Effect of Pre-processing on Using ANN and ANFIS

M.A. Moustafa Hassan*
Heliopolis University, Cairo, Egypt
*On Sabbatical Leave from Electrical Power Dept., Cairo University, Giza, Egypt (mmustafa@eng.cu.edu.eg)

Abstract— **Rotating machines are widely utilized in industrial lifecycle, since it represents a vital element in industrial processes. Therefore, quick detection of faults of rotating machines is necessary to sidestep the forced stopping for frequent maintenance in industrial progressions. Several condition monitoring and detecting procedures are used to diagnose the rotating machinery faults based on currents values, vibration signature analysis, temperature monitoring, noise signature analysis, lubricant signature analysis using Artificial Intelligence (AI) techniques. Many AI methods are in use for detection of faults of rotating machines. For instance, Fuzzy Inference System (FIS); Artificial Neural Networks (ANNs) and Adaptive Neuro-Fuzzy Inference System (ANFIS) are among AI techniques are advanced technologically for classifying and detection various rotating machinery faults. This research work describes a comprehensive methods to some extent for detecting and classification of rotating machines faults using two methods of artificial intelligence which are ANN and ANFIS. This study is implemented offline in MATLAB environment based on data preprocessing before applying ANN or ANFIS. The obtained data of rolling element were classified into different main conditions. The obtained data after preprocessing were imported to ANNs and ANFIS models. These studies work presents a comparison between the diagnosing based on ANNs and ANFIS. The input data were preprocessed before entering to ANNs and ANFIS models by using some techniques: the normalized data in range (0-1), the frequency domain analysis via discrete wavelet transform, the time domain features, and finally the Auto Regressive (AR) model. The accomplished outcomes of these preprocessing techniques give high accuracy results in faults detection and classification issues. The accomplished outcomes are encouraging and promising in the field of diagnosis of machinery faults.**

I. Introduction

The industrial technology has progressed toward high production rate, large scale machine and automated manufacturing machine where, the technological growth of modern equipment provide economic benefits. Rotating machines are the vital components of mechanical mechanisms that commonly used in various fields such as aviation industry, power generation division, machinery industry, metallurgical industry and petrochemical industry. The parts of rotating machines such as induction motors, gears, belts, and bearing are exposed to corrosion, wear, deficiencies in material, and fatigue due to the dynamic motion of mechanical mechanisms, in addition to wrong installation and assembly conditions, and service conditions such as inappropriate maintenance.

These significant topics found wide efforts from researchers. For ANNs, some research works introduced numerous AI techniques such as NNs, fuzzy logic, neural-fuzzy, Adaptive Neuro-Fuzzy Inference System (ANFIS) and Genetic Algorithms (GAs) to diagnose the faults of motors [1-8]. The numerous approaches such as the Artificial Intelligence (AI) techniques used to increase the accuracy for diagnosing of rotating machines faults and overcome the drawbacks of the traditional techniques. ANN and ANFIS are the common widely methods of Artificial Intelligence for diagnosing the faults of rotating machinery. The most public technique of monitoring the status of rolling element bearing is by using vibration signal analysis or current signals analysis as inputs to Induction motors. However, most of these techniques give reasonable results without data preprocessing. As presented in [3-7].

In this article; a common study for Artificial Intelligence techniques was given. It focus on neural networks properties, architecture and training are presented. A historical review of neural networks is introduced. Neural networks learning, generalization, parallelism, distribution, and pattern classification are presented. Architecture of neural networks, which is a number of layers, each layer, consists of number of neurons. Training neural networks depending on back propagation algorithm is detailed. In this chapter discusses a public study for fuzzy logic system properties. Fuzzy inference system algorism and defuzzification are displayed. In addition, this article introduces the adaptive neural fuzzy inference system properties, ANFIS architecture and learning algorithm of ANFIS. Furthermore, some of data preprocessing were introduced to enhance the obtained results.

This article is organized as follow: Section II introduces Artificial Intelligence techniques; whereas Section III presents Artificial Neural Networks and its architecture. While; Section IV summarizes the Fuzzy Inference System. Meanwhile; Section V illustrates Adaptive Neuro Fuzzy Inference System. Moreover; Section VI discusses the preprocessing for the input data of AI techniques. Also, Section VII discusses the rolling element bearing faults diagnosis methods. Furthermore Section VII illustrates the results of Diagnosis of Induction Motor faults. Finally, Section IX concludes the article and elaborates on the future work.

II. Artificial Intelligence Techniques

Artificial Intelligence (AI) technique is a high potential data processing tool that recreates fault diagnosis techniques with a significant impact. Artificial Neural Networks (ANN), Fuzzy Logic (FL), and Adaptive Neuro-Fuzzy Inference System

(ANFIS) are modern fault diagnosis and detection techniques arise under AI field as illustrated in [1]. ANN is imitation of the human brains intellectual progression. It has similar parallel processing, self-organizing, self-learning, classification and non-linear mapping abilities as given in many published research work. Fuzzy Logic (FL) can model subjective complexity nonlinear functions to a desired degree of accuracy as presented in [2]. In this research work, data pre-processing for ANN and ANFIS techniques will be investigated. Artificial Intelligence (AI) techniques are in use for classifying and detection various rotating machinery faults as illustrated in [3-7]. AI techniques are advanced technologically for classifying and detection various rotating machinery faults. AI techniques have been used for the diagnosis of ITSC and BRB faults of an IM as well as bearing defects diagnosis as presented in [3-8]. The supervised learning (gradient descent) algorithm is used here to train the weights to minimize the errors. In this section, application of ANN and ANFIS architecture takes into account the positive features of both the ANN and fuzzy logic technology for classifying and detection various rotating machinery faults. A brief discussion on ANN as well as ANFIS architectures will be elaborated.

III. Artificial Neural Networks

This section presents the neural networks properties. Furthermore it introduces a historical review of neural networks. It discusses learning and generalization and neural network architecture.

A) *Historical Review of Neural Networks.*

Neural networks are networks of nerve cell (neurons) in the brain. The human brain has billions of individual neurons and trillions of interconnections. It was found that the brain consists of large number of highly connected neurons which apparently can send very simple excitatory and inhibitory messages to each other and update their excitations on the basis of these simple messages. When the transmitters are released, the membrane of the target neuron is affected and its inclination to fire its own impulse is either increased or decreased according the condition whether the incoming signal is either excitatory or inhibitory as discussed in [9-12]. In 1943, McCulloch and Pitts proposed a mathematical model of the neurons and showed how neuronal-like networks could be computed. The schematic model of the McCulloch and Pitts neuron is described in [9, 11] and Figure (1).

The similarity between Biological and Artificial Neural Networks is presented in Table (1) as presented in [10].

The mathematical model of the neuron by McCulloch and Pitts is described as follows and explained in [10]:

$$net_j = \sum_{i=0}^{n} W_{ji}\, O_i + b_j \qquad (1)$$
$$O_i = f(net_i)$$

where:

- net_j is the net input to the neuron j
- b_i is a threshold value
- w_{ji} is the strength and sense of connection from a neuron i to a neuron j
- O_i is an output signal of a neuron i
- $f(x)$ is an output function or an activation function of a neuron.

