©2019 International Transaction Journal of Engineering, Management, & Applied Sciences & Technologies.

International Transaction Journal of Engineering, Management, & Applied Sciences & Technologies

http://TuEngr.com

ROLLER BEARING FAULT DETECTION USING EMPIRICAL MODE DECOMPOSITION AND ARTIFICIAL NEURAL NETWORK METHODS

Javad Zarekar [a], Mehrdad Nouri Khajavi [a*], Gholamhassan Payganeh [a]

[a] *Department of Solid Mechanics, Faculty of mechanical Engineering, Shahid Rajaee Teacher Training University, Tehran, IRAN*

ARTICLEINFO

Article history:
Received 15 June 2018
Received in revised form 20 December 2018
Accepted 05 January 2019
Available online 10 January 2019

Keywords:
ANN; EMD; EEMD; IMFs; Hilbert–Huang transform (HHT); Kurtosis coefficient non-stationary vibrations; time-frequency.

ABSTRACT

One of the methods for detection faults in structural and mechanical systems is processing vibrational signals extracted from the real system. The Hilbert–Huang transform (HHT) is a new and strong method for analyzing nonlinear and non-stationary vibrations based on time-frequency. This approach is based on decomposing a signal into empirical modes and Hilbert spectral analysis. In the current paper, first, vibrational signals of a roller bearing are decomposed into intrinsic mode functions (IMFs) using ensemble empirical mode decomposition (EEMD) method and IMFs sensitive to impulse are determined by Kurtosis coefficient. Then Kurtosis and standard deviation factors are extracted from the mentioned IMFs and used for training and validating the multi layers perceptron neural network. The results of network trial showed faulty or normal roller bearing and its fault type.

1. INTRODUCTION

The best method of maintaining rotating machinery is failure detection and their evaluation during working. Detecting a problem on time, when the faults are minor and have no effect on the machine operation is very useful if the problem incidence causes can be evaluated during work. Vibration analysis is the strongest and most common method to determine and detect mechanical systems' and bearings' faults when the fault begins to appear (Shiroishi et al., 1997). In addition, the existence of noises is considered a main problem in signal processing (Zhang and Randall, 2009). The reason for this issue can be known as better understanding of machine performance vibrating mechanisms and the possible relation of changes in vibrational signal with machine dynamic behavior and its faults. Moreover, vibrational signals of weak bearings are affected by stronger rotating parts such as rotors and interfere with them (Randal and Antoni, 2011). In this method, the vibrations caused by the existed faults in rotating machine parts are transferred to the body and structure and recorded by a sensor and analyzed by analyzer devices or computer software.

*Corresponding author (M.N. Khajavi). Tel: +98-9122813405 E-mail: mnouri@sru.ac.ir.

Time domain methods usually use statistical tools. Peak value, root mean square (RMS), Kurtosis (Ku) factor, crest factor (CF), etc., are some of the time domain indicators used frequently to detect rotating machinery faults (Randall, 2011). Frequency domain methods work based on obtaining frequency spectrum of time signal through Fourier transform. Frequency spectrum is very useful to find an overview of signal detection. However, as frequency spectrum shows vibration amplitude in the whole time interval of signal, it does not well reveal localized change in the signal.

2. THEORETICAL BACKGROUND

In recent years, the Hilbert–Huang transform (HHT) method has been used as one of the suggested and acceptable methods in the field of vibrational signals processing. This method was proposed by Huang el al., for the first time in 1998. The Hilbert–Huang transform is an adaptable time-frequency method for analyzing nonlinear and non-stationary signals. It is based on the empirical mode decomposition and the Hilbert spectrum analysis. Rai and Mohanty (2007) used EMD method, fast Fourier transform (FFT), and wavelet transform in their research to identify defective bearings. Considering nonlinear and non-stationary features in vibrational signal of roller bearings, Ali et al. (2015) used the method of feature extraction from empirical mode decomposition energy entropy and mathematical analysis to choose the most important intrinsic mode functions (IMFs).

De Moura et al. (2011) used artificial neural networks and principal components analysis (PCA) to detect damage severity to rolling bearings in the outer case. Using this method, they investigated four states of defect including a normal state and three different states in terms of the bearing's outer case damage severity and could differentiate between them. Jinde Zheng (2013) analyzed vibrational signal of roller bearings through generalized empirical mode decomposition, experimental envelope demodulation, and HHT methods.

