe you by words that is all I want say.

I predicate all continues support of my special supervisor and advisor Assist. Prof. Dr. Boran ŞEKEROĞLU, he is with every time even in the holidays, he is converting my mistakes to lessons and he make me a wise boy. Working with you is an honor, over all knowing you is everything I will proud which I have you.

I like to acknowledgment Dr.Saman Mirza my co-supervisor which he is helps me so much for learning methods and I took a lot of his time and predicate that all I reached this level all belongs to your continues support. I can't describe you here by words, I just say thanks for everything.

ABSTRACT

Machine learning techniques become a very wide tools that used for classification and clustering. There are many types of algorithms that utilized for different applications in multi-disciplinary area.

This work addresses is comprising among three different types of machine learning algorithms namely Artificial Neural Network, Radial Basis Function, and Support Vector Machine. The work applied different dataset to check the performance of the above mentioned algorithms through some performance indicators.

The found that each algorithm gives variety performance rate, as the value of true and false alarms are changed from and algorithm to another. The work concluded that finding perfect machine learning algorithm is not an easy task. Moreover, the work found that generalized a machine learning on many application as a better one is not correct.

Keywords: Artificial Neural Network (ANN); Radial Basis Function (RBF); Support Vector Machine (SVM); Machine learning; Comparison techniques

ÖZET

Makine Öğrenme teknikleri sınıflandırma ve kümeleme için çok yaygın kullanılan araçlar haline gelmiştir. Birçok çeşitli algoritmalar, farklı disiplinlerde kullanılmak için uygulanmıştır.

Bu tezde, üç farklı makine öğrenme algoritması, GYA, RBF ve DVM kıyaslaması yapılmıştır. Farklı veriler kullanılarak ve bazı performans göstergeleri dikkate alınarak bu algoritmaların kıyaslanması yapılmıştır.

Her algoritmanın farklı sonuçlar verdiği, doğru ve yanlış oranlarının da algoritmalar arasında farklılıklar gösterdiği tespit edilmiştir. En iyi algoritmanın seçiminin zor olduğu sonucuna varılmış ve sonuçlarla gösterilmiştir.

Anahtar Kelimeler: Yapay Sinir Ağları; RBF; Destek Vektör Makinaları; Makine öğrenimi; Kıyaslama teknikleri

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF ABBREVIATIONS

ML: Machine Language

ANN: Artificial Neural Network

BPNN: Backpropagation Neural Network

SVM: Support Vector Machine

RBF: Radial Basis Function

LIST OF TABLES

CHAPTER ONE
INTRODUCTION

1.1 Overview

Machine learning (ML) is a part of Artificial Intelligence (AI). The Machine Learning's objective usually is to recognize the construction of statistics and fit the statistics into models which recognized and used by others. Tough that in computers science Machine Learning is exist and main part, in traditional computing, it find the difference from traditional approaches. A group of explicitly programmed instructions which used by machines to solve a problem or calculate the files is called algorithms. Machine learning enables computers in constructing models from sample data according to automate processes making decision based on data inputs and the reason of that is machine learning algorithms as a replacement for computers input data to train and data analysis used to output values which fall within an exact range.

Nowadays, each user tools have advanced by machine learning. We can convert pictures into text that we can work on it by Optical character recognition (OCR), share photos and tag friends are done by social media platform from facial recognition technology. Which is all belongs to machine learning any many other things like self-driving and television. Because of that machine learning in developing continuously and keep in every developers and programmers mind to work on it and analyze the input of the process machine learning.

In this thesis, common machine learning methods of supervised learning were discussed, and common algorithmic approaches of machine learning, including the Backpropagation Artificial Neural Network (BPNN), Radial Basis Function (RBF), and Support Vector Machine (SVM). The work explores these techniques through Matlab programming language, which are the most popular and used in machine learning, providing some datasets and databases that used in different fields. Additionally, this work will test these three ML methods to check their ability on classifying unseen objects

1.2 Machine Learning

Nowadays, machine learning is one of massive need exhaustive explicit to be describe for learning and know how it learn, the main difference between humans and machines from learning is that humans decide to solve problems automatically depend on the past mistakes or want different output from the past outputs if they have a background on the problem it will be a strengthen to them but machines should prepare for solve a problem, computers are not carry about the output what will be it do it's process that what the problem related to which filed that machines programmed before to solve those kind of problems. The main important point in machine learning is to identify the problem and then do the process on the problem. Machines are able to learn and improve their performances by gathering more data and experience (DING Wenrui, 2008)(Drouhard, Sabourin, Robert, Godbout, & Mario, 1996). Machine learning are normally classified by their underlying learning strategies which are often identified by the amount of inference the companies program which able to perform. There are two input data for the machine learning. Arranged data and unarranged data, arranged data is the data it provide attributes. If it used in supervised learning because of meaning attached to the data and some sort of tag. Arrange attributes can be categorical or numerical. Categorical data are used in classification while the numerical data are used in regression to predict the value. Unarranged data is the data where are only data points and no labelling to assist because of this reason unarranged data used in unsupervised learning so that machine can identify the patterns or any structure present in the data set. There is a very important part in machine learning is how to know that which of the outputs were appropriate and contained mistakes. An example to the algorithm will be a computer program which tries to predict whether a customer in one e-commerce shop will perform a purchase or not. The evaluation of classification tasks is normally done by splitting the data set into a training data set and a test data set. The machine learning algorithm trained the first one, while the test data set is used to calculate performance indicators in order to evaluate the quality of algorithm (Zezhi Chen).

1.3 Problem Statements

There are many tools and methods used for machine learning (ML), and these methods are used for multi-disciplinary fields and topics. Therefore, it is necessary for researchers to test the performance of those ML techniques with different datasets. The gap that this work wants to address is measuring some performance indicators for ML tools and methods through applying a comparison among different ML methods and techniques utilizing a variety of datasets.

The thesis addresses some explanations and descriptions on the behaviors of ML techniques while they applied in different fields and for variety of topics. The thesis focuses on explaining the ability of each ML techniques in classifying objects into correct classes through interpreting the rate of accuracy and performance indicators.

1.4 Aim of the Thesis

The main objective that addressed by this work is testing the performance of the popular and most widely used ML techniques in classification projects and applications. The work can achieve the main objective through the following steps:

1. Studying and reviewing the concept of ML in the viewpoint of some popular and reputable algorithms.
2. Identifying different datasets and databases that will be applied for to the ML techniques that chosen in step one.
3. Building three types of ML model.
4. Training and testing the ML models that built in step three.
5. Identifying the performance indicators for analyzing the selected ML techniques.

1.5 Thesis Layout

This thesis has been laid out into the following chapters:

1. Chapter One.

In this chapter, the layout and the orientation of the work have been identified. The chapter presented the statements of this thesis with objectives.

2. Chapter Two.

This thesis covered many ML techniques and reviewed many related thesis. Chapter two covers the methods and algorithms that have been used by selected ML techniques. Based on the best knowledge of the author, the chapter covers the most important related works. Moreover, the chapter covers some explanations on utilized datasets and performance indicators.