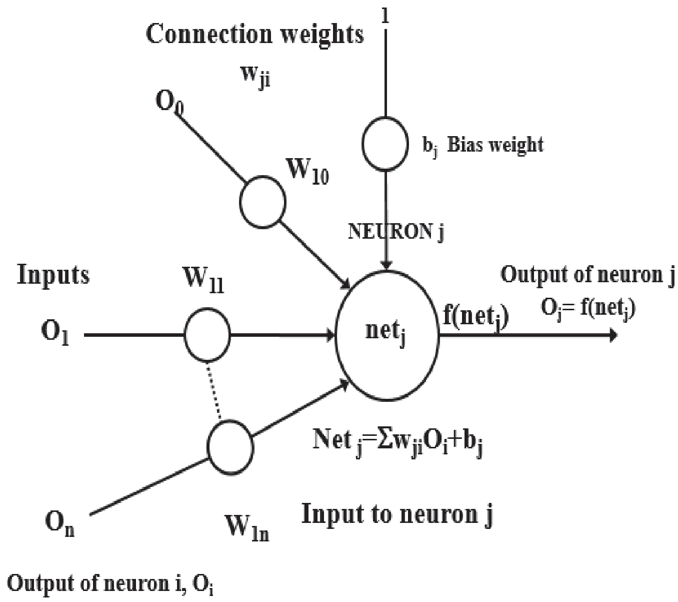


Fig.1. Schematic Model of a McCulloch and Pitts

TABLE I. ANALOGY BETWEEN BIOLOGICAL AND ANN

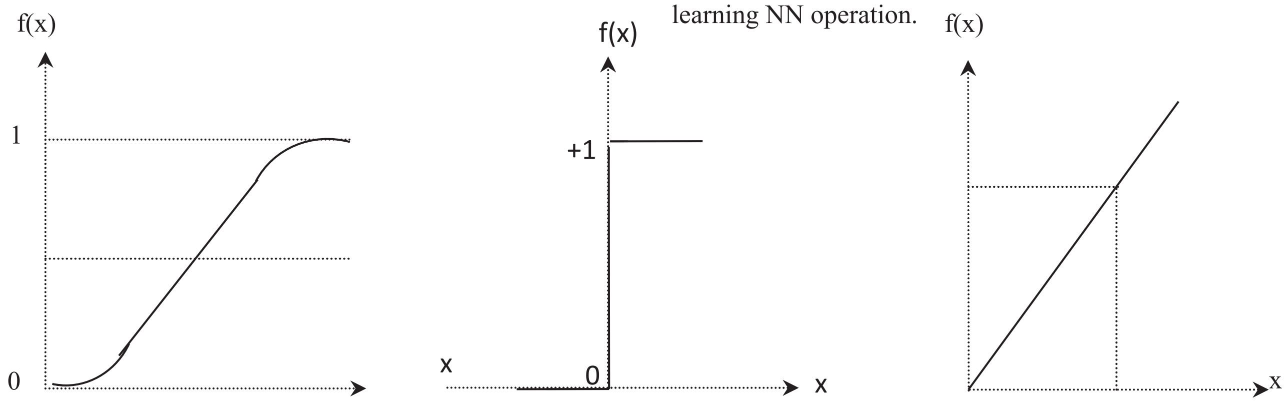
Biological Neural Network	Artificial Neural Network
Soma	Neuron
Dendrite	Input
Axon	Output
Synapse	Weight

Figure (2) displays different types of the neuron activation functions as reviewed in [9-12]. The famous activation functions are explained in [11, 12]:

i) *Hard limiter or sign function*
ii) *Sigmoid function*
iii) *Linear Function*

B) *Learning and Generalization*

The two most important properties of the neural networks are their capability to learn and to generalize. The neural networks learn to recognize certain patterns and give the correct output response to these patterns. The neural network generalization is that, the network ought to give the correct response even to patterns that it has not explicitly been trained and then the system is able to infer the general properties of the different classes of patterns from the given patterns as provided in [11, 12]. Neural network Learning methods can be classified into two categories, supervised and unsupervised.



Sigmoid logistic function. *Hard limiter function.* *Linear function.*

Fig. 2. Different Types of the Activation Functions

C) SUPERVISED LEARNING NEURAL NETWORK

The supervised Neural Network (NN) is trained offline at first with the input output data of the faulty and healthy cases. The training changes the network weights. The inputs data for training the NN are the system measured data and the desired output of NN. Perceptron learning rule, Widrow-Hoff learning rule, correlation and outstare learning rules are the examples of supervised learning of neural network as indicated in [13]. Figure (3) specifies the supervised learning NN operation.

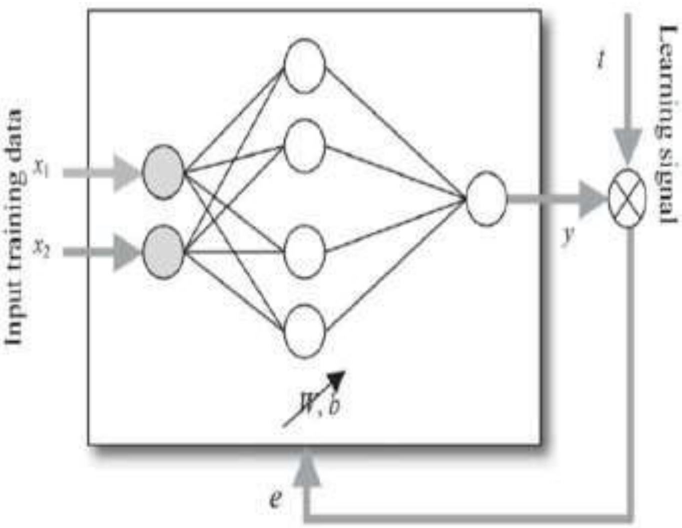


Fig.3. Supervised Learning Schematic

D) UNSUPERVISED LEARNING NEURAL NETWORK

This method of learning is not supervised. This means that target is unknown. It has internal representation to respond for different input patterns with the different parts of NN. This learning technique is based on input signal feature, regularities and trends for network configuration and weight adjustment. Hebbian learning rule, Winner-Take all learning rule are example of unsupervised learning. The unsupervised learning NN is suitable for data clustering and signal categorization as specified in [11, 12]. Figure (4) introduces the unsupervised learning NN operation.

E) PATTERN CLASSIFICATIONS

Neural networks can be used to explain a problem if that problem can be reduced to pattern classification. All neural networks are pattern classifier, but not all pattern classifier are neural networks due to the pattern classifiers may be empirical implementation. So that a neural network can be formed to solve many problems by means of pattern classification. These problems usually comprise the recognition of something, which is variable, which cannot be entirely described or predicted such as recognizing faces or fingerprints, identifying spoken words, and reading, and safety systems. All of these problems cannot be expected as detailed in [9, 10].

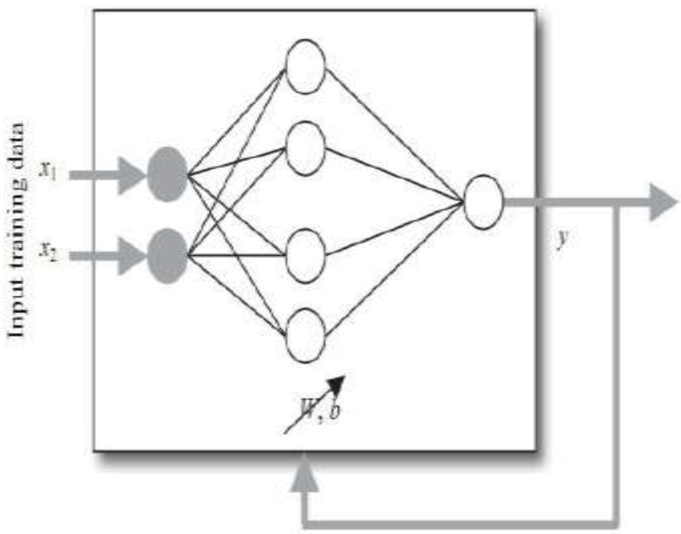


Fig.4. Unsupervised Learning Schematic

F) PARALLELISM AND DISTRIBUTION

The work of neural networks is in parallel and distributed among the processing elements. Therefore; ANN are described as a parallel-distributed processor. The parallelism in the ANN tends to be an advantage that is the prospective for very high processing speeds, and distributing process in neural networks.

Due to the parallelism and the distribution in neural networks, if a part of the network breakdowns, the whole system could still remain to operate. Although it will be less than perfectly which is preferable to complete failure as proposed in [9, 10].

G) Neural Network Architecture

Network architecture is defined by the basic processing elements and the way in which they are interconnected as displayed in [9-12]. The ways that have been developed of incorporating learning and generalization into electronic devices are divided into two categories. The first one is Boolean neural networks, which are a structure built from conventional Boolean logic gates, not biologically inspired. The second one is the biologically inspired neural networks which are a neural networks composed of elements, namely neuron, analogy with neurophysiology or perceptron, which direct descendants of the model of biological neuron are created by McCulloch and Pittsas as illustrated in [10]. These elements may be interconnected of several layers, multi-layer feedforward networks, which gives the possibility of more complex nonlinear mapping between the inputs and the outputs. A neuron in one layer receives inputs only from the previous layer. There is no feedback in such network as given in [10]. Backpropagation learning algorithm is used for automatic adjustment of network, perceptron and weights.

H) Neuron Model

An elementary neuron with n inputs is presented in [10-11] and Figure (5). Each input is weighted with an appropriate w. The weighted inputs are added to the bias to form the input of the transfer function f. Neurons may use any differentiable transfer function f to generate their output, linear, hard-limiter, or sigmoid function as discussed in [11].

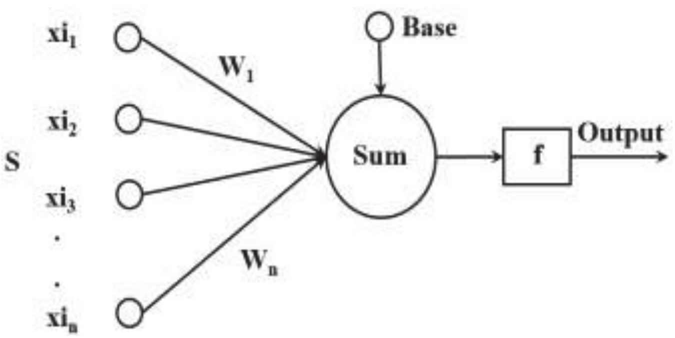


Fig.5. Artificial Neuron Model

I) Single Layer Perceptron

A single-layer feedforward network of m transfer function neurons having n inputs and w is the transformation matrices that transform the vector xi to the vector O, is detailed in Figure (6).
Single layer perceptron has a limitation that it is only able to implement linearly separable functions. This limitation can be overcome using multi-layered perceptron as described in [10].