2.1 EMPIRICAL MODES DECOMPOSITION (EMD) AND ENSEMBLE EMPIRICAL MODES DECOMPOSITION (EEMD)

Empirical modes decomposition method is an adaptable tool for nonlinear and non-stationary signals analysis. In this method, each signal is decomposed based on its local behavior. The results of this decomposition are intrinsic mode functions (original or manufacturer) each of which introduce as simple oscillation compared to a simple harmonic function. Data related to each signal include a variety of oscillatory modes which interfere with each other and produce complex data. Each intrinsic, linear or nonlinear mode is a simple oscillation which has extrema points (maximum and minimum) and the same zero crossing points. In the other words, oscillations around space mean are symmetric. The following algorithm is run to obtain IMFs (Rai and Mohanty, 2007):

1. The entire extrema of the original signal $x(t)$are identified.
2. Fitting curve on maximum and minimum points, two upper envelope curve e_{max} (maximum value envelope curve) and lower envelop curve e_{min} (minimum value envelope curve) are obtained.
3. The value of $m(t)$ is calculated.

$$m(t) = \frac{e_{max}+e_{min}}{2} \quad (1)$$

The value of h_1 is calculated using the following relation and this operation continues until $\mathrm{h}_1(\mathrm{t})$ value become less than the desirable amount d_1 ($d_1(t) < \in$).

$$h_1(t) = x(t) - \quad m_1 d_1(t) < \ \in \tag{2}$$

If $h_1(t)$ meets two following conditions, the first function is considered to be the intrinsic mode function IMF_1, otherwise the stages are repeated until the above condition is satisfied so that the first IMF is obtained.

First condition: In the entire data set, the number of extrema and the number of zero crossings of signal must differ at most by one.

Second condition: The mean value of the local maximum and minimum domain at any point of the signal must be equal.

4. The residual is calculated from the following relationship:

$$r_1(t) = x(t) - IMF_1(t) \tag{3}$$

By repeating this algorithm, signal components are calculated, and finally the initial signal can be calculated as follows:

$$x(t) = \ \sum_{i=1}^{n} c_i(t) + r_n(t) \tag{4}$$

Where, r_n is the residual component, n is the number of IMFs, and c_i is the intrinsic mode function (IMF). Through decomposing signal, the high-frequency and low-frequency components are obtained whose combination reconstructs the original signal. The created IMF components have a lower frequency at each stage compared to the previous one. Obtaining IMFs is known as "sifting process" or "screening". This process continues unless the standard deviation parameter restricts it. Standard deviation is obtained via two methods:

a: Through the following relation:

$$S_D = \ \sum_{t=0}^{T} \frac{(h_{k-1}(t) - \ h_k(t))^2}{h_{k-1}^2(t)} \tag{5}$$

b: The standard deviation value is usually chosen to be between 0.2 to 0.3.

The EMD method is capable to decompose complicated signal into a set of IMFs which are nearly orthogonal to each other. It seems that the most important weakness of EMD method is "modes mixing". In order to solve modes mixing problem the "ensemble empirical mode decomposition" (EEMD) method has been proposed. This method decomposes better the IMFs components (Yaguo Lei., 2011). Each trial includes the results of signal decomposition plus a white nose of finite amplitude.

The new method was extracted from statistical properties of white noise developed by Flandrin et al. (2004). It shows that EMD method applied with white noise is a self-tuning dyadic filter bank.

Moreover, investigating the obtained results by Flanderin et al. (2004), shows that noise can help with signal analysis in EMD method (Wu et al., 2009).

2.2 Ensemble empirical mode decomposition (EEMD) algorithm

1. A value is considered for ensemble M and white noise intensity and m is set to 1 ($m = 1$).

2. A set of white noise with certain amplitude is added to the signal under investigation and combined with it:

$$x_m(t) = x(t) + n_m(t) \qquad (6),$$

where n_m is the mth series of added white noise to the signal.

3. The signal $x_m(t)$ to which white noise added, is decomposed to I original mode functions using EMD method.
   a. If $m < M$, then go to stage 1 and set $m = m + 1$ and repeat stages. This process continues every time with different noise series.
   b. The ensemble mean of a_i from M trials is calculated based on the following relation for each original mode function and considered as the final IMF.

$$a_i = \frac{1}{M}\sum_{m=1}^{m} C_{i,m} \quad ,i = 1,2, \dots, I \qquad m = 1,2, \dots, M \qquad (7)$$

3. THE PROPOSED METHOD FOR ROLLER BEARING FAULT DETECTION

Using EEMD and artificial neural network, a method for roller bearing fault detection is proposed in this section.

3.1 DATA PREPARATION

The vibrational signals of the current study were provided from the bearing data center of Case Western Reserve University. In Figure 3, test bed consists of a 3-phase 2 HP (Horse Power) induction motor (left side), a torque sensor (center), and a dynamometer (right side) connected to a self-tuning coupling (center). The load amount is adjustable by dynamometer. The trial bearing keeps motor shaft on driver side. Balls, inner race (ring), and outer race defects have been separately created on bearing SKF6205-2RS with sizes of 0.177, 0.3556, and 0.5334 mm respectively (Case Western Reserve University Bearing Data Center Website). Graph of acceleration vs time obtained by this bearing is presented in Figure 4. Each sample consists of 10240 data points.

