FEBRUARY 2021 M.Sc. in Electronics and Electrical Engineering MUHANNAD ABDULRAHMAN

REPUBLIC OF TURKEY

GAZİANTEP UNIVERSITY

GRADUATE SCHOOL OF NATURAL & APPLIED SCIENCES

DESIGN AND SIMULATION OF DISTANCE PROTECTION RELAY FOR HIGH VOLTAGE TRANSMISSION LINE USING ARTIFICIAL NEURAL NETWORKS

M.Sc. THESIS

IN

ELECTRONICS AND ELECTRICAL ENGINEERING

BY

MUHANNAD ABDULRAHMAN

FEBRUARY 2021

DESIGN AND SIMULATION OF DISTANCE PROTECTION RELAY FOR HIGH VOLTAGE TRANSMISSION LINE USING ARTIFICIAL NEURAL NETWORKS

M.Sc. Thesis

in

Electronic and Electrical Engineering
Gaziantep University

Supervisor

Prof. Dr. Arif NACAROĞLU

by

Muhannad ABDULRAHMAN

February 2021

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

Muhannad ABDULRAHMAN

ABSTRACT

DESIGN AND SIMULATION OF DISTANCE PROTECTION RELAY FOR HIGH VOLTAGE TRANSMISSION LINE USING ARTIFICIAL NEURAL NETWORKS

ABDULRAHMAN, Muhannad
M.Sc. in Electronics and Electrical Engineering
Supervisor: Prof. Dr. Arif NACAROĞLU
February 2021
117 pages

This thesis presents an approach to design a distance protection relay that performs the function of detection, classification, and localization of the fault on high voltage transmission line. The method is based on combination of discrete wavelet transform as a technique for processing signal and artificial neural networks using the extracted features from current signals of fault by discrete wavelet transform as inputs to artificial neural networks that will be trained by using back-propagation algorithm to differentiate between non faulty and faulty cases in addition to, classification of fault type, and localization of the fault for power transmission line. Matlab/Simulink is utilized for designing the model of power transmission line and simulating the different types of faults for various fault resistance and distance as well as design, training, and testing of neural networks for detector, classifier, and locator.

Key Words: Transmission Line, Distance Protection Relay, Wavelet Transform, Artificial Neural Networks, Detection, Classification, Localization.

ÖZET

YÜKSEK GERİLİM HATLARINDA YAPAY SİNİR AĞLARI KULLANILARAK UZAKTAN ÖNLEME ANAHTARLARININ BENZETİMİ VE TASARIMI.

ABDULRAHMAN, Muhannad
Yüksek Lisans Tezi Elektrik ve Elektronik Mühendisliği
Danışman: Prof. Dr. Arif NACAROĞLU
Şubat 2021
117 Sayfa

Bu tezde, yüksek gerilim hatlarında meydana gelebilecek hataların algılanmasını, yerinin ve tipinin belirlenmesi uzaktan gerçekleştirecek bir tasarım gerçekleştirilmiştir. Çalışma temel olarak "Dalgacık Dönüşümü" ve "Yapay Zeka" yöntemlerinin birleştirilmesini esas almaktadır. İletim hatlarında oluşacak kısa devre veya açık devre hatalarının hattın herhangi bir noktasında oluşturacağı dalga şekli değişikliklerinin karakteristiği dalga dönüşümü yöntemi ile çıkartılmış ve bu bilgi yapay sinir ağı sistemine yüklenerek sistemin hata farklarını öğrenmesi sağlanmıştır. Güç iletim hatlarının benzetimi MATLAB/SIMULING gibi yazılımlar kullanılarak yapılmış ve değişik hata tipleri, değişik iletim hattı karakteristikleri gibi parametreler değiştirilerek hattın farklı uzaklıklarındaki hataların tespit edilmesi sağlanmıştır. Örnek olarak alınan iletim hattı farklı eşdeğer dirençlerinde benzetilerek olası değişik senaryolar çalışılmış ve çok az hatalar ile hata yeri tespiti sağlanmıştır.

Anahtar Kelimeler: Yüksek Gerilim 3-faz Güç İletim Sistemleri, Dalgacık Dönüşümü, Yapay Sinir Ağları, Hata Mesafe Tespiti, Sınıflandırılması.

''Dedicated to my family''

ACKNOWLEDGEMENTS

I would like to thank my supervisor, Prof. Dr. ARIF NACAROĞLU for his guidance and support throughout the study. I am thankful for his encouragement and motivation.

I would like to express my love and gratitude to my family for their support, always best wishes.

TABLE OF CONTENTS

Page

LIST OF TABLES

LIST OF FIGURES

Page

LIST OF SYMBOLS

$\boldsymbol{\alpha}$	Learning Rate
$\boldsymbol{\Omega}$	Ohm
$\boldsymbol{\psi}$	Wavelet Function
$\boldsymbol{a}$	Scaling
$\boldsymbol{b}$	Translation
$\bar{\boldsymbol{\mu}}$	Mean Value

LIST OF ABBREVIATIONS

VT	Voltage Transformer
CT	Current Transformer
HV	High Voltage
PS	Power System
EHV	Extra High Voltage
TL	Transmission Line
HVTL	High Voltage Transmission Line
FT	Fourier Transform
WT	Wavelet Transform
CWT	Continuous Wavelet Transform
DWT	Discrete Wavelet Transform
WTC	Wavelet Transform Coefficients
D	Details
A	Approximations
AI	Artificial Intelligence
ANN	Artificial Neural Network
BPNN	Backpropagation Neural Network
MLNN	Multi-Layer Neural Network
TF	Transfer Function
MSE	Mean Squared Error
SVM	Support Vector Machine
FD	Fault Detector
FC	Fault Classifier
FL	Fault Locator
SD	Standard Deviation

CHAPTER I

INTRODUCTION

1.1 Motivation of Thesis

The major motivation behind this thesis is the considerable impact of accurate detection and localization of the faults on reliable operation of power transmission system, in addition to the huge economic consequences, in terms of the amount of money and time that can be saved in case detecting the fault and isolating the faulty section of PS as soon as possible and pinpointing the location of fault on path of TL that are usually used to transmit power over a very long distances and secure the electric power for large area within the same country as well as it is used to transmit power from one country to the other. This thesis concentrates on reading and analysis of three-phase current signals of the TL at one end (at substation) during various fault conditions and utilizing the outputs of current signals analysis in the designing the distance protection relay performs the main functions effectively and accurately. The main objective of this thesis is to study and design a distance protection relay that can detect the fault, classify the fault type, and pinpoint the distance of fault in power transmission lines.

1.2 Literature Review

Discrete Wavelet Transform (DWT) has been applied to many studies for PS protection [1-4] and many Artificial Intelligence Techniques (AIT) have been employed for detection, classification, and localization of power TL, such as Fuzzy Logic Approach (FLA), Support Vector Machines (SVM), and Artificial Neural Network (ANN) [5-8].

The three techniques, namely wavelet, ANN, and fuzzy logic method, may be integrated to obtain hybrid techniques help to address disadvantages in applying one approach of which. The combined techniques can be categorized into [9]:

- Neuro-Fuzzy technique
- Wavelet-ANN technique
- Wavelet-Fuzzy-logic technique
- Wavelet-Neuro-Fuzzy technique

Many studies have employed the hybrid techniques for application of power TL protection as described below:

The study presented in [10] introduces a hybrid method based on DWT and ANN for fault detector, classifier, and locator for six phases lines utilizing single-end measured data only. The SD of voltage and current signals approximate coefficients acquired by DWT are utilized as inputs to the ANN for fault classification and pinpointing.

The paper [11] has proposed a technique of using (DWT) and (BPNN) based on Clarke's transformation for FC and FD. Simulation and training process for the NN are achieved by using PSCAD / EMTDC and MATLAB. Daubechies4 mother wavelet (DB4) is applied for fault signals decomposition. The WTC and wavelet energy coefficients (WEC) used as input to ANN for training using back-propagation (BPNN) method. The simulation also demonstrates that the proposed algorithm is more reliable and precise.

The application given in [12] utilizes high frequency information of fault currents at one end of a protected TL to classify transients on the TL. The feature vector of the fault signals can be determined from current signals without applying any signals related to voltages. DWT is employed to extract high frequency components. A feature vector is assembled using the wavelets details coefficients of one level of the aerial modes and utilized to train the ANN. Findings of classification were accurate.

The application of MR wavelet analysis combined with ANN for accurate classification was studied in [13]. The method utilizes energy of approximation coefficients for FC. DWT is applied to get approximation coefficients for fault current signals and thus classification of faults was precise. This work dealt with FC, this paper does not study localization of the fault distance. The proposed system and analysis were carried out in Matlab Simulink environment.

FD & FC in a long TL using ANN & DWT has suggested in [14]. The proposed method applies one cycle pre fault and one cycle post fault for the three-phase current signals to obtain the neutral current signal. Daubechies as the mother wavelet is employed while utilizing the DWT technique. The differential energy, is utilized as inputs for system, constructed for the FC. The features that are the energies gained from DWT of the current signals nominated are supplied to ANN for the purpose of FC. The findings showed that the proposed system may appropriately classify every potential fault with significant changes in system conditions.

A hybrid technique that is (DWT) and (SVM) for FC in the TL was studied in [15]. The DWT was employed to obtain the detailed coefficients of transient signal D8 and D9 using 50 kHz for sampling. The value of Root Mean Square (RMS) will be determined by the coefficients D8 and D9 that will be applied for training and testing ANN.

[16] has proposed a method to localize the zone of transient faults and to classify the fault type in Power distribution network utilizing DWT and Adaptive Neuro-Fuzzy Inference Systems (ANFIS). This method utilizes DWT as tool for signal processing and is applied on data sampled current signals to obtain the important attributes and features of the fault signals. Studies and simulations were implemented in an EMTP-RV environment. The simulation results demonstrated high accuracy in classifying the fault type and defining the fault zone, so that the maximum detected error was less than 2 %.

A hybrid approach based on based on DWT and ANNs have been applied to identify, classify, and localize several types of three-phase PS faults in [17]. The features obtained by DWT for voltage or current signals have been used as inputs to ANN. Simulation findings indicate that this methodology can effectively distinguish different faults in a typical three-phase PS.

1.3 Objective of Master Thesis

The prime objective of this Master Thesis is to study, successfully design, and test of performance a distance protection system that may detect, determine the fault type, and find the fault place in high voltage power TL by combining the DWT & ANNs techniques.

1.4 Research Methodology

The research considered in this Master Thesis was performed in several stages:

- Modelling of the high voltage power TL
- Reading and sampling of three-phase current signals
- Analyzing of sampled current signals by using DWT and extracting the features from wavelet coefficients.
- Applying the extracted features as input to the ANNs used to configure the distance protection relay that performs functions of detection, classification, and localization of the fault.
- Performance evaluation and testing the distance protection relay functions in (detection, classification, and localization) of the faults and proposal of recommendations for improvement of distance protection reliability and performance.

CHAPTER II

ELECTRICAL POWER SYSTEM

2.1 Introduction

The modern industry is highly dependent on consistent and stable electricity availability and a high efficiency. Most recent technologies do not operate without an extremely stable supply of energy such as communication networks, railway networks, and banking. In addition to this, in the event that the power supply is interrupted, the big number of domestic electricity consumers life is disrupted. Therefore, it is difficult to overemphasize the value of ensuring a constant supply of energy around the clock. There can be no power system built in such a way that it can never malfunction. So, you have to deal with those breakdowns that are considered faults in the language of electrical protection engineering. It is very important to know how the faults are avoided and how the effects of the faults are mitigated. The negative consequences of the faults are reduced by rapidly isolating the faulted components from the rest of the stable PS, limiting the disturbance impacts to a minimum in terms of time and place [18].

Electrical energy is the most common source of energy since it can be efficiently transmitted and distributed at a high efficiency and an affordable cost [19].

An electric PS is a group of electrical components and equipment used to generate, transfer, and distribute the electrical power [20]. It basically consists of generators, step-up power transformers, transmission lines, step-down power transformers, circuit breakers, measurement equipment (current measurement transformers and voltage measurement transformers), control system, and protection system.

The power systems today are expanding continuously and rapidly due to the requirements of the modern life and getting more complex.

The most important component that affect reliability of the power system is transmission line where many of faults occurs unexpected on the TL. The key requirement of the PS is reliability of operation that may be achieved by detecting, classifying, and isolating fault section as fast as possible.

It is well recognized that the electricity transmission networks run at HV level (in a kV range), this level of voltage is used for reducing the losses when the power is transferred across long distances; accordingly, when a fault happens, the extremely high currents flux through the power grid, which may affect equipment and devices severely. One fault can affect millions of customers that are supplied with electricity from the grid even if it lasts only a small part of a second, leading to enormous losses and breakdown of production processes in many industries [17].

Due to occurrences of unexpected faults in HVTL, which leads to equipment and devices deterioration, the reliable operation of PS is influenced by that seriously, in order to maintain the stable and safe functioning of the PS, it is necessary to execute an efficient and operative protection system that is accomplished through an progressing strategy of FD & FC that boosts the protection system of HVTL and improves its performance noticeably [21].

2.2 Electric Power System Structure

Figure 2.1. shows the standard model of the electrical power system divided into four main sections as described below.

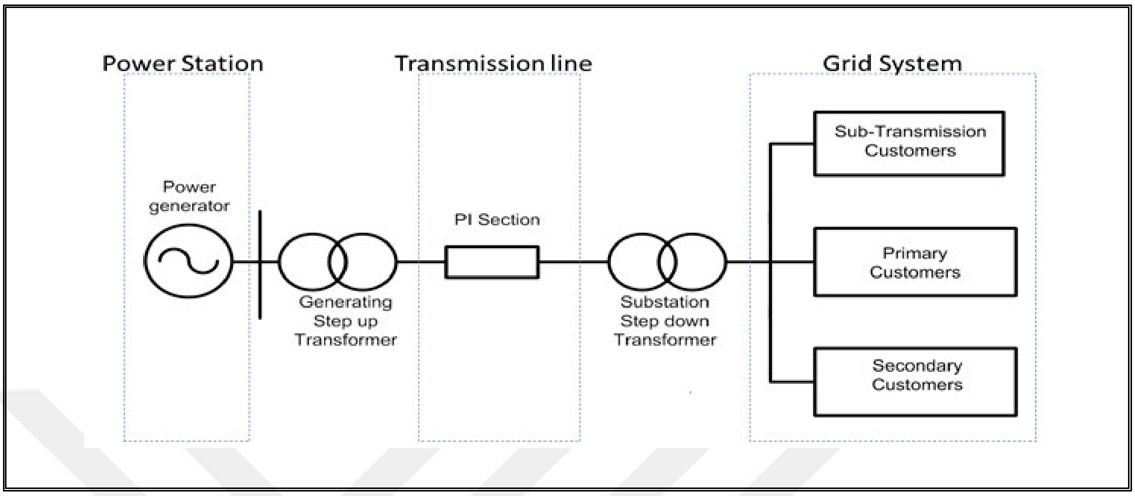


Figure 2.1 Typical model of PS

2.2.1 Generation System

Electricity is generated by transforming mechanical energy into electrical energy. In most situations, mechanical energy is either derived from thermal power plants or from hydropower plant. The key sources of thermal energy are oil, natural gas, nuclear power, and coal. The use of renewable energy such as wind, solar, and biogas in the production of electricity is also growing. Hydropower is the primary non-thermal source of mechanical energy used in the production of electricity. The transfer of mechanical to electrical energy is carried out using synchronous generation units in most power plants. Only Some systems of wind energy use induction generators. The power is normally produced at medium voltage from 11 to 35 kV and is then supplied by a step-up transformer to the transmission grid [20].

2.2.2 Transmission and Sub-Transmission System

The aim of an overhead transmission grid is to transmit electrical energy from power generation plants at different locations into the distribution network that supplies the various loads. The 69 kV, 115 k V, 138 kV, 161 kV, 230 kV, 345 kV, 500KV and 765 kV line-to-line are used as standard transmission voltages, Extra High Voltage (EHV) is normally known for voltage that is above 230 kV [19].

Electrical energy is produced at the generation stations and then transmitted through transformer substations and transmission lines over long distances to supply the points of load through distribution network. All generation stations are interconnected through substations and transmission lines that is considered spine of the power system.

The transmission lines operate at the high voltage levels, normally above of 275 kV, as the loss of power in a transmission line relates to the square of the line current ($|P_{loss} = R_{TL} * I^2$). Typically, the transmission network has an intertwined structure to provide alternative paths for flowing power from the generation stations to the loads. This would increase the reliability and stability of the PS.

At substations, as observed in Figure 2.2, the HV transmission lines are terminated. For industrial facilities that have large loads capacity, the electrical power may be supplied directly from these substations. The main purpose of step-down substations is to reduce voltage level and supply it into the sub-transmission system. The high voltage substations are connected to distribution substations by the step-down transformers. The voltage ranges for sub-transmission system are generally, between 66kV and 132kV [20].

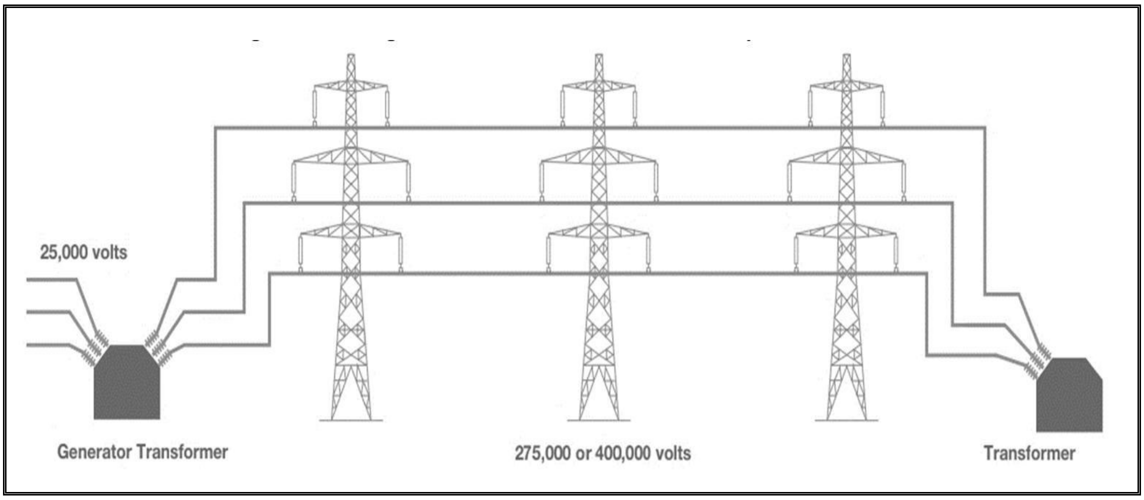


Figure 2.2 Demonstrative example of transmission system

2.2.3 Distribution System

As illustrated in Figure 2.3, distribution process is the final stage in the transmission of electrical power and delivery it to the customers. In general, the connection of the distribution network is in a radial model. The prime distribution voltage is usually between 11kV and 33kV. At this voltage level, some of industrial facilities are provided with main feeders. The secondary feeders provide 415/240V voltage for domestic and commercial use. Small generators installed along the loading points are normally directly connected to the sub-transmission system or distribution network [20].

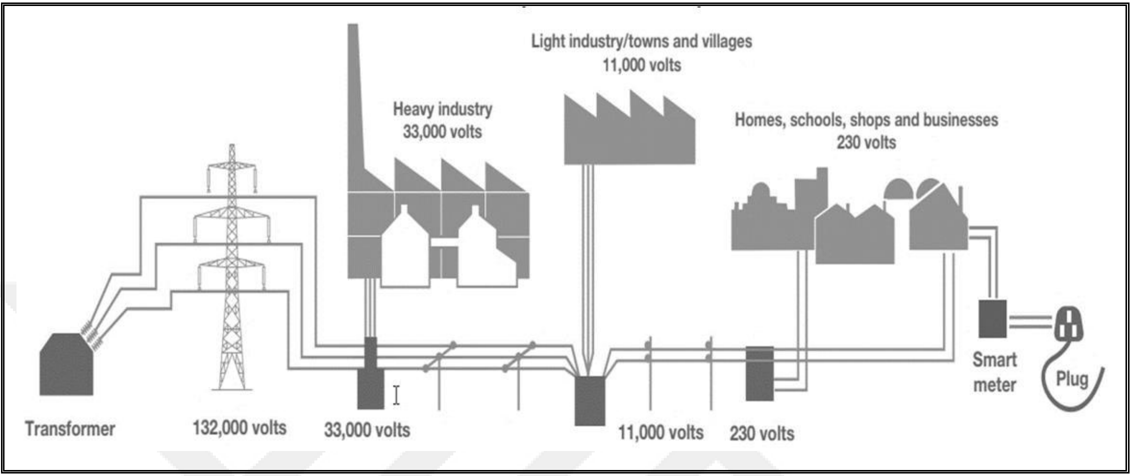


Figure 2.3 Typical structure of distribution system

2.2.4 Loads

The last part of the electrical PS is the load that delivers the power from the distribution network and consumes it as desired. In general, the loads are divided into residential, commercial, and industrial. Depending on the nominal capacity of the loads and their importance, can be fed directly either from transmission network or sub-transmission network, or main distribution network.

2.3 System Power Faults

Since a power system comprises of several complicated interactive components, faults continuously have the probability of occurrence. A fault is not preventable, but by using a proper protection, impacts resulting from faults may be reduced to a minimum.

The eventuality of faults on different components of the electric system is distinctive. The power lines, which are subjected to weather changes and different external conditions, have the most possibility to be faulted. As for internally installed equipment, there is less possibility to fail. The fault statistics are viewed in Table 2.1 [18].

Table 2.1 Fault statistics related to PS elements

Power System Element	Probability of faults (%)
Overhead lines	50 %
Underground cables	9 %
Transformers	10 %
Generators	7 %
Switchgear	12 %
CT, PT relays, control equipment, etc	12 %
Total	100 %

Faults of overhead power TLs are mainly divided into two types:

- **Open Circuit Faults (Series Faults):** which take place due to the failure of one or more phases and are characterized by increasing of voltage and frequency and decreasing of current for the faulted phases.
- **Short Circuit Faults (Shunt Faults):** which take place due to the failure of insulation between phases or/and between phases and ground and are characterized by increasing of current and decreasing of voltage and frequency. Short Circuit Faults are also categorized into two types:
  - **Symmetric Faults (balanced):** which occurs when all phases are shorted out at the same time.
  - **Unsymmetric Faults (unbalanced):** which occurs when one phase is shorted out to ground or when two phases are shorted out together or to ground.

The faults, which are between one of the phases and ground, are named ground faults. Faults that are between two or more phases, with or without ground, are known phase faults. For over decades, power facilities have been in service. Accumulated practice indicates that not all faults have the same possibility. The most likely fault is single line to ground fault (L-G), while the least likely fault related to the short circuit among all three phases, defined as a three-phase fault (L-L-L). This is clarified in Table 2.2 [18].

Table 2.2 Fault statistics related to fault type

Fault Type	Probability of Occurrence (%)	Severity
L-G	85 %	Least severe
L-L	8 %	
L-L-G	5 %	
L-L-L	2 %	Most severe
Total	100 %	

2.4 Power System Protection

PS protection is the system used to protect the power system against faults which may arise under critical situations with the aim of minimizing the impact of these faults on the components of electric system and consumers, therefore the protection system must keep the electric system reliable, continuous, and effective. So, certain requirements (such as below) should be applied for the construction of this system to minimize the impact of the faults (internal and external) to a minimum and maintain the power system properly operating under and after the fault happens [18].

Relays discover and localize faults by measuring of current and voltage values that are distinct under ordinary and fault conditions. The most critical function of protection relays is to protect individuals firstly, and to protect facilities equipment secondly by reducing the destruction and cost caused by insulation breakdowns which (above overloads) are referred to as 'faults'. These faults can arise because of insulation degradation or unexpected incidents, such as lighting strikes thunder, gale-force winds, or collision with trees [22].

2.4.1 Main Requirements of Protection System

Reliability: is the possibility that the system will operate perfectly when needed to take action. This reliability has two aspects: first, the system must operate in the presence of a fault that is inside its protection zone and second, it must abstain from operating unnecessarily for faults outside its protecting zone or in the absence of a fault.

Security: is the capacity of the system to avoid unnecessary function when the Power System operates under normal conditions.

Sensitivity: is the capacity of the system to identify an irregular situation that surpasses a nominal "pickup" or detection limit value, and which commences protective action when the sensed quantities surpass that limit.
Speed: is the capacity of system to operate at the required speed, because the longer the fault lasts for a longer period on the system, the greater the damage will be, and the system loses its stability [23].

2.4.2 Main Components of Protection System

- **Measurement Equipment:** that include current and voltage transformers. The main purpose of these equipment is to step down the current and voltage signals to be quantifiable values, within the operation range of electronic protection systems (1A -5A) for current transformers (CTs), and (110V-100V) for voltage transformers (VTs). Hence, protection equipment inputs can read and process these values.
- **Protection relay:** is electronic and digital device which receive measured signals from the secondary side of CTs and VTs, analysis these signals, based on results of analysis, a trip signal is sent by protection relays to the circuit breakers to disconnect the faulty section from PS if necessary.
- **Circuit Breaker:** Circuit Breaker acts upon trip signals sent by protective relays when faults are detected and close orders when faults are cleared. They can also be manually operated (opening or closing), to isolate a component of power system for maintenance.
- **Direct Current (DC) Supply:** The circuit breakers, protection relays, and monitoring and control system of electrical power systems are supplied by DC power for operating them. Because DC power can be stored in batteries and when total failure of incoming power occurs, all components of electric power system must continue to operate by the power stored in batteries till the normal situation is restored [22].

2.4.3 Transmission Line Protection

TL protection systems are designed to identify the place of faults and isolate only the faulted section from PS, TLs can be protected by many types of protection relays as below:

- Distance protection
- Differential protection
- Directional protection
- Over current protection

2.5 Distance Protection Relay

The distance protection relay is a digital device utilized for protection of power TLs and is responsible for detecting faults and making decision to isolate faulty section, therefore, it contributes to the security and reliability of the PS. It monitors and measures the current and voltage signals of the power system at each moment via CTs and VTs to detect any problem that may occur on the TL in the power system and issues the trip signal to the circuit breaker to disconnect the faulty line.

Based on data measured directly from measurement equipment (CTs & VTs transformers) and compared to the setting values, protection system has to make an accurate, fast and right decision during abnormal conditions.

If there is a fault, the distance protection system will isolate a faulty part from PS as quickly as possible to minimize the impact of fault on the equipment of the utilities and to avoid interruption of electric supply, but if there is no fault, the system has to stay ready for any fault may occur.

2.5.1 Main Functions of Distance Protection Relay

Detection, classification, and location

- **Detection:** is capability of distinguishing between faulty and non-faulty states.
- **Classification:** is process of determining and classifying the fault type.
- **Location:** is process of calculating, locating (estimating) the distance between a reference point (Station/Substation) and the location of the fault.

2.5.2 Principle of Distance protection Relay

Distance protection device measures the impedance of the short-circuit fault by using voltage and current signal values measured at substation where the distance relay is installed as depicted in Figure 2.4. The impedance of the fault is then compared to the impedance of the transmission line. If the calculated impedance is less than the impedance of the line, an internal fault is discovered, and a trip order is given to the circuit breaker to be disconnected.

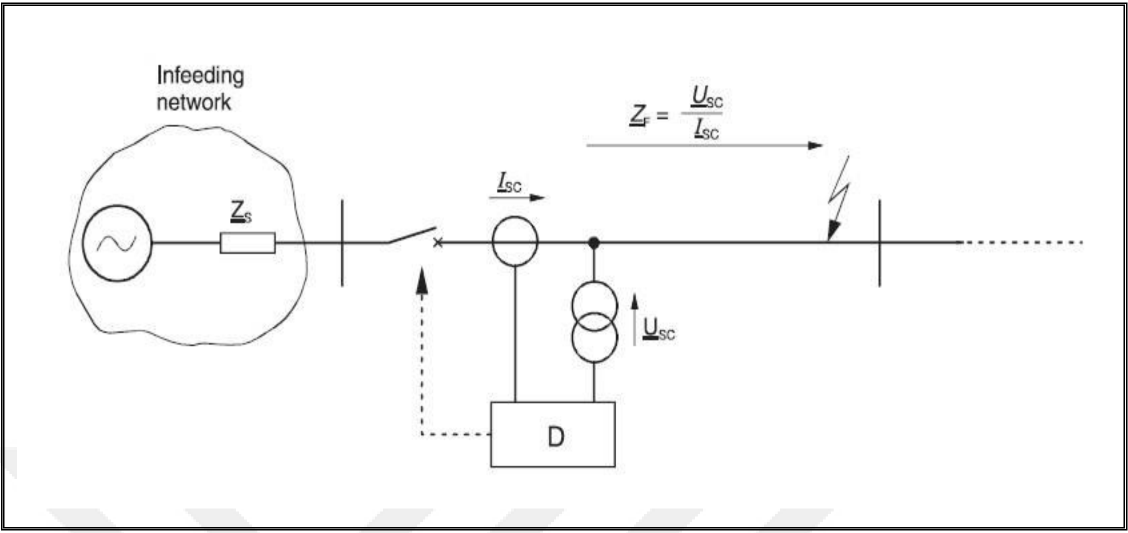


Figure 2.4 Distance protection principle

Distance protection device is provided with current and voltage signals measured by measurement transformers (CT and VT). The relay then calculates a secondary circuit impedance arising from the CT and VT transforming ratios:

$$Z_{sec} = \frac{I_{prim}/I_{sec}}{U_{prim}/I_{sec}} \tag{2.1}$$

The protection relay settings are performed with secondary circuit impedances since the testing of the relay is implemented with secondary circuit signals received from instrument transformers. Therefore, the impedance values must always be transformed by using Equation (2.1) [24].

CHAPTER III

WAVELET TRANSFORM

3.1 Introduction to Wavelet Transform

Wavelet analysis represents a windowing technique with variable-sized regions. Wavelet analysis allows the use of long-time intervals where we want more precise low-frequency information, and shorter regions where we want high-frequency information as illustrated in Figure 3.1 [25].

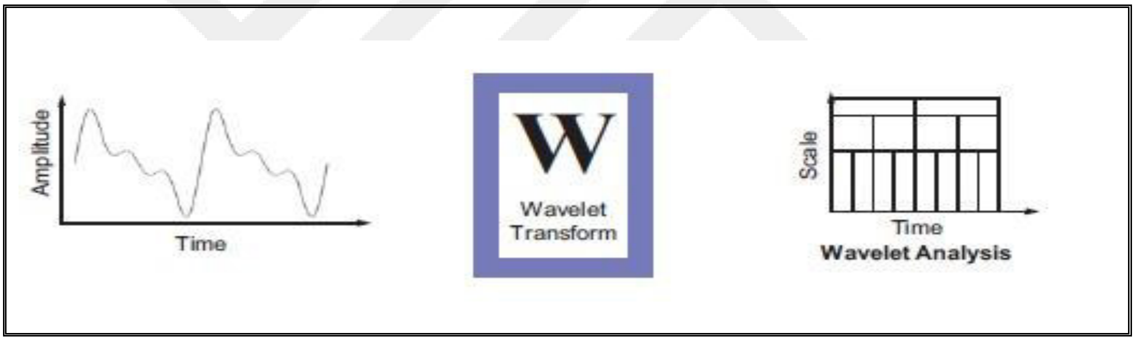


Figure 3.1 Wavelet analysis

A wavelet is a waveform of effectively limited duration that has an average value of zero as shown in Figure 3.2. One major advantage afforded by wavelets is the ability to perform local analysis -that is, to analyze a localized area of a larger signal. [25]

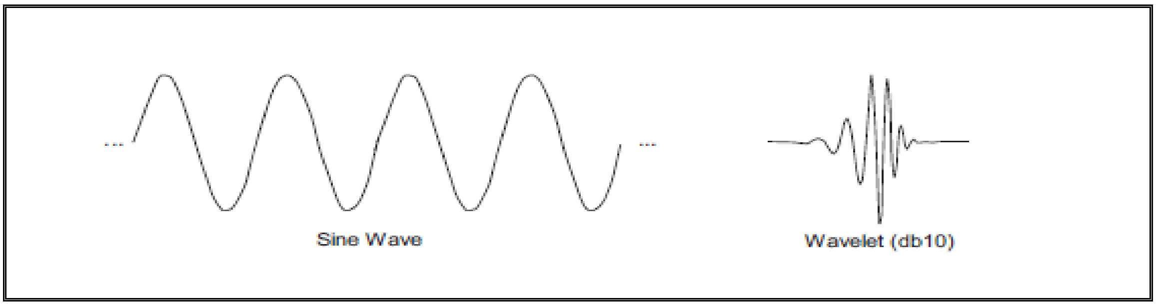


Figure 3.2 Comparison between wavelet & sine wave

Comparison between wavelets and sine waves, which are the basis of FT. Sinusoids do not have limited duration -they have range from minus to plus infinity. And where sinusoids are fine and predictable, wavelets have a tendency to be irregular and asymmetric. FT consists of decomposition a signal into sine waves of different frequencies. Likewise, WT is the decomposition of a signal into shifted and scaled versions of the original (or mother) wavelet.

WT is used as a mathematical tool for signal processing, it is considered a robust tool in the analysis of transient phenomena such as that associated with faults in PS, and has ability to analyze signal and extract information from the transient signals instantaneously in both time and frequency domain effectively.

Wavelet analysis beats the limitations of FT by localizing the fault signal both in time and frequency domain. FT does not provide information about the time of occurrence of the fault in the non-stationary current/voltage waveform of the PS. In general, WT is in two forms: a continuous and a discrete technique [21].

3.2 Continuous Wavelet Transform (CWT)

CWT is defined as the sum over all time of the signal multiplied by scaled, shifted versions of the wavelet function ψ:

$$C(scale, position) = \int_{-\infty}^{+\infty} f(t)\psi(scale, position, t).dt \qquad (3.1)$$

The continuous wavelet transforms equation of a continuous time signal $x(t)$ can be written as the following formula: [27].

$$CWT_{\psi}x(a,b) = \int_{-\infty}^{+\infty} x(t)\psi^{*}{}_{a,b}(t).dt \qquad (3.2)$$

where:

$$\psi^{*}{}_{b,a}(t) = \frac{1}{\sqrt{a}}\psi^{*}\left(\frac{t-b}{a}\right) \qquad (3.3)$$

a is scaling, b is translation (location of window), ψ is the mother wavelet.
The results of the CWT are many wavelet coefficients C, which are a function of

scale and position.

3.3 Discrete Wavelet Transform

The discrete wavelet transforms equation of the sampled signal $x(k)$ can be written as the following formula:

$$DWT_{\psi}x(m,n) = \sum_{k} x(k) \,.\, \psi^{*}_{m,n}(k) \tag{3.4}$$

where

$$\psi^{*}_{m,n}(k) = \frac{1}{\sqrt{a_0^m}} \psi^{*} \left(\frac{k - nb_0 a_0^m}{a_0^m} \right) \tag{3.5}$$

m is Scaling, n is Number of data points

3.3.1 Multiple-Level Decomposition

Multiple-Level Decomposition is one of the tools of DWT, which breaks down original, normally non-stationary signal into low frequency signals called approximations and high frequency signals called details, with different levels of resolution [27]. The breakdown process can be repeated, with sequential approximations being broken down in turn, so that one signal is analyzed into many lower resolution components. This is called the wavelet decomposition tree as illustrated in Figure 3.3 [25].

