Using MLP-GABP and SVM with Wavelet Packet Transform Based Feature Extraction for Fault Diagnosis of a Centrifugal Pump

Maamar Al Tobi[1], G. Bevan[2], P. Wallace[3], D. Harrison[4] and K. E. Okedu[5]

[1, 5] *National University of Science & Technology, College of Engineering, Oman*

[2, 3, 4] *School of Engineering and Built Environment, Glasgow Caledonian University, UK*

Abstract

In this paper, a hybrid training method for Multilayer Perceptron (MLP) is proposed based on combining Back Propagation (BP) and Genetic Algorithm (GA). The proposed scheme is compared with the Support Vector Machine approach to classify six fault conditions and the normal condition of a centrifugal pump. Two training algorithms were tested and compared. Features were extracted using Wavelet Packet Transform (WPT) with three levels decomposition, and two mother wavelets were used to investigate their effectiveness on feature extraction. Furthermore, GA is also used to optimize the number of hidden layers and neurons of MLP. The results obtained, show improved performance on the feature extraction, GA based hidden layers and neurons selection, training algorithm, and classification performance using the proposed scheme.

KEYWORDS: Back Propagation (BP), Wavelet Packet Transform (WPT), Centrifugal pump, Genetic Algorithm (GA), Multilayer Feedforward Perceptron (MLP), Support Vector Machine (SVM).

1 Introduction

Various techniques have been applied to fault detection of centrifugal pumps based on condition monitoring such as time domain analysis [1], and frequency domain analysis, where methods such as the Fast Fourier Transform (FFT) are applied [2, 3]. Also, a powerful multi-resolution technique called wavelet has been applied in rotating machinery fault detection and

E-mail: altubi555@gmail.com
E-mail: geraint.bevan@gcu.ac.uk
E-mail: p.wallace@gcu.ac.uk
E-mail: d.harrison@gcu.ac.uk
E-mail: Okedukenneth@nu.edu.om

has proved its ability to analyze non-stationary signals of machines [3-5]. Wavelet Transform (WT) is a method that has the ability to deal with non-stationary signals as it is known for its time-frequency domain resolution. However, looking for a more advanced and automatic fault diagnosis method, then, Artificial Neural Network (ANN) is a promising technique that has proved its ability as a fault classifier. ANN has been applied to centrifugal pumps as automatic fault diagnosis and classification systems in [6-10]. WT is a mathematical operation that converts a time domain signal into another form. To apply a wavelet transform, a wavelet function is required, which represents a small wave with oscillating wavelike characteristics and focuses on its short time energy. Wavelet transforms can be classified into three groups: continuous wavelet transform (CWT), discrete wavelet transform (DWT), and wavelet packet transform (WPT). Reference [11] proposed WPT and ANN (MLP) for helicopter gearbox fault detection. Eight different detection locations were identified for the vibration monitoring. WPT was used to de-noise and decompose the vibration signals, then the standard deviations were extracted from the decomposed four levels and used as inputs and resulted at an average rate of 99.25%. Reference [12], used three different classifications, namely, MLP, SVM and Radial Basis Function (RBF) to diagnose a fault of rub impact. WPT with db4 wavelet function is used for the feature extraction, and the classification rates are 82%, 99.3% and 98.6% for MLP, SVM and RBF.

Genetic Algorithm (GA) has been used for the selection of input features for machinery components like bearings and gears [13-15] and the number of neurons in the hidden layers [16] of MLP- ANN. Application of GA for the centrifugal pump have been applied with two WT methods; CWT with best classification rates of 99.5% and 94.64% using MLP and SVM respectively in [17], and DWT with best classification rates of100% and 99.8% using SVM and MLP respectively in [18]. It is also applied in training MLP combined with BP using both CWT and DWT based feature extraction, and the best rates are 88.5% and 89% respectively.

This paper investigates the classification performance of two artificial intelligence methods: MLP-BP along with GA based selection and SVM for different conditions of a centrifugal pump. The procedure consists of three main stages, namely, data collection, pre-processing and extraction, and fault classification. The feature extraction is implemented using WPT where the signals are decomposed into three levels and both the approximation (low frequency) and detail (high frequency) coefficients are extracted based on the decomposition tree. Two mother wavelet functions are selected and tested with WPT, namely, db4 and rbio1.5 to investigate their ability and impact on feature extraction. Classification and diagnosis of the centrifugal pump condition is implemented using two artificial intelligence classifiers, namely, MLP and SVM. MLP is implemented along with its traditional learning algorithm (Back-Propagation) and is also compared with a hybrid training algorithm (MLP-GABP). The network hidden layers and neurons are selected manually and also optimized using GA with comparable results. The flow chart of the diagnosis methods and training algorithm is shown in Figure 1. The performance is determined in terms of the number of hidden layers and neurons in the neural network, number of features, and the training and kernel methods. This paper is divided into six parts including this introduction. Section 2 presents a brief review of artificial intelligence systems including MLP-NN, SVM classifiers and GA. Section 3 illustrates the experimental setup. Section 4 outlines the method applied and procedures using WPT for feature extraction. Then, section 5 presents the results and discussion. Finally, a conclusion with remarks and recommendations is given in section 6.

