Diagnosis of Rotating Machines Faults Using Artificial Intelligence Based on Preprocessing for Input Data

Mostafa Metwally[1], Mohamed Moustafa Hassan[2], Galal Hassaan[3]
[1]Project engineer, EEHC, Cairo, Egypt (mostafaaoi@yahoo.com)
[2]Electrical Power Dept., Cairo University, Giza, Egypt (mmustafa@eng.cu.edu.eg)
[3]Mechanical Design and Production Dept., Cairo University, Giza, Egypt (galalhassaan@ymail.com)

***Abstract*—Rotating machines are extensively utilized in industrial life, since it represents a vital element in industrial processes. Therefore, early detection of faults of rotating machines is necessary to avoid the forced stopping for frequent maintenance in industrial processes. Various condition monitoring and detecting procedures are used to diagnose the rotating machinery faults based on vibration signature analysis, temperature monitoring, noise signature analysis, lubricant signature analysis, Artificial Intelligence (AI) techniques. many AI methods are in use for bearing defects diagnosis. For instance, Fuzzy Inference System (FIS); Artificial Neural Networks (ANNs) and Adaptive Neuro-Fuzzy Inference System (ANFIS). AI techniques are advanced technologically for classifying and detection various rotating machinery faults. This research work describes an inclusive method to some extent for detecting and classification of bearing faults using two methods of artificial intelligence which are ANN and ANFIS. The study is implemented offline in MATLAB environment. The proposed data were taken via using accelerometer which is mounted on the bearing housing in Machinery Fault Simulator (MFS) by M.Samy [1]. The obtained data of rolling element bearing were classified into four main conditions: Healthy, Outer Faulty, Inner Faulty, And Ball Faulty. The four conditions were imported to ANNs and ANFIS models. This work presents a comparison between the diagnosing based on Fast Fourier Transform (FFT), ANNs and ANFIS. The input data were preprocessed before entering to ANNs and ANFIS models by using three techniques as the following: the normalized data in range (0-1), the time domain features, and finally the Auto Regressive (AR) model. The accomplished outcomes of ANN and ANFIS models in case of AR model give high accuracy results in classification issue. The achieved outcomes are encouraging and promising in the field of diagnosis of mechanical 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 sector, machinery industry, metallurgical industry and petrochemical industry. The parts of rotating machines such as 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.

This significant topic found wide efforts from researchers. For ANNs, Gao et al (2001) introduced numerous AI techniques such as NNs, fuzzy logic, neural-fuzzy, and Genetic Algorithms (GAs) to diagnose the faults of motors [2]. Kalkat et al (2005) presented back-propagation neural networks to expect the vibration parameters of rotating systems amplitude, velocity and acceleration values in the vertical direction of the system [3]. Babu and Sekhar (2007) proposed ANNs and wavelet transform data to diagnose the cracks in shaft [4]. Zhang (2009) presented Back Propagation (BP) neural network to classify and diagnosis the different failure faults for fan [5]. Abdel-Hakam (2016) showed the two effective methods of AI techniques: the fuzzy inference system (FIS) and ANNS to diagnose the faults of rotating machinery and illuminate the features of these methods [6]. For ANFIS, Yang (2002) developed ANFIS technique for machinery fault diagnosis to detect multiple faults, at the one time through a Coactive ANFIS (CANFIS) where ANFIS is considered as a single output inference system, thus each ANFIS scheme can only effectively diagnose one kind of the defect for the machine [7]. Lei et al (2007) proposed that an ANFIS Gas are utilized together for fault detection of rotating equipment where it received the vibration signals after processed by using more than a few of pre-handling for data procedures, like EMD, filtration and demodulation to remove unnecessary knowledge and noise [8]. Yilmaz and Ayaz (2009) introduced effective method by using ANFIS for automatic bearing defect diagnosis of this type of failure where it was utilized three ANFIS models which select current, vibration and temperature as parameters and motor's state as output are associated [9]. Zhang et al (2014) applied ANFIS for diagnosing the faults of rolling element bearings and comparing with several other approaches [10]. Helmi and Forouzantabar (2018) investigated the features extraction of time domain and frequency domain as input of ANFIS model, ANFIS Network trained the test data sets to distinguish the rolling bearing faults of electric motor [11]. 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.

This article is organized as follow: Section 2 introduces components of the Experimental and presents to rolling element bearing faults frequencies and its calculation whereas Section 3 exhibits preprocessing for the input data of ai techniques Also, Section 4 discusses the rolling element bearing diagnosis methods. Furthermore Section 5 illustrates the results. Finally, Section 6 concludes the conclusions and future work.

II. EXPERIMENTAL SETUP FOR BEARING DEFECT FREQUENCIES CALCULATION

A. Components of the Experimental Setup

The test rig was utilized in the research is Machinery Fault Simulator (MFS). It consists of AC- motor (3 phase- 1HP), AC drive to changing the speed from 0-6000 RPM and the shaft is supported by the rolling element ball bearing under experiment. The MFS is exposed in Fig. 1, as stated in [1, 12].