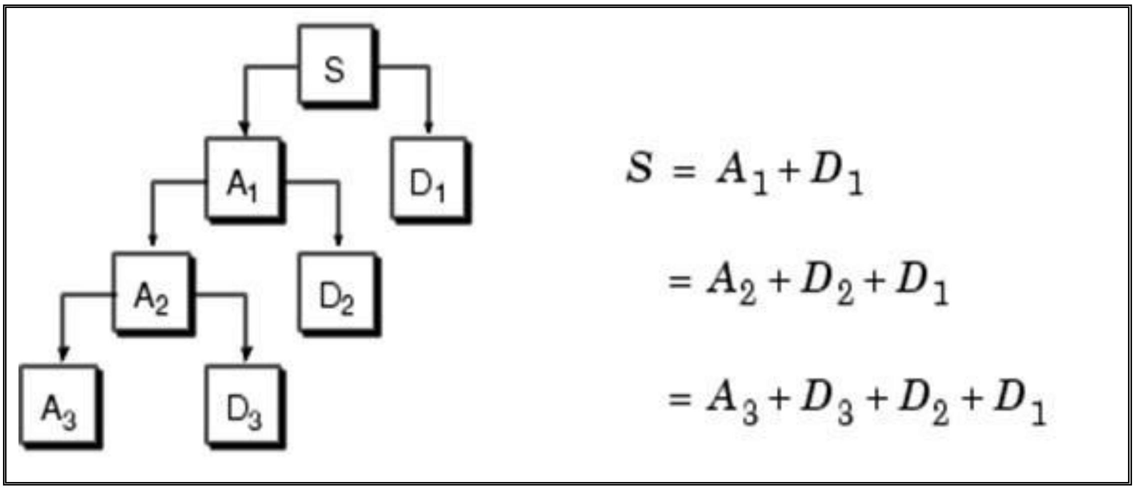


Figure 3.3 Diagram of DWT multiple-level decomposition

DWT is beneficial in analyzing transient signals such as that associated with faults on HVTL. The approximation and detail coefficients are defined as [26]:

$$c_j(n) = \sum_k h_0(k - 2n)\, c_{j-1}(k) \tag{3.6}$$

$$d_j(n) = \sum_k h_1(k - 2n)\, c_{j-1}(k) \tag{3.7}$$

where h_0 and h_1 represent the coefficients of the selected scaling and wavelet functions. Thus, the signal is represented by the following set of coefficients

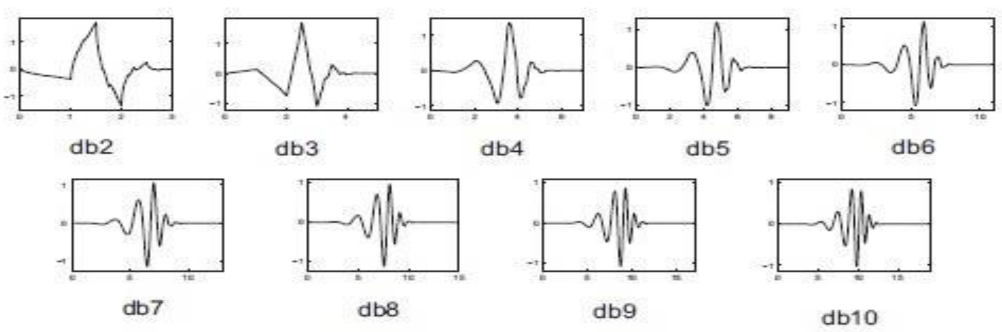
$$x_{signal} = \left[c_j\,, d_j\,, d_{j-1}\;,\; d_{j-2}\;\; \dots\dots\dots\dots, d_1\right]$$

3.4 Wavelet Functions

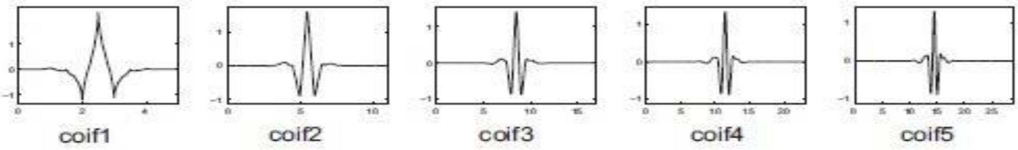
WT for original signals is set of functions generated from one single function, called a mother wavelet, by means of scaling and translating operations.

There are many types of mother wavelets with various characteristics, such as Symlets (sym), Daubechies (db), Coiflets (coif), etc. Some mother wavelets are shown in Figure 3.4. The prime difference between these wavelets is due to the different lengths of filters that define the wavelet and scaling functions. The selection of mother wavelet is decided according to the required application [25].

Daubechies Family



Coiflets Family



Symlets Family

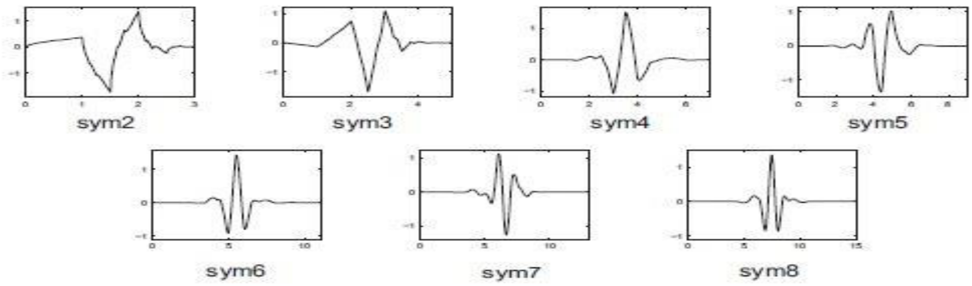


Figure 3.4 Some types of wavelet functions

An example of wavelet analysis using Daubechies family (db5) as mother wavelet at fifth level is illustrated in Figure 3.5.

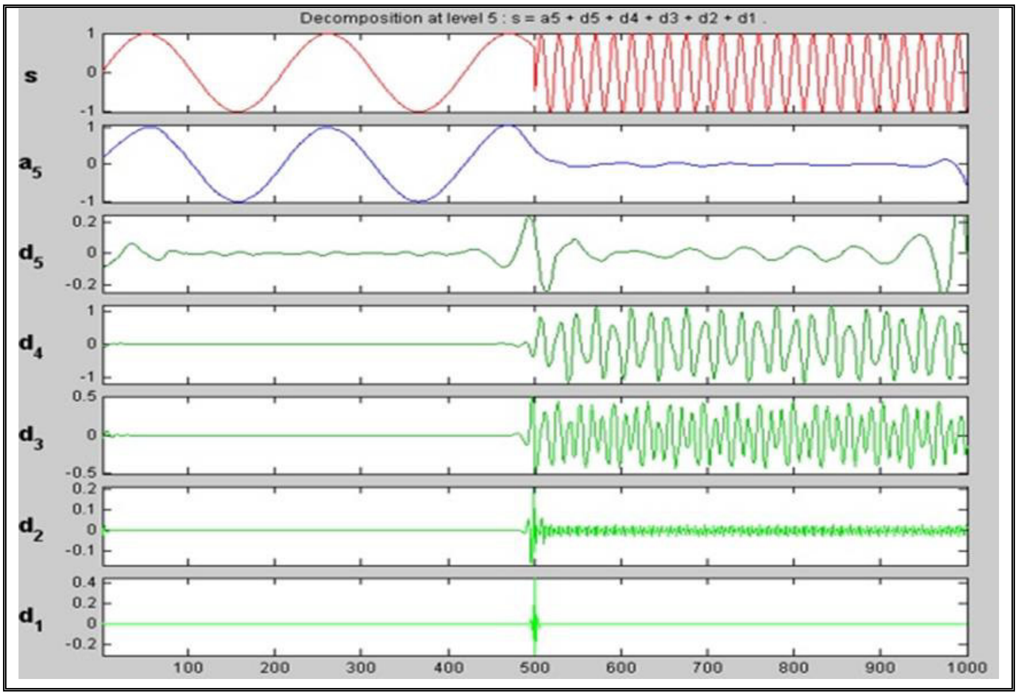


Figure 3.5 Example of wavelet analysis

Selection of a suitable wavelet function depends on the shape of waveform, among wavelet functions, the Daubechies are widely used for power system transient study, especially db4 therefore, in this work, db4 mother wavelet is chosen to analyze the faults signals.

3.5 Features Extraction

When the data size is very large, it is difficult to be processed completely and it is excrescent (much data, does not always mean much information) then the data is transformed into a reduced representation set of features (also named features vector). Transforming the data into the set of features is called features extraction. Thus, the extraction of discriminatory features in the signal boosts the reduction of the length of the data vector significantly. DWT is used to extract beneficial attributes from a signal by decomposition of signal into several levels as mentioned above. The essential feature extraction process consists of:

1. Decomposing the sampled signal by applying DWT into many levels to obtain the approximation and detailed coefficients
2. Extracting the features from the DWT coefficients.

In this work, the features extracted from the analyzed fault signal are as below:

- **Energy (E):**

  The energy of details and approximation coefficients are calculated using the following equations:

$$ED_i = \sum_{j=1}^{N_m} \left|d_{ij}\right|^2 \ , \ \ i = 1, 2, \dots\dots, m \tag{3.8}$$

$$EA_i = \sum_{j=1}^{N_m} \left|a_{ij}\right|^2 \ , \ \ i = 1, 2, \dots\dots, m \tag{3.9}$$

where:

EDi: is the energy of the detail at decomposition level i.

EAi: is the energy of the approximation at decomposition level i.

D: is the magnitude of the detail coefficient

A: is the magnitude of the detail coefficient

i: is the wavelet decomposition level

N: is the number of detail or approximation coefficients at each decomposition level.

- **Mean:**

The Mean is average that equals the sum of the values divided by the number of values. It is calculated using the following equations:

$$Mean = \frac{1}{N}\sum_{j=1}^{N}\left(d_{ij}\right) \tag{3.10}$$

- **Standard Deviation (SD):**

The SD is a statistical measure of the dispersion of a dataset relative to its mean. It is calculated using the following equations:

$$SD = \sqrt{\frac{1}{N-1}\left[\sum_{j=1}^{N}\left(d_{ij} - \bar{\mu}\right)\right]} \tag{3.11}$$

Where:

$\boldsymbol{d_{ij}}$: Value of the j^{th} point in the data set

$\boldsymbol{\bar{\mu}}$: Mean value of the data set

$\boldsymbol{N}$: Number of data points in the data set

- **RMS:**

RMS is defined as the square root of the mean square. It is calculated using the following equations:

$$RMS = \sqrt{\frac{1}{N}\sum_{j=1}^{N}\left(d_{ij}^{2}\right)} \tag{3.12}$$

The extracted features contain useful information that are fed as inputs to ANN.

In this work, the Discrete Wavelet Transform is applied on sampled values of three-phase fault current signals of fault for decomposition them up to fourth level and calculating wavelet coefficients (Details & Approximation) for four levels resulting from decomposition.

CHAPTER IV

ARTIFICIAL NEURAL NETWORK

4.1 Introduction

The realization that the way which human brain uses it in computation is an entirely different way from the traditional digital processor has promoted working on ANNs that process data in a highly complex, nonlinear, and parallel way simulates the human brain, and perform the tasks much faster than digital processors available today [28].

An ANN is a data processing unit that has specific performance attributes and have been created as generalizations of mathematical models of neural biology, based on the following assumptions:

1- Data processing occurs at many components called neurons.
2- Signals are passed between neurons over connection links.
3- Each connection link has a related weight, which, in a standard neural network, multiplies the signal transmitted.
4- Each neuron applies a TF (generally nonlinear) to its network input (sum of weighted input signals) to define its output signal [29].

The NN is typically created by using electronic items, the procedure followed to perform the learning process is called a learning algorithm, which sets weights and biases of the NN and updates them continuously until reaching the target output of NN, so effectively performs the purpose of the design [28].

4.2 Model of Neuron:

A neuron is data-processing unit that is considered key component of building and operating a NN. Figure 4.1 shows the model of a neuron [28].

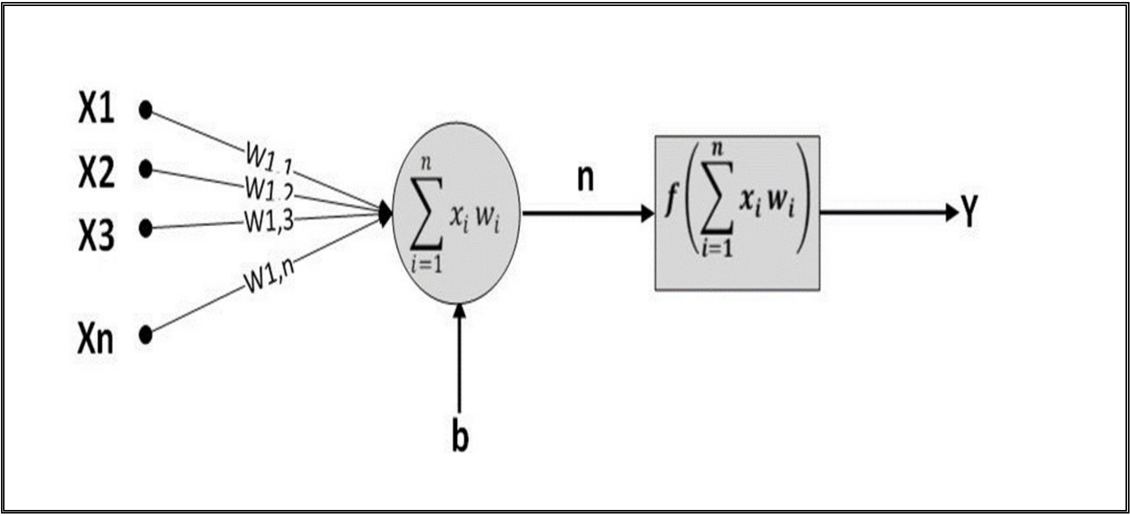


Figure 4.1 Basic neuron model

In mathematical terms, the neuron depicted in Figure 4.1 may present with the following equations:

$$n = [W]^T * [X] + [b] \tag{4.1}$$

$$n = w_1 * x_1 + w_2 * x_2 + .. + w_n * x_n + b \tag{4.2}$$

$$y = f(n) \tag{4.3}$$

where:

$\boldsymbol{X}$ is inputs vector, $\boldsymbol{W}$ is weights vector.

Weight: represents the strength of the connection among neurons of different layers.

Bias: it is an extra input to neuron and equals to one and has its own connection weight.

$\boldsymbol{w}$ and $\boldsymbol{b}$ are adjustable and modifiable parameters, and their values will be modified by learning rule so that the relationship between input and output of neuron meets specific goal.

4.3 Transfer Function (TF)

TF is mathematical equation that define the output of a neuron, each neuron in the NN has own TF, it also supports to standardize the output of each neuron to be within a scale between 1 and 0 or between -1 and 1. At each layer of a NN, the same TF is applied for all neurons. Selection of TF should be based on the problem required to be resolved by NN, the most appropriate TF is nominated by the designer.

4.3.1 Types of Transfer Function:

The TF may be in several types, a some of which are labeled and described below:

- **Hard Limit TF**

The ***hard limit TF***, shown in Figure 4.2, sets the output of the neuron to 0 if the input is less than 0, or 1 if its input is greater than or equal to 0. This function to create neurons that classify inputs into two different categories. [30]

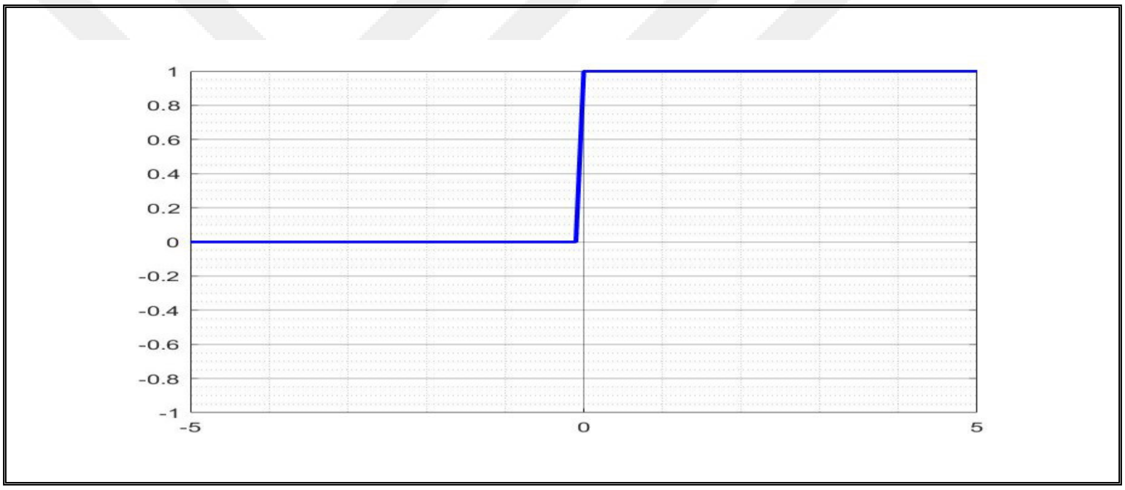


Figure 4.2 Hard limit TF

- **Linear TF**

The ***Linear TF***, shown in Figure 4.3, its output is equal to its input, according to the expression:

$$f(x) = x \tag{4.4}$$

Neurons with this TF are used in the Adaline networks.

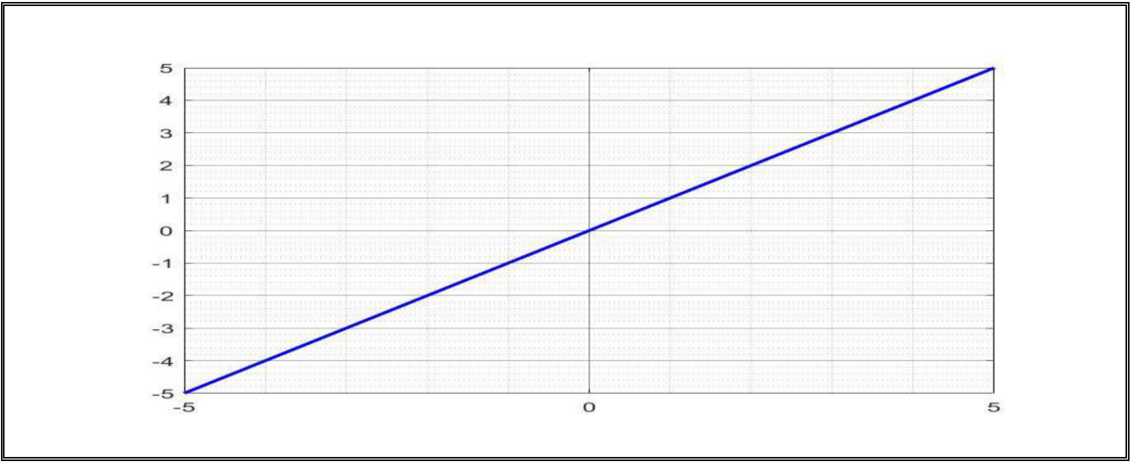


Figure 4.3 Linear TF

- **Log-Sigmoid TF:**

The ***Log-Sigmoid TF***, shown in Figure 4.4, it takes the input which may have any value between plus and minus infinity and converts the output into the range 0 to 1, according to the following equation:

$$f(x) = \frac{1}{1+e^{-x}} \tag{4.5}$$

This TF is commonly used in MLNN that are trained using the backpropagation algorithm, because it is differentiable [31].

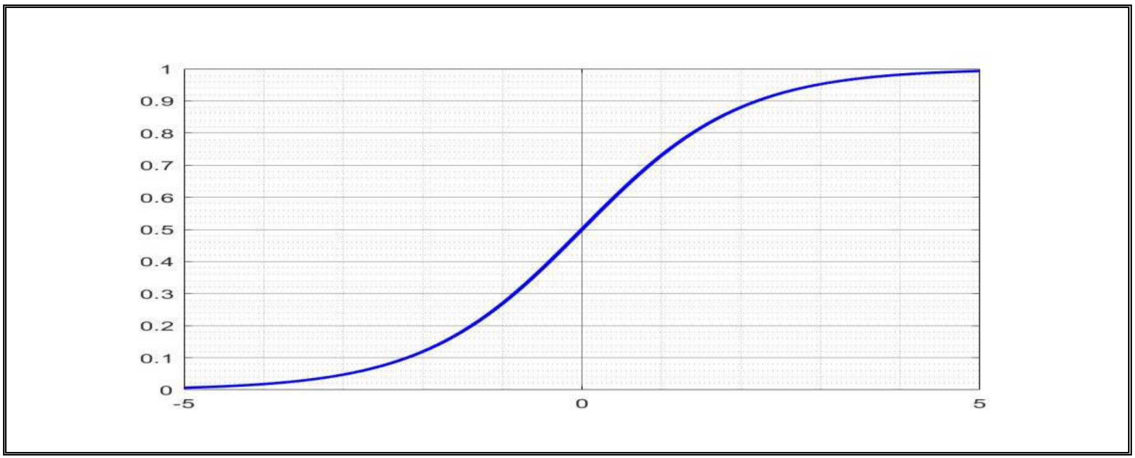


Figure 4.4 Log-sigmoid TF

- **Hyperbolic Tangent Sigmoid TF**

The ***Tangent-Sigmoid TF***, shown in Figure 4.5, takes the input which may have any value between plus and minus infinity and converts the output into the range -1 to 1, according to the expression:

$$f(x) = \frac{e^{x} - e^{\ x}}{e^{x} + e^{\ x}} \quad (4.6)$$

This TF is also commonly used in MLNN that are trained using the backpropagation algorithm, in part because it is differentiable [30].

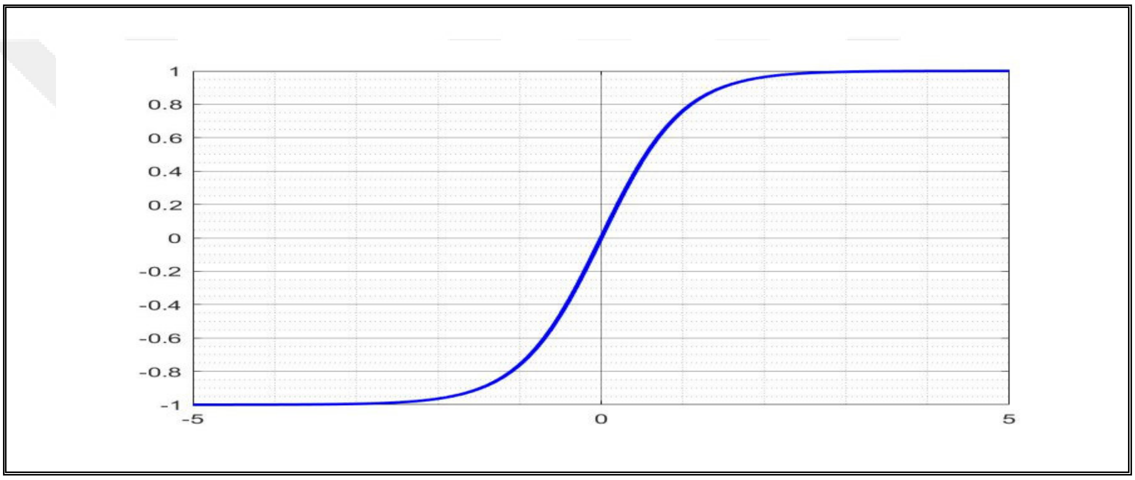


Figure 4.5 Tangent-sigmoid TF

4.4 Basic Structure of ANN:

ANN are comprised of multiple neurons. The neurons are interrelated and arranged in many layers and they interact with each other. The neurons receive input data and perform simple processes on that data. The result of these processes is forwarded to another group of neurons, the connections among neurons have weights value related with them. ANNs has capability to learn, which is performed by changing weight values [31].

Based on connections among neurons in a model, NN may be categorized into two types namely feed-forward and feedback NN. Unlike feed forward NN, the feedback NN has a feedback connection fed back into the network along with the inputs.

An Artificial Neural Network (ANN) helps to build predictive models from large databases. The basic structure of NN (input layer, hidden layers, and output layer) is illustrated in Figure 4.6.

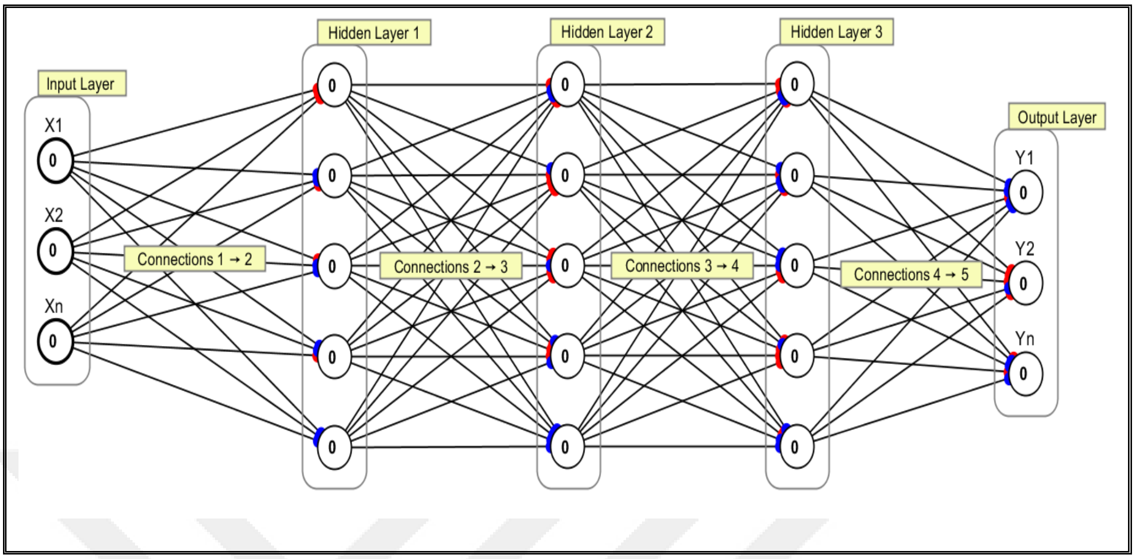


Figure 4.6 Basic neural network structure

4.4.1 Feed-Forward Neural Network

In this type of NN, information streams in feed forward direction i.e. from input layer to output layer so there is no feedback connection.

Figure 4.7 shows the structure of a feed forward MLNN with K_0 input and K_M output signals. In this type of NN, data passes in a single direction, from the input to the output. The in-between layer is called hidden because there are no connections between their neurons and outside information.

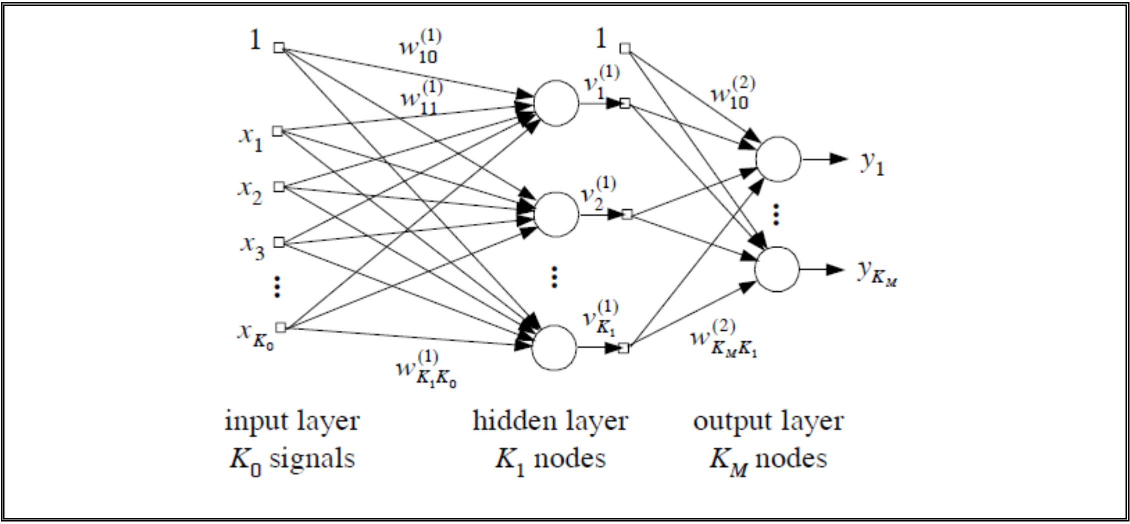


Figure 4.7 Structure of feed-forward multi-layer network

Signal processing in the ith network layer is performed according to the following equation:

$$v^{(i)} = f^{(i)}\left(W^{(i)} * q^{(i-1)}\right) \tag{4.7}$$

where:

$$v^{(i)} = [v_1^i \quad v_2^i \quad v_3^i \quad \dots \dots \dots \quad v_k^i] \tag{4.8}$$

$$W^{(i)} = \begin{bmatrix} w_{10}^i & w_{11}^i & \dots & w_{1k_{i-1}}^i \\ w_{20}^i & w_{21}^i & & w_{2k_{i-1}}^i \\ \vdots & & \ddots & \vdots \\ w_{k_{i0}}^i & w_{k_{i1}}^i & \dots & w_{k_i2k_{i-1}}^i \end{bmatrix} \tag{4.9}$$

$$q^{(i-1)} = \begin{cases} \begin{bmatrix} 1 \\ x \end{bmatrix} & for\ i = 1 \\ \begin{bmatrix} 1 \\ v^{(i-1)} \end{bmatrix} & for\ i = 2,3, \dots, M \end{cases} \tag{4.10}$$

$$y = v^{(M)} = [y_1 \quad y_2 \dots \dots \quad y_{K_M}]^T \tag{4.11}$$

$\boldsymbol{y}$: represents the output vector of the result of the network

4.4.2 Feed-Back Neural Network

In this type of NNs, signals pass in both the directions from input layer to output layer to generate the result, the error in output is then transferred back to previous layer. Neurons learn how much they contributed to the result being wrong. Weights are re-adjusted, and performance of NN is improved, during this process, the NN learns. Figure 4.8 shows the structure of a feedback NN [31].

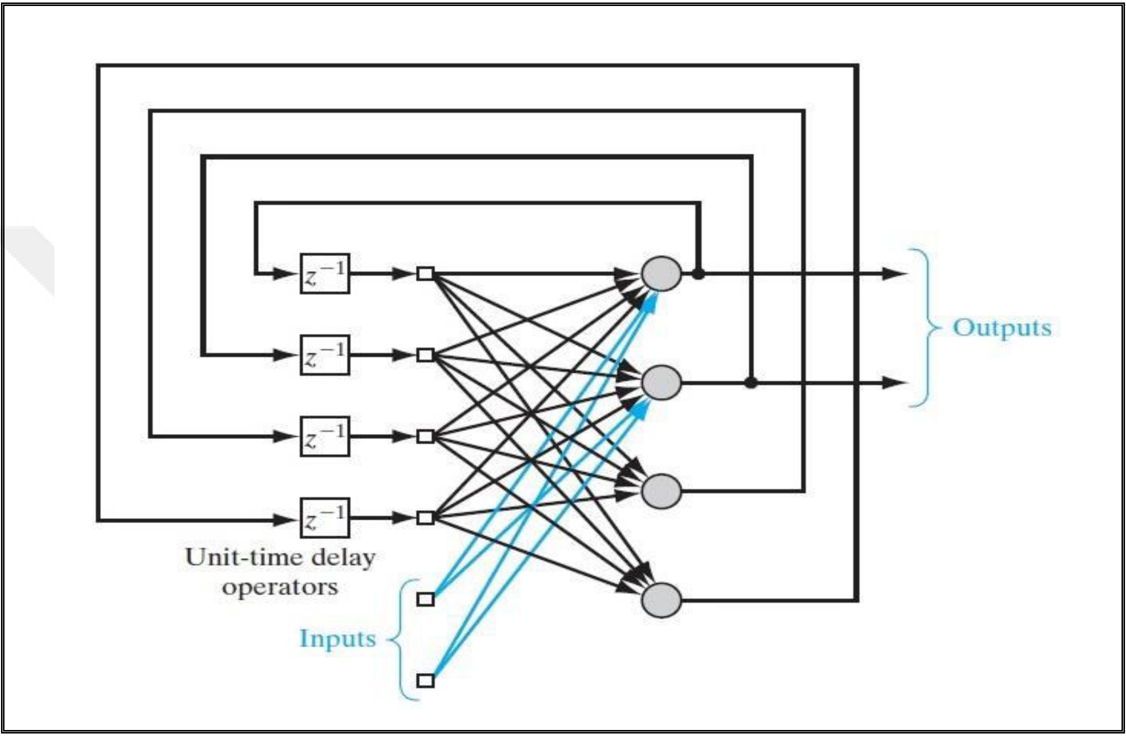


Figure 4.8 Structure of feed-back NN

4.5 Learning Process:

The basic issue in constructing of ANN is definition of the weights in order to able to achieve the desired output. This process is commonly called learning or training. Learning processes through which NNs learn may be categorized into learning with a instructor (supervised learning) and learning without a instructor, learning without a instructor may be subcategorized into unsupervised learning and reinforcement learning [28].

4.5.1 Supervised Learning

Figure 4.9 shows a block diagram that illustrates the supervised learning. In this type of learning, the training dataset contains input data and the output results required to be predicted by model. The model is trained until it finds relationship between the input and the outputs, the trained model is then tested with test dataset to verify the result of the training process and measure the accuracy.

