

Faults diagnosis of a centrifugal pump using multilayer perceptron genetic algorithm back propagation and support vector machine with discrete wavelet transform based feature extraction

Al Tobi, M.; Bevan, Geraint; Wallace, Peter; Harrison, David; Okedo, Kenneth Eloghene

Published in:
Computational Intelligence

DOI:
[10.1111/coin.12390](https://doi.org/10.1111/coin.12390)

Publication date:
2021

Document Version
Author accepted manuscript

[Link to publication in ResearchOnline](#)

Citation for published version (Harvard):

Al Tobi, M, Bevan, G, Wallace, P, Harrison, D & Okedo, KE 2021, 'Faults diagnosis of a centrifugal pump using multilayer perceptron genetic algorithm back propagation and support vector machine with discrete wavelet transform based feature extraction', *Computational Intelligence*, vol. 37, no. 1, pp. 21-46.
<https://doi.org/10.1111/coin.12390>

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

Take down policy

If you believe that this document breaches copyright please view our takedown policy at <https://edshare.gcu.ac.uk/id/eprint/5179> for details of how to contact us.

Faults diagnosis of a centrifugal pump using multilayer perceptron genetic algorithm back propagation and support vector machine with discrete wavelet transform based feature extraction

Al Tobi, M.; Bevan, Geraint; Wallace, Peter; Harrison, David; Okedo, Kenneth Eloghene

Published in:
Computational Intelligence

DOI:
[10.1111/coin.12390](https://doi.org/10.1111/coin.12390)

Publication date:
2020

Document Version
Peer reviewed version

[Link to publication in ResearchOnline](#)

Citation for published version (Harvard):

Al Tobi, M, Bevan, G, Wallace, P, Harrison, D & Okedo, KE 2020, 'Faults diagnosis of a centrifugal pump using multilayer perceptron genetic algorithm back propagation and support vector machine with discrete wavelet transform based feature extraction', *Computational Intelligence*. <https://doi.org/10.1111/coin.12390>

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

Take down policy

If you believe that this document breaches copyright please view our takedown policy at <https://edshare.gcu.ac.uk/id/eprint/5179> for details of how to contact us.

Faults Diagnosis of a Centrifugal Pump using Multilayer Perceptron Genetic Algorithm Back Propagation and Support Vector Machine with Discrete Wavelet Transform Based Feature Extraction

M. Al Tobi^a, G. Bevan^b, P. Wallace^c, D. Harrison^d and K. E. Okedu^e

^a*Mechanical & Industrial Engineering, College of Engineering, National University of Science and Technology, Oman*

^{b,c,d}*School of Engineering and Built Environment, Glasgow Caledonian University, UK*

^e*Electrical and Computer Engineering, College of Engineering, National University of Science and Technology, Oman*

Abstract: This paper presents a comparative study of two artificial intelligent systems, namely; Multi-layer Perceptron (MLP) and Support Vector Machine (SVM), to classify six fault conditions and the normal (non-faulty) condition of a centrifugal pump. A hybrid training method for MLP is proposed for this work based on the combination of Back Propagation (BP) and Genetic Algorithm (GA). The two training algorithms are tested and compared separately as well. Features are extracted using Discrete Wavelet Transform (DWT), both approximations, details, and two mother wavelets were used to investigate their effectiveness on feature extraction. GA is also used to optimize the number of hidden layers and neurons of MLP. In this study, the feature extraction, GA based hidden layers, neurons selection, training algorithm, and classification performance, based on the strengths and weaknesses of each method, are discussed. From the results obtained, it is observed that the DWT with both MLP-BP and SVM produces better classification rates and performances.

Keywords: Back Propagation (BP), Discrete Wavelet Transform (DWT), Centrifugal pump, Genetic Algorithm (GA), Multilayer Perceptron (MLP)

1. Introduction

Rotating machines, including centrifugal pumps, are dynamic machines that produce vibrations. Such vibrations can be monitored and analyzed for the purpose of condition monitoring of the machine, and diagnosing of faults [1-4]. Fault detection using vibration analysis in any rotating machinery requires two main steps: data acquisition and data analysis. For data acquisition, an accelerometer is commonly used to measure vibrations [5-11]. In data analysis, three main analysis methods have been applied in time domain [6, 10, 12], frequency domain [9-12], and time-frequency domain [14-17]. Fast Fourier Transform (FFT) is a frequency domain algorithm that has been used successfully for extracting statistical parameters, as it is able to present physical characteristics of frequency domain data [18]. However, a centrifugal pump is a complex machine that consists of many different parts and its vibration signals contain many non-stationary

signals, for which FFT is not well suited [14-16, 19]. On the other hand, Wavelet Transform (WT) which analyses the time-frequency domain has been proposed as an alternate method for non-stationary signals [16, 20].

Artificial Intelligent (AI) systems have been applied to centrifugal pumps as automatic fault diagnosis and classification systems in references [10, 21-23]. Automatic fault classification has been implemented and investigated using many machinery fault diagnostic methods, in order to improve precision and reduce mistakes caused due to human misinterpretation [10]. Previously, Discrete Wavelet Transform (DWT) has been applied by many researchers for applications of rotating machinery, accountability diagnosis combined with AI systems [7, 24-26]. Srinivas et al. in reference [27] proposed MLP and DWT based on Daubechies wavelet function to diagnose faults on rotor unbalance and shaft bent. The results showed the following classification success rates: unbalance at 99.78%, shaft bent at 99.81%, and combined faults of unbalance with shaft bent at 99.45%. Muralidharan and Sugumaran in reference [5] presented the application of the decision tree (J48 algorithm) in fault diagnosis of centrifugal pump with the features that were extracted using DWT families. In the literature, different families of DWT have been applied for fault classifications and it was observed that reverse biorthogonal wavelet (rbio1.5), is an appropriate scheme, as the best accuracy rate is 99.84%. DWT is tested with db4 and rbio1.5 wavelet functions for their popularity and effectiveness [28, 29] to investigate their impact on the classifiers' performance.

Genetic Algorithm (GA) has been used for the selection of input features for machinery components like bearings and gears [30-32] and the number of neurons in the hidden layers [30-31 and 33] of Multilayer Perceptron Artificial Neural Network (MLP-ANN). Application of GA based number of hidden layers and neurons selection for the centrifugal pump have been applied using the traditional Continuous Wavelet Transform (CWT) based feature extraction with the best classification rates of 99.5% using (Multilayer Perceptron Back Propagation (MLP-BP) and 88.5% using (Multilayer Perceptron Genetic Algorithm (MLP-GA) combined with Back Propagation (BP) based training [34].

This paper investigates the classification performance of two artificial intelligent methods: MLP-BP along with GA based selection and Support Vector Machine (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 DWT, where both the approximation and detail coefficients are extracted separately and then classification using the two forms is compared. Two mother wavelet functions are selected and tested with DWT, 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 intelligent 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 data sets used for the evaluation under GA are the same as those used for the MLP training. The performance of the system 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 intelligent systems including MLP-NN, SVM classifiers and GA. Section 3 illustrates the experimental setup. Section 4 outlines the method applied and procedures using DWT for feature extraction, while, section 5 presents the results and discussions. Finally, a conclusion with remarks and recommendations is given in section 6.

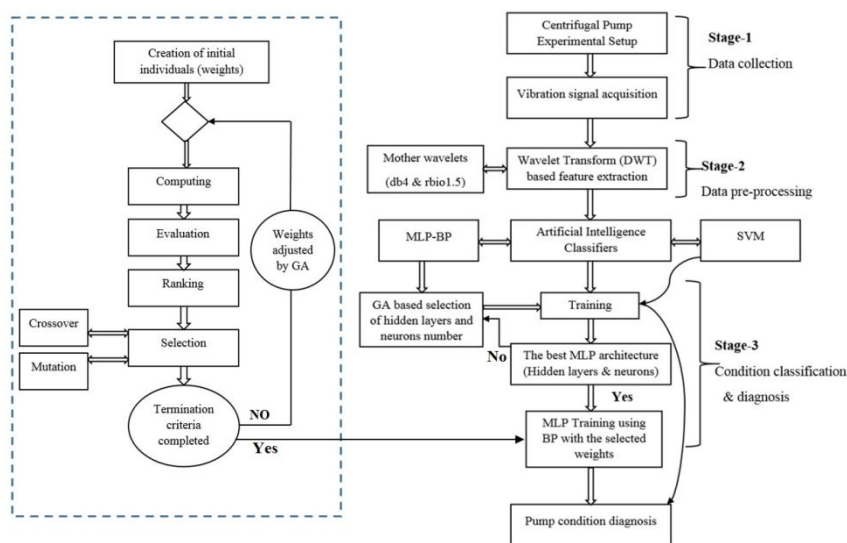


Fig. 1 Flow chart of diagnosis methods and training algorithm

2. Artificial Intelligent Systems

Automatic fault detection methods make use of AI which seeks to replicate mental capabilities with support of computational systems [35]. An Artificial Neural Network (ANN) was first introduced by McCulloch and Pitts [36], and Fuzzy Logic was first introduced by Zadeh [37]. ILott and Griffiths [38] proposed ANN with BP algorithm to diagnose pump faults. Then, Zouari et al. [39] 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 in the literature. Some of these are Back Propagation-Artificial Neural Network (BP-ANN) or MLP [30-33, 40-46] and Support Vector Machine (SVM) [30, 42, 44-47].

This section briefly discusses the basic principle and implementation of the MLP, SVM classifiers and GA based optimization and training.

A. Multilayer perceptron with back propagation

Multilayer perceptron 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 general behaviour 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 of the 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 [31]. 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.

Back propagation is a training algorithm that was developed by Rumelhart [48] and it has become a popular training method for MLP. The basic training principle of BP is based on the gradient decent method

that adjusts the weights to minimise Mean Square Error (MSE). The training process of BP can be initiated by multiplying the input vectors with the weights, as the biases and weights are summed in order to calculate the actual outputs. The desired outputs have to be determined and then compared with the actual outputs, continuing evaluation and weight modification until the process approaches the desired MSE value. MSE is also known as training or network error and represented mathematically as:

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

Where n is the total number of inputs, i is known as the index of summation, t_i is the desired (target) output from the output layer, and y_i is the actual output in the output layer.

Comparative studies have demonstrated the efficiency of MLP over other ANN types [31, 33]. However, a drawback of MLP is that, it is slow in training and needs longer computational time than other methods [31, 33, 43, 46]; but such weakness can be minimized by reducing the number of input features [40, 41].

B. Support Vector Machine

Support Vector Machine (SVM) was initially introduced by Corinna Cortes and Vladimir Vapnik [49] 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 working principle is based on separating (classify) two different classes (conditions); class A and class B as shown in Fig.2. From Fig. 2, the optimal hyper-plane can separate the two classes (A and B) with a wider margin compared to the non-optimal hyper-plane. The support vectors of both classes (A “+” & B “*”) indicates the margins. The optimal hyper plane (separator) is separating the two classes with a maximum width (margin), as the larger margin (width) between the two classes gives more generalization and eventually better linear classification. The support vectors indicate the margins and they are located beside the classes’ boundaries, where they have all the essential information of the classification. The linear classifier (hyper plane) is expressed as:

$$W^T X + b = 0 \quad (2)$$

If class A is assumed to be above the hyper plane, then it is > 0 , indicated as +1 and given by

$$W^T X + b = +1 \quad (3)$$

and if class B is < 0 , indicated as -1 and given by

$$W^T X + b = -1 \quad (4)$$

where W is the weight vector, X is the input and b is the bias.