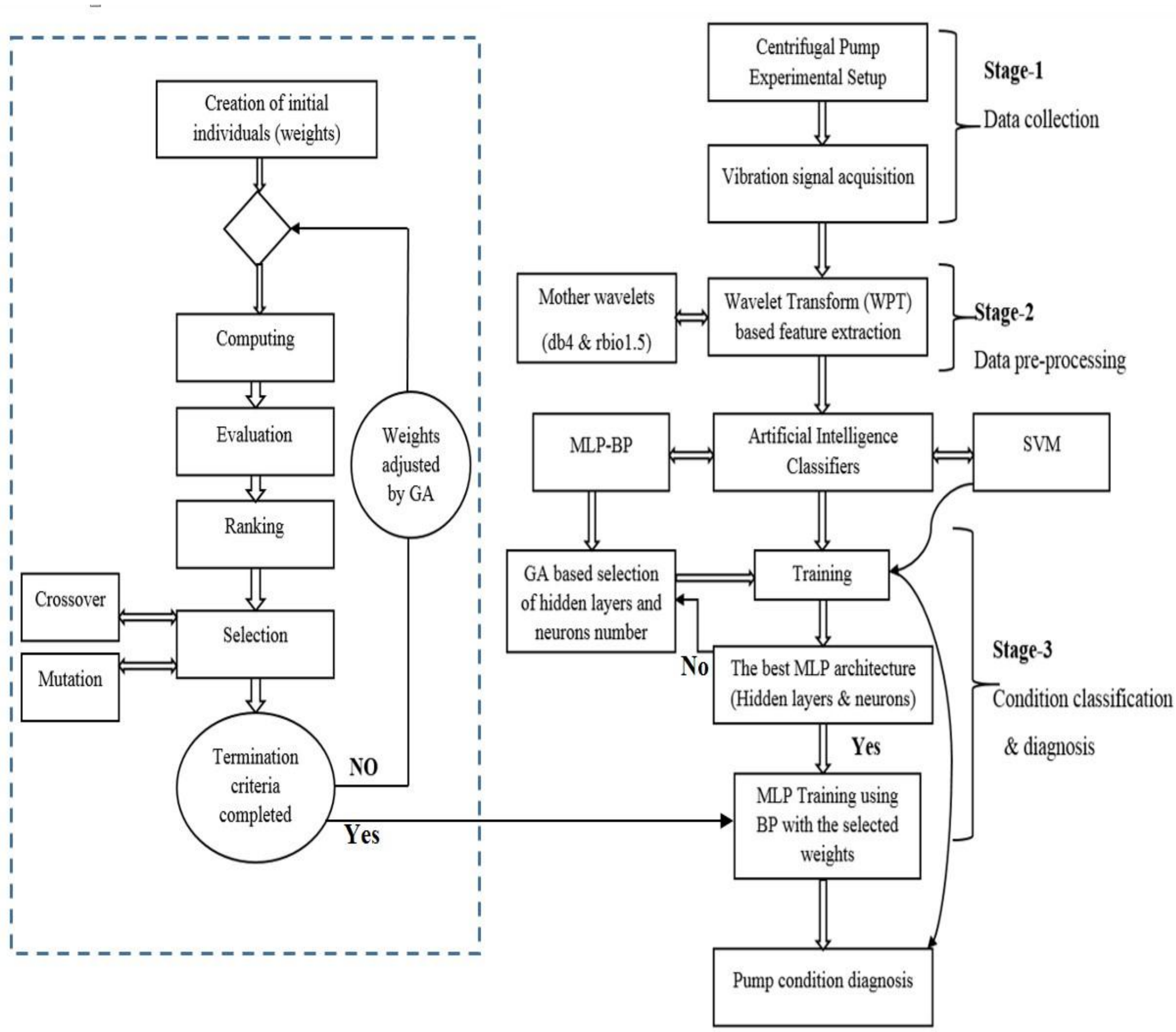


Fig. 1 Flow chart of diagnosis methods and training algorithm

2 Artificial Intelligence systems

Automatic fault detection methods make use of Artificial Intelligence (AI) which seeks to replicate mental capabilities with the support of computational systems [19]. An artificial Neural Network (ANN) was first introduced by [20], and Fuzzy Logic was first introduced by [21].

Artificial intelligence systems have been applied for centrifugal pump fault diagnosis using different methods for the feature extraction, starting from a simple method of statistical analysis [10, 11], later FFT [22-25], and also a wavelet transform has been applied using a time-frequency method [6, 27-30] proposed ANN with Back Propagation (BP) algorithm to diagnose pump faults. Then, [31] applied ANN and a fuzzy neural network to diagnose centrifugal pump faults; statistical methods of time and spectral analysis were used for the feature extraction.

There are many types of AI that have been applied as automatic fault diagnosis systems for different rotating machines and components such as Back Propagation Artificial Neural Network (BP-ANN) or Multilayer Perceptron (MLP) [13-16], and Support Vector Machine (SVM) [32-37].

2.1 Multilayer perceptron with back propagation

Multilayer Perceptron (MLP) consists of three layers, namely, input, hidden, and output layer of neurons. There may be several hidden layers between the input and output layers. The number of neurons in each section affects the generalizability of the system, while the number of neurons and hidden layers affects the efficiency. With a larger number, there is a possibility of over-fitting the training data and weak generalization of new data. Therefore, some methods might be used to select the appropriate number of hidden layers and neurons such as Genetic Algorithm [38]. The output layer can be more than one layer according to the required fault classifications. Each hidden layer has a number of neurons; the role of each is to calculate the weighted sum of its inputs and apply the sum as the input to an activation function that is usually a sigmoid. The Back Propagation algorithm has been widely used in training of MLP. It was first introduced by [39]. Comparative studies have demonstrated the efficiency of MLP over other ANN types [16, 38]. However, a drawback of MLP is that it is slow in training and needs longer computational time than other methods [35, 37]; but such weakness can be minimized by reducing the number of input features [32].

2.2 Support vector machine

Support Vector Machine (SVM) was initially introduced by [40], where it was used as a new approach for pattern recognition, employing non-linear projections of input features to a greater dimensional pattern area. The working principle of SVM is illustrated in Figure.2.

The SVM is working to separate (classify) two different classes (conditions); class A and class B to as shown in Fig. 2. The optimal hyper plane (separator) is separating the two classes with a maximum width (margin) as the larger margin (width) between the two classes the more generalization and eventually better linear classification. The linear classifier (hyper plane) is expressed as WTX+b= 0. If class A is assumed to be above the hyper plane, then it is >0 and indicated as +1 and given by WTX+b=+1, and class B is < 0 as -1 and given by WTX+b= -1. Where W is the weight vector and b is the bias.

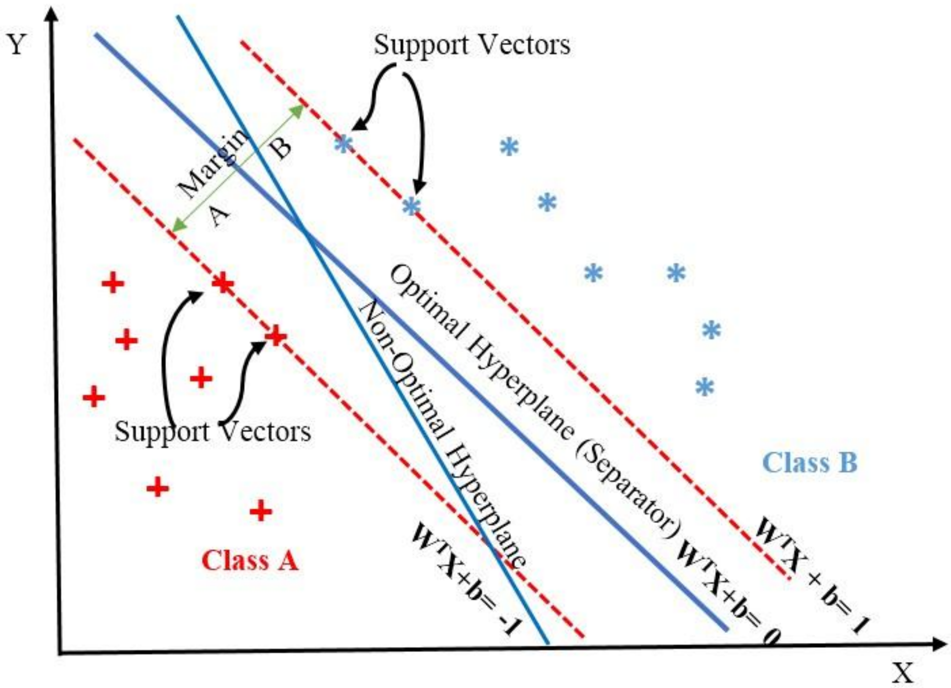


Fig. 2 Working principle of Support Vector Machine

SVM has been widely recommended for rotating machinery fault diagnosis as it has proved its high efficiency and out-performance over other AI classifiers e.g. MLP (ANN-BP) [12, 13, 15], and RBF [34]. In this work, SVM is used as a classifier and its performance is compared with MLP. MATLAB software is used to implement the classification stage, for which tools and codes were developed. Classifiers consist of two main processes: training of data, and testing, where an automatic classification is implemented for the different conditions. The performance of the AI classifier is measured according to the classification accuracy rates (%). Polynomial kernel function is selected for SVM.

2.3 Genetic algorithm

Genetic algorithm (GA) was introduced by [41]. It is based on the concept of a Darwinian-type fitness for survival that it is used to produce better individuals for the desired problem, as different possible solutions compete and match with each other. It is essentially a form of optimization, which can be applied to complex functions. GA has a similarity with chromosomes, in that individual terms are represented by means of a linear string [42]. The basic concept of GA processes is illustrated in Fig. 3. GA starts its process by initiating individual populations which are known as chromosomes where they then would be computed and evaluated individually based on fitness and then they would be ranked according to the higher fitness after which selection are based on the top survival individuals (their fitness). GA has two main operators, namely, crossover and mutation, and they operate to produce a new generation of individuals (chromosomes) and then would be sent to the first step of the process as the improper individuals are replaced with the new and good ones [42].

