ADAMA SCIENCE AND TECHNOLOGY UNIVERSITY

SCHOOL OF ELECTRICAL ENGINEERING AND COMPUTING

DEPARTMENT OF COMPUTING

A Master's Thesis on:

A Learning Analytical Model for Higher Educational Data Using Artificial Neural Network

SUBMITTED IN PARTIAL FULFILLMENT OF THE REQUIREMENTS OF THE DEGREE OF

M.Sc. in INFORMATION TECHNOLOGY

By:

Yeharerwork Solomon

February, 2017

Adama, Ethiopia

ADAMA SCIENCE AND TECHNOLOGY UNIVERSITY
SCHOOL OF ELECTRICAL AND COMPUTING ENGINEERING
DEPARTMENT OF INFORMATION SYSTEM

A LEARNING ANALYTICAL MODEL FOR HIGHER EDUCATIONAL DATA USING ARTIFICIAL NEURAL NETWORK

A THESIS SUBMITTED TO THE SCHOOL OF GRADUATE STUDIES OF ADAMA SCIENCE AND TECHNOLOGY UNIVERSITY IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF MASTER OF SCIENCE IN INFORMATION SYSTEM

BY:

YEHARERWORK SOLOMON

Name and Signature of Members of the Examining Board

Name	Title	Signature	Date
Dileep Kumar G.	Advisor,	_______________	22/02/2017
Nemerg Sing	Chairperson,	________________	22/02/2017
Dr. Wondwossen M.	External examiner's name,	________________	22/02/2017
Bahiru S.	Internal Examiner's name	________________	22/02/2017
Mohammed K.	Department Head	_________________	22/02/2017
Dereje R.	School Dean	_________________	22/02/2017

February 2017

ADAMA, ETHIOPIA

DECLARATION

I declared that this thesis is my original work and has not been presented for a degree in any other universities, and all sources of material used for the study have been accordingly acknowledged.

____

Yeharerwork Solomon woldeyohanes

February 2017

This thesis work has been submitted for examination with my approval as a university advisor.

____

Dileep Kumar G

February 2017

DEDICATION

This thesis is dedicated to my family especially my father (Ato Solomon W/yohaness) and my mother (W/o Selamawit Belte) and my beloved husband (Ato Getnet Tibebu). They played a great role in motivating and supporting me in various aspects.

ACKNOWLEDGMENT

First and for most I would like to thank the Almighty GOD for his unending helps and blessings.

I would like to deeply thank my advisor Dr. Dileep for shaping up the study and guiding my thoughts and steps to go on the right track. I admire his diligence, timely response and encouragement.

Finally, I would like to thank all my friends who provided me with moral support, idea and encouragements throughout my study.

Abstract

Learning analytics is an emerging educational technology that involves measuring, collecting, analyzing and reporting learner-related data to monitor and predict learner's academic performance. Learning analytics provides early interventions to improve learners' learning and performance. This study aims to design and develop A learning analytics model by analyzing learner-related information based on artificial neural network using MATLAB to predict the variable that determine the students' performance rate. A number of variables that may possibly influence the performance of students' in e-learning are outlined. Such variables are login into the system, scorm view, assignment attempt, resource view and quiz exam score of students' in the online course of introduction to civic and ethical studies in Law department students are used as input variables for the ANN model. A model based on the Multilayer Perceptron and feed-forward artificial neural network Topology to develop the architecture of the research.

The successful prediction rate of variables that determine the students' performance using ANN model (Multilayer Perceptron and feed-forward artificial neural network) was 99.3%.

Key Words: Artificial Neural Network, student performance, learning analytics, mean squared error, independent and dependent variable, Hidden layer, e-learning

Table of content

Content | Page

List of Tables

List of figures

List of Abbreviation and Acronyms

LA……………………………………………………………Learning Analytics

EDM…………………………………………………………Educational Data Mining

Vs.…………………………………………………………..Verses

HEI…………………………………………………………Higher Education Institution

LMS…………………………………………………………Learning Management System

E-learning……………………………………………………..Electronic Learning

ANN…………………………………………………………Artificial Neural Network

MATLAB……………………………………………………..MATLAB- Matrix Laboratory

MLP…………………………………………………………Multi-Layer Perceptron

MSE…………………………………………………………Mean Square Error

RBF…………………………………………………………Radial Basis Function

DB……………………………………………………………Data Base

ETL…………………………………………………………Extract,Transform,Load

CHAPTER ONE

INTRODUCTION

1.1 Background of the Study

Learning is a product of interaction. Depending on the epistemology underlying the learning design, learners might interact with instructors and tutors, with content and/or with other people. Many educators expend enormous amounts of effort to designing their learning to maximize the value of those interactions.

In the traditional educational model, instructors have the principal role in the learning process. Students are assumed to have basic knowledge and skills, while instructors are expected to share their knowledge and experience. Learning is tested by means of procedural exams and homework. Before the Internet era, there were several types of distance-education models based on TV programs, manuals or recorded audios/videos. Typically, instructors were available to solve doubts by phone or mail. Although they allowed learning from home and presented a flexible timetable, the lack of interactivity hindered the learning process [1].

The Internet has dramatically changed the system, since most institutions have become interested in providing online courses. Besides the fact that they do not require large investments, these courses are not restricted to a specific geographical location or timetable, which increases the number of potential students. As a result, universities dedicated only to online education have emerged and traditional universities have expanded their offer with b-learning (hybrid classroom and online learning) and e-learning (pure online learning) courses.

This e-learning education method is a very recent phenomenon in Ethiopia context. Some universities have applied e-learning system recently. For instance, Adama Science and Technology University (ASTU), formerly known as Nazareth College of Technical Teachers Education, it is the first university which applies e-Learning environment.

The e-Learning program, said the first of its kind for a higher education institution here in Ethiopia, where all of the learning materials have been developed in cooperation with the Adama e-Teaching

Competence Centre, lecturers and professors from the university, is also planned to be widely exported to other universities in the country. E-learning is an innovative learning style that gives students 24-hour access to interactive education material and the modern information society. After the successful accomplishment of the e-learning program at ASTU, the program has also been implemented to Jimma University, Bahir Dar University and Mekele University immediately. It was also be applied to Arbaminch University, Hawasa University and Addis Ababa University recently.

However, e-learning courses also present higher dropout rates due to the fact that distance education may create a sense of isolation in students, which can feel disconnected from the other students, the instructors and the university [2] E-learning courses may be provided through Learning Management Systems (LMS) such as Moodle, Sakai and ILIAS, or Learning Platforms such as Knew ton and Dream Box. Informing this data into new insights that can benefit students, teachers, and administrators.

Learning analytics research uses data analysis to inform decisions makers on every tier of the education system, leveraging student data to deliver personalized learning, enable adaptive pedagogies and practices, and identify learning issues in time for them to be solved [3].A Learning analytic use ANN to predict which assessment method in e-learning more determine or more value in students' performance rate.

1.2 STATEMENT OF THE PROBLEM

E-learning has been known to be a fundamental thing for the improvement of teaching learning process at HEIs in Africa. According to [4] adopting e-learning in Africa will increase education access and quality, as well as lower education cost. There are an increasing number of success stories with e-learning throughout the African continent. While e- learning is not a cure for all the problems related to education in Africa, it is clearly a tool that now must be taken into serious consideration by policy makers and donors. Introducing e-learning technologies in primary, secondary, and higher education in Africa will clearly present many challenges. There is now considerable potential for expanding public-private partnership activities in e-learning, including the delivery of services, private finance initiatives, demand-side initiatives and strategic partnerships.

In Ethiopian HEIs, the teaching learning processes are conducted in traditional method that is face-to-face teaching. However, this method is prone to lots of drawbacks. The problem of the traditional teaching learning process is mainly taking more time than electronic learning. [1]Because of this reason, e-learning method is advisable. Some of the universities have applied e-learning. However according to the experience and the knowledge of the current researcher, students' performance in e-learning method in various universities is not adequate. For this reason, conducting research is very significant. Therefore, this motivates the researcher to conduct this research. In addition, to improve the effectiveness of students in e-learning, applying A learning analytics model is necessary. Moreover, the application of using A learning analytics model by using ANN has not studied in our country yet.

1.3 RESEARCH QUESTION

To achieve the aforementioned objectives and arrive at a sound conclusion, the researcher has raised the following four major questions primarily and attempted to find their answers.

- What variables show students' performance rate in e-learning?
- How a learning analytics serve as a better system of e-learning?
- What kind of advantage will be presented for the e-learning system by using the proposed frame work?
- In e-learning system which assessment method more affect the students' performance rate?

1.4 OBJECTIVES OF THE STUDY

1.4.1 General Objective

The general objective of this study is proposing and designing A Learning Analytical Model for indicating students' performance of higher educational using artificial neural network.

1.4.2 Specific Objectives

In order to achieve the above stated general objective, the researcher has undertaken the following specific objectives.

- To identify some suitable factors that affect a student performance
- Find better solution in enhancing students' performance
- To evaluate the effectiveness of the model by using different parameters such as MSE and confusion matrix

- To develop the appropriate neural network architecture including number of hidden layers, number of hidden nodes, number of output nodes, transfer function, divide function and training function

1.5 SCOPE & LIMITATION

The scope of this research is limited to propose and design A learning analytical model to predict variables that determine the students' performance rate based on an artificial neural network by using available training algorithm, transfer function and divide function. Due to the time constraint, the researcher has decided to show and apply how A learning analytical model is necessary to indicate students' performance rate only.

1.6 SIGNIFICANCE OF THE STUDY

According to [4], e-learning is affordable, saves time, and produces measurable results. E-learning is more cost effective than traditional learning because less time and money is spent in traveling. Since e-learning can be done in any geographic location and there are no travel expenses, this type of learning is less costly than doing learning at a traditional institute. Flexibility is a major benefit of e-learning. E-learning has the advantage of taking class anytime anywhere. Education is available when and where it is needed. Thus, E-learning can be done everywhere and anytime like at the office, at home, on the road, 24 hours a day, and seven days a week. E-learning also has measurable assessments which can be created.

So, it is expected that the beneficiaries of this research are the university itself, i.e. the teachers, the students, the administration and totally the community of Ethiopian HEI. Moreover, the study serves as a springboard for other researchers who have interest to enhance the quality of education by understanding the concept of A-learning analytical model.

The current study shows A learning analytical model is very important for better education so as to indicate which assessment method most determine the students' performance rate in e-learning educational system by collecting and analyzing the data.

1.7 APPLICATION OF THE RESULTS

The result of this research can be used by Ethiopia Higher education institution to develop and increase the effectiveness of the student in teaching learning process by using online learning.

1.8 ORGANIZATION OF THE THESIS

This research report is organized into five chapters. The **first Chapter** briefly discusses background to the problem area, states the problem, the general and specific objectives of the study, the scope and limitation, and application of the results of the research.

Chapter two lays the foundation for other parts of this research. It thoroughly discusses important or related literatures in the area of learning analytics and ANN technology. The concepts pertaining to the learning analytics and ANN concept, different task of ANN and its application in the problem such like ANN techniques, types, architecture and model are reviewed in chapter two.

Chapter three explains about data preparation and data pre-processing. It also defines the methods, the techniques used in this paper. Moreover, the overall methodology used in this study also reviewed in this chapter.

Chapter four presents the experimentation phase of the study at hand. This chapter provides a detailed discussion about the experiment part of this work using MATLAB. The results of the experiments were also discussed here.

Finally, **Chapter five** provides conclusion of the research, and also presents recommendation for future work based on the research findings of the present study.

CHAPTER TWO

LITERATURE REVIEW

In this chapter, an attempt has been made to review literature on the concept of higher education institution, e-learning Vs traditional learning, A learning analytics and ANN technology and its application in higher educational institute in particular. The researcher believes that it is quite important to review so as to provide the background information about the model which is going to be built in this study.

2.1 Description of the Related Concepts

2.1.1 E-learning VS. Traditional learning

We consider two way of teaching learning system, traditional system and modern learning system (i.e. E-learning system).

Traditional learning: is the learning under the scope of classroom, viewed as teacher-center and static. The learning is conducted with the whole class participating, taking place with in classroom and the school. The instructor conducts the lesson according to the study program and curriculum. Moreover, the instructor dictates the structure of the lesson and the division of time [5].

The components of traditional learning include blackboard, books, instructor and students in a classroom. The instructor usually talks more than the students and the students learn "What" but not "How". The instructor is responsible to set all of tasks for students. When comparing E-Learning to the traditional learning, some of researcher rated E-Learning as more effective than traditional learning. However, there are some, rejected it with the reason of less social interaction, high cost investment, and technical problems in communication and computing technology.
[6] Says "E-Learning has a potential to improve students' performance, but to boost success in the digital economy, individuals and institutions of higher learning must use research to guide the adaptation and integration of new technology into the learning process."

E-Learning: is Internet-enabled learning. To provide a comprehensive understanding, E-Learning is defined as a new education concept by using the Internet technology [6]. It is defined as interactive learning in which students learn through the usage of computers as an educational medium [7]. Moreover, [6] pointed that “E-Learning covers a wide set of applications and processes, including multimedia online activities such as the web, Internet video CD-ROM, TV and radio.

The components of E-Learning include content delivery in multiple formats, management of the learning experience, a networked community of learners, and content developers and experts. E-Learning is personalized, focusing on the individual learner. Its environment includes self-paced training, many virtual events, mentoring, simulation, collaboration, assessment, competency road map, authoring tools, e-store, and learning management system [8]. According to [8], “E-Learning provides faster learning at reduced cost, increased access to learning, and clear accountability for all participants in the learning process.”

E-Learning includes many components that are familiar from traditional learning, such as: presentation of ideas by the students, group discussions, arguments and many other forms of conveying information and accumulating knowledge. The contents of the course’s curriculum might be organized according to subjects and in a serial manner.

E-Learning also includes advantages which are not found in traditional learning, such as: time for digesting the information and responding, enhanced communication among the learners, both as regards quality and as regards urgency, knowledge being acquired and transferred among the learners themselves, the ability to conduct an open discussion, where each learner gets more of an equal standing than in a face-to-face discussion, access to information and to discussion ability, responses may be made around the clock with no restrictions, a higher motivation and involvement in the process on the part of the learners.

The following table summarizes several opinions regarding the comparison between traditional learning and E-Learning:

	Traditional Learning	**E-Learning**
Classroom Discussions	The teacher usually talks more than the student	The student talks at least as much as or more than the teacher
Learning Process	The learning is conducted with the whole class participating; there is almost no group or individual study	Most of the learning process takes place in groups or by the individual student.
Subject Matter	The teacher conducts the lesson according to the study program and the existing curriculum	The student participates in determining the subject matter; the studying is based on various sources of information, including web data banks and net-experts located by the student.
Emphases in the Learning Process	The students learn “what” and not “how”; the students and the teachers are busy completing the required subject matter quota; the students are not involved in inquiry-based education and in solving problems, but rather in tasks set by the teacher.	The students learn “how” and less “what”; the learning includes research study which combines searching for and collecting information from web data banks and authorities on the communications network; the learning is better connected to the real

		world, the subject matter is richer and includes material in different formats.
Motivation	The students' motivation is low, and the subject matter is "distant" from them.	The students' motivation is high due to the involvement in matters that are closer to them and to the use of technology.
Teacher's Role	The teacher is the authority	The teacher directs the student to the information.
Location of Learning	The learning takes place within the classroom and the school	The learning takes place with no fixed location
Lesson Structure	The teacher dictates the structure of the lesson and the division of time	The structure of the lesson is affected by the group dynamics.

Table 1 Traditional Learning vs. E-Learning

2.1.2 Advantage and Disadvantage of E-learning

Understanding the e-learning advantages and disadvantages is important when considering how to make instructional and learning decisions.

E-learning Pros and Cons need to be considered in equal measure. Many organizations and institutions provide different forms of training and instruction to their employees or learners. Typically, they provide needed training by sending people to school, holding in-house training classes, or providing manuals and self-study guides. In some situations, it is advantageous for them to use other forms of e-learning instead of the traditional training.

Other times it is less useful. As with anything else, there are benefits and limitations, as well as pros and cons. There are many advantages to online and computer-based learning when compared to traditional face-to-face courses and lectures. There are a few disadvantages as well. Let us briefly compare them.