A decision function can be used to separate two different classes (i.e. A and B) and given by

$$f(f(x) = \text{sign}((W.X) + b) \quad (5)$$

The decision function can be given by

$$f(f(x) = \text{sign}\left(\sum_{i=1}^L v_i(X.X_i) + b\right) \quad (6)$$

Where L is the number of training data and v_i is applied as weighting factor to identify the suitable support vectors from the given inputs.

However, it is not always guaranteed that linear boundary is able to classify the two different classes. Therefore, in order to classify the two classes with better margin, SVM can map the non-linear training data into a higher dimension level which is known as the feature space s using a transformation $\Phi(X)$, with s given by [30]

$$s = \Phi(X) \quad (7)$$

Where $X \in R^N$ and $s \in R^Q$. Based on the above, the decision function can be defined as

$$f(f(x) = \text{sign}\left(\sum_{i=1}^L v_i(\Phi(X) \cdot \Phi(X_i)) + b\right) \quad (8)$$

Moreover, to provide such transformation into non-linear classification, a kernel function is used $k(X.Y)$ which is given by

$$k(k(X.Y) = \Phi(X) \cdot \Phi(Y) \quad (9)$$

where k indicates the kernel function.

The decision function for the non-linear classification can be given by

$$f(x) = \text{sign}(\sum_{i=1}^L v_i k(X, X_i) + b) \quad (10)$$

There are different kernel functions which can be employed with SVM for the purpose of non-linear classification, and such kernel functions like;

Polynomial kernel function is given by

$$k(X, Y) = k(X, Y)^d \quad (11)$$

Where d indicates the dimension.

The Radial Basis Function (RBF) is given by

$$k(X, Y) = \exp\left(-\frac{\|X - Y\|^2}{2\sigma^2}\right) \quad (12)$$

Where σ the width parameter of RBF kernel function, and the Sigmoid is given by

$$k(X, Y) = \tanh(\kappa(X, Y) + \Theta) \quad (13)$$

Where κ is known as gain parameter and Θ is offset parameter of sigmoid kernel function.

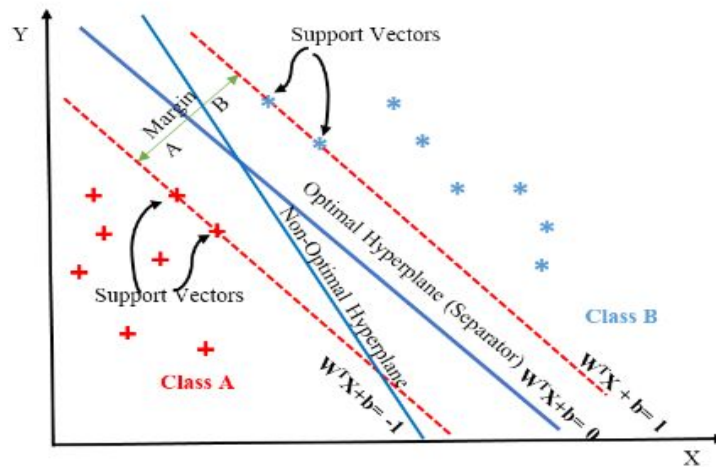


Fig. 2 Working principle of support vector machine

SVM has been widely recommended for rotating machinery fault diagnosis as it has proved its high efficiency and outstanding performance over other AI classifiers like MLP (ANN-BP) [27, 31, 40, 42, 45] and RBF [42].

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 the development of the codes. 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.

C. Genetic Algorithm

Genetic Algorithm was introduced by John Holland [50]. 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 [51]. GA starts its process by initiating individual populations which are known as chromosomes where they 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 [51].

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. The parameters of the GA based training and optimization are illustrated in Table 1.

Table 1 Genetic Algorithm based training and optimization architecture parameters

GA based training			
No. of generations	Size of population	TolFun	Mutation & crossover options
1000	1000	1E-60	Mutation-Gaussian Crossover-scattered
GA based optimization			
No. of generations	Size of population	TolFun	Mutation & crossover options
40	10	1E-60	Mutation-Gaussian Crossover-scattered

3. GA and GABP-Artificial Neural Network Training Methods

GA is used to train MLP along with BP by optimising the connection weights of MLP, and there are number of points to be considered while using GA in training MLP:

- Coding: the main purpose of coding is to encoding the weights into chromosomes, and there are two methods to encode the weights and threshold of a neural network. First one is binary encoding which presents each weight by coding them into fixed length of 1 and 0 string. There is a drawback using this method that the weights are within a predefined and fixed range of string. The second method is known as a real encoding which has shown better ability than the first methods [52, 53], as it presents weights with a real number in which the precision is increased, compared to the binary method. It also has more flexibility in using various operators (crossover and mutation) methods.
- Fitness function: fitness function of each chromosome has to be identified during the processing of GA. Using the real coding method, each individual weight of the neural network has to be restored from the chromosome. Then the neural network is totally identified, and the outputs to each of the test set is calculated and compared to the targets. Hence, the fitness value can be calculated through the training error on the test set. A low calculated error means better performance, and eventually a higher fitness value.
- Finding the GA-parameters and operators: in GA, there are number of parameters to be identified prior to optimisation. Such parameters like the population size, fitness normalisation, and selection operator (reproduction). In selection, two main operators are used; crossover and mutation to reproduce new populations for the next generation by selecting two chromosomes randomly and consider them as parents.

- Crossover: there are two chromosomes which act as parents, and then other chromosomes are produced from the parents. Therefore, the main purpose of crossover operator is to find and reproduce new chromosomes for the next generation. The number of chromosomes to be exchanged depends on the given crossover probability which is usually given in the range of 0.5 to 1 [54]. There are number of crossover operators like: single point crossover which is identified randomly by number like 1,1L, where 1L represents the length of chromosome, N-point crossover, and uniform crossover.
- Mutation: this is an essential operator that allows new genetic material to be added to a population. A mutation probability can be varied in the range of 0.001 to 0.01, which controls the random mutation for the chromosomes. The main purpose of mutation is to give a limited random search to provide a possibility of no zero probability during the random search.

4. Neural Network Training with Forward Propagation

The calculation of error determines if the calculated output is acceptable or not which eventually decides the neural network training completion. Therefore, the error cannot be sufficiently minimised, and it has to be further minimised by adjusting the weights which can be done using a traditional method which is also known as Forward Propagation (FP) given by:

$$W_{i(new)} = W_{i(old)} + \varepsilon[t_i - y_i]X_i \quad (14)$$

Where $W_{i(new)}$ represents the new adjusted weights, $W_{i(old)}$ represents the current weights, and ε is the learning rate. The steps of FP are shown in Figure 3.

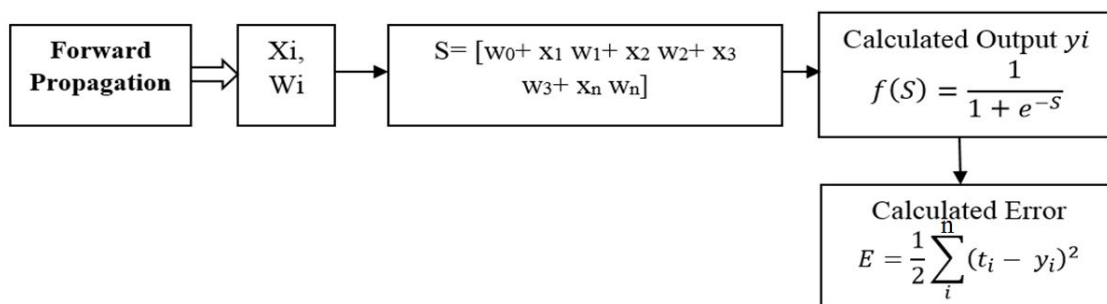


Fig. 3 Neural network training with Forward Propagation

However, FP has a drawback of being not able to give an accurate prediction and justification on why the updated weights can be better than the old ones, and on the impact of each weight over the calculated error. Hence, using Back Propagation (BP) provides more ability to understand the impact of weights over the error; to control the adjustment of the weights as shown in Figure 4.

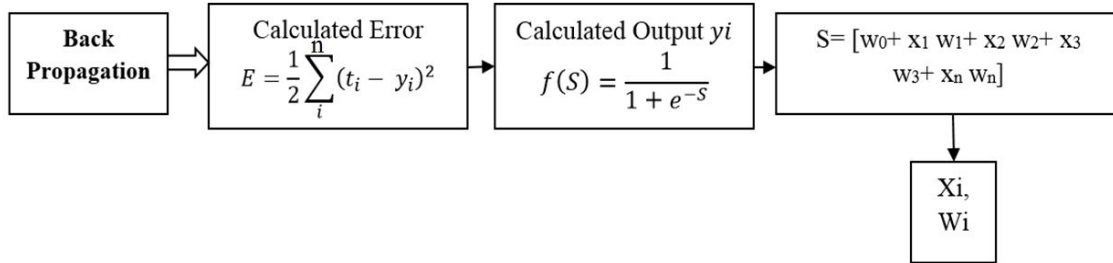


Fig. 4 Neural network training with Back Propagation by taking the reverse procedures of Forward Propagation method

From Equation (1), the derivative of the error shows how BP adjusts the weights which is given by [53]:

$$\Delta W_{ji} = -\varepsilon \cdot \frac{\partial E}{\partial W_{ji}} \quad (15)$$

momentum term μ is introduced into the training Equation (15), in order to improve the training process, and it is given by [53]:

$$\Delta W_{ji}(t+1) = -\varepsilon \cdot \frac{\partial E}{\partial W_{ji}}(t+1) + \mu \cdot \Delta W_{ji}(t) \quad (16)$$

Where W_{ji} = weight from the j^{th} unit to the i^{th} unit.

It is also possible to apply another training rule-based BP to adjust the weights and this is given by [53]:

$$\Delta W_{ji}(t+1) = -(1-\mu)\varepsilon \cdot \frac{\partial E}{\partial W_{ji}}(t+1) + \mu \cdot \Delta W_{ji}(t) \quad (17)$$

Where the factor $(1-\mu)$ is introduced to avoid any termination for the learning rate ε as the momentum term μ is increased.

5. Experimental Setup for the Study

The centrifugal pump experiment has been designed and assembled 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/min), control panel with speed controller (Schneider model Variable Frequency Drive (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 with 4 input channels. The accelerometer model is IMI 621B40 with sensitivity of 10 mV/g and frequency range from 3.4 to 18 kHz for ($\pm 10\%$) and 1.6 to 30 kHz for (± 3 dB). Figure 5 shows the centrifugal pump experimental setup.