3. Chapter Three.

The methodology of this work has been described in chapter three.

4. Chapter Four.

Coding and implementation for the selected ML techniques have been explained in chapter three. The chapter also covers the results of each round of implementation.

5. Chapter Five.

In chapter five, conclusions with some recommendation for future works have been presented.

CHAPTER TWO
THEORETICAL BACKGROUND

2.1 Introduction

This chapter introduces some related works to this study. The chapter gives a detailed explanation on each reviewed works to understand the process of the ML techniques and the ways thoroughly the techniques are measured.

The chapter described the structure and the framework of models that utilizing the ML methods. Moreover, the datasets and the databases that fed to the input stage of ML models have been studied and explained.

2.2 The Work Stages

This thesis has arranged in the following four stages:-

1. Collecting data and information regarding the ML techniques and methods. In this direction the work should review some previous works that focused on using ML techniques in different field and for variety of applications (this is explained in detail in subsequent sections of Chapter2).
2. Identifying the type of datasets that applied to selected ML techniques for training and testing phases.
3. Identifying performance indicators, then validating and testing the selected techniques.

2.3 Literature Review

In recent year, researchers indifferent fields and for a variety of purposes had moved towards utilizing machine learning approach to cluster, classify, and predict cases and events. One of the important method of learning known as supervised machine learning. With such machine learning process the aim is to near a function mapping that input map vectors to the output or target value for training data is used which arranged already. Though supervised learning, between the input and correct given output there is an association type to be conducted. According to this mapping build a model by supervised learning. This model can expect value of the output for any inputs that given. Supervised learning methodology is dividing into two groups, first, classification problems and second, regression. In case of modeling for regression problems the output which is expected is a real number value such as “electricity demand in KW”. However, problems of classification are almost those problems that output is a cluster or a group like normal and abnormal. The major supervised machine learning algorithms are known as logistic Regression, Naïve Bayes, Linear Regression, Decision Tree, Linear Discriminant analysis, KNearest Neighbor algorithm, Support Vector Machines and Neural Networks. In this work considered problems of binary classification only because of this purpose Support Vector Machines performance work analyzed and evaluated. Artificial Neural Network, and Radial Basis Function with the seven performance indicators: Accuracy, Precision, True positive rate, false positive rate, Misclassification, Specificity and F measure. The algorithm of Support Vector Machines is a non-probabilistic machines learning algorithm that make a model learns by classifying points in the space features (Kalpna Kashyap, 2013). ANN and Radial Basis Function in this work are presented (DING Wenrui, 2008).

There are many areas and fields where machine learning techniques have been utilized. Most applications performs computational learning, natural language processing and pattern recognition. Using methods of machine learning, force of computers are in the acting capability with non-explicit programmed. Systems be able to construct with the algorithms which be able to learn and teach from the data then can make driven-data predictions and decisions. While the last decades, bring machine learning into our life with a huge effect daily. The effect includes computer vision, self-driving systems, optical character recognition and efficient web search. More ever, by methods of machine learning

acceptance, the artificial intelligence (AI) human-level had been enhance better, when you look (Drouhard, Sabourin, Robert, Godbout, & Mario, 1996)(Abuhaiba & SI, 2007) they discuss more, nonetheless, the tools of human information processing it comes traditional methods for the machine learning performance which is from satisfactory is very var. Encouraged human speech view by the deep structure hierarchical and systems of production. At the ends of the 20th century introduced the deep learning algorithm idea. Deep belief network (DBN) is an architecture novel deep structure learning which planned by Hinton in 2006 which is called advances deep learning (Sing, et al., 2011). Deep learning techniques rapid development with impact significant an information processing and signal have been developed. Researches on methods of neuromorphic which also has support improvement deep network models (Kalpna Kashyap, 2013) Approaches of artificial learning and traditional machine learning constants, tools of the deep learning have been enormously progressing for natural language processing (NLP), speech processing, image analysis, information retrieval and computer vision (teddy surya guhawan, 2017)(Yogesh, VG, Patil, & Abhijit, 2014). Deep learning idea created by the study from artificial neural network (ANN). During the few past decades ANNs have become a dynamic research area (Ghiasi, 2015)(Wildes & P, 1997)(Moghaddam, 2000)(Morris & Anne, 2011). To build a typical Neural Network (NN). Basic thing is arrange neurons to produce activations real-value. Though updating the weights, the Neural Network like approximate. However, based on the problems, the training procedure an Neural Network might need long computational stages. Back propagation Neural Network is a gradient efficient descent algorithm that uses important rate of the Neural Network since 1980. If Artificial Neural Network train with essential approach which is supervised learning and accuracy of training should be high, backpropagation acting may be not acceptable in the time applying to the fast data. The scope for the training data should be big enough otherwise the Neural Network will be in the problem of over fitting. Thus, others machine learning algorithms effective like K-Nearest neighbor (KNN), boosting and support vector machines take onto the global obtain optimum with the consumption of lower power. In 2006, Hinton (Arya, S, Inamdar, & S, 2010). A new method proposed for training called layer-wise-greedy-learning that selected as a birth for deep learning techniques. There are a simple idea for the layer-wise-greedy-learning have to be perform

for network pertaining before layer-by-layer subsequent training. , for the features extraction buy inputs, the data is reduced and prepare for the next layer each layer exporting and importing data to the other layers.

For more information, the research for methods on deep learning has moved a massive attention deal, as such a deal, there are a chain of present results which reported in the literatures. In the past years especially since 2009 there are a lot researches on the ImageNet in the both side industry and academia. Hinton who was the led of the research group that won the ImageNet image classification competition. Deep learning approaches were used (Wu Qiao, 2017). Hinton's group entered the competition and their results were %10 for the first time. In the second place their results was better, Badiu and Google they accept deep learning architecture and update their search image engines on the Hinton's work. Great improvement in searching accuracy is the result of their work. Stanford University, his position is chief scientist. On the march 16, go game match was done in the South Korea by project of deep learning of Google which is called DeepMind. The match was between strongest world's players LEE se-dol and their AL player AlphaGO [14]. The game showed which AlphaGO utilized techniques of deep learning. Presented strength surprising and Lee se-dolwas beat by 4:1. Furthermore, systems of deep learning have also performance exposed great in predicting the activity of likely effects of mutations and drug molecules in non-coding DNA on gene expression

2.4 Artificial Neural Network (ANN)

Artificial neural network is a computational model or like a mathematical model which is stimulate by neural network biological. That sample network contain any neurons interrelated which can compute and process data although the approach connectionist. ANN neurons are divided into three kind of layers: input layer, hidden layer, and output layer. As shown in (Figure 2.1) every neuron accepts scalars vector and vector of weights which both of them multiplied. The output is to adding value of bias to the produce of vectors will be applied to get start of the function. That is transfer function called. Figure2.1 is neuron type drawing (Wudai Liao, 2017).