Advantages of E-Learning

- Class work can be scheduled around personal and professional work, resulting in flexible learning.
- Reduces travel cost and time to and from school
- Learners may have the option to select learning materials that meets their level of knowledge and interest
- Learners can study wherever they have access to a computer and Internet
- Self-paced learning modules allow learners to work at their own pace
- Flexibility to join discussions in the bulletin board threaded discussion areas at any hour, or visit with classmates and instructors remotely in chat rooms
- Different learning styles are addressed and facilitation of learning occurs through varied activities
- Development of computer and Internet skills that are transferable to other facets of learner's lives
- Successfully completing online or computer-based courses builds self-knowledge and self-confidence and encourages students to take responsibility for their learning

Disadvantages of E-Learning

- Unmotivated learners or those with poor study habits may fall behind
- Lack of familiar structure and routine may take getting used to
- Students may feel isolated or miss social interaction thus the need to understanding different learning styles and individual learner needs.
- Instructor may not always be available on demand

- Slow or unreliable Internet connections can be frustrating
- Managing learning software can involve a learning curve
- Some courses such as traditional hands-on courses can be difficult to simulate

But we come to Ethiopian context the student doesn't effective in e-learning.as many researcher advices to apply learning analytics to avoid some of the above disadvantage of e-learning and make the e-learning user are effective.

2.2 The Concept of Learning Analytics

"Analytics" is a term used in business and science to refer to computational support for capturing digital data to help informed decision-making. With the growth of huge data sets and computational power, this extends to designing infrastructures that exploit rapid feedback, to inform mealier interventions, whose impact can in turn be monitored. Organizations have increasingly sensitive 'digital nervous system' providing real me feedback on the external environment and the effects of actions.

In higher education, "analytics" can refer to a wide array of functions on both the academic and administrative sides, from improving student performance and instructor effectiveness to managing student enrollment and budgeting strategically. Efforts by colleges and universities to use data more systematically to better institutional decision making and operational efficiency can be viewed as a version of what the private sector calls "business intelligence." But many educators eschew that term, perhaps because it has too corporate a feel, opting instead for "institutional intelligence," "academic intelligence," or "operational intelligence". It's important to make the distinction clear between what has been referred to as either academic analytics or institutional analytics versus learning analytics. In the case of institutional analytics oblige calls academic analytics, that's an interest in trying to move the needle on some higher-order institutional metrics like attrition rates or year-to-year completion rates. On the other hand, learning analytics is focused at the level of the individual learner and on giving learners actionable information to make their decisions about study within a given course or set of courses.

[9]said about learning analytics, Learning analytics is defined as "the measurement, collection, analysis, and reporting of data about learners and their contexts, for purposes of understanding

and optimizing learning and the environments in which it occurs”. Learning analytics involves the collection and analysis of data to predict and improve student success. The author also said learning analytics allow us to suggest new learning opportunities or different courses of action to our students.

2.2.1 Types of Data in learning analytics

There are vast amounts of data that educational institutions have at their disposal to help faculty meet their goals of improving students’ performance and increasing retention. As an example, below is a table of the types of available data in the LMS that they use at one institution. The first column includes data that is automatically generated by the organization. The second column is an example of the types of data that might be generated by the instructor, much of which can be stored in the LMS.

Data Generated by LMS	**Data Generated by Instructor**
Number of Times Resource Accessed	Grades on Discussion Forum
Date and Time of Access	Grades on Assignment
Number of Discussion Posts Generated	Grades on Tests
Number of Discussion Posts Read	Final Grades
Types of Resource Accessed	Number (and Type) of Questions Asked in a Discussion Forum
	Number of Emails Sent to Instructor

Table 2 Types of Data Available for Learning Analytics

2.2.2 Benefits of Using Learning Analytics

There are various reasons for wanting to use learning analytics. Among these are to improve student success, increase retention, and improve accountability. But as an individual faculty member, why might you want to make use of learning analytics on a smaller, course-level scale? In this section, we highlight several of the benefits and advantages of using learning analytics. [9]Describe a multitude of benefits of using learning analytics for higher education. Several of these benefits are focused on an administrative level, such as improving decision-making and informing resource allocation, highlighting an institution’s successes and challenges, and increasing organizational productivity. From the perspective of a faculty member, they suggest that learning analytics can help faculty identify at-risk learners and provide interventions, transform pedagogical approaches, and help students gain insight into their own learning. Having data at hand and knowing what to do with it can allow us to realize these benefits. For example, if we learn (from a correlational analysis) that student performance on certain activities is not related to final grades, then we might consider modifying these activities. Similarly, we can use data from an LMS to build a model of successful student behaviors, which might include the frequency of LMS tool use, frequency of accessing discussion board posts, and the number of times taking quizzes. If we can build a model of successful student behaviors, then we can encourage (with data!) our students to engage in these behaviors. Alternatively, we can also identify at-risk students as ones who deviate from this model.

Likewise, [4]suggest that learning analytics can help faculty by informing them of the gaps in knowledge displayed by their students. Understanding these knowledge gaps can help faculty focus their attention on particular students or pieces of information.

2.3 Artificial Neural Networks

2.3.1 Basic Concept of Artificial Neural Network

Now we will briefly discuss the basic concepts of ANNs. An Artificial Neural Network (ANN) [10] is a mathematical model that tries to simulate the structure and functionalities of biological neural networks. Basic building block of every artificial neural network is artificial neuron, that is, a simple mathematical model (function). Artificial neuron is supposed to mimic the action of a biological neuron [11]. Artificial neuron is a basic building block of every artificial neural network. Its design and functionalities are derived from observation of a biological neuron that is basic building block of biological neural networks (systems) which includes the brain, spinal cord and peripheral ganglia. Similarities in design and functionalities can be seen in Fig.1. Where the top side of a figure represents a biological neuron with its synapse, dendrites and axon and where the bottom side of a figure represents an artificial neuron with its inputs, hidden layer and outputs.

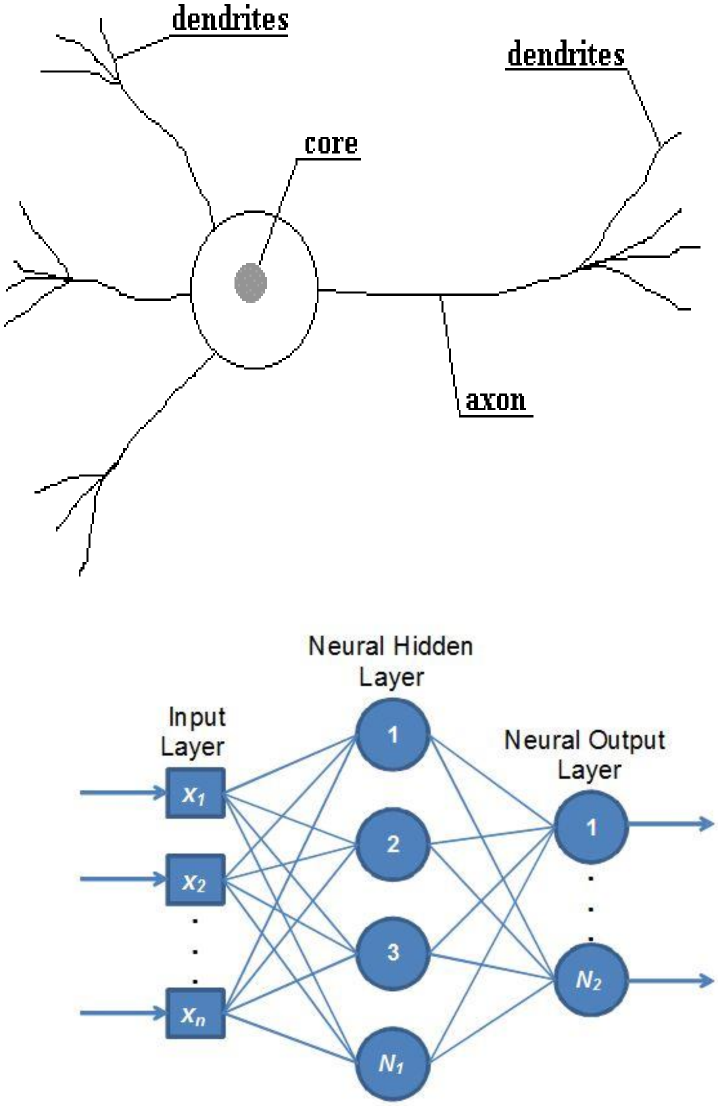

Figure 1: Comparison between the biological and artificial neuron

A neuron transmit input to the cell body of another neuron through the synaptic connections to the dendrites. The output signal or information (consisting of nerve impulses or electrical triggers) from a cell is transmitted through the connection wire (axon) to the synapses of other neurons.

2.3.2 Type of Architecture in Artificial Neural Network

The generic terms 'Structure' and 'Architecture' are used as synonyms for network topology [12]. A neural network topology represent the way in which neuron are connected to form a network. In other word, the neural network topology can be seen as the relationship between the neurons by means of their connection. The two most widely used architectures or topology of ANN are described below:

2.3.2.1 Feed-Forward Networks

In a feed-forward network, information flows in one direction along connecting pathways from input layer via the hidden layers to the final output layer. There is no feedback (loops) i.e. the output of any layer does not affect the preceding layer.

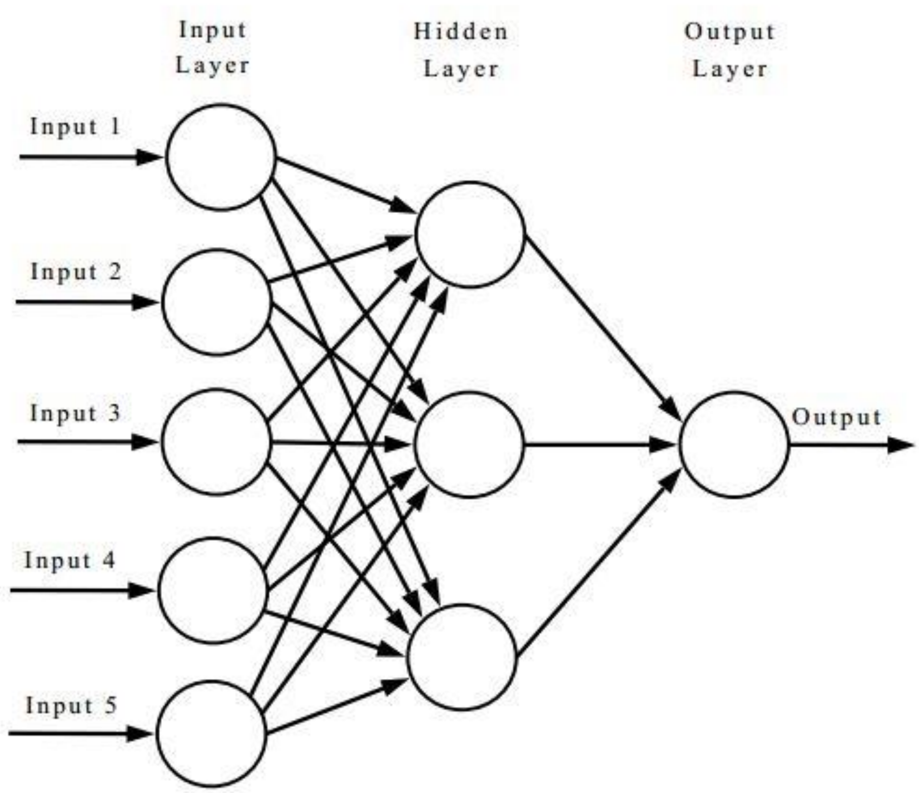

Figure 2 Feed-forward artificial neural network

2.3.2.2 Recurrent Artificial Neural Networks

These networks differ from feed-forward network architecture there is at least one feedback loop. Thus, in these networks, for example there could exist one layer with feedback connection as shown in figure below. There could also be neurons with self-feedback links, i.e. the output of neurons is feedback into itself as input.

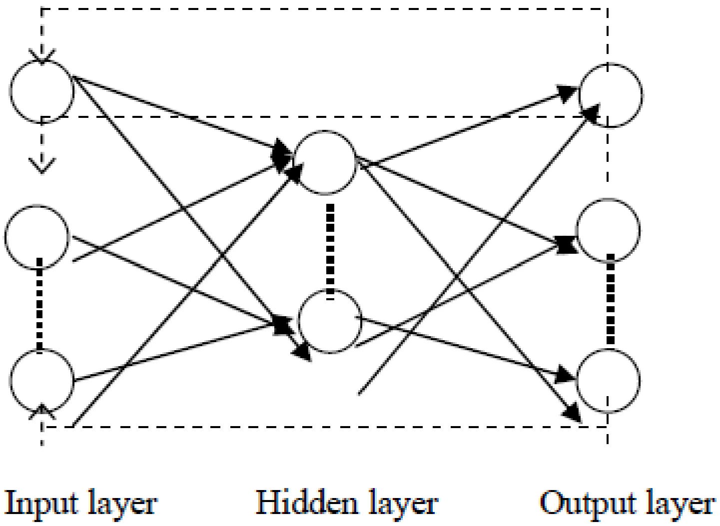

Figure 3 Recurrent artificial neural network

Comparison of the above type of Architecture:

Feed forward network	**Recurrent Artificial Neural Networks**
Information is propagated from the inputs to the outputs	Can have arbitrary topologies
Time has no role (NO cycle between outputs and inputs)	the output of neurons can be taken as feedback into itself as input
Training is easy and good.	Training is more difficult -Performance may be problematic Stable Outputs may be more difficult to achieve.

	-Unexpected behavior

Table 3 Comparison of 2 type of ANN Architecture

2.3.3 Types of Artificial Neural Network

The most common type of neural network are multilayer perceptron and radial basis function. Each type of artificial neural network has their own unique characteristics as described below.

2.3.3.1 Multilayer Perceptron

The multilayer perceptron is the most known and most frequently used type of neural network. [13]. MLP is [14] typically composed of several layers of nodes. Nodes that are no target of any connection are called input layer. A MLP that should be applied to input patterns of dimension n must have n input layer, one for each dimension. Input layer are typically enumerated as layer 1, layer 2, layer 3, … node that are no source of any connection are called output layer. A MLP can have more than one output neuron. The number of output neurons depends on the way the target values (desired values) of the training patterns are described. All nodes that are neither input neurons nor output neurons are called hidden layer. The input layer and output layer are connected by one or more intermediate hidden layers. Since the graph is acyclic, all node can be organized in layers, with the set of input layers being the first layer.

2.3.3.2 Radial Basis Function Networks

Feed forward neural networks with a single hidden layer that use radial basis activation function for hidden neurons are called radial basis function (RBF) network. A radial basis function neural network has an input, hidden and output layer [15].The connections exist between the input layer and the hidden layer, and between the hidden layer and the output layer. It uses radial combination functions in the hidden layer, based on the squared Euclidean distance between the input vector and the weight vector. It also uses exponential or softmax activation functions in the hidden layer, in which case the network is a Gaussian RBF network. MLPs are said to be distributed-processing networks because the effect of a hidden unit can be distributed over the entire input space. However Gaussian RBF networks are said to be local processing networks because the effect of a hidden unit is usually concentrated in a local area centered at the weight vector.

In this thesis Multilayer perceptron is used. The main reason to choose MLPs are the simplest and most commonly used neural network architecture that means it's much too selected architecture (feed forward network). [16]Artificial neural networks (ANN_s) of the feed forward type, normally called multilayer perceptron have been applied in these areas because they are excellent functional mappers.

2.3.4 Learning in Artificial Neural Network

Based on [10] there are three major learning paradigms; supervised learning, unsupervised learning and reinforcement learning. Usually they can be used by any given type of artificial neural network architecture. Each learning paradigm has many training algorithms.

2.3.4.1 Supervised learning

Supervised training is the most common form of neural network training [17].As supervised training proceeds the neural network is taken through several iterations, or epochs, until the actual output of the neural network matches the anticipated output, with a reasonably small error. Each epoch is one pass through the training samples.

Supervised learning is a machine learning technique that sets parameters of an artificial neural network from training data. The task of the learning artificial neural network is to set the value of its parameters for any valid input value after having seen output value. The training data consist of pairs of input and desired output values that are traditionally represented in data vectors. Supervised learning can also be referred as classification, where we have a wide range of classifiers, each with its strengths and weaknesses. In order to solve a given problem of supervised learning various steps has to be considered. In the first step we have to determine the type of training examples. In the second step we need to gather a training data set that satisfactory describe a given problem. In the third step we need to describe gathered training data set in form understandable to a chosen artificial neural network. In the fourth step we do the learning and after the learning we can test the performance of learned artificial neural network with the test (validation) data set. Test data set consist of data that has not been introduced to artificial neural network while learning.

2.3.4.2 Unsupervised learning

Unsupervised training is similar to supervised training [17]. Except that no anticipated outputs are provided. Unsupervised training usually occurs when the neural network is to cluster the inputs into several groups. The training progresses through many epochs, just as in supervised training.