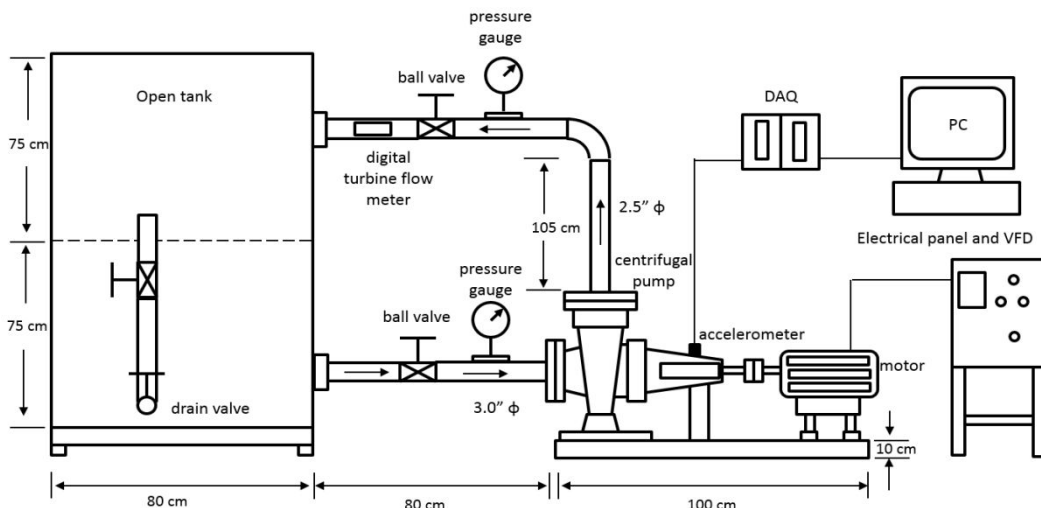


Fig. 5 Model system for experimental test rig

The vibration signals are measured under healthy and faulty conditions. Firstly, the signal of normal condition is acquired when the pump is healthy, without any faults. The faulty conditions are divided into two main categories: five mechanical faults (bearing, misalignment, unbalance, impeller, and looseness), and a 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. A DAQ is used to read the signals, where the signals are amplified and noise filtered out before digitization and filtering is used with a

bandwidth of 2.5 kHz; and then transmitted to a computer which is equipped with a digital/analogue converter card (D/A) in order to convert the analogue signals to digital. Data are acquired for a period of 2.4s at a sampling rate of 16 kHz, resulting in acquisition of 38400 samples. Averaging is applied with 10 as number of averages. 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 at a speed of 20 Hz (1200 RPM).

6. 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 [55]. A reduced number of features can result in higher rates of successful classification when only relevant and non-redundant features are used [56, 57]; and also helps to reduce memory usage, network size and processing time [58]. Application of a feature selection algorithm specifically to reduce the number of features has indeed been observed to achieve higher accuracy rates compared to non-reduced feature sets [59, 60].

Wang and Chen [7] 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 [7]. This difficulty has driven researchers to apply wavelet transform analysis in order to perform noise cancellation for the feature extraction [61]. Reference [62] proposed CWT for the feature extraction from the vibration signals of a centrifugal pump.

DWT was first introduced and developed by Mallat [63], where DWT is the discretization of CWT [16] which is given by:

$$DWT(j,k) = 1/(\sqrt{2^j}) \int_{-\infty}^{\infty} s(t) \varphi^* ((t - 2^j k)/2^j) dt \quad (18)$$

Where j and k are integers that represent scale and translation processes respectively, and 2^j and 2^{jk} are scale and time parameters. DWT also differs from CWT in terms of selection of values for scale (a) and

time (b) parameters; there are no constraints in selection with CWT [64], but with DWT, restrictions apply for the selection of (a, b) parameters [65].

In this work, DWT is applied with decomposition to 6 levels for the feature extraction, where the approximation and detail coefficients are extracted from DWT from seven different pump cases. In each case, a signal of length 34800 samples was recorded. These signals were each divided into 10 segments, of length 3840 samples. From each of the 10 segments, a signal is decomposed using DWT into 6 levels and produced 60 approximation features and 60 detail 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. Feature extraction is implemented once with db4 mother wavelet and also with rbio1.5 function. Figure 6 (a-d), shows the detailed (left) and approximated (right) sixth levels of decomposed signals including the original signal of each condition using 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), as this prominently illustrated in signals like imbalance, cavitation and misalignment. It is also remarked that approximation reveals successively less noisy signals by reducing the high-frequency information which could be resulting in extracting better features than the ones from detail. The main methodology of DWT in this work is depicted in Figure 7. The 60 features of both approximation and detail are used separately to train the MLP-BP. In addition, 30 approximation features are also considered to train the MLP-BP.

For the SVM, two cases are considered; first using 2 parameters and 60 features; and then with the number of features reduced to just 30. The effectiveness (sensitivity) of each parameter against all conditions are plotted in Figure 8. Normally, when healthy (blue +) is the lowest, it indicates good effectiveness of the parameter. Therefore, peak and RMS are selected for SVM due to their ability in distributing and distinguishing the conditions effectively.

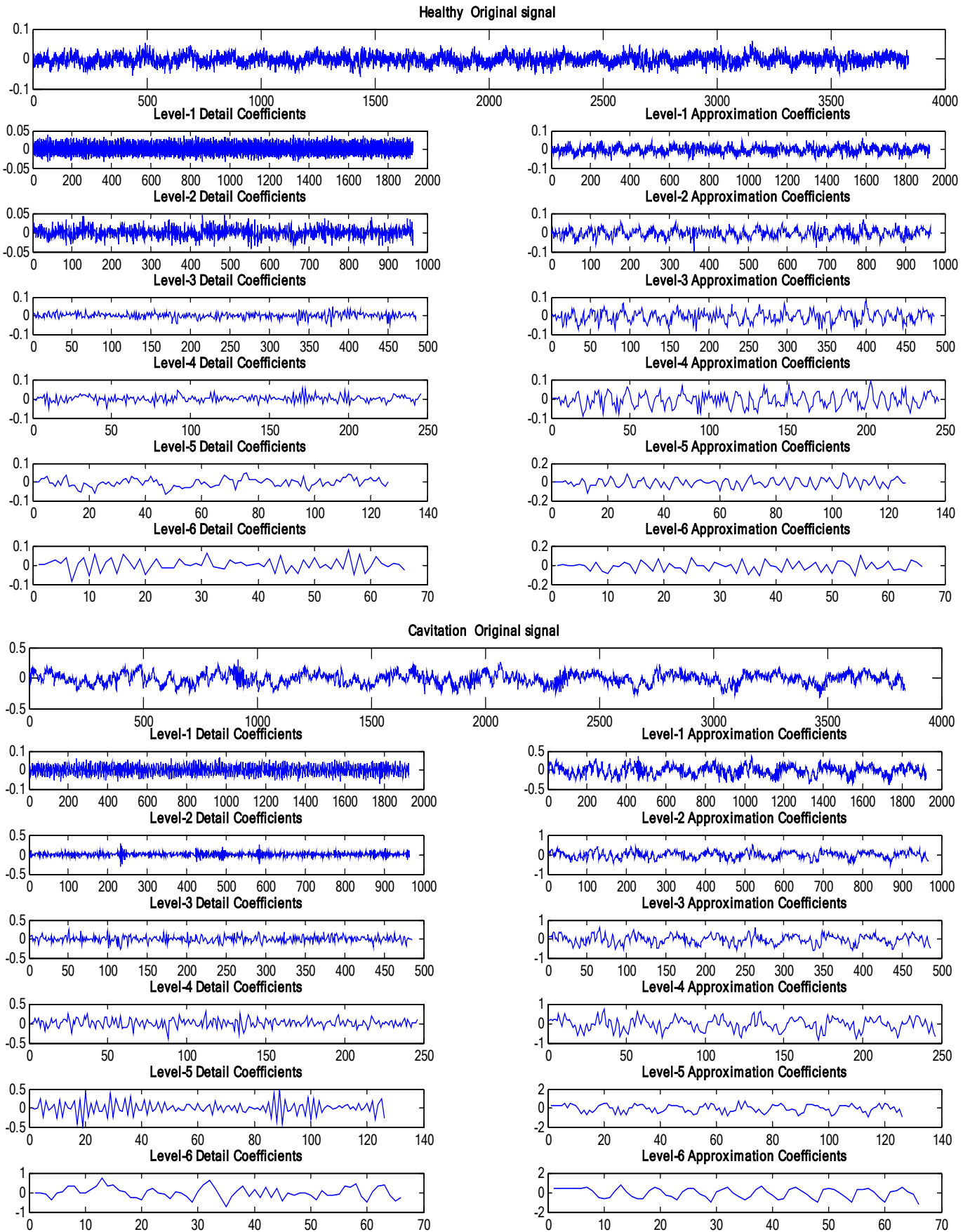


Fig. 6a Detail and approximation decomposed 6 level signals of the 7 pump conditions (db4 mother wavelet)

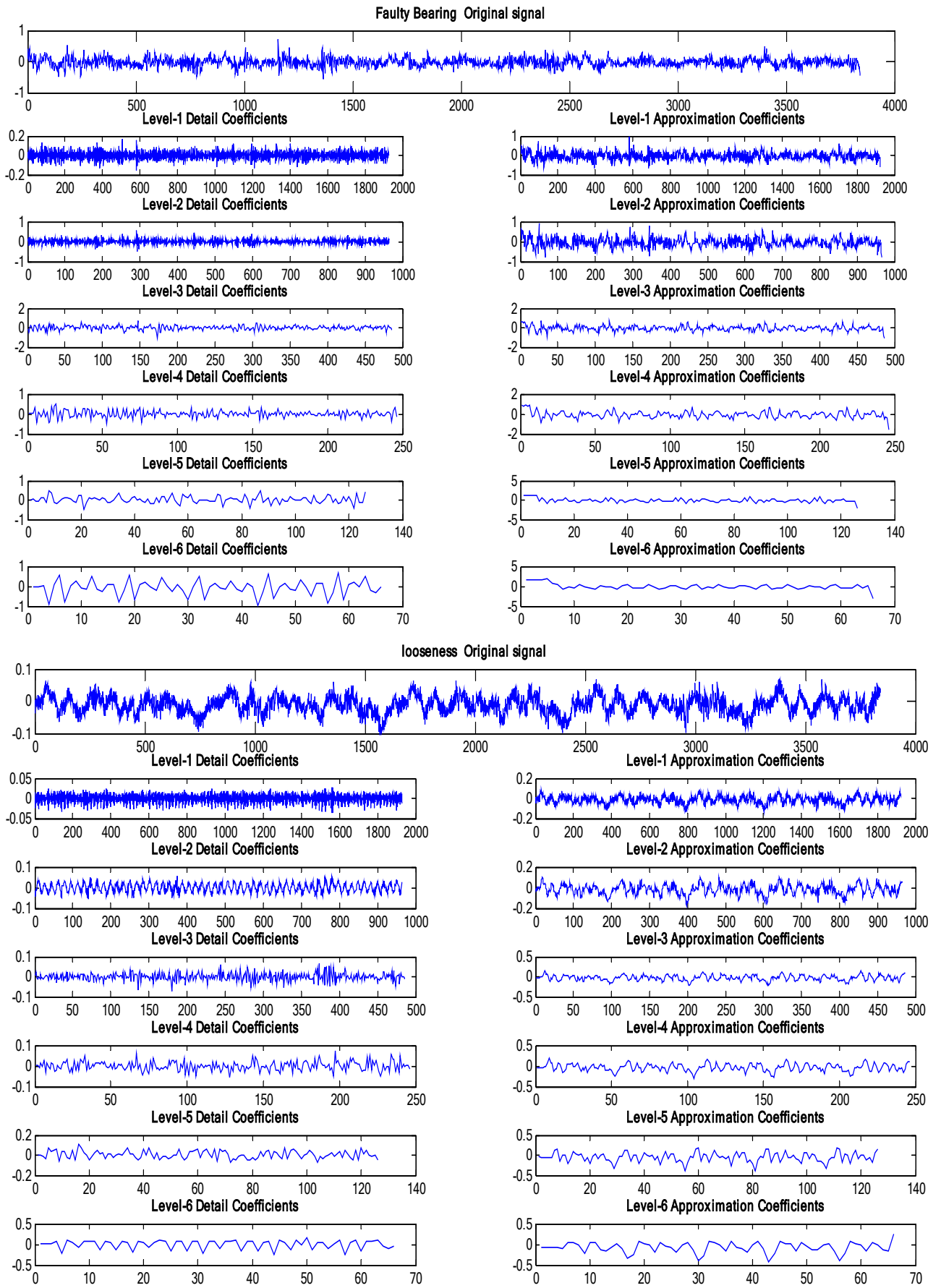


Fig. 6b Detail and approximation decomposed 6 level signals of the 7 pump conditions (db4 mother wavelet)

