

Detection and Classification of Wear in Rotary Systems based on ANN and GA

M.Barghi Mehmandari¹, A.Sadeghi Bakhshayesh², C.Teymuri Sindi³

Student, Department of Mechanical Engineering, Islamic Azad University, Tabriz Branch, Tabriz, Iran²

Assistant Professor, Department of Mechanical Engineering, Islamic Azad University, Tabriz Branch, Tabriz, Iran³

ABSTRACT: The subject matter of this paper refers to the improvement of the acoustic emission (AE) method when used for detection and classification of wear in rotor and stator. Firstly, time-domain signal analysis methods are used for feature extraction of the recorded AE signals. Then, for classifying AE signals based on extracted features, a Feedforward Neural Network (FFNN) is used. Two different training algorithms (Bayesian Regularization and Levenberg Marquardt) are applied to train the Neural Network. This method is used for classifying 450 signals (300 with wear and 150 without wear) into two classes. The results stated the effectiveness of the Bayesian Regularization training algorithm as oppose to Levenberg Marquardt. GA, in combination with Levenberg Marquardt, improved the performance of this algorithm, too.

KEYWORDS: Acoustic Emission (AE), Artificial Neural Network (ANN), Genetic Algorithm (GA), Classification, Rotary systems, wear detection

I. INTRODUCTION

In most engineering areas such as manufacturing, aerospace, and transportation, rotary mechanical equipment is commonly used. Malfunction in rotors and stators is the most important reason for unexpected mechanical failures include mounting failure, low lubrication, fatigue, and wear and tear. These may result in formidable challenges in the accurate diagnosis of faults. Therefore, rotary systems should be maintained at highly reliable conditions to avoid any unexpected defects. Real-time monitoring is a solution for tackling this problem which leads to early fault detection and prevention of serious failures in components. Comparing to the other non-destructive testing techniques, AE has the advantage of real-time, continuous monitoring of in-service structures [1].

The main problem with the AE technique is noise contamination from the other machine components rather than to the relatively low AE signal magnitudes. Therefore, signal processing is the next important issue after signal registration. Signal processing methods are divided into two categories: time-domain and frequency domain analysis. In the time domain methods, statistical parameters of AE signals such as mean value, standard deviation, Root Mean Square (RMS), and peak value have been used. The frequency domain is another explanation of a signal and these features are based on the Fourier transform of the AE signals such as power spectral density and frequency spectrum [2]. In this paper, the AE method is used in order to register elastic waves for monitoring wear conditions of a rotor and stator in a rotary machine. Time-domain analysis techniques are used to extract features from the AE signals. Then, a neural network is employed to classify the input signals.

The reminder of this paper is organized as follows: Section 2 mentions some related works in this field. Feature Extraction method is introduced in section 3. The principle of ANNs and GAs are mentioned in section 4. Section 5 describes the experimental procedure. In section 6, results and evaluation of proposed methods are given. Finally, section 7 provides a conclusion.

International Journal of Innovative Research in Science, Engineering and Technology

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Visit: www.ijirset.com

Vol. 8, Issue 10, October 2019

II. RELATED WORK

In last few decades, AE techniques have increasingly used in various fields. For instance, Lei Lu, Jihong Yan, and Clarence W. de Silva [3] proposed a technique based on AE, GA, and SVM for fault diagnosis in rotary mechanisms. Mba and Rao [2] have demonstrated an enhanced detection ability of the AE technique over the conventional vibration analysis in rotating machines, especially at low rotational speeds. Mohammad Reza Asadi Asad Abad and others used acoustic condition monitoring and artificial neural network technique for Fault diagnosis of a differential [5].

III. FEATURE EXTRACTION

Feature extraction is a process of dimensionality reduction by which an initial set of raw data is reduced to more manageable groups for processing. A characteristic of these large data sets is a large number of variables that require a lot of computing resources to process. Feature extraction is the name for methods that select and combine variables into features, effectively reducing the amount of data that must be processed, while still accurately and completely describing the original data set. The process of feature extraction is useful when reducing the number of resources used for processing without losing important or relevant information is needed. It can also reduce the amount of redundant data for a given analysis. Furthermore, the reduction of the data and the machine's efforts in building variable combinations (features) facilitate the speed of learning and generalization steps in the machine learning process [6].