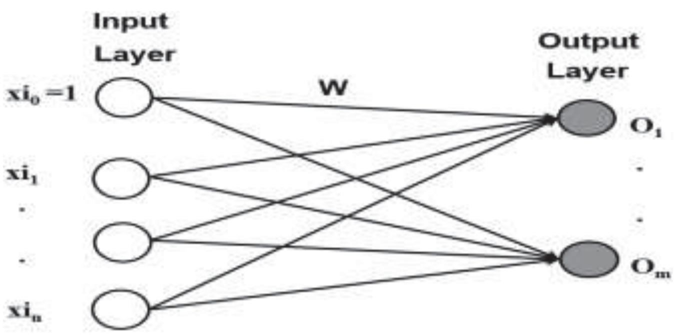


Fig.6. Single-Layer Network

J) Multi-Layered Perceptron

Neural networks with one or more hidden layers (it is called hidden layer because neither the input nor the outputs of the processing elements of these layers can be seen from the outside) are called multilayer neural networks or Multi-Layer Perceptron (MLP) as reviewed in [10-12]. Normally, each hidden layer of a network uses the same type of activation function. The activation function of output is either sigmoidal or linear. The MLP can be used in cases where the shapes of the class boundary are complex and linearly not separable. Minimum amount of neurons and number of instances are necessary to program given task into MLP as explained in many literatures. There is no analytical method for finding the neurons number in the hidden layer. Therefore it only found by trial and error as studied in [6,7]. There is no clear and exact rule due to complexity of the network mapping. Neurons depend on the function to be approximated and its degree of nonlinearity affects the size of network. Big number of neurons and layers may cause over fitting and may cause decrease in the generalization ability as defined in literature. The MLPs have been termed a universal approximator as depicted in [10]. Figure (7) describes general multilayer perceptron architecture with one hidden layer.
Where Wa and Wb are transformation matrices that transform the $n \times 1$ vector xi to h and h to the $m \times 1$ vector O, respectively as provided in [10].

K) Overview of Neural Network Training Methodology

The methodology, which is followed when training a neural network is indicated in [114] and Figure (6.8).First, data which will be used for training and testing the neural network must be collected or generated. The training data set should cover the input space or should at least cover the space in which the network will be expected to operate. If there is not training data for certain conditions, the output of the network should not be trusted for those inputs. The training data should be set small so that training is fast, but also it is wanted to exercise the input space well, which may require a large training data set. Once the training data set is selected, neural network architecture must be chosen. There are two lines of thought here. Some designers choose to start with a fairly large network that is sure to have enough degrees of freedom to train to the desired Steady State Error (SSE) goal. Then, once the network is trained, they try to shrink the network until the smallest network that trains remains. Other designers choose to start with a small network and grow it until the network trains and its error goal is met. After the network architecture is chosen, the weights and biases are initialized and the network is trained based on as specified in [9]. Once the smallest network that trains to the desired error goal is found, it must be tested with the test data. If the error goal is met, training is complete. If the error goal is not met, the incomplete test data is causing the poor performance, the test patterns that have high error levels should be added to the training data set, a new test set should be chosen, and the network should be retrained. If there is not enough data left for training and testing, data may need to be collected again or be regenerated as detailed in [9-10].

IV. Fuzzy Inference System

A detailed background on fuzzy sets and membership functions and linguistic variables could be found in many literatures as [13]. It also describes the inference system construction. Furthermore, it explains the defuzzification methods. The Fuzzy Logic (FL) emulates the human brain logic. It is a type of expert system. The fuzzy linguistic variable is used to present the human compare logic instead of rigorous mathematical formulas. FL is capable to transform the linguistic variable to numerical values based on the fuzzy rules and membership functions. FL structure main items are the input/output membership functions, linguistic variable and fuzzy rules as studied in [14].The Fuzzy Inference System (FIS) is composed of five functional blocks as defined in [14] and Figure (9).

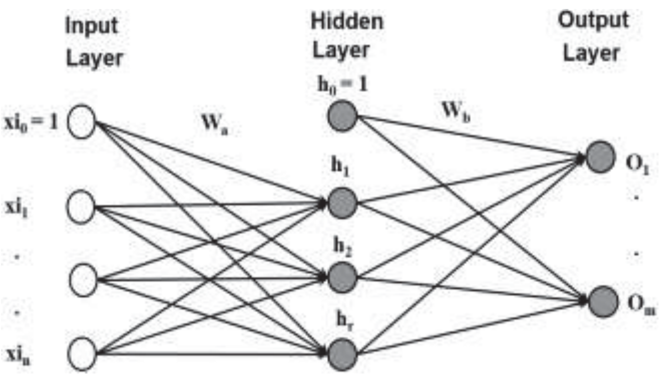


Fig.7. Multi-Layer Network A

Defuzzification

A fuzzy inference system maps an input vector to a crisp output value. The defuzzification process is used to obtain a crisp output. The defuzzification methods are given in [14].

V. Adaptive Neural Fuzzy Inference System

The combination between NN and fuzzy system is to introduce the learning ability to fuzzy system. This combination is called Adaptive Neuro Fuzzy Inference System (ANFIS); as presented in [15]. The linguistic rules and/or memberships functions can learn after the combination. The fuzzy system is an expert system based on rules, on the other side the NN knowledge is based on training with historical data as provided in [16-21].This section introduces a brief idea about the architecture and learning procedure of the Adaptive Neural Fuzzy Inference System (ANFIS).

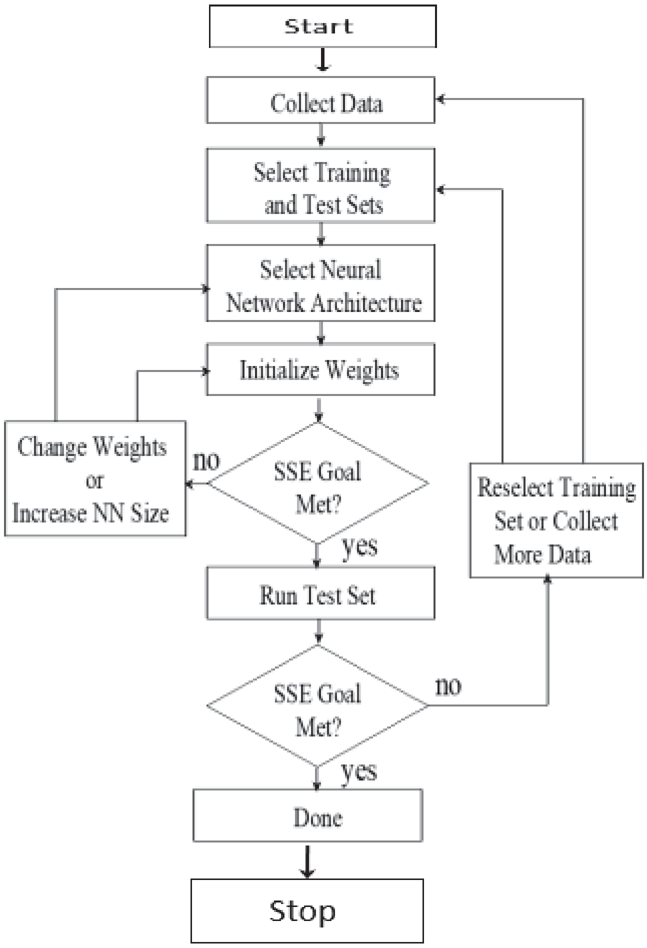


Fig.8. Neural Network Training Methodology

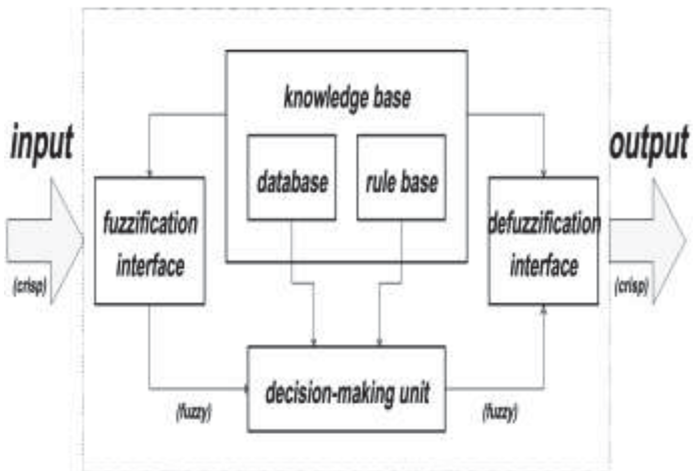


Fig.9. Fuzzy Inference System Schematic

A. ANFIS Architecture

ANFIS consists of *IF-then rules*, training and learning algorithms. For the FIS, consider a system with two inputs (x, y) and one output (f) as indicated in [16 -21], the fuzzy rules based on 1st order Sugeno type are:

Rule1: IF x is A_1 and y is B_1 Then $f_1=p_1x+q_1y+r_1$,
Rule2: IF x is A_2 and y is B_2 Then $f_2=p_2x+q_2y+r_2$,

where $A_{1,2}$, $B_{1,2}$ are the fuzzy set, $f_{1,2}$are the system outputs within the specified fuzzy rules and the $p_{1,2}$, $q_{1,2}$ and $r_{1,2}$ are the design parameters based on the ANFIS training as specified in [16 -21]. The ANFIS network consists of five layers as presented in Figure (10). An extra normalization layer is added to the Neuro-Fuzzy Network as detailed in [16 -21].