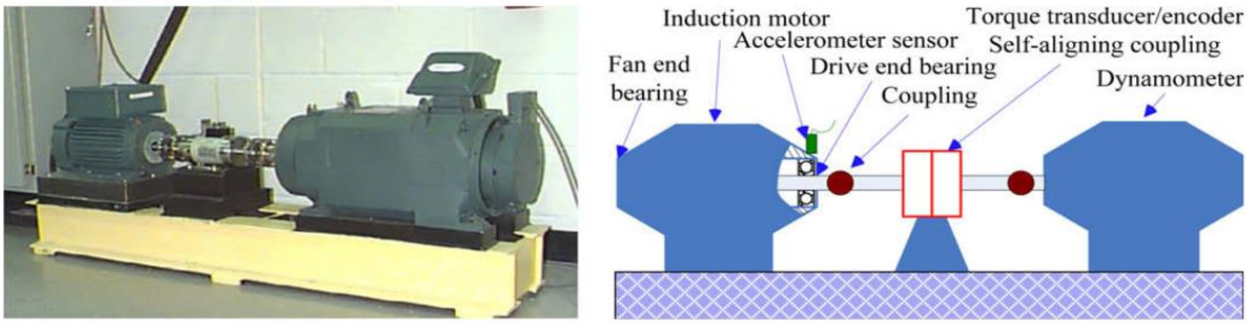

Figure 3: (a) Experimental setup (left) and
(b) Schematic diagram of the experimental setup (right)
(Available from the Case Western Reserve University Bearing Data Center Website)

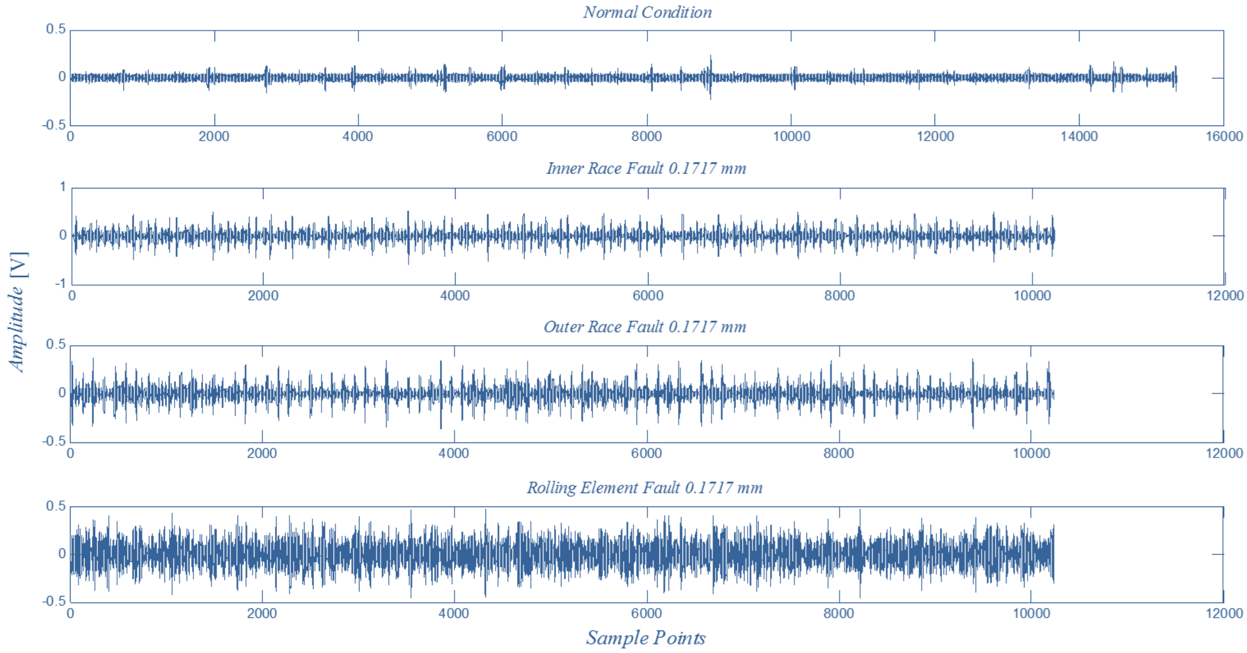

Figure 4: Vibrational signals in four conditions: (a) normal Bearing,
(b) defective bearing with inner race fault,
(c) defective bearing with outer race fault and
(d) defective bearing with rolling elements (balls) fault.

3.2 VIBRATIONAL SIGNAL DECOMPOSITION USING EEMD METHOD:

After preparing and classifying samples at previous stage, the sample signal is decomposed to 15 IMFs using EEMD method. Different faults modes are seen in Figures 5, 6, 7, and 8.