IV. PRINCIPLE OF ANN AND GA

ANN

ANN is a branch of artificial intelligence (AI). An artificial neuron network (ANN) is a computational model based on the structure and functions of biological neural networks. Information that flows through the network affects the structure of the ANN because a neural network changes - or learns, in a sense - based on that input and output. ANN takes data samples rather than entire data sets to arrive at solutions, which saves not only time but also money. ANNs are considered fairly simple mathematical models to enhance existing data analysis technologies. ANNs have three layers that are interconnected. The first layer consists of input neurons. Those neurons send data on to the second layer, which in turn sends the output neurons to the third layer.

Generally, an ANN can be defined as a system or mathematical model that consists of many nonlinear artificial neurons running in parallel and may be generated as one-layered or multilayered. Most ANNs have three layers: input, output, and hidden. The function of the hidden layer is to intervene between the external input and the network output in some useful manner. Detailed theoretical information about ANNs can be found in [7]. An ANN structure is shown in fig 1.

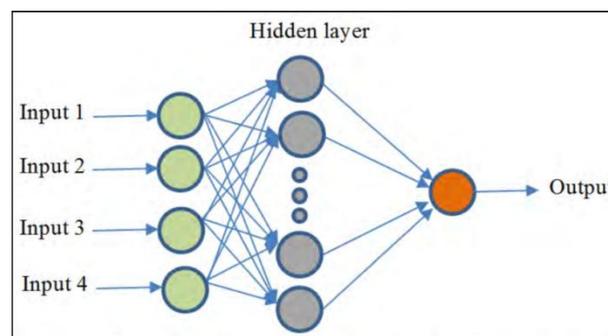


Fig1, ANN structure

Feed forward multilayer neural network (FFNN) is dominantly used. Two kinds of signals are identified in this network: The function signals (also called input signals) that come in at the input of the network, propagate forward (neuron by

International Journal of Innovative Research in Science, Engineering and Technology

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Visit: www.ijirset.com

Vol. 8, Issue 10, October 2019

neuron) through the network and reach the output end of the network as output signals; The error signals that originate at the output neuron of the network and propagate backward (layer by layer) through the network. The output of the neural network is described by equation 1:

$$y = F_o\left(\sum_{j=0}^M w_{oj} \left(F_h\left(\sum_{i=0}^N w_{ji} x_i\right)\right)\right) \quad (1)$$

Where w_{oj} represents the synaptic weights from neurons in the hidden layer to the single output neuron, x_i represents the i th element of the input vector, F_h and F_o are the activation function of the neurons from the hidden layer and output layer, respectively, w_{ji} are the connection weights between the neurons of the hidden layer and the inputs.

The learning phase of the network proceeds by adaptively adjusting the free parameters of the system based on the meansquare error, described by equation 2 between predicted and measured path loss for a set of appropriately selected training examples.

$$E = \frac{1}{2} \sum_{i=1}^m (y_i - d_i)^2 \quad (2)$$

Where y_i the output value is calculated by the network and d_i represents the expected output. When the error between network output and the desired output is minimized, the learning process is terminated and the network can be used in a testing phase with test vectors. At this stage, the neural network is described by the optimal weight configuration, which means that theoretically ensures the output error minimization [8].

GA

The genetic algorithm is a method for solving both constrained and unconstrained optimization problems that are based on natural selection, the process that drives biological evolution. The genetic algorithm repeatedly modifies a population of individual solutions. At each step, random individuals are selected from the current population to be parents and produce the children for the next generation. Over successive generations, the population "evolves" toward an optimal solution. This algorithm can be applied to solve a variety of optimization problems.

It starts with an initial population (which may be generated at random or seeded by other heuristics), select parents from this population for mating. Then, crossover and mutation operators are applied to the parents to generate new off-springs. Finally, these off-springs replace the existing individuals in the population and the process repeats. In this way, genetic algorithms try to mimic human evolution to some extent. A typical GA loop is shown in fig 2, [9].