Unsupervised learning is a machine learning technique that sets parameters of an artificial neural network based on given data and a cost function which is to be minimized. Cost function can be any function and it is determined by the task formulation. Unsupervised learning is mostly used in applications that fall within the domain of estimation problems such as statistical modelling, compression, filtering, blind source separation and clustering. In unsupervised learning we seek to determine how the data is organized. It differs from supervised learning and reinforcement learning in that the artificial neural network is given only unlabeled examples. One common form of unsupervised learning is clustering where we try to categorize data in different clusters by their similarity. Among above described artificial neural network models, the Self-organizing maps are the ones that the most commonly use unsupervised learning algorithms.

2.3.4.3 Reinforcement learning

Reinforcement learning is a machine learning technique that sets parameters of an artificial neural network, where data is usually not given, but generated by interactions with the environment. Reinforcement learning is concerned with how an artificial neural network ought to take actions in an environment so as to maximize some notion of long-term reward. Reinforcement learning is frequently used as a part of artificial neural network's overall learning algorithm.

After return function that needs to be maximized is defined, reinforcement learning uses several algorithms to find the policy which produces the maximum return. Naive brute force algorithm in first step calculates return function for each possible policy and chooses the policy with the largest return. Obvious weakness of this algorithm is in case of extremely large or even infinite number of possible policies. This weakness can be overcome by *value function approaches* or *direct policy estimation*. Value function approaches attempt to find a policy that maximizes the return by maintaining a set of estimates of expected returns for one policy; usually either the current or the optimal estimates. These methods converge to the correct estimates for a fixed policy and can also be used to find the optimal policy. Similar as value function approaches the direct policy estimation can also find the optimal policy. It can find it by searching it directly in policy space what greatly increases the computational cost.

Reinforcement learning is particularly suited to problems which include a long-term versus short-term reward trade-off. It has been applied successfully to various problems, including robot control, telecommunications, and games such as chess and other sequential decision making tasks.

[17] Supervised training or learning is better than the other. Because supervised learning is accomplished by giving the neural network a set of sample data along with the anticipated outputs from each of these samples. Supervised training is the most common form of neural network training. As supervised training proceeds the neural network is taken through several iterations, or epochs, until the actual output of the neural network matches the anticipated output, with a reasonably small error. Each epoch is one pass through the training samples. According to this the researcher used supervised training or learning.

2.6 Review of Related Literature

The effectiveness of the student in higher education has great impact on the society and the development of the country. Thus, applying ANN to model learning analytics for higher educational data records can help to understand which variable affects the student performance rate and to find possible solutions. This can help the decision makers to formulate better teaching learning process.

Some ANN researches on the analysis of higher educational data have been done globally and no locally. Reviewing this related works that were done learning analytics by using ANN tools and techniques at different place and time in the same problem domain gave the researcher an in depth insight for this paper. The researcher has summarized some of the most important works as follows.

Naser A. etal [17] presented Artificial Neural Network (ANN) model, for predicting the performance of a sophomore student enrolled in engineering majors in the Faculty of Engineering and Information Technology in Al-Azhar University of Gaza. A model based on the Multilayer Perceptron Topology was developed and trained using data spanning five generations of graduates from the Engineering Department of the Al-Azhar University, Gaza. It consists of an input layer with 10 neurons, a hidden layer with 6 node and an output layer with 4 node. They used total of 150 sophomore students' records in the analysis. A success rate of 84.6% was achieved.

Naser,A. [18]Employed Artificial Neural Networks and expert systems to obtain knowledge for the learner model in the Linear Programming Intelligent Tutoring System to be able to determine the academic performance level of the learners in order to offer the learner the suitable difficulty level of linear programming problems to solve. Feed forward Back propagation algorithm was trained with a group of learner's data to predict their academic performance. The accuracy of predicting the performance of the learners were very high (92%) and thus states that the Artificial Neural Network is skilled enough to make suitable predictions.

Stamos T. and Andreas V. [16]Presented a model using an artificial neural for predicting student graduation outcomes. The network was developed as a three-layered perceptron or multilayer perceptron and was trained using the back propagation principles. The first layer (input level) comprised of 11 neurons (processing element), the third layer (output level) comprised of 2 neurons and the second layer (hidden level) used 4 neurons. For training and testing various experiments were executed. In these experiments, a sample of 1,407 profiles of students was used. The sample represented students at Waubonsee College and it was divided into two sets. The first set of 1,100 profiles was used for training and the remaining 307 profiles were used for testing. The average predictability rate for the training and test sets were 72% and 68%, respectively.

Sultan A. etal [19]proposed a prototype of a Decision Support System (DSS) for providing the knowledge for optimizing the newly adopted e-learning education strategy in educational institutions. The proposed DSS involves Database (DB) engine, Data Mining (DM) engine and Artificial Intelligence (AI) engine. All these engines work integrally together in order to extract the knowledge necessary to improve the effectiveness of any institution strategy, including the e-learning. The DSS integrates ETL (Extract, Transform, Load) tool for extracting a task–relevant data from operational database, and transforming it to suitable formats and loading into the DSS's data store. This paper put the design but not study detail as well or arguments are not supported by practical evidence.

Bahadir E. [20] Presented the results of an experimental comparison study of Logistic Regression Analysis (LRA) and Artificial Neural Network (ANN) for predicting prospective mathematics teachers' academic success when they enter graduate education. The model has been implemented using MATLAB and SPSS .A sample of 372 student profiles was used to train and test the model. The strength of the model can be measured through Logistic Regression Analysis (LRA). The average correct success rate of students for ANN was higher than LRA. The successful prediction rate of the back-propagation neural network (BPNN, or a common type of ANN was 93.02%, while the success of prediction of LRA was 90.75%. The results show that artificial neural network could be useful for identifying the factor that affect the student performance.

Czerkawski B. [21] In this paper explain Learning analytics is about better understanding of the learning and teaching process and interpreting student data to improve their success and learning experiences. The purpose of this paper is to provide an overview to learning analytics in higher education and more specifically, in e-learning and explore some of the issues and challenges around it and discuss the e-learning and learning analytics relationship. The author concludes with a brief discussion of major implications of learner analytics for e-learning.

This paper put the overview but not study detail in learning analytical as well or arguments are not supported by practical evidence. My paper differs from in this paper the work in detail that means it add the implementation phase or arguments are supported by practical evidence using ANN, it works more widely in prediction analytics also the technique uses to predict have some differences. After that depends on the predict analytics result the new system predict the factor that affect the student performance in e-learning.

Uhler B.&Hurn J. [22]The main aim of their paper is to define learning analytics, how it has been used in educational institutions, what learning analytics tools are available, and how faculty can make use of data in their courses to monitor and predict student performance. Finally, they discuss several issues and concerns with the use of learning analytics in higher education.

They prepare a simple case study and tool to support their paper. They assessed many institutions that applied the use of learning analytics to improve student success and retention. Below is a table that highlights some of these success stories. As the evidence from the table shows below, many of the successful institutions have used or designed learning analytics tools that often provide a “dashboard” indicator to both students and faculty. For example, Purdue University created SIGNALS, which extracts data and provides a dashboard for both students and faculty to track student progress. Other institutions, such as UMBC, make use of a learning analytics tool built in to their institutions' LMS, which allows them to track student progress. As indicated in the third column, most of these institutions are using these data to help their students perform better in a course.

Institution	**Learning Analytic Tool**	**Uses of Data**
University of Central Florida	EIS (Executive Information System)	Data management
Rio Salado Community College	PACE (Progress and Course Engagement)	Track student progress in course; intervention
Northern Arizona University	GPS (Grade Performance System)	Student alerts for academic issues and successes
Purdue University	Course Signals System	Student alerts for academic issues; intervention
Ball State University	Visualizing Collaborative Knowledge Work	Enhance knowledge building work
University of Michigan	E^2Coach	Student support and intervention
University of Maryland Baltimore County (UMBC)	ANN	Track performance and predict student success
Graduate School of Medicine, University of Wollongong	BIRT (Business Intelligence and Reporting Tools)	Reveal continuity of care issues

Table 4 Institutions and Learning Analytics Tools *[22]*

And they conclude their paper by suggesting many benefits to learning analytics; most notably that it can inform how we help our students succeed. In educational institutions, we have an enormous amount of data at our disposal. Our ability to harness this data and use it to inform what we do in the classroom, whether face-to-face or online, is at the heart of learning analytics.

Strength of this paper the authors try to support the study by applying simple case study and tools, Arguments and the themes in the introduction part are indicated precisely, the conclusion summarizes the significance and relevance of the paper well. The weakness of the paper the authors didn’t use clear theoretical framework related to the topic. This type of problem avoided in my paper.

The related papers stated above [19] [21] [22] provide the idea of e-learning and learning analytic idea. However, the researchers in those studies didn’t show practical examples in their research. But, the researcher of this study has discussed the concept of A learning analytical model by supporting practical evidence in depth. The other related paper mentioned above [20] Compare logistic regression strategy and ANN by using database engine, data mining engine and artificial intelligence engine. In addition, the related papers above [16] [17] [18] explain about predicting students’ performance using different ANN model and data. However, my study differ from the above related papers since my argument are supported by good and practical evidence. Furthermore, this study is not only indicating or predicting the students’ performance but also it is identifying the factors that affect students’ performance. In addition to this, the criteria to predict the students’ performance used by this researcher were not used by any of the above researchers. As it is mentioned in the earlier section of the study, the main objective of this paper is identifying the factors that affect the students’ performance based on student assessment method in e-learning.

CHAPTER THREE

METHODOLOGY

Methodology is defined as the steps and procedures that one follows to achieve the objectives and research questions of a certain study. It is something that shows the direction in which a research follows to attain the end. In order to achieve the specified objectives of the current research, the researcher has reviewed different related literatures (books, journal articles, Conference proceeding papers, and the Internet) in order to have detailed understanding on the present research.

This chapter discusses about research design, data collection, data preparation or pre-processing, tool selection tasks of the analytics. This embraces detail description from which the research data has been collected. This chapter also describes the methods and tools used in this study and describes conceptual and architectural analysis of student performance which is important in the model development process by including determination of ANN architecture, ANN training, and post training analysis.

3.1 Research Design

Research designs are the plans and procedure that cover the decision from broad assumption to detailed methods of the research. Below shows the conceptual diagram of the proposed system.

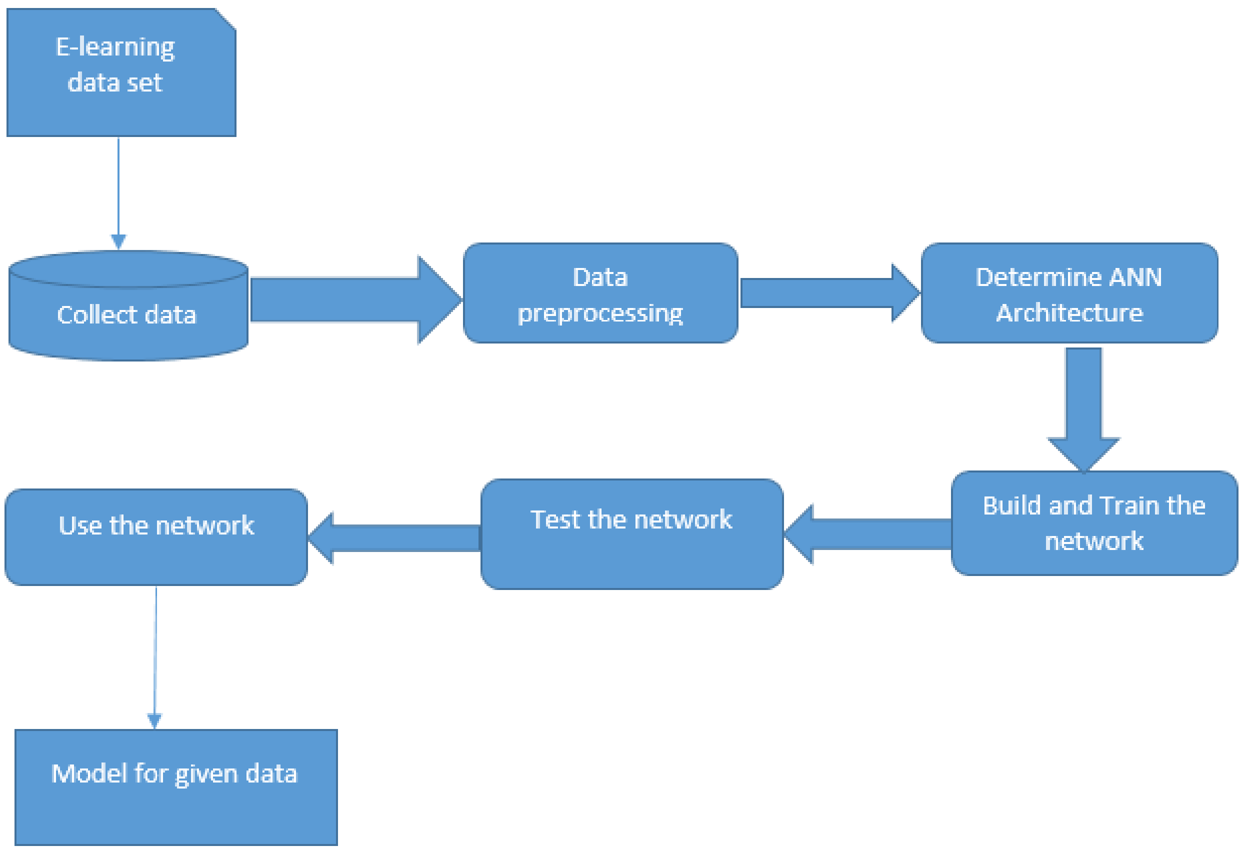

Figure 4 Diagram of the proposed model

3.2 Data Preparation

3.2.1 Data collection

Data collection is a critical initial step. Therefore, the researcher has got the data set from ASTU E-learning administrator data base. Number of frequency when the student login in to the e-learning system, corm view , assignment attempt , resource view and quiz result are independent variable of the model.

3.2.2 Sample Size

The researcher conducted this study in one university (ASTU) which is applying E-learning. From this university the researcher has got 1200 recorded data, and the researcher has got the variables that determine the students' performance rate in e-learning system.

3.2.3 Data Analysis

In order to reach to the final output and meet the objectives of the study, the collected data were analyzed. The analysis technique or approach used in this study are ANN, the researcher follows feed forward model and a tool used in this study is MATLAB. MATLAB is used for the implementation of this study. It is used for the analysis of the collected data and Data Analysis and help to define and differentiate the behavior of the collected data. The researcher's main reason for using MATLAB is that it is flexible, it has many prepared functions and proper performance to ANN simulation. Moreover, the researcher also used MS-Excel for organizing the data.

3.2.4 Data Preprocessing

The researcher preprocesses the dataset into input data, which is suitable to use in the neural network. Some duplicate records are removed by using the technique "remove constant rows" which is available in MATLAB. The independent variables of the model are represented by 1, 2 and 3.That means quiz result takes 1, 2, 3 (1=0-4, 2=5-7, 3=8-10), scorm view takes 1, 2, 3(1=0-15, 2=16-30, 3=31-40), resource view takes 1, 2, 3(1=0-20, 2=21-60, 3=61-114) ,assignment attempt takes 0,1 and 2 and also number of frequency login in to the system takes 1, 2, 3 (1=1-20, 2=21-40&3=41-255). Dependent variables are also represented by 1, 2, and 3. 1 for low performance, 2 for medium performance and 3 for high performance.

3.3 ANN Architecture Determination

The researcher use three types of layers. These are: The input layer, hidden layer and output layer.

3.3.1 Number of Input Nodes

The input layer should represent the condition for which the neural network is trained for the output layer. Every input neuron should represent some independent variable that has an influence over the output of the neural network. An important step in developing ANN models is [23] to select the model input (independent) variables that have the most significant impact on model performance. The number of nodes in the input layers is [24] limited by the number of model inputs.

The current researcher took scorm view, assignment attempt, resource view, number of frequency login in to the system, and quiz result as independent variables (inputs variable) of the model in this study as the list below shows:

1. Scorm View
2. Assignment Attempt
3. Resource View
4. Number of Frequency Login in to the System
5. Quiz result

S/N	Input Variable	Domain	
1	Scorm View	1-15 16-30 31-40	1 2 3
2	Assignment Attempt	0 1 2	
3	Resource View	1-20 21-60 61-144	1 2 3
4	Number of Frequency Login in to the System	1-20 21-40 41-255	1 2 3
5	Quiz result	0-4 5-7 8-10	1 2 3

Table 5 Input Data Transformation

3.3.2 Number of Hidden Layers and Nodes

The hidden layer and nodes [25] play very significant roles for many successful applications of neural networks. It is the hidden nodes in the hidden layer that allow neural networks to detect the feature, to capture the pattern in the data, and to perform complicated nonlinear mapping between independent and dependent variables. Single hidden layer is used in this study. The main reason is influenced by theoretical works which show that a single hidden layer is [26] [27] enough for ANNs to approximate any complex nonlinear function with any desired accuracy. In addition to this [26] [27] also states that one hidden layer is adequate to approximate any continuous function provided that sufficient connection weights are given. Hidden node starting from n/2 up to 2n+1 are used in this study. n refers number of input variable.in one process the system uses one hidden layer at a time. The main reason is several rules of thumb have also [25] been proposed to decide the number of hidden nodes in the hidden layer. This study use all hidden layer to come the result more accurate.