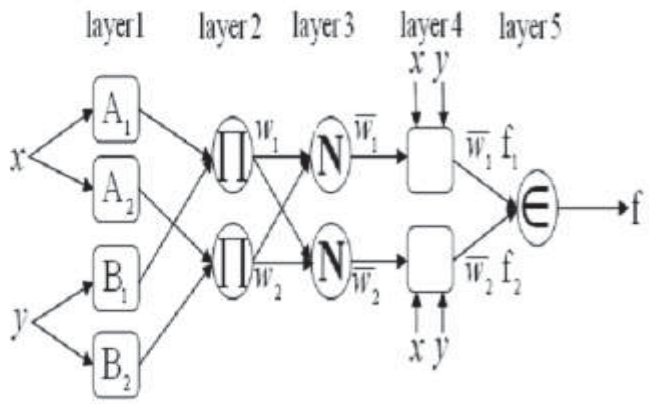


Fig.10. Equivalent ANFIS Architecture

Layer1 is the fuzzification layer adaptive nodes with bell membership function.
Layer 2 is the layer of rules where its output is located as the fire strength of each node.
Layer 3 is the normalization layer and its output is the normalized fire strength:
Layer 4 represents the consequent layer and its output is the product of normalized firing strength and the fuzzy rules consequent polynomial:
Layer 5 is the defuzzification layer which has only one node (output node) and its output is the overall ANFIS output; Which is Summation of the layer 4 outputs.

B. ANFIS Learning Algorithm

Tuning all the adaptable parameters is the task of the learning algorithm for this architecture in order to make the training data match the ANFIS output. When the premise parameters a_i and c_i of the membership function are fixed, the output of the ANFIS model can be written as Equation (6.23):

C. Fuzzy Model Optimization

The ANFIS is the final model for the Root Mean Square Error (RMSE) is the minimum. The consequent parameters of the fuzzy model are updated using the least squares estimation algorithm. This updating leads to the optimization of the premise parameters of the fuzzy membership functions to give the final fuzzy model as proposed in [16-21].

RMSE value nearer to zero shows the accuracy of predicted values as displayed in Equation (2) based on [20-21].

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(x-y)^2} \quad (2)$$

Where x is the actual value, y is the estimated value and n is the number of samples.

VI. PREPROCESSING FOR THE INPUT DATA OF AI TECHNIQUES

There are many techniques to prepare the input data to the AI techniques. The importance of data preparation lies in its ability to improve significantly the network results. Most AI technique requires more information or pri-knowledge about the input data where the original data are only unidimensional data whereas the AI techniques require many dimensional data to improve the performance system. So, it should be changed the data from one-dimensional to multidimensional data. In this section; pre-processing techniques of input data will be discussed. These techniques are classified as:

i) *Data normalization,*
ii) *Frequency domain features*
iii) *Time domain features, and*
iv) *Auto Regressive (AR) model.*

Three of these pre-processing techniques were applied on two different examples. These examples are:
a) *Rolling element bearing detection techniques;*
b) *Diagnosis Induction Motor (IM) faults.*

The preprocessing Techniques will be discussed in the following Sub Sections.

A. Data Normalization

In the most cases, the data for training have extraordinary value, so it is always advisable to normalize network inputs. The normalized data have many advantages. Among these advantages that is all inputs of Network has converged values, hence no input is larger than another. It is preferred that the normalized input values changes in range from 0 to 1 or from -1 to 1 [6]. Equation (3) refers how to apply the normalized data in range from 0 to 1.

$$\text{norm\_data} = (X - \text{minVal}) / (\text{maxVal} - \text{minVal}) \quad (3)$$

Where:

X(t)	The original Inputs data
in Val	Smallest value of the original data
max Val	Max. value of the original data
norm_data	The data after normalization

B. Frequency Domain Features

Typically, time-frequency domain analysis is the most common method to analyze non-stationary signals. One of the standard methods to analyzing such signals include Short Time Fourier Transform (STFT) as illustrated in [26]. But this method has shortcomings, it presents constant resolution for all frequency and it uses the same window with constant width for the analysis the whole signal. This means that if a good frequency resolution using windows is desired, this would compromise the time resolution. Furthermore, there are no orthogonal bases existing for the STFT. Thus; it is difficult to obtain a fast and effective algorithm to calculate the STFT as given in [26]. More recent development in non-stationary signals analysis have been use of the wavelet transforms. Wavelet transforms can often be applied for analysis of signals through dilation and translation in order to extract the time-frequency of signals more effectively as presented in [26].

i) Wavelet Transforms

Wavelet Transform (WT) analyzes a signal concurrently in time and frequency domains. The WT is very useful in analyzing non-stationary signal. Wavelet transform can be applied to both continuous and discrete signals. In the following subsections, different forms of wavelet transforms and their mathematical formulations are briefly presented.

ii) Continuous Wavelet Transform

The wavelet transform of a continuous signal $\mathrm{x(t)}$ with respect to the wavelet function $\psi(\mathrm{t})$ can be defined as follows based on [26]:

$$\mathrm{T(a,b)} = \frac{1}{\sqrt{\mathrm{a}}} \int_{-\infty}^{\infty} \mathrm{x(t)}\, \psi^{*}\left(\frac{\mathrm{t-b}}{\mathrm{a}}\right) \mathrm{dt} \tag{5.1}$$

Where

$\psi^{*}(\mathrm{t})$ is the complex conjugates of the mother wavelet $\psi(\mathrm{t})$.

a is the distance between the center of the wavelet function and its crossing on the time axis.

b is used to govern the movement of the wavelet function along the time axis.

The normalized wavelet function is defined more compactly as:

$$\psi_{\mathrm{a,b}} = \frac{1}{\sqrt{\mathrm{a}}} \psi\left(\frac{\mathrm{t-b}}{\mathrm{a}}\right) \tag{5.2}$$

Now the transform integral equation can be rewritten as:

$$\mathrm{T(a,b)} = \int_{-\infty}^{\infty} \mathrm{x(t)}\, \psi^{*}_{\mathrm{a,b}}(\mathrm{t})\mathrm{dt} \tag{5.3}$$

The mother wavelet is dilated and translated continuously on a real number system in continuous wavelet transform as described in [26].

iii) Discrete Wavelet Transform

Continuous wavelet transform was introduced as a sampled version for implementation on the computer, but required a considerable amount of resources. It also suffered redundancy when attempting to reconstruct the signal. The Discrete Wavelet Transform (DWT) was presented to address the weaknesses of continuous wavelet transform. DWT was used in the 2nd example only; while it will be considered as Future work for the 1st example.

C. Time Domain Features

The captured vibration signal in time domain analysis is one of common preprocessing techniques. It is characterized by simplicity and efficiency. There are a lot of many significant statistical features were utilized to express the raw signal. The most important of these features in time domain (The famous applied statistical parameters as discussed in [22, 25]) are:

i) *Root Mean Square (RMS) Value (X_{RMS}),*
ii) *Peak to Peak Value (X_{PP}),*
iii) *Mean (X_{Mean}),*
iv) *Root Sum of Squares (X_{RSS}),*
v) *Crest Factor (X_{CF}),*
vi) *Impulse Factor (X_{IF}),*
vii) *Shape Factor (X_{SF}),*
viii) *Margin Factor (X_{MF}),*
ix) *Peak to Average Power ratio (X_{PAP}),*
x) *Energy (X_E),*
xi) *Variance (X_V),*
xii) *Skewness Value (X_{SV}) and*
xiii) *Kurtosis Value (X_{KV}) is given in [131].*

These features and the related Equations are presented in details in [22-25].

Six features were selected from the above 13 features acquired from vibration signals. These signals are obtained from the test utilized rig in the research as Machinery Fault Simulator (MFS) as explained in [6, 7] for te 1st example (*Rolling element bearing detection techniques)*. These six features are: *Root Mean Square Value; Peak to Peak; Kurtosis Value; Crest Factor; Impulse Factor; Energy* in Time Domain. While five features were selected from the above 13 features obtained from Iq and Id signals for the 2nd example (*Diagnosis Induction Motor (IM) faults)* as illustrated in [3-5; 8]. These features are: *Root Mean Square (RMS) Value (X_{RMS}); Peak to Peak Value (X_{PP}); Mean (X_{Mean}); Energy (X_E); and Variance (X_V).*
The above chosen features were used in the studied examples. The results will be presented and discussed in a later Section.