International Journal of Innovative Research in Science, Engineering and Technology

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Visit: www.ijirset.com

Vol. 8, Issue 10, October 2019

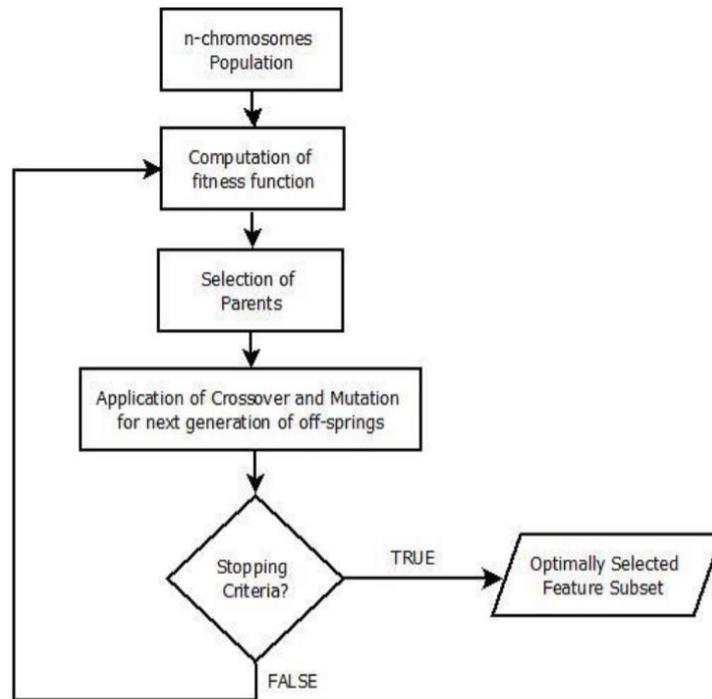


Fig 2, A typical GA loop

V.EXPERIMENTAL PROCEDURE

In this section first, a general block diagram of the proposed method is given in fig 3. Then, the following subsections will explain each block in detail.



Fig 3, Block Diagram of the proposed method

Data Acquisition

In this section, the process of building a simulator and data acquisition is proposed. 8 blades were made and installed on a disk mounted on the rotor. Then, the rotor was assembled on an electromotor via a flexible coupling (Fig 4).

International Journal of Innovative Research in Science, Engineering and Technology

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Visit: www.ijirset.com

Vol. 8, Issue 10, October 2019

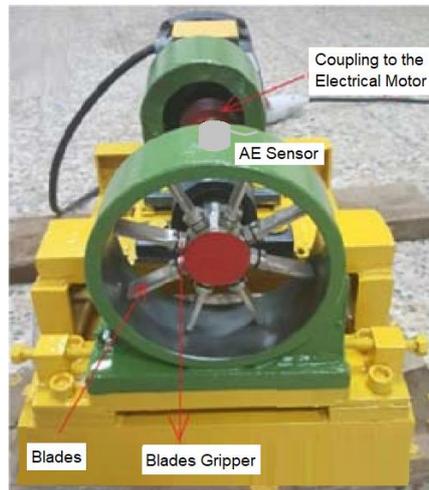


Fig 4, Experimental setup used for wear examinations

AE signals are acquired using a Physical Acoustics Corporation (PAC) PCI-2 system with AE software AEWIn. Two AE sensors (placed on the external side of the stator), with resonant frequency of 150 kHz, 20 dB pre-amplification and threshold level of the filter 40 dB are used in order to record AE's. In order to provide good acoustic coupling between die and AE sensors, a thin layer of silicone grease is applied at stator and sensor interface.

Table 1, Different types of wear

Wear type	Actions to do
All blade full wear	Removing the gap between all of the blades and the body
All blade partial wear	Removing the gap between all of the blades and the body and unbalancing the body and the rotor
One blade full wear	Removing the gap between one of the blades and the body
One blade partial wear	Removing the gap between one of the blades and the body and unbalancing the body and the rotor
Without wear	There is a gap between the blades and the body

This paper aims to classify the recorded signals into two classes called with wear and without wear. 300 signals of different types of wear and 150 signals without wear were selected to analyse in this paper. Before classification, the most important features of signals should be extracted. To obtain this goal, the most important time-domain feature parameters, stated in the next subsection, were selected.