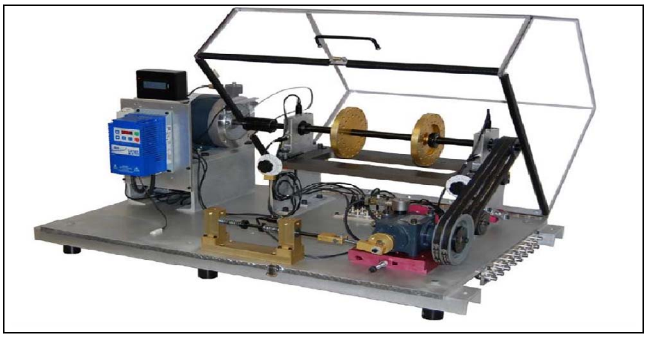

Fig.1. Machinery Fault Simulator.

The accelerometer is the sensing element in this experiment with sensitivity 100 mV/g. The accelerometer senses the vibration signal of bearing when the balls rotate at running speed that produce different frequencies in Spectrum in X and Y directions. The sensor is attached with Data Acquisition system (DAQ). The DAQ involves of four channels, each channel takes 51.2 kHz sampling rate. The function of DAQ system is capturing the signal from the sensor and processes it for presenting. It symbolizes the interface system between the Machine and the computer.

B. Bearing fault frequencies

There are several reasons of bearing drawbacks. Those drawbacks are existed in the outer race, cage, ball, and the inner race. When a rolling element rotates, these drawbacks produce a chains of vibrations each time a running rolling element spins over the surfaces of the flaws. These vibrations create frequencies called Bearing Fundamental Defect Frequencies (FDF) [1], [12]. Fig.2. illustrates the Sectional view of a rolling element bearing.

Fundamental defect frequencies rely on [1],[12]:

i) Shaft speed.

ii) Geometry of the bearing. Fig. 3 displays sectional view of a rolling element bearing with key bearing parameters [1], [12].

iii) The defect location.

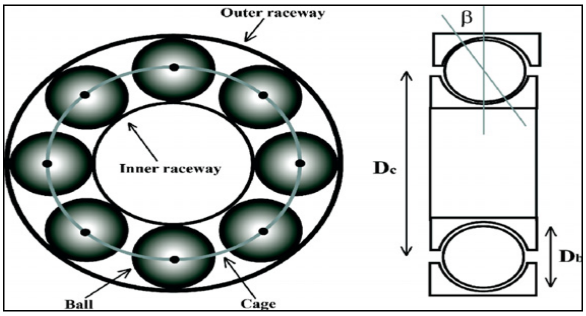

Fig.2. Sectional view of a rolling element bearing

The next Equations (Formulas 1 to 4) are utilized to diagnose the bearing faults frequencies [1], [12]:

FTF: Fundamental Train Frequency

$$F_{cage} = \frac{1}{2}\left(1 - \frac{D_{ball}}{D_{cage}}\cos\beta\right)F_r \quad (1)$$

BPFI: Ball Pass Frequency of the Inner race

$$F_{inner} = \frac{N_{ball}}{2}\left(1 + \frac{D_{ball}}{D_{cage}}\cos\beta\right)F_r \quad (2)$$

BPFO: Ball Pass Frequency of the Outer race

$$F_{outer} = \frac{N_{ball}}{2}\left(1 - \frac{D_{ball}}{D_{cage}}\cos\beta\right)F_r \quad (3)$$

BSF: Ball Spin Frequency

$$F_{ball} = \frac{D_{cage}}{D_{ball}}\left(1 - \frac{D_{ball}^{2}}{D_{cage}^{2}}\cos\beta^{2}\right)F_r \quad (4)$$

Where:

N_{ball}	The number of balls
F_r	Rotational shaft frequency.
D_{cage}	The cage diameter (pitch diameter)
D_{ball}	The ball diameter.
β	The contact angle.
F_{cage}	Cage frequency.
F_{inner}	Ball Pass Frequency of the Inner race.
F_{outer}	Ball Pass Frequency of the Outer race.
F_{ball}	Ball Spin Frequency.

C. Bearing Frequencies Calculation

The bearing flaw frequencies were computed before starting in the experiment. The kind of used bearing is M-BFK-1", with next details [1, 12] as given in Table I.

By applying the tested bearing data in the former Equation (1 to 4) in above subsections; one could obtain the following table (Table II). The steps of experiment are explained in details [1].

III. PREPROCESSING FOR THE INPUT DATA OF AI TECHNIQUES

There are many techniques to prepare the input data of the neural network. 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 requires many dimensional data to improve the performance system. So, it should be changed the data from one-dimensional to multidimensional data. In this section will be discussed the normalization data, time domain features, and Auto Regressive (AR) model.

TABLE I. BEARING DEFECT FREQUENCIES AND HARMONICS (HZ)

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
F_r	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

A. *Normalization Data*

In the most cases, the data for training have high 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 (5) refers how to apply the normalized data in range from 0 to 1.

$$norm\_data = (X - minVal)/(maxVal - minVal) \quad (5)$$

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. *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 vibration signal. The most important of these features are Root Mean Square (RMS) Value, Peak to Peak Value, Kurtosis, Crest factor, Impulse factor and Energy in time domain. It will review Root Mean Square (RMS), Peak to Peak value, Kurtosis factor, Crest factor, Impulse and Energy for following Equations ((6) – (12)).