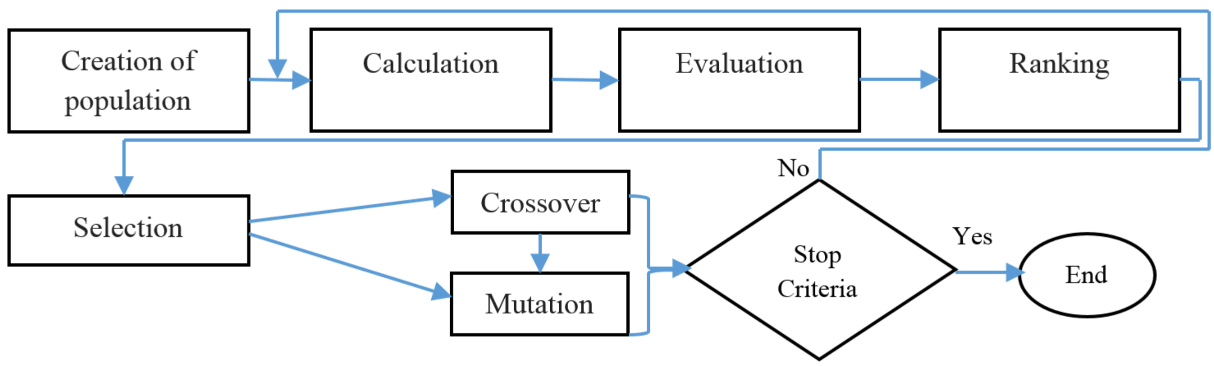


Fig. 3 Flow chart of genetic algorithm

In this work, GA is used to optimize the number of hidden layers and neurons to select the optimal architecture of the neural network using MATLAB software. The GA based selection and optimization has been developed as a MATLAB code where the range of constraints and parameters are exploring a parameter space ranging from 1 to 4 layers with up to 30 neurons per layer, 20 generations, and population size is 10 individuals to avoid long computational time. GA is also used to train MLP along with BP using 1000 generations and 1000 population size.

3 Experimental setup

The centrifugal pump experiment has been designed and assembled specifically for this research where it consists of several parts including: centrifugal pump which is coupled with a motor (Saer company, Italy, model: NCBZ-2P-50-125C, 2.2 kW, 3-phase, 420 V, head 8-17 m and flow rate 500-1000 L/M), control panel with speed controller (Schneider model VFD with speed controller and display screen, switch (OFF/ON) and emergency shutdown), digital turbine flow meter (USA-TM model, 2 inch diameter), pressure gauges, vacuum pump and clear PVC pipes; and spare parts: a rolling element bearing, mechanical seal, gasket and impeller. A data acquisition system (DAQ) and accelerometers from National Instruments (NI) are used. The DAQ system comprises SCXI-1000 and SCXI-1530 models. The accelerometer model is IMI 621B40 with sensitivity of 10 mV/g and frequency range from 3.4 Hz to 18 kHz for (±10%) and 1.6 Hz to 30 kHz for (±3 dB). Figure 4 shows the centrifugal pump experimental setup with the faulty impeller and bearing.

Fig. 4 Experimental setup

The vibration signals are measured under two conditions, namely healthy and faulty. Firstly, the signal of normal condition is acquired when the pump is healthy, without any faults. Secondly, the faulty conditions are divided into two main categories; mechanical faults (bearing, misalignment, unbalance, impeller, and looseness), and one hydraulic fault (cavitation). These faults are created and simulated one by one. Signals are acquired from the pump using an accelerometer which is mounted on its bearing housing. This sensor transfers the vibrational data to the data acquisition device (DAQ) where the signals have to be amplified and noise filtered out; and then moved to a computer which is equipped with a digital/analogue converter card (D/A) in order to convert the analogue signals to digital. The sampling rate applied for data acquisition is 16 KHz and 2.4 s as a sampling time with 38400 number of samples. Finally, these signals are captured via LabVIEW software where raw signals are saved in order to use them in the second stage for further processing.

All data of the pump conditions are acquired with the speed of 20 Hz (1200 RPM).

4 Feature extraction

The purpose of feature extraction is to extract some characteristics from the vibration signals to be implemented in a neural network. It is important to ensure good feature extraction and selection otherwise weak classification performance might result [43]. Reference [44] recommended that the extracted features have to be strongly relevant to the machine faults. However, they stated that feature extraction methods have difficulty with vibration signals that contain strong noise which conceals the important information. This difficulty has driven researchers to apply wavelet transform analysis in order to perform noise cancellation for the feature extraction [45].

WPT was introduced by [46] and is a multi-stage filtering method that de- composes a signal into packets or levels of approximation which are denoted with A, and details coefficients which are denoted with D, as illustrated in Figure 5 [34, 47, 48]. The WPT is defined as:

$$W_{(j,k)}(t) = 2^{\frac{j}{2}} w\left(2^{j}x - k\right) j\epsilon \mathrm{Z} \quad (1)$$

WPT is similar to DWT except WPT provides higher and finer decomposition tree, where both approximation (A) and detail (D) can produce pairs of packets (second level of approximation and detail), but DWT does not have such ability (i.e. the next or second level of approximation and detail can be split by the approximation (A) only). WPT has been applied for other types of rotating machinery [43, 48, 49-51]. In this work, WPT using two mother wavelets (db4 and rbio1.5) is applied for the preprocessing and feature extraction. Three cases are considered and they are as follows:

4.1 Case 1

The signals are decomposed to 3 levels for the feature extraction where the approximation and detail coefficients are extracted from 7 different pump cases. In each case, a signal of length 34800 samples was recorded. These signals were each divided into 5 segments, of length 7680 samples.

The five segments produce a total of 60 features. From these 60 features, 6 parameters (Kurtosis, RMS, Peak, Crest Factor, Shape Factor and Impulse Factor) are computed for the signal from each case. Figure 5 shows the description of the WPT tree decomposition to three levels, where A denotes the approximation, and D refers to the detail. Figure 6 illustrates the third level tree decomposition of imbalance condition using the db4 function, where the general sinusoidal pattern of the signals has better representation with the approximation decomposition, and it is also preserved in successive approximation levels (but not the detail levels). It is also remarked that approximation reveals successively less noisy signals by reducing the high-frequency information in which could be resulting in extracting better features than the ones from detail. The best approximation decomposed signals (A1, AA2 and AAA3) are considered for this work, as they have successively less noise. Therefore 60 features of both approximation and detail are used to train the MLP-BP, where the desired number of features (60) are completed with considering 14 features from each segment (except the fifth segment, as the total number of 60 features are extracted with discarding the last 10 features from the fifth segment. It is also considered that the network is trained with the best three approximations per each segment, except the fifth segment (due to the intention of considering the required number of features only with discarding the unnecessary ones, where the first two approximations from the fifth segment are considered only to have totally 14 features from each condition.

4.2 Case 2

The signals are also decomposed to 6 levels with db4 only, and from 5 segments of each signal, the first 3 approximation packets of each level are selected. The total features per condition and parameter are 30.

4.3 Case 3

The signals are finally again decomposed in 3 levels, and each signal is divided into 8 segments with a length of 4800 samples. The first 3 approximation packets of each level are selected. The total features per condition and parameter are 24.

These three cases are analyzed to determine the best number of features and types of coefficients for classification accuracy. For the SVM, one case is considered; using 2 parameters and 14 features. The extracted features are normalized. The effectiveness (sensitivity) of each parameter against all conditions are plotted in Fig. 7. Normally, when healthy (blue) is the lowest, it indicates good effectiveness of the parameter. Therefore, peak and RMS are selected for SVM due their ability in distributing and distinguishing the conditions effectively.