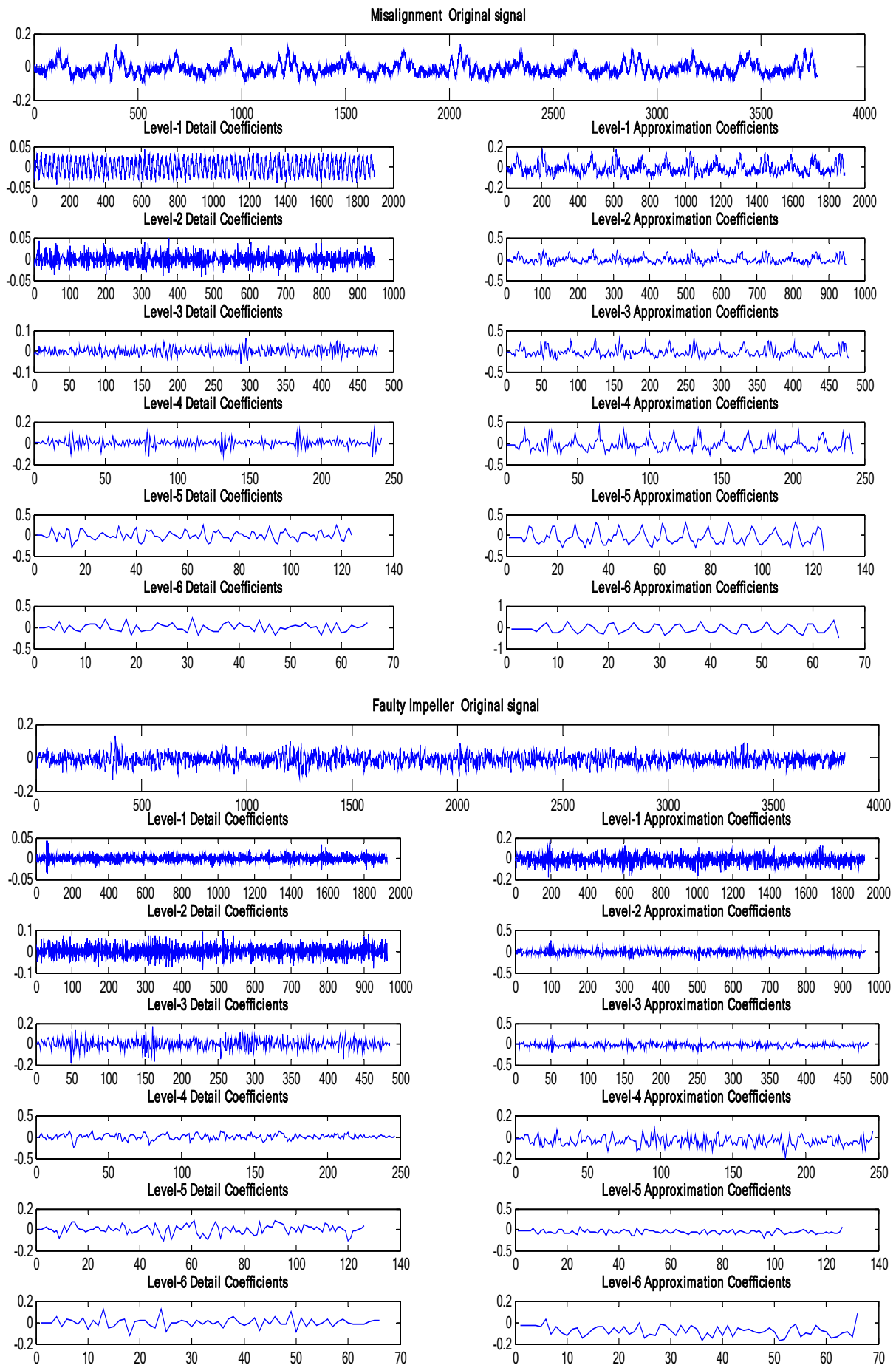


Fig. 6c Detail and approximation decomposed 6 level signals of the 7 pump conditions (db4 mother wavelet) 17

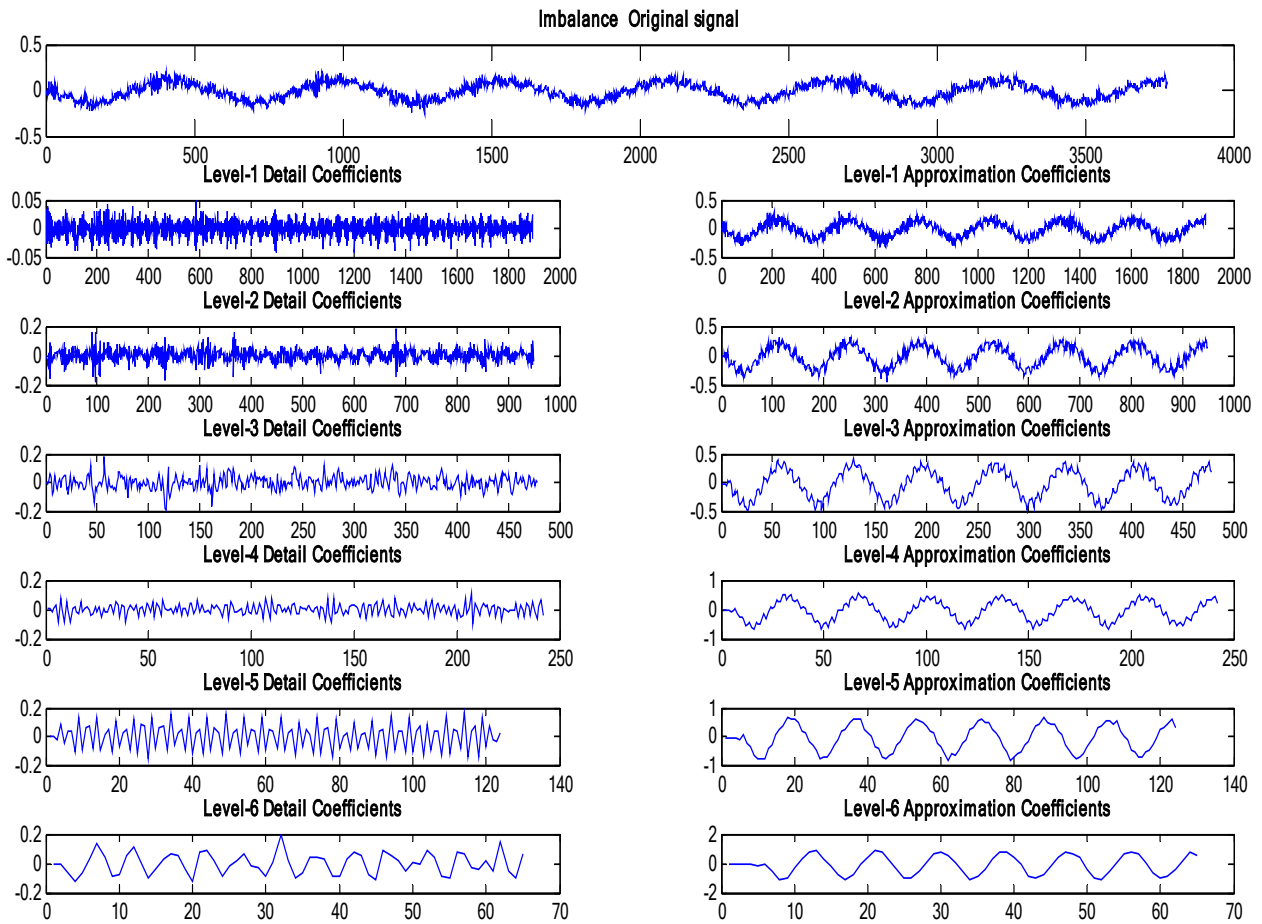


Fig. 6d Detail and approximation decomposed 6 level signals of the 7 pump conditions (db4 mother wavelet)

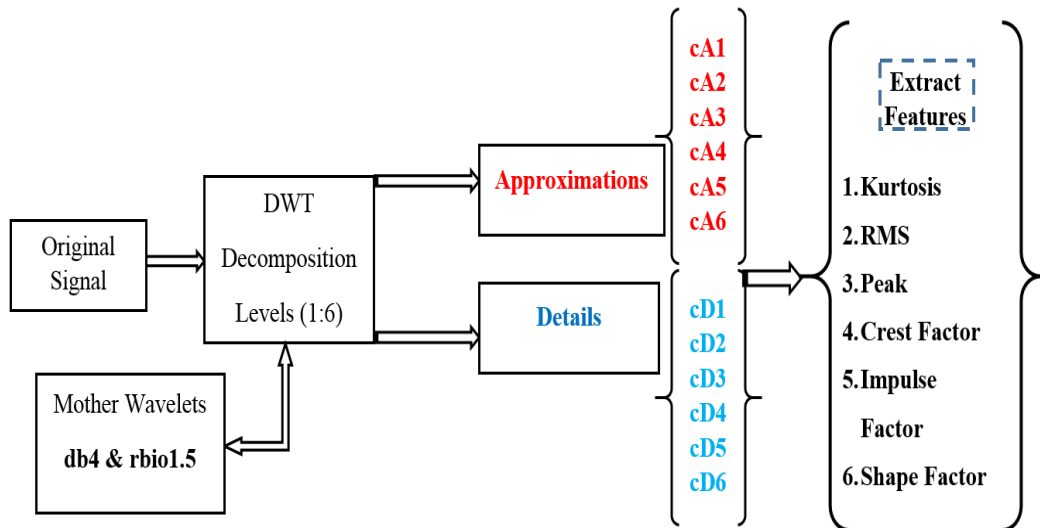


Fig. 7 DWT working processes for feature extraction based on decomposition procedure (6 levels) to produce two type of coefficients, namely, approximation (low frequency), and detail (high frequency). Features are extracted as approximations and also details using 6 different parameters.