The Peak Value is the easiest statistical technique to detect the faults in time domain which observed the peak value in the vibration signal.

RMS is different from the peak value in the way of detection of the defects in the vibration signal which based on the standard deviation of the vibration signal. RMS can be calculated from the following Equation (6) as shown in [5], [13]:

$$RMS = \sqrt{\frac{1}{N}\sum_{N=0}^{N-1}(x(n) - x^{-})^2} \quad (6)$$

Where x- is the mean value of the signal

$$x^{-} = \frac{1}{N}\sum_{n=0}^{N-1} x(n) \quad (7)$$

Peak to Peak value illustrates the variance between the biggest value and the smallest value in a time interval. The Equation is stated as:

$$X_{pp} = X_{max} - X_{min} \quad (8)$$

Kurtosis is a feature employed to get the peak of Probability density distribution curve's mean value. The formula is specified as [1,13].

$$k = \frac{\sum_{i=1}^{N}(X_i - X^{`})^4}{(N-1)S^4} \quad (9)$$

Where S represents the standard deviation of X.

Crest factor of employed to define a waveform. It displays the relation of peak value to the RMS the waveform. The formula is listed as [1,13].

$$C = \frac{peak\ value}{RMS} \quad (10)$$

Impulse factor shows the relation of peak value to the mean value. The formula is stated as [1,13]:

$$\mathrm{i} = \frac{peak\ value}{(\sum_{i=1}^{N} X_i)/N} \quad (11)$$

Energy is a time domain feature. It is utilized to acquire the power of the waveform in an assured period. The formula is set as [1], [13].

$$E = \left(\frac{\sum_{i=1}^{N}\sqrt{|x_i|}}{N}\right)^2 \quad (12)$$

C. *Auto Regressive Model for Input Data*

Auto Regressive (AR) model is a commonly utilized model in applied study 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, which is the Linear Prediction Coding (LPC) [5]. 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 [5], [13]:

$$X(t) = \sum_{k=1}^{n} \varphi_k x(t-k) + \mathrm{e}\,(\mathrm{t}) \tag{13}$$

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):

$$\begin{bmatrix} R_1 \rightarrow & R_p \\ \downarrow \;\square & \downarrow \\ R_p \leftarrow & R_1 \end{bmatrix} \begin{bmatrix} a_2 \\ \downarrow \\ a_{p+1} \end{bmatrix} = \begin{bmatrix} -R_1 \\ \downarrow \\ -R_{p+1} \end{bmatrix} \tag{14}$$

This is AR model's mathematical formula, which is also called Yule-Walker function [12]. 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.

Final Prediction Error (FPE) [12].

$$FPE\,(n) = \frac{N + \mathrm{n}}{N - \mathrm{n}}\,\sigma^2 \tag{15}$$

An Information Criterion (AIC)

$$AIC\,(n) = \mathrm{N}\ln\sigma^2 + 2\mathrm{n} \tag{16}$$

Where: -

- N The number of samples.
- n The number of AR model parameters.
- σ2 Model's residual variance.

IV. ROLLING ELEMENT BEARING DIAGNOSIS METHODS

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 [1, 12]. 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 as specified in [12]

B. *Artificial Intelligence Techniques*

Scientists and researchers had accomplished unconventional approaches such as Artificial Intelligence (AI). AI techniques are simulation of human behavior and the brain manner for making complex and critical decisions while taking a feature of the high speed of the computer to process the numerical information that cannot be processed by human. In the light of scientific and technological development, researchers introduced the other techniques to simulate the human brain such as the Artificial Neural Networks (ANNs), Fuzzy Inference System (FIS) and ANFIS. It is the best frequently empolyed for AI systems [12].

i) Artificial Neural Networks

ANNs is created to emulate the human neural networks whence the processing techniques of the data. The researchers assume that the secret behind the high response of human neurons due to the data processing method which may be realized in parallel, which improve its high speed. ANN architecture is illustrated in Fig. 3 (as presented in [12]). ANN involves of three main nodes layers: input, hidden, and output layers [12].

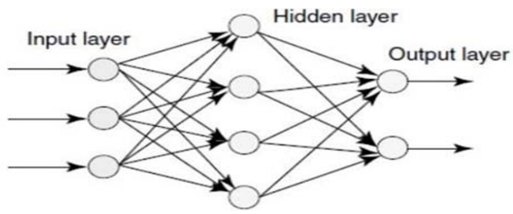

Fig.3 The Artificial Neural Architecture

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. 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 [5],[12].

$$I < \sqrt{(m + n)} + a \tag{17}$$

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 [14]:

1) restart the training with a new values of network weights and biases

2) Increase the number of hidden layer's neurons.

3) The number of training vectors should be increased.

4) The knowledge and information associated with input data should be increased to improve the performance.