D. Auto Regressive Model for Input Data

Auto Regressive (AR) model is a commonly utilized model in applied studies which is a series of linear mathematical calculations. The information of output signal of DAQ device handling have a significant role in AR model where a P-order AR model is corresponding to a P-order linear predictor. According to the output of the AR model, the estimation factors are computed. These factors are calculated using Prediction Coding (LPC) as explained in [6, 24]. Early, the LPC was utilized in speech coding or (voice coding). The AR model's coefficients can be extracted through linear prediction analysis. Suppose acquired bearing data is X (t), its AR model is as exposed in [6, 24, 25]:

$$X(t)=\sum_{k=1}^{n}\varphi_k x(t-k)+\mathrm{e}(\mathrm{t}) \tag{3}$$

Where t is time; φk (k=1, 2… n) are AR model's factors. e (t) is model's residuals. The following formula refers to AR model's factors and its autocorrelation R (k):

$$\begin{bmatrix} R_1 \rightarrow & R_p \\ \downarrow \quad \square & \downarrow \\ R_p \leftarrow & R_1 \end{bmatrix}\begin{bmatrix} a_2 \\ \downarrow \\ a_{p+1} \end{bmatrix}=\begin{bmatrix} -R_1 \\ \downarrow \\ -R_{p+1} \end{bmatrix} \tag{4}$$

This is AR model's mathematical formula, which is also called Yule-Walker function [24]. AR model's parameters can be computed by using many algorithms. In this methodology, Forward-backward approach is predominantly employed to compute AR model's features. There are two effective methods to choose the best value for parameter of (AR) model.

a) Final Prediction Error (FPE) [6].

$$FPE(n)=\frac{N+\mathrm{n}}{N-\mathrm{n}}\sigma^2 \tag{5}$$

b) An Information Criterion (AIC)

$$AIC(n)=\mathrm{N}\ln\sigma^2+2\mathrm{n} \tag{6}$$

Where: -

N The number of samples.
n The number of AR model parameters.
σ Model's residual variance.

This pre-processing technique was tested on the two studied examples.

VII. ROLLING ELEMENT BEARING FUALTS DIAGNOSIS

This section includes the Time domain, Frequency domain and Artificial Intelligence System (AIS) used in rolling element bearing detection techniques. AIS involves Artificial Neural Network (ANN), Fuzzy Inference System and Adaptive Neuro-Fuzzy Inference System (ANFIS). The following sections describe these techniques briefly.

A. Time Domain and Frequency Domain Analysis

Time domain, frequency domain, and time-frequency vibration signal analyses are the best public utilized techniques. Time domain analysis is the analysis of mathematical functions; it is exposing the raw signal correlated with the time. The waveform exposes the raw signal whether continuous series time or discrete time along with the measured signal [6, 24]. However, the frequency domain analysis displays the frequencies which are occurred in the signal. It is possible that time domain analysis is converted to frequency domain analysis by Fast Fourier Transform (FFT) as specified in [24]. Table I presents the Bearing Parameters and Faults Frequencies.

Multilayer Feed-Forward Neural Networks are the most important forms of network architecture. It involves different layer as an input and output layers, the only two layers connected to the outside world surrounding the network.

TABLE I. BEARING PARAMETERS AND FAULTS FREQUENCIES

Number of balls (N_{ball})	9
Ball diameter (D_{ball})	0.3125
Pitch diameter (D_{cage})	1.516
Contact angle of bearing (β)	0°
F_{cage}	0.4 Fr
F_{inner}	5.43 Fr
F_{outer}	3.572 Fr
F_{ball}	4.633 Fr
Fr	35 Hz

TABLE II. BEARING DEFECT FREQUENCIES AND HARMONICS (HZ)

Type of fault	1X Hz	2X Hz	3X Hz
F_{cage}	14.04	28.08	42.12
F_{inner}	190.593	381.186	571.779
F_{outer}	125.377	250.7544	376.1316
F_{ball}	162.864	325.728	488.592

In addition to, the network involves one hidden layer or more at least. The hidden layer is used to improve the system performance. The number of hidden layer's neurons can be stated by the functions presented in Equation 17 [6, 24].

$$I<\sqrt{(m+n)}+a \tag{7}$$

Where:-

I *The number of hidden layer's nodes.*
n *The number of input nodes.*
m *The number of output nodes.*
a *a constant number between 1and 10.*

For the classification issues and pattern recognition, the performance of neural network may be improved according to the next steps as specified in [6]:

a) restart the training with a new values of network weights and biases
b) Increase the number of hidden layer's neurons.
c) The number of training vectors should be increased.
d) The knowledge and information associated with input data should be increased to improve the performance.
e) The training algorithm may be inappropriate. So, change this algorithm.

The flaws of the ANNs are the incapability to select or change a single synaptic weight as a separated portion of knowledge where all the interrelated handling the network nodes adjust at the same time with the stream of information and the adaptive rules. Consequently, any variation in the inputs may cause unexpected outcomes. The main weakness of ANNs for classification problems could be present poor capability to treat the raw data which take uni-diementional data while the ANN requires multi-dimensional data [6; 24]. On the other hand; Adaptive Neuro-Fuzzy Inference System (ANFIS) plays an important tool as classification technique. However; it should be identified the reasons that prompted to create the ANFIS technique as stated in [18]:

a) No obvious criterion approaches exist for converting human information or practice into rule base and knowledge base of FIS.

b) There is a requirement for active procedures for smooth setting the membership function (MFs) so as to reduce the output error degree or increase performance index to the best.

It is clear from the above that the ANN and the FIS aren't enough only to solve the complicated and uncertainly problems especially if more precision is needed. Thus; ANFIS is generated to be the hybridization of those methods for getting the better results. ANFIS is the approach of mixture between the ANN and the FIS to overcome the flaws in one of the technology during its application. ANFIS architecture is similar to Fuzzy Inference System structure which consists of five layers feed forward networks. Hence, it is the capability of education as neural network and employs fuzzy logic to emulate inference such as human brain [16, 12].

B. Results and Discussion on the bearing fualts detection

This section refers to the outcomes of implementation the research and shows the result in the waveform and the spectrogram in X and Y direction. Where, the data are implemented in FFT, ANN, and ANFIS. ANNs and ANFIS techniques utilize to categorize the healthy and the faulty bearings. Clearly from the results in [6]; that the analyses of vibration signals under the four operating status are quite complex. It is not easy to identify and detect of computed frequencies for bearing defect that is because the noise from nearby machines or the machine itself is overlapped with the main signal. Consequently it will be difficult to evaluate and detect the defects of bearing. Thus, it is motivation to apply the pre-processing methods as stated in [6; 24].

C. Applying Artificial Neural Network for research

For this practical work, the used neural network is a three layers feed-forward network with sigmoid hidden and output nodes is employed to categorize classes' vectors. The used learning algorithms are Error Back Propagation Neural Network (EBPNN).

The performance of neural network model widely depends on the input data which entering to pattern recognition for training. So, it is useful for data to be processed before it is entered into the model.

The used pre-processing methods for this work are normalization data, time domain features, and Auto Regressive (AR) model.

A. Bearing Fualts Detection and classification Based on Normalized Data

The input data for training are different large value, so it is preferred to normalize model inputs. For this work, it is used the normalized input values changes in range from 0 to 1 by using Equation (3). The input data to model represents the raw vibration signal in time domain in x-y direction. Thus, the number of inputs is two inputs whereas the outputs have four classes which represent the four of bearing conditions (Healthy- Outer Faulty- Inner Faulty- Ball faulty). It was taken 52500 samples for each bearing status of four bearing statuses (Healthy- Outer Faulty- Inner Faulty- Ball faulty). These samples represent 35 revolutions of rotating shaft. The samples are used to train the Back Propagation Neural Network (BPNN). The bearing vibration signals from four different statuses are classed to four labels. The target output values- bearing indices- were specified in Table III. The second stage is to determine the number of the hidden layer's neurons. the number of hidden layer's nodes varies within range 4 to13 is calculated related to Equation (7). The neurons' number of hidden layer has a significant effect on the performance of BPNN. Table IV describes the relationship between the numbers of hidden layer's neurons and the accuracy of the network.

TABLE III. TARGET OUTPUT VALUES

Bearing's condition	*Target output values*
Healthy	*1000*
Outer Faulty	*0100*
Inner Faulty	*0010*
Ball Faulty	*0001*

It is obvious from the above results, that the number of hidden layer neurons was chosen as 10 which presents the best accuracy of the model. On the other hand, the accuracy of classification model is low because of the formerly mentioned causes. Table IV presents the accuracy versus No. of hidden layers using normalized data as preprocessing.

TABLE IV. ACCURACY OF THE NETWORK VERSUS VARIED NUMBER OF HIDDEN LAYER NEURONS with Normalized Data

No. of hidden layer	accuracy of classification (%)
4	67.7
5	69.5
6	70.1
7	70.2
8	70.4
9	70.4
10	70.7
11	73.1
12	72.8
13	72.9

B. Bearing Fualts Detection and classification Model Based on Time Domain Features extraction

Table V presents the accuracy versus No. of hidden layers using features extraction (Time Domain) as preprocessing before applying ANN.

Although the overall classification results are rather improved but it didn't improve enough for ensuring that all the data samples of BPNN are properly classified. Consequently, it was motivated to use the Auto- Regressive (AR) model for preprocessing the inputs data before its entering to the ANN.