3.3.3 Number of Output Nodes

The output layer of the neural network is [25] what actually presents an output of the neural network to the external environment. Whatever pattern is presented by the output layer can be directly traced back to the input layer. The number of output neurons should directly related to the type of work that the neural network is to perform. To consider the number of neurons to use in output layer one must consider the intended use of the neural network. If the neural network is to be used to classify items into groups, then it is often suitable to have one output neurons for each group that the item is to be assigned into. Low performance, medium performance and high performance are dependent variables (outputs) of the model for this study.

Now I can design neural network architecture of the proposed model based on the above explanation.

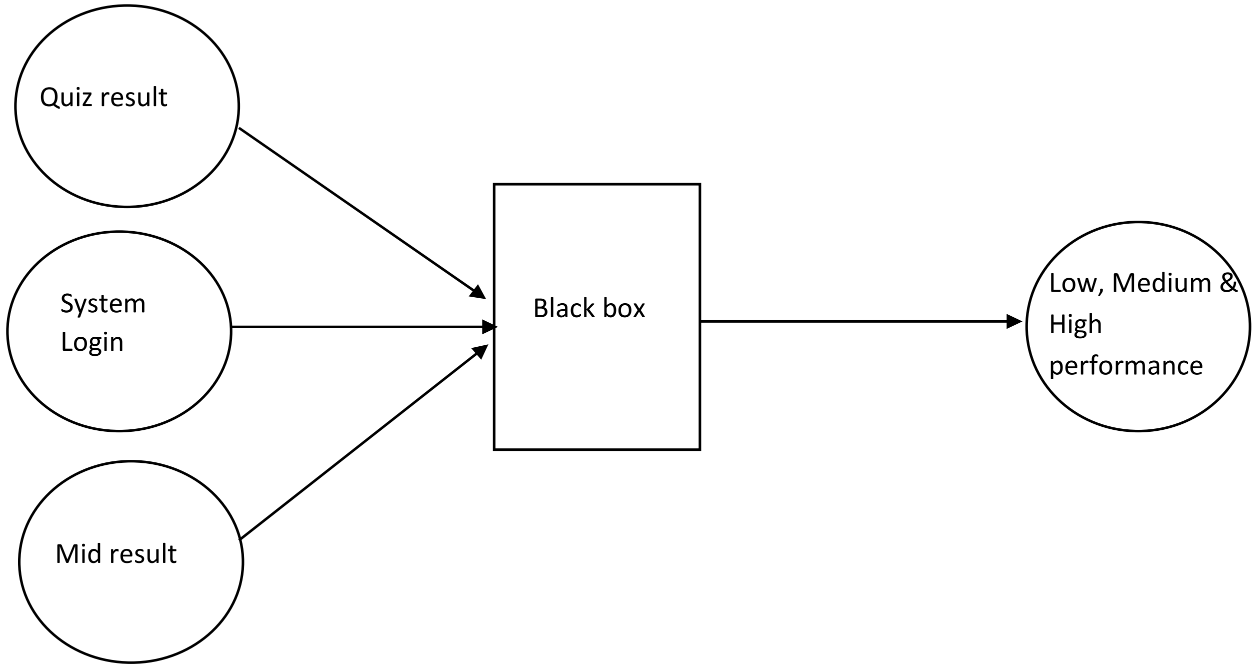

Figure 5 Neural network architecture of the proposed model

3.4 ANN Training

Train the network by using three function. Each function has different algorithm.

3.4.1 Transfer Function

The activation function is [25] also called the transfer function. It determines the relationship between inputs (independent variables of model) and outputs (dependent variables of model) of a node and a network. An elementary neuron with R inputs is [28] shown below on Figure 6. Each input is weighted with an appropriate w. The sum of the weighted inputs and the bias forms the input to the transfer function f. Neurons may use any differentiable transfer function f to generate their output.

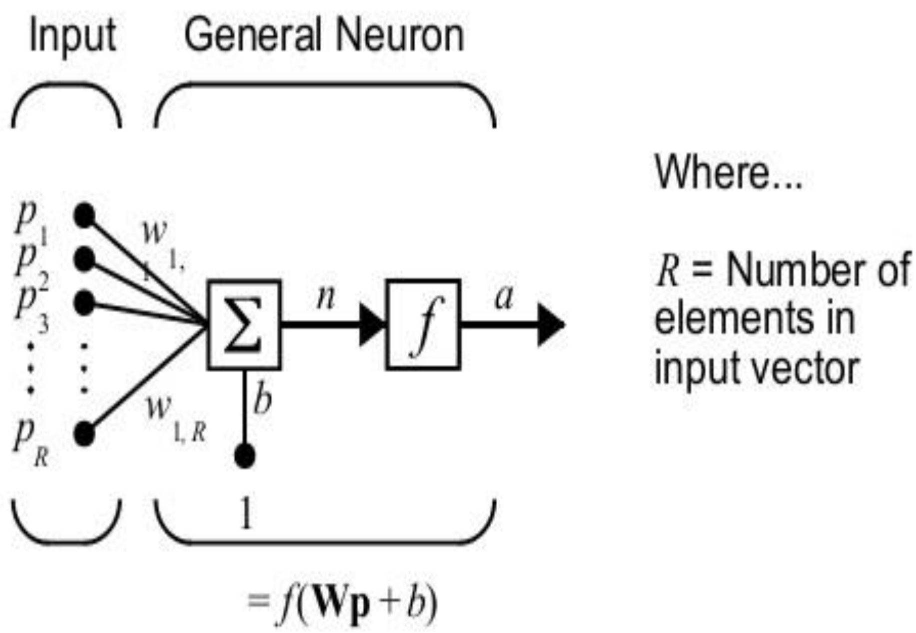

Figure 6 Elementary neuron with R inputs

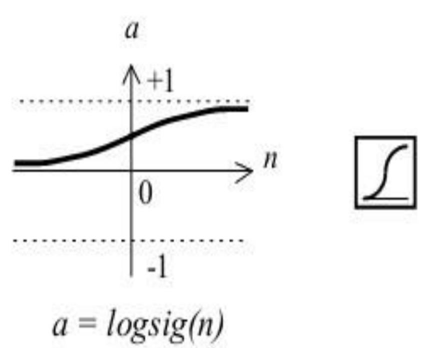

Figure 7 Log-Sigmoid Transfer Function

As shown on Figure 7 above the function logsig generates outputs between 0 and 1 as the neuron's net input goes from negative to positive infinity. Alternatively, multilayer networks may use the tan-sigmoid transfer function tansig.

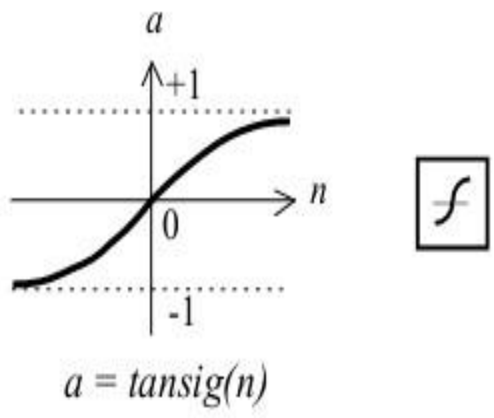

Figure 8 Tan-Sigmoid Transfer Function

Infrequently, the linear transfer function purelin is used in back propagation networks.

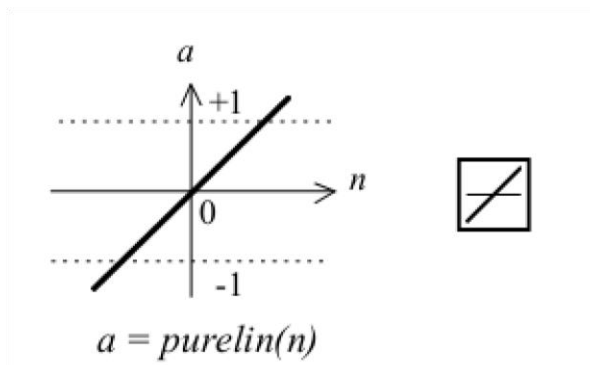

Figure 9 Linear Transfer Function

If the last layer of a multilayer network has sigmoid neurons, then the outputs of the network are limited to a small range. If linear output neurons are used the network outputs can take on any value. In backpropagation it is important to be able to calculate the derivatives of any transfer functions used. Each of the transfer functions above, logsig, tansig, and purelin, can be called to calculate its own derivative. In this thesis the optimal transfer function is selected through experiments.

3.4.2 Divide Function

There are four functions provided [28] for dividing data into training, validation and test sets. They are dividerand (the default), divideblock, divideint, and divideind. If net.divideFcn is set to `dividerand' (the default), then the data is randomly divided into the three subsets. If net.divideFcn is set to `divideblock', then the data is divided into three subsets using three contiguous blocks of the original data set (training taking the first block, validation the second and testing the third). If net.divideFcn is set to `divideint', then the data is divided by an interleaved method, as in dealing a deck of cards. It is done so that different percentages of data go into the three subsets. When net.divideFcn is set to `divideind', the data is divided by index. The optimal divide function for this study is selected through experiments.

3.4.3 Training Function

There are several different [29] back propagation training algorithms. They have a variety of different computation and storage requirements, and no one algorithm is best suited to all locations. Table 1 below shows the list of the training algorithms with their characteristics.

Function	Description
Traingd	Basic gradient descent. Slow response, can be used in incremental mode training.
Traingdm	Gradient descent with momentum. Generally faster than traingd. Can be used in incremental mode training.
Traingdx	Adaptive learning rate. Faster training than traingd, but can only be used in batch mode training.
Trainrp	Resilient backpropagation. Simple batch mode training algorithm with fast convergence and minimal storage requirements.
Traincgf	Fletcher-Reeves conjugate gradient algorithm. Has smallest storage requirements of the conjugate gradient algorithms.
Traincgp	Polak-Ribiére conjugate gradient algorithm. Slightly larger storage requirements than traincgf. Faster convergence on some problems.
Traincgb	Powell-Beale conjugate gradient algorithm. Slightly larger storage requirements than traincgp. Generally faster convergence.
Trainscg	Scaled conjugate gradient algorithm. The only conjugate gradient algorithm that requires no line search. A very good general purpose training algorithm.

Trainbfg	BFGS quasi-Newton method. Requires storage of approximate Hessian matrix and has more computation in each iteration than conjugate gradient algorithms, but usually converges in fewer iterations.
Trainoss	One step secant method. Compromise between conjugate gradient methods and quasiNewton methods.
Trainlm	Levenberg-Marquardt algorithm. Fastest training algorithm for networks of moderate size. Has memory reduction feature for use when the training set is large.
Trainbr	Bayesian regularization. Modification of the Levenberg-Marquardt training algorithm to produce networks that generalize well. Reduces the difficulty of determining the optimum network architecture.

Table 6 list of training algorithms

Trainlm, trainrp, trainscg, traincgb, traincgp, traincgf, trainoss training functions are used for this study. The main reason is these training functions are [29] used for classification purpose.

3.4.4 Stopping Criteria

The networks are trained according to the following values of training parameters. The values of parameters that are listed below are default values in MATLAB used as stopping criteria.

Maximum number of epochs to train = 1000

Performance goal (mse) = 0

Minimum performance gradient = 1e-5

Maximum validation failures = 6

Training stops when any of the following conditions occur:

1) The maximum number of epochs (repetitions) is reached.

2) Performance has been reduced to the goal.

3) The performance gradient moves below the minimum.

Validation performance has increased more than the maximum validation failures.

3.5 Post Training Analysis

Recognition accuracy, speed of training, correctness [30] are parameters for performance evaluation of training functions. One of the most important parameter is the mean squared error.

$$MSE = \frac{\sum_{i=1}^{N}(y_i - o_i)^2}{N}$$

Where y_i is the target and o_i is the observed output and N is the number of data set. The other parameter for performance evaluation of training functions is:

- ✓ The confusion matrix gives the percentage of correct and incorrect classifications in the simulated feed-forward back propagation networks.

In this study, MSE and confusion matrix are used for the performance evaluation of the model.

CHAPTER FOUR

EXPERMENTATION, RESULT AND DISCUSSION

This chapter present the process of experimentation, result and discussion. The experiment is performed based on the concept discussed in the previous chapter. Different experiments are conducted to find optimal number of hidden layer neurons, divide function, transfer function and training functions for the model of this study by comparing their result in terms MSE. First hidden layer neurons starting from 3 up to 11 with single hidden layer are tested by using default parameters such as training function trainlm, transfer function tansig, and divide function dividerand. After the completion of the first step the optimal hidden layer neurons are selected and applied for each of the divide function such as dividerand, divideint, divideind, and, divideblock in order to find the optimal divide function by using default parameters such as training function trainlm, and transfer function tansig. Both the optimal hidden layer neurons and the optimal divide function applied for each of the transfer functions such as tansig, logsig, and purelin to find the optimal transfer function by using the default parameter such as training function trainlm. Finally the optimal hidden layer neurons, optimal divide function, and optimal transfer function applied for each of the training functions such as trainlm, trainscg, trainrp, trainsgp, traincgb, traincgf, and trainoss to find the optimal training function. MSE and confusion matrix are used for performance evaluation of the training functions.

4.1 Experiments to Find Optimal Hidden Layer Neurons

To find optimal hidden layer neurons experiments are performed by using default parameters (training function= “trainlm”, transfer function= “tansig”, divide function= “dividerand”). Here the performance result of the hidden neurons starting from 3 up to 11 are compared. Table 7 below shows the comparison among the neurons.

Number of neurons	Best validation performance (MSE)	Epoch
3	0.016368	15
4	0.012989	108

5	0.0051375	51
6	0.012506	52
7	0.004462	17
8	0.0059544	32
9	0.004474	32
10	0.0055065	37
11	0.0052861	66

Table 7 Comparison among hidden neurons

As we can see from table 7 above hidden neurons 5, 7,9,10 & 11 are selected I can select due to their optimal MSE. So these neurons are applied for each of the divide, transfer and train functions in the experiments that are performed to find the optimal divide, transfer and train function.

4.2 Experiments to Find Optimal Divide Function

To find the optimal divide function experiments are performed by using default parameters (training function= “trainlm”, transfer function= “tansig”). Here the performance result of the divide functions are compared. Table 8 below shows the comparison among divide functions using 3 hidden neurons.

Number of neurons	Divide function	Best validation performance (MSE)	Epoch
5	dividerand	0.0051375	51
5	divideint	0.025666	12
5	divideind	4.9665e-09	1
5	divideblock	0.023877	10

Table 8 Comparison among divide functions with 5 hidden neurons

As we can see from table 8 above the best validation performance 0.0051375 at epoch 51 is the optimal with the divide function dividerand. Figure 10 below shows the best validation performance for dividerand.

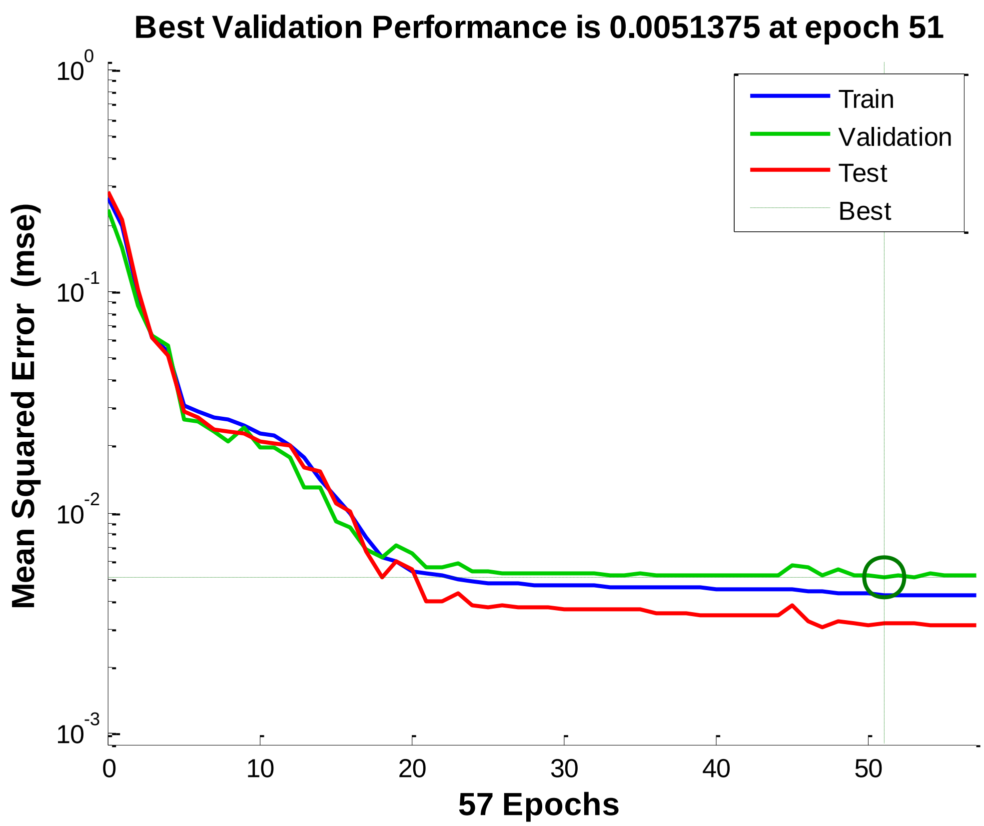

Figure 10 Best validation performance for dividernd with 5 hidden neurons

As we can see from Figure 10 above the best validation performance is 0.0051375 at iteration 51.