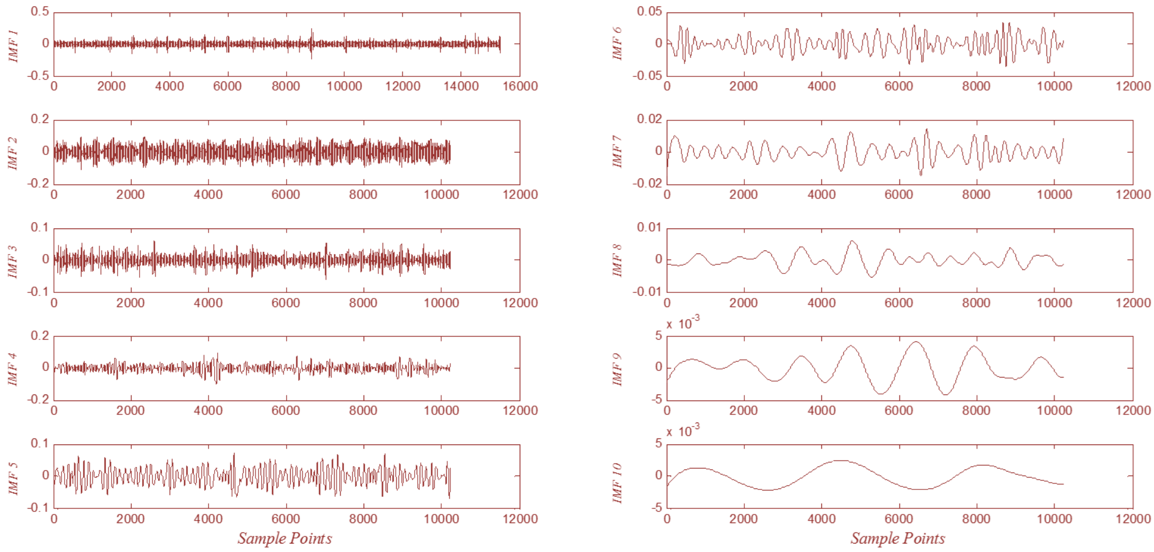

Figure 5: IMFs of normal bearing

*Corresponding author (M.N. Khajavi). Tel: +98-9122813405 E-mail: mnouri@sru.ac.ir.

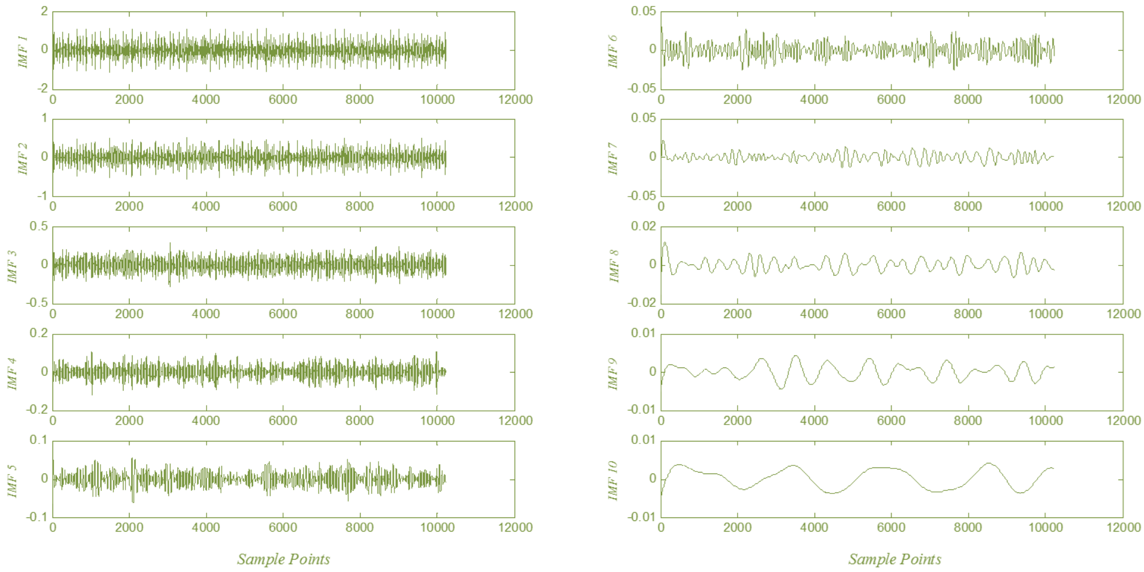

Figure 6: IMFs of inner race fault with fault diameter of 0.1778 mm.