5) 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 be cause unanticipated results. The main weakness of ANNs for classification problems be present poor capability to treat the raw data which take uni-diementional data while the ANN requires multi diementional data [12].

ii) Fuzzy Inference System

In 1965, Lotfi Zadeh has created one of the most common and a significant AI method is fuzzy logic. The traditional techniques depend on classical logic where classic sets are expressed either 0 or 1. In contrast, the fuzzy logic built on Fuzzy sets is represented according to the degree of membership function between 0 and 1. Zadeh replaced the numerical variable with linguistic variable. Zadeh expressed the simple and complex relations in fuzzy algorithms [12], [15]. The FIS basically determines a nonlinear assigning of the input data vector into a numerical output. The entered input data to the FIS model are mapped by using fuzzy rules. The mapping procedure comprises three main processes: fuzzification, rule base (if - then), and defuzzification. To apply fuzzy logic approach in the practical work, the following three procedures are needed as presented in [6].

iii) Adaptive Neuro-Fuzzy Inference System (ANFIS)

Before exhibiting the architecture of the ANFIS, it should be identified the reasons that prompted Yang to create the ANFIS technique as stated in [16]:

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 need, so 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].

V. Results and Discussion

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. The preprocessing for input data is employed to improve the performance of ANNs and ANFIS models.

A. The vibration signal in frequency domain

The acquired vibration signal from data acquisition (DAQ) in time domain is plotted. The following figures are the vibration signal in frequency domain after converting the vibration signal from time domain to frequency domain.

Figs. 4 and 5 show the Spectrum in X and Y direction under the healthy condition. While Figs. 6 and 7 clarify the spectrum in X and Y direction under the Outer Fault Bearing. Also, Figs. 8 and 9 illustrate the spectrum in X and Y direction under the Inner Fault Bearing. Furthermore, Figures 10 and 11 reveal the spectrum in X and Y direction under the Ball Fault Bearing [12].

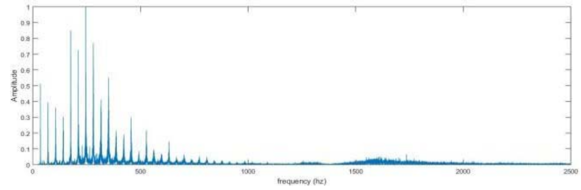

Fig. 4 The Spectrum in X-Direction (Healthy Bearing).

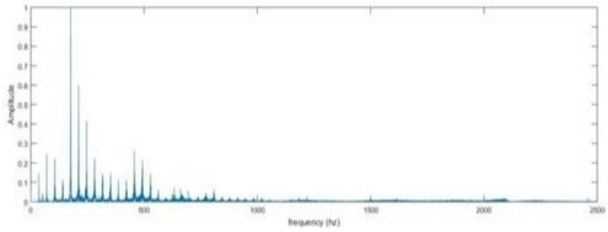

Fig. 5 The Spectrum in Y-Direction (Healthy Bearing).

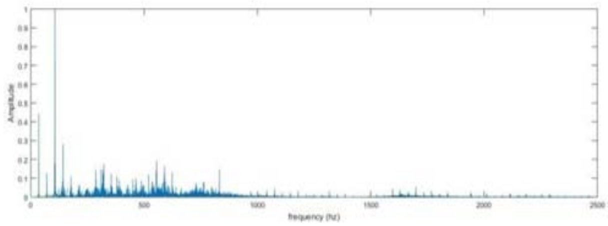

Fig. 6 The Spectrum in X-D4irection (Outer Fault Bearing).

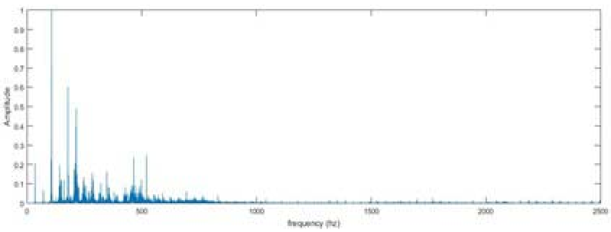

Fig. 7 The Spectrum in Y-Direction (Outer Fault Bearing).

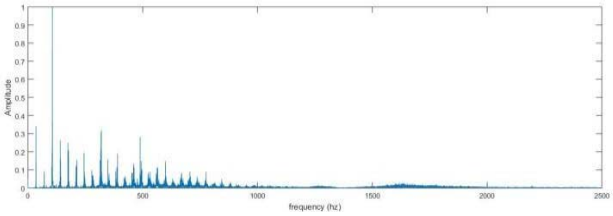

Fig. 8 The Spectrum in X-direction (Inner Fault Bearing).

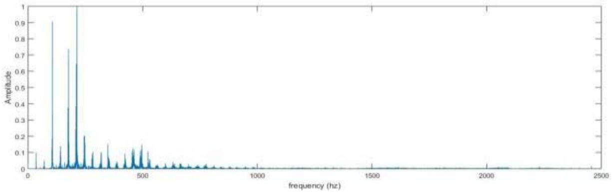

Fig. 9 The Spectrum in Y-Direction (Inner Fault Bearing).

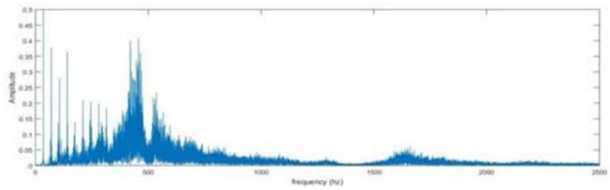

Fig. 10 The Spectrum in X-Direction (Ball Fault Bearing).

Clearly from the previous figures, that the analyses of vibration signals under the four operating status are quite complex. It isnot 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 next methods [12].