Table 9 below shows the comparison among divide functions using 7 hidden neurons.

Number of neurons	Divide function	Best validation performance (MSE)	Epoch
7	dividerand	0.004462	17
7	divideint	0.021467	19
7	divideind	0.22847	0
7	divideblock	0.014562	14

Table 9 Comparison among divide functions with 7 hidden neurons

As we can see from table 9 above the best validation performance 0.004462 at epoch 17 is the optimal with the divide function dividerand. Figure 11 below shows the best validation performance for dividerand.

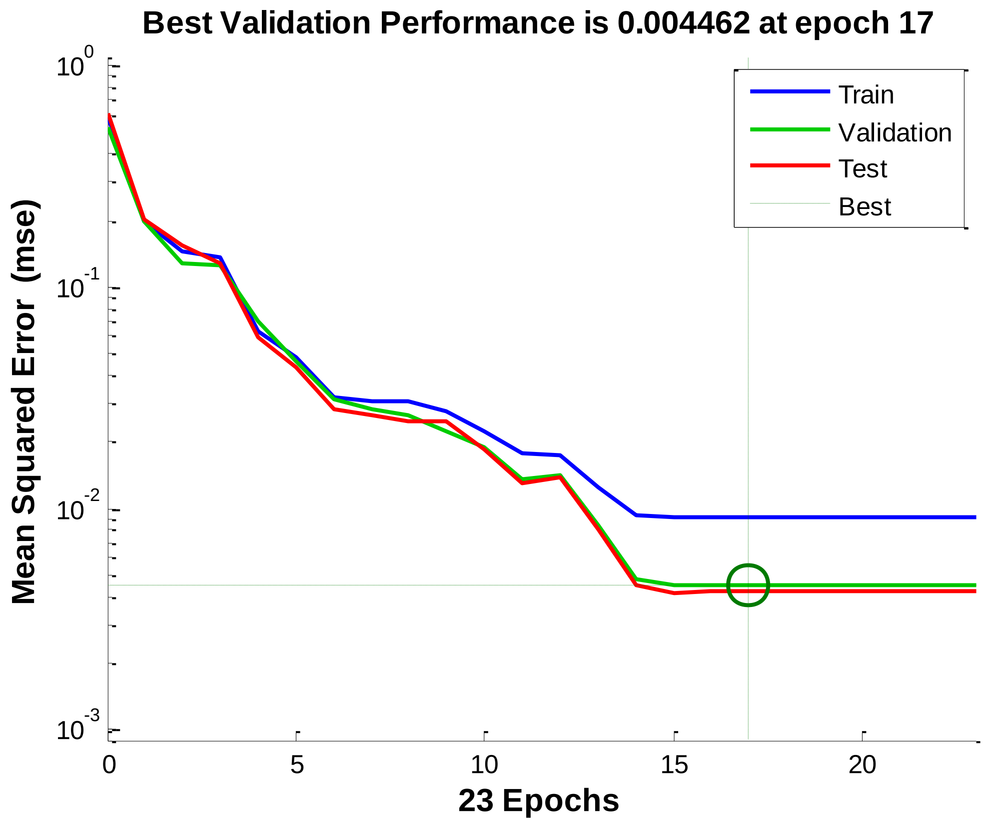

Figure 11 Best validation performance for dividernd with 7 hidden neurons

As we can see from Figure 11 above the best validation performance is 0.004462 at iteration 17.

Table 10 below shows the comparison among divide functions using 9 hidden neurons.

Number of neurons	Divide function	Best validation performance (MSE)	Epoch
9	dividerand	0.004474	32
9	divideint	0.014409	22
9	divideind	2.4165e-05	1
9	divideblock	0.018438	25

Table 10 Comparison among divide functions with 9 hidden neurons

As we can see from table 10 above the best validation performance 0.004474 at epoch 32 is the optimal with the divide function dividerand. Figure 12 below shows the best validation performance for dividerand.

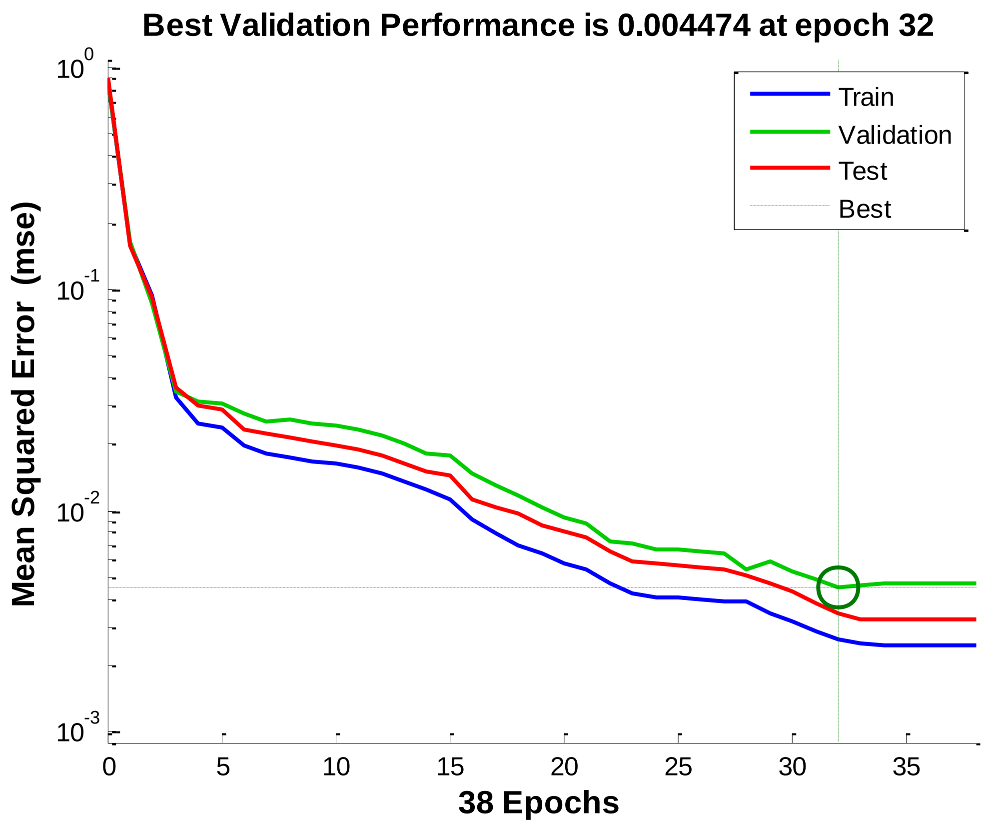

Figure 12 Best validation performance for dividerand with 9 hidden neurons

As we can see from Figure 12 above the best validation performance is 0.004474 at iteration 32.

Table 11 below shows the comparison among divide functions using 10 hidden neurons.

Number of neurons	Divide function	Best validation performance (MSE)	Epoch
10	dividerand	0.0055065	37
10	divideint	0.018168	11
10	divideind	7.7742e-06	1
10	divideblock	0.0059425	11

Table 11 Comparison among divide functions with 10 hidden neurons

As we can see from table 11 above the best validation performance 0.0055065 at epoch 37 is the optimal with the divide function dividerand. Figure 13 below shows the best validation performance for dividerand.

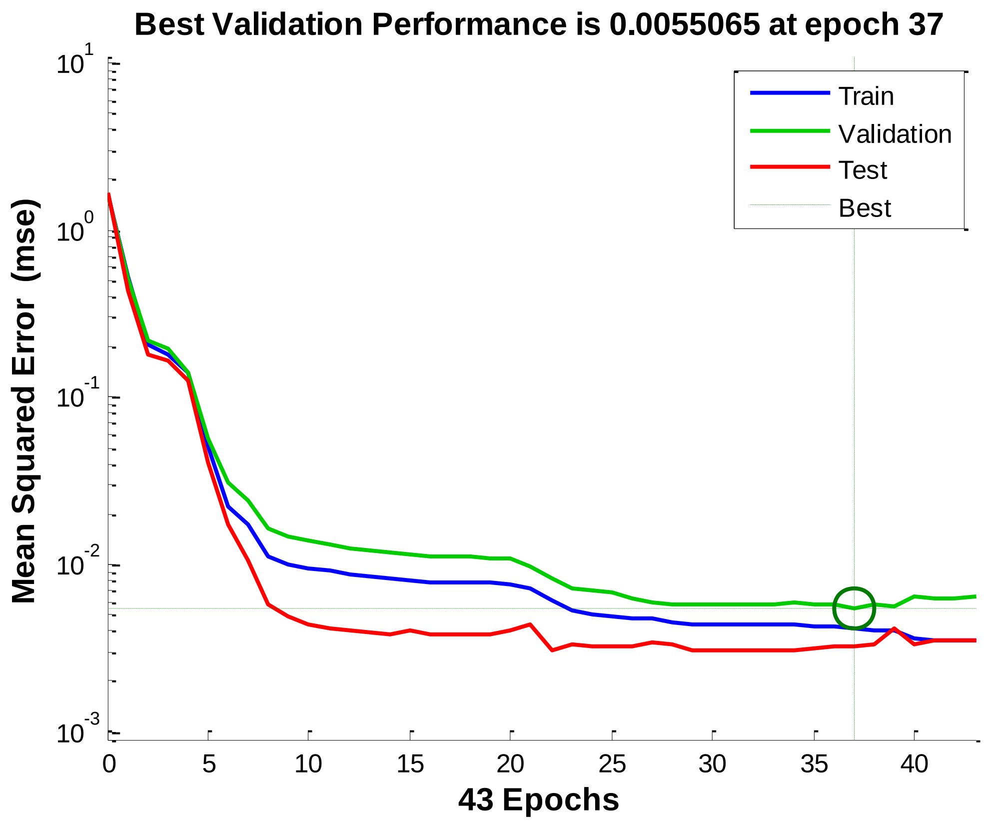

Figure 13 Best validation performance for dividerand with 10 hidden neurons

As we can see from Figure 13 above the best validation performance is 0.0055065 at iteration 37.

Table 12 below shows the comparison among divide functions using 11 hidden neurons.

Number of neurons	Divide function	Best validation performance (MSE)	Epoch
11	dividerand	0.0052861	66
11	divideint	0.017891	14
11	divideind	2.494e-06	1
11	divideblock	0.0059415	45

Table 12 Comparison among divide functions with 11 hidden neurons

As we can see from table 12 above the best validation performance 0.0052861 at epoch 66 is the optimal with the divide function dividerand. Figure 14 below shows the best validation performance for dividerand.

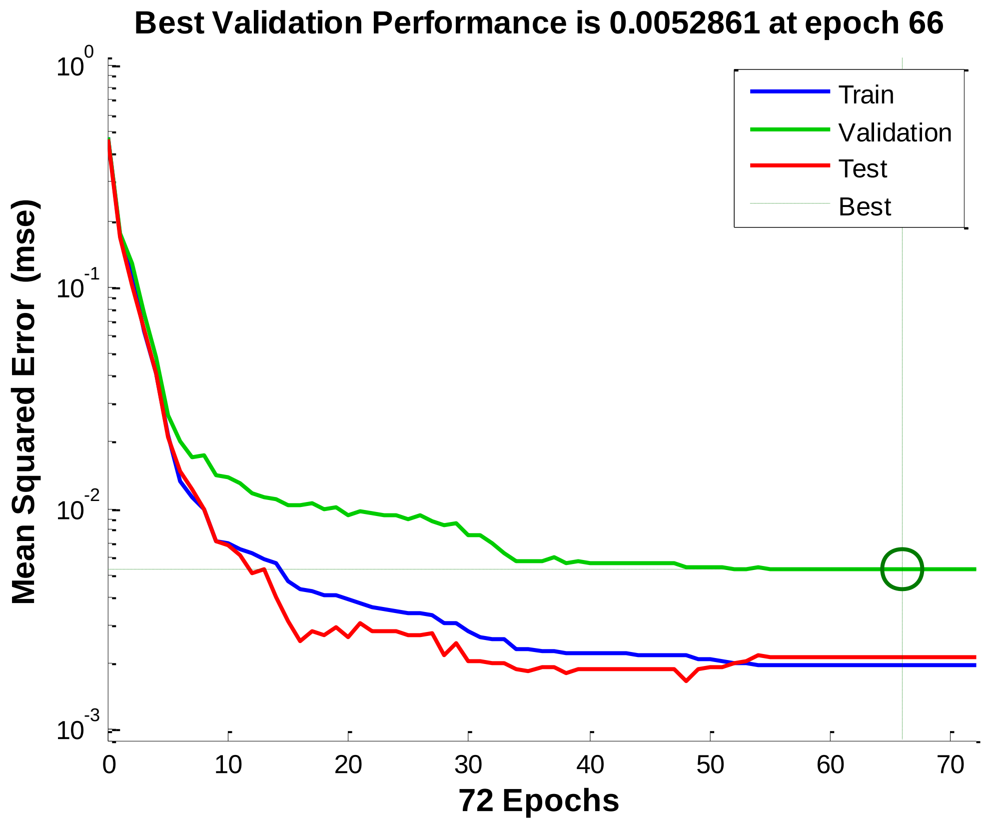

Figure 14 Best validation performance for dividerand with 11 hidden neurons

Table 13 below shows the comparison among divideblock, divideint and dividerand with optimal MSE.

Number of neurons	Divide function	Best validation performance (MSE)	Epoch
5	dividerand	0.0051375	51
7	dividerand	0.004462	17
9	dividerand	0.004474	32
10	dividerand	0.0055065	37

11	dividerand	0.0052861	66

Table 13 Comparison among divideblock, divideint and dividerand with optimal MSE

As we can see from table 13 above the best validation performance 0.004462 at epoch 17 is the optimal with the divide function dividerand. So dividerand is applied for each of transfer functions in order to find the optimal transfer function for this model.

4.3 Experiments to Find Optimal Transfer Function

To find the optimal transfer function experiments are performed by using default parameter (training function= “tansig”). Here the performance result of the transfer functions are compared.

Table 14 below shows the comparison among transfer functions using 5 hidden neurons.

Number of neurons	Transfer function	Best validation performance (MSE)	Epoch
5	tansig	0.0051375	51
5	logsig	0.0045386	40
5	purelin	0.11622	3

Table 14 Comparison among transfer functions with 5 hidden neurons

As we can see from table 14 above the best validation performance 0.0045386 at epoch 40 is the optimal with transfer function logsig. Figure 15 below shows the best validation performance for logsig.