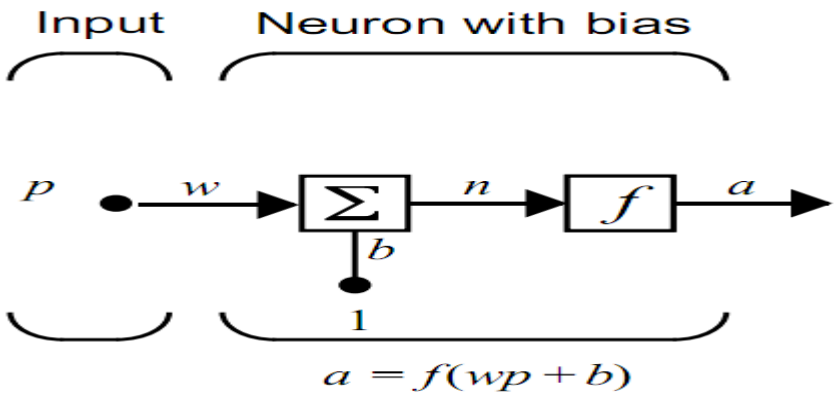

Figure 2.1: Input and output of an ANN neuron**.**

As you see in the Figure, numerous natures of the function transfer in ANN prototypes which be able to be in employment. Every function could adequate a special application. Logistic function uses the exacting project (Figure2.2-b) for the neurons of all layers for the classification purposes, and for the FBS rate diagnosis logistic and linear functions are used. This function is carefully chosen as it is a different function above range of the infinite. Furthermore, the outcome of this exacting function laid in the range of [0, 1], which comes suitable with the performance of binary classifier models.

Hyperbolic tangent transfer function is shown in the (Figure2.2-a), It will have outcome in the duration of -1 to +1. The function belongs to the sigmoid bipolar in the artificial neural networks. Another function is called Pure-line function (linear function), which has Linear transfer function properties. The function as shown in the (Figure2.2-c) also called as linear simple function which gives the same outcome as its input and like it is inputs(Lesne, 2006)

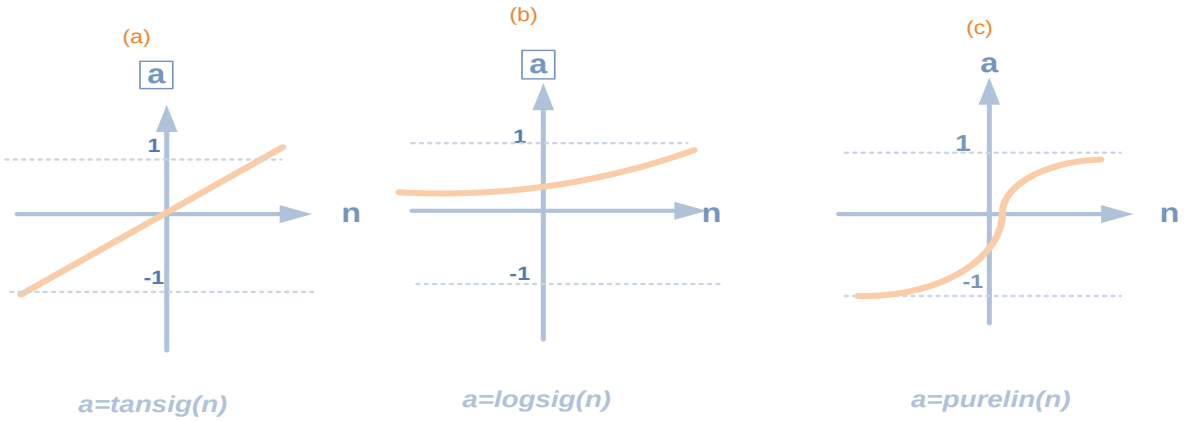

Figure 2.2: Three main types of transfer function in ANN

There are a variety ways to connect neurons inside a layer and between consequent layers. However, the simplest method is to join neurons from the layer straightly a current layer above to those below in the layer. Such connections can be seen in the case of Feed Forward Back propagation Neural Network (FF-BPNN). With such structure, a neuron can accepts its inputs from the neurons below, and then it sends its output to neurons above in a forward manner. Such structure is called feed forward as no reverse connections are between neurons from different layers. **Figure 2.**3, illustrates a classic structure of an FF-BBNN. The figure shows the details of the output of any neurons in such a type of neural network (Vinita Bhandiwad, 2017).

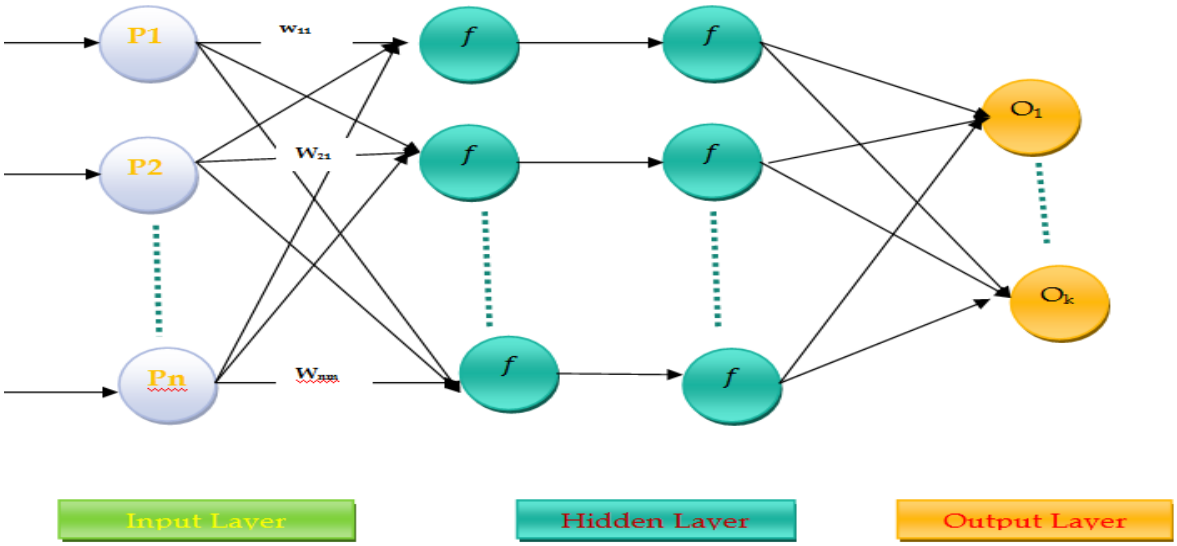

Figure 2.3: An *n* hidden layer structure of FFBP-NN with i[th] neurons at output

$$O_k(t+1) = F_k(2.1) \qquad (2.1)$$

$O_k(t+1)$ represents the k^{th} output of the neuron, $O_j(t)$expressing the output of the j^{th} neurons progressed for the k^{th} neuron, and $b(t)$ bias value for the j^{th} node, and ω_{jk} is the value of the weight which finds the effect neuron *j* on neuron *k*. Each neuron accepts like an input for the outside resources or adjacent neurons. Then, it calculates the output by using the function of active. After that the neuron straight on the result to the next adjacent neurons. The actions of ANN neurons consist of training through forwarding an input set and computing the desired output.