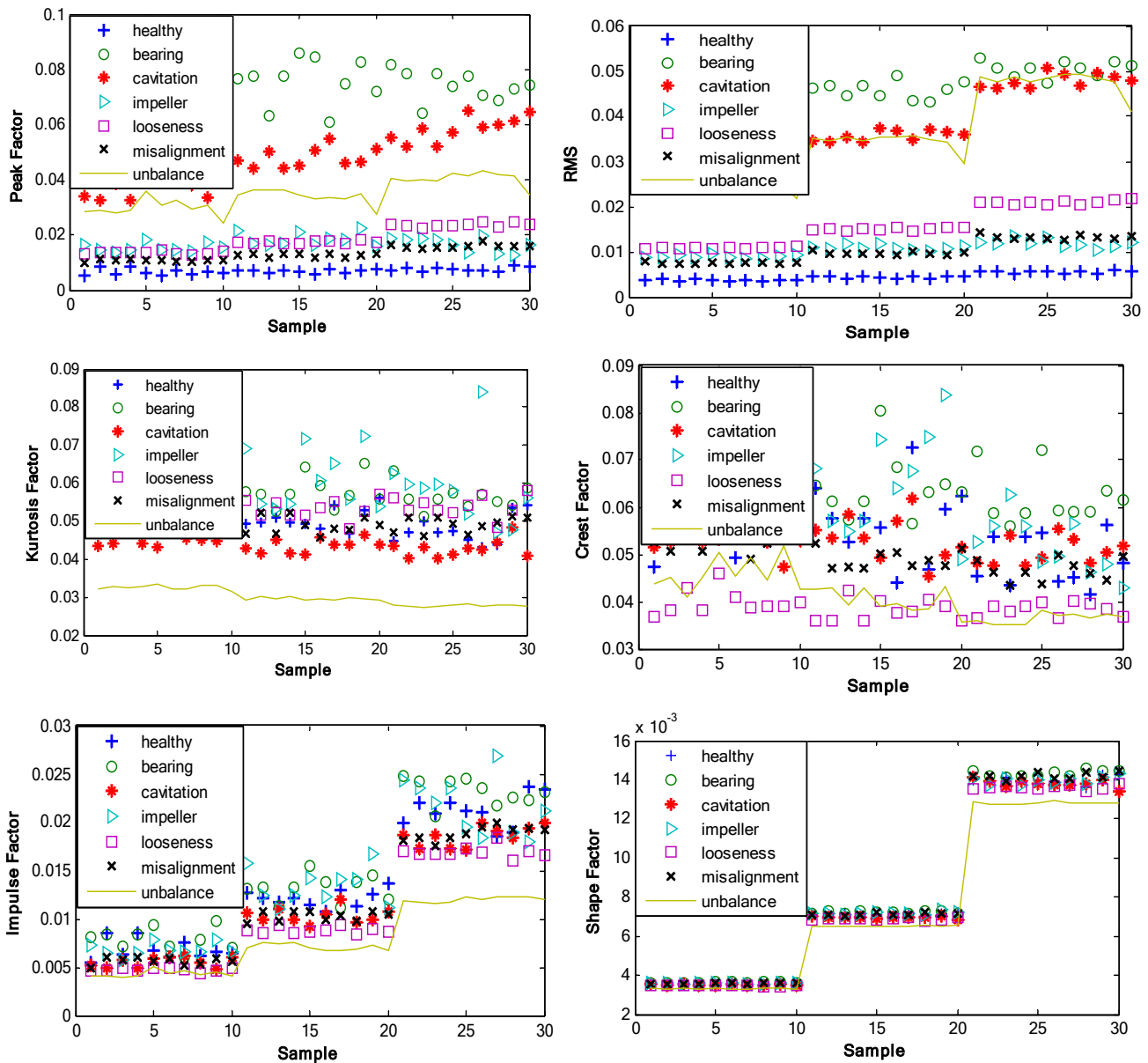


Fig. 8 The effectiveness of each parameter against all conditions (db4 mother function). RMS and Peak parameters show better distribution of the samples (features) compared to the other parameters.

7. 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 DWT. The number of hidden layers and neurons is 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. Two cases are considered, 60 and 30 features (normalized and non-normalized) per condition with a total of 420 and 210 input features for all conditions per parameter and forwarded to the MLP-ANN which results in a matrix of size $[6 \times 420]$ and $[6 \times 210]$ respectively. The input vectors are divided into three datasets (training has 70%, test has 15% and validation has 15% [66]. The target for training was a Boolean matrix of size $[7 \times 420]$ (60 features) and $[7 \times 210]$ (30 features); with the rows corresponding to the 7 conditions: (cases) in which seven digits-coding and each digit represents a block of size (1x60) and (1x30) given as:

Healthy [1 0 0 0 0 0 0], Bearing fault [0 1 0 0 0 0 0], Cavitation [0 0 1 0 0 0 0], Impeller fault [0 0 0 1 0 0 0], Misalignment [0 0 0 0 1 0 0], Looseness [0 0 0 0 0 1 0], Imbalance [0 0 0 0 0 0 1].

The neural network details and structure are shown in Table 2.

Table 2 Multilayer Feed-Forward Perceptron neural network architecture

Transfer Function		MLP-ANN details		
		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 1000	MSE 10E-20

SVM classification has been applied using two cases which are 60 features and 30 (normalized and non-normalized) ones with approximation features and again with detail features. For the two cases, all the seven conditions were tested and classified against each other, and all results are compared.

The SVM has been investigated using 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 be most effective for distinguishing between the conditions.

8. Results and Discussions

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 and using normalized or non-normalized features. Alternatives and improvements are proposed to enhance the classification accuracy rates. The results obtained using these three AI methods are discussed as follows.

A. MLP-BP

Classification rates are remarked for all cases with db4 mother wavelet function using GA based selection which suggested four hidden layers containing [24 21 24 23] neurons. The overall classification rates using 60 normalized approximation features of 99.8%, 60 non-normalized approximation features of 99.5%, 60 normalized detail features of 92.1%, and 60 non-normalized detail features of 85%. Therefore, the best accuracy rate is achieved using 60 normalized approximation features, as the overall confusion matrix and best performance graph for training, testing, and validation are illustrated in Figure 9.

The lower right blue square shows the overall classification rates, where overall, 99.8% of the classifications are correct and 0.2% are incorrect classifications. Based on the accuracy rate of each class (pump conditions) the following results were observed; healthy (case 1) 100%, bearing fault (case 2) 100%,

cavitation (case 3) 98.4%, impeller fault (case 4) 96.8%, misalignment (case 5) 100%, mechanical looseness (case 6) 100% and imbalance (case 6) 100%. On the other hand, classification rates with rbio1.5 mother functions, using 60 normalized approximation features presented the same overall classification accuracy rate as db4 of 99.8%. Therefore, using the 60 approximation normalized features based db4 and rbio1.5 mother wavelets would give the same overall performance rate. However, the overall test performance rates using the two mother wavelets would achieve 100% with rbio1.5 and 98.4% with db4 as illustrated in Figure 10.

Although, with the use of 30 normalized approximate features, the overall classification rates present better performance compared to using 60 features, where MLP-BP using both wavelet functions of db4 and rbio1.5 performed 100%.

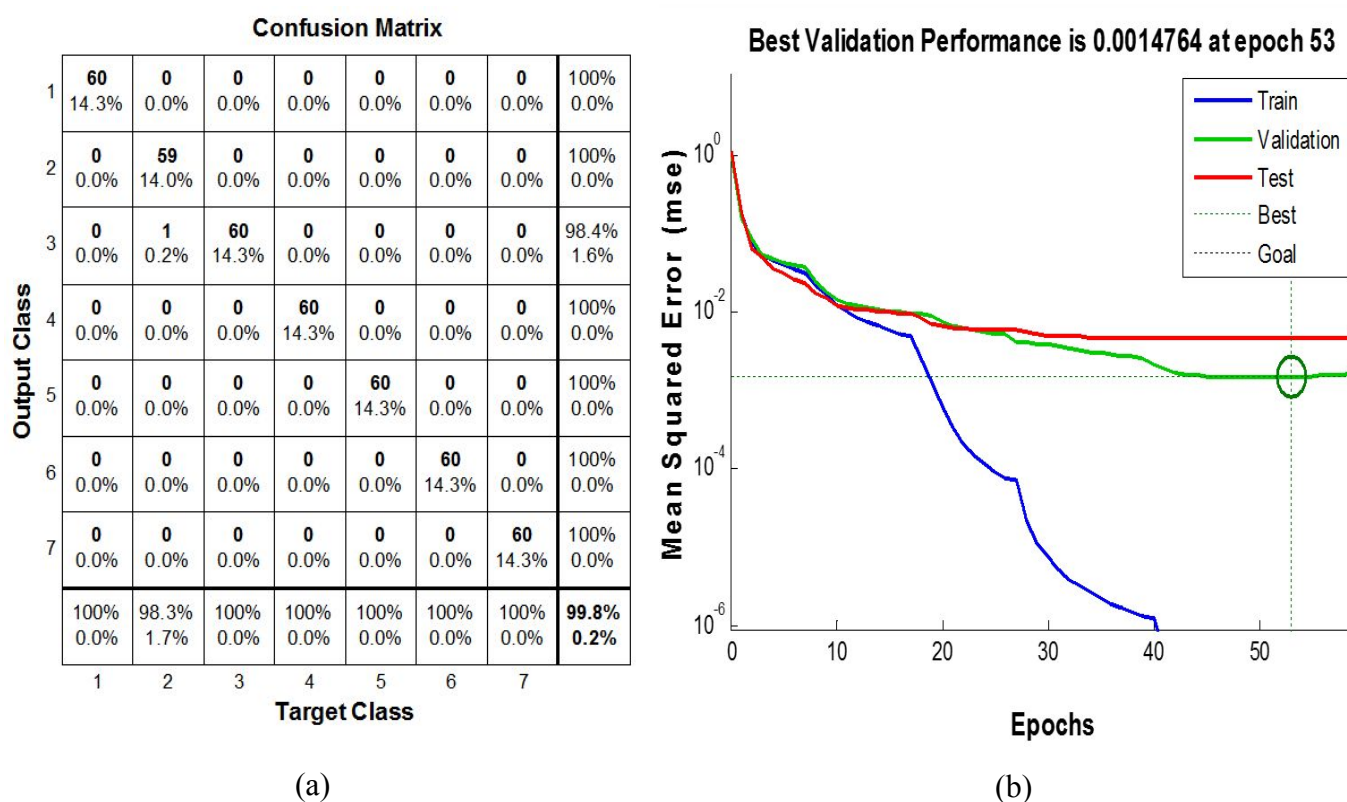


Fig. 9 (a) The overall confusion matrix of classification accuracy rate and (b) performance MLP-BP using 60 normalized features (db4 mother wavelet)