TABLE V. ACCURACY OF THE NETWORK VERSUS NUMBER OF HIDDEN LAYER NODES WITH TIME DOMAIN FEATURES

No. of hidden layer	accuracy of classification (%)
4	84.7
5	80.7
6	81.6
7	83.2
8	83.5
9	82.4
10	82.9
11	83.8
12	84
13	83.1

C. Bearing Fualts Detection and classification Model Based on Auto Regressive Parameters

The autoregressive model is effective technique for preprocessing the input signals to the neural network. The AR model depends on widely extraction the hidden features of signal with attenuating the noise as possible via auto correlation method [6, 24]. The number of the parameters of Auto Regressive (AR) model is 20 parameters which are considered the features of AR model for each 3000 samples of 510 k samples are used to compute their features. These features allow for using it to classify bearing conditions. However; the BPNN takes more iterations for training. Thus it takes a lot of time and computation. Using feature vectors as a replacement for the raw vibration signal will be successful method to minimize the calculation, which makes the training process faster. The structure of ANN is 20 nodes for the input layer and 4 nodes for the output layer whereas the nodes of hidden layer change in range from 7 to17 nodes according to Equation (7). The bearing indices are the identical to the last methods. Table VI introduce the relationship between the accuracy and the number of hidden layer's nodes using AR as preprocessing before applying ANN. It is observed that the inputs data sets in the first three confusion matrices and the overall confusion matrix were classified to be 100 % which mean all entered features (170 sets) to pattern recognition tool were classed in the right class. The preprocessing data via AR model introduces 100 % correct classification in the neural network model which means high accuracy. This model is perfect compared to the neural network models based on the normalized data and time domain features.

D. Comparison of the Results of the Preptocessing on ANN

From the previous, there are three varied techniques were used for preprocessing the raw vibration data before entering the Pattern Recognition Tool. The pre-processed data were employed as input to help the ANN for training. Thus, the different pre-processing data aim to build the network model in the best possible structure. It is clear that the results of normalized data model and time domain feature don't give the well results compared to the Auto Regressive model. The traditional methods such as only normalized data and time domain features don't train the neural network properly. If the information of acquired data is not enough due to lack of test equipment's, it would be difficult to get the perfect classification by traditional method. The comparison of the results is exposed in Table VII and Fig. 11.

TABLE VI. ACCURACY OF THE NETWORK VS. NUMBER OF HIDDEN LAYER NODES WITH AR

No. of hidden layer	accuracy of classification (%)	Performance error
7	99.7	0.000386
8	99.6	0.000104
9	100	0.001273
10	100	3.08E-06
11	99.9	0.001774
12	100	1.34E-06
13	100	0.000165
14	99.9	4.40E-05
15	100	1.06E-05
16	100	8.71E-05
17	100	0.00069

TABLE VII. COMPARISON BETWEEN DIFFERENT PREPROCESSING TECHNIQUES FOR ANNS

Technique	Classification	Class 1	Class 2	Class 3	Class 4	Overall
Normalized data	Correct	82.1	35.1	68.9	96.7	70.7
	Incorrect	17.9	64.9	31.1	3.3	29.3
Time Domain Data	Correct	84.1	88.2	77.1	89.4	84.7
	Incorrect	15.9	11.8	22.9	10.6	15.3
AR Model	Correct	100	100	100	100	100
	Incorrect	0	0	0	0	0

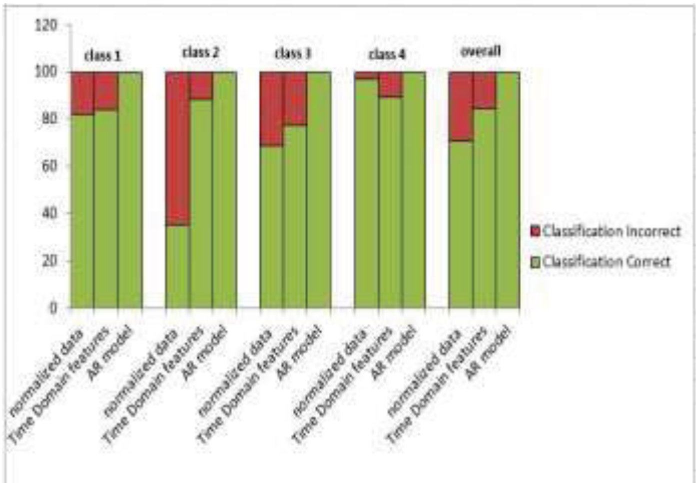


Fig.11. Comparisons between Preprocessing Data for ANN

The used algorithms for learning are hybrid algorithms of the least-squares estimator (LSE) technique and the back propagation gradient descent technique for training and checking. The hybrid learning algorithms give a high accuracy in classification problem where error decreases by using two phase (forward pass and backward pass).

E. Applying ANFIS for Bearing Fualts Detection and Classification

For this practical study, the structure of ANFIS model is a five-layer feed-forward network as explained before.

The performance of ANFIS model widely relies on the input data which entering to pattern recognition for training. In the next sections, it will be clarified about the effect of preprocessing data which incoming to ANFIS model to develop the model performance. For this study, the used model is Sugeno FIS model, it was loaded 3000 samples for each bearing condition of four bearing conditions (Healthy- Outer Faulty- Inner Faulty- Ball faulty). These samples represent 2 revolutions of rotating shaft. The samples are used to train and check the ANFIS model. The indices of four different statuses are classed to four tags as 1,2,3,4. The target output values were specified in Table VIII. These indices are used for next case (Normalized Data, Time Domain Features, and AR models).

TABLE VIII. INDEX FOR THE BEARING CONDITIONS

Bearing's condition	Target output values
Healthy	1
Outer Faulty	2
Inner Faulty	3
Ball Faulty	4

F. ANFIS Model Based on Normalized Data

For this work, it is used the normalized input values varies in range from 0 to 1.

The inputs data to model are the acquired vibration signal in time domain in x-y direction. So, the number of inputs is two inputs whereas the outputs have four classes which represent the four of bearing conditions (Healthy- Outer Faulty- Inner Faulty- Ball faulty). ANFIS model has only one output to display the right status of the different four statuses in the case of the results evaluation. The second step is the determination of the subtractive clustering method's factors (range of influence, quash factor, accept ratio, reject ratio) to get the best performance of ANFIS model, before starting the FIS training. The training algorithm to be chosen as Hybrid method, the error tolerance is zero, and the Number of epoch is 100.

G. ANFIS Model Based on Time Domain Features

The ANFIS model for classification relying on time domain features is more active techniques compared to the last approach. There are several time domain features such as. For Sugeno FIS model, it was taken only six features (*Root Mean Square Value; Peak to Peak; Kurtosis Value; Crest Factor; Impulse Factor; Energy*) for preprocessing the input data to ANFIS model. It was chosen 3000 samples as a period where it is calculated the time domain every 3000 samples of 510 K samples. The numbers of inputs are 6 for each of X and Y dimensions. The index bearing conditions are the same index in the last technique.

H. ANFIS Model Based on Auto Regressive Model

ANFIS Model Based on Auto Regressive Model was also studied. For Sugeno FIS model, it took hidden features for preprocessing the input data to ANFIS model.

I. Comparison between the ANNs Models and ANFIS Models Based on Preprocessing Data

Table IX illustrates the comparison between the results of ANN models and ANFIS models. It was chosen 3000 samples as a stage where it is determined the Auto Regressive features each 3000 samples of 510 K samples. The numbers of inputs are 20 for both of X and Y dimensions. The index bearing conditions are the same index.

TABLE IX.THE COMPARISON BETWEEN THE RESULTS OF ANN MODELS AND ANFIS MODELS

Type of preprocessing data	Overall accuracy of classification ANNs (%)	ANFIS Accuracy (%)	
		Training accuracy	Checking accuracy
Normalized data	70.7	50.819	51.055
Time domain features	84.7	75.01	73.823
Auto Regressive (AR)Model	100	99.9986	99.9984

VIII. DIAGNOSIS OF INDUCTION MOTOR FAULTS

This Section presents the tests and experiments that have been done on the system. It explains experimental implementation of Inter Turns Short Circuit (ITCS) faults Detection and classification. It discusses also the diagnosis of experimental implementation of Broken Rotor Bar (BRB) faults. Experimental implementation of combined ITSC faults and BRB faults is investigated. The IM under the test has the presented parameters in Table X.

Table X presents the Tested Squirrel Cage IM Specifications.

TABLE X. TESTED SQUIRREL CAGE IM SPECIFICATIONS

IM Specifications	**Unit**	**Value**
Power	HP	1.5
Voltage	Volt	380
Rated current	Amp	2.8
Rated speed	RPM	1400
Frequency	Hz	50
Number of rotor bars	--	28
Number of turns per phase	--	348

In this section; three pre-processing techniques of input data were applied for Diagnosis of Induction Motor faults. These techniques are classified as [3], [5], [8]:

i) *Frequency domain features (Discrete Wavelet Transform)*
ii) *Time domain features, and*
iii) *Auto Regressive (AR) model.*

Table XI presents The ANFIS and DWT Performance for Diagnosis 2% ITSC Fault.

Therefore the preprocessing for the data using DWT is encouraging and promising.