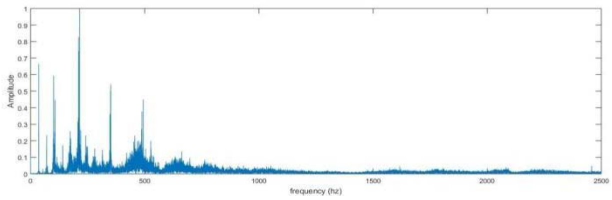

Fig. 11 The Spectrum in Y-Direction (Ball Fault Bearing).

B. *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 preprocessing methods for this work are normalization data, time domain features, and Auto Regressive (AR) model.

A. *Pattern Recognition Model 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 (5). 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 indice- 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 (17). 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*

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

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

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. Figure 12 offers the confusion matrices for the training results of four bearing conditions at hidden layer neurons are 10. On the other hand, the accuracy of classification model is low because of the formerly mentioned causes. The primary three matrices represent the outcomes of the training set, validation set and testing set and the final matrix illustrates the outcomes of the collective data set. The training set is about 70 % (147000 of 210000) of the total samples while both of the validation set and test set is about 15 % (31500 of 210000) for each set.

B. Pattern Recognition Model Based on Time Domain Features

The pattern recognition model for classification depending on time domain features is more effective techniques compared to the last technique. There are many time domain features such as RMS, Peak to Peak; Mean, Standard deviation, Peak value, Crest factor, Skewness, Kurtosis, Impulse, and Energy [1,13]. It was taken only six features for preprocessing the input data to Back Propagation Neural Network (BPNN). For this research, the sampling rate is 51. kHz where the number of samples is 510000 samples in every segment of time series data where the experiment was continued for ten seconds and rotating speed is 2100 rpm (35 Hz).

The number of samples is calculated according to the following Equation:

$$No.\ \text{of samples} = \text{time} * \text{sampling rate} \tag{18}$$

Fig. 12 The confusion matrix for 10 hidden layer nodes

Depending on the sample rate and the rotating speed of the bearing system, 3000 samples were chosen as a period to calculate their features according to the Equations from 6 to 12. The reason of the selection for 3000 samples as a segment of time series data which represent two complete revolutions of the rotating bearing at rotation speed frequency 35 Hz. The structure of ANN is 6 neurons for the input layer and 4 neurons for the output layer whereas the hidden layer varies from 4 to 13 according to Equation (17). The bearing indice are the same from the last method. Table V refer to Accuracy of the Network versus Number of Hidden Layer Nodes whereas Figure 13 exhibits the confusion matrices of the BPNN.

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 network. 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 attenuate the noise as possible via auto correlation method [5]. 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 (17). The bearing indices are the identical to the last methods. Table VI introduce the relationship between the accuracy and the number of layer's nodes.

TABLE V. ACCURACY OF THE NETWORK VERSUS NUMBER OF HIDDEN LAYER NODES

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

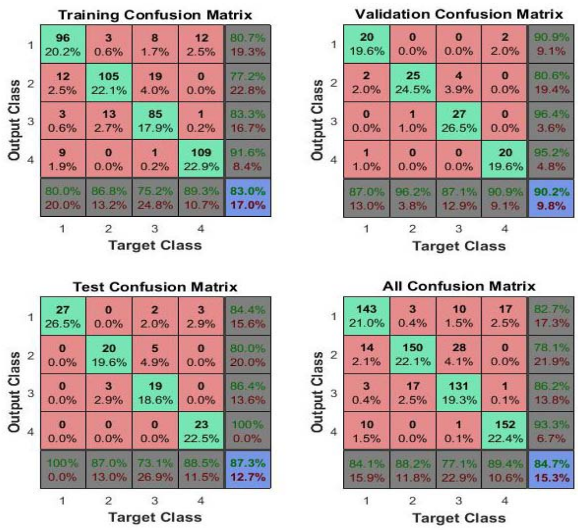

Fig. 13 The Confusion Matrices for Time Domain Features Data Model

C. Pattern Recognition Model Based on (AR) Model

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.

Fig. 14 presents the results of confusion matrix for bearing vibration signal by using 12 hidden layer's nodes.

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

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

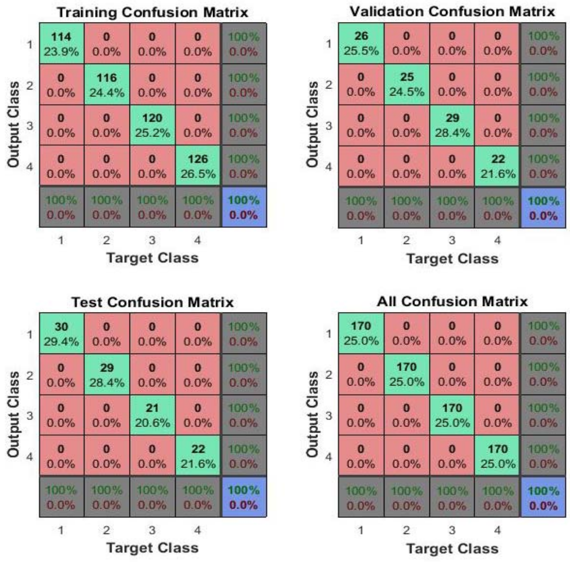

Fig. 14 The Confusion Matrices for Auto Regressive (AR) Model

D. *Comparison of the Results of the Previous Neural Network Models*

From the previous, there are three varied techniques were used for preprocessing the raw vibration data before entering the Pattern Recognition Tool. The processed data was employed as input to help the neural network for training. Thus, the different preprocessing 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. 15.