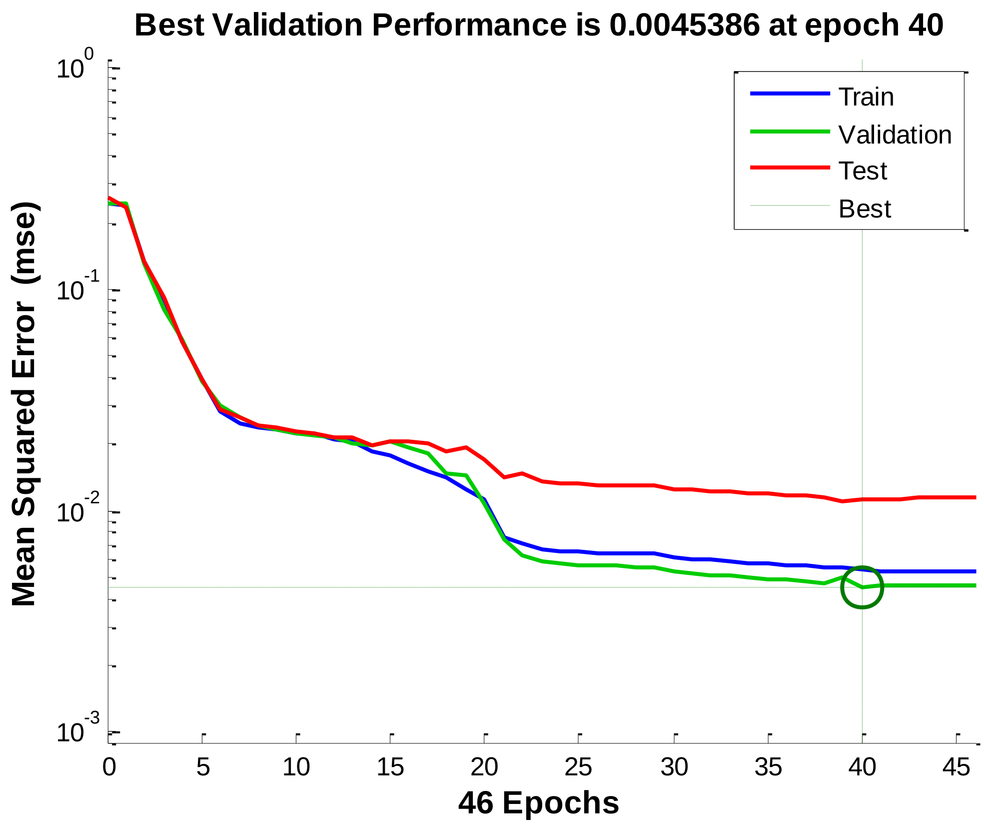

Figure 15 Best validation performance for logsig with 2 hidden neurons

As we can see from Figure 15 above the best validation performance is 0.0045386 at iteration 40.

Table 15 below shows the comparison among transfer functions using 3 hidden neurons.

Number of neurons	Transfer function	Best validation performance (MSE)	Epoch
7	tansig	0.004462	17
7	logsig	0.0054496	60
7	purelin	0.11677	3

Table 15 Comparison among transfer functions with 7 hidden neurons

As we can see from table 15 above the best validation performance 0.004462 at epoch 17 is the optimal with transfer function tansig. Its figure is similar to dividerand because tansig value also take the default value so its figure is similar to dividerand at hidden neurons 7.

Table 16 below shows the comparison among transfer functions using 9 hidden neurons.

Number of neurons	Transfer function	Best validation performance (MSE)	Epoch
9	tansig	0.004474	32
9	logsig	0.0018561	74
9	purelin	0.12648	3

Table 16 Comparison among transfer functions with 9 hidden neurons

As we can see from table 16 above the best validation performance 0.00118561 at epoch 74 is the optimal with transfer function logsig. Figure 16 below shows the best validation performance for logsig.

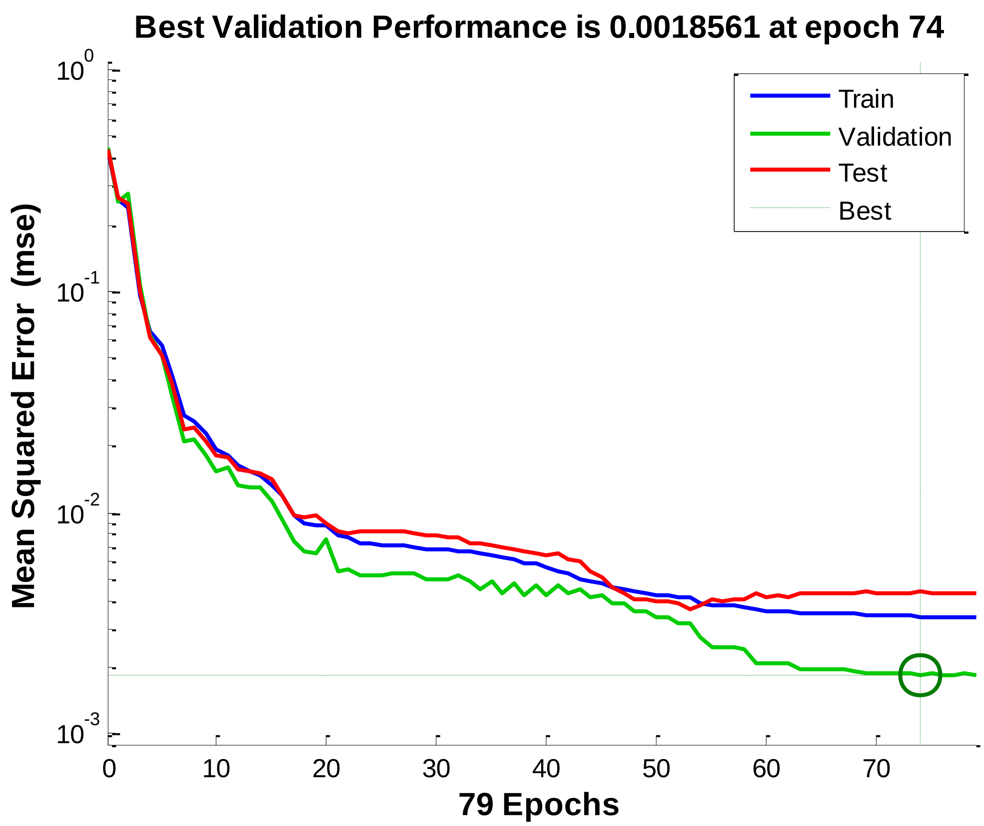

Figure 16 Best validation performance for logsig with 4 hidden neurons

As we can see from Figure 16 above the best validation performance is 0.0018561 at iteration 74.

Table 17 below shows the comparison among transfer functions using 10 hidden neurons.

Number of neurons	Transfer function	Best validation performance (MSE)	Epoch
10	tansig	0.0055065	37
10	logsig	0.011269	14
10	purelin	0.12261	2

Table 17 Comparison among transfer functions with 10 hidden neurons

As we can see from table 17 above the best validation performance 0.0055065 at epoch 37 is the optimal with transfer function tansig. Its figure is similar to dividerand because tansig value also take the default value so its figure is similar to dividerand at hidden neurons 10.

Table 18 below shows the comparison among transfer functions using 11 hidden neurons.

Number of neurons	Transfer function	Best validation performance (MSE)	Epoch
11	tansig	0.0052861	66
11	logsig	0.0022395	68
11	purelin	0.11677	3

Table 18 Comparison among transfer functions with 11 hidden neurons

As we can see from table 18 above the best validation performance 0.0022395 at epoch 68 is the optimal with transfer function logsig. Figure 16 below shows the best validation performance for logsig.

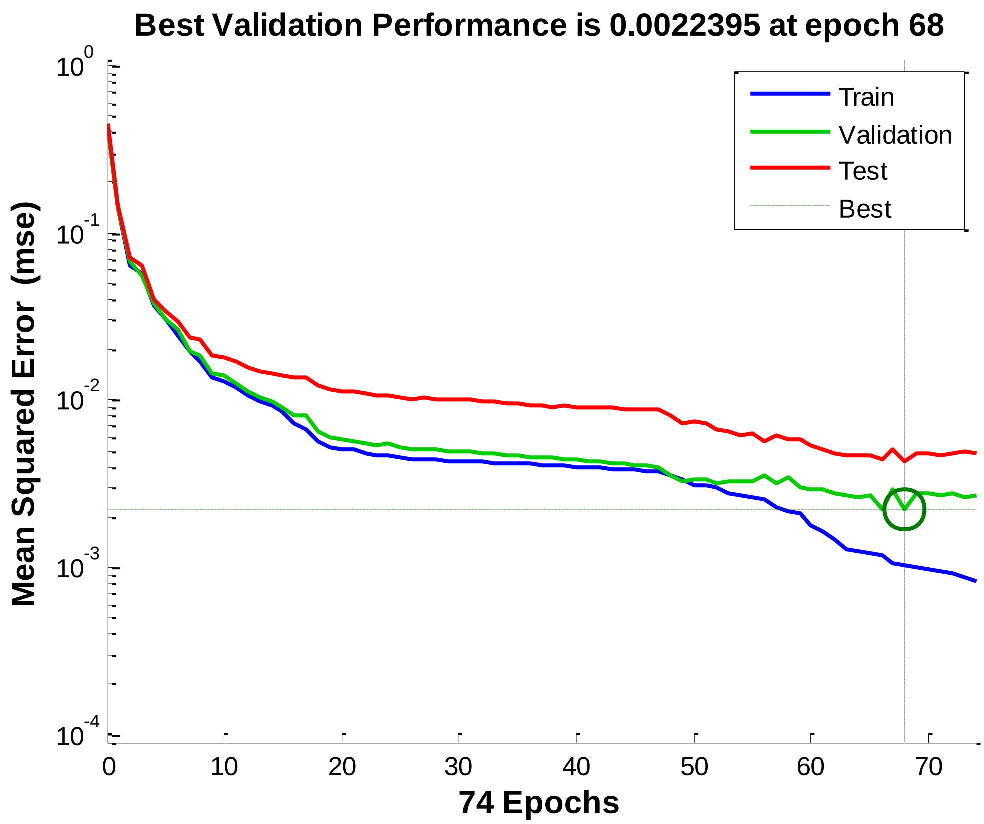

Figure 17 Best validation performance for logsig with 11 hidden neurons

Table 20 below shows the comparison between logsig and tansig with optimum MSE.

Number of neurons	Transfer function	Best validation performance (MSE)	Epoch
5	logsig	0.0045386	40
7	tansig	0.004462	17
9	logsig	0.0018561	74
10	tansig	0.0055065	37
11	tansig	0.0022395	68

Table 19 Comparison between logsig and tansig with optimal MSE

As we can see from table 20 above the best validation performance 0.0018561 at epoch 74 is the optimal with the transfer function logsig. So logsig was applied for each of training functions in order to find the optimal training function for this model.

4.4 Experiments to Find Optimal Training Function

To find the optimal training function experiments are performed by using the optimal hidden neurons (5, 7, 9, 10, and 11) from the experiments that are done to find the optimal hidden layer neurons, the optimal divide function (dividerand) from the experiments that are done to find the optimal divide function and, the optimal transfer function (tansig) from the experiments that are done to find the optimal transfer function. Here the performance result of the training functions such as

- ✓ trainlm,
- ✓ trainscg,
- ✓ trainrp,
- ✓ traincgb,
- ✓ traincgp,
- ✓ traincgf, and
- ✓ trainoss are compared.

Table 22 below shows the comparison for trainlm using optimal hidden neurons.

Training function	Number of neurons	Best validation performance(MSE)	Epoch
trainlm	5	0.0051375	51
	7	0.004462	17
	9	0.004474	32
	10	0.0055065	37
	11	0.0052861	66

Table 20 Comparison for trainlm with optimal hidden neurons

As we can see from table 21 above the best validation performance 0.004462 at epoch 17 is the optimal with 7 hidden neurons. Its figure is similar to dividerand because trainlm value also take the default value so its figure is similar to dividerand at hidden neurons 7.

Table 22 below shows the comparison for trainscg using optimal hidden neurons.

Training function	Number of neurons	Best validation performance(MSE)	Epoch
trainscg	5	0.028013	59
	7	0.022742	52
	9	0.017008	139
	10	0.025323	63
	11	0.009578	178

Table 21 Comparison for trainscg with optimal hidden neurons

As we can see from table 22 above the best validation performance 0.009578 at epoch 178 is the optimal with 11 hidden neurons. Figure 18 below shows the best validation performance for trainscg.

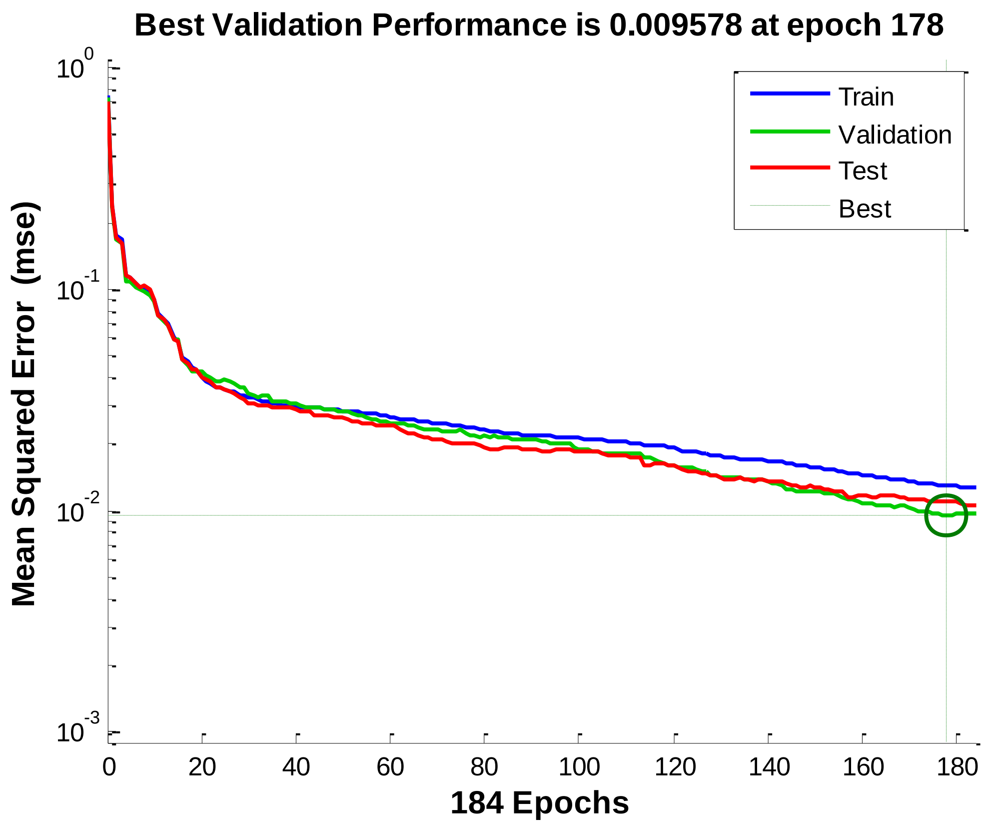

Figure 18 Best validation performance for trainscg with 11 hidden neurons

As we can see from Figure 18 above the best validation performance is 0.009578 at iteration 178.

Table 23 below shows the comparison for trainrp using optimal hidden neurons.

Training function	Number of neurons	Best validation performance(MSE)	Epoch
trainrp	5	0.028523	71
	7	0.017088	2049
	9	0.019618	173
	10	0.014423	179
	11	0.016447	321

Table 22 Comparison for trainrp with optimal hidden neurons

As we can see from table 24 above the best validation performance 0.014423 at epoch 179 is the optimal with 10 hidden neurons. Figure 29 below shows the best validation performance for trainrp.

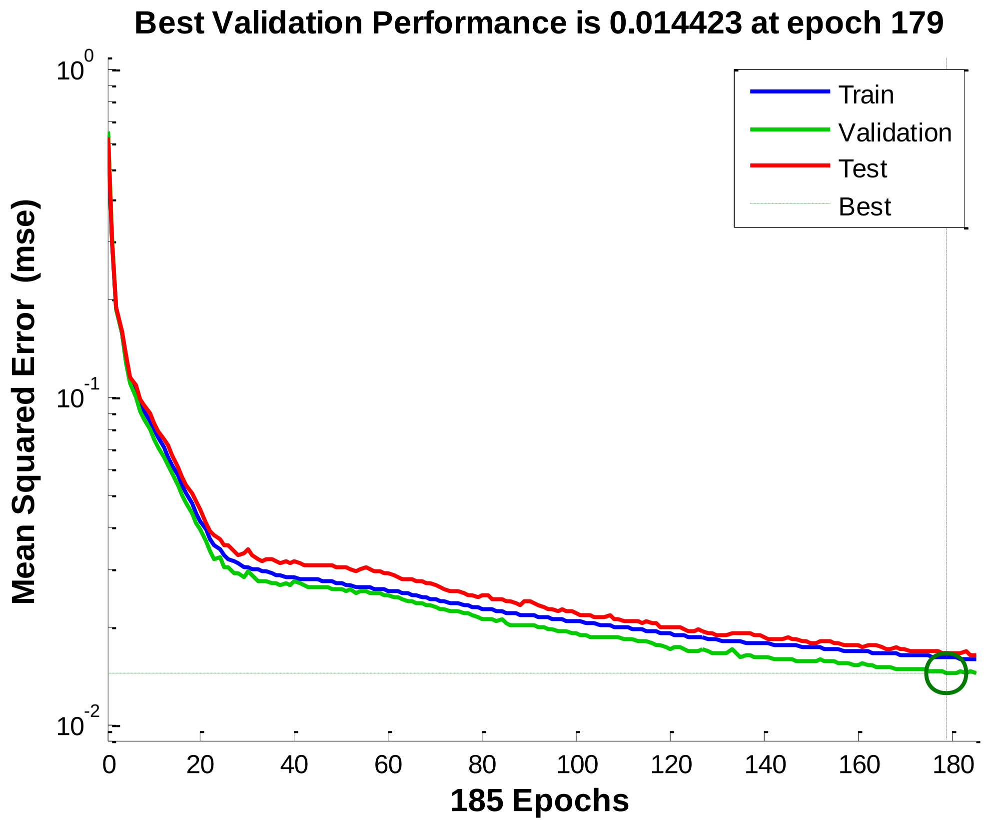

Figure 19 Best validation performance for trainrp with 10 hidden neurons

As we can see from Figure 19 above the best validation performance is 0.014423 at iteration 179.