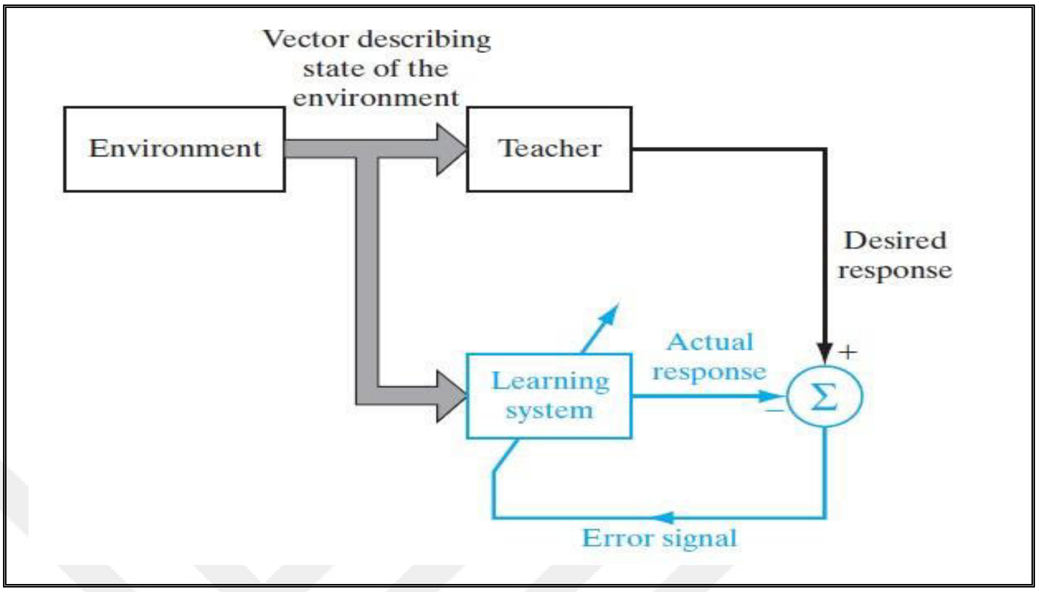


Figure 4.9 Block diagram of supervised learning

During training stage, the NN parameters (weights and biases) are adjusted under the collective influence of the training data and the error signal. The error signal is defined as the discrepancy between the target output value and the actual output value of the NN. This adjustment is performed iteratively in a step-by-step with the purpose of eventually making the NN reaches the best performance. In this way, training data set that presents knowledge of the environment available is transferred to the NN through training and stored in the form of weights and biases. When this condition is reached, we may leave the NN interact with the environment absolutely by itself, therefore the supervised-learning process represents a closed-loop feedback system [28].

Supervised learning classified into two categories of algorithms:

- **Classification**: A classification issue is when the model categorizes the output variable such as type or color.
- **Regression**: A regression issue is when the model produces a true value such as price or distance.

4.5.2 Unsupervised Learning

In unsupervised, that is also called self-controlled, there is no participation from an external instructor to supervise the learning process, as indicated in Figure 4.10.

Unsupervised learning is the training of machine by using data that is neither classified nor labeled and allowing the model to act based on that data without supervising. In this case, the machine function is to cluster unsorted dataset according to similarities, patterns, and discrepancies without any previous training of data. Unlike supervised learning, no instructor is provided, that means there will not be training process performed to the machine. Therefore, machine is constrained to discover the hidden structure in unlabeled data by itself [28].

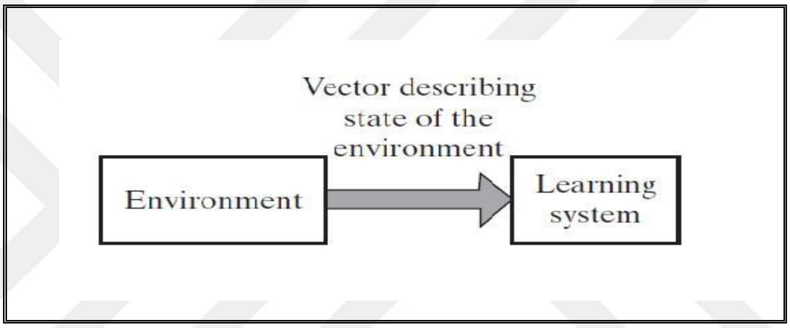


Figure 4.10 Block diagram of unsupervised learning

Unsupervised learning classified into two categories of algorithms:

- **Clustering:** A clustering is function of gathering a group of data in a way so data in the same group is more identical, such as grouping employees by skills they have.
- **Association:** An association is a method for discovering attractive links between variables in significant dataset [32].

4.5.3 Reinforcement Learning

In reinforcement learning, the learning of an input‐output mapping is performed through continued interaction with the environment to minimize a scalar index of performance. Figure 4.11 shows the block diagram of one form of a reinforcement-learning system built around a critic that converts a primary reinforcement signal received from the environment into a higher quality reinforcement signal called the heuristic reinforcement signal, both of which are scalar inputs. The system is designed to learn under delayed reinforcement, which means that the system observes a temporal sequence of stimuli also received from the environment, which eventually result in the generation of the heuristic reinforcement signal [28].

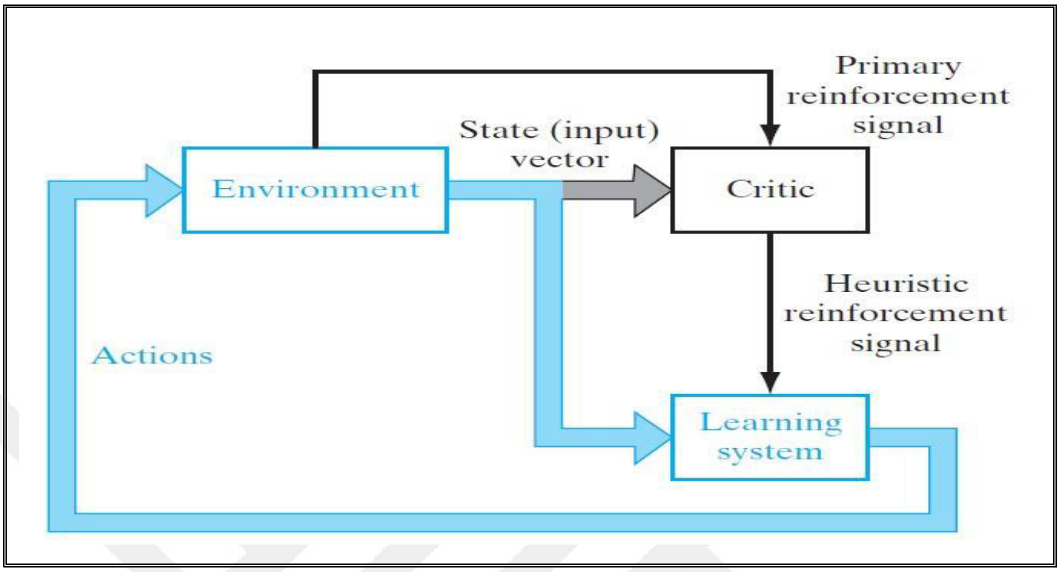


Figure 4.11 Block diagram of reinforcement learning

4.6 Back-Propagation Neural Network:

The learning algorithm is the mechanism that turns a data set into a model. The most general algorithm used for supervised learning of multilayered neural network is Back-Propagation using gradient descent.

In the BPNN, for each input pattern, the network must be also provided with a target output pattern, the output is feedback to the input to calculate the change in the values of weights.

When using gradient descent for training a NN, a performance function is computed, which indicates the difference between the predictions of the NN and the true labels. Backpropagation helps us to compute the gradient of the performance function for each weight of ANN. Therefore, every weight is separately updated to gradually reduce the performance function through several training iterations.

The error for each iteration and for each point is calculated by initiating from the last step and by sending calculated the error backwards. The weights of the back-error-propagation algorithm for the neural network are chosen randomly, feeds back in an input pair and then obtain the result. After each step, the weights are updated with the new ones and the process is repeated for entire set of inputs-outputs combinations available in the training data set provided by developer. This process is repeated until the network converges for the given values of the targets for a pre-defined value of error tolerance. The entire process of back propagation can be understood by Figure 4.12 & Figure 4.13. The proposed algorithm uses the MSE technique for calculating the error in each iteration [5].

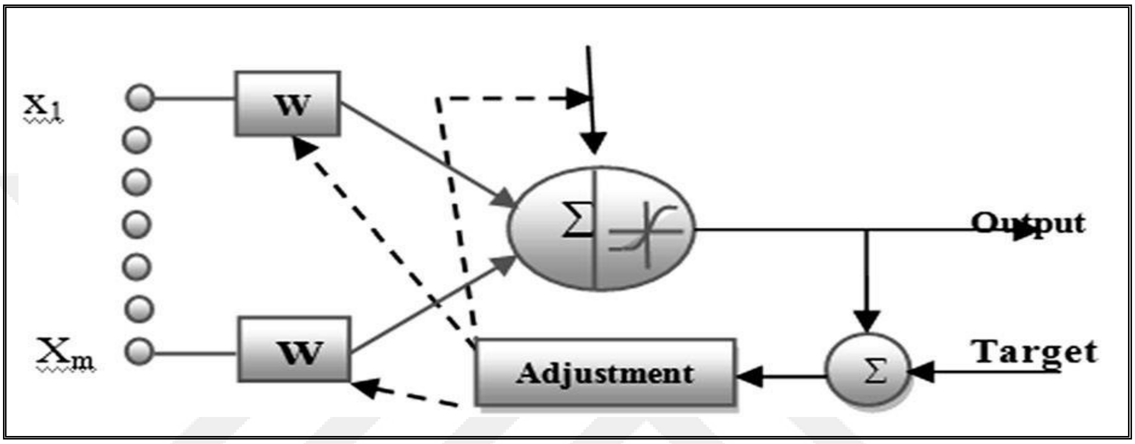


Figure 4.12 Single line of back-propagation learning.

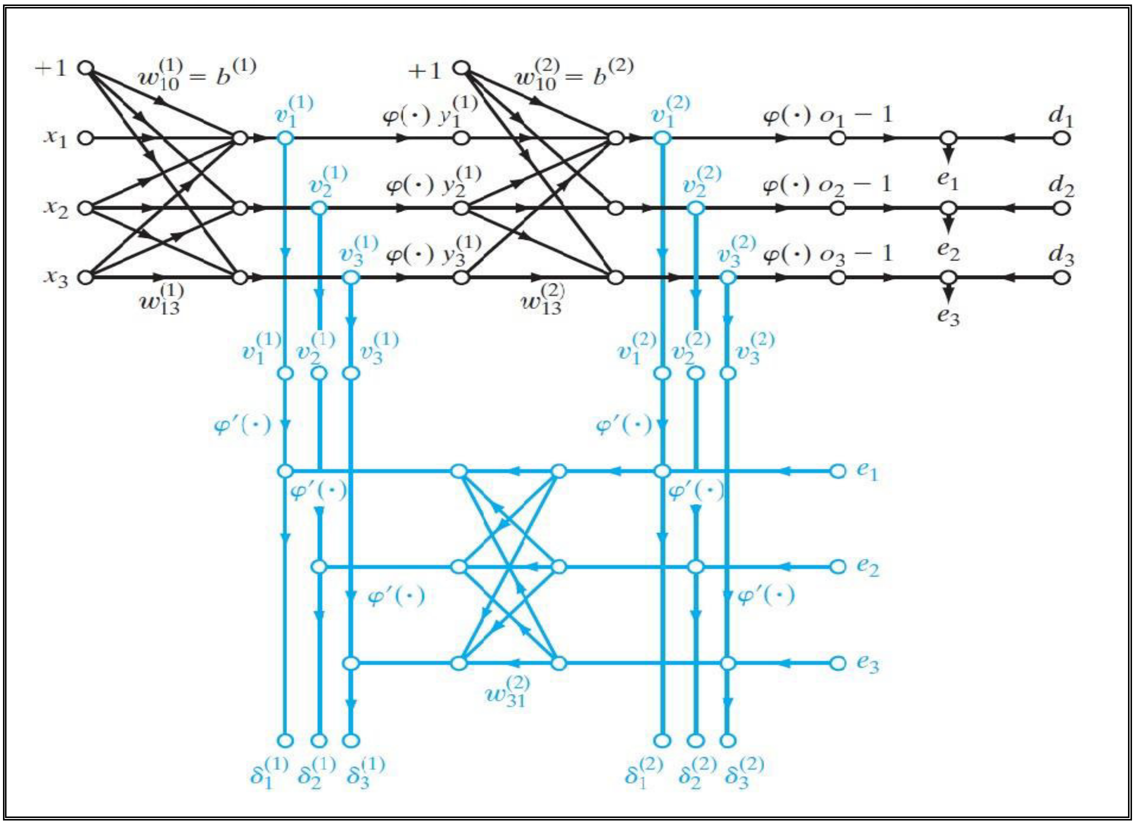


Figure 4.13 Signal-flow graphical summary of back-propagation learning

The algorithm of BP is as follows:

- **Initialization:** initial weights are applied to all the neurons.
- **Forward propagation**: the inputs from a training set are passed through the neural network and an output is computed. The output for neuron *j* in layer *l* is:

$$v_j^l(n) = \sum_i w_{ji}^l(n) y_i^{l-1}(n) \tag{4.12}$$

where $y_i^{l-1}(n)$ is the output signal of neuron i in the previous layer $l-1$ at iteration n, and $w_{ji}^l(n)$ is the synaptic weight of neuron j in layer l that is fed from neuron i in layer $l-1$. For $i=0$, we have $y_0^{l-1}(n) = +1$, and $w_{j0}^l(n) = b_j^{(l)}(n)$ is the bias applied to neuron *j* in layer *l*. The output signal of neuron j in layer l is:

$$y_j^{(l)} = f_j\left(v_j(n)\right) \tag{4.13}$$

If neuron *j* is in the first hidden layer, set ($l=1$)

$$y_j^{(1)}(n) = x_j(n) \tag{4.14}$$

where $x_j(n)$ is the *jth* element of the input vector $x(n)$.

If neuron *j* is in the output layer, set ($l =$ L) (where L is referred to as the depth of the network).

$$y_j^{(L)}(n) = O_j(n) \tag{4.15}$$

The error (between actual output and target output) is:

$$e_j(n) = d_j(n) - O_j(n) \tag{4.16}$$

where $d_j(n)$ is the *jth* element of the desired response vector $d(n)$.

- **Backward propagation:**

Compute the δ_s (i.e., local gradients) of the network, defined by:

$$\delta_j^{(l)}(n) = \begin{cases} e_j^{(L)}(n)\varphi_j'\left(v_j^{(L)}(n)\right) & \text{for neuron } j \text{ in output layer } L \\ \varphi_j'\left(v_j^{(L)}(n)\right)\sum_k \delta_k^{(l+1)}(n)w_{kj}^{(l+1)}(n) & \text{for neuron } j \text{ in hidden layer } l \end{cases} \quad (4.17)$$

where the prime in φ_j' denotes differentiation with respect to the argument. Adjust the synaptic weights of the network in layer *l* according to the generalized delta rule:

$$w_{ji}^{(l)}(n+1) = w_{ji}^{(l)}(n) + \alpha\left[\Delta w_{ji}^{(l)}(n-1)\right] + \eta\left[\delta_j^{(l)}(n)y_j^{(l-1)}(n)\right] \quad (4.18)$$

where $\boldsymbol{\eta}$ is the learning-rate parameter and $\boldsymbol{\alpha}$ is the momentum constant.

- **Iteration:** Iterate the forward and backward computations by presenting new epochs of training data to the network until the chosen stopping criterion is met [28].

4.7 Training of Neural Network:

The training process of NN is setting weight values to decrease the error between actual and target output. It is started with initial values (selected randomly) for the NN parameters (weights and biases) and compute the error between actual and target output, then the parameters are updated by backpropagation algorithm until reach the lowest MSE, that is employed as performance index then the weight values are the best. The MSE for each output in each iteration is computed by equation:

$$MSE = \frac{1}{N}\sum_1^N (E_i - E_o)^2 \quad (4.19)$$

where:

N: is number of iterations, Ei: is actual output.

Eo: is output of the model.

ANN can be applied to FD and FC effectively because it is a technique that depends principally on programming, capable to solve the nonlinear problems easily. The

ANNs are able to learn with experiences. Therefore, they are commonly accepted and applied in the problem of FD and FC because of the following characteristics:

- Various configurations of transmission lines (short, long, single circuit, double circuit, etc.)
- Simulation of the network with different PS conditions.
- The ANN output is very rapid, reliable, and accurate.

Learning algorithm utilized is supervised learning, so all data is labeled, and the target is getting identified value and is associated with each input in the training set [33].

In this work, artificial neural networks have been used for designing the structure of detector, classifier, and locator. Number of hidden layers, number of neurons in each hidden layer, transfer functions, weights, and biases are defined to get the best performance and least mean square error during training.

CHAPTER V

MODELING AND SIMULATION

In this chapter, the methodology of design of fault detector, classifier and locator is demonstrated, and it illustrates how ANNs can be applied as a tool for designing the distance protection relay.

Conventional distance protection systems have been based upon parameters setting for fault currents and voltages to trip the circuit breaker that isolates the faulty section of power system accordingly. However, to detect and localize the faults more intelligently, Artificial Neural Networks (ANN) were introduced as effective tool for designing the FD, FC, and FL.

The data set used for training of ANN utilized in designing the protection system can be collected either on-line measurement data directly obtained from the field, or off-line by accessing fault and disturbance registers in the substations. Since interesting and critical cases do not occur frequently, and to obtain a sufficient training data, simulations of relevant power system at many fault cases are performed by changing parameters of faults through simulation software.

The design process of fault detector, classifier, and locator is performed according to the following steps:

- Creating of the PS model, data acquisition of three phase current signals.
- Setting up the system parameters, simulating the system under different situations, and acquiring current signals of fault.
- Sampling the measured current signals
- Application of D.W.T on the sampled current signals and computation of approximation and detail coefficients.

- Extracting of the analyzed current signal features by using wavelet coefficients
- Selection of suitable ANN structures for detector, classifier, and locator.
- Training of ANNs and validation of the trained ANNs by utilizing test data set. [34]. Figure 5.1 shows flow chart of design process

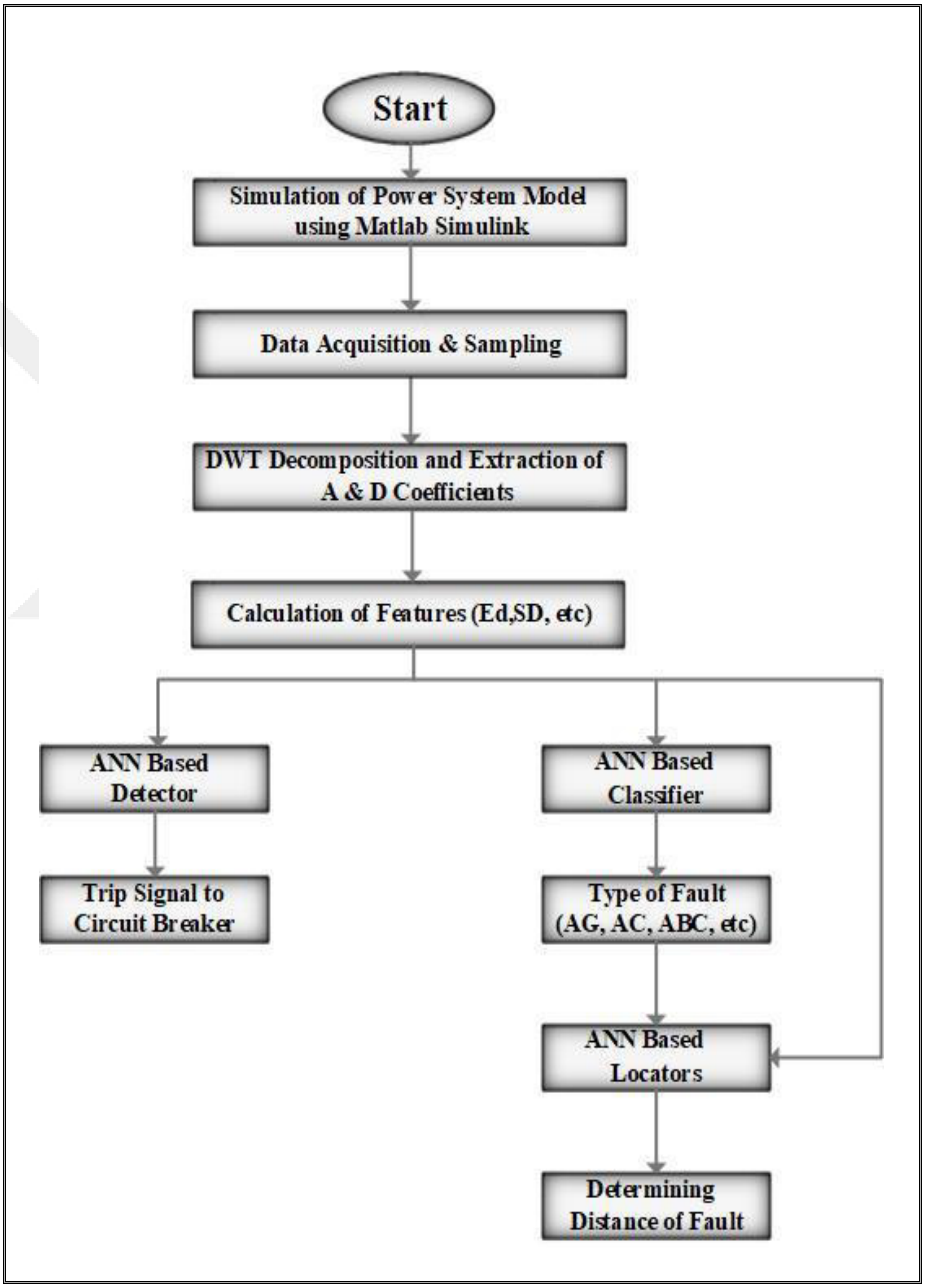


Figure 5.1 Flowchart of Design Process

5.1 Transmission Line Model

A three phase, single circuit, 50 Hz, 400kV HVTL has been employed to develop and carry out the proposed approach by using ANNs, connecting between two stations of length 300 km. Figure 5-2 illustrates the single line diagram of the power TL that has been used throughout the research as well as measurement equipment and distance protection relay.

At each end of the transmission line, it gets into the substation through measurement equipment and circuit breaker which receives the trip order from electronic protection system, measurement equipment read and step down the voltage and current signals to be suitable for electronic system, these signals are transferred to protection system that analyzes them, and according to the analysis results it takes the correct decision if trip order should be issued or not.

The model line has been constituted by employing the distributed parameters so that it more appropriately characterizes a long power TL. The standard values of distributed parameters utilized in the ***Simulink Matlab*** have been used in our model.

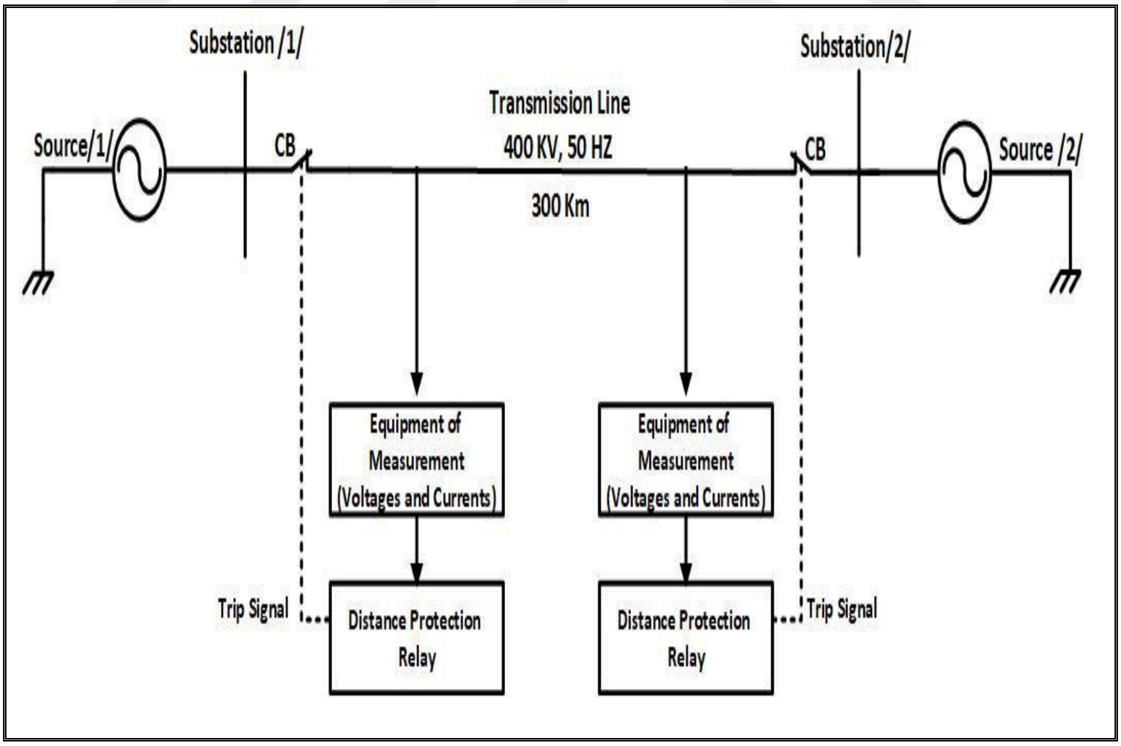


Figure 5.2 Single line diagram for transmission line and measurement and distance relay equipment.

5.2 Simulation of Modeled Transmission Line

The single line diagram shown in Figure 5.2 has been modeled by using the SimPowerSystems toolbox in MATLAB and simulated in MATLAB/Simulink environment for different types of faults at varying distances along the TL with various fault resistances to collect the training and test data sets for various fault cases. Figure 5.3 illustrates the modeled transmission line in MATLAB Simulink environment, the three-phase fault block is to make faults under different conditions through which we can set parameters of faults like type of faults, resistance, and time of faults.

The three phase V-I measurement block is used to measure the values of the three-phase currents at the end A. The three-phase fault block is used to make fault and set parameters of the fault like type, resistance, and time of fault.

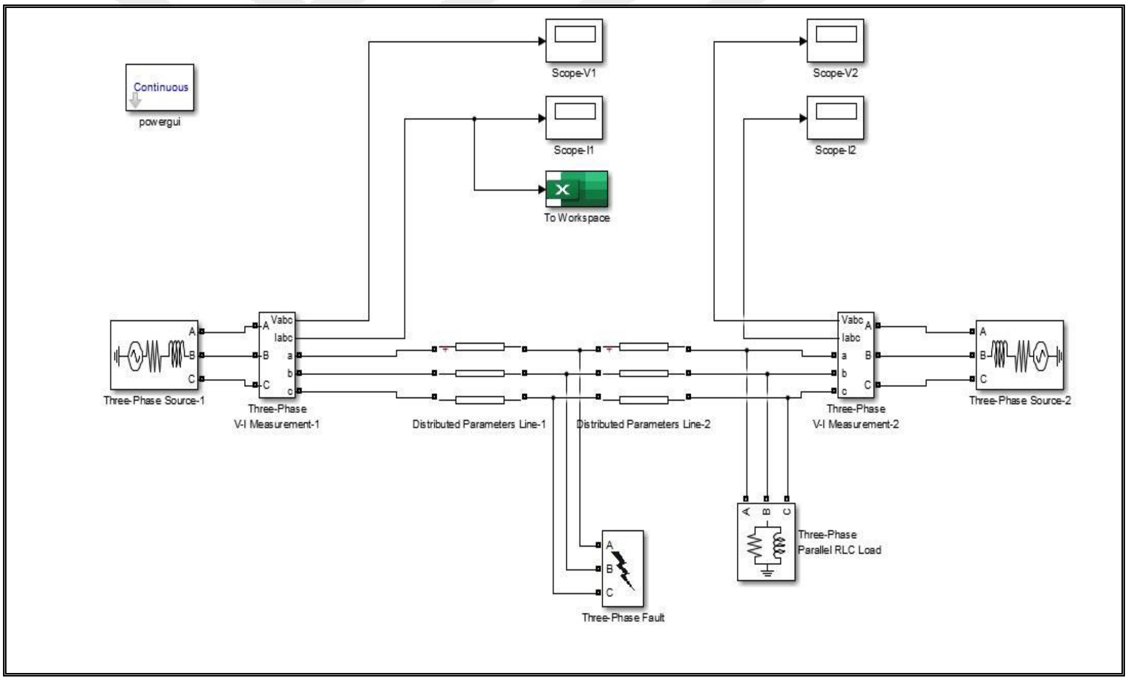


Figure 5.3 Model of high voltage transmission line

Utilized parameters for simulation of power system model illustrated in Table 5.1.

Table 5.1 Power system model data

Parameter	Value
Transmission Line	Three Phase
	Single circuit
Source Voltage/1/	400 kV
Source Voltage/2/	400 kV
Frequency	50 Hz
Load	100 MW Active, 10 MVAR Reactive
Transmission Line Length	300 km
Positive Sequence Resistance	0.01273 (Ohms/km)
Zero Sequence Resistance	0.3864 (Ohms/km)
Positive Sequence Inductance	$0.9337 \text{x} 10^{-3}$ (H/km)
Zero Sequence Inductance	$4.1264 \text{x} 10^{-3}$ (H/km)
Positive Sequence Capacitance	$12.74 \text{x} 10^{-9}$ (F/km)
Zero Sequence Capacitance	$7.751 \text{x} 10^{-9}$ (F/km)

5.3 Data Generation

Data Generation is the most important step to any method employing machine learning techniques. The designed model needs to be trained, tested, and evaluated in advance to be able to predict the output in a real environment for a data set that has not been previously trained on. Machine learning models learn from data. Therefore, it is very important to be trained on the appropriate data for the problem being addressed. Moreover, the type of the data collected controls the selection of algorithm to be applied to obtain the best findings. For the FD, FC, and FL problem, the data generation is performed through simulation of the transmission line faults using Matlab model in Simulink environment for three varying main fault parameters which are:

- **Type of the fault:** The type of simulated short circuit faults is as following:
  - Line to ground (A-G, B-G, C-G).
  - Line to line (A-B, B-C, A-C).
  - Line to line to ground (A-B-G, B-C-G, A-C-G).
  - Line to line to line (A-B-C).
  - Line to line to line to ground (A-B-C-G).
- **Distance of the fault:** the distance between relay at station and point of fault along the transmission line has been changed at an incremental factor of every 5 km.

- **Resistance of the fault:** The fault resistance has been changed as follows: 0.01 ohm, 0.05 ohm, 0.1 ohm, 1 ohm, 5 ohm, 10 ohm, 25 ohm, 50 ohm, 100 ohm.

Three-Phase fault current signals have been generated through the simulation of the power system according to mentioned above.

To simulate the faults, the time for when the fault can occur can be set up, which phase(s) has a fault and where the fault occurs in the Matlab model as illustrated in Figure 5.4. Thus, it is possible to simulate several fault cases, e.g. various fault types, different fault distances and resistances. To perform the simulation process easily and effectively, Matlab Script has been used to generate faults and read fault data automatically. After the completion of simulation for each case, the fault current data is also saved automatically in a specified file.