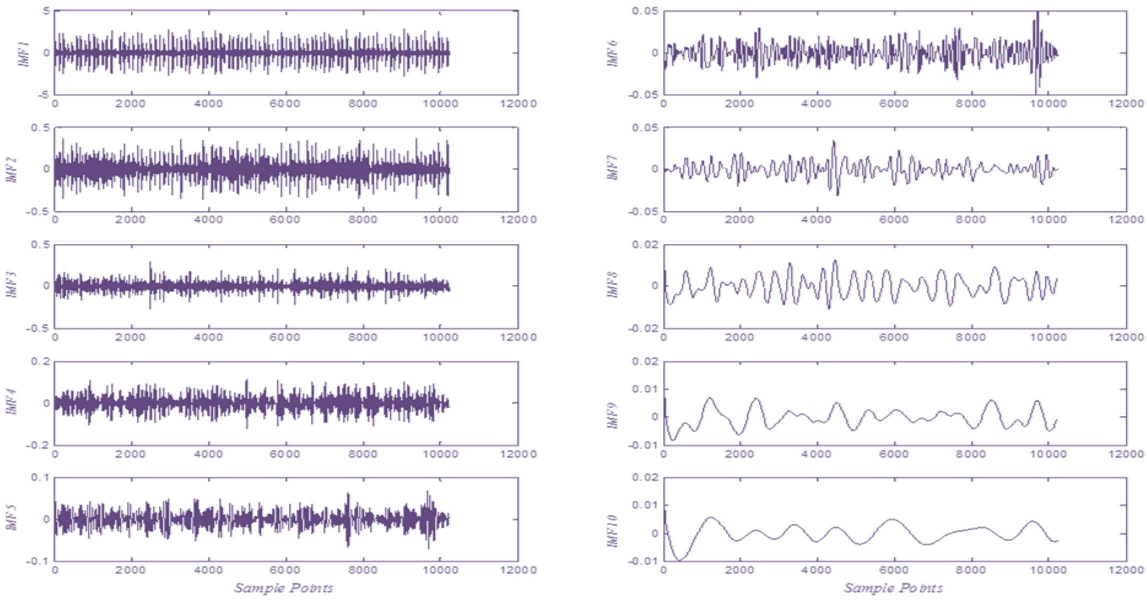

Figure 7: IMFs of outer race fault with fault diameter of 0.1778 mm

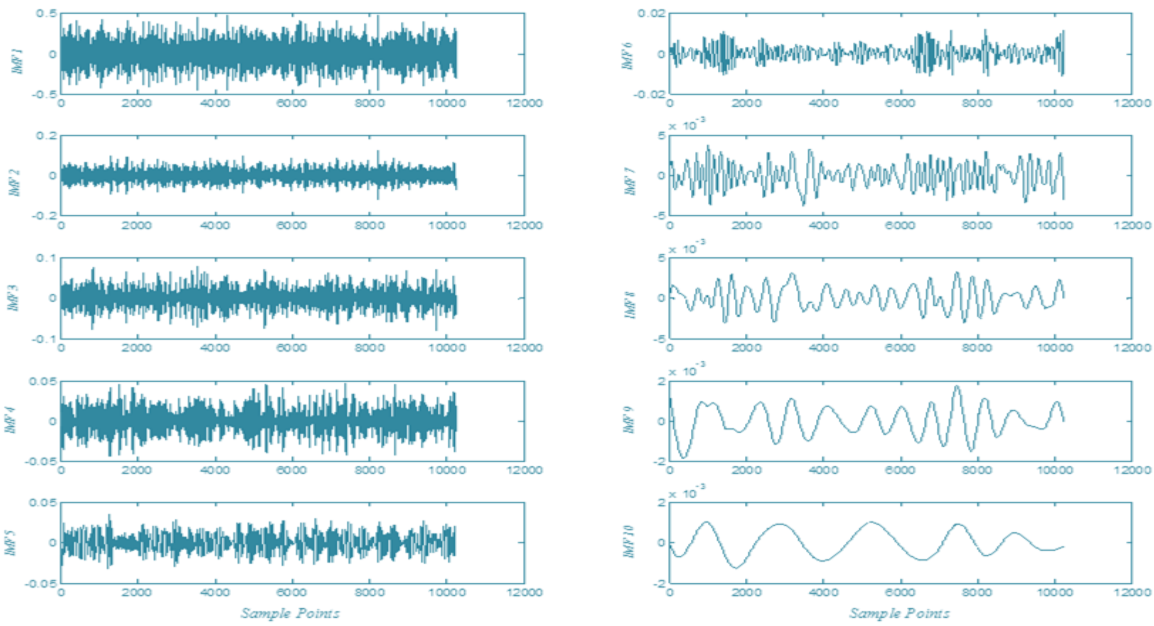

Figure 8: IMFs of defective bearing with ball fault and fault diameter of 0.11717 mm

3.3 CHOOSING SENSITIVE MODE

After obtaining IMFs for each sample, kurtosis coefficient was calculated for all 15 modes and then the more sensitive mode to impulse with higher kurtosis coefficient is chosen. Kurtosis coefficient is very sensitive to impulse and is more suitable to detect and choose the more sensitive to impulse mode function. Kurtosis coefficient is obtained from (Wasserman, 1989):

$$x_{kur} = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(x_i - \bar{x})^2} \quad (8).$$

3.4 EXTRACTING TIME DOMAIN CHARACTERISTICS:

At this stage, first, data were divided into 40 categories each of which consists of 256 data points. Then from each category, two characteristics of kurtosis and standard deviations were calculated and arrays in the form of 2×40 and 2×60 were made. Standard deviation is calculated (Wasserman, 1989):

$$\sigma = \frac{\sum_{i=1}^{N}(x-\bar{x})^4}{(N-1)\sigma^4} \quad (9),$$

where x is the nth measurement in the ith IMF x; $\bar{x}$ is the average of x, N is the number of data points of IMF x, and i is the number of IMFs.