E. *Applying ANFIS for the Study*

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

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

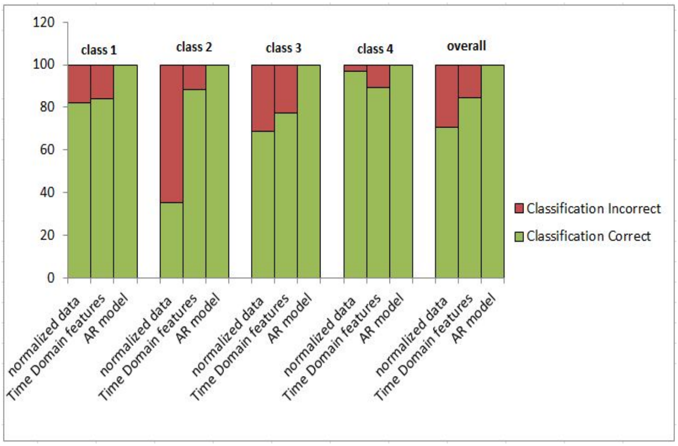

Fig. 15 Comparison between Preprocessing Data

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 method to be chosen as Hybrid method, the error tolerance is zero, and the Number of epoch is 100, then press train now. Table IX clarifies the relationship between the Root Mean Square Error (RMSE) of training and checking and different the subtractive clustering method's factors.

TABLE IX. THE RELATION BETWEEN ACCURACY AND SUBTRACTIVE CLUSTERING METHOD'S FACTORS

Trial No.	Range of influence	Squash factor	Accept ratio	Reject ratio	Number of Rules	Training RMSE	Checking RMSE
1	0.5	1.5	0.5	0.15	3	0.54279	0.53148
2	0.2	2	0.3	0.15	4	0.50834	0.48578
3	0.2	1.5	0.4	0.15	4	0.50829	0.49508
4	0.4	0.8	0.4	0.15	6	0.50693	0.50601
5	0.2	0.9	0.2	0.15	5	0.50531	0.48808
6	0.25	0.8	0.25	0.15	7	0.50149	0.48988
7	0.1	0.9	0.1	0.15	11	0.49532	0.49156
8	0.2	0.6	0.2	0.15	16	0.49244	0.48831
9	0.25	0.5	0.25	0.15	20	0.49181	0.48945
10	0.5	1.25	0.5	0.15	4	0.59846	0.61112

It is obvious in the above results that the best performance of the ANFIS model is trail No. 9 which gives 0.49181 and 0.48945 as Training RMSE and Check RMSE, respectively at range of the influence is 0.25, quash factor is 0.5, accept ratio is 0.25, and reject ratio is 0.15.

Training RMSE is Root Mean Square Error of training which represents the dissimilarity between the value of training data output, and the output of the FIS model parallel to the same training data input value.

Checking RMSE is Root Mean Square Error of Checking which represents the dissimilarity between the value of checking data output, and the output of the FIS model parallel to the same checking data input value.

The number of rules is 20 rules for each input. Each rule represents cluster. So, it has 20 clusters for each input. Fig. 16 denotes checking data against the trained FIS model under normalized data.

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 RMS, Peak to Peak; Mean, Standard deviation, Peak value, Crest factor, Skewness, Kurtosis, Impulse, and Energy. For Sugeno FIS model, it was taken only six features 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. These features are calculated according to the Equations from 6 to 12. The cause of the choice for 3000 samples as a portion of time series data to denote two complete revolutions of the spinning bearing at rotating speed frequency 35 Hz. Table X explains the relationship between the RMSE of training and checking under several the subtractive clustering method's factors.