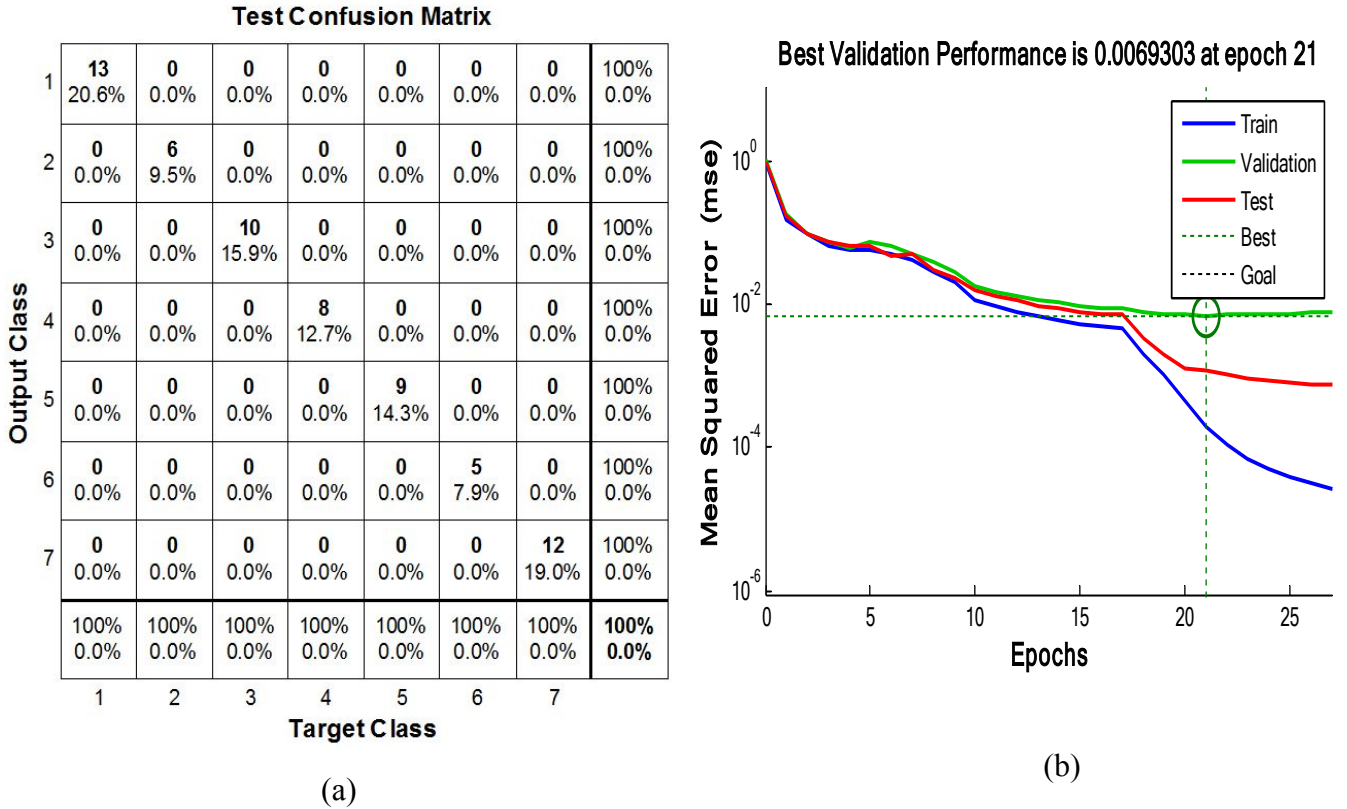


Fig. 10 (a) The test confusion matrix of classification accuracy rate and (b) performance MLP-BP 60 normalized features (rbio1.5 mother wavelet)

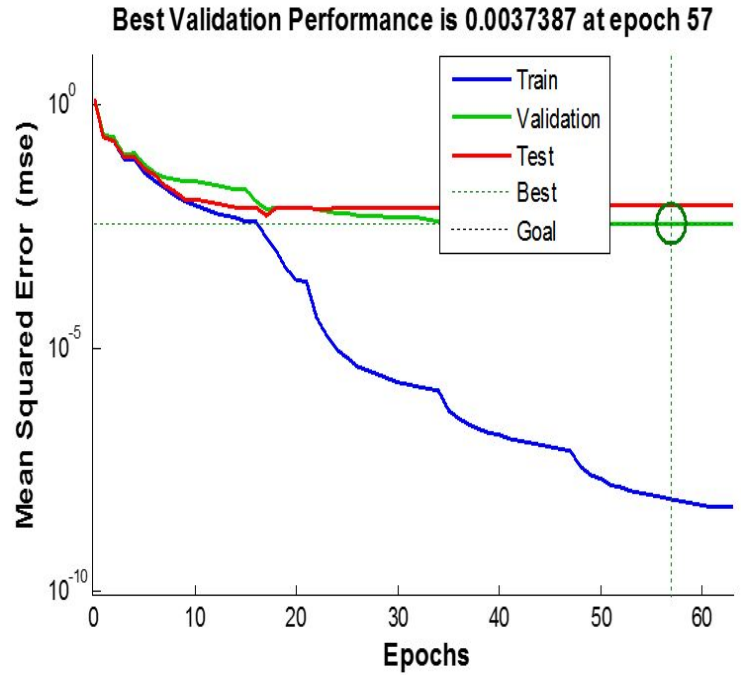
B. MLP-GABP

MLP-GABP is implemented for two cases of 60 normalized approximation features with db4 mother wavelet and also rbio1.5. MLP-GABP illustrated lower performance compared to MLP-BP in terms of computational time and overall classification accuracy rate which is 84.3% for db4 and 89% for rbio1.5. Four hidden layers containing [24 21 24 23] neurons are used in MLP as per selection of GA and weights of neural network have been adjusted and selected using GA. GA based optimization and training using db4 and rbio1.5 is terminated after 600 and 424 generations with best fitness function of 0.157143 and 0.109524 as shown in Figures 11-14 respectively. The best fitness function denotes the best minimized Mean Square Error (MSE). Table 3 shows a comprehensive MLP performance of both BP and GA for all cases.

Confusion Matrix

	1	2	3	4	5	6	7	
1	57 13.6%	0 0.0%	0 0.0%	13 3.1%	0 0.0%	6 1.4%	0 0.0%	75.0% 25.0%
2	1 0.2%	46 11.0%	1 0.2%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	95.8% 4.2%
3	0 0.0%	7 1.7%	56 13.3%	0 0.0%	0 0.0%	0 0.0%	2 0.5%	86.2% 13.8%
4	2 0.5%	0 0.0%	0 0.0%	28 6.7%	0 0.0%	2 0.5%	0 0.0%	87.5% 12.5%
5	0 0.0%	0 0.0%	0 0.0%	9 2.1%	57 13.6%	0 0.0%	0 0.0%	86.4% 13.6%
6	0 0.0%	0 0.0%	0 0.0%	10 2.4%	2 0.5%	52 12.4%	0 0.0%	81.3% 18.8%
7	0 0.0%	7 1.7%	3 0.7%	0 0.0%	1 0.2%	0 0.0%	58 13.8%	84.1% 15.9%
	95.0% 5.0%	76.7% 23.3%	93.3% 6.7%	46.7% 53.3%	95.0% 5.0%	86.7% 13.3%	96.7% 3.3%	84.3% 15.7%
	1	2	3	4	5	6	7	

(a)



(b)

Fig. 11 (a) The confusion matrix of overall classification accuracy rate and (b) performance of MLP-GABP (db4 mother wavelet)

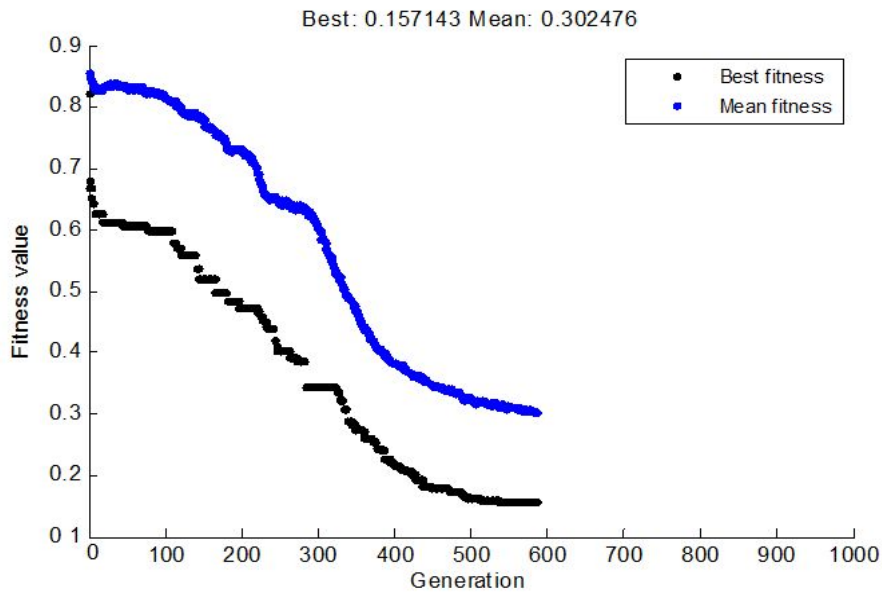
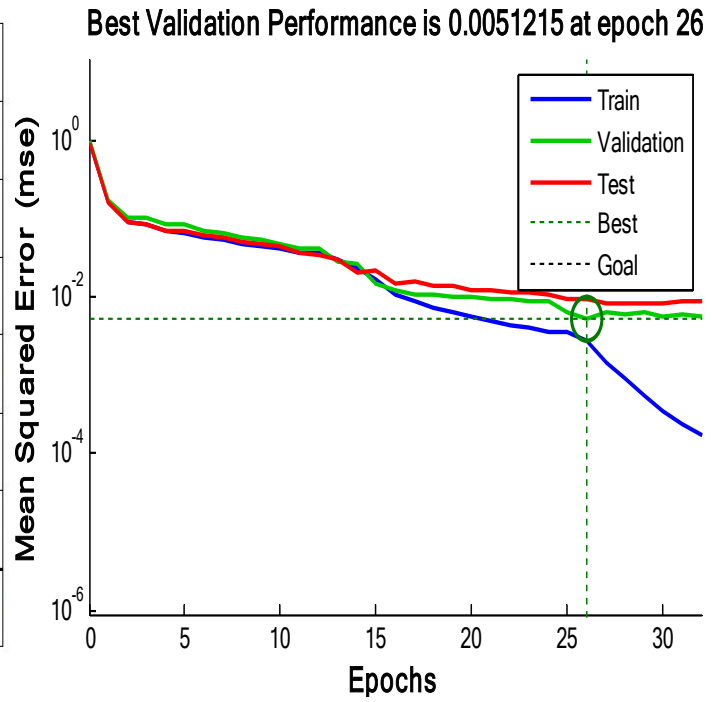


Fig. 12 Best score value and mean score versus generation (db4 function)

Confusion Matrix

1	59 14.0%	0 0.0%	0 0.0%	1 0.2%	0 0.0%	0 0.0%	0 0.0%	98.3% 1.7%
2	0 0.0%	58 13.8%	10 2.4%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	85.3% 14.7%
3	0 0.0%	2 0.5%	50 11.9%	2 0.5%	0 0.0%	0 0.0%	0 0.0%	92.6% 7.4%
4	1 0.2%	0 0.0%	0 0.0%	27 6.4%	0 0.0%	0 0.0%	0 0.0%	96.4% 3.6%
5	0 0.0%	0 0.0%	0 0.0%	0 0.0%	60 14.3%	0 0.0%	0 0.0%	100% 0.0%
6	0 0.0%	0 0.0%	0 0.0%	30 7.1%	0 0.0%	60 14.3%	0 0.0%	66.7% 33.3%
7	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	60 14.3%	100% 0.0%
	98.3% 1.7%	96.7% 3.3%	83.3% 16.7%	45.0% 55.0%	100% 0.0%	100% 0.0%	100% 0.0%	89.0% 11.0%
	1	2	3	4	5	6	7	
	Target Class							