3.5 ARTIFICIAL NEURAL NETWORK DESIGN AND FORMATION

Artificial neural networks are based on biological learning process models of human brain. Artificial neural networks are widely used in data analysis, patterns identification and control (Haykin, 1999). Multilayer perceptron (MLP) is the most common neural network. In condition monitoring, this kind of network is used in more than 90 % of cases. Multilayer perceptron neural network is composed of an input layer with source nodes, one or multiple hidden layers with computational nodes or neurons and an output layer. The number of nodes in the input and output layer depends on the input and output values. For less hidden layers and neurons, the performance may be inadequate. On the other hand, with a lot of hidden nodes the risk of over fitting in training data and weak generalization to new data exists. There are a variety of manual and systematic methods to choose the number of hidden layers and nodes.

In the current research two hidden layers with 25 and 13 neurons were used. The number of input layer nodes is two, which is equal to the number of statistical parameters (Kurtosis coefficient and standard deviation). It consists of a normal condition, three conditions related to inner race fault, three conditions related to outer race fault and three conditions related to rolling elements (balls) fault. The network output layer includes four nodes indicating normal condition, inner race fault, outer race fault and rolling elements (balls) fault respectively (Figure 9). Therefore network structure is considered to be [2:25:13:4]. Of course, different structures were prepared and tested by writing script code in MATLAB® software and finally the mentioned structure was chosen. 35% of data set was used for training, 10 % was used for validation and 55% for neural network testing. In the neural network design, the Levenberg-Marquardt (LM) algorithm and sigmoid- tangent activation function were used for hidden layer and linear activation function in the output layer. In order to determine the network efficiency the mean squared error (MSE) was used based on relation 8 (Wasserman, 1989).

$$MSE = \frac{1}{n}\sum_{i=1}^{n}(y_i - \bar{y})^2 \quad (10)$$

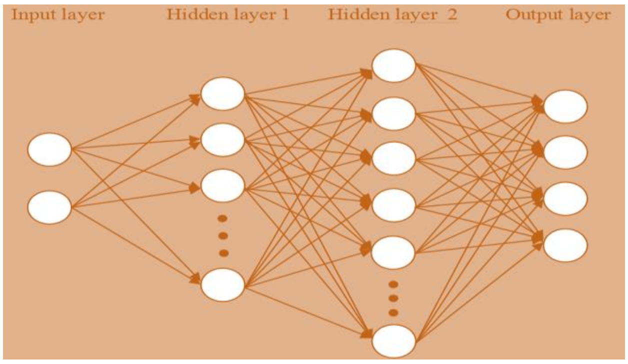

Figure 9: Schematic diagram of the designed MLP.

The input data for each part of neural network is based on Table 1. 35% of data set was used for training, 10 % was used for validation and 55% for neural network testing.

3.6 SYSTEM PERFORMANCE ANALYSIS USING CONFUSION MATRIX

In general, in the classification systems, the confusion matrix is used for defect detection in order to determine the success rate and efficiency of these systems. For analyzing the confusion matrix, in terms of the classification and bearing fault detection, four conditions are considered, including, (a) normal bearing, (b) defective bearing with inner race fault, (c) defective bearing with outer race fault and (d) defective bearing with rolling elements (balls) fault. Each of these values is shown in the turbulence matrix.