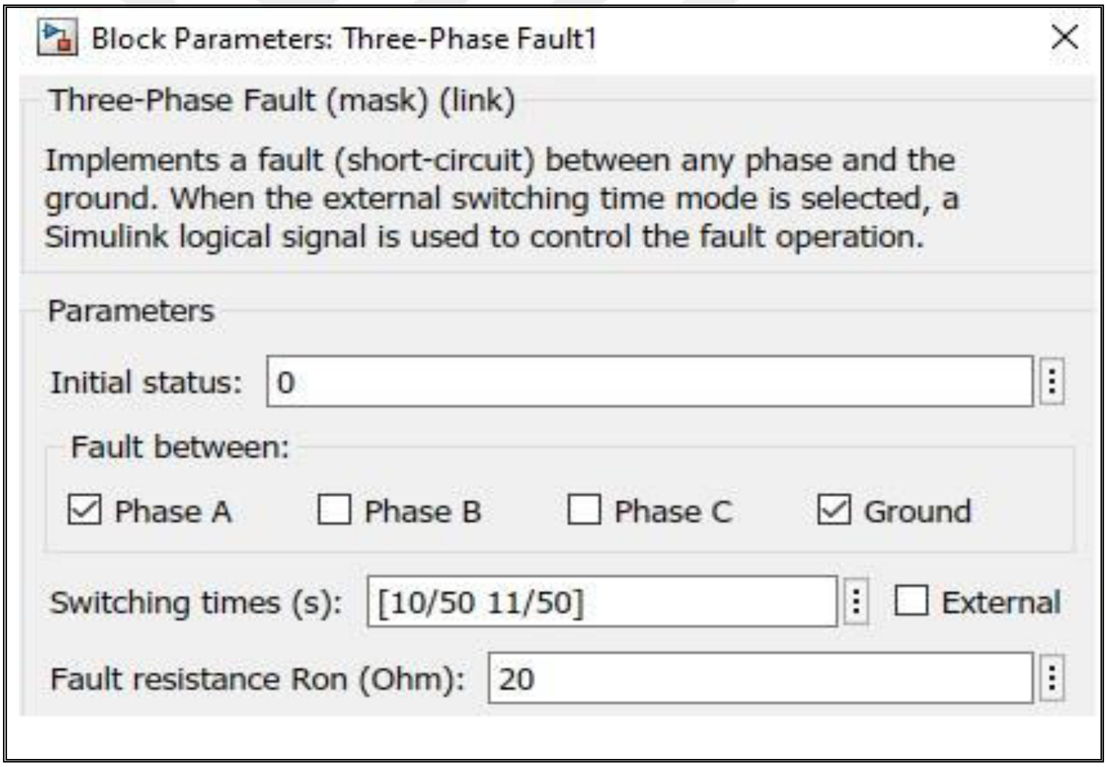


Figure 5.4 Setup block of fault parameters in Matlab

The figures 5.5 - 5.10 show the current waveforms at different types of faults:

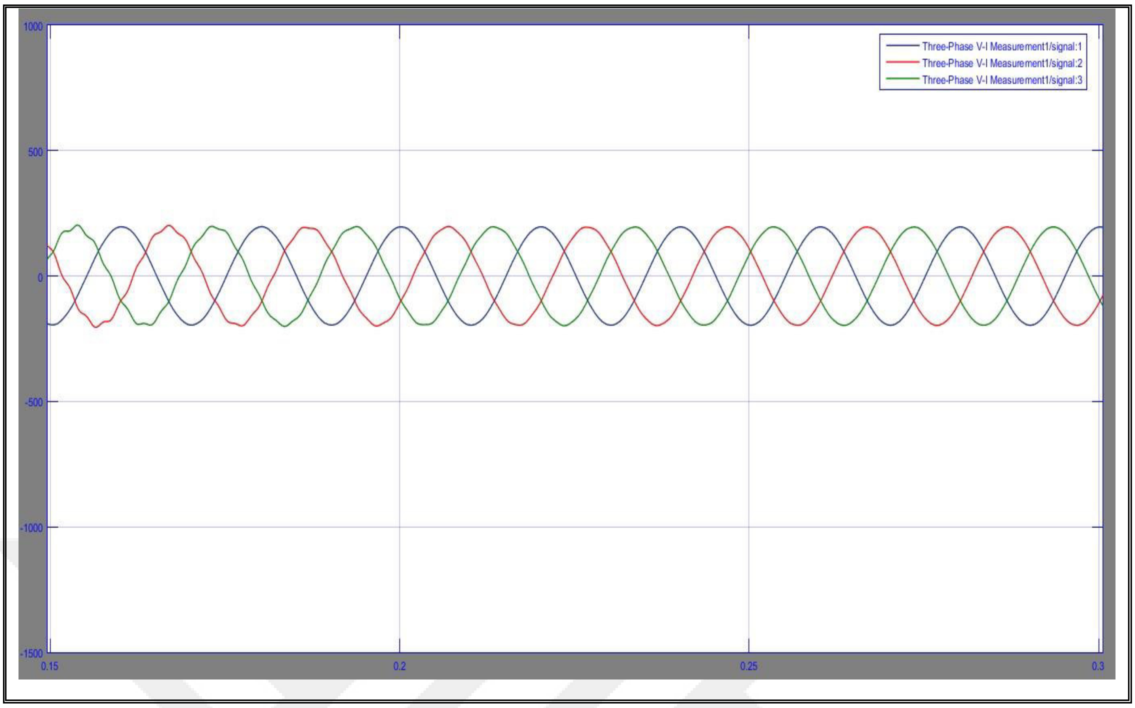


Figure 5.5 Current wave shape of normal conditions

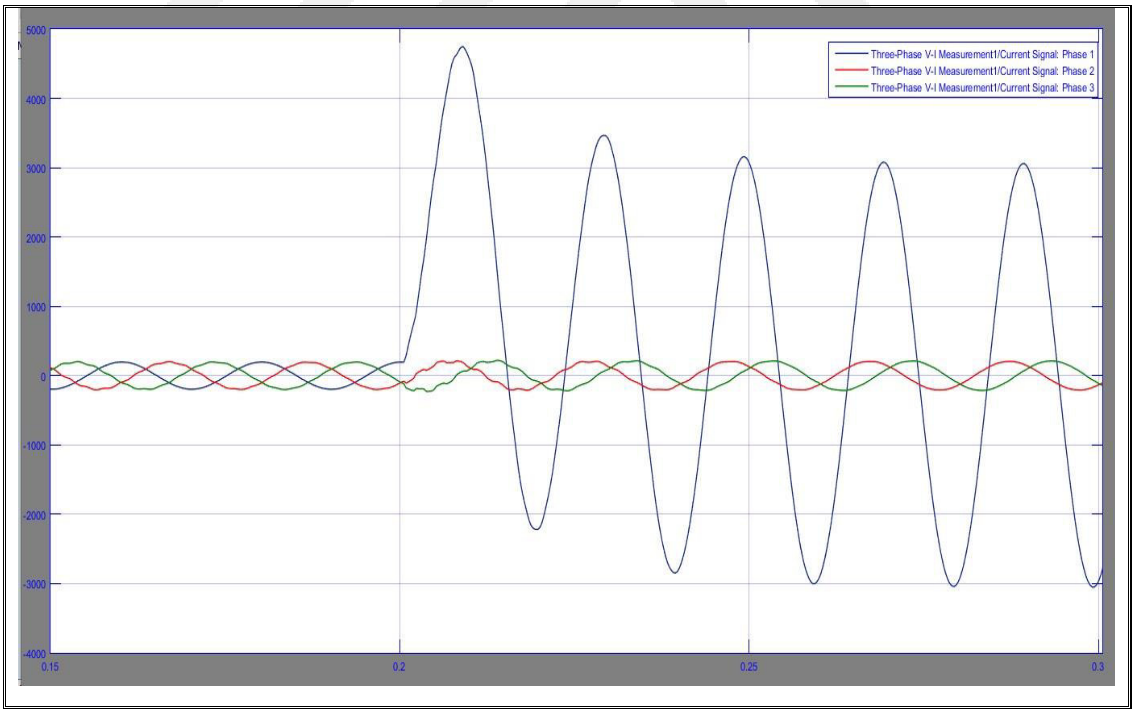


Figure 5.6 Current wave shape of phase to ground fault (A-G)

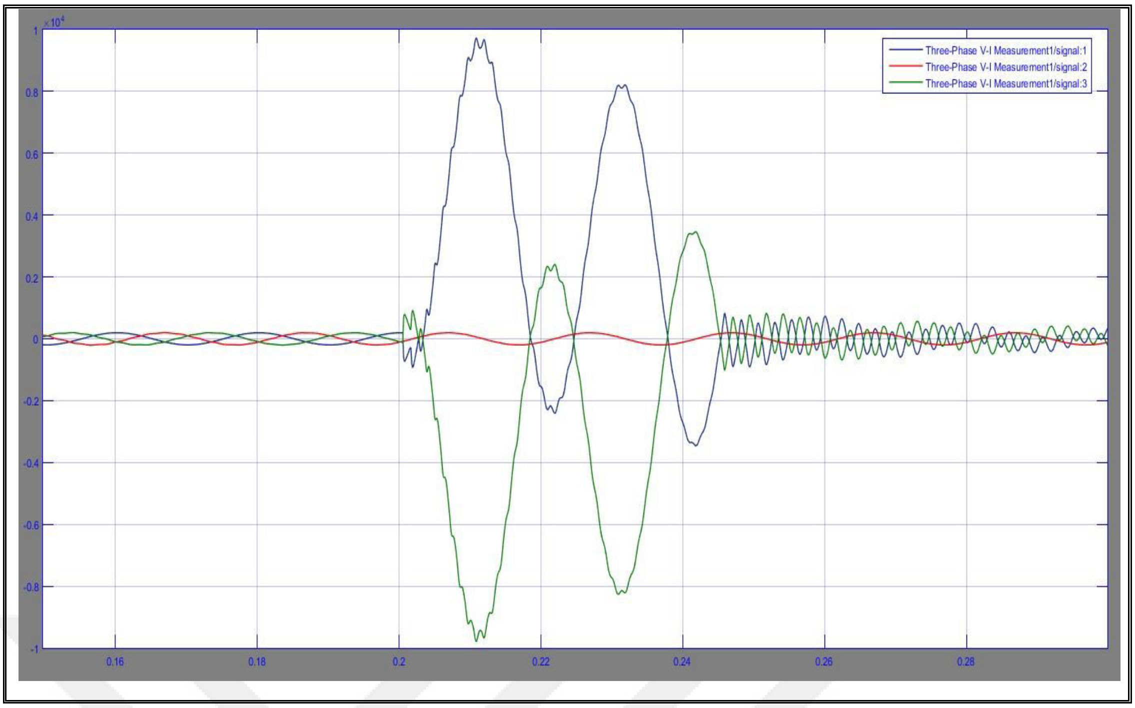


Figure 5.7 Current wave shape of phase to phase fault (A-C)

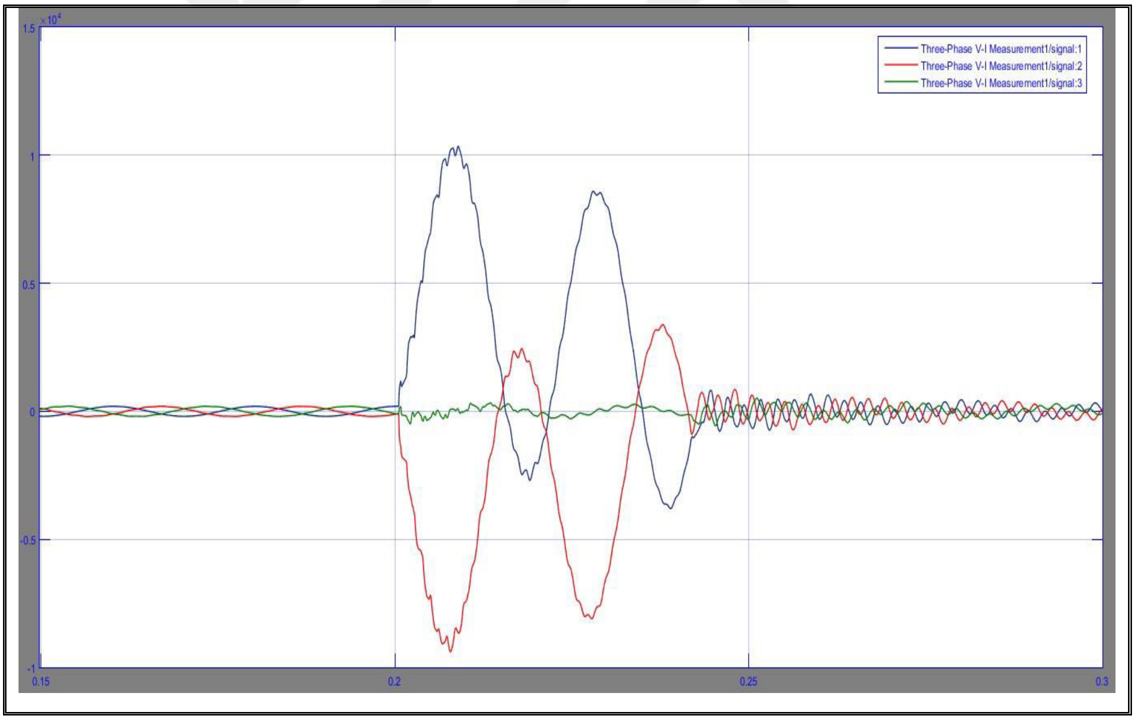


Figure 5.8 Current wave shape of 2 phase to ground fault (AB-G)

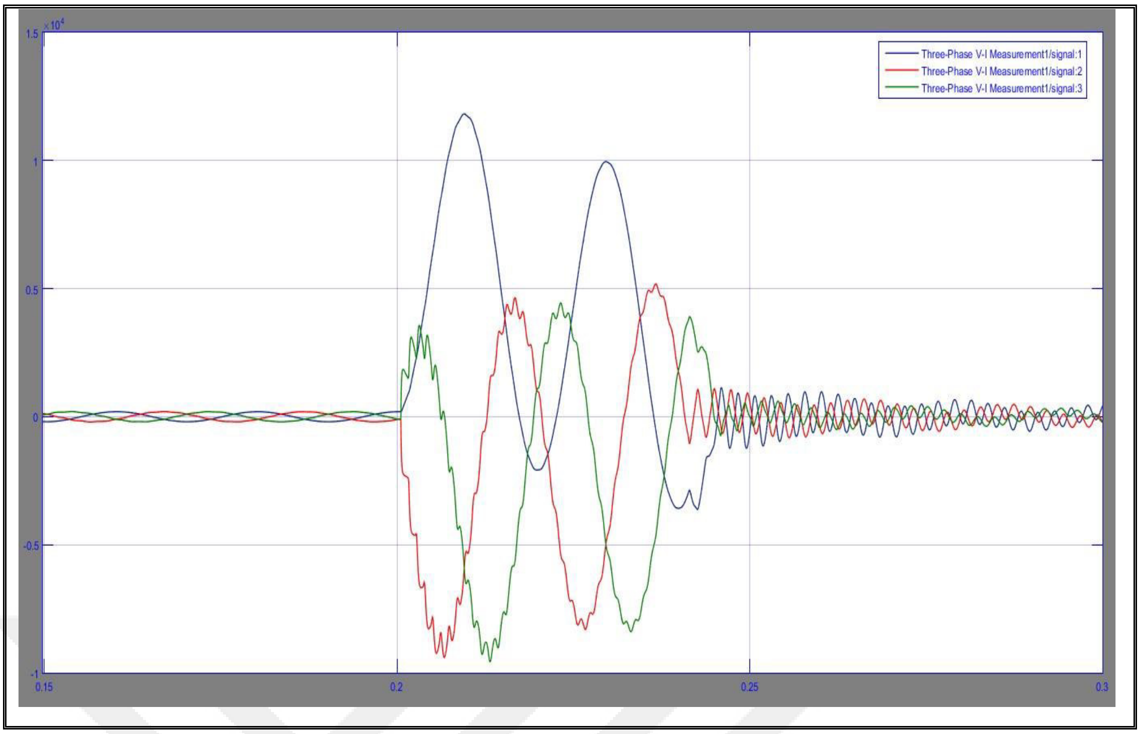


Figure 5.9 Current wave shape of 3 phase fault (ABC)

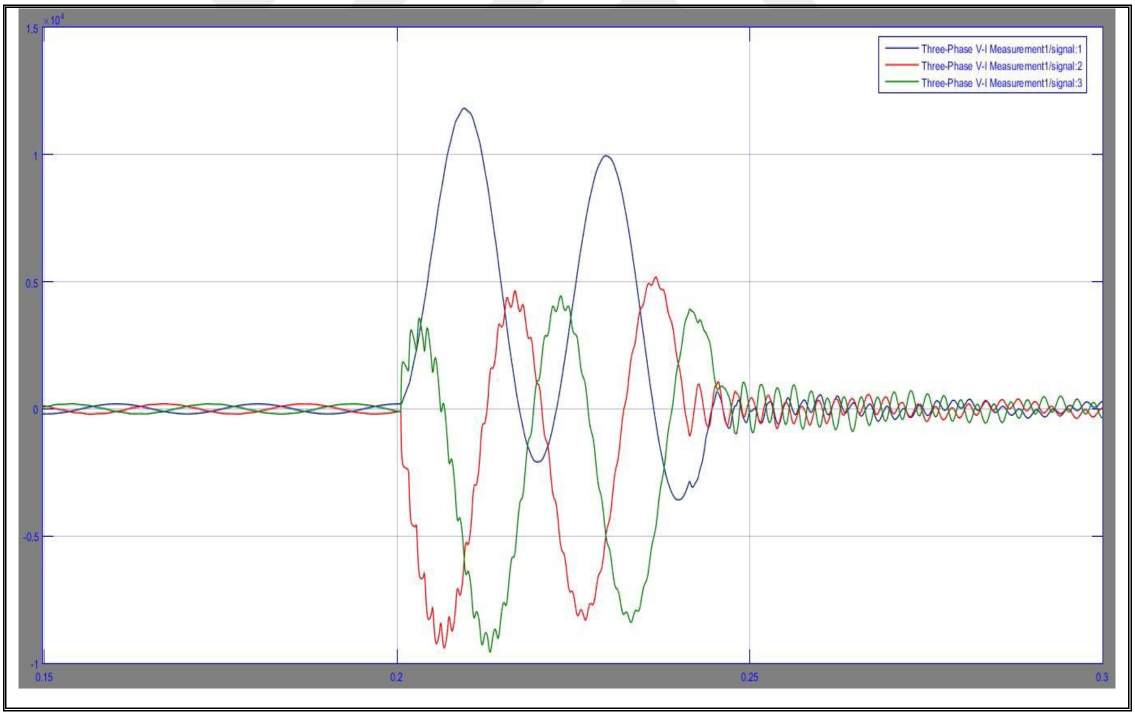


Figure 5.10 Current wave shape of 3 phase to ground fault (ABC-G)

For Training: The number of simulations performed to obtain the training dataset are 6270 (11 type of faults * 57 different places of faults * 10 various values of fault resistance) as illustrated in Table 5.2.

Table 5.2 Generation of training data set

Parameter	Training Data Set
Fault Type	**AG, BG, CG, AB, BC, AC, ABG, BCG, ACG, ABC, ABCG**
Fault Resistance (Ω)	**0.01, 0.05, 0.1, 0.5, 1, 5, 10, 25, 50, 100**
Fault Location (Km)	**5,10, 15, 20, 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, 95, 100, 105, 110, 115, 120, 125, 130, 135, 140, 145, 150, 155, 160, 165, 170, 175, 180, 185, 190, 195, 200, 205, 210, 215, 220, 225, 230, 235, 240, 245, 250, 255, 260, 265, 270, 275, 280, 285**
Training Cases	**6270**

The faults cases generated according to Table 5.2 are used to form training dataset used for training detector and classifier, for locater the fault cases generated for each type of fault are increased to be 1160 case to get better results during training.

For Testing: The number of simulations performed to obtain the test data are 616 (11 type of faults * 8 different places of faults * 7 various values of fault resistance) as illustrated in Table 5.3.

Table 5.3 Generation of testing data set

Parameter	Test Data Set
Fault Type	**AG, BG, CG, AB, BC, AC, ABG, BCG, ACG, ABC, ABCG**
Fault Resistance (Ω)	**0.03, 0.07, 1.2, 8, 30, 75, 90**
Fault Location (Km)	**13, 27, 48, 72, 143, 177, 222, 277**
Test Cases	**616**

5.4 Processing of Data

Sampling: Three-Phase fault current signals have been then sampled at sampling frequency of 12.8 KHz. Thus, there are 256 samples for per cycle (12,800 Hz/50 = 256 samples per cycle). The main reasons for selecting the 12.8 kHz sampling rate are:

Firstly, if the sampling rate is higher, there will be more burden on the digital microprocessor and more powerful microprocessor is needed.

Secondly, if the sampling rate is lower, the required high frequency information contained in transient signals cannot be properly captured.

The sampling process for fault signals is performed for each case during simulation through setup sampling rate and saving the sampled data to be processed later.

Discrete Wavelet Transform:

The DWT has been applied for analyzing the fault signals to capture the features contained in these signals. One cycle of the sampled current signals (x), after occurrence of the fault, have been decomposed by using DWT to compute the A and D coefficients of the four WT decomposition levels *(d1, d2, d3, d4, and A4)* as shown in Figure 5.11.

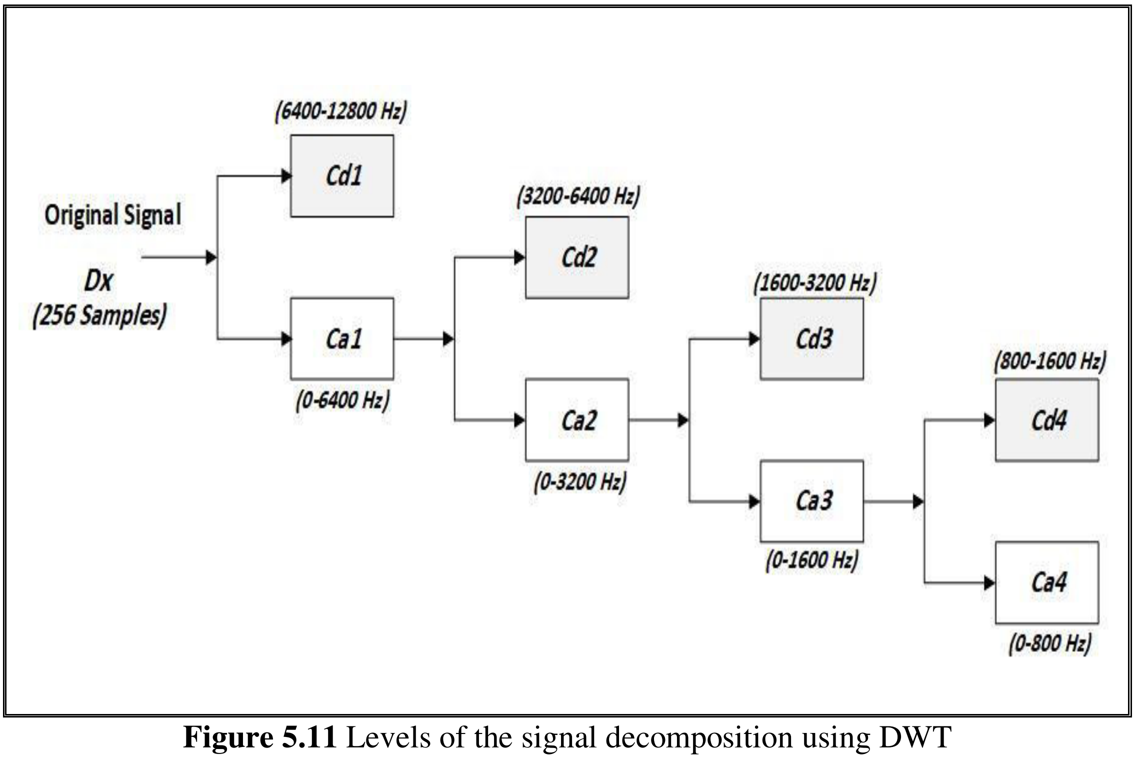


Figure 5.11 Levels of the signal decomposition using DWT

The mother wavelet used is the Daubechies 4 “db4”, since it is a type of Wavelet function that suits excellently when it is applied to identify the high frequency transients. Details of fault signals analysis are given in Table 5.4. [35].

Table 5.4 Details of wavelet transform

Parameter (N8)	Value
Mother wavelet	Daubechies 4 “db4”
Sampling frequency	12.8 KHz
Information analyzed	Detail at level 4 D4
Frequency band of D4	1.6 to 0.8 KHz
Number of samples per cycle	256
Data window length analyzed	One cycle/20 msec

Table 5.5 Frequency bandwidth for approximation and details

Decomposition Level	Frequency Band (Approximations)	Frequency Band (Details)
1	0 to 6.4 KHz	6.4 to 12.8 KHz
2	0 to 3.2 KHz	3.2 to 6.4 KHz
3	0 to 1.6 KHz	1.6 to 3.2 KHz
4	0 to 0.8 KHz	0.8 to 1.6 KHz

The analysis of the relevant signals is performed using MATLAB with the assistance of the “wavedec” instruction. It performs a multilevel wavelet analysis in a single dimension (1-D), for which it uses a specific wave type “wname” or a specific set of wavelet decomposition filters.

[C, L] = wavedec (X, N, ‘wname’) yields the wavelet analysis of the signal X, for a analysis level N, where N is a strictly positive integer value and a specific wavelet type “wname” [36].

The outcome of the WT analysis is a vector consisting of the wavelet coefficients, which is then used to calculate the energy of the WT coefficients of the (Ds) and the (As) for each of sampled current signal, in addition to all other features such as R.M.S, Mean, Max, Min, and SD values.

Features Extracting:

A reduction of the hidden layers and number of neurons for each hidden layer play a critical part an in improving the ANN performance, this can be performed by extracting features contained on useful information.

After implementing the DWT analysis, WC are obtained, but the WC are not distinct enough to be directly applied in the decision-making process related to certain fault; therefore, the signal features are firstly extracted from these coefficients and computed using MATLAB script with the assistance of the next instructions:[33]

[Ea,Ed] = wenergy(C,L) computes Ea, which is the proportion of energy relating to the approximation and Ed, which is the vector containing the proportions of energy relating to the details [37].

Standard Deviation = std(A) computes the SD of the elements of A.

Mean = mean(A) computes the mean of the elements of A.

Max = max(A) computes the maximum elements of A.

Min = min(A) computes the minimum elements of A.

RMS = rms(A) computes the root-mean-square (RMS) of A.

The DWT decomposition steps, are repeated for various types of faults, as listed in Table 5.3, then the features for each fault type (i.e. *Ed, Ea,* etc.) are calculated and saved in one array, which contains all fault types features.

CHAPTER VI

EXPERIMENTAL FINDINGS AND DISCUSSION

6.1 Neural Networks Application

NN is the tool applied to detect, classify, and localize the faults in this thesis. The simulation results are applied as inputs to ANN to move toward the desired results. ANN structures must be modified by changing number of hidden layers, number of neurons at each hidden layer, training functions, learning rate, weights, and biases to obtain the best performance and least mean square error. Each of detector, classifier, and locator has a special structure of ANN.

The NN must be trained using data resulting from the HVTL model simulation before applying to perform the required performance. After identifying of NN structure, different set of model simulation outputs are applied to the neural network to test and evaluate its performance.

In this chapter, an integral method based on NN has been explained in detail for designing the device that performs distance protection functions in the event of faults on TL in a PS. To perform that, the problem has been addressed in three major steps, namely FD, FC, and FL.

6.2 Overview of the Training and Testing Process

The procedure needed for the utilization of ANNs for any purpose involves two essential steps which are training and testing, each of which requires a separate dataset.

Training: is a method of ANN learning using a dataset of training, then adjusting weights and biases accordingly. The training dataset consists of a group of input-output couples that are fed into the NN, and during the training process, the weights and biases of the NN are modified to reduce the performance function to a minimum and increase the accuracy. The performance function commonly employed for BPNN is MSE and will be applied in this work.

The complete set of training data generated in previous chapter is randomly divided by Matlab Script into three different sets namely the training set (70%), validation set (15%), and the testing set (15%).

Testing: the next essential step to be implemented before the operation of NN is to test a trained NN following completion of the training process. Testing the ANN is a very critical stage in order to guarantee that the learned network can generate the ideal results as new data is entered into it. The set of testing data generated in previous chapter is applied to the trained NNs

The applied techniques for evaluating and testing the trained NN are:

- Plotting linear regression between the actual output values and target output values, this technique is mainly used for FL.
- plotting the confusion matrix that clarifies number of cases classified correctly and number of cases classified wrongly, this technique is mainly used for FD & FC.

6.3 Fault Detection (FD)

For detecting fault, various structures of MLNN have been examined. The factors that play a role in designing the best configuration are the NN structure, the learning methodology employed and the training dataset size.

The back-propagation algorithm has been chosen to be utilized for training. The selection of the NN structure is very critical due to this does not only reduce the training time but also greatly boosts the ability of the NN to exemplify the problem that is in the process of being addressed. There is no certain principle that can instruct the number of hidden layers and the number of neurons for each hidden layer; therefore, they are determined by experimentation and testing, considering some general rules to be followed.

6.3.1 Training of Fault Detector Neural Network

To train NN to be able to detect the occurrence of the fault on TL, training dataset generated in fifth chapter for 6270 different cases for different types of faults at several locations, and various fault resistances and some of no-fault cases have been used as inputs to the NN. Training data set is labeled in ninth inputs and one output for each case as it is illustrated in Figure 6.1. The output of the NN states whether a fault has been detected or not: a value of '1' means a faulty condition, while a value of '0' points out an unfaulty condition, Table 6.1 shows the target output for detector.

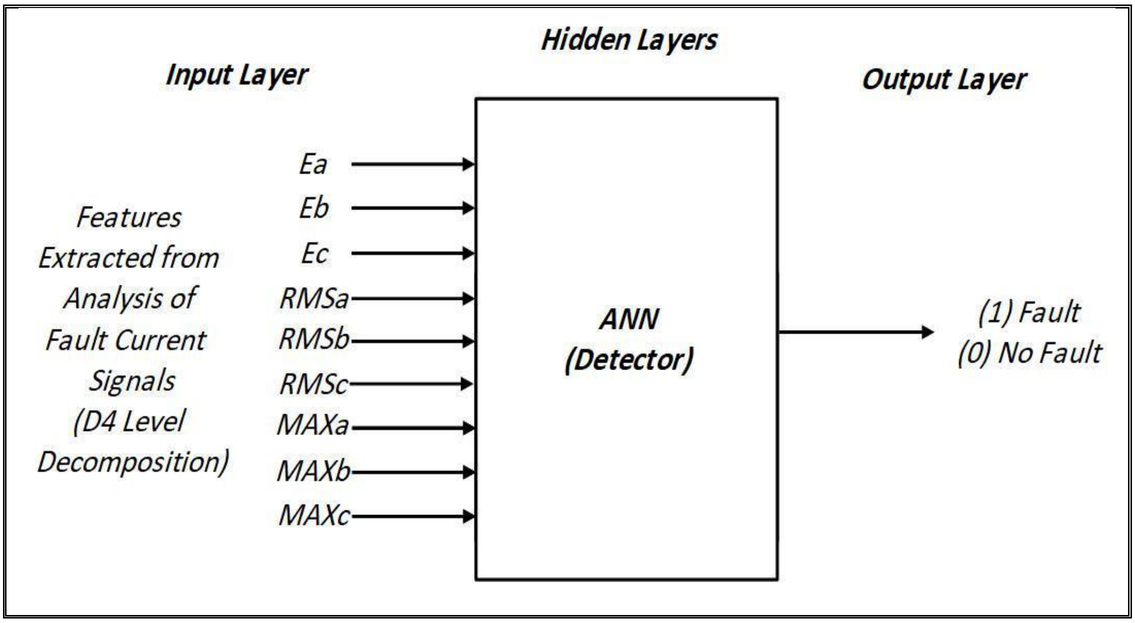


Figure 6.1 Single line of fault detector

Table 6.1 Target output of fault detector

Fault Type			Target Output (Detection)
1	**No Fault**	**NF**	0
2	**Fault**	**F**	1

To obtain the best performance, several simulations for many NN structures (with varying number of hidden layers and neurons at each hidden layer) have been performed, then it has been found that the desired network to give satisfied results has three hidden layers with 6 neurons at each hidden layer. Figure 6.2 shows the ANN structure for the FD with its performance. The MSE of 6.88×10^{-5} is achieved. This value is considered acceptable and delivers very good results.

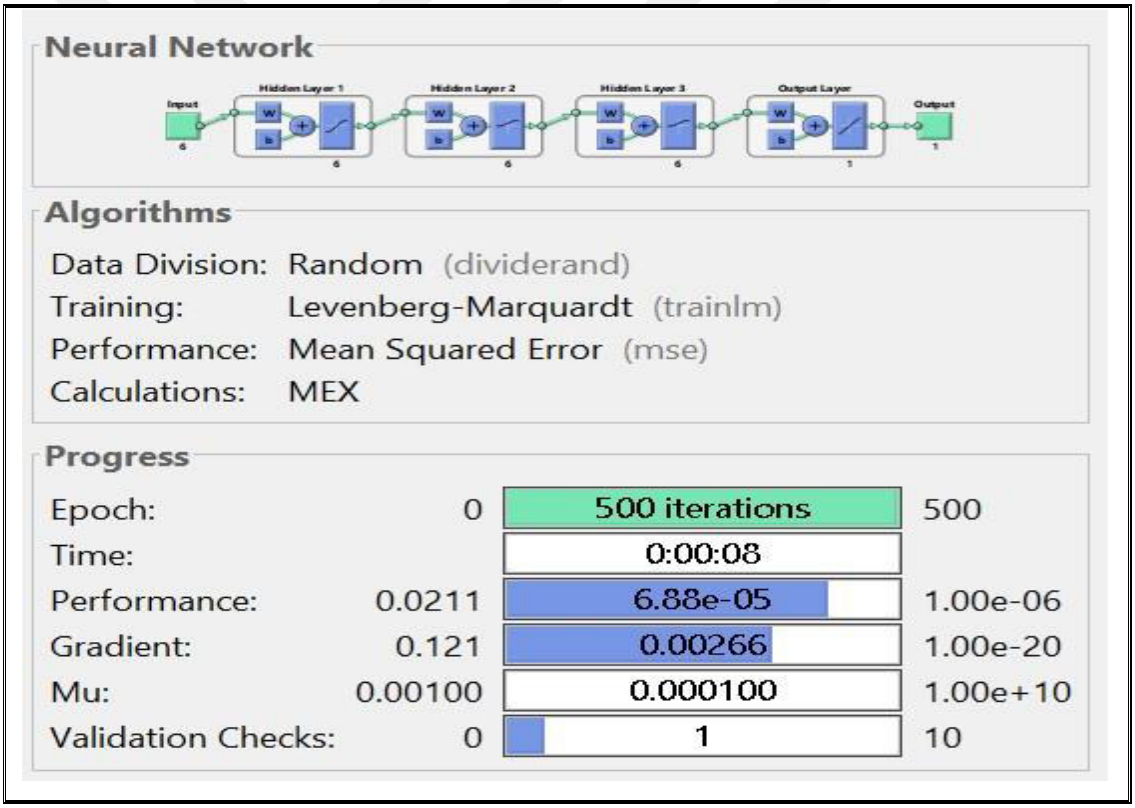


Figure 6.2 Structure of ANN for FD (6-6-6-6-1)

To increase accuracy, the number of features used as inputs to ANN has been increased to be nine which are values of Energy, RMS, and Max for the fourth level detail coefficients resulting from analysis of the three phase current signals by using DWT. Figure 6.3 & Figure 6.4 show the selected 9-6-6-6-1 NN structure and its performance when employing the Levenberg-Marquardt training function. With the selected structure, the MSE of 9.66×10^{-7} were performed.

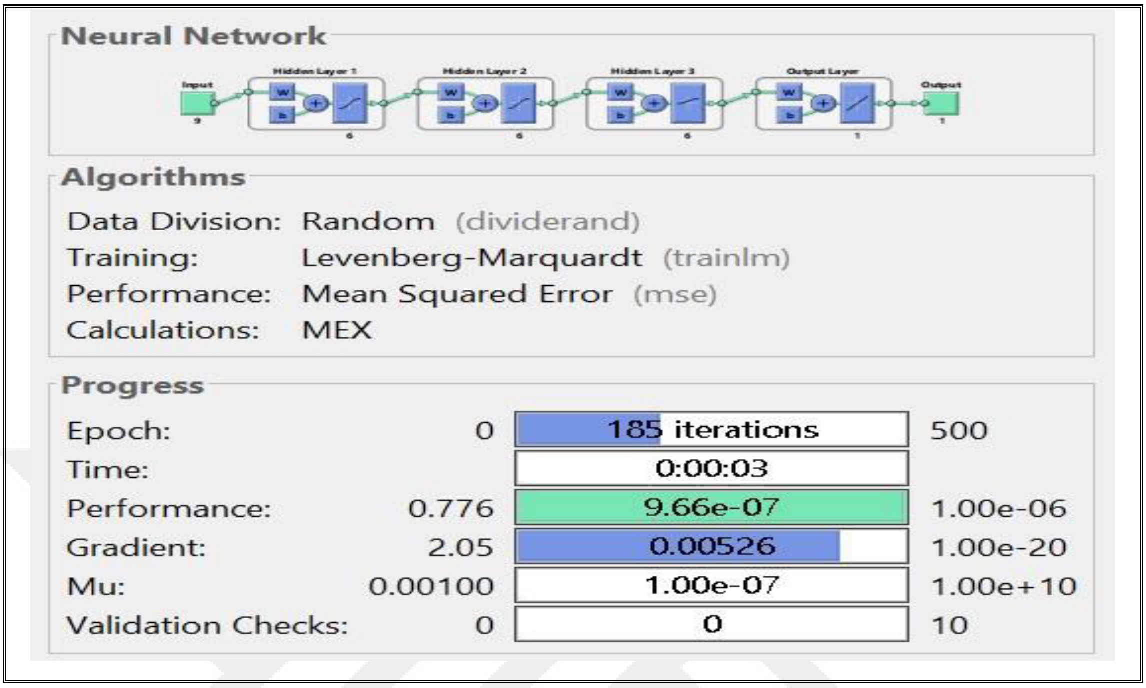


Figure 6.3 Structure of ANN for FD (9-6-6-6-1)

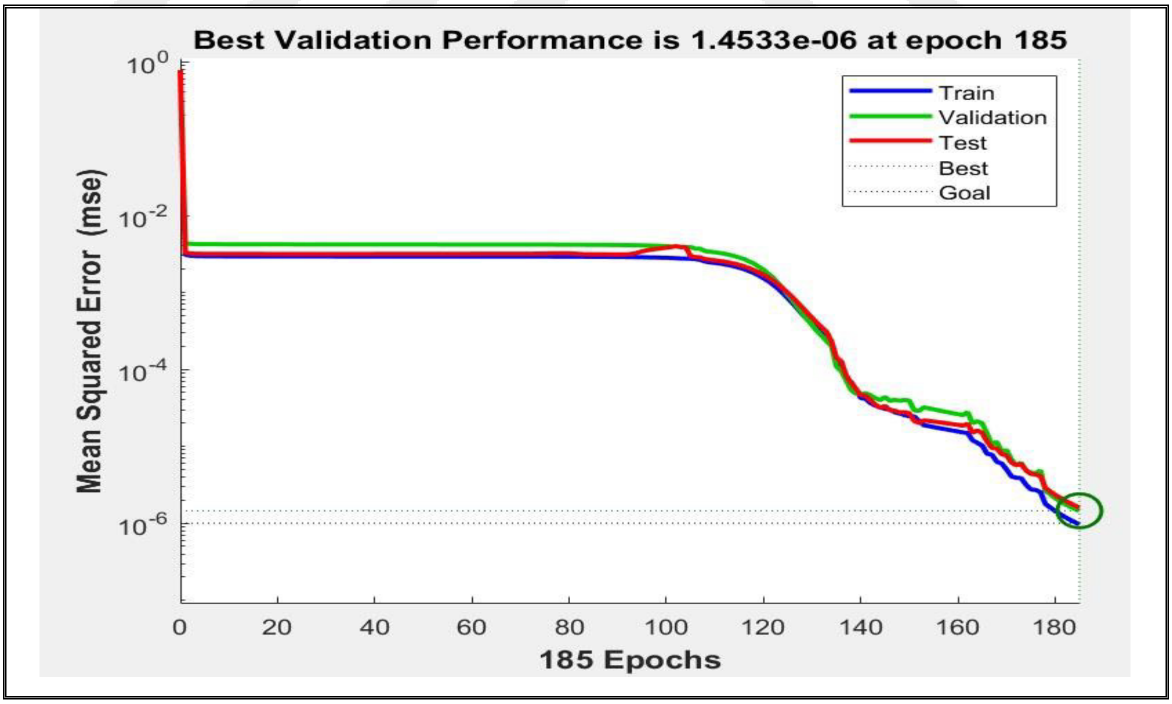


Figure 6.4 Training performance curve of ANN for FD (9-6-6-6-1)

Note: MATLAB script has been used for training of ANNs, through which data is divided into three groups; training data; to train and find appropriate NN weights and biases, validation data; used to validate the trained NN, and test data; used to test NN performance (represents by 70%, 15%, and 15% of entire data, respectively).