For that reason, the period "back propagation" defines the method which an ANN obtains training. The process of the training regulates the weights for the all ANN therefore it can achieve a specific request.

To achieve ANN configuration weight we need use various methods. Prior knowledge is the one way to initialize and set the value of weight and there are other ways like train the ANN by teaching and feeding it to update the weights by the learned rules. There are two methods for training data and learning data they are classified into two groups, unsupervised and supervised learning methods. Both learning methods result in the

arrangement of values for the weight between neuron (Wudai Liao, 2017)(Wu Qiao, 2017).

Backpropagation id one of the most important learning methods, that training function is the basic of backpropagation (Chinu Sayal). Backpropagation including supervised learning, and backpropagation depends on the delta rule, the difference and updating the weights the actual outputy_k and the desired d_k output, and is to delta rule do, as presented in equation 2.2. The delta rule is to find errors between the desired and actual output for every neuron, and γ shows learning rate and$y_j$is the output of the neurons. However, errors in the hidden layer are founded by back propagating the errors of the output layer neurons.

$$\Delta\omega_{jk} = \gamma y_j(d_k - y_k) \qquad (2.2)$$

Because of this reason, there are two phases for the backpropagation. The input data are fed for the input layer, for ANN in the first phase. After that, they will toward propagate the output. Later, signal errors are generate between the real outcome and the wanted outcome, in the other phase reverse action through the network are included. Although this action, the signal error sends to the network and to the every neuron, finally, considered the changes of the appropriate weight(Ghiasi, 2015).

The value for the error between the real and the wanted outcomes minimized. This minimization will be obtained through adjusting the value of the weight for the neuron. The whole story of training the ANN is to select a present input into the desired output. Such mapping capability makes ANN an efficient approach for many classification applications and pattern recognition.

2.5 Radial Basis Function

Multi-layer perceptron (MLP) is a model of network which is popular architecture network that used in the many applications area [1]. In the multi-layer perceptron, sum of the inputs is the weight and passed bias term to the level activation inside function transfer to introduce the output (i.e., $p(\vec{x}) = w_0 + \sum_{i=1}^{D} w_i x_i$ where $\vec{x}$ is an input vector of dimensionD, wi ,i = 1, 2, . . . , D are weights, w0 is the bias weight, and $f_{so} = \frac{1}{1+e^{-ao}}$, a is the slope) in the layered feed-forward topology arranged the units which is called forward neural network by the backpropagation learning algorithm. An multi-layer perceptron with backpropagation learning which is also called backpropagation neural network (BPNN). Training speed is the common barrier in the BPNN (ex. Number of layers are increasing and number of neurons are grow in the training speed) (Jain, et al., 2000).

To bypass this issue another worldview of simpler neural arrange structures with just a single hidden layer has been entered to numerous application territories with name of Radial Basis Function Neural Networks (RBFNs)[4, 6– 14]. RBFNs were first presented by Powell [15– 18]to take care of the addition issue in a multi-dimensional space requiring the same number of focuses as information focuses. Later Broomhead and Lowe (Drouhard, Sabourin, Robert, Godbout, & Mario, 1996)expelled the 'strict' restriction and utilized less focuses than information tests, so allowing many practical RBFNs applications in which the quantity of samples is high. A critical element of RBFNs is the existence of a quick, straight learning calculation in a network capable of speaking to complex non-direct mapping. At the same time it is additionally imperative to enhance the generalization properties of RBFNs (Drouhard, Sabourin, Robert, Godbout, & Mario, 1996)(Sing, et al., 2011).

Today RBFNs have been a focal point of concentrate not just in numerical investigation but also in machine learning specialists. Being acquired from the idea of organic open (Larsen, et al., 2008) and followed, Park and Sandberg demonstrate, "RBFNs were competent to build any non-direct mappings amongst jolt and response"(Verma, Chandra, Saha, Dipankar, & Saikia, 2013). The development of RBFNs investigate have been

consistently increased and across the board in various application territories. The examination in RBFNs are assembled into three classes (Samuel, Daramola, Samuel, & Ibiyemi, 2010): I) development of learning instrument (incorporate heuristics and no heuristics),ii) outline of new parts, and iii) zones of application. Recently, Patrikar [24] has contemplated which RBFNs can be identified by multi-layer perceptron with quadratic inputs (MLPQ). Let the MLPQ with one concealed layer comprising of M units and the yield unit with straight activation function. Let w0 (o) be the predisposition weight for the yield unit and let wi(o) be the weights related with the connection between concealed unit i and yield unit o. the output of the MLPQ can be processed as:

$$Q(\vec{x}) = w_0(o) + \sum_{i=1}^{M} \frac{w_i(0)}{1+exp} \qquad (2.3)$$

Where for every unit's hidden*i*,
w_{io} replacement for weight bias
w_{Lij} replacement for linear terms weightsx_j, and
w_{Qij} quadric terms associated weights replacementx_{2j} .

With the same hidden layer, an RBFN has an output, which given by:

$$R(\vec{x}) = \sum_{i=1}^{M} w_i exp \qquad (2.4)$$

Where w_i are the weights for unit *i*,
μ_i is the center path of unit *i*, and
$\beta$is the parameter. Using approximation, Equation above can be identified as

$$R(\vec{x}) = \sum_{i=1}^{M} w_i \qquad (2.5)$$

Where *c*,d, and *H* are constants.

2.6 Support Vector Machine

Support Vector Machine (SVM), is in the top of machine learning processes that was appeared in 1990's. Pattern recognition is the main area that SVM used for, but it also uses in many classification and pattern classification like a speech recognition, faulty card detection, text categorization, face detection and image recognition, etc.… the goals of pattern recognition is for classifying the data which is based on priori either information or data statistical which is extracted from the row data, that is a great approach in the separation of the data in many works. The type of the SVM is the supervised machine learning system and marked each one for their groups which come from many groups. Training Support Vector Machine process makes system which forecasts the group of the produce instances. Support Vector Machine in generalize the problems has the better ability especially when the objective is statistical learning.

The statistical learning concept provides a framework for learning the difficult of acquisition information, building forecasts, building choices from a group of statistic. In learning statistical concept the problem of supervised learning is expressed like follows. set of training data were given$\{(x_1, y_1)...(x_n, y_n)\}$ in R_n *x R*etested according to unidentified probability distribution P(x, y), and a function loss $V(y, f(x))$ which checks the error, for a given x, $f(x)$ is "foreseen" in its place of the real value y.

The issue comprises in finding a function f that limits the desire of the mistake on new information i.e., finding a capacity f that limits the normal blunder: $\int V(y, f(x)) P(x, y) dx dy$(PeiYun Zhang, 2018) . Early machine learning calculations intended to learn portrayals of basic capacities. Henceforth, the objective of learning was to yield a speculation that played out the right order of the preparation information and early learning calculations were intended to discover such a precise fit to the data (Ramachandra, et al., 2009). The capacity of a theory to effectively order information not in the preparation set is known as its speculation. SVM performs better in term of not over speculation when the neural systems may wind up finished summing up effortlessly (Rigoll, Gerhard, & Kosmala, 1998) .