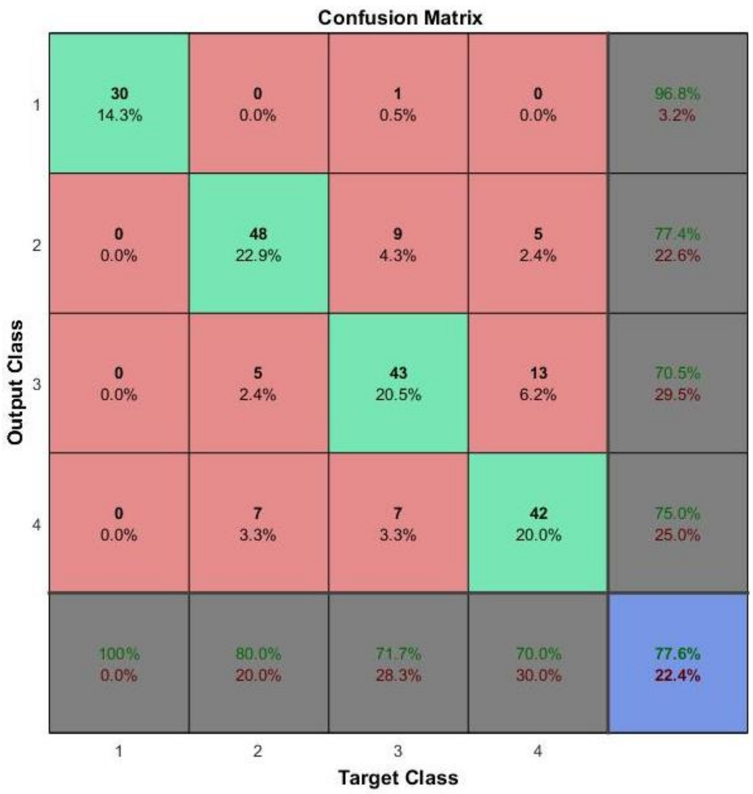

Figure 10: Confusion matrix using a test set data (neural network output)

As shown in Figure 10, confusion matrix is in fact a classified table for examining the correct and false detection modes of faults. There are four classes for output and four classes for the target. The number of data in these classes is properly classified in their own classes; and, the network test accuracy in classifying the given classes is expressed as a percentage.

Table 1: The confusion matrix

		Actual class				
		normal	inner race fault	outer race fault	rolling elements fault	Network training
Predicted class	normal	**30**	**0**	**1**	**0**	**96.8%**
	Inner race fault	**0**	**48**	**9**	**5**	**77.4%**
	Outer race fault	**0**	**5**	**43**	**13**	**70.5%**
	Rolling elements fault	**0**	**7**	**7**	**42**	**75.0%**
	Network test	**100%**	**80%**	**71.1%**	**70%**	**77.6%**

In neural networks, the network performance is measured based on the mean square error. In this study, the selected neural network performance was 77.6%, which was 15 in Epoch and the selected network performance was obtained with the test data set. The network performance is displayed in the Table 1. The network efficiency in Table 1 shows that with 15 repetitions, the lowest error rate occurs within the network, the difference between the output matrix and the target reaches its lowest value.

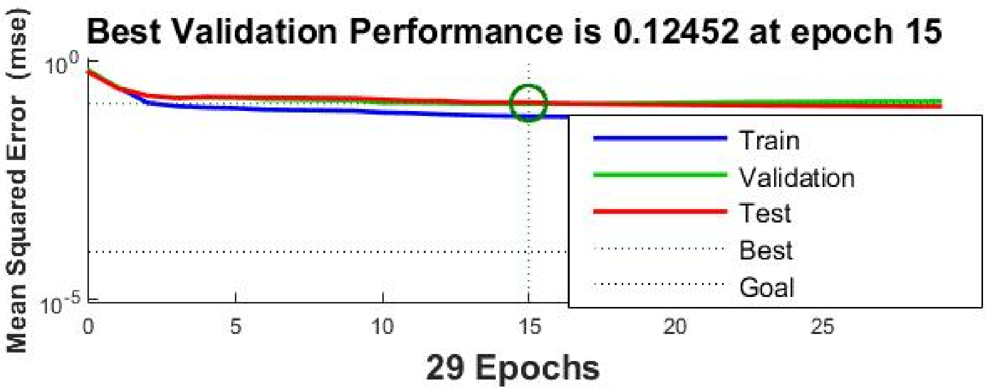

Figure 11: The least square error curve in the learning, testing, and validation of the MLP neural network.

Table 2: Description of bearing data set.

Bearing status	Fault size (mm)	Number of training and validation samples	Number of test samples
Normal		27	33
Inner race fault	0.177	18	22
	0.3556	18	22
	0.5334	18	22
Outer race fault	0.177	18	22
	0.3556	18	22
	0.5334	18	22
Rolling element or ball fault	0.177	18	22
	0.3556	18	22
	0.5334	18	22

Finally, the results of this neural network are presented in Table 2. The network accuracy of the test data in the normal state is high as 96.8%. Due to the fact that the data analyzed in this study were obtained from an accelerometer sensor on the roller bearing of actuator shaft, the vibration of the inner race fault directly affects the amplitude (due to the contact of the shaft and the inner race).

Consequently, as expected, the network test accuracy for this data set is high as 77.40%; however, it is reduced to 75.00 and 70.50 for defective bearing with rolling elements (balls) faults and defective bearing with outer race fault as in Table 2.

4. CONCLUSION

Using EEMD and MLP methods and based on time domain characteristics of vibrational signal, an approach to detect the roller bearing status was represented in the current paper. The vibrational signal data include 4 status (normal, defective inner race, defective outer race, and defective rolling element or ball) with diameters of 0.177, 0.3556, and 0.5334 mm were decomposed using EEMD method and IMFs were obtained. Considering the roller bearing performance method, Kurtosis coefficient was used for choosing and filtering the sensitive to impulse modes from the entire IMFs. Finally, two features including standard deviation and Kurtosis were extracted from these sensitive to impulse modes and given to the entry of neural network. The number of layers , transform functions, and the number of different neurons in the middle layer were investigated and the best result of neural network with 2:25:13:4 structure was obtained. This network has the mean accuracy of 78.3 % and not only detects the normal or defective state but also determine the fault type (defective inner race, defective outer race, defective roller element or ball). This algorithm is proposed to extract frequency domain or time frequency domain characteristics from sensitive modes. In addition, to increase network detection accuracy it is recommended to use this algorithm with neural networks such as support vector machine (SVM) or radial basis function (RBF).