(a)



(b)

Fig. 13 (a) The confusion matrix of overall classification accuracy rate and (b) performance of MLP-GABP (rbio1.5 mother wavelet)

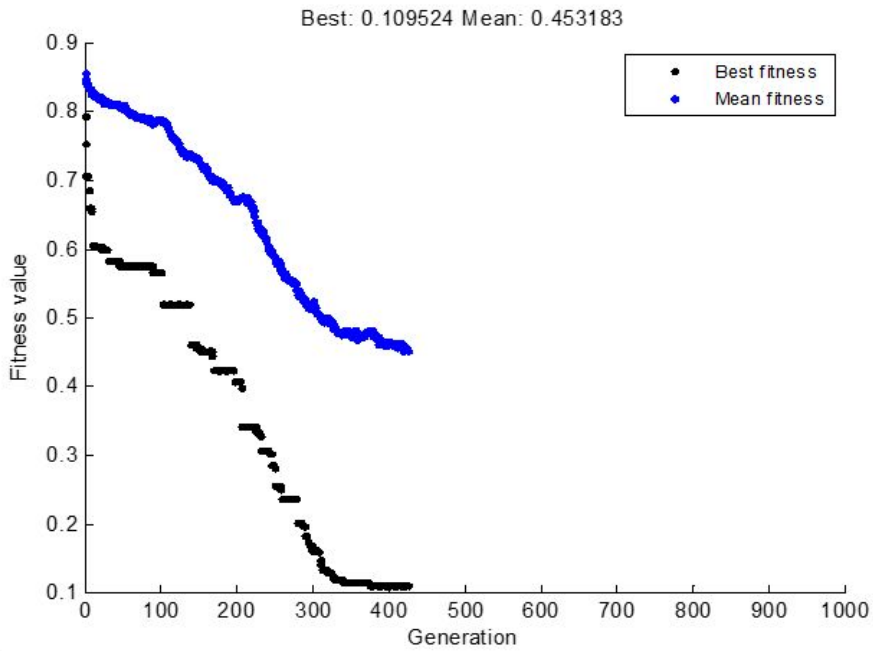


Fig. 14 Best score value and mean score versus generation (rbio1.5 function)

Table 3 MLP overall performance based DWT

Method	Features	Hidden layers & neurons	Training time (hh:mm:ss)	Test rate (%)	Validation rate (%)	Training rate (%)	Overall classification rate (%)
MLP-BP & GA selection (db4)	30 approx. (normalised)	[24 21 24 23]	00:00:17	100	100	100	100
MLP-BP & GA selection(rbio1.5)	30 approx. (normalised)	[24 21 24 23]	00:01:15	100	100	100	100
MLP-BP & GA selection(rbio1.5)	60 approx. (normalised)	[24 21 24 23]	00:00:43	100	98.4	100	99.8
MLP-BP & GA selection (db4)	60 approx. (normalised)	[24 21 24 23]	00:01:27	98.4	100	100	99.8
MLP-BP & GA selection (db4)	60 approx. (non-normalised)	[24 21 24 23]	00:01:00	98.4	98.4	100	99.5
MLP-BP & GA selection (rbio1.5)	60 approx. (non-normalised)	[24 21 24 23]	00:00:40	98.4	98.4	100	99.5
MLP-BP & GA selection (db4)	60 detail (normalised)	[24 21 24 23]	00:00:58	74.6	84.1	97.6	92.1
MLP-GABP & GA selection (rbio1.5)	60 approx. (normalised)	[24 21 24 23]	MLP: 00:00:51, GA: 15:00:00	90.5	88.9	88.8	89
MLP-BP & GA selection (db4)	60 detail (non-normalised)	[24 21 24 23]	00:00:39	77.8	61.9	91.5	85
MLP-GABP & GA selection (db4)	60 approx. (normalised)	[24 21 24 23]	MLP: 00:01:36, GA: 22:00:00	81	85.7	84.7	84.3

C. SVM

Classification accuracy rates using polynomial kernel function of the cases of db4 mother wavelet (approximation features) 60 normalized, 60 non-normalized, 30 normalized, and 30 non-normalized features are 96%, 88.39%, 99.8% and 98.8% respectively. The case of using 30 normalized features achieved 99.8% and it is found to be the best in terms of hyper-plane with wider margin and better classification accuracy rates as shown in Table. 4. On the other hand, Classification rates using rbio1.5 mother wavelet for the same 4 cases are 96.96%, 90%, 100%, and 99.8%. The best overall classification rate among both mother wavelets is 100% and is achieved with 30 normalized features using rbio1.5. Table 5 shows the SVM classification performance rates of all cases.

Table 4 The SVM classification accuracy rates with 30 normalized features (db4 mother function)

Classification (%)	Imbalance	Bearing	Cavitation	Impeller	Looseness	Misalignment
Healthy	100	100	100	100	100	100
Imbalance	—	100	100	100	100	100
Bearing	—	—	98	100	100	100
Cavitation	—	—	—	100	100	100
Impeller	—	—	—	—	100	100

Table 5 The SVM classification accuracy rates with 30 and 60 normalized features (db4 & rbio1.5 wavelet mother functions)

Method	Features	Overall classification rate (%)
SVM (rbio1.5)	30 approx. (normalised)	100
SVM (db4)	30 approx. (normalised)	99.8
SVM (rbio1.5)	30 approx. (non-normalised)	99.8
SVM (db4)	30 approx. (non-normalised)	98.8
SVM (rbio1.5)	60 approx. (normalised)	96.96
SVM (db4)	60 approx. (normalised)	96
SVM (rbio1.5)	60 approx. (non-normalised)	90
SVM (db4)	60 approx. (non-normalised)	88.39

9. Conclusion

The feature extraction and classification of the pump conditions based on MLP-BP, MLP-GABP and SVM are conducted successfully using DWT considering features extraction with two mother wavelet functions (db4 and rbio1.5). In MLP-BP, it has been observed that the best classification rate in terms of the overall classification rates are achieved using both mother wavelets, but in terms of the overall test classification rate, rbio1.5 outperformed db4, where 100% accuracy rate is performed using rbio1.5 and 98.4% using db4. However, with reduced number of features (30), the performance of MLP-BP is improved, and 100% is achieved with both wavelet functions (db4 and rbio1.5).

For the SVM, it has been also found that rbio1.5 using reduced number of features (30 normalized features) resulted in better classification rate than MLP-BP of an overall classification rate 100%. Peak and

RMS have been identified as the best parameters for SVM due to the good distribution of the features between the different conditions.

It is remarked that rbio1.5 mother wavelet provided better results compared to db4 mother wavelet for both MLP-BP and SVM, as shown with the first case of 60 features. Extracted features as approximations provided better classification rates than detail features. Also, it was reported 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 lower performance compared to MLP-BP with the case of 60 normalized features, as MLP-BP achieved a performance rate of 99.8%, with both wavelet functions and MLP-GABP achieved 89% and 84.3% with rbio1.5 and db4 respectively. In terms of computational time, MLP-BP outperformed MLP-GABP, as the longest time with MLP-BP using db4 is 1 minute and 27 seconds, but with MLP-GABP, 22 hours.

Finally, this work can be further investigated in future by implementing different architecture parameters to the GA for the optimization of MLP-BP. This would help to determine the best options to improving the classification accuracy rates, other optimization methods and more suitable mother wavelet functions for wavelet transform based feature extraction. Moreover, other AI methods would be applied and compared with MLP and SVM. In addition, as part of future work, different speeds for the pump will be considered for the experiments.

References

- [1] K. K. McKee, G. Forbes, I. Mazhar, R. Entwistle, and I. Howard, A review of major centrifugal pump failure modes with application to the water supply and sewerage industries, in ICOMS Asset Management Conference Gold Coast, QLD, Australia: Asset Management Council (2011).
- [2] B. K. N. Rao, Handbook of condition monitoring, 1st ed: Elsevier advanced technology, 1996.
- [3] H. Bendjama, K. Gherfi, D. Idiou, and M. S. Boucherit, Condition monitoring of rotating machinery by vibration signal processing methods, International Conference on Industrial Engineering and Manufacturing Batna University Algeria, (2014) 297-301.
- [4] A. Aherwar and M. S. Khalid, Vibration analysis techniques for gearbox diagnostic: a review, International Journal of Advanced Engineering Technology, 3 (2012).
- [5] V. Muralidharan and V. Sugumaran, Selection of discrete wavelets for fault diagnosis of monoblock centrifugal pump using the J48 algorithm, Applied Artificial Intelligent, 27 (2013).
- [6] A. Suhane, "Experimental Study on Centrifugal Pump to Determine the Effect of Radial Clearance on Pressure Pulsations, Vibrations and Noise," International Journal of Engineering Research and Applications vol. 2, pp. 1823-1829, 2012.
- [7] H. Q. Wang and P. Chen, Fault Diagnosis of Centrifugal Pump Using Symptom Parameters in Frequency Domain, Agricultural Engineering International: the CIGR E-journal, IX (2007).
- [8] J. F. Gulich, Centrifugal pumps: Springer, 2008.
- [9] A. Al-braik, O. Hamomd, F. Gu, and A. D. Ball, "Diagnosis of Impeller Faults in Centrifugal Pump Based on Spectrum Analysis of Vibration Signals," in Eleventh International Conference on Condition Monitoring and Machinery Failure Prevention Technologies Manchester UK, 2014.
- [10] S. Farokhzad, Vibration based fault detection of centrifugal pump by fast fourier transform and adaptive neuro-fuzzy inference system, Journal of mechanical engineering and technology, 1 (2013) 82-87.
- [11] W. Abdulkarem, R. Amuthakkannan, and K. F. Al-Raheem, "Centrifugal Pump Impeller Crack Detection Using Vibration Analysis," in 2nd International Conference on Research in Science, Engineering and Technology Dubai (UAE), 2014.
- [12] M. C. Pan and P. Sas, "Transient analysis on machinery condition monitoring," in 3rd international conference on signal processing. vol. 2 Beijing: IEEE, 1996, pp. 14-18 Oct 1996.
- [13] M. M. Stopa, B. J. C. Filho, and C. B. Martinez, "Incipient Detection of Cavitation Phenomenon in Centrifugal Pumps," IEEE TRANSACTIONS ON INDUSTRY APPLICATIONS, vol. 50, pp. 120-126, 2014.
- [14] A. Prakash, V. K. Agarwal, A. Kumar, and B. Nand, A review on machine condition monitoring and fault diagnostics using wavelet transform, International Journal of Engineering Technology, Management and Applied Sciences, 2 (2014) 84-93.
- [15] N. Mehala and R. Dahiya, A comparative study of FFT, STFT and wavelet techniques for induction machine fault diagnostic analysis, International conference of computational intelligent, man-machine systems cybernetics (CIMMACS '08), India, (2008).
- [16] Z. K. Peng and F. L. Chu, Application of the wavelet transform in machine condition monitoring and fault diagnostics: a review with bibliography, Mechanical Systems and Signal Processing, 18 (2004) 199-221.
- [17] M. Al-Tobi and K. F. Al-Raheem, Rolling element bearing faults detection, wavelet de-noising analysis, Universal journal of mechanical engineering, 5 (2015) 47-51.
- [18] S. Farokhzad, H. Ahmadi, and A. Jafary, Fault classification of centrifugal water pump based on decision tree and regression model, Journal of Science and today's world, 2 (2013) 170-176.
- [19] M. A. S. Al-Tubi, K. F. Al-Raheem, and W. Abdul-Karem, Rolling element bearing element faults detection, power spectrum and envelope analysis, International conference on applications and design in mechanical engineering Penang, Malaysia, (2012).