The Figure 2.4 is the typical exemplary for representing SVM method. The model contains two patterns dissimilar and the SVM's goal is to distinct these two patterns. This model contains three type of diverse lines. The line w.x-b=0 is called like marginal line or margin of separation.

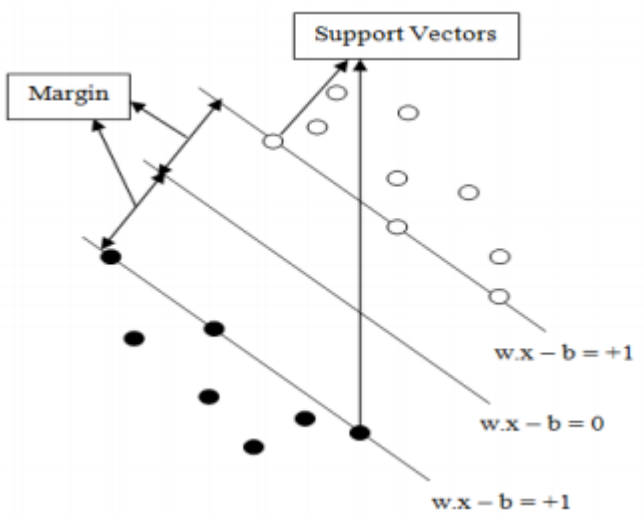

Figure 2.4: Simple model of SVM

The lines w.x - b = 1 and w.x – b = - 1 were those lines on other lines on the lines in the edge. The three above lines all together build hyper plane which isolates the given examples and the example that lies in the edges of the hyper plane is known vector of bolster. The opposite departure between the line of edge and the edges of hyper plane is called as edge. One of the destinations of SVM for precise grouping is to boost this edge for better order. The bigger the estimation of edge or the opposite separation, the better is the arrangement procedure and subsequently limiting the event of error.

The SVM typically manages design arrangement that implies this calculation is utilized for the most part to classify the diverse kinds of examples. Presently, there is distinctive sort of examples i.e. Straight and non-direct. Straight examples are designs that are effortlessly discernable or can be effectively isolated in low measurement though non-direct examples are designs that are not effortlessly recognizable or can't be effectively isolated and

consequently these sort of examples should be additionally controlled with the goal that they can be effectively isolated. Essentially, the fundamental thought behind SVM is the development of an ideal hyper plane, which can be utilized for arrangement, for directly divisible examples. The ideal hyper plane is a hyper plane chose from the arrangement of hyper planes for ordering designs that augments the edge of the hyper plane i.e. the separation from the hyper plane to the closest purpose of every example. The principle goal of SVM is to expand the edge with the goal that it can effectively group the given examples i.e. bigger the edge measure all the more accurately it characterizes the examples. The condition appeared underneath is the hyper plane portrayal.

$$\textit{Hyper plane, } aX + bY = C \qquad (2.6)$$

The Figure 2.5 illustrated the basic concept of the hyper plane telling how it is like when two patterns dissimilar are divided by hyper plane, in a three dimension. Principally, this plane contains of three lines which splits two dissimilar in 3-D space.

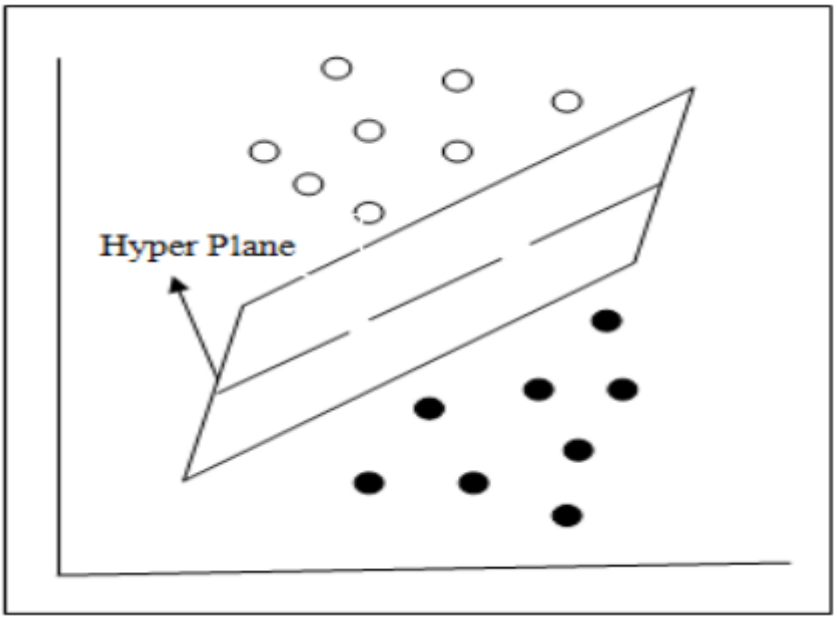

Figure 2.5: Hyper plane in of SVM

For non-straight divisible examples, the given example by drawing it into new space ordinarily a higher measurement space thus that in higher measurement space, the example turns out to be directly distinguishable. The given example can be recorded into higher measurement space utilizing part work, Φ(x). i.e. x Φ(x) Selecting diverse piece work is a vital perspective in the SVM-based characterization, generally utilized part works incorporate LINEAR, POLY, RBF, and SIGMOID.

2.7 Utilized Datasets

For checking the performance of some ML techniques, it is necessary to train and test them with a variety of datasets. This work utilized the following datasets.

2.7.1. Wine Dataset

Wine data set has 178 instances and 13 attributes, Multivariate is a data set characterizes, attribute characterizes are integer and real, and this dataset are used in classifications methods and this dataset donated on 01-07-1991. And it was created by Forina, and the donor is Stefan Aeberhard.

This work uses same 3D dimensions to find the output on this dataset and we apply this dataset to RBF, SVM and Backpropagation

2.7.2. Sonar Dataset

Sonar dataset has 208 instances and 60 attributes, Multivariate is a data set characterizes, attribute characterizes are real, and this dataset are used in classifications methods. And it was created by Terry Sejnowski, and we use binary classification for this dataset unlike above because it is contain only two classes that is the reason and the results is better than the others because it is binary.

2.7.3. Spect Hart Dataset

This dataset has 267 instances and 22 attributes, Multivariate is a data set characterizes, Categorical is the attribute characterizes. And this dataset is used in classification methods, and this dataset donated 01-10-2001. And this dataset created by Krzysztof J. Cios, Lukasz A. Kurga, and this dataset is separate training and test to classify this dataset.

2.7.4 Iris dataset

The Iris Dataset. This data sets consists of 3 different types of irises' (Setosa, Versicolour, and Virginica) petal and sepal length, stored in a 150x4 numpy.ndarray. The rows being the samples and the columns being: Sepal Length, Sepal Width, Petal Length and Petal Width. The below plot uses the first two features

2.8 Summary

This chapter starts with an introduction about the ML techniques and the methodology stages. The chapter introduced the theories and algorithms that utilized in this work, such as Back-Propagation Neural Network, Radial Bias Function.

The next chapter shows how these theories and tools will be combined together to build and simulate the proposed model.