Table 24 below shows the comparison for traincgb using optimal hidden neurons.

Training function	Number of neurons	Best validation performance(MSE)	Epoch
traincgp	5	0.025069	48
	7	0.030948	31
	9	0.026974	50
	10	0.01692	68
	11	0.027971	54

Table 23 Comparison for traincgp with optimal hidden neurons

As we can see from table 24 above the best validation performance 0.01692 at epoch 68 is the optimal with 10 hidden neurons. Figure 20 below shows the best validation performance for traincgp.

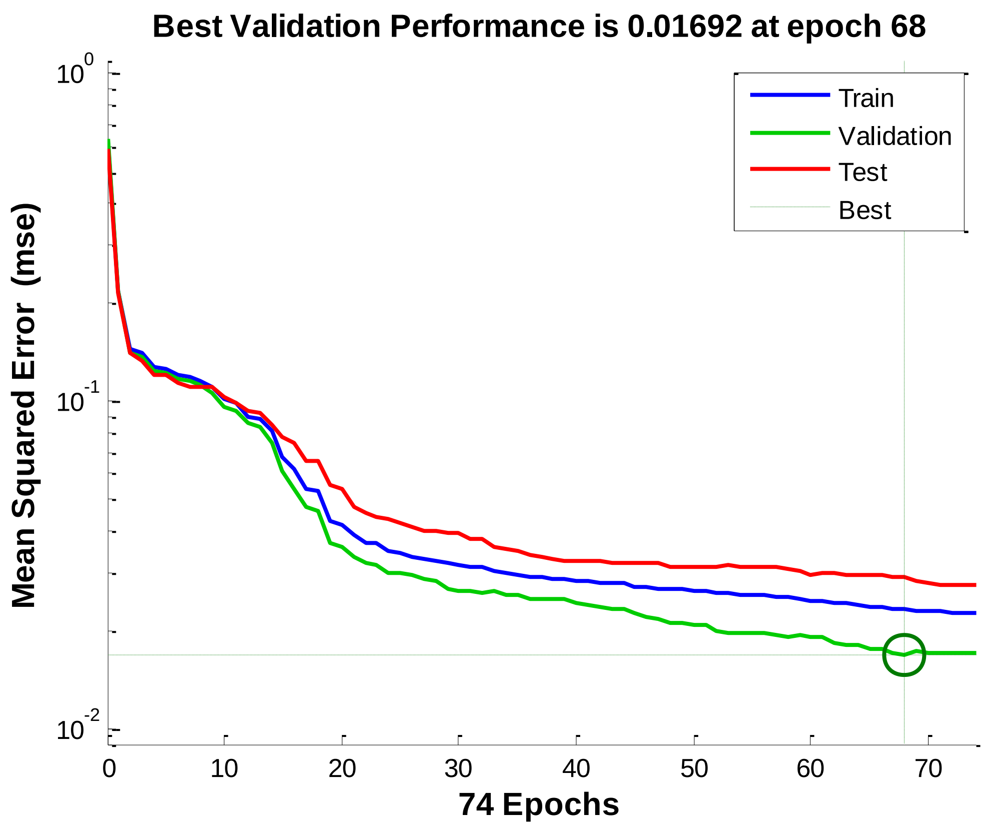

Figure 20 Best validation performance for traincgp with 10 hidden neurons

As we can see from Figure 20 above the best validation performance is 0.01692 at iteration 68. Table 25 below shows the comparison for traincgb using optimal hidden neurons.

Training function	Number of neurons	Best validation performance(MSE)	Epoch
Traincgb	5	0.036154	44
	7	0.028026	69
	9	0.010053	218
	10	0.029845	36

	11	0.022577	105

Table 24 Comparison for traincgb with optimal hidden neurons

As we can see from table 26 above the best validation performance 0.010053 at epoch 54 is the optimal with 9 hidden neurons. Figure 21 below shows the best validation performance for traincgb.

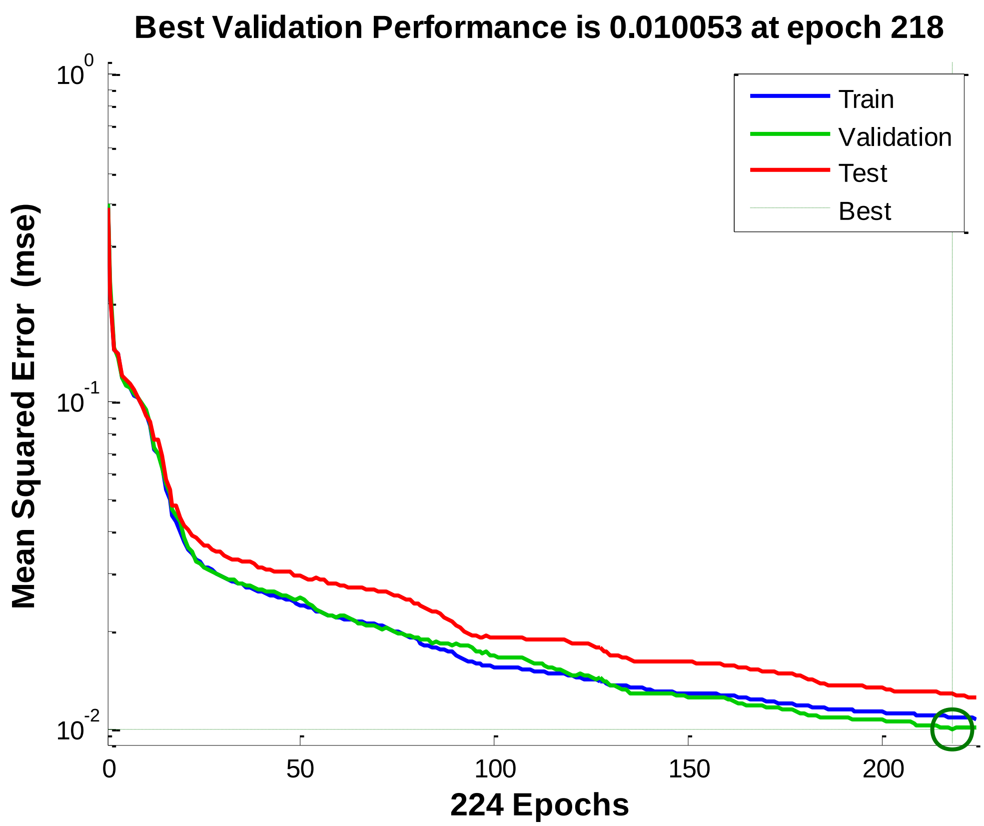

Figure 21 Best validation performance for traincgp with 9 hidden neurons

As we can see from Figure 21 above the best validation performance is 0.010053 at iteration 218. Table 26 below shows the comparison for traincgf using optimal hidden neurons.

Training function	Number of neurons	Best validation performance(MSE)	Epoch
traincgf	5	0.028105	173
	7	0.02084	133
	9	0.027321	139

	10	0.026759	140
	11	0.022075	204

Table 25 Comparison for traincgf with optimal hidden neurons

As we can see from table 26 above the best validation performance 0.02084 at epoch 133 is the optimal with 7 hidden neurons. Figure 22 below shows the best validation performance for traincgf.

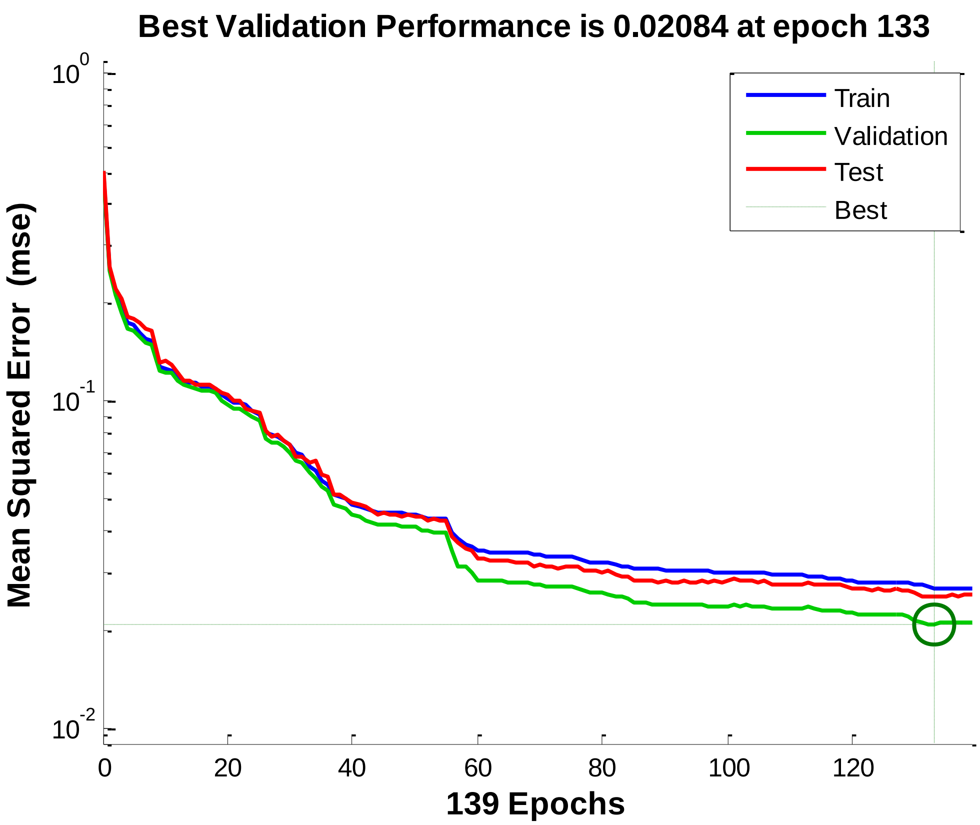

Figure 22 Best validation performance for traincgf with 7 hidden neurons

As we can see from Figure 22 above the best validation performance is 0.02084 at iteration 133.

Table 28 below shows the comparison for trainoss using optimal hidden neurons.

Training function	Number of neurons	Best validation performance(MSE)	Epoch
trainoss	5	0.022526	25
	7	0.057089	27

	9	0.036783	56
	10	0.013767	193
	11	0.030479	54

Table 26 Comparison for trainoss with optimal hidden neurons

As we can see from table 28 above the best validation performance 0.013767 at epoch 193 is the optimal with 10 hidden neurons. Figure 23 below shows the best validation performance for trainoss.

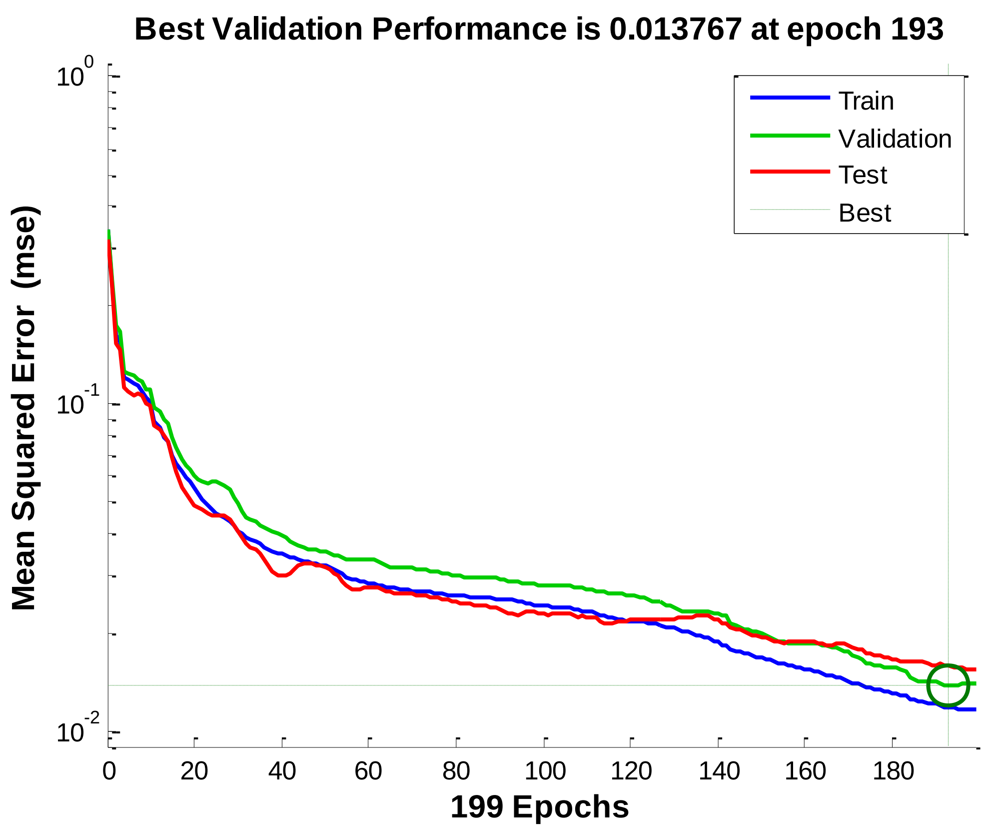

Figure 23 Best validation performance for trainoss with 10 hidden neurons

As we can see from Figure 23 above the best validation performance is 0.013767 at iteration 193.

Table 28 below shows the comparison among the 7 training functions.

Training function	Number of neurons	Best validation performance (MSE)	Epoch
trainlm	7	0.004462	17
trainscg	11	0.009578	178
trainrp	10	0.014423	179
traincgp	10	0.01692	68
traincgb	9	0.010053	218
traincgf	7	0.02084	133
Trainoss	5	0.022526	25

Table 27 Comparison among training functions

As we can see from table 28 above trainlm is the optimal training function with best validation performance of 0.004462 at epoch 17 for the model of this study. Figure 24 below shows the neural network architecture of the selected model which contains training function trainlm, transfer function tansig, divide function dividerand, and 7 neurons in the hidden layer.

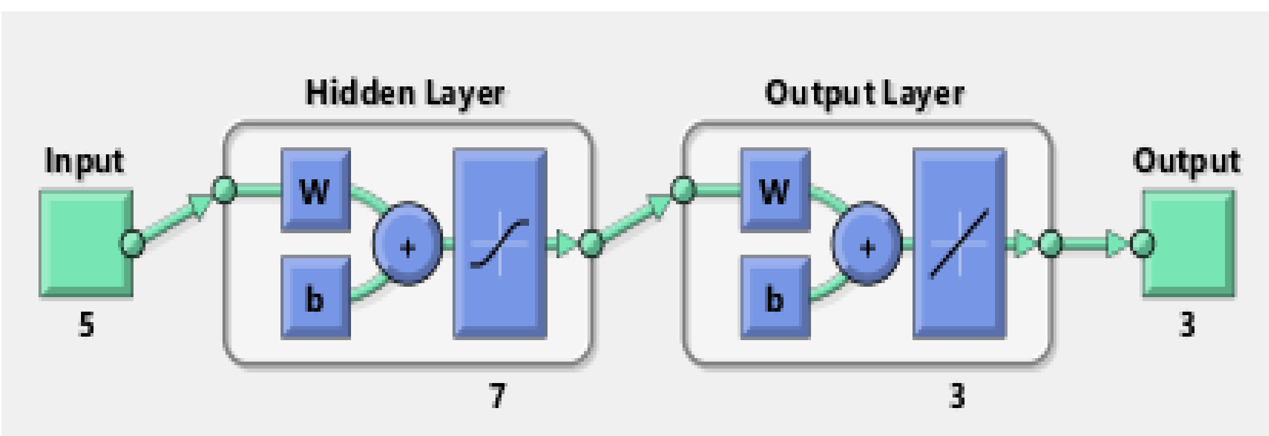

Figure 24 Neural network architecture of the selected model

In addition to MSE we can evaluate the performance of the selected model using confusion matrix.