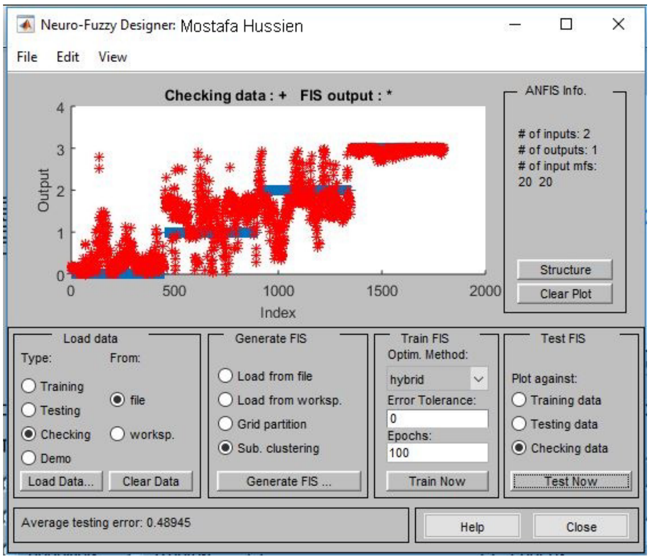

Fig. 16 Checking Data against the Trained FIS Model for normalized data

TABLE X. THE RELATION BETWEEN ACCURACY AND SUBTRACTIVE CLUSTERING METHOD'S FACTORS

Trial number	Range of influence	Squash factor	Accept ratio	Reject ratio	Number of Rule	Training RMSE	Checking RMSE
1	0.5	1.5	0.25	0.15	4	0.48529	0.50707
2	0.3	1	0.3	0.15	8	0.46693	0.49984
3	0.25	1.25	0.25	0.15	8	0.462248	0.490056
4	0.3	0.9	0.2	0.15	9	0.43092	0.4608
5	0.2	0.8	0.5	0.15	12	0.38527	0.40858
6	0.2	0.5	0.5	0.15	14	0.37508	0.39647
7	0.2	0.7	0.3	0.15	13	0.3685	0.38556
8	0.25	0.6	0.25	0.15	20	0.33002	0.3445
9	0.2	0.6	0.3	0.15	16	0.32139	0.33173
10	0.25	0.5	0.25	0.15	33	0.2499	0.26177

It is obvious in the previous results that the best performance of the ANFIS model is trail No. 10. It is noted that the numbers of rules are 33 rules for each input where the numbers of inputs are greater than before. Consequently, it is required to have more rules. Each rule denotes to cluster. So, it has 33 clusters for each input. The accuracy of model is improved. Fig. 17 illuminates checking data versus the trained FIS model for time domain features model.

H. *ANFIS Model Based on Auto Regressive Model*

For Sugeno FIS model, it took hidden features for preprocessing the input data to ANFIS model. 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 XI depicts the relationship between the RMSE for training and checking at various the subtractive clustering method's parameters.

VI. THE RELATION BETWEEN RMSE AND SUBTRACTIVE CLUSTERING METHOD'S PARAMETERS

The Relation between RMSE and Subtractive Clustering Method's parameter is presented in Table XI.

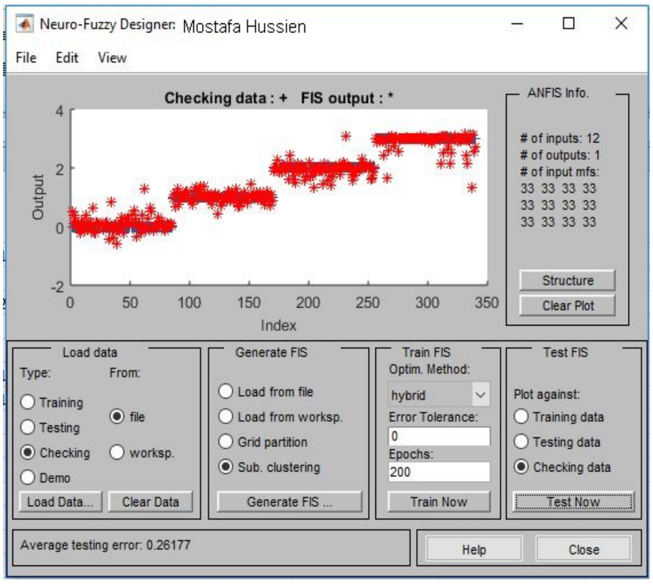

Fig.17 Checking Data vs. the Trained FIS Model

TABLE XI. THE RELATION BETWEEN RMSE AND SUBTRACTIVE CLUSTERING METHOD'S PARAMETERS

Trial No.	Range of influence	Squash factor	Accept ratio	Reject ratio	Number of Rule	Training RMSE	Checking RMSE
1	0.5	1.25	0.5	0.15	13	0.018821	0.02
2	0.5	[illegible]	0.5	0.15	7	0.047154	0.051293
3	0.45	1.25	0.5	0.15	19	0.000578	0.00067858
4	0.5	2	0.5	0.15	8	0.0438	0.046467
5	0.4	1.25	0.9	0.15	33	1.40E-05	1.56E-05

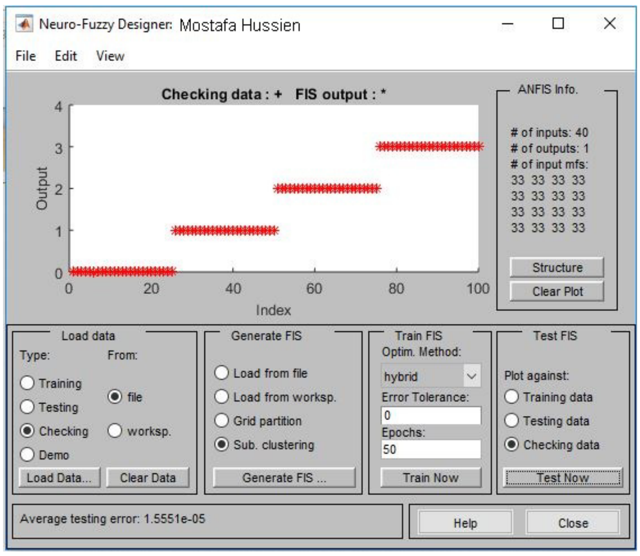

Fig.18 Checking Data versus the Trained FIS Model for the AR model

It is observed in the former method that the optimum performance of the ANFIS model is realized in trail No. 5 which gives 1.4E-5 and 1.56E-5 as training RMSE and checking RMSE, respectively at range of influence is 0.4, quash factor is 1.25, accept ratio is 0.9, and reject ratio is 0.15. It is noted that the number of rules became 33 rules for each input. The reason of the number of rules is the number of inputs which is increased. Consequently, it is necessary to have more rules. Each rule is a cluster. So, it has 33 clusters for every input. The accuracy of model is optimum. Fig. 18 describes the checking data versus the trained FIS model for the AR model.