5. REFERENCES

Ali, J. B., Fnaiech, N., Saidi, L., Chebel-Morello, B., & Fnaiech, F. (2015). Application of empirical mode decomposition and artificial neural network for automatic bearing fault diagnosis based on vibration signals. *Applied Acoustics*, 89, 16-27.

Shiroishi, J.Y.S.T., Li, Y., Liang, S., Kurfess, T., Danyluk, S., (1997). Bearing condition diagnostics via vibration and acoustic emission measurements. *Mechanical systems and signal processing*, *11*(5), 693-705.

Zhang, Y., & Randall, R. B. (2009). Rolling element bearing fault diagnosis based on the combination of genetic algorithms and fast kurtogram. *Mechanical Systems and Signal Processing*, *23*(5), 1509-1517.

Randall, R. B., & Antoni, J. (2011). Rolling element bearing diagnostics—A tutorial. *Mechanical systems and signal processing*, *25*(2), 485-520.

Randall, R. B. (2011). *Vibration-based condition monitoring: industrial, aerospace and automotive applications*. John Wiley & Sons.

Huang, N. E., Shen, Z., Long, S. R., Wu, M. C., Shih, H. H., Zheng, Q., Yen, N.C., Tung, C.C., & Liu, H. H. (1998). The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis. In *Proceedings of the Royal Society of London A: mathematical, physical and engineering sciences* (Vol. 454, No. 1971, pp. 903-995). The Royal Society.

Rai, V. K., & Mohanty, A. R. (2007). Bearing fault diagnosis using FFT of intrinsic mode functions in Hilbert–Huang transform. *Mechanical Systems and Signal Processing*, *21*(6), 2607-2615.

Jaouher Ben Ali , Nader Fnaiech, Lotfi Saidi, Brigitte Chebel-Morello." Application of empirical mode decomposition and artificial neural network for automatic bearing fault diagnosis based on vibration signals".*Applied Acoustics* , pp. 16-27, 2015.

De Moura, E. P., Souto, C. R., Silva, A. A., & Irmao, M. A. S. (2011). Evaluation of principal component analysis and neural network performance for bearing fault diagnosis from vibration signal processed by RS and DF analyses. *Mechanical Systems and Signal Processing*, *25*(5), 1765-1772.

Zheng, J., Cheng, J., & Yang, Y. (2013). Generalized empirical mode decomposition and its applications to rolling element bearing fault diagnosis. *Mechanical Systems and Signal Processing*, *40*(1), 136-153.

Lei, Y., He, Z., & Zi, Y. (2011). EEMD method and WNN for fault diagnosis of locomotive roller bearings. *Expert Systems with Applications*, *38*(6), 7334-7341.

Flandrin, P., Rilling, G., & Goncalves, P. (2004). Empirical mode decomposition as a filter bank. *IEEE signal processing letters*, *11*(2), 112-114.

Wu, Z., & Huang, N. E. (2009). Ensemble empirical mode decomposition: a noise-assisted data analysis method. *Advances in adaptive data analysis*, *1*(01), 1-41.

Case Western Reserve University. (No date). Bearing Data Center ⟨http://csegroups.case.edu/bearingdatacenter/home⟩. Accessed June 2018.

Wasserman, P. D. (1993). *Advanced methods in neural computing*. John Wiley & Sons, Inc.

Haykin, S. (1999). Neural networks: a comprehensive foundation", 2nd Ed. Prentice Hall, New Jersey.

.

Javad Zarekar is a PhD student of Mechanical Engineering, Department of Solid Mechanic, Faculty of Mechanical Engineering, Shahid Rajaee Teacher Training University, Tehran, IRAN. He is interested in Artificial Intelligence, Signal Processing and Vibration Condition Monitoring.

Dr.Mehrdad Nouri Khajavi is an Associate Professor at Shahid Rajae Teacher Training University, Tehran Iran. He has been head of automotive engineering group for 6 years. He received his PhD degree in Mechanical Engineering from Amir Kabir University. His research interests includes: Artificial Intelligence, Signal Processing, Vibration Condition Monitoring and Vehicle Dynamics.

Dr.Gholamhasan Payganeh is an Associate Professor at Shahid Rajae Teacher Training University, Tehran Iran. He is Head of Mechanical engineering Faculty. He received his PhD degree in Mechanical Engineering from Istanbul Technical University. His research interests includes: Fatigue on the gears, Signal Processing, Dynamics of vibrations, Designing machine components, Dynamics and vibrations.