6.3.2 Evaluating of Fault Detector Neural Network

After completing the training process, the performance of ANN is evaluated through the following criteria:

- plotting the linear regression scheme that links the target output to the actual output of ANN as shown in Figure 6.5.

  The degree of association between target and actual values of ANN output is measured by a correlation coefficient (r), it is measured on a scale that varies from + 1 through 0 to - 1. Complete correlation between two variables is represented by 1 and complete absence of correlation is represented by 0. From Figure 6.5, the correlation coefficient equals to 0.9999 in our case which indicates excellent correlation between target and actual output values.

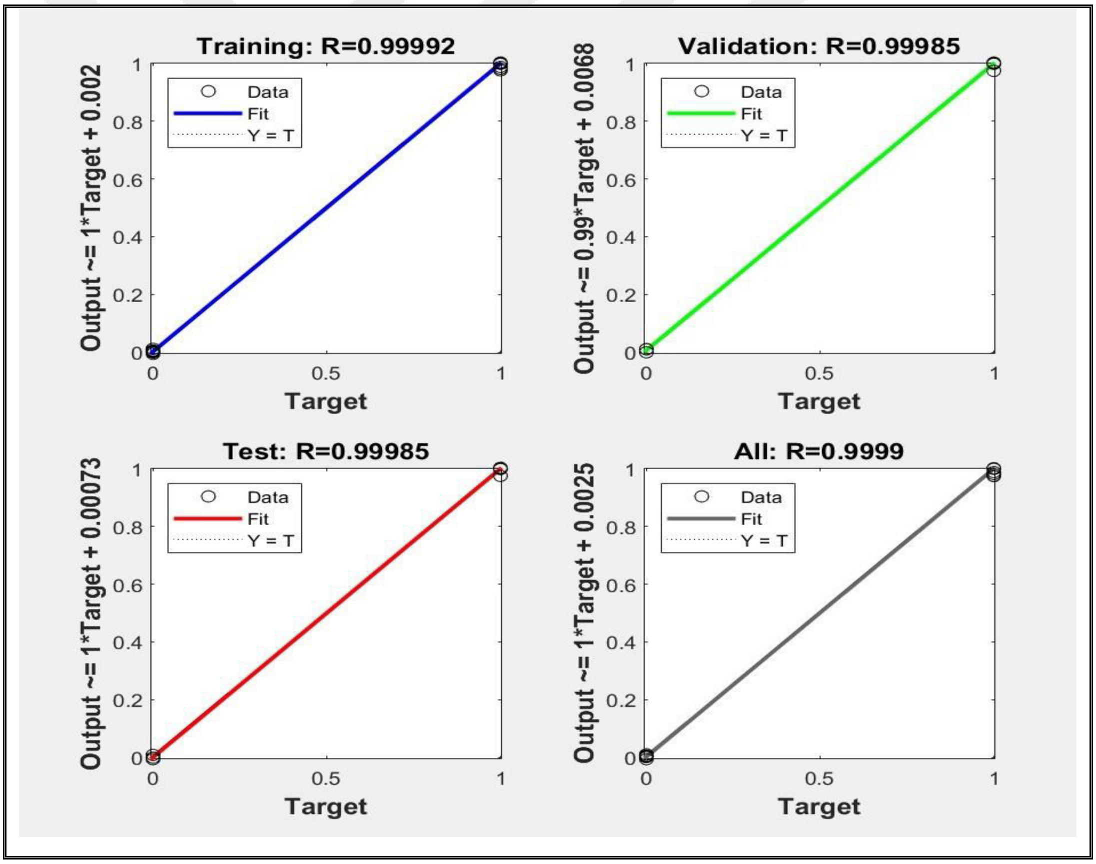


Figure 6.5 Regression scheme of training process for FD (9-6-6-6-1).

- plotting all confusion matrix that includes the three stages of training, validation, and testing as shown in Figure 6.6. The diametrical cells in green point out the numeral of cases classified properly by the NN and the off-diametrical cells which are in red point out the numeral of cases classified mistakenly by the ANN. The last cell in gray points out the overall percentage of cases classified properly written in green and the overall percentage of cases classified wrongly written in red. From Figure 6.6, it is very clear that the selected NN has 100 percent accuracy in detecting the faults.

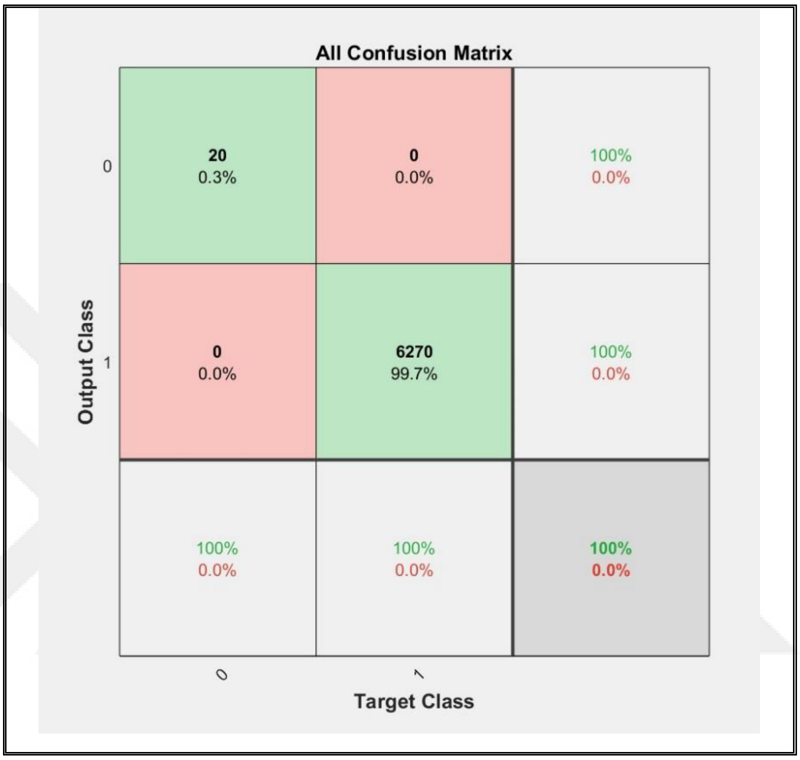


Figure 6.6 All confusion matrix for training data of FD.

6.3.3 Testing of Fault Detector Neural Network

Once the training of the selected NN is completed, its performance based on training data set is evaluated, it is tested using a separate data that is called testing data set, testing data set generated in fifth chapter for 616 different fault conditions and some of no-fault cases that ANN have not been trained are applied as inputs for testing the NN.

The ANN structure is simulated for 626 cases (at different locations for various resistances of faults). Figure 6.7 illustrates the testing confusion matrix and the regression plot of the relation between actual and target outputs of FD for testing data set.

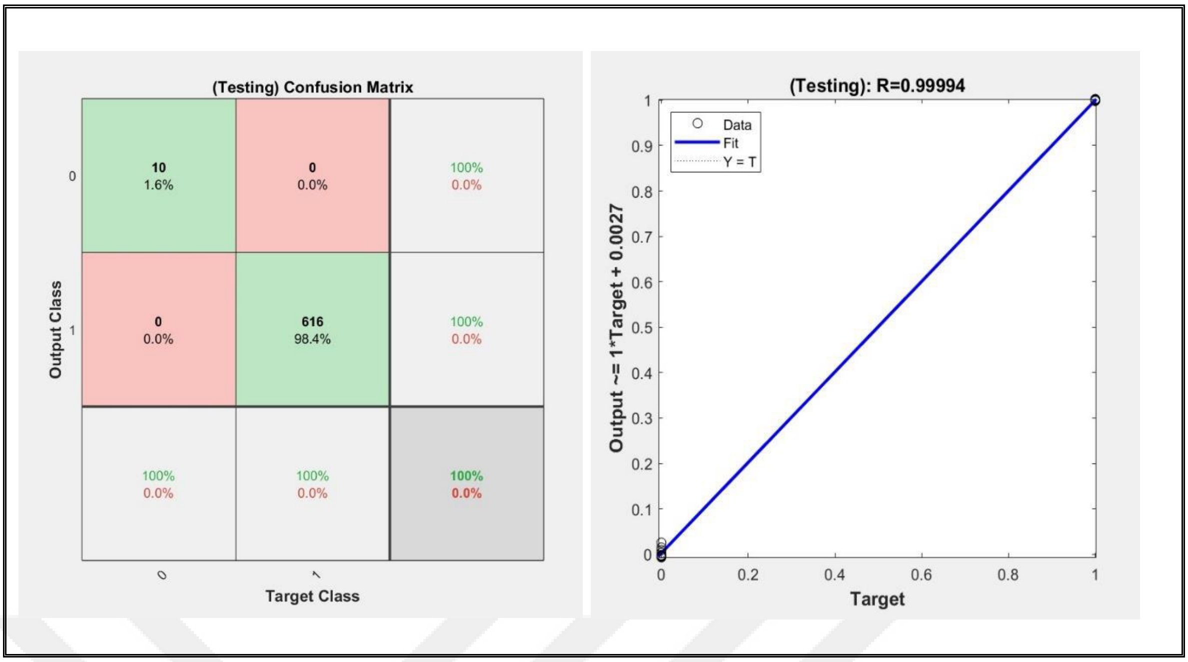


Figure 6.7 All confusion matrix and regression scheme for testing data of FD

The error rate between NN actual and the target output values was very small, the NN output can be approximated to get error equals zero for all tested cases. Thus, the NN can, with very high accuracy, distinguish a normal situation from a fault situation on a HVTL.

The Table 6.2 shows examples of ANN output selected for FD when testing data set is fed as inputs to the FD.

Table 6.2 Examples of ANN output selected for FD

Fault Conditions					Target Output (Detection)	Actual Output of ANN (Detector)
N.	**Type of Fault**		**Fault resistance (Ω)**	**Fault Distance (Km)**		
1	No Fault	NF	-	-	0	0.0253
2		NF	-	-	0	0.0147
3		NF	-	-	0	0.0088
4		NF	-	-	0	-0.0020
5		NF	-	-	0	-0.0046
6		NF	-	-	0	-0.0056
7	1 phase to Ground	A-G	0.03	27	1	0.9988
8		A-G	1.2	13	1	0.9986
9		B-G	0.03	277	1	0.9999
10		B-G	1.2	177	1	0.9999
11		C-G	1.2	27	1	0.9999
12		C-G	30	13	1	0.9999
13	2 Phase	AB	1.2	177	1	1.0003
14		AB	30	48	1	0.9995
15		BC	0.07	177	1	1.0001
16		BC	8	222	1	1.0003
17		AC	8	27	1	0.9998
18		AC	75	48	1	1.0003
19	2 Phase to Ground	AB-G	0.07	13	1	1.0001
20		AB-G	30	143	1	1.0001
21		BC-G	75	72	1	0.9998
22		BC-G	90	72	1	0.9998
23		AC-G	0.07	277	1	1.0000
24		AC-G	8	27	1	1.0002
25	3 Phase	ABC	1.2	13	1	1.0001
26		ABC	1.2	143	1	1.0007
27		ABC	8	13	1	1.0003
28		ABC	30	13	1	0.9994
29		ABC	75	222	1	1.0002
30		ABC	90	72	1	0.9994
31	3 Phase to Ground	ABC-G	1.2	13	1	1.0006
32		ABC-G	8	13	1	1.0002
33		ABC-G	30	13	1	0.9994
34		ABC-G	75	177	1	1.0002
35		ABC-G	90	13	1	0.9997
36		ABC-G	90	222	1	1.0001

6.4 Fault Classification (FC)

Classifying the fault type is very important for HVTL for the following reasons:

- Recording the fault type in station faults record to find out the number of times the same fault occurs during a specific period, and in the event that the same fault occurs repeatedly, the causes of this must be investigated and addressed by the technical team.
- Knowing the type of fault facilitates and saves time for detecting, and thus significantly reduces the time needed to carry out the required maintenance.
- Accurately identifying the type of fault, reduces the number of technical team required to check the faulty TL path, thus decreases the cost considerably.

6.4.1 Training of Fault Classifier Neural Network

After the fault on the TL is detected, the next step is to identify the type of fault. Ten types of line faults are studied and each type is represented by special code. Table 6.3 illustrates the fault types and their own codes.

To train NN to classify the type of the fault on TL, the same data set used for training the detector generated in the fifth chapter for 6270 various cases for distinct types of faults at several locations, and various fault resistances were employed for training the classifier ANN. Training data set is labeled in twelve inputs and four outputs as it is illustrated in Figure 6.8, three of which represent the state of three phase (A, B, &C) and one output represent state of ground line, the output is either a 0 or 1, a value of '1' means a faulty condition while a value of '0' indicates an unfaulty condition. According to state each output and based on coding indicated in Table 6.3, the type of fault is identified. The output layer gives ten cases, each of them classifies certain fault.

Multilayer NN and the back-propagation learning strategy are used in designing and developing the classifier so that the NN can accurately distinguish among the ten possible types of fault.

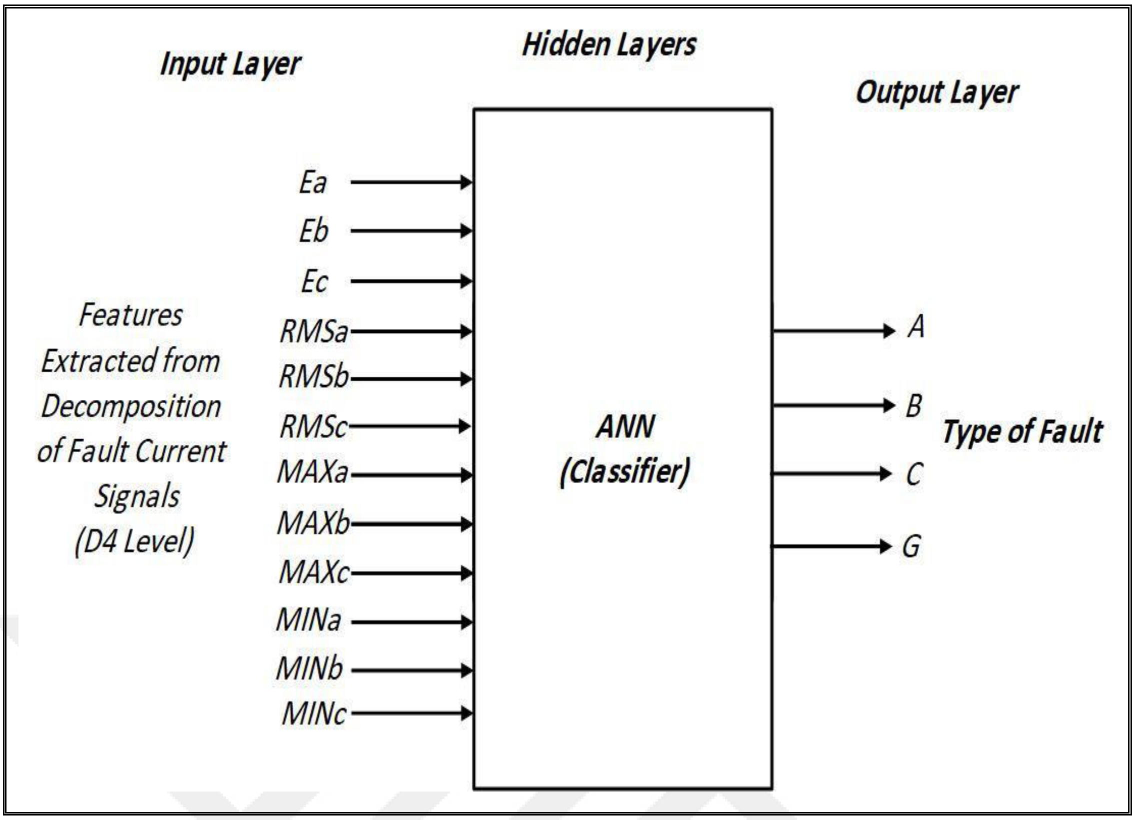


Figure 6.8 Single line of fault classifier

Table 6.3 Target output of fault classifier

Fault Type			**Target Output (Classification)**			
			A	**B**	**C**	**G**
1	**1 phase to Ground**	**AG**	1	0	0	1
2		**BG**	0	1	0	1
3		**CG**	0	0	1	1
4	**2 Phase**	**AB**	1	1	0	0
5		**BC**	0	1	1	0
6		**AC**	1	0	1	0
7	**2 Phase to Ground**	**ABG**	1	1	0	1
8		**BCG**	0	1	1	1
9		**ACG**	1	0	1	1
10	**3 Phase**	**ABC**	1	1	1	0
11	**3 Phase to Ground**	**ABCG**	1	1	1	0

Figure 6.9 & Figure 6.10 show the selected 12-10-10-10-4 NN structure and its performance curve for classifier 12-10-10-10-4 (12 neurons in the input layer, 3 hidden layers with 10 neurons for each and four neurons in the output layer). It can be noticed that the superior performance in terms of the MSE value by the end of the training process is 0.00170 which is considered an acceptable result.

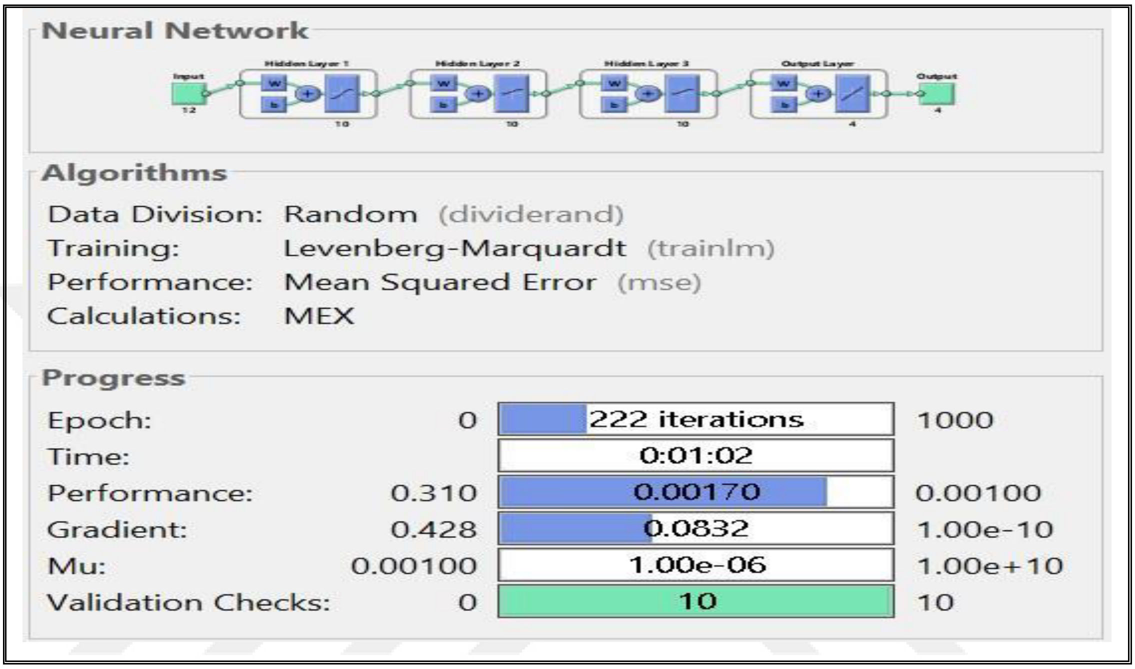


Figure 6.9 Structure of ANN for FC (12-10-10-10-4)

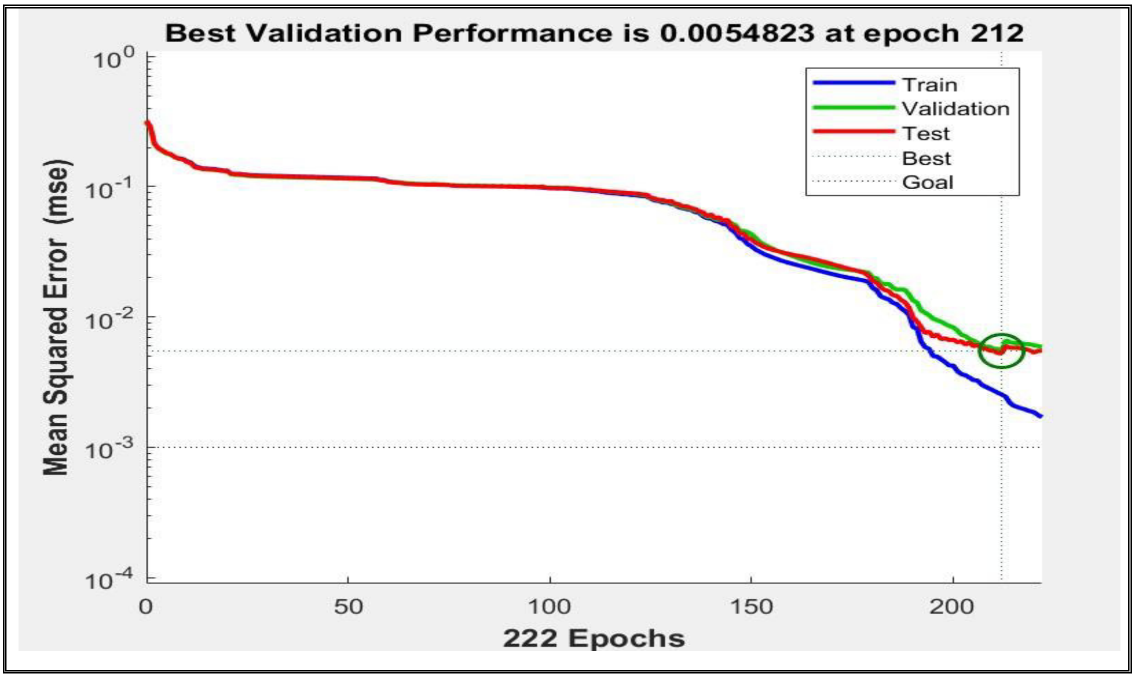


Figure 6.10 Training performance curve of ANN for FC (12-10-10-10-4)

To increase the accuracy, after implementing many trails and modifications to the structure of the NN, so the number of features used as inputs to ANN has been increased to be fifteen which are values of Energy, RMS, Max, Min, and SD for the fourth level detail coefficients, also the number of neurons at each hidden layer has been increased to be twelve for each. Figure 6.11 & Figure 6.12 show the selected 15-12-12-12-1 NN structure and its performance curve when employing the Levenberg-Marquardt training function. With the selected structure, the MSE of 9.63×10^{-4} were achieved.

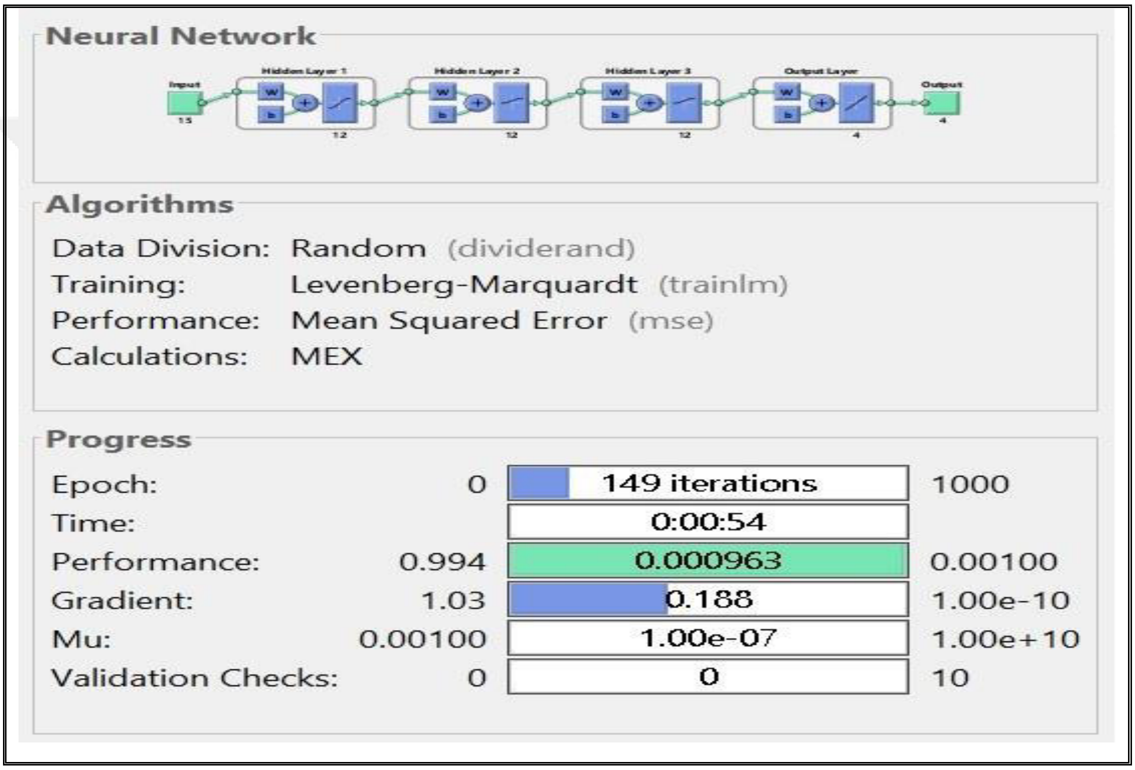


Figure 6.11 Structure of ANN for FC (15-12-12-12-4)

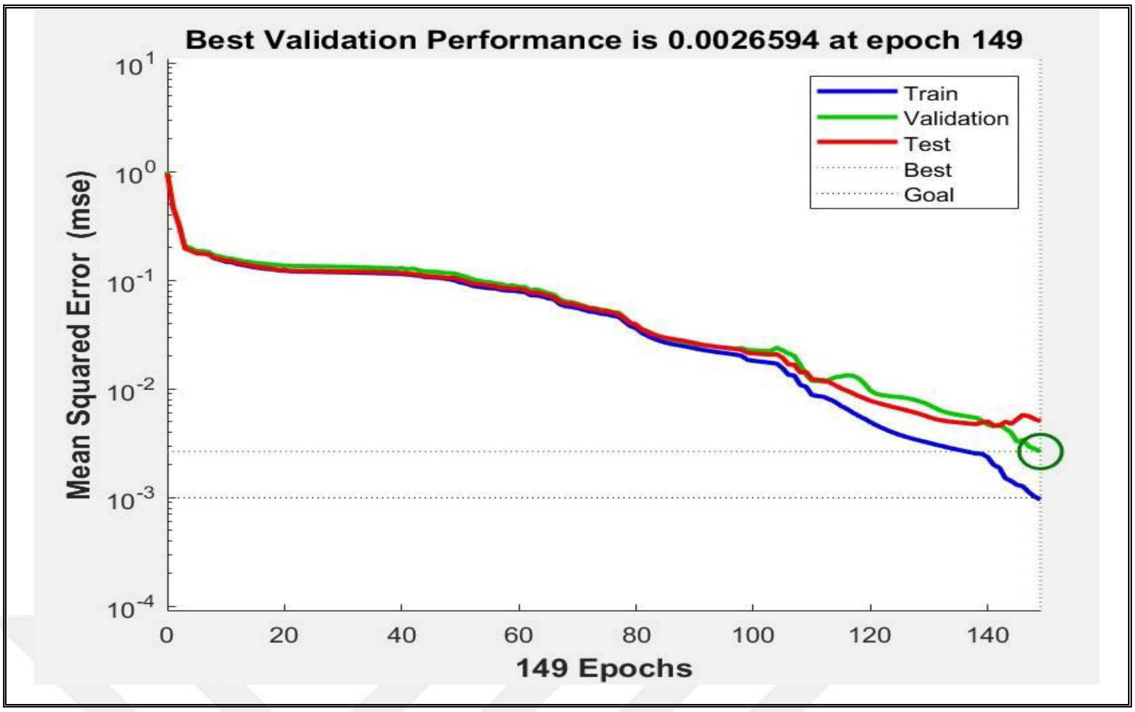


Figure 6.12 Training performance curve of ANN for FC (15-12-12-12-4)

6.4.2 Evaluating of Fault Classifier Neural Network

After completing the training process, the performance of ANN for classifier is evaluated through the following criterion:

- plotting all confusion matrix that includes the three phases of training, validation, and testing as shown in Figure 6.13.

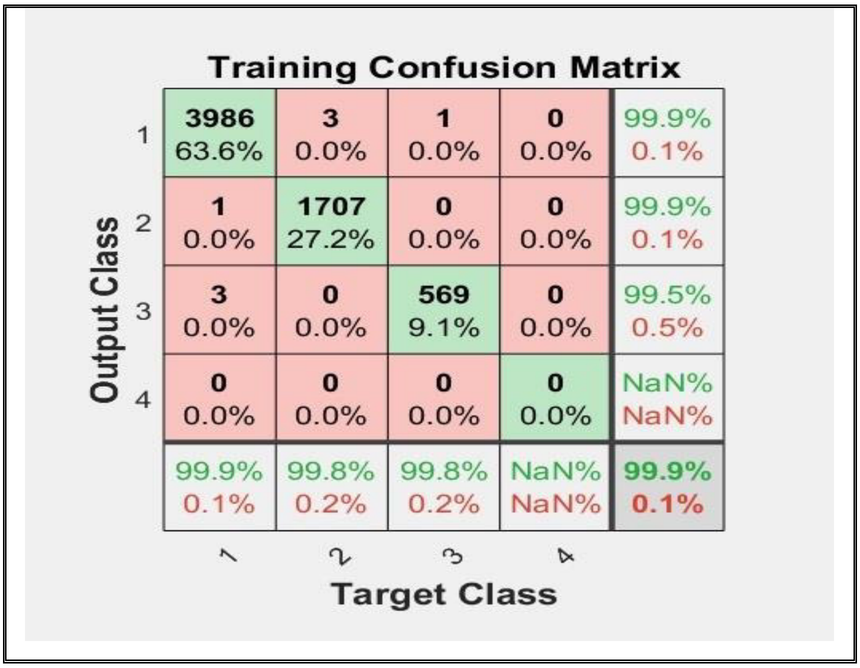


Figure 6.13 Training data confusion matrix of ANN for FC (15-12-12-12-4)

6.4.3 Testing of Fault Classifier Neural Network

Once the training process of the selected NN is completed, its performance based on training data group is evaluated, it is tested using a separate data that is called testing data group, generated in fifth chapter for 616 different conditions from ones used to generate training data is used as inputs for testing the NN.

To achieve more accurate and stable results for the FC, the hard limit TF is used to force a neuron output to be 1 if its input reaches a threshold, otherwise it outputs is 0. The threshold employed is o.5 so:

- if the output is less 0.5, then the output is reset to be 0
- if the output is more than 0.5, then the output is reset to be 1

The ANN structure is simulated for 616 cases (at different locations for various resistances of faults). Figure 6.14 illustrates the confusion matrix between actual and target outputs of FC for testing data.

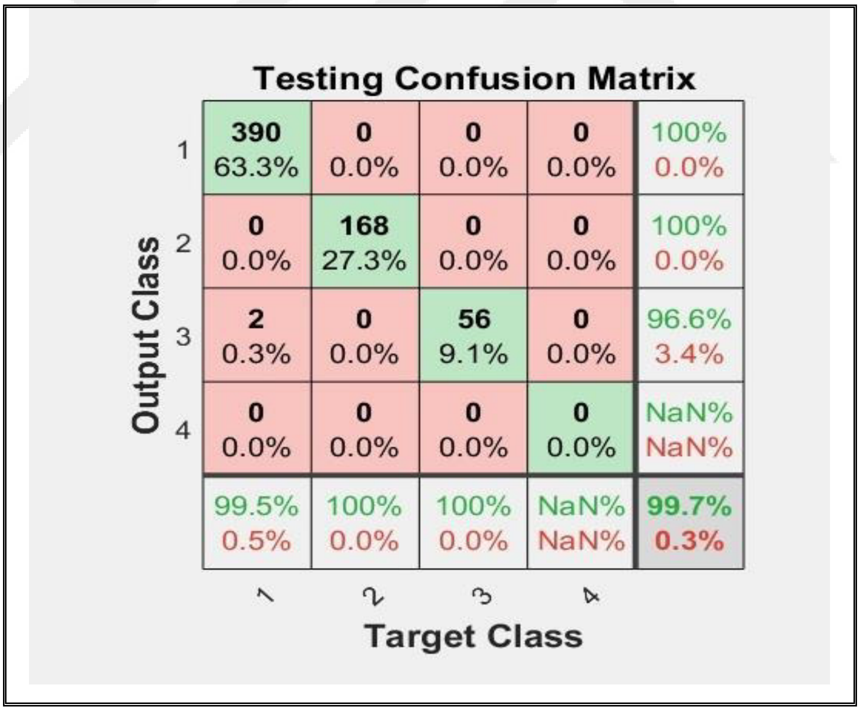


Figure 6.14 Testing data confusion matrix of ANN for FC (15-12-12-12-4).

Table 6.4 shows samples of ANN output selected for FC when testing data is fed as inputs to the FC.