Feature Extraction

Among different feature extraction techniques, in this study, time-domain signal analysis methods were selected. Table (2) proposes the time-domain feature parameters which were used in this investigation. A Matlab code was written for computing these parameters. As they were extracted for all 450 signals, a 7*450 matrix was the input of the neural network for classifying.

International Journal of Innovative Research in Science, Engineering and Technology

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Visit: www.ijirset.com

Vol. 8, Issue 10, October 2019

Table 2. Time-domain feature parameters, $y(n)$ is a signal series for $n = 1, 2, \dots, N$.

$RMS = \sqrt{\frac{\sum_{n=1}^N (y(n))^2}{N}}$ (3)	Standard deviation = $\sqrt{\frac{\sum_{n=1}^N (y(n) - Mean)^2}{N-1}}$ (7)
Mean = $\frac{\sum_{n=1}^N y(n)}{N}$ (4)	ClearanceFactor = $\frac{Peak}{Standard\ Deviation}$ (8)
Peak = $\max y(n) $ (5)	Skewness = $\frac{\sum_{n=1}^N (y(n) - Mean)^3}{(N-1)S^3}$ (9)
Crest Factor = $\frac{Peak}{RMS}$ (6)	

Classification

Feedforward Neural Network (FFNN) was designed and used for classifying the output into two classes. The number of extracted features and targets were selected as the number of input and output neurons of the neural network, respectively. This neural network was trained once via Bayesian Regularization algorithm and once via Levenberg Marquardt algorithm. Two different structures of the neural network, with a different number of hidden layers and neurons, shown in table 3 were used for classification.

Table 3. Structures of neural networks

Training Algorithm	Structure 1		Structure 2	
	LevenbergMarquardt	Bayesian Regularization	LevenbergMarquardt	Bayesian Regularization
Number of input layers	1		1	
Number of output layers	1		1	
Number of hidden layers	2		3	
Number of first hidden layer neurons	10		3	
Number of 2nd hidden layer neurons	5		2	
Number of 3rd hidden layer neurons	-		4	

VI. RESULT AND VALIDATION

In the field of machine learning a confusion matrix and a Receiver Operating Characteristic (ROC) curve are measurement that visualizes the performance of supervised learning algorithm. Each column of the matrix represents the instances in a predicted class while each row represents the instances in an actual class.

A Receiver Operating Characteristic (ROC) curve is a graphical plot that illustrates the performance of a classifier. The curve is created by plotting the True Positive Rate (TPR) against the False Positive Rate (FPR) at various threshold settings [10].

For evaluating the proposed algorithm based on the confusion matrix, the metrics of sensitivity, specificity, miss classification, and accuracy were used. Equations 10 to 13 show these criteria.

$$Accuracy = \frac{TP + TN}{Total} \quad (10)$$

$$sensitivity = \frac{TP}{Actual\ yes} \quad (11)$$

$$Specificity = \frac{TN}{Actual\ no} \quad (12)$$

International Journal of Innovative Research in Science, Engineering and Technology

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Visit: www.ijirset.com

Vol. 8, Issue 10, October 2019

$$\text{Miss Classification} = \frac{FP + FN}{\text{Total}} \quad (13)$$

Results of classification are given in the following subsections.

Bayesian Regularization Algorithm

As mentioned above, each training algorithm was applied to two structures with a different number of hidden layers and neurons. For both of them with this training algorithm, the accuracy of classification was somewhere in 97.5%. The results are shown in fig 7 and 8.