- [20] F. Al-Badour, L. Cheded, and M. Sunar, "non-stationary vibration signal analysis of rotating machinery via time-frequency and wavelet techniques," in 10th international conference on information science, signal processing and their applications (ISSPA 2010) Malaysia Kuala Lumpur: IEEE, 2010, pp. 21-24.
- [21] V. Muralidharan, V. Sugumaran, and V. Indira, Fault diagnosis of monoblock centrifugal pump using SVM, *Engineering Science and Technology, an International Journal*, 17 (2014) 152-157.
- [22] N. R. Sakthivel, B. B. Nair, M. Elangovan, V. Sugumaran, and S. Saravanmurugan, 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 (2014) 30-38.
- [23] N. R. Sakthivel, Binoy.B.Nair, and V.Sugumaran, Soft computing approach to fault diagnosis of centrifugal pump, *Applied Soft Computing*, 12 (2012) 1574–1581.
- [24] G. M. Gharehbolagh, S. Farokhzad, M. R. A. A. Abad, and M. Ranjbarkohan, Fault diagnosis of crown wheel and pinion in differential using acoustic signals with discrete wavelet transform and neural network, *Technical Journal of Engineering and Applied Sciences*, 3 (2013) 1358-1364.
- [25] J. Sanz, R. Perera, and C. Huerta, Gear dynamics monitoring using discrete wavelet transformation and multi-layer perceptron neural networks, *Applied Soft Computing*, 12 (2012) 2867–2878.
- [26] N. Saravanan and K. I. Ramachandran, Incipient gear box fault diagnosis using discrete wavelet transform (DWT) for feature extraction and classification using artificial neural network (ANN), *Expert Systems with Applications*, 37 (2010) 4168–4181.
- [27] H. K. Srinivas, K. S. Srinivasan, and K. N. Umesh, Application of Artificial Neural Network and Wavelet Transform for Vibration Analysis of Combined Faults of Unbalances and Shaft Bow. *Advances In theoretical and applied mechanics*, 3, (2010) 159 - 176.
- [28] Muralidharan V., and Sugumaran V., "Rough set based rule learning and fuzzy classification of wavelet features for fault diagnosis of monoblock centrifugal pump", *Measurement*, vol. 46, (2013b) pp. 3057-3063, DOI 10.1016/j.measurement.2013.06.002.
- [29] Muralidharan V., Sugumaran V., Shanmugam P., Sivanathan K. "Artificial neural network based classification for monoblock centrifugal pump using wavelet analysis", *International Journal of Mechanical Engineering*, Vol. 1, (2010) pp. 28-37, ISBN ISSN 0976 – 6359.
- [30] B. Samanta, "Gear fault detection using artificial neural networks and support vector machines with genetic algorithms," *Mechanical Systems and Signal Processing*, 18 (2004) 625–644.
- [31] B. Samanta, K. R. Al-Balushi, and S.A. Al-ArAIMI, "Artificial Neural Networks and Genetic Algorithm for Bearing Fault Detection," *Soft Computing*, 10 (2006) 264-271.
- [32] Z. Yang, W. I. Hoi, and J. Zhong, Gearbox Fault Diagnosis based on Artificial Neural Network and Genetic Algorithms, in *International Conference on System Science and Engineering*, Macau, China (2011) 37-42.
- [33] K. F. Al-Raheem and W. Abdul-Karem, Rolling bearing fault diagnostics using artificial neural networks based on laplace wavelet analysis, *International journal of Engineering, Science and Technology*, 2 (2010) 278-290.
- [34] M. A. S. AL Tobi, G. Bevan, P. Wallace et al., Fault diagnosis of a centrifugal pump using MLP-GABP and SVM with CWT, *Engineering Science and Technology, an International Journal*, vol. 22 (3), (2019) 854-861 <https://doi.org/10.1016/j.jestch.2019.01.005>.
- [35] E. Charniak and D. McDermott, *Introduction to artificial intelligent: addison wesley Longman Inc*, 2000.
- [36] W. S. McCulloch and W. Pitts, A logical calculus of the ideas immanent in nervous activity, *bulletin of mathematical biophysics*, 5 (1943) 115-133.
- [37] L. A. Zadeh, Fuzzy Sets, *Information and Control*, 8 (1965) 338-353.
- [38] P. W. Hott and A. J. Griffiths, Fault diagnosis of pumping machinery using artificial neural networks, *Journal of Process Mechanical Engineering*, 211 (1997) 185-194.

- [39] R. Zouari, S. Sieg-Zieba, and M. Sidahmed, Fault detection system for centrifugal pumps using neural networks and neuro-fuzzy techniques, *Surveillance* 5 Cetim Senlis (2004).
- [40] K. F. Al-Raheem, A. Roy, K. P. Ramachandran, D. K. Harrison, and S. Grainger, Application of the Laplace- Wavelet Combined With ANN for Rolling Bearing Fault Diagnosis, *Journal of Vibration and Acoustics*, 130 (2008).
- [41] Z. Zhang, Y. Wang, and K. Wang, 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 (2013) 763-773.
- [42] Ł. Jedlinski and J. Jonak, Early fault detection in gearboxes based on support vector machines and multilayer perceptron with a continuous wavelet transform, *Applied Soft Computing*, 30 (2015) 636–641.
- [43] D. Y. Wang, W. Z. Zhang, and J. G. Zhang, Fault Bearing Identification Based on Wavelet Packet Transform Technique and Artificial Neural Network, in *International conference on system science, engineering design and manufacturing informatization*, Yichang, (2010) 11-14.
- [44] L. Yanjun, M. Fanlong, and L. Yibo, 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*, (2009) 4806-4810.
- [45] W.-T. Sui and D. Zhang, Rolling Element Bearings Fault Classification Based on SVM and Feature Evaluation Proceedings of the Eighth International Conference on Machine Learning and Cybernetics, Baoding (2009).
- [46] Y. Liu and T. Liu, Rotating Machinery Fault Diagnosis Based on Support Vector Machine, *International Conference on Intelligent Computing and Cognitive Informatics Kuala Lumpur: IEEE*, (2010) 71-74.
- [47] J. Zhong, Z. Yang, and S. F. Wong, Machine Condition Monitoring and Fault Diagnosis based on Support Vector Machine, *International conference of Industrial Engineering Management Macao: IEEE*, (2010) 2228-2233.
- [48] D. E. Rumelhart, G. E. Hinton, and R. J. William, Learning representations by back-propagating errors, *Nature*, 323 (1986).
- [49] C. Cortes and V. Vapnik, Support-vector networks, *Machine Learning*, 20 (1995) 273-297.
- [50] J. H. Holland, *Adaptive in natural and artificial systems*. Ann Arbor: The university of Michigan press, 1975.
- [51] A. J. F. V. Rooij, L. C. Jain, and R. P. Johnson, *Neural network training using genetic algorithm*. Singapore: World Scientific Publishing Co. Pte. Ltd, 1996.
- [52] C. Z. Janikow and Z. Michalewicz, An experimental comparison of binary and floating-point representations in genetic algorithms. In *ICGA (1991)*, pp. 31-36.
- [53] A. J. F. V. Rooij, L. C. Jain, R. P. Johnson, *Neural network training using genetic algorithm*. Singapore, World Scientific Publishing Co. Pte. Ltd, (1996), ISBN 981-02- 2919-4.
- [54] H. Mahmoudabadi, M. Izadi, M. B. MenhaJ, A hybrid method for grade estimation using genetic algorithm and neural networks, *Computational Geoscience*, 13 (2009), pp. 91-10, DOI 10.1007/s10596-008-9107-9
- [55] J. Rafiee, P. W. Tse, A. Harifi, and M. H. Sadeghi, A novel technique for selecting mother wavelet function using an intelligent fault diagnosis system, *Expert Systems with Applications*, 36 (2009) 4862–4875.
- [56] K. Akilandeswari and G. M. Nasira, Multi-layer perceptron neural network classifier with binary particle swarm optimization based feature selection for brain-computer interfaces, *World Academy of*

Science, Engineering and Technology, International Journal of Computer, Electrical, Automation, Control and Information Engineering, 9 (6), (2015), pp. 1615-1621, ISSN 1307-6892.

[57] M.T. Wong, X. He, H. Nguyen, W. C. Yeh, Mass classification in digitized mammograms using texture artificial neural network”, In International Conference on Neural Information Processing, (2012), pp. 151-158, Springer, Berlin, Heidelberg, ISBN 978-3-642-34500-5.

[58] E. M. Karabulut, S. A. Özel, Ṫbrikçi, A comparative study on the effect of feature selection on classification accuracy, Procedia Technology, 1 (2012), pp. 323-327.

[59] M. H. Song, J. Lee, S. P. Cho, K. J. Lee, S. K. Yoo, Support Vector Machine based Arrhythmia classification using reduced features, International Journal of Control, Automation, and Systems, 3 (2005), pp. 571-579.

[60] S. Shilaskara and A. Ghatolb, Feature selection for medical diagnosis: Evaluation for cardiovascular diseases, Expert Systems with Applications, 40 (2013), pp. 4146-4153.

[61] B. K. Alsberg, A. M. Woodward, and D. B. Kell, An introduction to wavelet transforms for chemometricians: A time-frequency approach, Chemometrics and Intelligent Laboratory Systems, 37 (1997) 215-239.

[62] V. Muralidharan, V. Sugumaran, V. Indira, Fault diagnosis of mono block centrifugal pump using SVM, Engineering Science and Technology, an International Journal, 17 (2014), pp. 152-157.

[63] S. G. Mallat, A theory for multiresolution signal decomposition: the wavelet representation, IEEE Transactions on Pattern Analysis and Machine Intelligent, 11 (1989) 674-693.

[64] Barbosa A.O.G., Diaz D.R.A., Vellasco M.M.B.R., Meggiolaro M.A., Tanscheit R. (2009) Mental Tasks Classification for a Noninvasive BCI Application. In: Alippi C., Polycarpou M., Panayiotou C., Ellinas G. (eds) Artificial Neural Networks – ICANN 2009. ICANN 2009. Lecture Notes in Computer Science, vol 5769. Springer, Berlin, Heidelberg

[65] Suuny, S., Peter, S. D., & Jacob, K. P. (2013). Performance of different classifiers in speech recognition. International Journal of Research in Engineering and Technology, 2(04), 590-597.

[66] Wang, Z. Q., & Tashev, I. (2017). Learning utterance-level representations for speech emotion and age/gender recognition using deep neural networks. ICASSP.