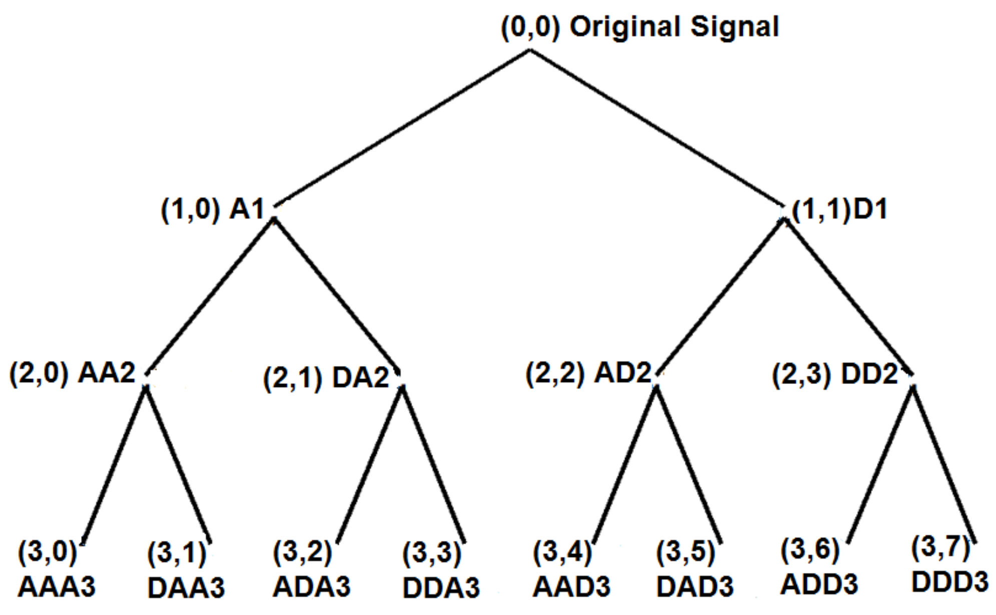


Fig. 5 WPT tree decomposition to three levels

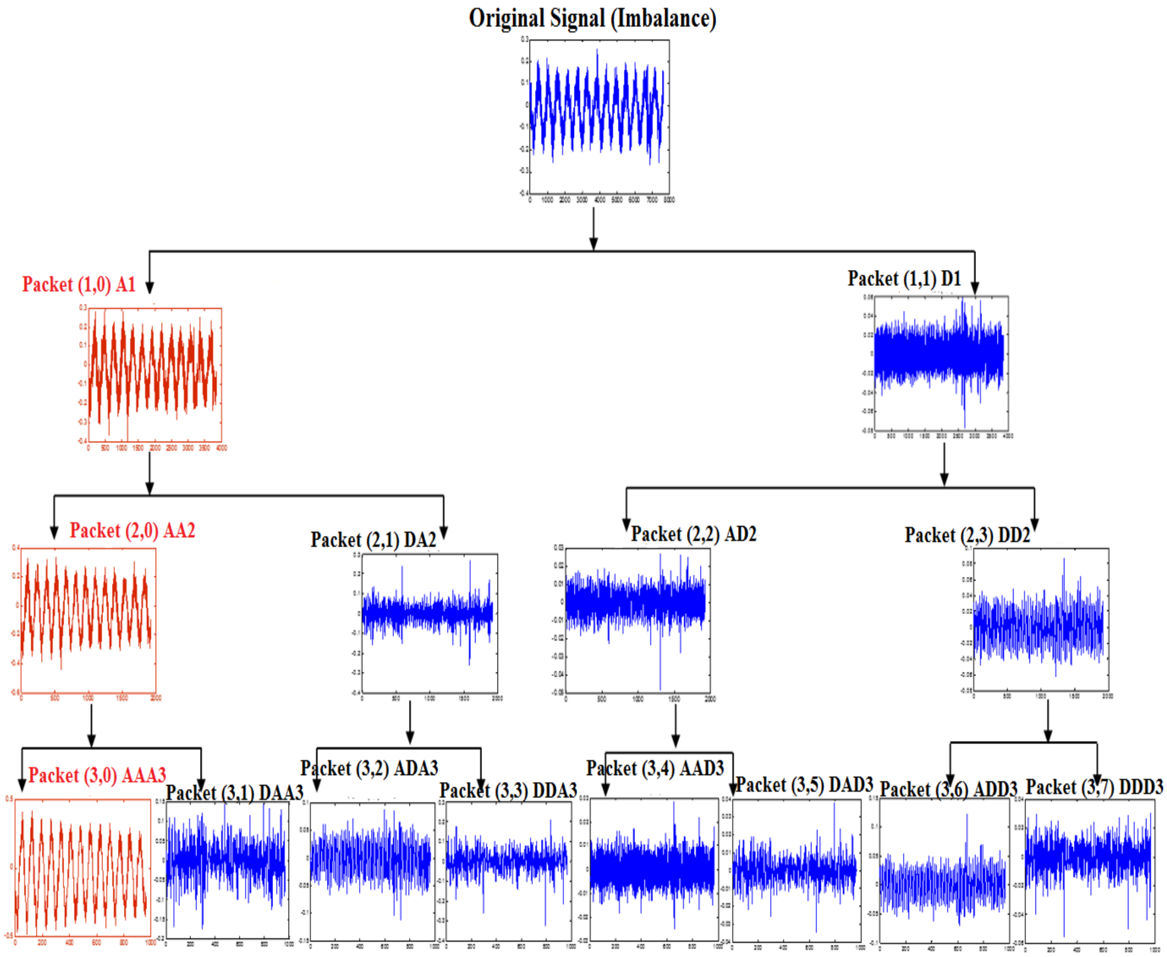


Fig. 6 The third level tree decomposition of imbalance condition using db4 function