TABLE XI. THE ANFIS AND DWT PERFORMANCE FOR DIAGNOSIS 2% ITSC FAULT

RMSE	ANFIS with DWT	ANFIS without DWT
Training RMSE	5.3805×10^{-8}	9.6072×10^{-5}
Testing RMSE	5.4303×10^{-8}	0.7586

From Table XII; the training and testing RMSEs is developed in case of using an ANFIS. Hence it proves that the ANFIS and DWT technique is better than applying ANFIS without DWT. Additionally, it can diagnose the ITSC faults of an induction motor more accurately.

TABLE XII. THE ANFIS PERFORMANCE FOR DIAGNOSIS DIFFERENT ITSC FAULTS AT NO LOAD

ITSC Condition	Training RMSE	Testing RMSE
5%	7.9322×10^{-8}	7.917×10^{-8}
7%	1.0055×10^{-7}	1.0033×10^{-7}
10%	1.2531×10^{-7}	1.2493×10^{-7}

Furthermore; Table XIII illustrates The ANFIS Performance for Diagnosis Different ITSC Faults at Different Loading conditions. Moreover; the data pre-processing was tested on different faults levels at different loading conditions. While; Table XIV presents BRB Fault Index. Table XV depicts The ANFIS and DWT Performance for Diagnosis One BRB Fault. While, Table XVI displays The ANFIS and DWT Performance for Diagnosis of two BRB Fault under different loading conditions.

TABLE XIII. The ANFIS PERFORMANCE FOR DIAGNOSIS DIFFERENT ITSC FAULTS AT DIFFERENT LOAD

Loading Condition	Training RMSE	Testing RMSE
10%	3.5697×10^{-7}	3.5133×10^{-7}
20%	9.8329×10^{-6}	9.6491×10^{-6}
30%	1.4849×10^{-7}	1.4944×10^{-7}
40%	6.1396×10^{-5}	6.0249×10^{-5}
50%	1.7464×10^{-6}	1.7146×10^{-6}
60%	4.1448×10^{-6}	4.0674×10^{-6}
70%	2.4306×10^{-6}	2.386×10^{-6}
80%	2.0457×10^{-7}	2.0562×10^{-7}
90%	8.1223×10^{-7}	7.9737×10^{-7}
100%	2.0518×10^{-7}	2.0397×10^{-7}

Combined faults; ITSC and BRB; are implemented and preprocessing was applied before ANN as well as ANFIS. The obtained results are illustrated in the following Tables.

Table XVII presents the Combined BRB and ITSC Fault Indices.

TABLE XIV. BRB FAULT INDEX

Condition	Index of Fault
Healthy Motor	0
One BRB Fault	1
Two BRB Faults	2

TABLE XV. THE ANFIS AND DWT PERFORMANCE FOR DIAGNOSIS ONE BRB FAULT

RMSE	ANFIS with DWT	ANFIS without DWT
Training RMSE	7.748×10^{-9}	4.75×10^{-8}
Testing RMSE	7.7182×10^{-9}	0.0218

TABLE XVI. THE ANFIS PERFORMANCE WITH DWT FOR DIAGNOSIS TWO BRB FAULTS AT DIFFERENT LOAD

Loading Condition	Training RMSE	Testing RMSE
10%	1.9489×10^{-8}	1.9482×10^{-8}
20%	2.9369×10^{-8}	2.9557×10^{-8}
30%	2.5007×10^{-8}	2.5004×10^{-8}
40%	1.6245×10^{-8}	1.625×10^{-8}
50%	1.405×10^{-8}	1.4012×10^{-8}
60%	1.4039×10^{-8}	1.4×10^{-8}
70%	1.6559×10^{-8}	1.6505×10^{-8}
80%	1.4207×10^{-8}	1.4169×10^{-8}
90%	1.5918×10^{-8}	1.5866×10^{-8}
100%	1.3893×10^{-8}	1.3853×10^{-8}

TABLE XVII. COMBINED BRB AND ITSC FAULT INDICES

Condition	Index of Fault
Healthy Motor	0
One BRB Fault	1
Two BRB Faults	2
2% ITSC Fault	3
Combined Two BRB and 2% ITSC Faults	4
5% ITSC Fault	5
Combined Two BRB and 5% ITSC Faults	6
7% ITSC Fault	7
Combined Two BRB and 7% ITSC Faults	8
10% ITSC Fault	10
Combined Two BRB and 10% ITSC Faults	11

The experimental research work considers:

- Diagnosis of Different BRB Faults and 2% ITSC Fault in IM at No Load
- Diagnosis of Different BRB Faults and 2% ITSC Fault in IM without using DWT
- Diagnosis of Different BRB Faults and 5% , 7% and 10% ITSC Faults in IM at No Load

- Diagnosis of Different BRB Faults and Different ITSC Faults in IM at No Load
- Diagnosis of Two BRB Faults and Different ITSC Faults in IM at Different Load

Table XVIII illustrates The ANFIS and DWT Performance for Diagnosis Different of BRB Faults and 2% ITSC Fault

TABLE XVIII. THE ANFIS AND DWT PERFORMANCE FOR DIAGNOSIS DIFFERENT BRB FAULTS AND 2% ITSC

Fault RMSE	ANFIS with DWT	ANFIS without DWT
Training RMSE	5.6073×10^{-7}	0.5327
Testing RMSE	5.5547×10^{-7}	2.1271

Table XIX clarifies The ANFIS and DWT Performance for Diagnosis of two BRB Faults and one of ITSC Fault.

TABLE XIX. The ANFIS Performance for Diagnosis of Two BRB and ITSC Faults at Different Loading Conditions

Loading Condition	Training RMSE	Testing RMSE
10%	6.7152×10^{-5}	6.7421×10^{-5}
20%	5.3322×10^{-4}	5.3791×10^{-4}
30%	5.7229×10^{-5}	5.7915×10^{-5}
40%	5.1395×10^{-5}	5.1772×10^{-5}
50%	5.7158×10^{-4}	5.7992×10^{-4}
60%	$1{,}2401\times 10^{-3}$	$1{,}2283\times 10^{-3}$
70%	7.1056×10^{-5}	7.0383×10^{-5}
80%	3.9184×10^{-5}	3.8812×10^{-5}
90%	1.2478×10^{-5}	1.2409×10^{-5}
100%	$1{,}0214\times 10^{-4}$	10117×10^{-4}

Table XX presents the Evaluation of ANN and Machine Learning (ML) Classifiers. Here ANN was studied using preprocessing of data based on the features extraction. It is clear that limiting these features to five gives better results. However, the selection of these features is based on the obtained accuracy for each feature when it is used alone. The used ML classifiers are classified as:

i) Decision Tree (DT);
ii) Gradient Boosting (GB);
iii) K-Nearest Neighbours (KNN);
iv) Naïve Bayes (NB);
v) Random Forest (RF); and
vi) Support Vector Machine (SVM).

The experiment is implemented with ML algorithms: (DT, GB, KNN, NB, RF and SVM). The accuracy is calculated for each algorithm to evaluate the performance of the classifier. A comparative study is performed between ANN and machine learning algorithms using all feature extraction and five feature extraction as presented in Table XX.

It was observed that all classifiers provide high performance once using five feature extractions. The proposed method, DT, KNN and RF achieve better performance and obtain the highest scores of accuracy of 100%,

Table XXI illustrates the diagnosis results using the proposed approach compared with DT, KNN and RF classifiers. Proposed method achieved a high overall accuracy of 99.67%. While DT obtained a low overall accuracy of 85.17%.

TABLE XX. EVALUATION OF ANN AND ML CLASSIFIERS

Classifier	All Features	Five Features
	Accuracy (%)	Accuracy (%)
ANN	93.33	100
DT	95	100
GB	81.67	98.33
KNN	91.67	100
NB	30	31.67
RF	93.33	100
SVM	81.67	88.33

TABLE XXI. EVALUATION OF ANN, DT, KNN AND RF CLASSIFIERS

Load Condition	ANN	DT	KNN	RF
	Accuracy	Accuracy	Accuracy	Accuracy
10%	100	86.67	86.67	90
20%	100	88.33	86.67	96.67
30%	98.33	85	90	95
40%	98.33	86.67	86.67	96.67
50%	100	81.67	90	93.33
60%	100	85	86.67	98.33
70%	100	86.67	80	95
80%	100	88.33	91.67	95
90%	100	88.33	86.67	96.67
100%	100	75	88.33	93.33
Overall	99.67	85.17	87.34	95

Hence, ANN based statistical features extraction gives better performance as compared to all other approaches for ITSC fault diagnosis.

Auto Regressive Model as preprocessing for input data to the ANFIS gives bad results. This experiment is done for 2%, 5%, 7% and 10% ITSC faults. The obtained Training RMSE = 0.00048 while; testing RMSE = 0.43768. Therefore; AR as preprocessing is not recommended for ITSC and BRB faults in IM.