CHAPTER THREE
MACHINE LEARNING MODELS

3.1 Introduction

This chapter presents the process of methodology that followed by this work to build an AI model that can classify objects into groups and classes. The chapter presents in step by step form, how such kind of work can start from a scratched point and progress its stages. To visualize sections of the methodology of this work, at each step required figures are illustrated to make the progress of this work more understandable.

3.2 Framework of Machine Learning Models

There are a basic rules for design a model. First of all, you have to know what is the model and what is the goal of this model and how it works and in how many works it used and how to implement it and how to apply and use it. Second step, you have to gathering the data to inter data to your models in some projects they use datasets and databases which somebody else already gathering the data and ready for use which it is more useful for classification methods, but for verification models you have collecting the data because verification models they are locally methods and offline. After gathering the data you have to arrange the data according to the system and arrange how the system need this data and how it is useful. After that you have to enter the data to the model and see the output. After first result you have to selecting and transforming variables to find the fit output for the model after some transforming variables you will see deferent outputs the important point here is you run the system many times and finally find the best results and choose that result to the model. When you find the best result you have to processing and evaluate the model which is the most important part of other parts, this part introduce several methods for processing and evaluating the model, in this part we have to capture information and shows the data and result and shows the ability of the system which this part must most explicit part of other parts. Before, final step you have to validating the model and the model should perform well on the development data, and the model if you enter data

randomly should run and gives the best results and if run many times should give you similar results and the model perform all datasets are similar or near to the model.

This part are discussed about three great methods for creating models. Perform your data in the real campaign is the best way to show your massive of your system by scoring alternate data. If you estimate your system you have to make trust intervolve by resampling techniques which is uses by bootstrapping. The important factors market are calculates by the key variable analysis like they are effected to the model. Finally, should applying and keeping data, this sections starts with labeling a new data set with a new model many checking procedures are demonstrates, best practices are tracking.

The organization of this chapter is coming in three main parts. The first part is about dataset that utilized as input data. The second part is regarding the tools and techniques that used to build the proposed model. The last part is about performance evaluation. Figure 3.1 illustrates, in general, the process of building the proposed model.

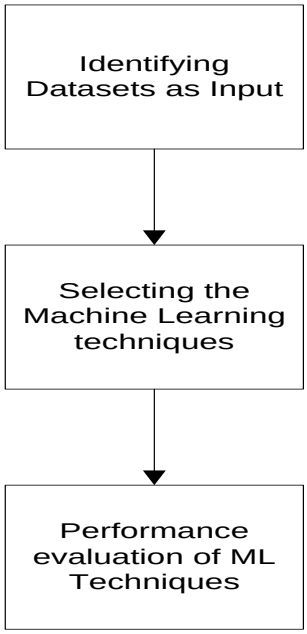

Figure 3.1: Frameworks of ML models

3.3 Comparison Methods

Many intelligent agents are performed and each one work in one section, and sometimes more than one intelligent agents can work in a same section so we practice all intelligent agents and test and find which one is the best for this section according to the results we are find the best one and this is called comparison.

For the classification sub-model, this work depends on checking the rate of false positive, false negative, true positive and true negative. Then using the equation 3.6, the accuracy of any classifier models can be computed (Salman Taseen Haque, 2017).

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \qquad (3.1)$$

Where:

TP, is the number of Type-I DM patients that classified correctly,

TN; is the number of Type-II DM patients that classified correctly,

FP; is the number of Type-I DM patients that classified as Type- II, and

FN; is the number of Type-II DM patients that classified as Type- I.

RECALL: Recall which is also known as sensitivity, recall return how many correct answers are return and labeled, recall is the number of true results divided by the number have to returned.

$$Recall = TP / (TP + FN) \qquad (3.2)$$

3.4 Summary

This chapter presented the methodology of the work, which comes in three main stages. The first stage of the work started with identifying the datasets that utilized by this work. The second stage of the work methodology is about designing the sub-part of the ML models; BP ANN, RBF, and SVM. The final main part is about evaluating the works and the performance, which depend on Confusion Matrix, Accuracy, and Recalls.

CHAPTER FOUR
MODEL IMPLEMENTATION AND MODELS

4.1 Introduction

This chapter presented the implementation of the work, which comes with three different types of techniques. The chapter explains the codes of each technique as well as the results of each code. The final part is about visual the performance evaluation of the methods.

4.2 Back Propagation Neural Network (BPNN)

In this section we will show all the results on the BPNN that we have done on the four datasets:

$$\text{Accuracy: } (TP + TN)/(TP+TN+FP+FN) \tag{4.1}$$

$$\text{Error Rate} = 1 - \text{Accuracy} = (FP+FN)/(TP+TN+FP+FN) \tag{4.2}$$

$$\text{Precision} = TP / (TP + FP) \tag{4.3}$$

$$\text{Recall} = TP / (TP + FN) \tag{4.4}$$

$$\text{Specificity} = TN / (TN + FP) \tag{4.5}$$

$$\text{F measure} = ((1 + \beta)^2 \times \text{Recall} \times \text{Precision}) / (\beta^2 \times \text{Recall} + \text{Precision}) \tag{4.6}$$

4.2.1 Training Phase

A. Training phase for iris dataset:

Iris dataset has 147 rows and 14 columns which we used 73 rows for training phase and we input all rows to test it we use 13 columns as predictor and last column as response.

B. Training phase for wine dataset:

Wine dataset has 14 columns and 178 rows we used first column as response and others predictors and we used the 89 rows to train the dataset and then use all rows to test the dataset.

C. Training phase for spect heart dataset:

spect heart dataset has 22 columns and 278 rows we use the first 21 column as a predictor and last one as response for the train we use 139 rows to train the dataset and then use all rows to test the dataset.

D. Training phase for sonar dataset:

Sonar dataset is a big dataset which consists from 208 columns and 60 rows and use first 59 columns as a predictor and last one as a response and we use the 104 rows to train the data and in the next step use all dataset to test the dataset. And sonar dataset consist of two groups mines and rocks.

4.2.2 Testing phase

A. Testing phase for iris dataset:

Table 4.1: Testing phase for Backpropagation for iris dataset.

	Iris Setosa	Iris Versicolour	Iris Virginica
Accuracy	0.97	0.97	0.97
Error rate	0.03	0.03	0.03
Precision	0.96	0.97	1.00
Recall	0.97	0.96	1.00
Specificity	0.97	1.00	0.76

In the previous table there is results about the testing the backpropagation neural network in the testing phase we use all attributes for testing in iris dataset as showed above accuracy which is more important than others is %97 which is good result which means that for iris dataset Back Propagation is better.

B. Testing phase wine dataset:

Table 4.2: Testing phase for Backpropagation for wine dataset.

	1	**2**	**3**
Accuracy	0.98	0.98	0.98
Error rate	0.02	0.02	0.02
Precision	0.96	1.00	1.00
Recall	1.00	0.96	1.00
Specificity	0.98	1.00	1.00

For wine dataset testing phase we have results for Back Propagation which is accuracy %98 it shows us we can depend on Back propagation Neural Network for wine dataset because the error rate %0.02.