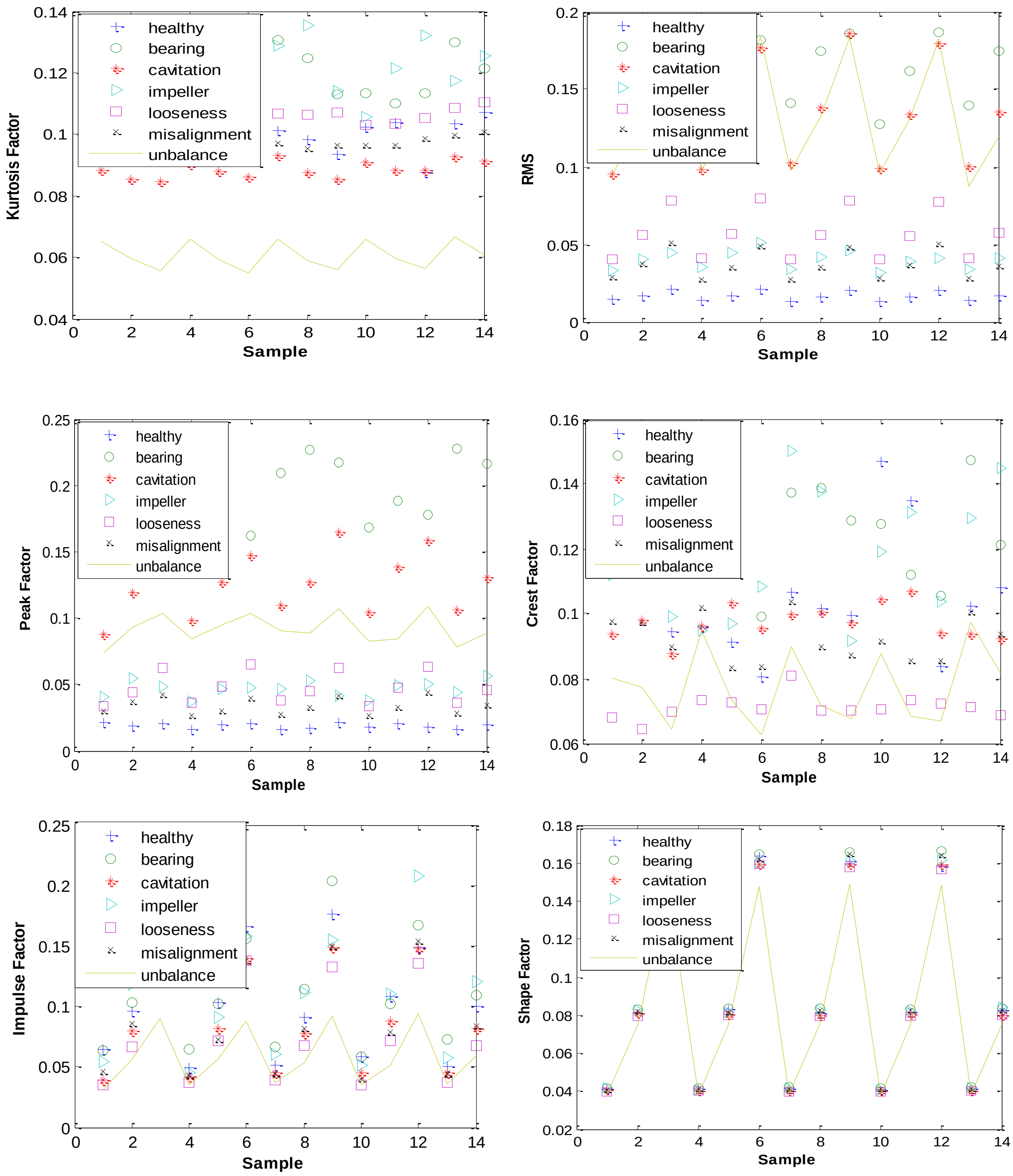


Fig. 7 The effectiveness of each parameter against all conditions (db4 mother function)

5 Classification methods

The extracted features were used as input vectors that were forwarded to the neural network classifier and SVM. In this work, MLP consists of three layers, namely, input layer, hidden layer, and output layer. The input layer consists of 6 neurons which represent the extracted and normalized features for each parameter that are pre-processed using WPT. The number of hidden layers and neurons were optimized and then selected using GA. The output layer consists of 7 neurons; one for each tested pump condition; one neuron for a healthy condition, and six neurons for six different fault conditions. The network is trained using Levenberg-Marquardt (LM) function which is a back propagation algorithm to update weights and biases. As shown in section 4, there are three cases for the extracted features based decomposition levels and number of signal segments, 60, 30, 24, and 14 features (60 and 14 features are used as normalized and non-normalized) per condition with a total of 420, 210, 168 and 98 input features for all conditions per parameter are forwarded to the MLP-ANN which results in a matrix of size [6 × 420], [6 × 210], [6 × 168] and [6 × 98] respectively. The input vectors are divided into three datasets (training has 70%, test has 15% and validation has 15%. The target for training is a Boolean matrix of size [7x420] (60 features), [7x210] (30 features), [7x168] (24 features) and [7x98] (14 features); with the rows corresponding to the 7 conditions: (cases) in which seven digits-coding and each digit represents a block of size (1x60), (1x30), (1x24) and (1x14) respectively given as follows:

1-	Healthy	[1 0 0 0 0 0 0]
2-	Bearing fault	[0 1 0 0 0 0 0]
3-	Cavitation	[0 0 1 0 0 0 0]
4-	Impeller fault	[0 0 0 1 0 0 0]
5-	Misalignment	[0 0 0 0 1 0 0]
6-	Looseness	[0 0 0 0 0 1 0]
7-	Imbalance	[0 0 0 0 0 0 1]

Neural network details and structure are shown in Fig. 8 and Table 1, respectively.

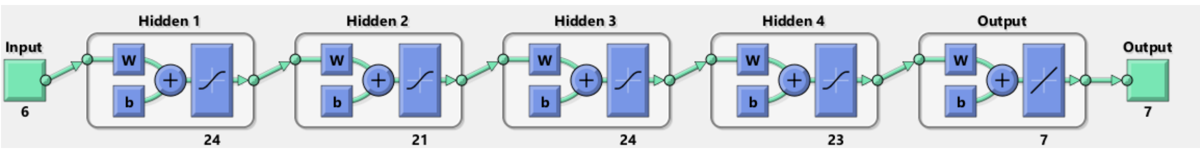


Fig. 8 Network structure

Table 1 Multilayer Feed-Forward Perceptron neural network architecture parameters

MLP-ANN details				
Transfer Function		No. of input neurons	No. of hidden layers and neurons (GA based selection)	No. of output neurons
Hidden Layer	Output Layer			
Sigmoid	Linear	6	4 layers: [24 21 24 23]	7
ANN Training Parameters				
Training Algorithm		Learning Rate	Training Stop Criteria	
LM (BP) & GA		0.56	Max. epoch	MSE
			1000	10E-20

SVM classification has been applied using 14 features (normalized) which are representing the best three level approximations. All the seven conditions were tested and classified against each other, and all results are compared.

The SVM has been investigated using a polynomial kernel. The parameter C (width) set to 3. A randomly selected input dataset is divided into a training set and a test set. Two parameters (RMS and Peak) are used for both conditions since both parameters are found to distinguish well between the conditions.

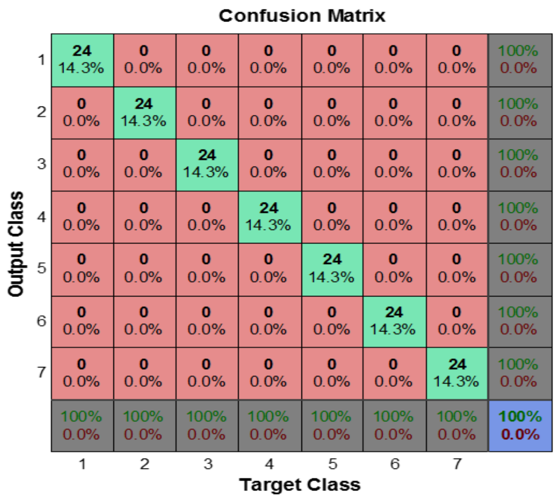
6 Results and discussion

The results of using these AI methods (MLP-BP, MLP-GABP and SVM) show the strengths and drawbacks of each method including the impact of mother wavelet selection, using approximation detail features, and normalized or non-normalized features.

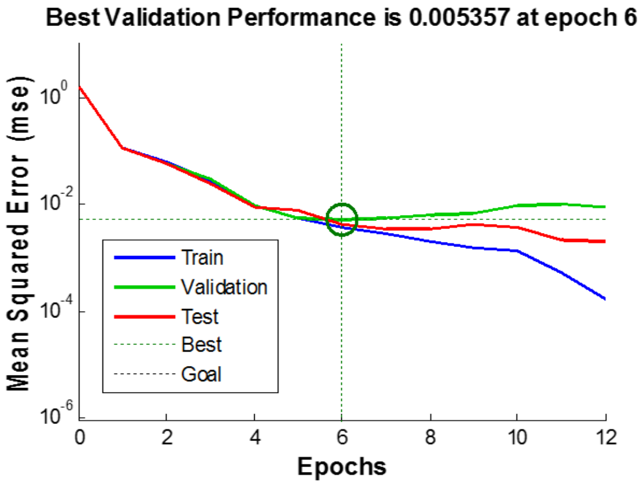
6.1. MLP-BP

Classification rates are remarked for all cases with db4 mother wavelet using GA based selection which suggested four hidden layers containing [24 21 24 23] neurons. Using 14 approximation normalized and non-normalized features, presented overall classification rates of 100% and 98% respectively, but it is remarked that the test classification is successfully conducted for 6 cases out of 7. To avoid such misclassification, higher number of features are used. Classification rates of 75.5% using 60 normalized features, 71.2% using 60 non-normalized features, 97.6% using 30 normalized 6 level approximation features and 100% using 24 normalized features were obtained. However, only 6 out of 7 cases are classified (test) with 14 approximation normalized features of 100, and 14 approximation non-normalized features of 98%.