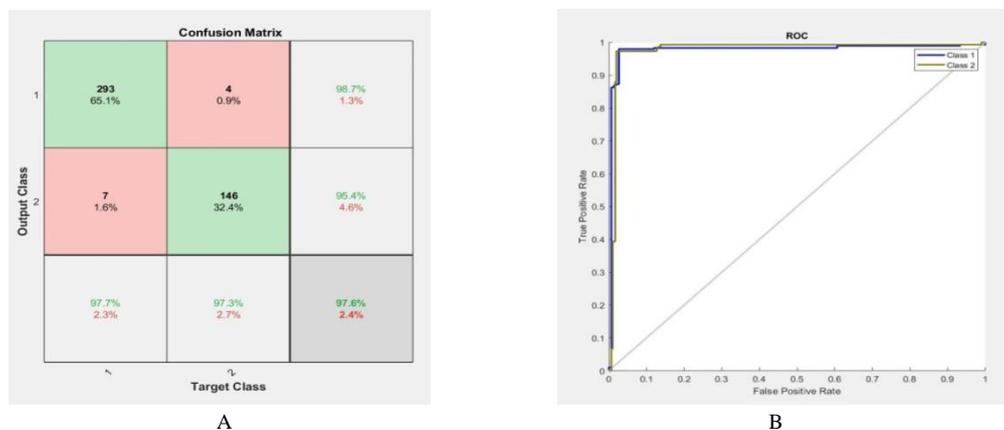


Fig 7, Classification with BR algorithm – 1st Structure. A) Confusion Matrix, B) ROC

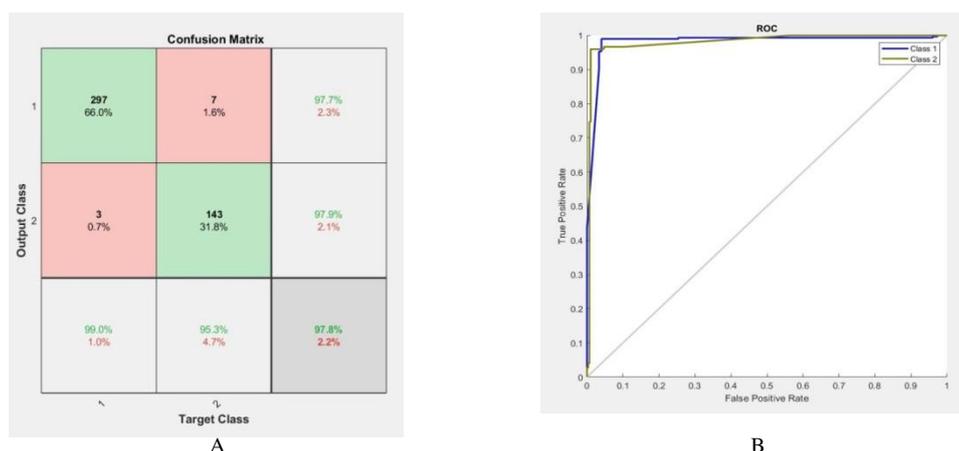


Fig 8, Classification with BR algorithm – 2nd Structure. A) Confusion Matrix, B) ROC

Levenberg Marquardt algorithm

Here, the results of classification for the two mentioned structures via the Levenberg Marquardt training algorithm are given. The accuracy of this method was approximately 20% lower than the prior one. Fig 9 and 10 present the results.

International Journal of Innovative Research in Science, Engineering and Technology

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Visit: www.ijirset.com

Vol. 8, Issue 10, October 2019

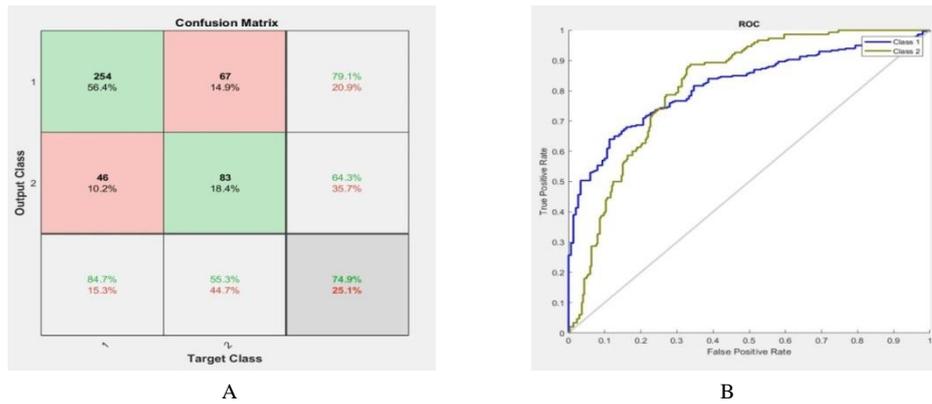


Fig 9, Classification with LM algorithm – 1stStructure. A) Confusion Matrix, B) ROC