Table 6.4 Samples of ANN output selected for classifier

Fault Conditions					Actual Output of ANN (Classifier)			
N.	**Type of Fault**		**Fault resistance (Ω)**	**Fault Distance (Km)**	**A**	**B**	**C**	**G**
1	1 phase to Ground	AG	0.03	27	0.9905	-0.0226	-0.0013	1.0118
2		AG	1.2	13	1.0059	-0.0232	0.0255	0.9411
3		BG	0.03	277	0.0046	1.0007	-0.0003	1.0028
4		BG	1.2	177	0.0021	1.0004	-0.0018	1.0039
5		CG	1.2	27	0.0044	0.0041	0.9959	0.9958
6		CG	30	13	0.0096	0.0063	0.9937	0.9937
7	2 Phase	AB	1.2	177	1.0041	0.9986	-0.0125	-0.0185
8		AB	30	48	1.0074	0.9929	-0.0060	-0.0114
9		BC	0.07	177	0.0183	1.0000	1.0004	-0.0124
10		BC	8	222	0.0299	0.9988	1.0022	0.0028
11		AC	8	27	0.9873	-0.0096	1.0112	-0.0014
12		AC	75	48	1.0802	-0.0009	1.0231	-0.1090
13	2 Phase to Ground	ABG	0.07	13	0.9981	1.0016	-0.0006	1.0010
14		ABG	30	143	1.0001	0.9985	0.0026	0.9976
15		BCG	75	72	0.0047	1.0025	0.9992	1.0019
16		BCG	90	72	0.0052	1.0019	0.9993	0.9999
17		ACG	0.07	277	1.0027	-0.0027	0.9999	1.0008
18		ACG	8	27	1.0055	-0.0027	0.9971	0.9980
19	3 Phase	ABC	1.2	13	1.0430	1.0012	1.0019	-0.0078
20		ABC	1.2	143	0.9796	1.0015	1.0013	-0.0091
21		ABC	8	13	1.0424	1.0012	1.0019	-0.0087
22		ABC	30	13	1.0198	1.0015	1.0014	-0.0088
23		ABC	75	222	0.9594	1.0010	1.0001	-0.0038
24		ABC	90	72	0.9995	1.0008	0.9991	-0.0056
25	3 Phase to Ground	ABCG	1.2	13	1.0314	1.0014	1.0016	-0.0078
26		ABCG	8	13	1.0460	1.0012	1.0020	-0.0087
27		ABCG	30	13	1.0209	1.0015	1.0014	-0.0089
28		ABCG	75	177	0.9999	0.9955	0.9892	-0.0033
29		ABCG	90	13	0.9918	1.0016	1.0007	-0.0020
30		ABCG	90	222	0.9241	1.0016	1.0008	0.0062

6.5 Fault Localization (FL)

The localization of the fault is extremely important stage, given the importance of pinpointing the fault distance from reference point (substation where measurement equipment and protection relay are installed). The importance of this stage is for the following reasons:

- Recording the fault distance in station faults record to find out the number of times the same fault occurs at the same location during a specific period, and if the same fault occurs at the same location repeatedly, the causes of this must be investigated and addressed by the specialized technical team.
- The power transmission lines extend over long distances of hundreds of kilometers therefore, the accurate localization of the fault distance will save time and cost and leads to recover electric services for affected area as soon as possible.

To localize the distance of the fault on TL, the ANN-based fault locator will be utilized for each type of the faults that is simulated at different distances and for various values of fault resistance. Training dataset is labeled in fifteen inputs and one output as it is illustrated in Figure 6.15.

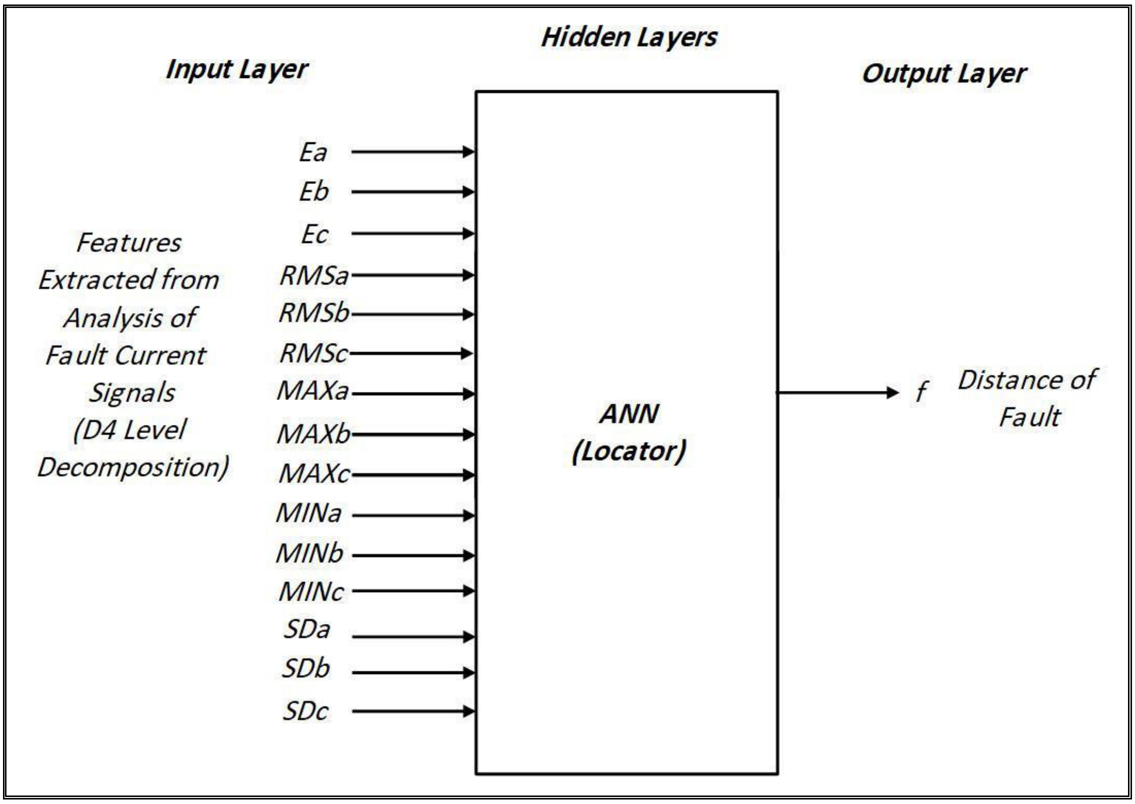


Figure 6.15 Single line of fault locator

The features used as inputs to ANN of the locator are values of Energy, RMS, Max, Min, and SD for the fourth level detail coefficients, MLNN and the back-propagation learning strategy are applied in designing and developing the locator so that the NN can accurately pinpoint the distance of fault. The NN-based fault locators for each type of faults will be designed, developed, implemented, and tested.

6.5.1 Localization of Phase to Ground Faults

For phase to ground faults, three types of faults will be studied which are (A-G, B-G, and CG)

6.5.1.1 Training of A-G Phase Fault Locater

To train the NN-based locater to determine the distance of the fault from the substation in case of A phase to ground fault happened on the TL, several faults on the model are simulated at each 2.5 Km on 300 Km long TL between 5 Km and 285 Km for varying fault resistance as depicted in Table 6.5 in addition to, two critical cases at 0 Km & 300 Km, to generate the training data set that consists of 1160 cases. In each of these cases, the three-phase current signals are read, sampled, and decomposed by using DWT, then the features are extracted from wavelet coefficients and used as inputs to the NN of locater. The output of the NN is the distance of the fault from station/substation.

Table 6.5 Generation conditions of training dataset for A-G locator

Parameter	Training Data Set
Fault Type	**AG**
Fault Resistance (Ω)	**0.01, 0.05, 0.1, 0.5, 1, 5, 10, 25, 50, 100**
Fault Location (Km)	**Each 2.5 Km between 5 Km and 287.5 Km, in addition to two critical cases at 0 Km & 300 Km**
Training Cases	**1160**

Figure 6.16 & Figure 6.17 show the selected 15-15-15-15-1 NN structure to determine the A-G fault distance and its performance. This NN has MSE of 0.219.

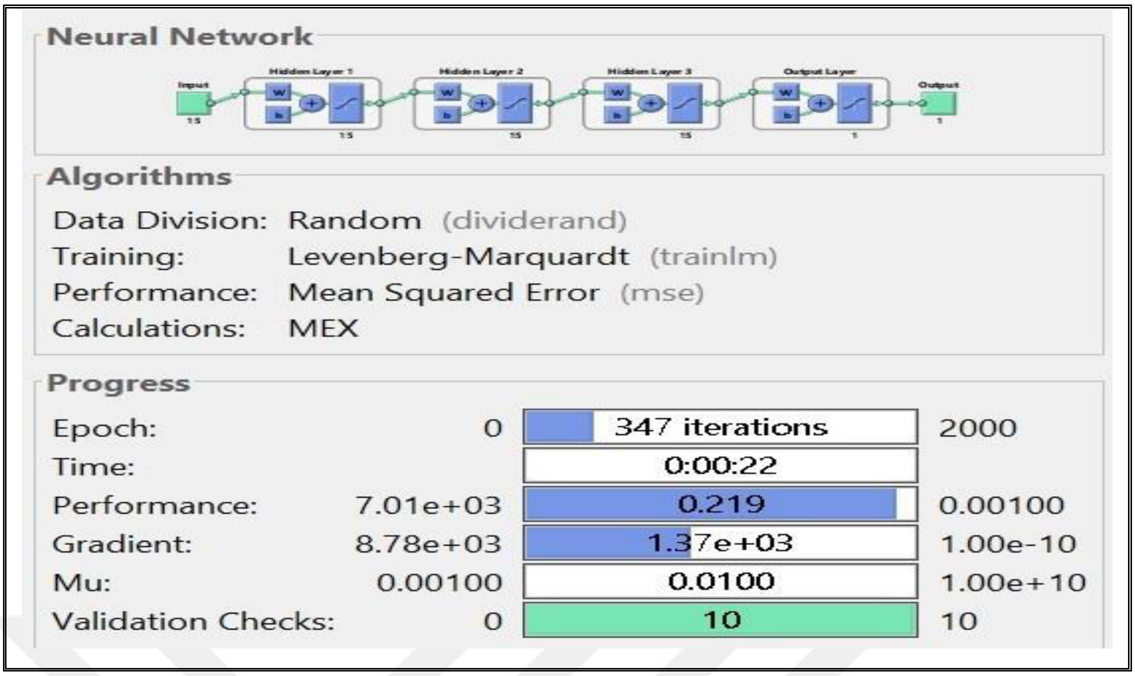


Figure 6.16 Structure of ANN for A-G FL (15-15-15-15-1)

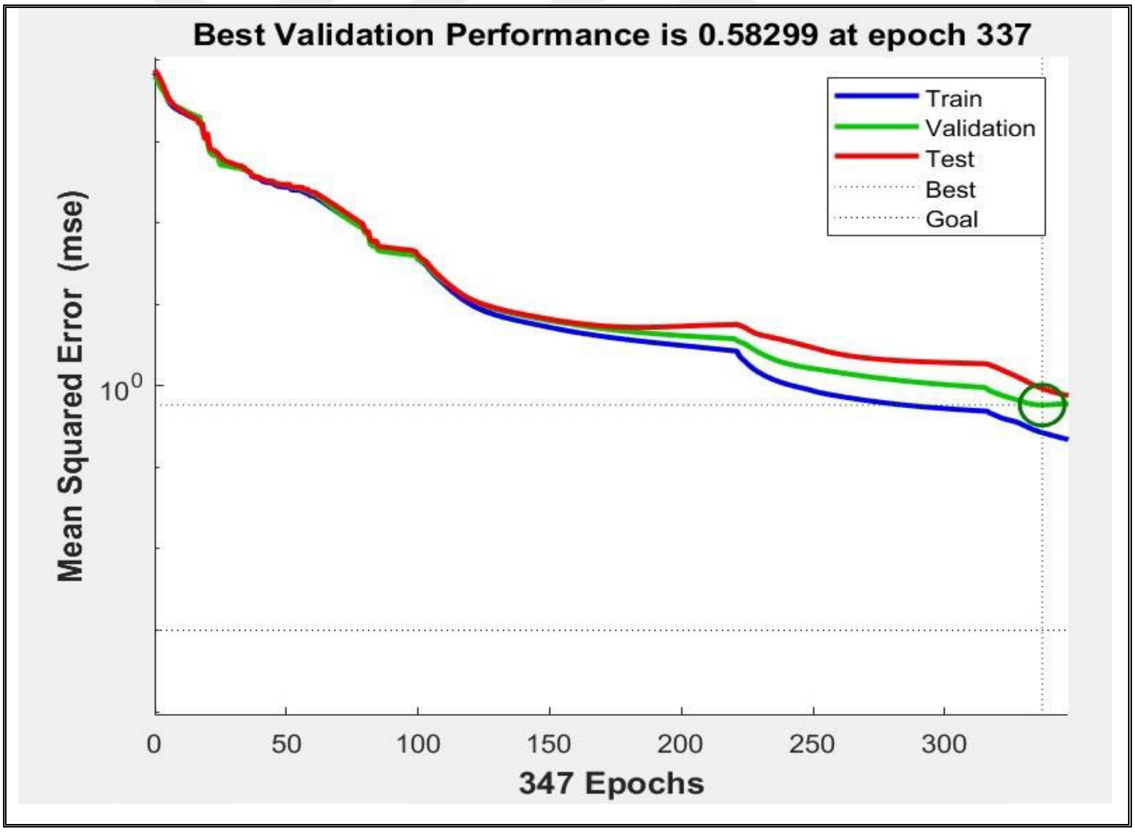


Figure 6.17 Training performance curve of ANN for A-G FL (15-15-15-15-1)

6.5.1.2 Evaluating of A-G Phase Fault Locater

After completing the training process, the performance of ANN is evaluated through the following criterion:

- plotting the linear regression scheme that relates the target output to the actual output of ANN as shown in Figure 6.18. which indicates that all the NN actual outputs have values near their target outputs, the actual value for all outputs can be represented with the following equation:

$$\text{Actual Output} \approx 1 \times \text{Target output} + 0.047 \text{ km} \tag{6.1}$$

From Figure 6.18, the correlation coefficient equals to 0.9999 which indicates excellent association between target and actual output of the locator.

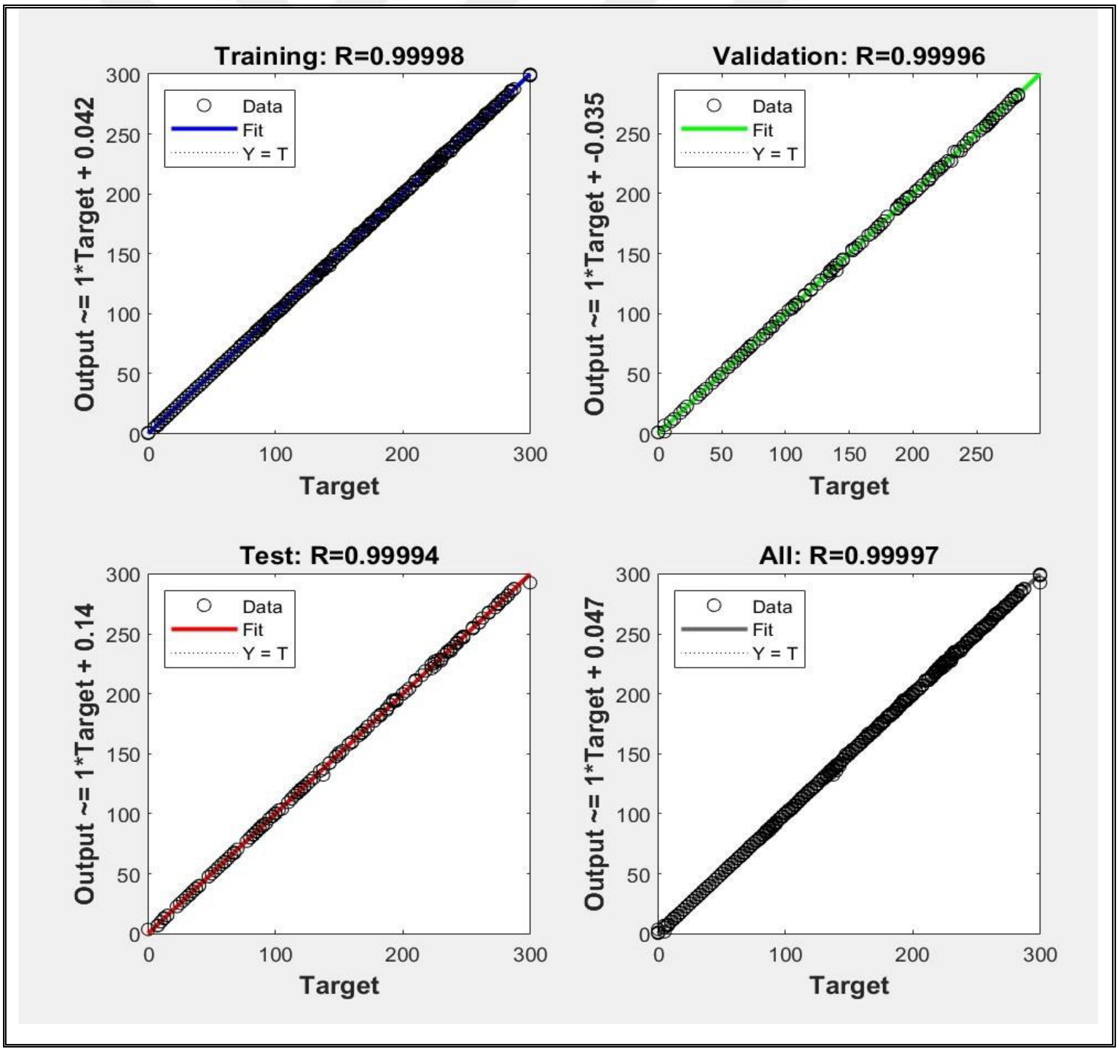


Figure 6.18 Regression scheme of training process for A-G FL

6.5.1.3 Testing of A-G Phase Fault Locator

To test the effectiveness and performance of the designated NN of A-G fault locator, the ANN structure has been simulated and tested for 56 scenarios (at different distances for various values of faults resistance). Figure 6.19 illustrates the regression scheme of the relation between actual and target output of A-G fault locator for testing dataset.

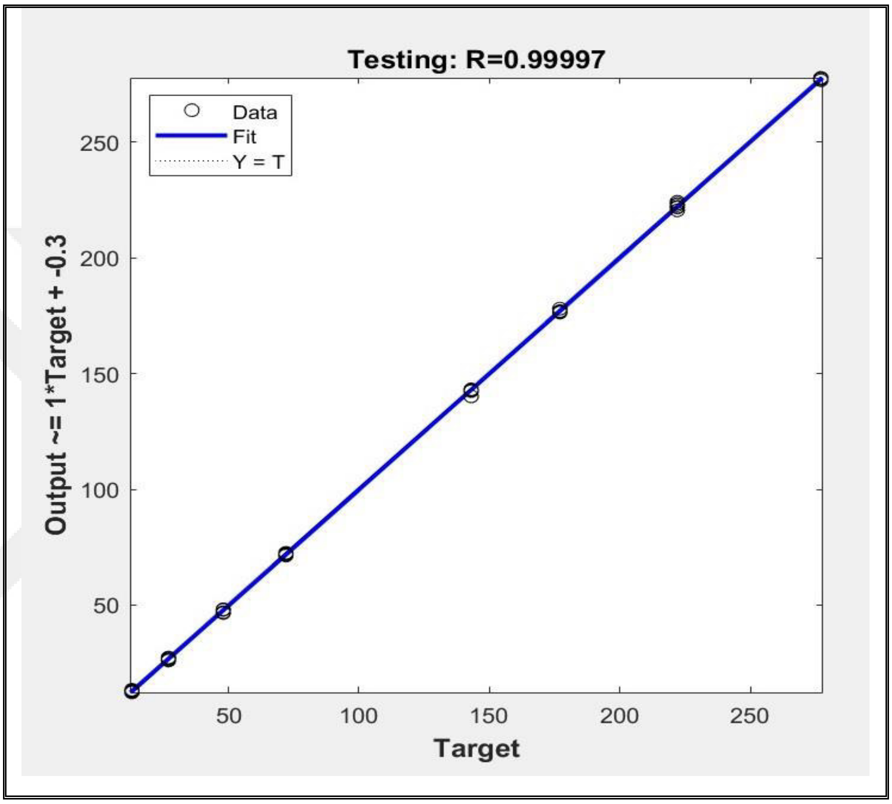


Figure 6.19 Regression scheme of testing process for A-G FL

From Figure 6.19, the actual output of NN for testing data can be represented by the following equation:

$$\text{Actual Output} \approx 1 \times \text{Target Output} \pm 0.3 \text{ km} \tag{6.2}$$

Output values in this Figure approaches their target values. These results are acceptable since the errors are considered very small compared to the HVTL length.

Table 6.6 shows samples of these actual and target values of ANN output and percentage of error under different conditions.

Table 6.6 Percentage error of ANN selected for phase A-G fault locator.

Fault Conditions				Actual Output of ANN	Error Value	Percentage Error (%)
N.	Fault Type	Fault resistance (Ω)	Fault Distance (Km)			
1	AG	0.07	27	27.017	0.017	0.06%
2		0.07	72	71.755	0.245	0.34%
3		0.07	143	143.057	0.057	0.04%
4		0.07	177	176.413	0.587	0.33%
5		0.07	222	221.927	0.073	0.03%
6		0.07	277	277.557	0.557	0.20%
7		1.2	27	27.228	0.228	0.84%
8		1.2	72	72.053	0.053	0.07%
9		1.2	143	143.192	0.192	0.13%
10		1.2	177	176.981	0.019	0.01%
11		1.2	222	224.277	2.277	1.03%
12		1.2	277	277.572	0.572	0.21%
13		8	27	26.971	0.029	0.11%
14		8	72	72.509	0.509	0.71%
15		8	143	140.179	2.821	1.97%
16		8	177	176.742	0.258	0.15%
17		8	222	222.814	0.814	0.37%
18		8	277	277.356	0.356	0.13%
19		30	27	27.311	0.311	1.15%
20		30	72	72.567	0.567	0.79%
21		30	143	143.212	0.212	0.15%
22		30	177	176.896	0.104	0.06%
23		30	222	223.500	1.500	0.68%
24		30	277	277.415	0.415	0.15%
25		75	27	26.037	0.963	3.57%
26		75	72	71.405	0.595	0.83%
27		75	143	142.633	0.367	0.26%
28		75	177	178.076	1.076	0.61%
29		75	222	220.540	1.460	0.66%
30		75	277	277.813	0.813	0.29%
Mean Value of Percentage Error						**0.53%**
Standard Deviation						**0.007**

Table 6.6 provides information about Percentage Error of fault distance determined by locator for faults under many conditions, the best scenario was at distance 177 Km for fault resistance at 1.2 Ω whereas the worst scenario was at distance 27 Km for fault resistance at 75 Ω. The value of SD is 0.007 that indicates the lower the dispersion, and error readings are close to the average.

6.5.1.4 Training of B-G Phase Fault Locater

To train the NN-based B-G Phase Fault locater, the dataset generated under the same conditions for the A-G Phase Fault Locater will be used that consists of 1160 cases. Figure 6.20 & Figure 6.21 show the selected 15-16-16-16-1 NN structure to determine the B-G fault distance and its performance scheme. This NN has MSE of 0.128.

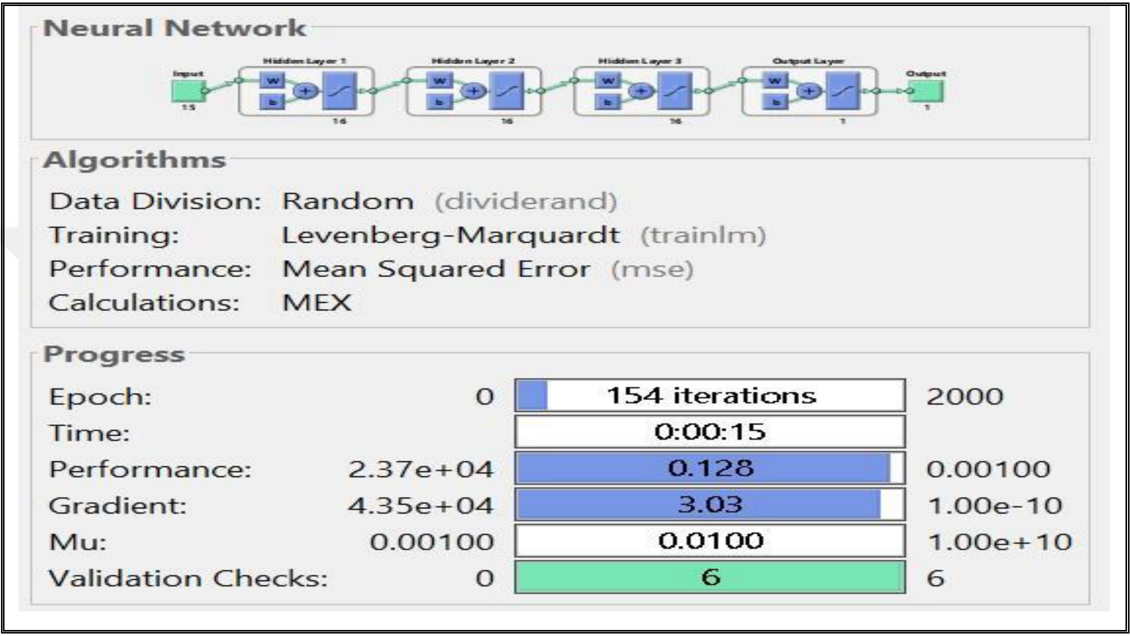


Figure 6.20 Structure of ANN for B-G FL (15-16-16-16-1)

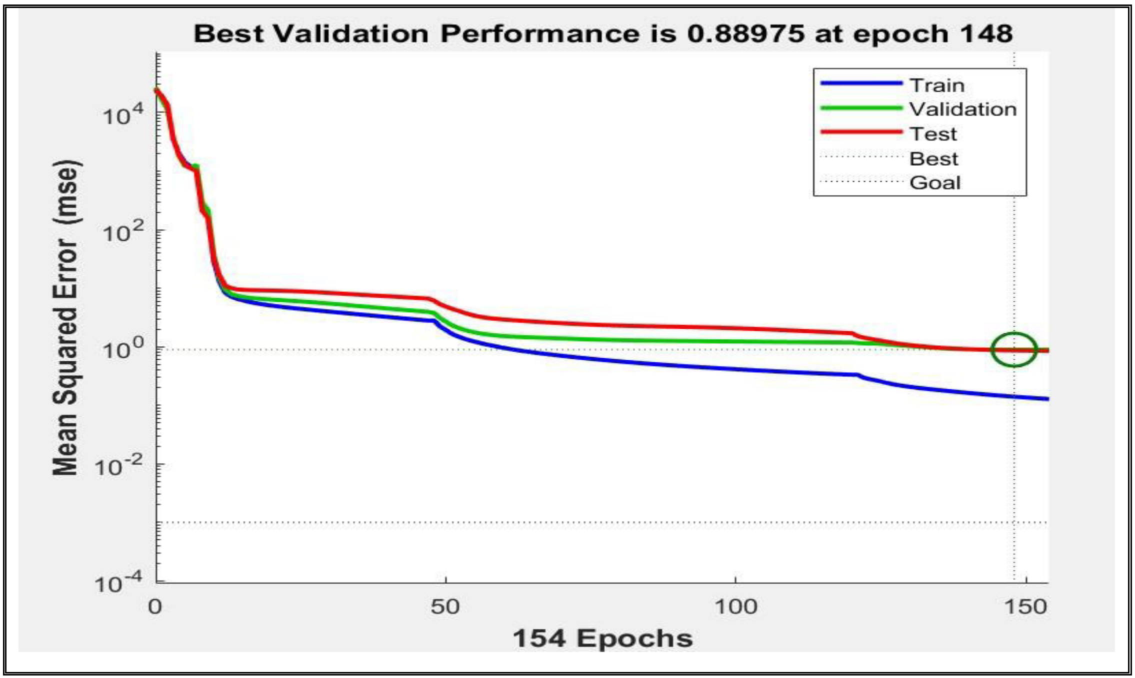


Figure 6.21 Training performance curve of ANN for B-G FL (15-16-16-16-1)

6.5.1.5 Evaluating of Phase B-G Fault Locater

After completing the training process, the performance of ANN is evaluated through plotting the linear regression scheme as shown in Figure 6.22. which indicates that all the NN actual outputs have values near their target outputs, the actual value for all outputs can be represented with the following equation:

$$\text{Actual Output} \approx 1 \times \text{Target Output} + 0.044 \text{ km} \quad (6.3)$$

From Figure 6.22, the correlation coefficient equals to 0.9999 which indicates excellent association between target and actual output of locator.

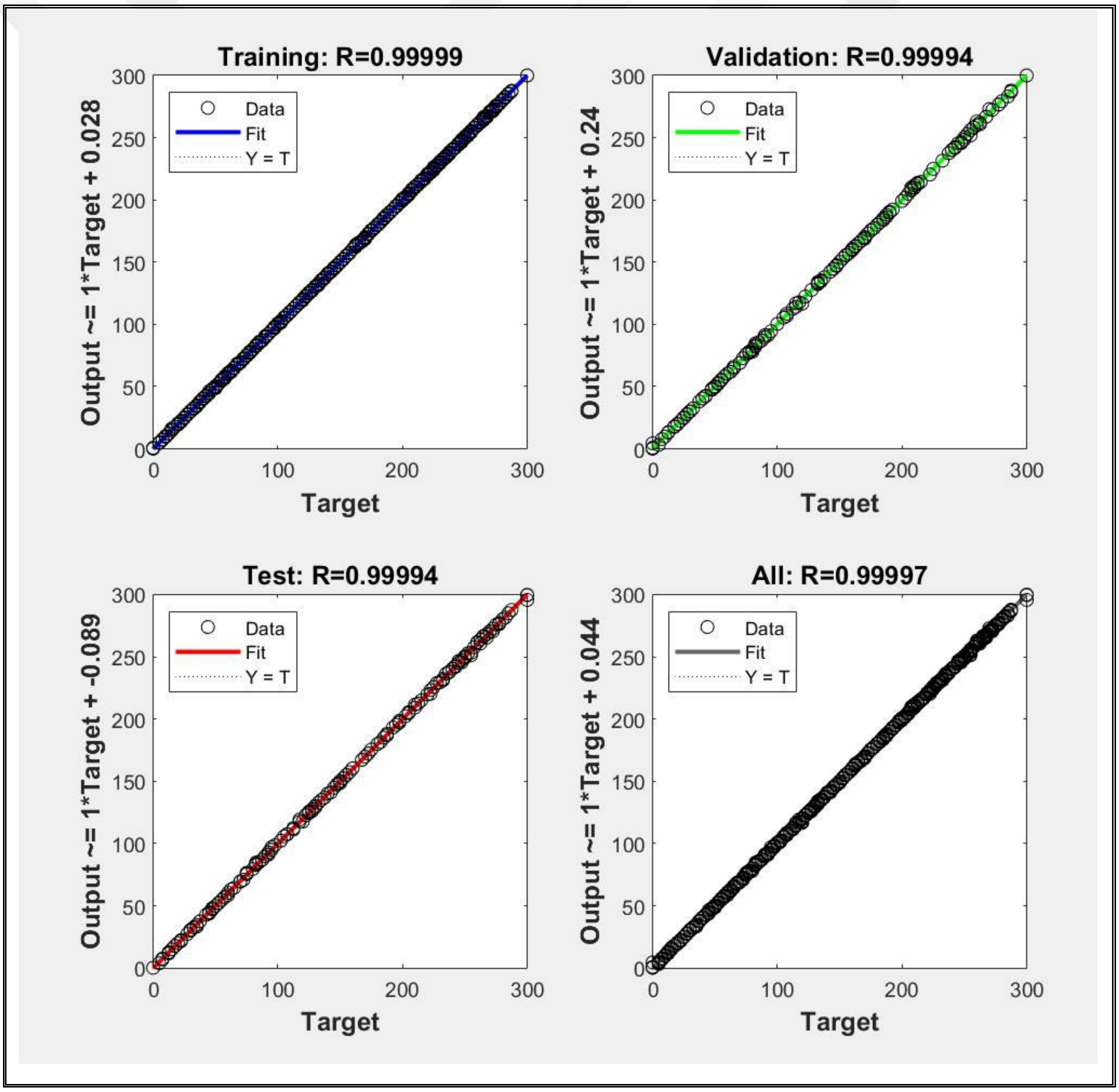


Figure 6.22 Regression scheme of training process for B-G FL

6.5.1.6 Testing of B-G Phase Fault Locator

To test the performance of the selected neural network of B-G fault locator, the ANN structure has been simulated for the cases generated under the same conditions used for Phase A-G fault locator. Figure 6.23 illustrates the regression scheme of the relation between actual and target outputs of B-G fault locator for testing dataset.

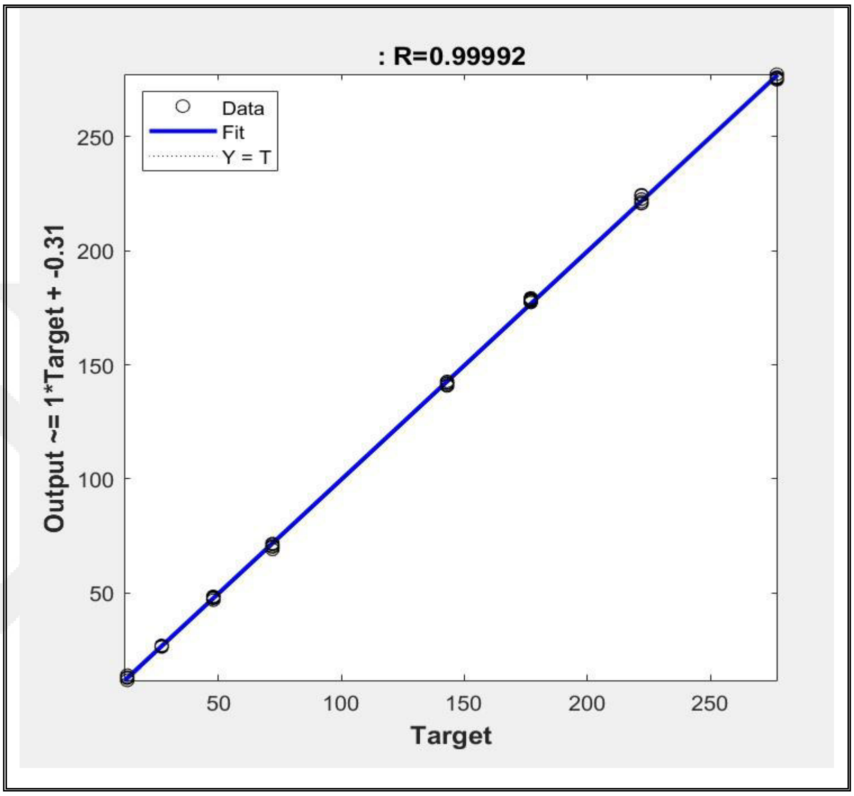


Figure 6.23 Regression scheme of testing process for B-G FL

From Figure 6.23, the actual output of NN for testing data can be represented by the following equation:

$$\text{Actual Output} \approx 1 \times \text{Target Output} \pm 0.31 \text{ km} \quad (6.4)$$

Output values in this Figure approaches their target values. These results are acceptable since the errors are considered very small compared to the HVTL length.

Table 6.7 shows samples of these actual and target values of ANN output and percentage of error under different conditions.