Based on the results using db4, classification rates with rbio1.5 mother functions is conducted using 14 and 24 approximation normalized features only and classification rates are 100% for 6 out of 7 cases (validation) and 100% for all 7 cases respectively. Therefore, the best accuracy rate is achieved using 24 approximation normalized features; the overall confusion matrix for training, testing, and validation is illustrated in Fig. 9. The lower right blue square shows the overall classification rates, where overall, 100% (in green) of the classifications are correct and 0% are incorrect classifications. Taking each classes' accuracy rate (pump conditions) individually; healthy (case 1) has the accuracy rate of 100%, bearing fault (case 2) has an accuracy rate of 100%, cavitation (case 3) scored 100%, impeller fault (case 4) has an accuracy rate of 100%, misalignment (case 5) has 100%, mechanical looseness (case 6) has 100% and imbalance (case 6) has shown an accuracy rate of 100%.



(a) The overall confusion matrices for classification accuracy rate



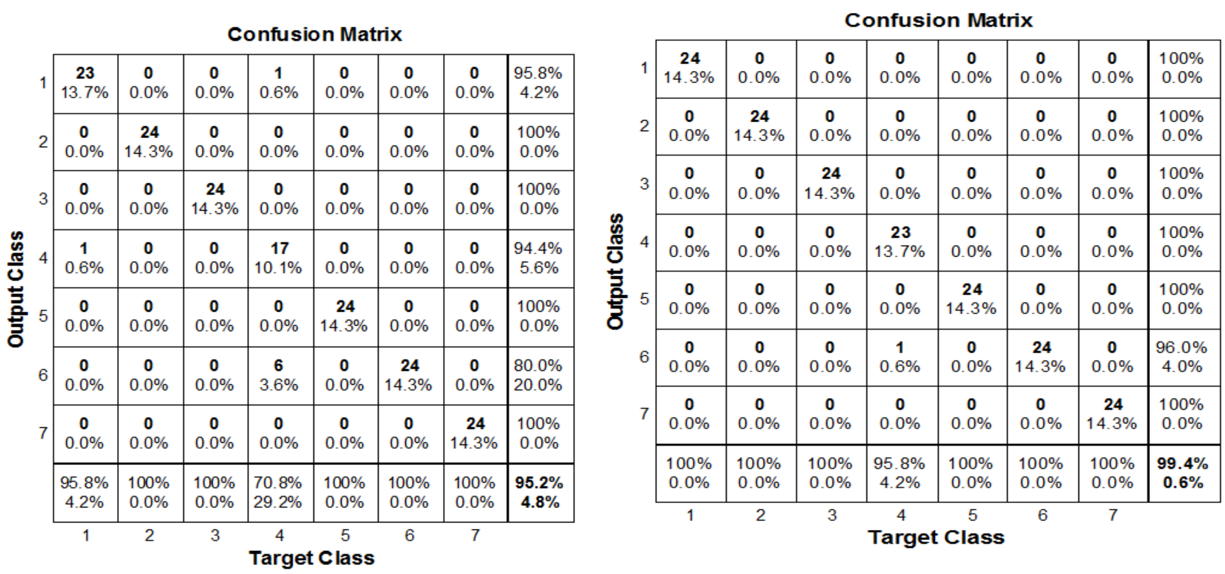
(b) Performance of MLP-BP using 24 normalized features (rbio1.5 mother wavelet)

Fig. 9 Validation of MLP-BP scheme

6.2. MLP-GABP

MLP-GABP is implemented for two cases of 14 and 24 approximation normalized features with db4 mother wavelet and also rbio1.5. MLP-GABP illustrated lower performance comparing MLP-BP in terms of computational time, overall classification accuracy rates, and number of classified cases, where using 14 features, overall rates of 100% for db4 and 98% for rbio1.5, but only 6 cases out of the 7 cases are classified with test and validation classifications using rbio1.5, and with test using db4. Although, using 24 normalized features, accuracy rates with rbio1.5 and db4 are presented an overall of 99.4% and 95.2% respectively, but only 6 cases out of 7 are classified in validation for both wavelets as shown in Fig. 10.

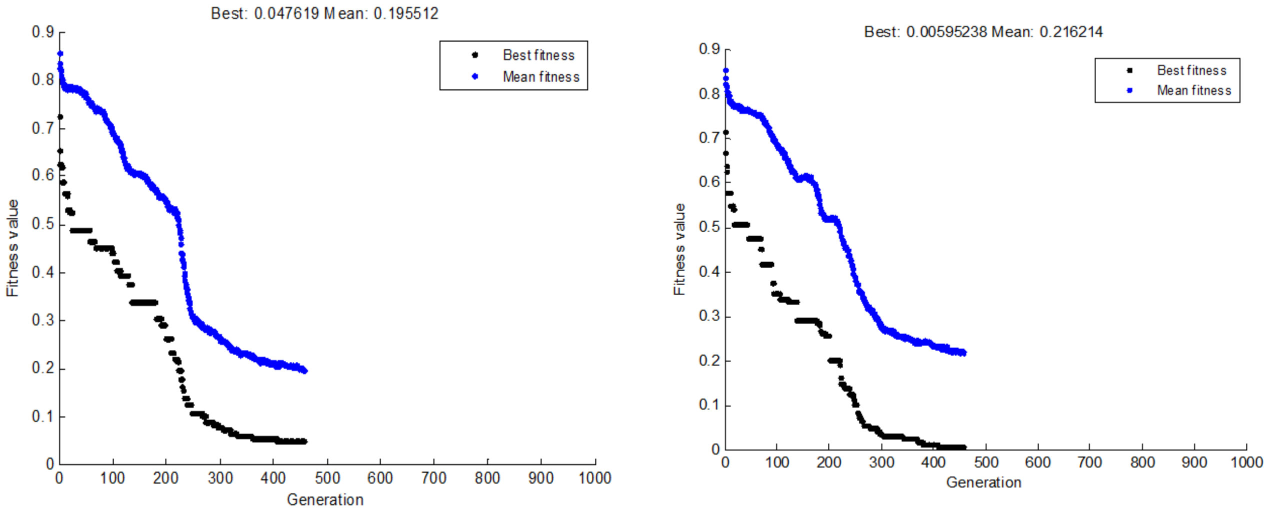
Four hidden layers containing [24 21 24 23] neurons are used in MLP as per as selection of GA and weights of neural network have been adjusted and selected using GA. In the case of using 14 features, GA based optimization and training using db4 and rbio1.5 are terminated after 576 and 403 generations with best fitness functions of 0.020482 and 0 respectively. Whereas, using 24 features, terminations with db4 and rbio1.5 are 457 and 548 and with best fitness functions of 0.047619 and 0.00595238 respectively as shown in Figure 11. The best fitness function denotes the best minimized mean square error (MSE).



(a) db4 function

(b) rbio1.5 function

Fig. 10 Overall classification accuracy rate using 24 normalized approximation features using GA based hidden neuron and layer selection with MLP-GABP (WPT)



(a) db4 function (b) rbio1.5 function

Fig. 11 Best score value and mean score versus generation based MLP-GABP with WPT using 24 normalized approximate features

6.3.SVM

Classification accuracy rates using polynomial kernel function of the cases of db4 and rbio1.5 mother wavelets (approximation features) 14 normalized, are 100% for both cases. Table 2 gives the overall performance of the AI methods employed in this work.