C. Testing phase sonar dataset:

Table 4.3: Testing phase for Backpropagation for sonar dataset.

Accuracy	0.90
Error rate	0.10
Precision	0.90
Recall	0.88
Specificity	0.91

Like the others above this table shows us the all results for the sonar dataset, sonar dataset is two groups so it has just one accuracy and for others same each one has one, for the sonar dataset the accuracy is %90 so it is good but we have to test other models is the accuracy is higher and error rate is lower we choose them.

D. Testing phase Spect Heart dataset:

If you look at the table below, there are listed for all values of spect heart dataset, spect heart dataset has two groups also as mentioned before the accuracy for spect heart dataset is %81 which is good but we have to comparison this result with other models which one has lower error rate and higher accuracy we choose it.

Table 4.4: Testing phase for Backpropagation for spect heart dataset.

Accuracy	0.81
Error rate	0.19
Precision	0.63
Recall	0.25
specificity	0.83

4.3 Radial Basis Function (RBF)

4.3.1 Training Phase

A. Training phase for iris dataset:

Iris dataset has 147 rows and 14 columns which we used 73 rows for training phase and we input all rows to test it we use 13 columns as predictor and last column as response.

B. Training phase for wine dataset:

Wine dataset has 14 columns and 178 rows we used first column as response and others predictors and we used the 89 rows to train the dataset and then use all rows to test the dataset.

C. Training phase for spect heart dataset:

spect heart dataset has 22 columns and 278 rows we use the first 21 column as a predictor and last one as response for the train we use 139 rows to train the dataset and then use all rows to test the dataset.

D. Training phase for sonar dataset:

Sonar dataset is a big dataset which consists from 208 columns and 60 rows and use first 59 columns as a predictor and last one as a response and we use the 104 rows to train the data and in the next step use all dataset to test the dataset. And sonar dataset consist of two groups mines and rocks.

4.3.2 Testing Phase

A. Testing for the iris dataset:

Table 4.5: Testing phase for Radial Basis function for iris dataset.

	Iris Setosa	Iris Versicolour	Iris Virginica
Accuracy	1.00	1.00	1.00
Error rate	0.00	0.00	0.00
Precision	1.00	1.00	1.00
Recall	1.00	1.00	1.00
Specificity	1.00	1.00	1.00

In the above table there are results for the iris dataset which shows us the accuracy is %100 so it means the error rate is %0 so for any work or any comparison for iris dataset the RBF is the best model and we can use it and depending on it very well.

B. Testing for the wine dataset:

Table 4.6: Testing phase for Radial Basis function for wine dataset.

	1	**2**	**3**
Accuracy	1.00	1.00	1.00
Error rate	0.00	0.00	0.00
Precision	1.00	1.00	1.00
Recall	1.00	1.00	1.00
Specificity	1.00	1.00	1.00

In the previous page results of the wine dataset is shown which for all the classes are show us the accuracy is %100 and the error rate is %0. According to these results we can use RBF model for all works that belongs to wine dataset.

C. Testing for the sonar dataset:

Table 4.7: Testing phase for Radial Basis function for sonar dataset**.**

Accuracy	1.00
Error rate	0.00
Precision	1.00
Recall	1.00
Specificity	1.00

In the table 1.7 all results for the sonar dataset is shown, and according to the results same above datasets RBF is the best model for the sonar dataset, sonar dataset has two classes and so it has one accuracy and it is %100 which is shows the best result for our model.

D. Testing for spect heart dataset:

Table 4.8: Testing phase for Radial Basis function for spect heart dataset.

Accuracy	1.00
Error rate	0.00
Precision	1.00
Recall	1.00
Specificity	1.00

In the table 1.8 the results for the spect heart dataset is shown, the spect heart dataset has two classes and each class has many attributes, and according to this results also for the spect heart dataset the RBF intelligent agent is the best agent for the spect heart dataset.

4.4 Support Vector Machine (SVM)

4.4.1 Training Phase

A. Training phase for iris dataset:

Iris dataset has 147 rows and 14 columns which we used 73 rows for training phase and we input all rows to test it we use 13 columns as predictor and last column as response.

B. Training phase for wine dataset:

Wine dataset has 14 columns and 178 rows we used first column as response and others predictors and we used the 89 rows to train the dataset and then use all rows to test the dataset.

C. Training phase for spect heart dataset:

spect heart dataset has 22 columns and 278 rows we use the first 21 column as a predictor and last one as response for the train we use 139 rows to train the dataset and then use all rows to test the dataset.

D. Training phase for sonar dataset:

Sonar dataset is a big dataset which consists from 208 columns and 60 rows and use first 59 columns as a predictor and last one as a response and we use the 104 rows to train the data and in the next step use all dataset to test the dataset. And sonar dataset consist of two groups mines and rocks.

4.4.2 Testing Phase

A. Testing phase for iris dataset:

Table 4.9: Results of the testing phase for iris dataset.

Accuracy	0.967
Num. of true positive	145/150

B. Testing phase for wine dataset:

Table 4.10: Results for the testing phase for wine dataset.

Accuracy	0.978
Num. of true positive	174/178

C. Testing phase for spect heart dataset:

Table 4.11: Results for the testing phase for spect heart dataset.

Accuracy	0.816
Num. of true positive	218/267

D. Testing phase for sonar dataset:

Table 4.12: Results for the testing phase for sonar dataset.

Accuracy	0.988
Num. of true positive	197/208

For the comprising we present all results in one table:

Table 4.13: Results for all intelligent agents:

Data set	Accuracy			Precision			Recall			Specificity		
	BPNN	RBF	SVM	BPNN	RBF	SVM	BPNN	RBF	SVM	BPNN	RBF	SVM
Iris	0.97	1.00	0.96	0.97	1.00	0.97	0.97	1.00	0.97	0.97	1.00	0.97
Wine	0.98	1.00	0.97	0.96	1.00	0.92	1.00	1.00	0.96	1.00	1.00	0.94
sonar	0.90	1.00	0.98	0.90	1.00	0.96	0.88	1.00	0.92	0.91	1.00	0.91
Spect	0.81	1.00	0.81	0.63	1.00	0.84	0.25	1.00	0.83	0.83	1.00	0.84

4.5 Summary

This chapter illustrated the implementation and the results that have been obtained during evaluating three different types of ML techniques. The chapter gave enough explanations about the results so that readers can understand the important of designing and building such an enhanced model. More conclusion and discussion about the results and the whole works will be given in the next chapter.

CHAPTER FIVE
CONCULSIONS

The main direction of this work was to make a compassion between some popular ML techniques. To achieve that, the first step was doing some literature review on a variety of ML application. Then, different types of datasets and three main type of ML techniques have been selected to achieve this comparison tests. Finally, visualizing performance indicators for all three techniques have been illustrated.

The last point that needed to be discussed in this part is to argue the applicability of the proposed model. To show that, it is necessary to present here the accuracy that obtained for each types model. All figures and results showed that ML techniques have good ability to learn and understand classifying objects into correct classes. However, some accuracy rates have been obtained, which changed with the change of dataset.