Table 6.7 Percentage error of ANN selected for B-G fault locator.

Fault Conditions				Actual Output of ANN	Error Value	Percentage Error (%)
N.	Fault Type	Fault resistance (Ω)	Fault Distance (Km)			
1	BG	0.07	27	26.957	0.043	0.16%
2		0.07	72	70.637	1.363	1.89%
3		0.07	143	141.888	1.112	0.78%
4		0.07	177	178.541	1.541	0.87%
5		0.07	222	224.151	2.151	0.97%
6		0.07	277	275.637	1.363	0.49%
7		1.2	27	26.644	0.356	1.32%
8		1.2	72	71.304	0.696	0.97%
9		1.2	143	142.215	0.785	0.55%
10		1.2	177	178.056	1.056	0.60%
11		1.2	222	222.613	0.613	0.28%
12		1.2	277	274.906	2.094	0.76%
13		8	27	27.086	0.086	0.32%
14		8	72	71.710	0.290	0.40%
15		8	143	142.787	0.213	0.15%
16		8	177	178.939	1.939	1.10%
17		8	222	222.704	0.704	0.32%
18		8	277	276.126	0.874	0.32%
19		30	27	26.904	0.096	0.36%
20		30	72	71.690	0.310	0.43%
21		30	143	141.941	1.059	0.74%
22		30	177	177.297	0.297	0.17%
23		30	222	220.634	1.366	0.62%
24		30	277	277.371	0.371	0.13%
25		75	27	26.674	0.326	1.21%
26		75	72	70.081	1.919	2.67%
27		75	143	140.721	2.279	1.59%
28		75	177	179.362	2.362	1.33%
29		75	222	220.866	1.134	0.51%
30		75	277	275.426	1.574	0.57%
Mean Value of Percentage Error						**0.75%**
Standard Deviation						**0.005**

Table 6.7 provides information about Percentage Error of fault distance determined by locator for faults under many conditions, the best scenario was at distance 277 Km for fault resistance at 30 Ω whereas the worst scenario was at distance 72 Km for fault resistance at 75 Ω. The value of SD is 0.005 that indicates the lower the dispersion, and error readings are close to the average.

6.5.1.7 Training of C-G Phase Fault Locater

To train the NN-based C-G Phase Fault locater, the dataset generated under the same conditions for the A-G & B-G Phase Fault Locater will be used that consists of 1160 cases. Figure 6.24 & Figure 6.25 show the selected 15-16-16-16-1 NN structure to determine the C-G fault location and its performance scheme. This NN has MSE of 0.0853.

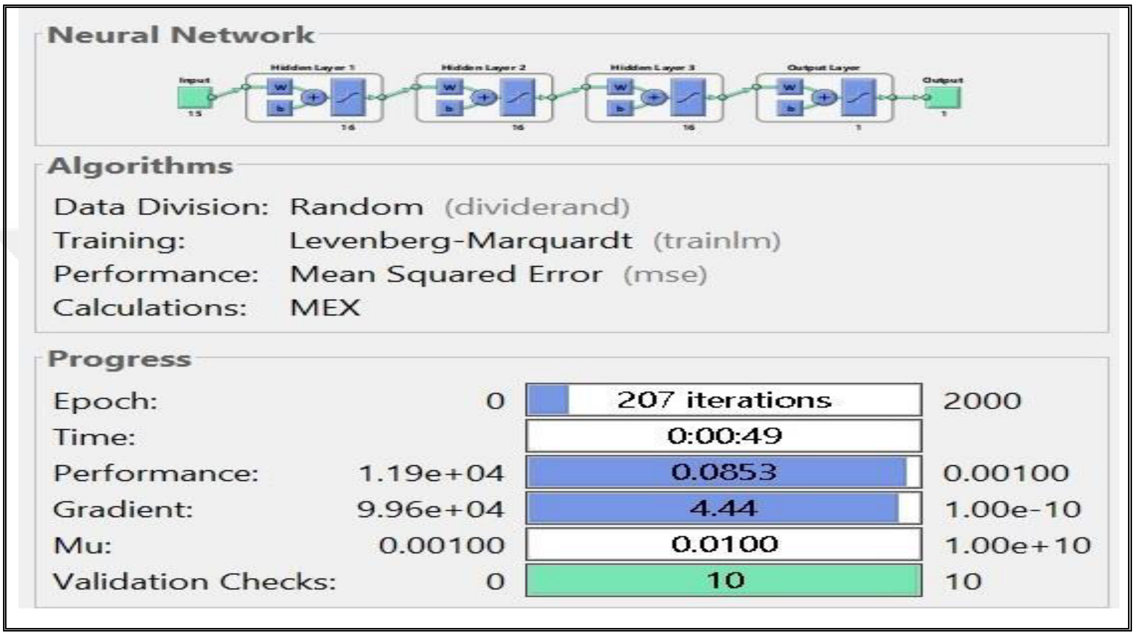


Figure 6.24 Structure of ANN for C-G FL (15-16-16-16-1)

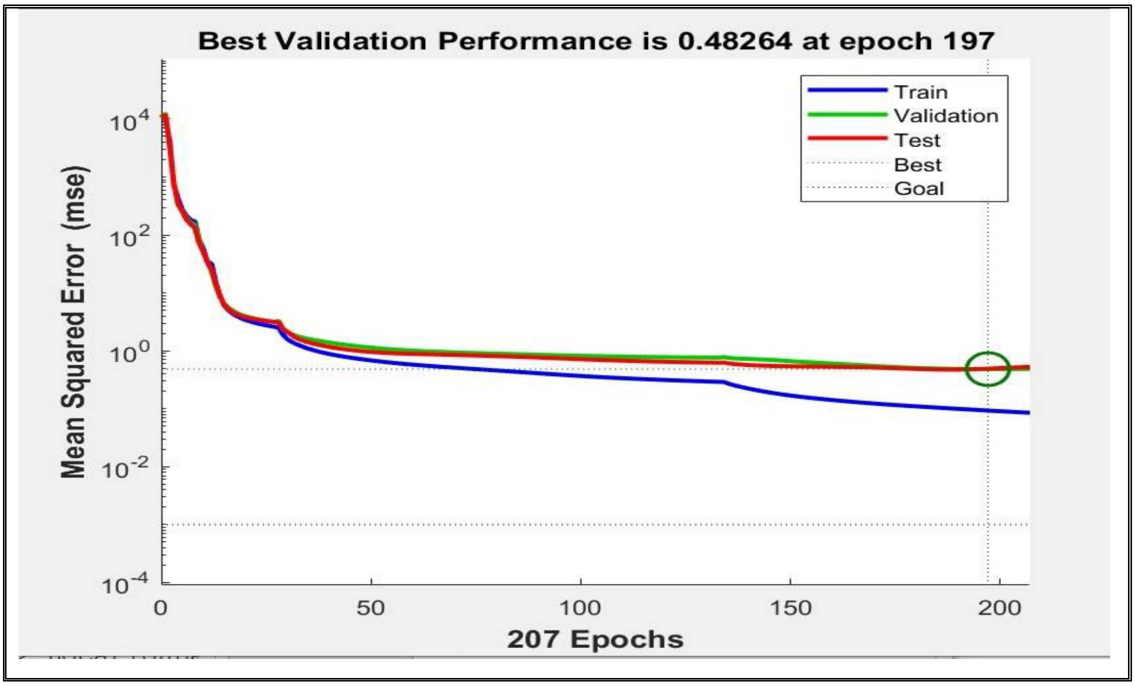


Figure 6.25 Training performance curve of ANN for C-G FL (15-16-16-16-1)

6.5.1.8 Evaluating of Phase C-G Fault Locater

After completing the training process, the performance of ANN is evaluated through plotting the regression scheme as shown in Figure 6.26. which indicates that all the NN actual outputs have values near their target outputs, the actual value for all outputs can be represented with the following equation:

$$\text{Actual Output} \approx 1 \times \text{Target Output} + 0.0085 \text{ km} \tag{6.5}$$

From Figure 6.26, the correlation coefficient equals to 0.9999 which indicates excellent association between target and actual output of locator.

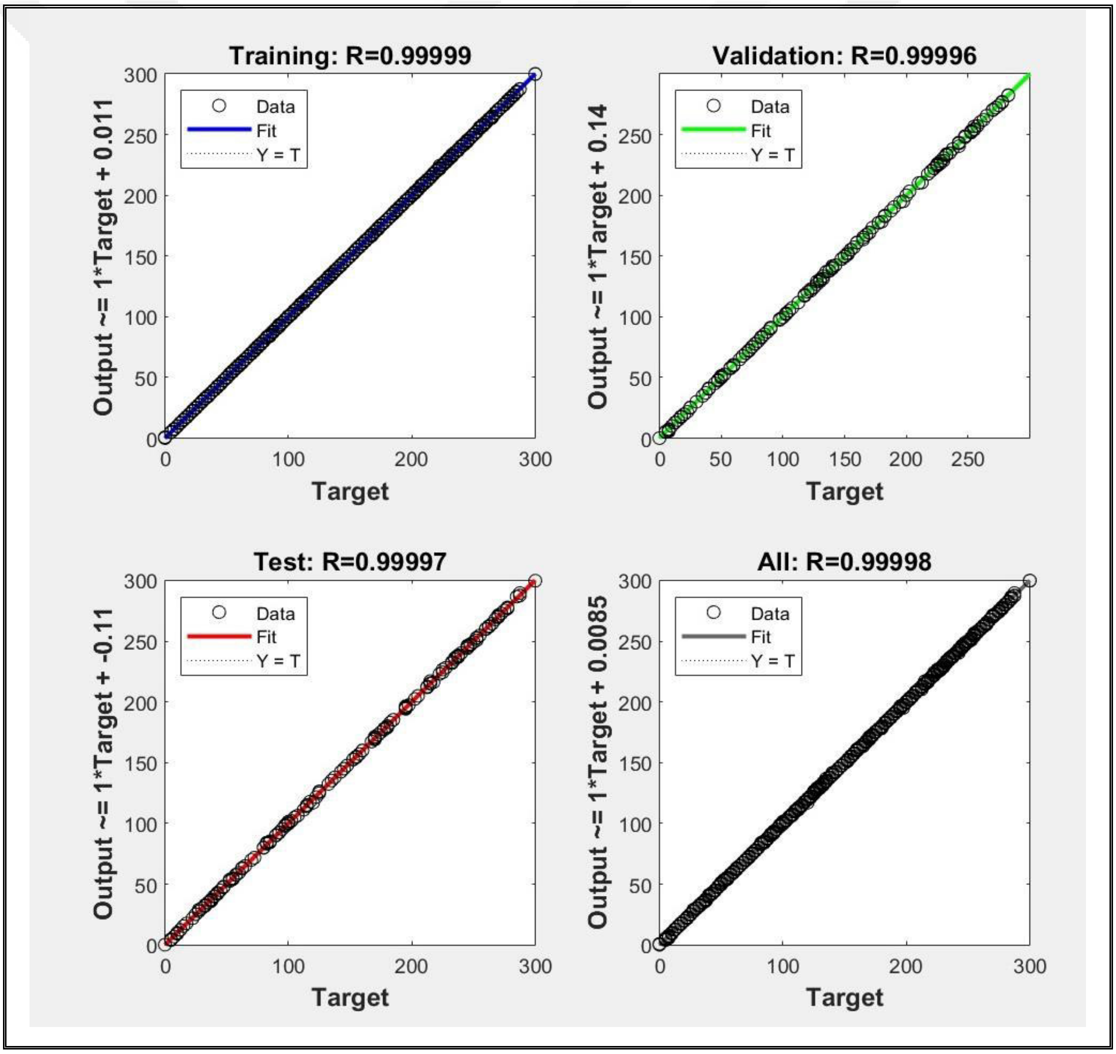


Figure 6.26 Regression scheme of training process for C-G FL

6.5.1.9 Testing of C-G Phase Fault Locator

To test the performance of the selected NN of C-G fault locator, the ANN structure has been simulated for the fault cases generated under the same conditions used for testing of the A-G & B-G fault locator. Figure 6.27 illustrates the regression scheme of the relation between actual and target outputs of C-G fault locator for testing data.

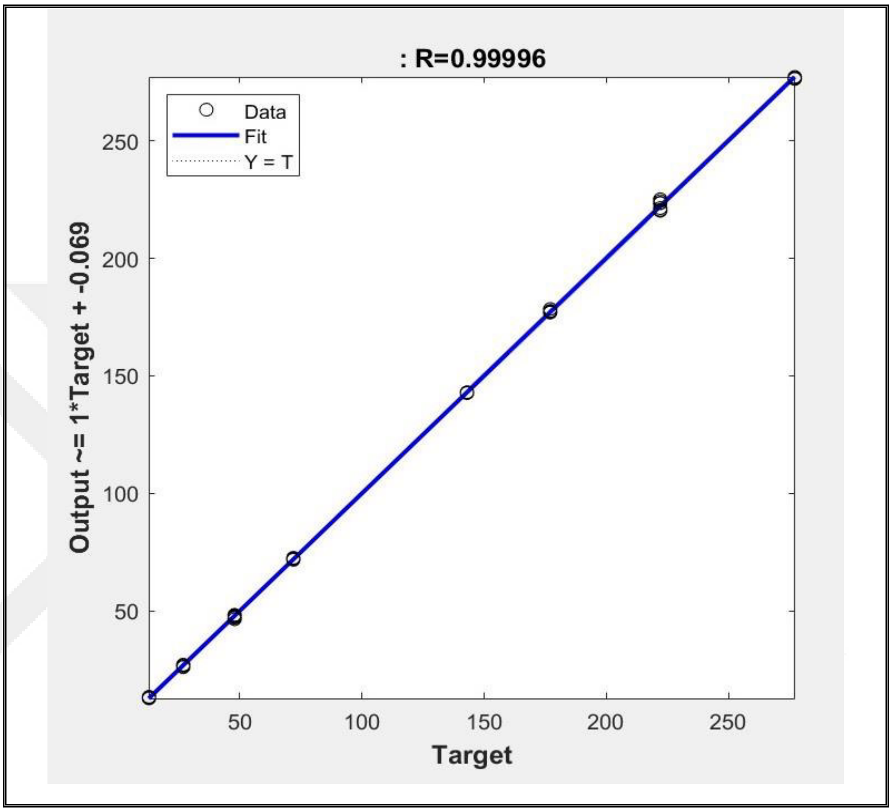


Figure 6.27 Regression scheme of testing process for C-G FL

From Figure 6.27, the actual output of NN for testing data can be represented by the following equation:

$$\text{Actual Output} \approx 1 \times \text{Target Output} \pm 0.069 \text{ km} \qquad (6.6)$$

Output values in this figure approaches their target values. These results are acceptable since the errors are considered very small compared to the HVTL length.

Table 6.8 shows examples of these actual and target values of ANN output and percentage of error under different conditions.

Table 6.8 Percentage error of ANN selected for C-G fault locator.

Fault Conditions				Actual Output of ANN	Error Value	Percentage Error (%)
N.	Fault Type	Fault resistance (Ω)	Fault Distance (Km)			
1	CG	0.07	27	27.157	0.157	0.58%
2		0.07	72	71.967	0.033	0.05%
3		0.07	143	142.978	0.022	0.02%
4		0.07	177	177.259	0.259	0.15%
5		0.07	222	223.441	1.441	0.65%
6		0.07	277	276.288	0.712	0.26%
7		1.2	27	26.931	0.069	0.26%
8		1.2	72	72.204	0.204	0.28%
9		1.2	143	142.866	0.134	0.09%
10		1.2	177	177.176	0.176	0.10%
11		1.2	222	225.132	3.132	1.41%
12		1.2	277	276.388	0.612	0.22%
13		8	27	26.345	0.655	2.43%
14		8	72	72.656	0.656	0.91%
15		8	143	142.935	0.065	0.05%
16		8	177	177.366	0.366	0.21%
17		8	222	224.399	2.399	1.08%
18		8	277	277.063	0.063	0.02%
19		30	27	26.846	0.154	0.57%
20		30	72	72.393	0.393	0.55%
21		30	143	142.732	0.268	0.19%
22		30	177	176.914	0.086	0.05%
23		30	222	221.392	0.608	0.27%
24		30	277	277.035	0.035	0.01%
25		75	27	26.250	0.750	2.78%
26		75	72	71.795	0.205	0.28%
27		75	143	142.999	0.001	0.00%
28		75	177	178.473	1.473	0.83%
29		75	222	220.196	1.804	0.81%
30		75	277	277.122	0.122	0.04%
Mean Value						**0.51%**
Standard Deviation						**0.006**

Table 6.8 provides information about Percentage Error of fault distance determined by locator for faults under many conditions, the best scenario was at distance 143 Km for fault resistance at 75 Ω whereas the worst scenario was at distance 27 Km for fault resistance at 75 Ω. The value of SD is 0.006 that indicates the lower the dispersion, and error readings are close to the average.

6.5.2 Localization of Phase-to-Phase Faults

For phase-to-phase faults, due to limitation of space, just one type of these faults will be studied, which is A-B phase fault.

6.5.2.1 Training of A-B Phase Fault Locater

To train the NN-based A-B Phase Fault locater, the dataset used for training consists of 1160 cases generated under the same conditions of Ph-G faults will be used. Figure 6.24&Figure 6.25 show the selected 15-16-16-16-1 NN structure to determine the C-G fault location and its performance scheme. This NN has MSE of 2.28.

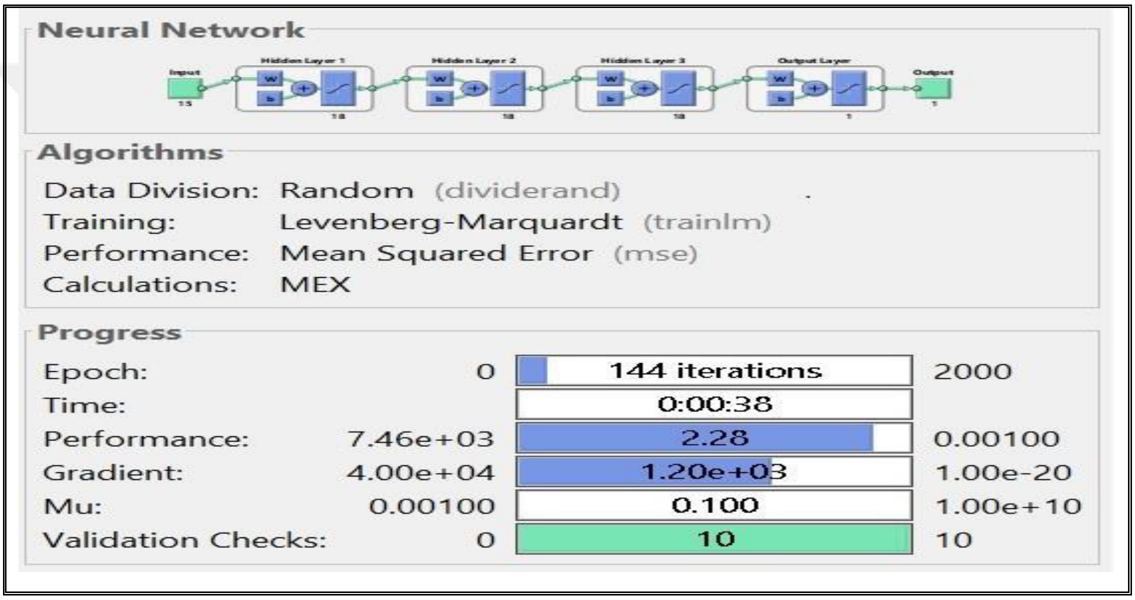


Figure 6.28 Structure of ANN for A-B FL (15-16-16-16-1)

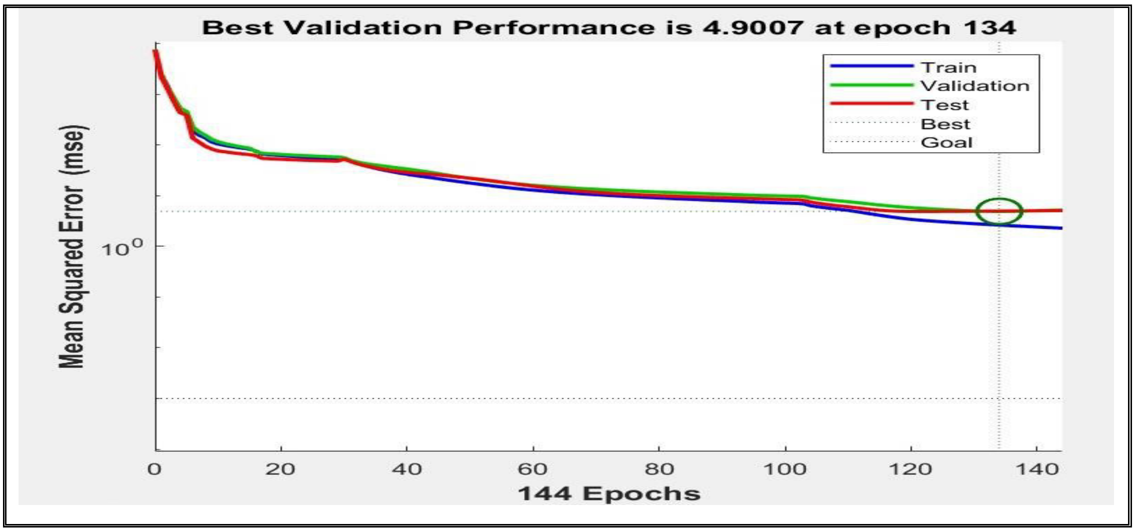


Figure 6.29 Training performance curve of ANN for A-B FL (15-16-16-16-1)

6.5.2.2 Evaluating of A-B Phase Fault Locater

After completing the training process, the performance of ANN is evaluated through plotting the regression scheme as shown in Figure 6.30. which indicates that all the NN actual outputs have values near their target outputs, the actual value for all outputs can be represented with the following equation:

$$\text{Actual Output} \approx 1 \times \text{Target Output} + 0.2\ \text{km} \tag{6.7}$$

From Figure 6.30, the correlation coefficient equals to 1 which indicates excellent association between target and actual output of locator.

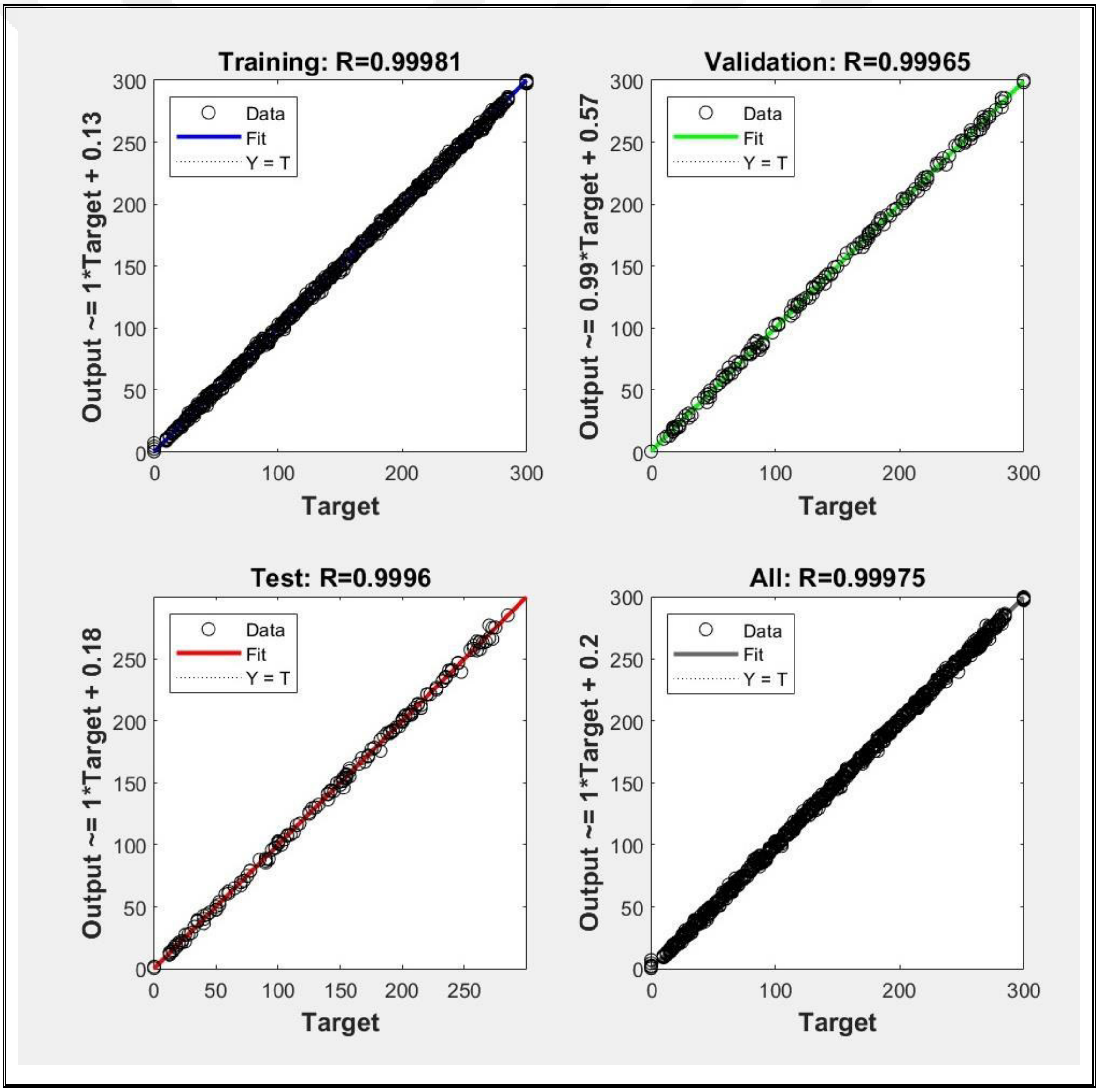


Figure 6.30 Regression schemes of training process for A-B FL

6.5.2.3 Testing of A-B Phase Fault Locator

To test the performance of the selected NN of A-B fault locator, the ANN structure has been simulated for the same cases used for testing of A-G & B-G & C-G fault locators. Figure 6.31 illustrates the regression scheme of the relation between actual and target outputs of A-B fault locator for testing data.

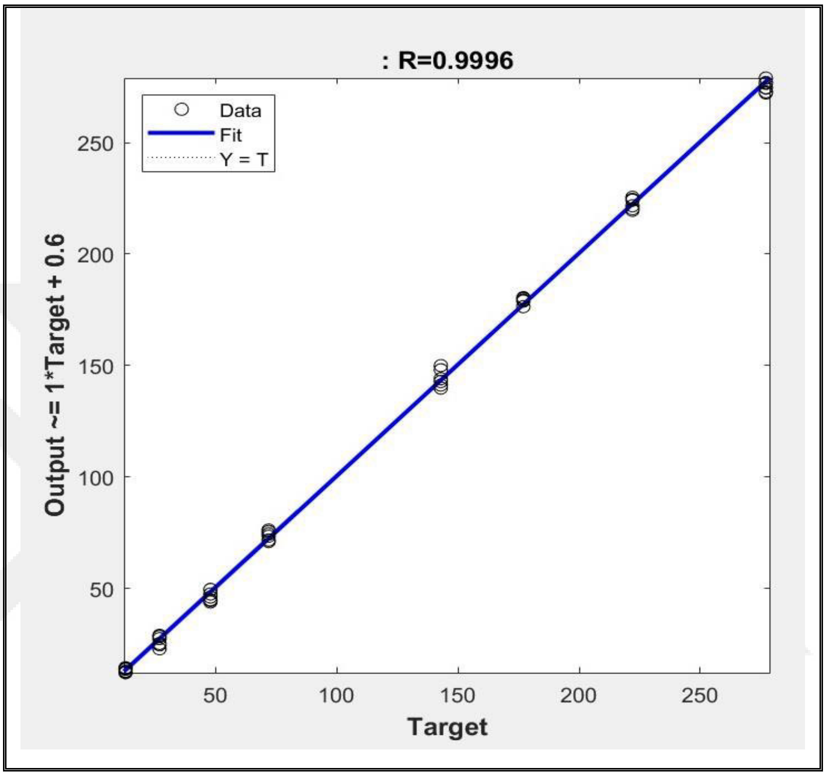


Figure 6.31 Regression scheme of testing process for A-B FL

From Figure 6.31, the actual output of NN for testing data can be represented by the following equation:

$$\text{Actual Output} \approx 1 \times \text{Target Output} + 0.6\ \text{km} \tag{6.8}$$

Output values in this Figure approaches their target values. These results are acceptable since the errors are considered very small compared to the HVTL length.

Table 6.9 shows examples of these actual and target values of ANN output and percentage of error under different conditions.

Table 6.9 Percentage error of ANN selected for A-B fault locator.

Fault Conditions				Actual Output of ANN	Error Value	Percentage Error (%)
N.	Fault Type	Fault resistance (Ω)	Fault Distance (Km)			
1	AB	0.07	27	28.584	1.584	5.87%
2		0.07	72	74.513	2.513	3.49%
3		0.07	143	148.133	5.133	3.59%
4		0.07	177	180.517	3.517	1.99%
5		0.07	222	224.522	2.522	1.14%
6		0.07	277	274.612	2.388	0.86%
7		1.2	27	27.634	0.634	2.35%
8		1.2	72	75.588	3.588	4.98%
9		1.2	143	150.008	7.008	4.90%
10		1.2	177	180.150	3.150	1.78%
11		1.2	222	224.501	2.501	1.13%
12		1.2	277	276.536	0.464	0.17%
13		8	27	25.455	1.545	5.72%
14		8	72	76.307	4.307	5.98%
15		8	143	144.126	1.126	0.79%
16		8	177	178.985	1.985	1.12%
17		8	222	224.166	2.166	0.98%
18		8	277	278.906	1.906	0.69%
19		30	27	29.118	2.118	7.85%
20		30	72	73.559	1.559	2.17%
21		30	143	141.322	1.678	1.17%
22		30	177	176.440	0.560	0.32%
23		30	222	220.354	1.646	0.74%
24		30	277	276.900	0.100	0.04%
25		75	27	23.068	3.932	14.56%
26		75	72	71.846	0.154	0.21%
27		75	143	139.967	3.033	2.12%
28		75	177	178.957	1.957	1.11%
29		75	222	221.653	0.347	0.16%
30		75	277	272.745	4.255	1.54%
Mean Value						**1.42%**
Standard Deviation						**0.030**

Table 6.9 provides information about Percentage Error of fault distance determined by locator for faults under many conditions, the best scenario was at distance 277 Km for fault resistance at 30 Ω whereas the worst scenario was at distance 27 Km for fault resistance at 75 Ω. The value of SD is 0.03 that indicates the lower the dispersion, and error readings are close to the average.

6.5.3 Localization of Phase-to-Phase to Ground Faults

For phase-to-phase to ground faults, due to limitation of space, just one type of faults will be studied, which is A-B phase to G fault.

6.5.3.1 Training of AB phase to-G Fault Locater

To train the NN-based AB phase-G Fault locater, the dataset consisted of 1160 cases generated under the same conditions will be used. Figure 6.32 & Figure 6.33 show the selected 15-18-18-18-1 NN structure to determine the AB-G fault location and its performance scheme. This NN has MSE of 0.259.

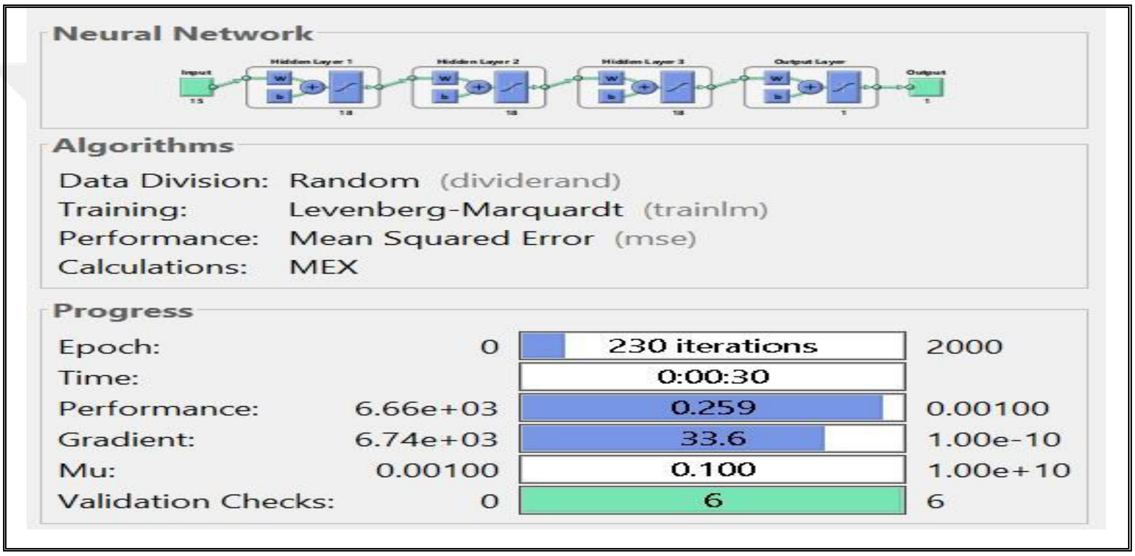


Figure 6.32 Structure of ANN for AB-G FL (15-18-18-18-1)

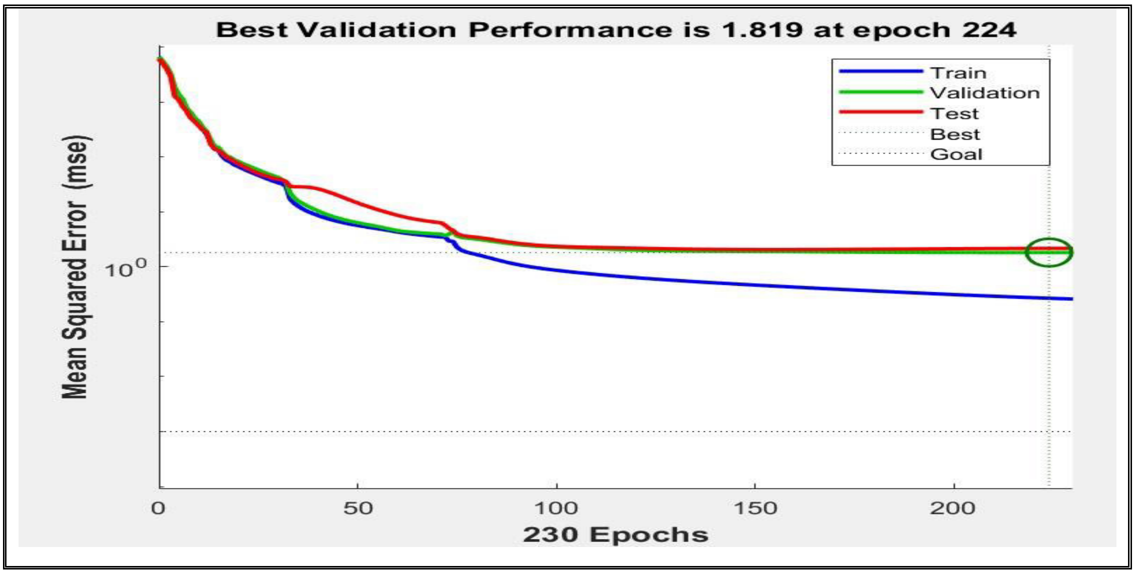


Figure 6.33 Training performance curve of ANN for AB-G FL (15-18-18-18-1)

6.5.3.2 Evaluating of Phase AB-G Fault Locater

After completing the training process, the performance of ANN is evaluated through plotting the regression scheme as shown in Figure 6.34. which indicates that all the NN actual outputs have values near their target outputs, the actual value for all outputs can be represented with the following equation:

$$\text{Actual Output} \approx 1 \times \text{Target Output} + 0.022 \text{ km} \tag{6.9}$$

From Figure 6.34, the correlation coefficient equals to 0.999 which indicates excellent association between target and actual output of locator.