Table 2 Artificial Intelligence methods overall performance

Method	No. of features per condition and parameter	No. of hidden layers & neurons	Training Computational time (hh:mm:ss)	Test rate (%)	Validation rate (%)	Training rate (%)	Overall classification rate (%)	Remarks
SVM (rbio1.5)	14 approx. (normalized)	____	____	____	____	____	100	
SVM (db4)	14 approx. (normalized)	____	____	____	____	____	100	
MLP-BP with GA based selection (db4)	24 approx. (normalized)	[24 21 24 23]	00:01:29	100	100	100	100	
MLP-BP with GA based selection (rbio1.5)	24 approx. (normalized)	[24 21 24 23]	00:00:10	100	100	100	100	
MLP-BP with GA based selection (db4)	14 approx. (normalized)	[24 21 24 23]	00:00:13	100	100	100	100	6 classes are classified (test)
MLP-BP with GA based selection(rbio1.5)	14 approx. (normalized)	[24 21 24 23]	00:00:16	100	100	100	100	6 classes are classified (validation)
MLP-GABP with GA based selection (db4)	14 approx. (normalized)	[24 21 24 23]	MLP: 00:00:18 , GA: 08:00:00	100	100	100	100	6 classes are classified (test)
MLP-GABP with GA based selection (rbio1.5)	24 approx. (normalized)	[24 21 24 23]	MLP: 00:00:12 , GA: 13:00:00	100	100	99.2	99.4	6 classes are classified (validation)
MLP-BP with GA based selection (db4)	14 approx. (non-normalized)	[24 21 24 23]	00:00:10	100	86.7	100	98	6 classes are classified (test)
MLP-GABP with GA based selection (rbio1.5)	14 approx. (normalized)	[24 21 24 23]	MLP: 00:00:24 , GA: 15:00:00	100	100	97.1	98	6 classes are classified (test & validation)
MLP-BP with GA based selection(db4)	30 approx. (normalized)	[24 21 24 23]	00:00:20	100	96.9	97.3	97.6	
MLP-GABP with GA based selection (db4)	24 approx. (normalized)	[24 21 24 23]	MLP: 00:00:10 , GA: 12:00:00	96	92	95.8	95.2	6 classes are classified (validation)
MLP-BP with GA based selection (db4)	60 (normalized)	[24 21 24 23]	00:00:54	58.7	60.3	82.3	75.5	
MLP-BP with GA based selection (db4)	60 (non-normalized)	[24 21 24 23]	00:00:43	63.5	57.1	75.9	71.2	

7 Conclusion

The feature extraction and classification of the pump conditions using MLP-BP are conducted successfully for all 7 cases using WPT with normalized 60 and 30 features of 75.3% and 97.6% respectively. It has been observed that using 6 levels approximations only (30 features) rather than using all decomposed 3 levels approximations and details (60 features) provided a better classification rate. On the other hand, with reduced number of features, MLP-BP has successfully achieved training, validation of 100% for all the seven cases, but in test, 100% of accuracy is achieved for the classification of 6 cases out of the seven cases. It can be remarked that MLP-BP can loss its ability in classifying all the cases using insufficient features. Therefore, the number of features has to be carefully selected and reduced. However, it has been found that using more 3 level approximations (24 features) achieved an overall accuracy rate of 100% using both db4 and rbio1.5 for all 7 cases which outperformed the 6 level approximations (30 features). For the SVM, it has been also found that both db4 and rbio1.5 using reduced number of features (14 normalized features) resulted in an overall classification rate of 100%. Extracted features as approximations provided better classification rates than using all features of approximation and detail. The good selection of approximation features has a positive impact on the classification rate using all features. It has been remarked that MLP-BP and SVM performed better using normalized fewer parameters and features. GA has shown a good ability in optimizing and selecting the number of hidden layers and neurons, as the best performance was scored using 4 hidden layers containing 24, 21, 24 and 23 neurons respectively. However, GA needs longer computational time and the risk of getting stuck in a local minimum. On the other hand, GA along with BP based MLP training, presented a good performance but slightly lower comparing MLP-BP using 14 approximate features of 100% and 98% as an overall rate with db4 and rbio1.5 respectively with 6 classified cases. In addition, MLP-GABP with 24 approximate features classified all 7 cases with an accuracy rate of 99.4% and 95.2% using rbio1.5 and db4 respectively, but 6 cases only are classified with validation using both wavelets.

Finally, this work showed that MLP-BP classification accuracy can be improved if the neural network architecture is optimized using GA, a suitable mother wavelet for wavelet transform based feature extraction, good selection for the approximation features is achieved and with fewer number of features. It is also concluded that WPT with both MLP-BP and SVM produces better classification rates than CWT.

References

[1] AL-TUBI, M. A. S. & AL-RAHEEM, K. F. (2010) Rolling element bearing faults detection, a time domain analysis. Caledonian Journal of Engineering, 6.

[2] AL-TUBI, M. A. S., AL-RAHEEM, K. F. & ABDUL-KAREM, W. (2012) Rolling element bearing element faults detection, power spectrum and envelope analysis. International conference on applications and design in mechanical engineering. Penang, Malaysia.

[3] MEHALA, N. & DAHIYA, R. (2008) A Comparative Study of FFT, STFT and Wavelet Techniques for Induction Machine Fault Diagnostic Analysis. International conference of computational intelligence, man-machine systems cybernetics (CIMMACS '08). India.

[4] PENG, Z. K. & CHU, F. L. (2004) Application of the wavelet transform in machine condition monitoring and fault diagnostics: a review with bibliography. Mechanical Systems and Signal Processing, 18, 199-221.

[5] AL-TOBI, M. & AL-RAHEEM, K. F. (2015) Rolling element bearing faults detection, wavelet de-noising analysis. Universal journal of mechanical engineering, 5, 47-51.

[6] MURALIDHARAN, V., SUGUMARAN, V. & INDIRA, V. (2014) Fault diagnosis of monoblock centrifugal pump using SVM. Engineering Science and Technology, an International Journal, 17, 152e157.

[7] YAN, R., GAO, R. X. & CHEN, X. (2014) Wavelets for fault diagnosis of rotary machines: A review with applications. Signal Processing, 96, 1-15.

[8] FAROKHZAD, S. (2013) Vibration based fault detection of centrifugal pump by fast fourier transform and adaptive neuro-fuzzy inference system. Journal of mechanical engineering and technology, 1, 82-87.

[9] SAKTHIVEL, N. R., NAIR, B. B., ELANGOVAN, M., SUGUMARAN, V. & SARAVANMURUGAN, S. (2014) Comparison of dimensionality reduction techniques for the fault diagnosis of mono block centrifugal pump using vibration signals. Engineering Science and Technology, an International Journal 17, 30-38.

[10] SAKTHIVEL, N. R., BINOY.B.NAIR & V.SUGUMARAN (2012) Soft computing approach to fault diagnosis of centrifugal pump. Applied Soft Computing, 12, 1574–1581.

[11] LIU, L.-S., YANG, Y.-H., LI, Z.-Y. & YU, W. (2011) Condition Monitoring for Helicopter Main Gearbox Based on Wavelet Packet Transform and Wavelet Neural Network. International conference on quality, reliability, maintenance, and safety engineering. Xi'an, IEEE.

[12] YANJUN, L., FANLONG, M. & YIBO, L. (2009) Research on Rub impact Fault Diagnosis Method of Rotating Machinery Based on Wavelet Packet and Support Vector Machine International Conference on Measuring Technology and Mechatronics Automation. Changchun, IEEE.

[13] SAMANTA, B. (2004) Gear fault detection using artificial neural networks and support vector machines with genetic algorithms. Mechanical Systems and Signal Processing 18, 625–644.

[14] AL-RAHEEM, K. F., ROY, A., RAMACHANDRAN, K. P., HARRISON, D. K. & GRAINGER, S. (2008) Application of the Laplace- Wavelet Combined With ANN for Rolling Bearing Fault Diagnosis. Journal of Vibration and Acoustics, 130.

[15] YANG, Z., HOI, W. I. & ZHONG, J. (2011) Gearbox Fault Diagnosis based on Artificial Neural Network and Genetic Algorithms. International Conference on System Science and Engineering. Macau, China

[16] AL-RAHEEM, K. F. & ABDUL-KAREM, W. (2010) Rolling bearing fault diagnostics using artificial neural networks based on laplace wavelet analysis. International journal of Engineering, Science and Technology, 2, 278-290.