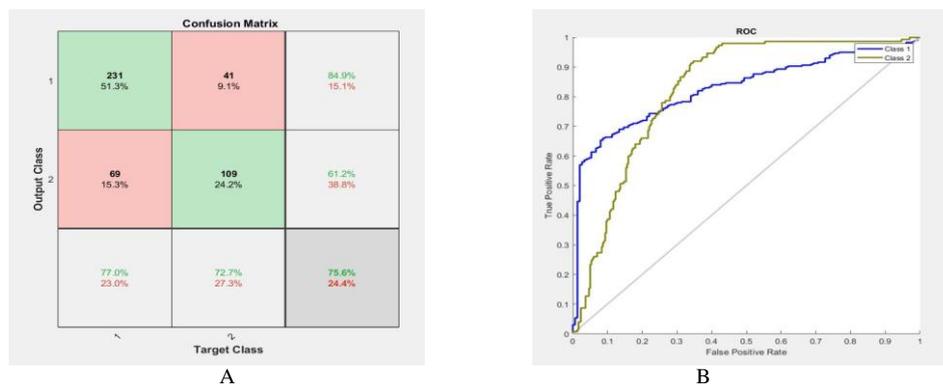


Fig 10, Classification with LM algorithm – 2ndStructure. A) Confusion Matrix, B) ROC

Results indicated the high accurate classification for the Bayesian Regularization algorithm, while outputs were not classified via the Levenberg Marquardt algorithm as accurate as of the latter one. GA was implemented for improving the performance of the Levenberg Marquardt algorithm. The results are given in fig 11 and 12. This can be derived from figures that the combination of GA and the Levenberg Marquardt algorithm improved the efficiency of classification.

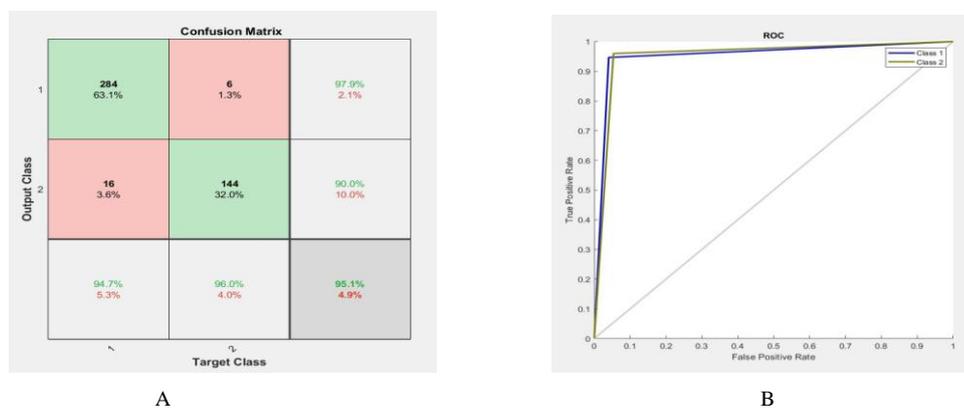


Fig 11, Classification with LM algorithm and GA- 1st Structure. A) Confusion Matrix, B) ROC