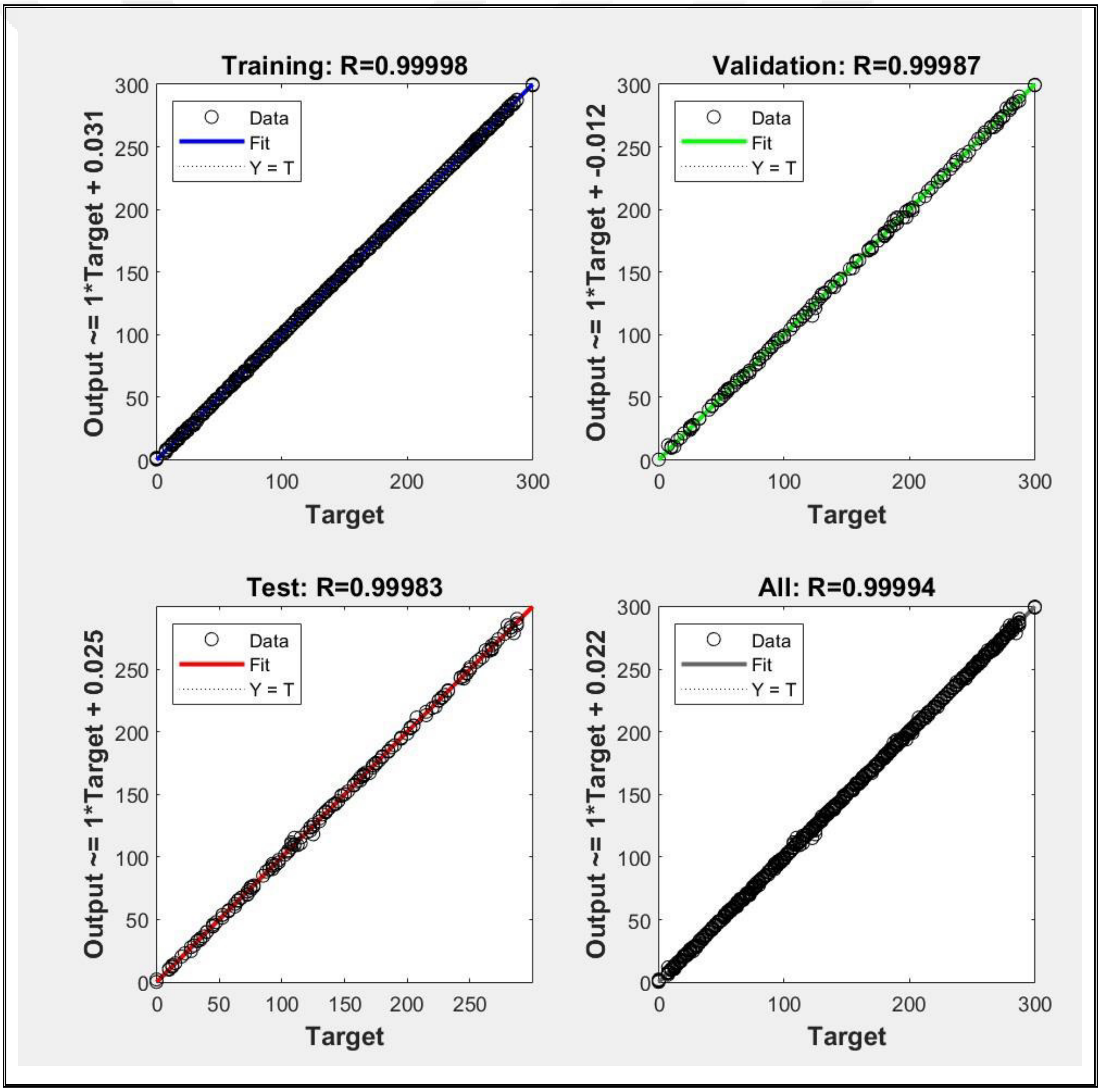


Figure 6.34 Regression scheme of training process for AB-G FL

6.5.3.3 Testing of AB-G Phase Fault Locator

To test the performance of the selected NN of AB-G fault locator, the ANN structure has been simulated for the fault cases generated under the same conditions used for testing the other types of locators. Figure 6.35 illustrates the regression scheme of the relation between actual and target outputs of AB-G fault locator for testing data.

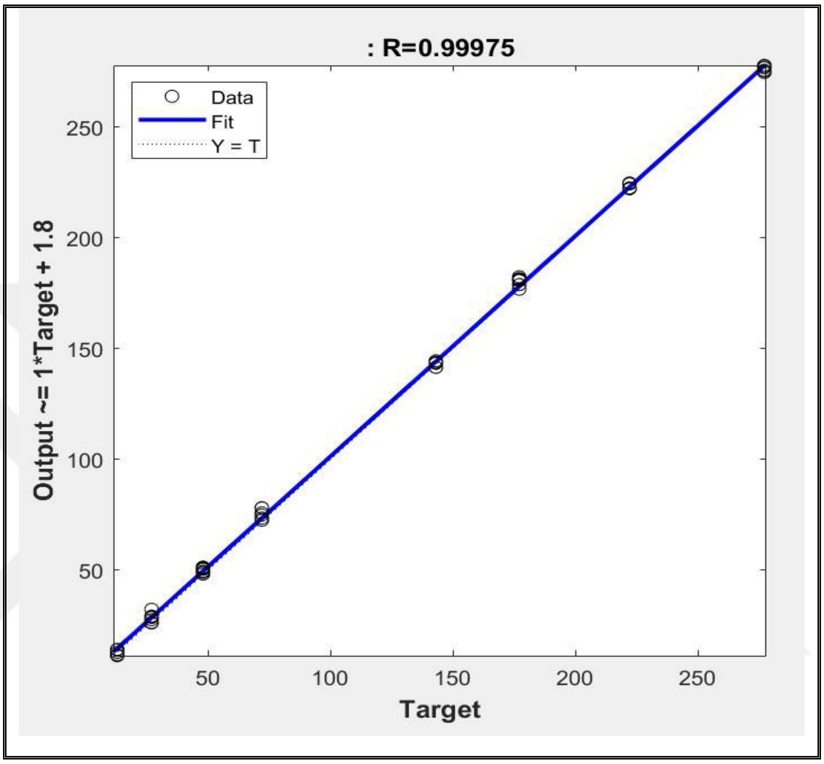


Figure 6.35 Regression scheme of testing process for AB-G FL

From Figure 6.35, the actual output of NN for testing data can be represented by the following equation:

$$\text{Actual Output} \approx 1 \times \text{Target Output} + 1.8 \text{ km} \tag{6.10}$$

Output values in this Figure relatively approaches their target values. These results are acceptable since the errors are considered very small compared to the HVTL length. Table 6.10 shows examples of these actual and target values of ANN output and percentage of error under different conditions.

Table 6.10 Percentage error of ANN selected for AB phase-G fault locator.

Fault Conditions				Actual Output of ANN	Error Value	Percentage Error (%)
N.	Fault Type	Fault resistance (Ω)	Fault Distance (Km)			
1	AB-G	0.07	27	25.335	1.665	6.17%
2		0.07	72	73.418	1.418	1.97%
3		0.07	143	144.786	1.786	1.25%
4		0.07	177	177.112	0.112	0.06%
5		0.07	222	221.705	0.295	0.13%
6		0.07	277	277.722	0.722	0.26%
7		1.2	27	26.403	0.597	2.21%
8		1.2	72	72.703	0.703	0.98%
9		1.2	143	144.505	1.505	1.05%
10		1.2	177	177.747	0.747	0.42%
11		1.2	222	221.669	0.331	0.15%
12		1.2	277	276.360	0.640	0.23%
13		8	27	26.248	0.752	2.79%
14		8	72	73.112	1.112	1.54%
15		8	143	143.227	0.227	0.16%
16		8	177	178.008	1.008	0.57%
17		8	222	221.356	0.644	0.29%
18		8	277	276.052	0.948	0.34%
19		30	27	26.462	0.538	1.99%
20		30	72	75.246	3.246	4.51%
21		30	143	142.680	0.320	0.22%
22		30	177	177.348	0.348	0.20%
23		30	222	221.607	0.393	0.18%
24		30	277	281.452	4.452	1.61%
25		75	27	26.741	0.259	0.96%
26		75	72	72.991	0.991	1.38%
27		75	143	142.660	0.340	0.24%
28		75	177	176.502	0.498	0.28%
29		75	222	219.095	2.905	1.31%
30		75	277	280.062	3.062	1.11%
Mean Value						**0.62%**
Standard Deviation						**0.013**

Table 6.10 provides information about Percentage Error of fault distance determined by locator for faults under many conditions, the best scenario was at distance 177 Km for fault resistance at 0.07 Ω whereas the worst scenario was at distance 27 Km for fault resistance at 0.07 Ω. The value of SD is 0.013 that indicates the lower the dispersion, and error readings are close to the average.

6.5.4 Localization of Three Phase Faults

6.5.4.1 Training of Three Phase Fault Locater

To train the NN-based three Phase Fault locater, the dataset generated under the same conditions will be used that consists of 1160 cases. Figure 6.36 & Figure 6.37 show the selected 15-16-16-16-1 NN structure and its performance scheme to determine the three-phase fault location. This NN has MSE of 1.35.

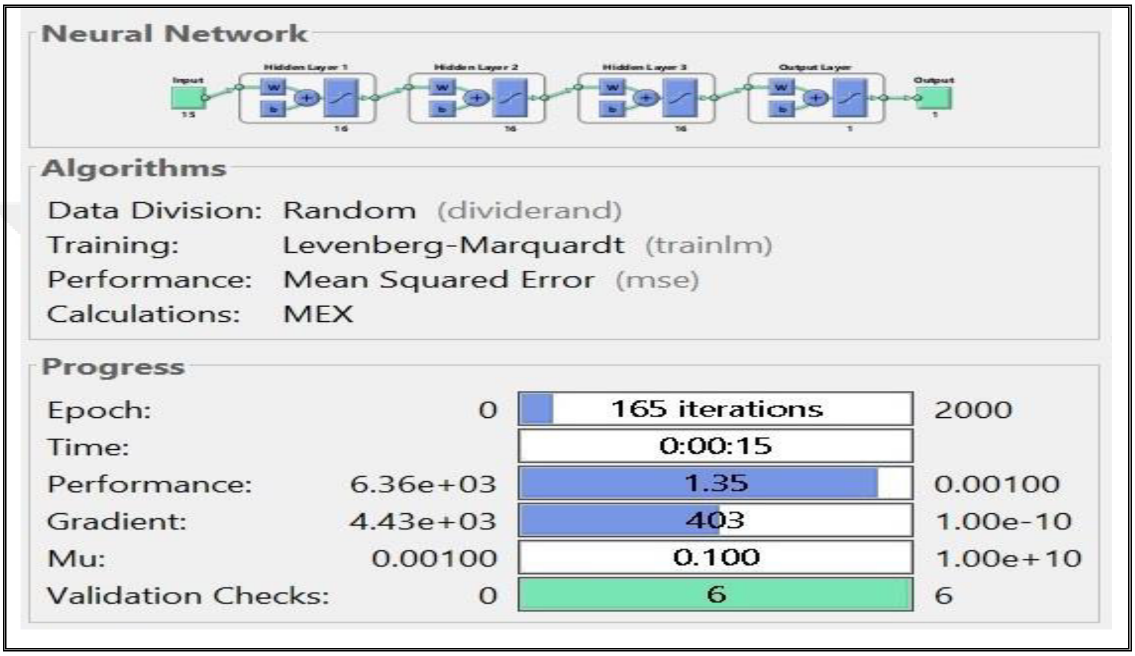


Figure 6.36 Structure of ANN for 3-phase FL (15-16-16-16-1)

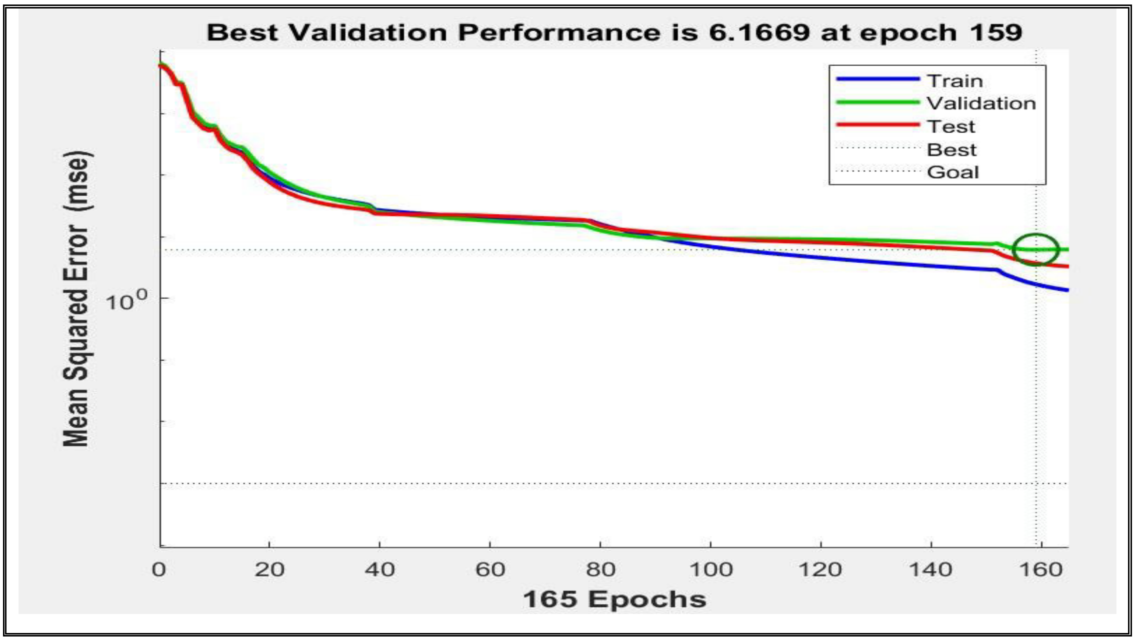


Figure 6.37 Training performance curve of ANN for 3-phase FL (15-16-16-16-1)

6.5.4.2 Evaluating of Three Phase Fault Locater

After completing the training process, the performance of ANN is evaluated through plotting the regression scheme as shown in Figure 6.38. which indicates that all the NN actual outputs have values near their target outputs, the actual value for all outputs can be represented with the following equation:

$$\text{Actual Output} \approx 1 \times \text{Target Output} \pm 0.012 \text{ km} \tag{6.11}$$

From Figure 6.38, the correlation coefficient equals to 0.999 which indicates excellent association between target and actual output of locator.

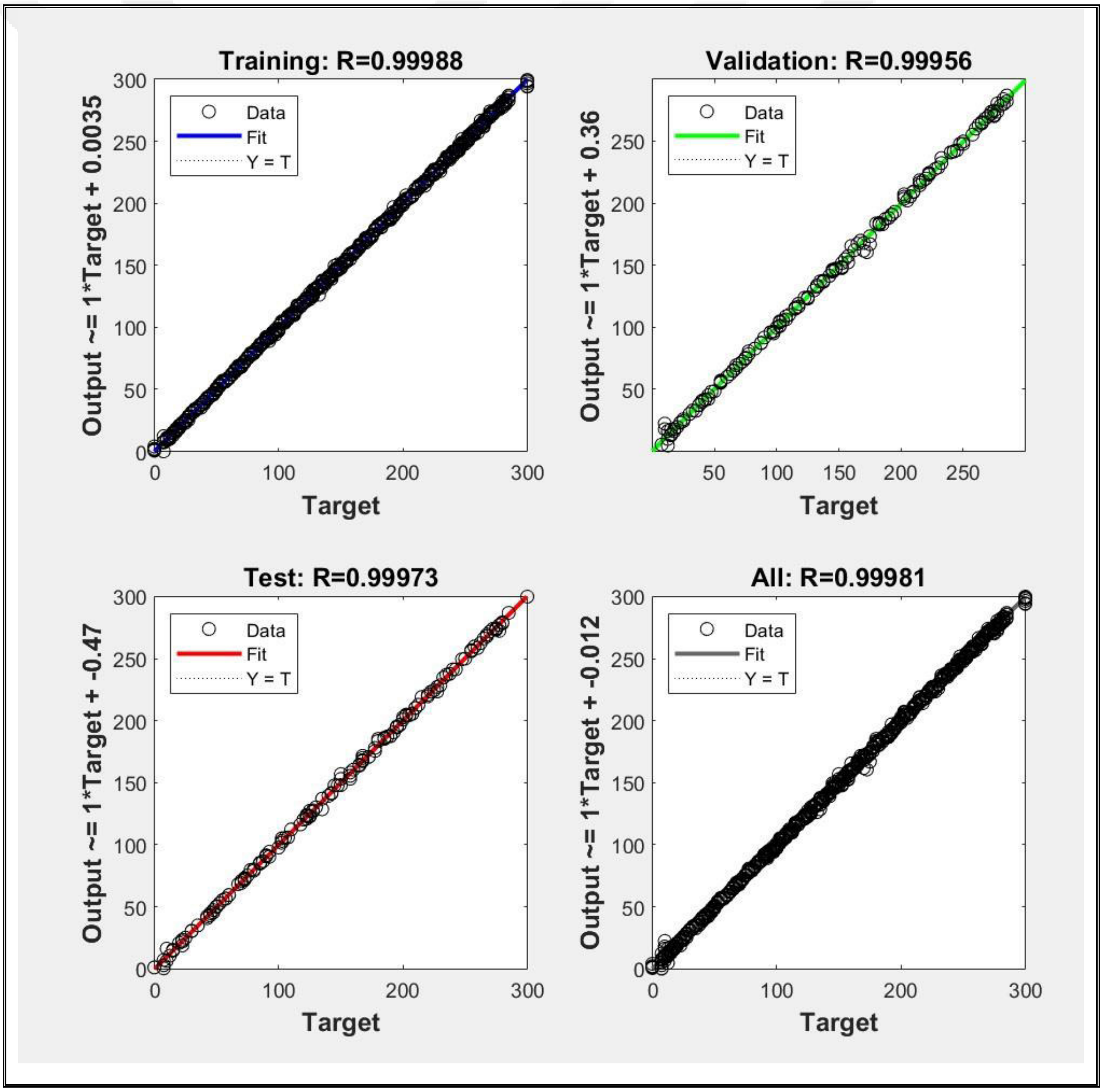


Figure 6.38 Regression scheme of training process for 3-phase FL

6.5.4.3 Testing of Three Phase Fault Locator

To test the performance of the selected NN of three phase fault locator, the ANN structure has been simulated for the same cases used for testing of the other locators. Figure 6.39 illustrates the regression scheme of the relation between actual and target outputs of three phase fault locator for testing data.

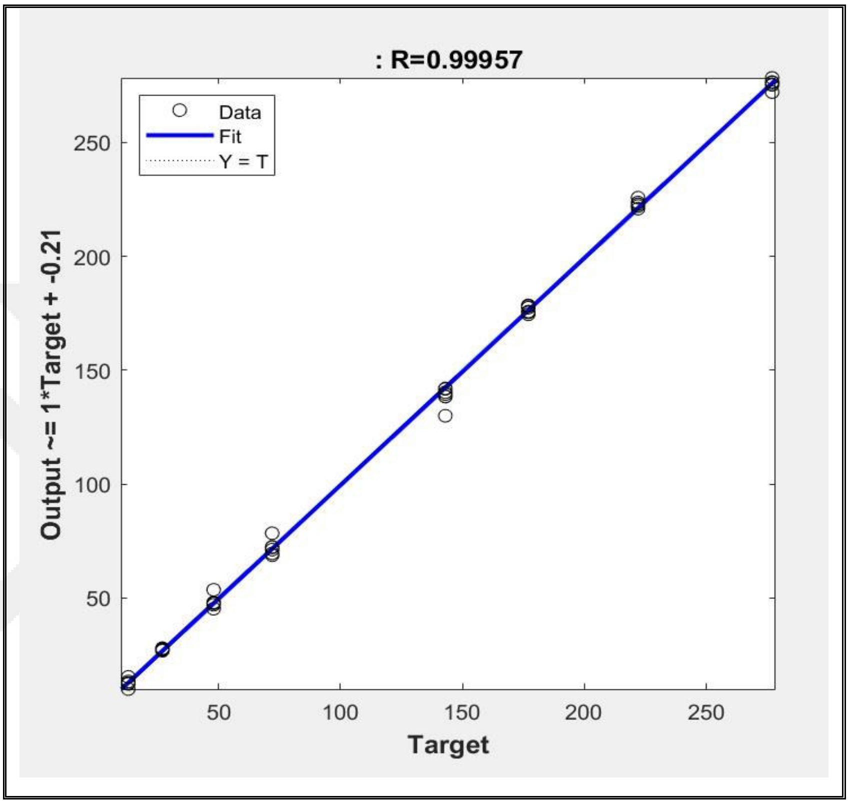


Figure 6.39 Regression scheme of testing process for 3-phase FL

From Figure 6.39, the actual output of NN for testing data can be represented by the following equation:

$$\text{Actual Output} \approx 1 \times \text{Target Output} \pm 0.21 \text{ km} \qquad (6.12)$$

Output values in this Figure relatively approaches their target values. These results are acceptable since the errors are considered very small compared to the HVTL length.

Table 6.11 shows examples of these actual and target values of ANN output and percentage of error under different conditions.

Table 6.11 Percentage error of ANN selected for three phase fault locator.

Fault Conditions				Actual Output of ANN	Error Value	Percentage Error (%)
N.	Fault Type	Fault resistance (Ω)	Fault Distance (Km)			
1	ABC	0.07	27	27.129	0.129	0.48%
2		0.07	72	72.217	0.217	0.30%
3		0.07	143	141.759	1.241	0.87%
4		0.07	177	177.526	0.526	0.30%
5		0.07	222	222.891	0.891	0.40%
6		0.07	277	276.492	0.508	0.18%
7		1.2	27	27.382	0.382	1.42%
8		1.2	72	72.231	0.231	0.32%
9		1.2	143	140.339	2.661	1.86%
10		1.2	177	178.169	1.169	0.66%
11		1.2	222	223.660	1.660	0.75%
12		1.2	277	272.150	4.850	1.75%
13		8	27	26.834	0.166	0.62%
14		8	72	71.219	0.781	1.08%
15		8	143	138.418	4.582	3.20%
16		8	177	178.485	1.485	0.84%
17		8	222	223.689	1.689	0.76%
18		8	277	278.383	1.383	0.50%
19		30	27	27.949	0.949	3.51%
20		30	72	69.583	2.417	3.36%
21		30	143	138.537	4.463	3.12%
22		30	177	174.570	2.430	1.37%
23		30	222	225.791	3.791	1.71%
24		30	277	276.370	0.630	0.23%
25		75	27	27.711	0.711	2.63%
26		75	72	78.424	6.424	8.92%
27		75	143	130.022	12.978	9.08%
28		75	177	175.420	1.580	0.89%
29		75	222	220.969	1.031	0.46%
30		75	277	275.293	1.707	0.62%
Mean Value						**0.93%**
Standard Deviation						**0.022**

Table 6.12 provides information about Percentage Error of fault distance determined by locator for faults under many conditions, the best scenario was at distance 277 Km for fault resistance at 0.07 Ω whereas the worst scenario was at distance 143 Km for fault resistance at 75 Ω. The value of SD is 0.022 that indicates the lower the dispersion, and error readings are close to the average.

CHAPTER VII

CONCLUSION AND RECOMMENDATIONS

7.1 Conclusion

Hybrid approach based on combination between wavelet transforms and artificial neural network by employing back propagation algorithm have been proposed, trained, evaluated, and tested for detecting occurrences of faults, classifying the type of fault, and localizing the distance of fault on HV power transmission line connecting two stations. This technique can be applied in digital protection systems utilized at substations for power transmission lines.

The model applied to generate data required to train and test the NN is a 300 Km long, 400 kV, single circuit, overhead line, SIMULINK/MATLAB software has been utilized for modeling the TL. The measurement of fault current for one cycle for various types of faults under different conditions in terms of faults resistance, fault type, and position of fault have been performed at one end. The measured values have been decomposed by using wavelet transform employed as tool for processing the transient signals, from wavelet coefficients, the features employed as inputs to ANNs have been extracted, these features are values of Energy, RMS, Max, Min, and SD for the fourth level detail coefficients.

For fault detector, three features (Energy, RMS, and Max) have been used to train ANN with three layers and six neurons for each. The output of detector was with no error for all tested cases and accuracy of fault detection was 100%.

For fault classifier, five features (Energy, RMS, Max, Min, and SD) have been used to train ANN with three layers and twelve neurons for each. The output of classifier was very accurate, and accuracy of fault classification was 99.7%.

To localize the distance of the fault on TL after detecting and classifying of the fault, the ANN-based fault locator has been utilized for each type of the faults, four main categories of faults have been studied.

For category of phase to ground faults, three types of faults have been studied which are (A-G, B-G, and CG), five features (Energy, RMS, Max, Min, and SD) have been used to train ANNs with three layers and fifteen neurons for each for A-G locator and sixteen neurons for B-G & C-G locators. The output of locators for this category was very accurate with mean value of percentage error less than 0.8% for tested cases.

For category of phase-to-phase faults, just one type of faults has been studied which is (A-B), five features (Energy, RMS, Max, Min, and SD) have been used to train ANN with three layers and eighteen neurons for each. The output of locator for this type was accurate with mean value of percentage error equals 1.42% for tested cases.

For category of phase-to-phase to Ground faults, just one type of faults has been studied which is (AB-G), five features (Energy, RMS, Max, Min, and SD) have been used to train ANN with three layers and eighteen neurons for each. The output of locator for this type was accurate with mean value of percentage error equals 0.62% for tested cases.

For category of three phase faults, just one type of faults has been studied which is (ABC), five features (Energy, RMS, Max, Min, and SD) have been used to train ANN with three layers and sixteen neurons for each. The output of locator for this type was accurate with mean value of percentage error equals 0.93% for tested cases.

The test results illustrate very high accuracy in detecting and classifying transmission line faults, and high accuracy in determining the fault distance for phase to ground, phase-to-phase to Ground, and three phase faults, while for phase-to-phase faults, the accuracy was favorable.

7.2 Recommendations:

This work can be extended by increasing the decomposition levels of wavelet transform, and/or changing the sampling frequency, and/or employing possible neural network architectures, to perform a comparative analysis about impacting each of these factors on accuracy and performance, and how we can improve accuracy especially for determining the distance of fault for category of phase-to-phase faults.

REFERENCES

[1] Valsan, S. P. and K. S. Swarup. (2009). Wavelet Transform Based Digital Protection for Transmission Lines, *International Journal of Electrical Power & Energy Systems*. **31(7-8)**, 379-388.

[2] Saravanababu, K., Balakrishnan, P., & Sathiyasekar, K. (2013). Transmission Line Faults Detection, Classification, and Location Using Discrete Wavelet Transform, *International Conference on Power, Energy and Control (ICPEC)*. 233-238.

[3] Eid, G., & Mohamed, S. H. (2016). A New Approach of Power System Protection Based on Wavelet Transformation, *Journal of Electrical and Electronics Engineering*. **9(1),** 13.

[4] Matarweh, J., Mustaklem, R., Saleem, A., & Mohamed, O. (2019). The Application of Discrete Wavelet Transform to Classification of Power Transmission System Faults, *IEEE Jordan International Joint Conference on Electrical Engineering, and Information Technology (JEEIT).* 699-704.

[5] Jamil, M., Sharma, S.K. & Singh, R. (2015). Fault Detection and Classification in Electrical Power Transmission System Using Artificial Neural Network, *SpringerPlus.* **4(1),** 1-13.

[6] Tağluk, M. E., Mamiş, M. S., Arkan, M., & Ertuğrul, Ö. F. (2015). Detecting Fault Type and Fault Location in Power Transmission Lines by Extreme Learning Machines, *23nd Signal Processing and Communications Applications Conference (SIU)*. 1090-1093.

[7] Akmaz, D., Mamiş, M. S., Arkan, M., & Tağluk, M. E. (2018). Transmission Line Fault Location Using Traveling Wave Frequencies and Extreme Learning Machine, *Electric Power Systems Research*. **155**, 1-7.

[8] Resmi, R., Vanitha, V., Aravind, E., Sundaram, B. R., Aswin, C. R., & Harithaa, S. (2019). Detection, Classification and Zone Location of Fault in Transmission Line Using Artificial Neural Network, *IEEE International Conference on Electrical, Computer and Communication Technologies (ICECCT).* 1-5.

[9] Maheshwari, A., & Sharma, S. (2018). Classification of Fault Identification Techniques in Power System Transmission Lines, *International Journal of Emerging Technologies and Innovative Research.* 89-98.

[10] Koley, E., Verma, K. & Ghosh, S. (2015). An Improved Fault Detection Classification and Location Scheme Based on Wavelet Transform and Artificial Neural Network for Six Phase Transmission Line Using Single End Data only, *SpringerPlus.* **4(1),** 1-22.

[11] Saini, M., bin Mohd Zin, A.A., Bin Mustafa, M.W., Sultan, A.R., Rahimuddin. (2016). Transmission Line Using Discrete Wavelet Transform and Back-Propagation Neural Network Based on Clarke's Transformation, *Applied Mechanics and Materials.* **818**, 156–165.

[12] Abdullah, A. (2017). Ultrafast Transmission Line Fault Detection Using a DWT-Based ANN, *IEEE Transactions on Industry Applications*. **54(2)**, 1182-1193.

[13] Rao, Y.S., Kumar, G.R., & Rao, G.K. (2017). A New Approach for Classification of Fault in Transmission Line with Combination of Wavelet Multi Resolution Analysis and Neural Networks, *International Journal of Power Electronics and Drive Systems.* **8(1)**, 505-512.

[14] Ray, P., Mishra, D. P., Dey, K., & Mishra, P. (2017). Fault Detection and Classification of a Transmission Line Using Discrete Wavelet Transform & Artificial Neural Network, *International Conference on Information Technology (ICIT)*. 178-183.

[15] Zakri, A. A., Darmawan, S., Usman, J., Rosma, I. H., & Ihsan, B. (2018). Extract Fault Signal via DWT and Penetration of SVM for Fault Classification at Power System Transmission, *2nd International Conference on Electrical Engineering and Informatics (ICon EEI)*.191-196.

[16] Khaleghi, A., Sadegh, M. O., Ghazizadeh-Ahsaee, M., & Rabori, A. M. (2018). Transient Fault Area Location and Fault Classification for Distribution Systems Based on Wavelet Transform and Adaptive Neuro-Fuzzy Inference System (ANFIS), *Advances in Electrical and Electronic Engineering*. **16(2),** 155–166.

[17] Malla P., Coburn W., Keegan K., Yu XH. (2019). Power System Fault Detection and Classification Using Wavelet Transform and Artificial Neural Networks, *Advances in Neural Networks*. 266-272.

[18] Paithankar, Y. G., Bhide, S. R. (2003). *Fundamental of power system protection*. Prentice-Hall: india.

[19] Hadi Saadat. (2010). *Power system analysis*. Third Edition, PSA.

[20] Gonzalez-Longatt, Francisco. (2019). *Chapter 1. Introduction to power systems.*

[21] Aker, E., Othman, M.L., Veerasamy, V., Aris, I., Wahab, N.I.A., Hizam, H. (2020). Fault Detection and Classification of Shunt Compensated Transmission Line Using Discrete Wavelet Transform and Naive Bayes Classifier, *Energies*. **13(1)**, 243.

[22] Leelaruji, R., & Vanfretti, L. (2011). Power System Protective Relaying: basic concepts, industrial-grade devices, and communication mechanisms, *KTH Royal Institute of Technology Internal Report*.

[23] Anderson. P. M. (1999). *Power system protection*. Institute of Electrical and Electronics Engineers-Inc: New York.

[24] Gerhard Ziegler. (2011). *Numerical distance protection*. Erlangen: Germany.

[25] Matlab Program, Wavelet Toolbox™ 4 User Guide.

[26] Erişti, H., Uçar, A., & Demir, Y. (2010). Wavelet-Based Feature Extraction and Selection for Classification of Power System Disturbances Using Support Vector Machines, *Electric power systems research*. **80(7),** 743-752.

[27] Pallavi, V. P., Pallavi, R. N., Abhiruchi, R. V. (2018). Transmission Line Fault Detection, Classification and Location by Using DWT (Discrete Wavelet Transform), *International Journal of Innovative Research in Electrical, Electronics, Instrumentation and Control Engineering*. **6(3)**, 51-56.

[28] Haykin, S. S. (2009). *Neural networks and learning machines*, New York.

[29] Fausett, L. V. (2006). *Fundamentals of neural networks: architectures, algorithms and applications.* Pearson Education: India.

[30] Hagan, M. T., Demuth, H. B., Beale, M. H., & De Jesús, O. (2014). *Neural network design.* Martin Hagan: Oklahoma.

[31] Msatechnosoft. 2018. Artificial Neural Network, Types, Feed Forward, Feedback, Structure, Perceptron, Machine Learning, Applications. Msatechnosoft Intl.: www.msatechnosoft.in., 21.05.2018.

[32] Geeksforgeeks. 2020. Supervised and Unsupervised learning. Geeksforgeeks Int2.: www.geeksforgeeks.org., 19.12.2020.

[33] Thwe, E.P. and Oo, M.M. (2016). Fault Detection and Classification for Transmission Line Protection System Using Artificial Neural Network, *Journal of Electrical and Electronic Engineering*. **4(5)**, 89-96.

[34] Hasabe, R. P., & Vaidya, A. P. (2014). Detection and Classification of Faults on 220 KV Transmission Line Using Wavelet Transform and Neural Network, *International Journal of smart Grid and Clean Energy*. **3(3)**, 283-290.

[35] Patel, M., & Patel, R. N. (2015). Fault Analysis in Transmission Lines Using Neural Network and Wavelets, *International Conference on Signal Processing and Integrated Networks (SPIN).* 719-724

[36] Mathworks. 2020. Multilevel 1-D wavelet decomposition-MATLAB wavedec. Int2.: www.mathworks.com., 2020

[37] Mathworks.2020. Energy for 1-D wavelet or wavelet packet decomposition. Int2.: www.mathworks.com., 2020