[17] ALTOBI, M. A. S., BEVAN, G., WALLACE, P., HARRISON, D. & RAMACHANDRAN, K. P. (2017) Faults Diagnosis of a Centrifugal Pump Using MLP-GABP and SVM with CWT Based Feature Extraction

[18] ALTOBI, M. A. S., BEVAN, G., WALLACE, P., HARRISON, D. & RAMACHANDRAN, K. P. (2017) Faults Diagnosis of a Centrifugal Pump Using MLP-GABP and SVM with DWT Based Feature Extraction.

[19] CHARNIAK, E. & MCDERMOTT, D. (2000) Introduction to artificial intelligence, addison wesley Longman Inc.

[20] MCCULLOCH, W. S. & PITTS, W. (1943) A logical calculus of the ideas immanent in nervous activity. BULLETIN OF MATHEMATICAL BIOPHYSICS, 5, 115-133.

[21] ZADEH, L. A. (1965) Fuzzy Sets. Information and Control, 8, 338-353.

[22] NASIRI, M. R., MAHJOOB, M. J. & VAHID-ALIZADEH, H. (2011) Vibartion signature analysis for detecting cavitation in centrifugal pump using neural networks. IEEE international conference on mechatronics (ICM). Istanbul, Turkey, IEEE.

[23] FAROKHZAD, S., AHMADI, H. & JAFARY, A. (2013) FAULT CLASSIFICATION OF CENTRIFUGAL WATER PUMP BASED ON DECISION TREE AND REGRESSION MODEL. Journal of Science and today's world, 2, 170-176.

[24] FAROKHZAD, S., AHMADI, H. & JAEFARI, A. (2012) Artificial Neural Network Based Classification of Faults in Centrifugal Water Pump. Journal VIBROENGINEERING 14.

[25] LIU, L.-S., YANG, Y.-H., LI, Z.-Y. & YU, W. (2011) Condition Monitoring for Helicopter Main Gearbox Based on Wavelet Packet Transform and Wavelet Neural Network. International conference on quality, reliability, maintenance, and safety engineering. Xi'an, IEEE.

[26] MURALIDHARAN, V. & SUGUMARAN, V. (2013) SELECTION OF DISCRETE WAVELETS FOR FAULT DIAGNOSIS OF MONOBLOCK CENTRIFUGAL PUMP USING THE J48 ALGORITHM. Applied Artificial Intelligence, 27.

[27] CHEN, H. W. P. (2009) Intelligent diagnosis method for a centrifugal pump using features of vibration signals. Neural Comput & Applic, 18, 397–405.

[28] WANG, H. Q. & CHEN, P. (2007) Fault Diagnosis of Centrifugal Pump Using Symptom Parameters in Frequency Domain. Agricultural Engineering International: the CIGR Ejournal, IX.

[29] MURALIDHARAN, V., SUGUMARAN, V., SHANMUGAM, P. & SIVANATHAN, K. (2010) Artifical neural network based classification for monoblock centrifugal pump using wavelet analysis. International journal of mechanical engineering, 1, 28-37.

[30] IIOTT, P. W. & GRIFFITHS, A. J. (1997) Fault diagnosis of pumping machinery using artificial neural networks. Journal of Process Mechanical Engineering, 211, 185-194.

[31] ZOUARI, R., SIEG-ZIEBA, S. & SIDAHMED, M. (2004) Fault detection system for centrifugal pumps using neural networks and neuro-fuzzy techniques. SURVEILLANCE 5 CETIM Senlis.

[32] NASIRI, M. R., MAHJOOB, M. J. & VAHID-ALIZADEH, H. (2011) Vibartion signature analysis for detecting cavitation in centrifugal pump using neural networks. IEEE international conference on mechatronics (ICM). Istanbul, Turkey, IEEE.

[33] ZHANG, Z., WANG, Y. & WANG, K. (2013) Intelligent fault diagnosis and prognosis approach for rotating machinery integrating wavelet transform, principal component analysis, and artificial neural networks. The international journal of advanced manufacturing technology, 68, 763-773.

[34] JEDLINSKI, Ł. & JONAK, J. (2015) Early fault detection in gearboxes based on support vector machines and multilayer perceptron with a continuous wavelet transform. Applied Soft Computing, 30, 636–641.

[35] WANG, D. Y., ZHANG, W. Z. & ZHANG, J. G. (2010) Fault Bearing Identification Based on Wavelet Packet Transform Technique and Artificial Neural Network. International conference on system science, engineering

design and manufacturing information, Yichang, IEEE.

[36] SUI, W.-T. & ZHANG, D. (2009) Rolling Element Bearings Fault Classification Based on SVM and Feature Evaluation. Proceedings of the Eighth International Conference on Machine Learning and Cybernetics. Baoding, IEEE.

[37] LIU, Y. & LIU, T. (2010) Rotating Machinery Fault Diagnosis Based on Support Vector Machine International Conference on Intelligent Computing and Cognitive Informatics. Kuala Lumpur, IEEE.

[38] SAMANTA, B., AL-BALUSHI, K. R. & S.A.AL-ARAIMI (2006) Artificial Neural Networks and Genetic Algorithm for Bearing Fault Detection. Soft Computing, 10, 264-271.

[39] RUMELHART, D. E., HINTON, G. E. & WILLIAM, R. J. (1986) Learning representations by back-propagating errors. Nature, 323.

[40] CORTES, C. & VAPNIK, V. (1995) Support-vector networks. Machine Learning, 20, 273-297.

[41] HOLLAND, J. H. (1975) Adaptive in natural and artificial systems, Ann Arbor, The university of Michigan press.

[42] ROOIJ, A. J. F. V., JAIN, L. C. & JOHNSON, R. P. (1996) Neural Network Training Using Genetic Algorithm, Singapore, World Scientific Publishing Co. Pte. Ltd.

[43] RAfiEE, J., TSE, P. W., HARIfi, A. & SADEGHI, M. H. (2009) A novel technique for selecting mother wavelet function using an intelligent fault diagnosis system. Expert Systems with Applications, 36, 4862–4875.

[44] SABERI, M., AZADEH, A., NOURMOHAMMADZADEH, A. & PAZHOHESHFAR, P. (2011) Comparing performance and robustness of SVM and ANN for fault diagnosis in a centrifugal pump. 19th International Congress on Modelling and Simulation. Perth, Australia.

[45] ALSBERG, B. K., WOODWARD, A. M. & KELL, D. B. (1997) An introduction to wavelet transforms for chemometricians: A time-frequency approach. Chemometrics and Intelligent Laboratory Systems, 37, 215-239.

[46] COIFMAN, R. R., MEYER, Y. & WICKERHAUSER, V. (1992) Size properties of wavelet packets IN RUSKAI, M. B. (Ed.) Wavelets and their applications. Boston, Jones and Bartlett.

[47] SAKTHIVEL, N. R., BINOY.B.NAIR & V.SUGUMARAN (2012) Soft computing approach to fault diagnosis of centrifugal pump. Applied Soft Computing, 12, 1574–1581.

[48] BENDJAMA, H., GHERFI, K., IDIOU, D. & BOUCHERIT, M. S. (2014) Condition monitoring of rotating machinery by vibration signal processing methods. International Conference on Industrial Engineering and Manufacturing. Batna University Algeria.

[49] ZHONG, J., YANG, Z. & WONG, S. F. (2010) Machine Condition Monitoring and Fault Diagnosis based on Support Vector Machine. International conference of Industrial Engineering Management Macao, IEEE.

[50] WUMING, H., PEILIANG, W. & QIANGGUO, Y. (2010) Fault Diagnosis of Traction Machine for Lifts Based on Wavelet packet algorithm and RBF Neural Network. International conference on information sciences. Chengdu, IEEE.

[51] CHEN, C.-H., SHYU, R.-J. & MA, C.-K. (2007) Rotating Machinery Diagnosis Using Wavelet Packets-Fractal Technology and Neural Networks. Journal of Mechanical Science and Technology, 21, 1058-1065.