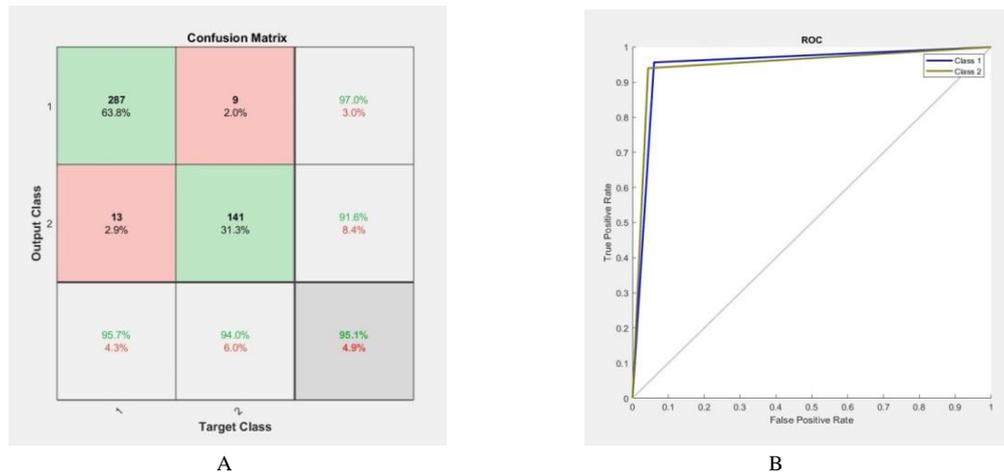


Fig 11, Classification with LM algorithm and GA- 2ndStructure. A) Confusion Matrix, B) ROC

Evaluation Criteria are calculated and given in table 4.

Table 4. Evaluation Criteria

Classification Evaluation Criteria	Bayesian Regularization		Levenberg Marquardt		GA and Levenberg Marquardt	
	1 st Structure	2 nd Structure	1 st Structure	2 nd Structure	1 st Structure	2 nd Structure
Accuracy (%)	97.6	97.8	74.9	75.6	95.1	95.1
Sensitivity (%)	97.7	99	84.7	77	94.7	95.7
Specificity (%)	97.3	95.3	55.3	72.7	96	94
Miss Classification (%)	2.4	2.2	25.1	24.4	4.9	4.9

As can be seen, GA improved the efficiency of the Levenberg Marquardt algorithm, but Bayesian Regularization, even without GA, is more accurate.

VII.CONCLUSION

This paper proposed a novel method of a binary classification for detecting wear between rotor and stator in rotary devices. First, the AE technique was used for recording signals generated from a simulator. Then, the most important time-domain features of the signals were extracted. Finally, a FFNN, with two training algorithms were applied to classify the input signals. Results indicated the high accuracy of the Bayesian Regularization algorithm, while the Levenberg Marquardt algorithm was not so efficient. To tackle the problem, GA was combined with this algorithm. This technique improved the accuracy considerably.

REFERENCES

- [1] ASTM E 569-597. "Standard practice for acoustic emission monitoring of structures during controlled stimulation".
- [2] HosseiniSadegh, ... "Classification of acoustic emission signals generated from journal bearing at different lubrication conditions based on wavelet analysis in combination with artificial neural network and genetic algorithm". *Tribology International*; 95: 426-434, 2016
- [3] Lei Lu, Jihong Yan, Clarence W. de Silva. "Dominant feature selection for the fault diagnosis of rotary machines using modified genetic algorithm and empirical mode decomposition". *Journal of Sound and Vibration*; 344: 464-483, 2015.
- [4] Mba D, Rao RBKN. "Development of acoustic emission technology for the condition monitoring and diagnosis of rotating machines: bearings, pumps, gearboxes, engines and rotating structures". *Shock Vib Dig*; 38(1):3-16, 2006.
- [5] Mohammad Reza AsadiAsadAbad, ... "Fault diagnosis of a differential using acoustic condition monitoring and artificial neural network technique". *Journal of Current Research in Science*, 1(6):575, 2013.



ISSN(Online): 2319-8753
ISSN (Print): 2347-6710

International Journal of Innovative Research in Science, Engineering and Technology

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Visit: www.ijirset.com

Vol. 8, Issue 10, October 2019

- [6] Saniya Ansari, "Optimized and efficient feature extraction method for devanagari handwritten character recognition". IEEE, International Conference on Information Processing (ICIP), 2015.
- [7] Haykin, S. "Neural networks and learning machines (3rd ed.)". New Jersey: Pearson Prentice Hall; 2008.
- [8] El-Sayed A, ... "Computer-aided diagnosis of human brain tumor through MRI: A survey and a new algorithm". Expert Systems with Applications; 41: 5526–5545, 2014.
- [9] Anam Fatima, "Android Malware Detection Using Genetic Algorithm based Optimized Feature Selection and Machine Learning". IEEE, 2019.
- [10] Chen D, Wang WJ. "Classification of wavelet map patterns using multi-layer neural networks for gear fault detection". MechSyst Signal Process 2002; 16 (4):695–70.