Beside the conclusion of this work, below are some expectation that could be used to extend the work in future:

1. There are many ML techniques that need evaluation, such as Decisions tree, fuzzy logic, linear regression, Genetic Algorithm.

2. There is a field about finding optimum solution among some possible solutions. ML techniques could be used for that purses.

3. ML techniques could be used for a specific filed and application, not generally as done in this work.

4. ML becomes a part of Robotic Engineering Field, which is also part in Internet of Robotic Things. It means ML will be a hart of the field that dealing with Smart objects and things.

In this thesis three intelligent agents are used which they called BPNN, RBF, SVM. According to the results that we got form all intelligent agents RBF is better than other intelligent agents because the results show us no errors and the agent got full accuracy and other two intelligent are good and maybe for other works they are better than RBF but for the datasets which we work on it, RBF is best.

For the future work, the thesis try to use more different datasets including databases like voice, fingerprint, face and images for also more intelligent agents and find the results and comparison for all datasets, and the thesis will choose to each dataset which agent is good and how we work with each dataset and how use this dataset, and this thesis work will work on systems and give the quality for the systems and find the lake for each system.

REFERENCES

Abuhaiba, & SI, I. (2007). Offline signature verification using graph matching. *Turkish Journal of Electrical Engineering & Computer Sciences*, 89--104.

Al-Omari, M, Y., Abdullah, Sheikh, S. N., Omar, & Khairuddin. (2011). State-of-the-art in offline signature verification system. 59--64.

Arya, S, M., Inamdar, & S, V. (2010). A preliminary study on various off-line hand written signature verification approaches. *International Journal of Computer Applications*, 50--56.

Asmara, R. A. (n.d.). Gender Classification Using Human Gait Based on Skeleton-Model-Based Method Through Joint Angle Estimation and Model-Free Method with Gait Energy Motion Derived from Wavelet.

Batista, Luana, Rivard, Dominique, Sabourin, R., Granger, . . . Patrick. (2008). State of the art in off-line signature verification. *Pattern Recognition Technologies and Applications: Recent Advances*, 39--62.

Chinu Sayal, R. K. (n.d.). Analysis of Gait Recognition Algorithm Models with SURF and SVM. *ijariit*, 4.

Ding Wenrui, T. Y. (2008). Image and Video Quality Assessment Using Neural Network and SVM. *IEEE*, 112-116.

Drouhard, J.-P., Sabourin, Robert, Godbout, & Mario. (1996). A neural network approach to off-line signature verification using directional PDF. *Pattern Recognition*, 415--424.

Ghiasi, M. (2015). Complexity revisited. In *Application of Information and Communication Technologies (AICT), 2015 9th International Conference on* (pp. 553--557).

Impedovo, Donato, Pirlo, & Giuseppe. (2008). Automatic signature verification: the state of the art. *Systems, Man, and Cybernetics, Part C: Applications and Reviews, IEEE Transactions on*, 609--635.

Jain, K, A., Duin, PW, R., Mao, & Jianchang. (2000). Statistical pattern recognition: A review. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 4--37.

Jiansheng Huang, Z. J. (2017). SVM-based PQ disturbance recognition system. *IET*, 7.

Kalpna Kashyap, M. Y. (2013). FINGERPRINT MATCHING USING NEURAL NETWORK TRAINING. *ijecs*, 2041-2044.

Korotaev, M. W.-M. (2017). Evolutionary radial basis function network for gestational diabetes data analytics. *computational science*.

Larsen, K, P., Simonsen, B, E., Lynnerup, & Niels. (2008). Gait analysis in forensic medicine. *Journal of forensic sciences*, 1149--1153.

Lesne, A. a.-M. (2006). Chromatin fiber functional organization: some plausible models. Springer.

Maidstone, R. (2012). *Wavelets in a Two-Dimensional Context.* May.

Maltoni, Davide, Maio, Jain, D. a., Anil, Prabhakar, & Salil. (2009). Handbook of fingerprint recognition.

Moghaddam, B.-H. Y. (2000). Gender classification with support vector machines. IEEE.

Morris, & Anne, E. (2011). Gait analysis techniques to understand the effect of a hip strength improving program on lower-limb amputees.

Parmar, N, D., Mehta, & B, B. (2014). Face Recognition Methods & Applications. *arXiv preprint arXiv:1403.0485*.

PeiYun Zhang, S. S. (2018). An Online Fault Detection Model and Strategies Based on SVM-Grid in Clouds. *IEEE*, 12.

Qi Yang, Y. T. (2011). Gait recognition based on KDA and SVM. *IEEE* (pp. 160-163). IEEE.

Ramachandra, A. a., Srinivasa, J., Raja, B, K., Venugopla, Patnaik, K. a., & LM. (2009). Robust Offline signature verification based on global features. 1173--1178.

Rigoll, Gerhard, & Kosmala, A. (1998). A systematic comparison between on-line and off-line methods for signature verification with hidden Markov models. 1755--1757.

Saikia, Hemanta, Sarma, & Chandra, K. (2012). Approaches and issues in offline signature verification system. *Int J Comput Appl*, 45--52.

Salman Taseen Haque, M. S. (2017). Single Image Super-Resolution using Back propagation neural network. *IEEE* (pp. 1-5). IEEE.

Samuel, Daramola, Samuel, & Ibiyemi. (2010). Novel feature extraction technique for off-line signature verification system. *International Journal of Engineering Science and Technology*, 3137--3143.

Sing, Chandrani, Gopal, Arpita, Mishra, & Santosh. (2011). Extraction and analysis of faculty performance of management discipline from student feedback using clustering and association rule mining techniques. 94--96.

Teddy Surya Guhawan, N. m. (2017). Development of offline handwritten signature Authentication using ANN. *IEEE* (p. 4). IEEE.

Verma, Chandra, A., Saha, Dipankar, & Saikia. (2013). Forgery Detection in Offline Handwritten Signature using Global and Geometric Features. *IJCER*, 182--188.

Vinita Bhandiwad, B. T. (2017). Face Recognition and Detection using Neural Networks. *IEEE* (pp. 879-882). IEEE.

Wayman, James, Jain, Anil, Maltoni, Davide, . . . Dario. (2005). An introduction to biometric authentication systems.

Wildes, & P, R. (1997). Iris recognition: an emerging biometric technology. *Proceedings of the IEEE*, 1348--1363.

Wu Qiao, X. X. (2017). Error Prediction method of Electronic Transformers Based on RBF Neural Network. *IEEE* (pp. 567-571). IEEE.

Wudai Liao, W. L. (2017). Inverse Dynamics Control of A Parallel Robot Based on RBF Neural Network. *IEEE* (p. 4). IEEE.

Xinrong Liang, S. L. (2017). Highway Traffic Density Control Based on RBF Neural Network. *IEEE* (p. 3931_3936). IEEE.

Yogesh, VG, Patil, & Abhijit. (2014). Offline and Online signature Verification Systems: A Survey. *International Journal of Research in Engineering and Technology*, 328--332.