IX. CONCLUSIONS

As science and technology progresses, the need of fault diagnosis has also increased. The study covered most types of bearing faults relied on mechanical vibration signals as well as Inter Turn Short Circuit and Broken Rotor Bar faults in Induction Motor. Vibration frequency spectrum as well as input currents were useful in case of the acquired data without noise. The preprocessing of input data had a significant role in classification issues where it was revealed in the performance of the selected model for diagnosing the faults of bearing as well as Inter Turn Short Circuit and Broken Rotor Bar faults in Induction Motor. It is clear; that the training both ANNs and ANFIS models were improving when theses AI models were provided with more information about the data to be trained both in ANNs and ANFIS models. So; before applying the training in classification models, the appropriate preprocessing methods for the input data to AI models should be chosen. From the previous results, it is obvious that the best preprocessing of input data was the Auto-Regressive (AR) model which gave the optimum performance in both of ANNs and ANFIS models for bearing faults. The accuracy of ANNs model was very high under Auto-Regressive (AR) model as preprocessing of input data, which generated the classification of various bearing faults were more easily. The number of hidden layer's neurons helped to improve the accuracy for bearing faults analysis; but its effect was in a limited range. The ANFIS model automatically adjusted the rules. The accuracy of ANFIS model was very high considering Auto-Regressive (AR) model as preprocessing of input data for bearing faults. However, AR as data preprocessing in Induction motor faults analysis gives bad results. On the other hand data preprocessing using features extraction gives distinct results in Induction motor faults analysis. In the future studies, some works could be considered as follows:

Other types of faults such as gearbox faults may be considered. Extra accelerometers can be used to get more vibration details and database in the program for data storage. Implementation of Wavelet Transform (WT) for preprocessing of input data and comparing it to time domain features and Auto Regressive (AR) model. Implementation both experiments as online instead of offline should be considered.

ACKNOWLEDGEMENTS

The Authors would like to acknowledge the Sponsor and financial support of Heliopolis University (Cairo - Egypt) for this Research work and presenting it in the international Conferences.

REFERENCES

[1] Zhang, W., Jia, M. P., Zhu, L., & Yan, X.A."Comprehensive overview on computational intelligence techniques for machinery condition monitoring and fault diagnosis," Chinese Journal of Mechanical Engineering. Vol. 30, No. 4, pp. 782-795, 2017.

[2] Dernoncourt, Franck. "Introduction to fuzzy logic," Massachusetts Institute of Technology. Vol. 21, 2013.

[3] Menshawy A.Rady Mohamed; "Detection and Classification of Broken Rotor Bars Faults in Induction Motor Using Adaptive Neuro - Fuzzy Inference System"; M. Sc. Thesis, Graduate Studies, Faculty of Engineering, S. Valley University, Aug. 2015.

[4] Menshawy A. Mohamed and M. A. Moustafa Hassan; "Prediction of the Amount of Water Production and Estimate the Needs of Water Using ANFIS Techniques"; IWA Water and Development Congress & Exhibition, Dead Sea, Jordan 2015.

[5] Menshawy A. R. Mohamed, Mahmoud A. Sayed, E. H. A. Hameed, and M. A. Moustafa Hassan; "Detection and Classification of Broken Rotor Bars Faults in Induction Motor Using Adaptive Neuro-Fuzzy Inference System"; MEPCON 2014, Cairo, Egypt, Dec. 23-25, 2014.

[6] Metwally; Mostafa H. ; "Diagnosis OF Rotary Machines Faults Using Artificial Intelligence" ; M. Sc. Thesis, Graduate Studies, Faculty of Engineering, Cairo University, Feb. 2020.

[7] Mostafa. H. Metwally; M.A.Moustafa Hassan; Galal A. Hassaan; "Mechanical Machinery Faults Detection and Classification Based on Artificial Intelligence Techniques"; The 10th IEEE International Conference on Intelligent Data Acquisition and Advanced Computing Systems: Technology and Applications (IDAACS'2019); September 18- 21, 2019; Metz, France.

[8] Menshawy A.Mohamed, Al-Attar Ali Mohamed, Mohamed Abdel-Nasser, Essam E. M. Mohamed & M. A. Moustafa Hassan (2019): Induction motor broken rotor bar faults diagnosis using ANFIS-based DWT, International Journal of Modelling and Simulation, DOI:10.1080/02286203.2019.1708173; pp. 1-14.

[9] S. Omatu, M. Khalid and R. Yusof. "Neuro-Control and Its Applications, " Book, Springer-verlag London Limited, 1996.

[10] P. Picton. "Introduction to Neural Networks," Macmillan Press LTD, 1994.

[11] Rajasekaran S, Vijayalakshmi Pai GA. "Neural Networks, Fuzzy logic, and Genetic Algorithms Synthesis and Applications," Prentice Hall of India, New Delhi. 2007.

[12] Mehrotra K, Mohan CK, Ranka S."Elements of Artificial Neural Networks," Penram International Publishing Pvt. Ltd, India, 1997.

[13] Zadeh, L. A., "Outline of a new approach to the analysis of complex systems and decision processes", IEEE Transactions on systems, man, and cybernetics, Vol.3, No.1, pp. 28-44, Jan 1973. network-based fuzzy inference system, IEEE Transactions on Systems,". Man, and Cybernetics, Vol. 23, No. 3, (1993).

[14] G. Chen and T. T. Pham. "Introduction to Fuzzy Sets, Fuzzy Logic and Fuzzy Control Systems," CRC Press LLC, 2001.

[15] Jang, J.S.R."Adaptive-network-based fuzzy inference system, IEEE Transactions on Systems,". Man, and Cybernetics, Vol. 23, No. 3, (1993).

[16] Ali, M. M. Ismail, and M. A. Moustafa Hassan. "Speed sensorless field–oriented control of a six–phase saturated model of induction motors drive with online stator resistance estimation using ANFIS," International Journal of Modeling, Identification and Control. Vol. 17, No. 4, pp. 334-347, 2012.

[17] Aziz, M. S. Abdel, M. A Moustafa Hassan, and E. A. Zahab. "High-*impedance Faults Analysis in Distribution Networks Using an Adaptive Neuro Fuzzy Inference System," Electric Power Components and Systems. Vol. 40, No. 11, pp. 1300-1318, 2012.*

[18] J. Jang. "ANFIS: Adaptive-Network Based Fuzzy Inference System," IEEE Transactions on Systems, Man and Cybernetics. Vol. 23, No. 3, pp. 665-685, 1993.

[19] H.A.Taha Hussein, M.E.Ammar, M.A.Moustafa Hassan. "Three Phase Induction Motor's Stator Turns Fault Analysis Based on Artificial Intelligence," International Journal of System Dynamics Applications (IJSDA). Vol. 6, No. 3, pp. 1-19, 2017.

[20] Menshawy A Mohamed, Al-Attar Ali Mohamed, Mohamed Abdel-Nasser, Essam EM. Mohamed and M. A. Moustafa Hassan "Induction motor broken rotor bar faults diagnosis using ANFIS-based DWT," International Journal of Modelling and Simulation pp. 1-14, 2020,

[21] Zarei, J., Hassani, H., Wei, Z., & Karimi, H. R. "Broken rotor bars detection via Park's vector approach based on ANFIS," In 2014 IEEE 23rd International Symposium on Industrial Electronics (ISIE). pp. 2422-2426, 2014.

[22] Patel J., Patel V., Patel A.. "Fault diagnostics of rolling bearing based on improve time and frequency domain features using artificial neural networks," International Journal for Scientific Research & Development, Vol. 1, No. 4, pp. 781-788, 2013.

[23] Quiroz J. C., Mariun N., Mehrjou M. R., Izadi M., Misron N., Radzi M. A. M.; " Fault detection of broken rotor bar in LS-PMSM using random forests"; Measurement Vol. 116, pp. 273-280, 2018.

[24] Zhang, BO,"Condition monitoring of a fan using Neural Networks", M.Sc. Thesis, The School of Graduate Studies, Laurentian University, Sudbury, Ontario, Canada, 2009.

[25] Faghidi, H., “Non-parametric and Non-filtering Methods for Rolling Element Bearing Condition Monitoring”, D.SC., Ottawa-Carleton Institute for Mechanical and Aerospace Engineering, Laurentian University, Ottawa, Canada, 2014.
[26] B.S Prabhu. "Dynamic Analysis of Rotating Systems and Appplications," Brentwood: Multi Science Publishing. 2008.
[27] Ali, S.M., et al.“A comparative experimental study on the use of machine learning approaches for automated valve monitoring based on acoustic emission parameters,” IOP Conference Series: Materials Science and Engineering. Vol. 328, No. 1, pp. 012032, 2018.
[28] Kankar PK, Sharma SC, Harsha SP. "Fault diagnosis of ball bearings using machine learning methods," Exp Syst Appl. Vol. 38, No. 3, pp. 1876–1886, 2011.
[29] Martin-Diaz I., Morinigo-Sotelo D., Duque-Perez O., Romero-Troncoso R. J. "An experimental comparative evaluation of machine learning techniques for motor fault diagnosis under various operating conditions," IEEE Transactions on Industry Applications. Vol. 54, No. 3, pp. 2215-2224, 2018.
[30] Kotsiantis SB. "Supervised machine learning: A review of classification techniques," Informatica. Vol. 31, pp. 249-268, 2007