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

Table XII describes the comparison between the results of ANN models and ANFIS models.

VII. CONCLUSIONS

As science and technology progresses, the need of fault diagnosis has also increased. The subsequent is the core conclusions that can be extensive vision from the present work: The study covered most types of bearing faults relied on mechanical vibration signals. Vibration frequency spectrum was useful in case of the acquired data without noise. The preprocessing of input data had a significant role in classification issues where it was reflected in the performance of the selected model for diagnosing the faults of bearing. It is clear; that the training both ANNs and ANFIS models were improving when the 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. The advantages of ANNs model are the following: firstly, the training time was very low. It took a few minutes compared to ANFIS model. Secondly; the accuracy of ANNs model was very high under Auto-Regressive (AR) model as preprocessing of input data, which generated the classification of various faults were more easily. The disadvantages of ANNs model are that when the model was trained, the model created different results in each time. The number of hidden layer's neurons helped to improve the accuracy but its effect was in a limited range. The advantages of ANFIS model are the following: the ANFIS model overcame on the drawbacks of the ANNs model that when the model was trained, the model generated the same results in each time. Unlike the FIS model, 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. The disadvantages of ANFIS model are the following: firstly, unlike the ANNs model the training time needed a big time especially when the number of rules were large. Secondly, the bad choosing of subtractive clustering's factors leaded to bad performance for ANFIS model. In the future studies, some works could be considered as follows:

- Other types of faults such as gearbox faults and motor faults may be considered.

- Extra accelerometers can be used to get more vibration details and database in the program for data storage.
- Execution Wavelet Transform (WT) for preprocessing of input data instead of time domain features and Auto Regressive (AR) model.
- Implementation the experiment as online instead of offline.
- Realizing other techniques to detect the faults of bearing such as Support Vector Machine (SVM) and decision tree.

TABLE XII. 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	0.50819	0.51055
Time domain features	84.7	0.7501	0.73823
Auto Regressive (AR)Model	100	0.999986	0.999984

Acknowledgement:

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.

VIII. REFERENCES

[1] Samy, M. ; “Real-Time Fault Detection and diagnosis of Machinery”, M.SC., Faculty of Engineering, Helwan University, 2016.

[2] Gao, X.Z., and Ovaska, S.J., “Soft computing methods in motor fault diagnosis", Applied Soft Computing, Vol. 1, pp. 73–81, 2001.

[3] Kalkat, M., Yildirim, S., and Uzmay, I., "Design of artificial neural networks for rotor dynamics analysis of rotating machine systems", Journal of Mechatronics, Vol. 15, pp. 573–588, 2005.

[4] Babu, T. R., and Sekhar, A. S., "Shaft crack identification using artificial neural networks and wavelet transform data of a transient rotor". 13th National Conference on Mechanisms and Machines, Bangalore, India, December, pp. 327-332, 2007.

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

[6] Abdel-Hakam, K., “Machinery Fault Diagnosis Using Advanced Techniques,”, M.Sc. Thesis, Faculty of Engineering. Minia University, Minia, Egypt, 2016.

[7] Yang, H., Mathew, J., and Ma, L., “Intelligent Diagnosis of Rotating Machinery Faults-A Review ", 3rd Asia-Pacific Conference on Systems Integrity and Maintenance, 2002.

[8] Lei, Y., He, Z., Zi, Y., Hu, Q., “Fault diagnosis of rotating machinery based on multiple ANFIS combination with GAs.”, Mechanical Systems and Signal Processing Vol. 21, No. 5, pp. 2280–2294, 2007.

[9] Yilmaz, M.S., Ayaz, E., “Adaptive Neuron-Fuzzy Inference System for Bearing Fault Detection in induction motors using temperature, current, vibration data.” , 3rd International Conference on electrical engineering and communications., EUROCON, pp. 1140-1145, Russia, 2009.

[10] Zhang, J., Ma, W., and Ma, L., “A Fault Diagnosis Method Based on ANFIS and Bearing Fault Diagnosis ", International Conference on Information Science, Electronics and Electrical Engineering, pp. 1274-1278, Sapporo, 2014.

[11] Helmi, H., and Forouzantabar, A., “Rolling bearing fault detection of electric motor using time domain and frequency domain features extraction and ANFIS", International Conference on Information Science, Electronics and Electrical Engineering, Vol.13, No. 5, pp. 662 – 669, 2018.

[12] Metwally, M. H., Hassaan, G. A., and Moustafa Hassan, M. A., “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., Metz; France; 2019.

[13] 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.

[14] Beale, M.H., Hagan, M.T., Demuth, H.B., Neural Network Toolbox™ Getting Started Guide, The MathWorks, Inc., 2013.

[15] 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).

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