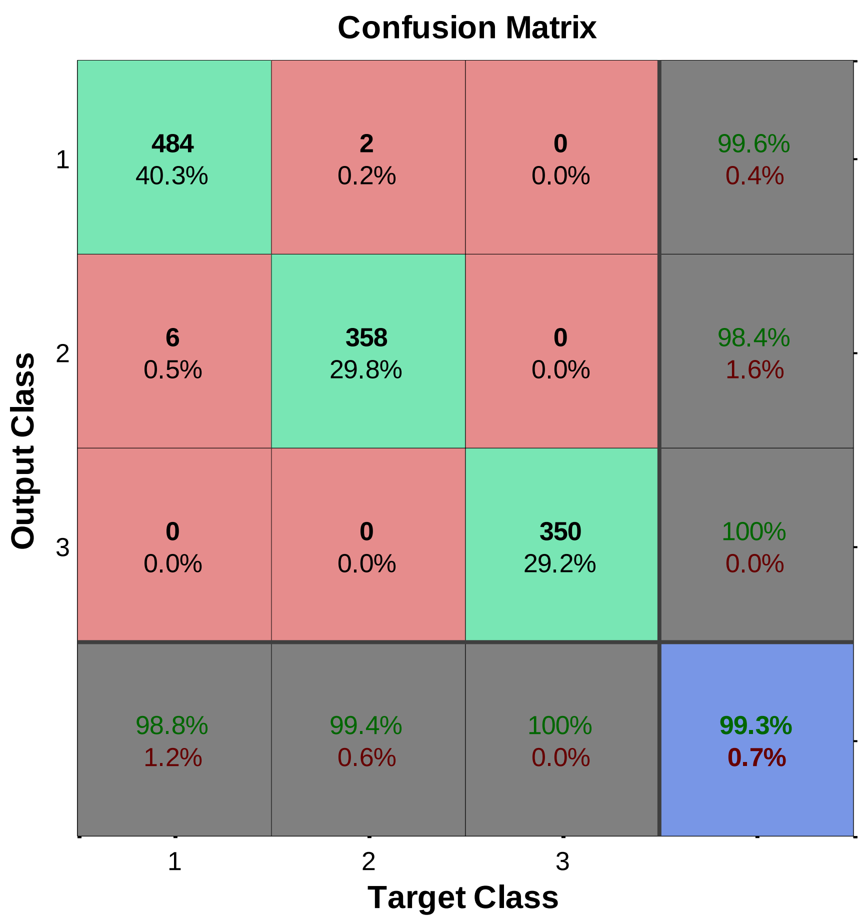

Figure 25 confusion matrix of the selected model.

As we can see from Figure 25 above for testing 484 instances are classified as true low performance but 2 instance is classified as false medium performance and no instances are classified as false high performance. For testing 358 instances are classified as true medium performance but 6 instances that are classified as false low performance and no instances are classified as false high performance. For testing 350 instances are classified as true high performance but there are no instances that are classified as false low performance and no instances are classified as false medium performance. Finally as we can see this model 99.3% approved it is good research result. This study showed the potential of the artificial neural network for predicating the factor that affect the student performance

4.5 Determine Relative Importance of Input Variables

Significance level or rank of input variables is performed using the selected model. Figure 26 below shows the relative importance of input variables.

KEY: 1 Login into the system

2. Scorm view

3. Quiz result

4. Resource view

5. Assignment attempt

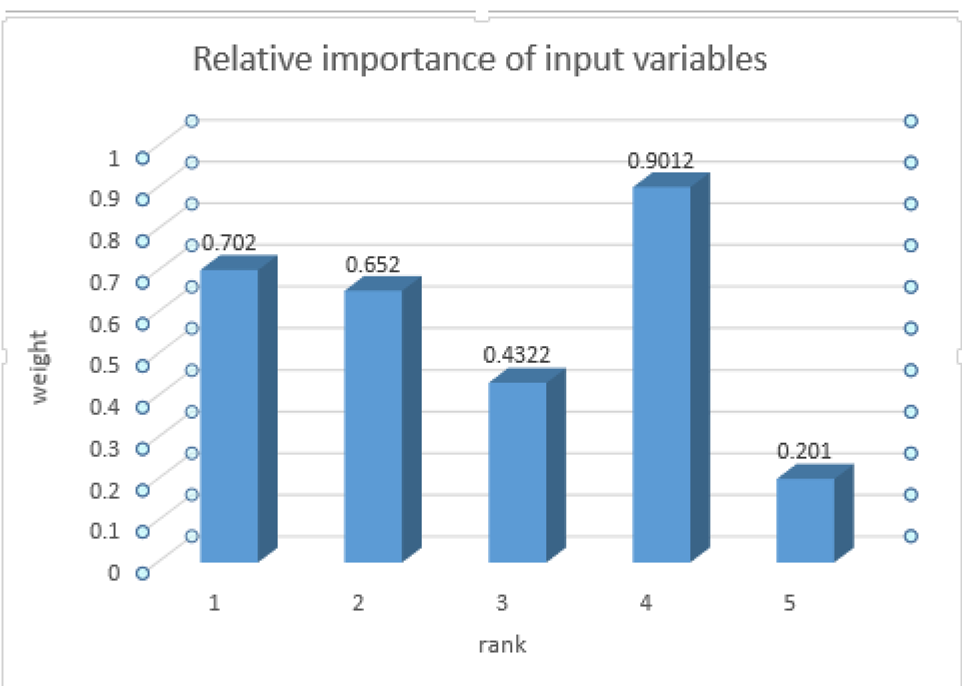

Figure 26 Relative importance of input variables

As we can see from Figure 26 above resource view is the first influential variable with the weight of 0.9012, the second determinant variable is login in to the system with the weight 0.702 , the third determinant variable with the weight of 0.652 is scorm view ,the fourth and the fifth determinant variables are quiz result and assignment attempt with the weight of 0.4322 and 0.201 respectively.

The main goal of this research is identifying the variable that determine the students' performance rate because identification of these variables is subjective for the instructor and it has main role to students' performance. Table 30 below shows the result of the study to suggest the significance levels of input variables. The rank according to the selected model is identified based on a scientific research procedure i.e. weight based identification. This shows the result obtained from this study is trustworthy than domain experts, because there is no weight based identification of input variables by the domain experts and significance of input variables is done based on their experience.

Symptoms	Rank of the selected model	Weight of the selected model
Assignment attempt	5	0.2010
No of login in to the system	2	0.702
Resource View	1	0.9012
Quiz result	4	0.4322
Scorm View	3	0..652

Table 28 Rank of selected model

This study identified the significance levels of the variables such as resource view, scorm view, Quiz result, assignment attempt and No of login in to the system. The Resource view weighted higher than the other variables and No of login in to the system, scorm , Quiz result and assignment are second, third, fourth and fifth rank respectively for this study.

Based on this result the instructor can easily identify the students' performance rate. By doing so, the teacher can help students by preparing additional or supplementary learning materials based on the observed results.

CHAPTER FIVE

CONCLUSION AND FUTUER WORK

5.1 Conclusion

In this thesis, a model is developed to predict the determinant variables that affect the students' performance rate based on artificial neural network. The researcher consider the appropriate neural network architecture that means feed-forward networks and multilayer perceptron model it uses different algorithms such as hidden layer neurons, divide function, transfer function and training function, in order to achieve the optimal result by comparing their performance in terms of MSE to reach the optimal possible answer. By selecting trainlm training function, dividerand divide function, tansig transfer function and one hidden layer with 7 nodes, the model reached an optimal solution. The significance levels of input variables is done using the model which contains optimal hidden layer neurons, divide function, transfer function, and training function. From the input variables resource view is the first influential variable with weight of 0.0912, login in to the system is the second influential variable with weight of 0.702,scorm view is the third influential variable with the weight of 0.652,quiz result is the fourth influential variable with weight of 0.4322 and assignment attempt is the fifth influential variable with weight of 0.2010.

The basic target of this research as it was briefly discussed in the previous section of this study is to identify the variables that determine the students' performance rate in e-learning technology by proposing a better learning analytical model. Moreover, this new system predict the variable that determine students' performance rate based on ANN. Based on this prediction, teachers can easily identify those students who need special support and may also prepare additional support for students who have high performance rate.

Finally as we can see in chapter four this research model was tested and the overall result was 99.3% approved it is good research result.

5.2 Future Work or Recommendations

Based on the results obtained from the study, the researcher makes the following recommendations. Other researchers can extend this research by taking the following future research directions

- One can develop a prototype as a senior project by using the optimal model.
- One can develop a prototype that give some notification to the instructor this notification give some idea to the instructor which student is need special help based on the output of this research.
- Future studies can have their directions set in determining an accurate, reliable prediction network for helping instructors and students in decision-making to increase student performance. Although limited data sets were used in this study, future studies might continue to collect information on students' grades and more profile about the student from different university and different course.

References

[1] Berhe H., "The challenges and current status of e-learning in Ethiopian higher educationinstitutions: case of Mekelle University," 2011.

[2] J. A. L.-L. F. &. J. Daradoumis T., "Using Collaboration Strategies to Support the Monitoring of Online Collaborative Learning Activity.," 2010.

[3] Brandon K., "Investing in Education: The American Graduation Initiative. Office OfSocial Innovation and Civic Participation.," 2009.

[4] P. A. Linan L., "Educational Data Mining and Learning Analytics: Open University of Catalonia (UOC)," 2015.

[5] Rashty D., "Traditional Learning vs. Elearning," 2015. [Online]. Available: http://www.researchtrail.com. [Accessed May 2016].

[6] Cheng Z., "Introduction to E-Learninng," 2016. [Online]. Available: http://www.chengzhi.net/english/. [Accessed may 2016].

[7] Tamsamani Y., "Elearning versus Traditional Classroom," 2012. [Online]. Available: http://www.morocco. [Accessed June 2016].

[8] Papanis E., "Traditional Teaching versus e-learnning," 2015. [Online]. Available: http://www.academic.edu/958026. [Accessed May 2016].

[9] L. P. Siemens G., "Penetrating the fog: Analytics in learning and education," pp. 31-40, 2011.

[10] B. J. A. Krenker A., "Introduction to the Artiificial Neural Networks," 2011.

[11] Z. J., "Introduction to Artificial Neural Network (ANN) Methods:what thy are and how to use them.," 2014.

[12] Fiesler E., Arteficial Neural Network, 2014.

[13] Cuza E., "Multilayer Perceptron and Neural Networks," vol. 8, no. 7, 2009.

[14] Riedmiller M., "Machine Learning: Multi-Layer Perceptrons," p. 53/61, 2015.

[15] Rupp M. etal, "Comparison Between Multilayer Feedforward Neural Networks and a Radial Basis Function Network to Detect and Locate Leaks in Pipelines Transporting Gas," vol. 32, 2013.

[16] Stamos T.&Andreas V., "An Artificial Neural Network for Predicting Student Graduation Outcomes," 2008.

[17] Naser A. etal, "Predicting Student Performance Using Artificial Neural Network: in the Faculty of Engineering and Information Technology," vol. 8, pp. 221-228, 2015.

[18] Naser A., "predicting learners performance using artificial neural networks in linear programming intelligent tutoring system," *International Journal of Artificial Intelligence & Applications,* vol. 3, 2012.

[19] Sultan A. etal, "A PROTOTYPE DECISION SUPPORT SYSTEM FOR OPTIMIZING THE EFFECTIVENESS OF E-LEARNING IN EDUCATIONAL INSTITUTIONS," *International Journal of Data Mining & Knowledge Management Process,* vol. 1, 2011.

[20] Bahadir E., "Using Neural Network and Logistic Regression Analysis to Predict Prospective Mathematics Teachers' Academic Success upon Entering Graduate Education," 2016.

[21] Czerkawsk B., "When Learning Analytics Meets E-Learning," 2014.

[22] Uhler B.&Hurn J., "Using Learning Analytics to Predict Student Success: A Faculty Perspective," 2013.

[23] Faraway J.&Chatfiel C., "Time series forecasting with neural networks: A comparative study using the airline data," vol. 47, no. 2, pp. 231-250, 1998.

[24] Shahin M. etal, "State of the Art of Artificial Neural Networks in Geotechnical Engineering," *Journal of Geotechnical Engineering,* 2008.

[25] Zhang G. etal, "Forecasting with artificial neural networks: The state of the art," *International Journal of Forecasting,* vol. 14, pp. 35-62, 1997.

[26] Cybenko G., "Approximation by super positions of a sigmoidal function." Mathematics of control," vol. 3, pp. 303-314, 1989.

[27] Hornik etal, "Multilayer feed-forward networks are universal approximators," vol. 2, pp. 359-366, 1989.

[28] Beale m. etal, Neural Network Toolbox™ User's Guide, 2006.

[29] Demuth H. and Beale M., "Neural Network Toolbox," vol. 4, 2001.

[30] Sharma B. and Prof. Venugopalan K., "Comparison of Neural Network Training Functions for Hematoma Classification in Brain CT Images," *IOSR Journal of Computer Engineering ,* vol. 16, no. 1, pp. ISNN: 2278-8727, 2014.

[31] Gharehchopogh S. etal, "using artificial neural network in diagnosis of thyroid disease: a case study," *International Journal on Computational Sciences & Applications (IJCSA),* vol. 14, no. 4, 2013.

Appendix

Sample Code

The prototype of the proposed architecture was developed and the sample MATLAB code includes:

//MATLAB code for train number of neurons

Number of Nueros

```
X=[];

Y=[];

I=X';

O=Y';

NW1=newff(I,O,3);

NW1=train(NW1,I,O);

NW2=newff(I,O,4);

NW2=train(NW2,I,O);
```

//MATLAB code for divide function

//MATLAB code for divideint

```
NW10=newff(I,O,5);

NW10.dividefcn='divideint';

[NW10,tr1]= train(NW10,I,O);
```

//MATLAB code for divideblock

```
NW15=newff(I,O,5);

NW15.dividefcn='divideblock';

[NW15,tr6]= train(NW15,I,O);
```

//MATLAB code for divideind

```
NW20=newff(I,O,5);

NW20.dividefcn='divideind';

[NW20,tr11]= train(NW20,I,O);
```

//The MATLAB code for dividerand in divide function is default.so it take the first code that means train neurons.

//MATLAB code for transfer function

//MATLAB code for logsig by using best of best from divide function are "dividerand"

NW50=newff(I,O,5);

NW50.layers{1}.transferfcn='logsig';

[NW50,tr40]= train(NW50,I,O);

//MATLAB code for purelin using best of best from divide function are "dividerand"

NW55=newff(I,O,5);

NW55.layers{1}.transferfcn='purelin';

[NW55,tr45]= train(NW55,I,O);

//The MATLAB code for tansig in transfer function is default.so it take the first code that means train neurons as dividerand.

//MATLAB code for train function

//MATLAB code for trainscg by using best of best from divide and transfer function so I take "dividerand" and "logsig".

NW60=newff(I,O,5);

NW60.layers{1}.transferfcn='logsig';

NW60.trainfcn='trainscg';

[NW60,tr50]= train(NW60,I,O);

//MATLAB code for trainrp by using best of best from divide and transfer function so I take "dividerand" and "logsig".

```
NW65=newff(I,O,5);

NW65.layers{1}.transferfcn='logsig';

NW65.trainfcn='trainrp';

[NW65,tr55]= train(NW65,I,O);
```

//MATLAB code for traincgp by using best of best from divide and transfer function so I take "dividerand" and "logsig".

```
NW70=newff(I,O,5);

NW70.layers{1}.transferfcn='logsig';

NW70.trainfcn='traincgp';

[NW70,tr60]= train(NW70,I,O);
```

//MATLAB code for traincgb by using best of best from divide and transfer function so I take "dividerand" and "logsig".

```
NW75=newff(I,O,5);

NW75.layers{1}.transferfcn='logsig';

NW75.trainfcn='traincgb';
```

```
[NW75,tr65]= train(NW75,I,O);
```

//MATLAB code for traincgf by using best of best from divide and transfer function so I take "dividerand" and "logsig".

```
NW80=newff(I,O,5);

NW80.layers{1}.transferfcn='logsig';

NW80.trainfcn='traincgf';

[NW80,tr80]= train(NW80,I,O);
```

//MATLAB code for trainoss by it can take the default value dividerand and logsig..

// MATLAB code for Trainlm-default in train function is default.so it take the first code that means train neurons.

```
NW85=newff(I,O,5);

NW85.layers{1}.transferfcn='logsig';

NW85.trainfcn='trainoss';

[NW85,tr85]= train(NW85,I,O);
```

// MATLAB code to generate the Architecture in neuron 7 because it’s the best optimal value.

```
view(NW5);
```

```
simpleclassOutputs = sim(NW5,I);

plotconfusion(O,simpleclassOutputs);//to design the confiusion matrix by using best of best hiden layer network value
```

// MATLAB code to generate rank

```
[rank,weight]= relieff(I,O,1200);
bar(weight(rank));
bar(rank);
rank
weight
```