Available online at www.sciencedirect.com
ScienceDirect
Procedia Engineering 144 (2016) 343 – 350
Procedia Engineering
www.elsevier.com/locate/procedia

12th International Conference on Vibration Problems, ICOVP 2015

# Bearing Fault Diagnosis Using Feature Ranking Methods and Fault Identification Algorithms

V.Vakharia[a]*,V.K.Gupta[b],P.K.Kankar[b]

[a]*Medi-Caps Group of Institutions, Indore, 453331,India*
[b]*PDPM IIITDM, Jabalpur, 482005, India*

**Abstract**

Diagnoses of bearing faults are important to avoid catastrophic failures in rotating machines. This paper presents a methodology to detect various bearing faults from the measured vibration signal. Features such as kurtosis, skewness, mean, root mean square and complexity measure such as Shannon entropy are calculated from time domain ,frequency domain and discrete wavelet transform. In total 40 features are calculated from bearing conditions such as Healthy bearing, Inner race fault, Outer race fault and Ball fault. Feature ranking methods such as Chisquare, ReliefF method are used to select most informative feature and subsequently to reduce size of feature vector. Comparison has been made between feature ranking methods and classifiers to obtain best diagnosis result with reduce feature set. Our results shows good fault identification accuracy with minimum number of features.


Peer-review under responsibility of the organizing committee of ICOVP 2015

*Keywords:* *Bearing;Fault diagnosis; Feature ranking method; Classifiers*

## 1. Introduction

The fault diagnosis technique of complex rotating components using vibration analysis has gained considerable attention from researchers across the globe. Majority of problems in rotating machinery are caused by faulty gears, bearings etc. Failure in bearing is one of the primary causes of breakdown in rotating machines. Such breakdowns can lead to expensive shutdowns, drifts in production and even human casualties.

* Corresponding author. Tel.: +91-835-988-1214.
*E-mail address:*vinayvakharia4343@gmail.com


Peer-review under responsibility of the organizing committee of ICOVP 2015
doi:10.1016/j.proeng.2016.05.142

**Nomenclature**

| | |
|---|---|
| Yj | number of observations in category j |
| $\mu_j$ | expected value of Yj |
| $p_{io}$ | known probability of occurrence |
| $f_{t,i}$ | value of instance $x_i$ on feature $f_i$ |
| P | distance measurement |

In data driven techniques for condition monitoring, vibration signals are obtained from rotor bearing test rig and this signals are processed to extract relevant features which contains information about state of the system. A review of several vibration and acoustic measurement methods used to detect defects in rolling element bearings was presented by Tandon and Choudhury [1]. Advantage of vibration based fault diagnosis technique is that there is no stoppage of machinery required during maintenance. Signals extracted from rolling element bearing for the fault diagnosis are broadly categorized in to three domains: time domain, frequency domain and time-frequency domain. Time domain signal generally gives information how signal amplitude is varied with respect to time. The drawback of features calculated by time domain method is that it is unable to detect faults at early stage [2].Frequency domain method is another technique for fault diagnosis of bearing. Every bearing component has its own characteristic frequency. With the help of Fast Fourier Transform (FFT) defects in bearing component can be identify. Disadvantage of Frequency domain method lies in its inability to analyze the non-stationary signals which are generally related to component/machinery defects [3,4].

In recent years, the time-frequency methods such as Wavelet Transform (WT) have been suggested by authors to extract very weak signals [5,6].To enhance fault related information and to reduce noise, DWT based denoising technique was used by Du et al. [7] for bearing condition monitoring. Discrete wavelet transform produces wavelet coefficients after transforming original time domain signal in to the wavelet domain. DWT is applied to discrete data sets and it produces discrete outputs. While using wavelet transform challenge is to choose most appropriate wavelet. Therefore, mother wavelet selection methodologies were proposed based on maximum energy to Shannon entropy ratio and multiscale permutation entropy [5, 8]. Estimating the quality of features is an important issue in the field of machine learning. The criteria used to select the useful features entirely depend on the nature of the feature ranking technique used. In feature ranking method effectiveness of each individual feature is calculated and then the analyst selects features which are appropriate from a given dataset.

In present study, a generalized approach to select optimal number of feature set using gain ratio and ReliefF feature ranking method has been proposed and is evaluated with different classifiers. The combination of feature ranking technique and classifier is used to select the optimum number of feature set which gives maximum efficiency. Fig.1 shows the proposed methodology for fault diagnosis using feature ranking method.

## 2. Machine Learning Techniques

### *3.1 Artificial Neural Network (Multilayer Perceptron)*

Artificial intelligence techniques such as fuzzy logic, artificial neural network (ANN) have been continuously and successfully applied for bearing fault detection and diagnosis. ANN [9] are made up of interconnected processing units known as neurons and it is adaptively changes its structure during learning phase. ANN usually consists of inputs which are multiplied by weights where weights denote the strength of signal and the computation is done by a mathematical function which denotes the activation of neuron. Based on the signal received neuron computation will be different. Thus higher the weight of artificial neuron stronger the input and by adjusting the weights of a neuron we can able to obtain desired output for a pre-specified inputs. ANN is a type of supervised learning methods which can be trained by supplying data. Multilayer perceptron algorithm is used for testing purpose during which weights are adjusted for error minimization between ANN predictions and outputs [10].

### *3.2 Random Forest (RF)*

Random forest is a type of ensemble learning method based on decision trees used for classification and regression purpose. The random forest algorithm was developed by Breiman [11].In the initial phase, the training sets are divided into in-bag and out-bag set. In the next phase for prediction (in-bag) samples, two third portion of training samples are used and the remaining one third portion is used for prediction accuracy validation (out-bag).The process is repeated for several times on the constructed feature set to produce multiple in-bag sets and out-bag set subsets. Finally the predictions are obtained from out-bag value from the entire training dataset. Here each individual decision tree cast a vote for one class and these can be used to calculate the generalization capability of classifier. In case of multiclass prediction the class is identified by gaining maximum vote [12].

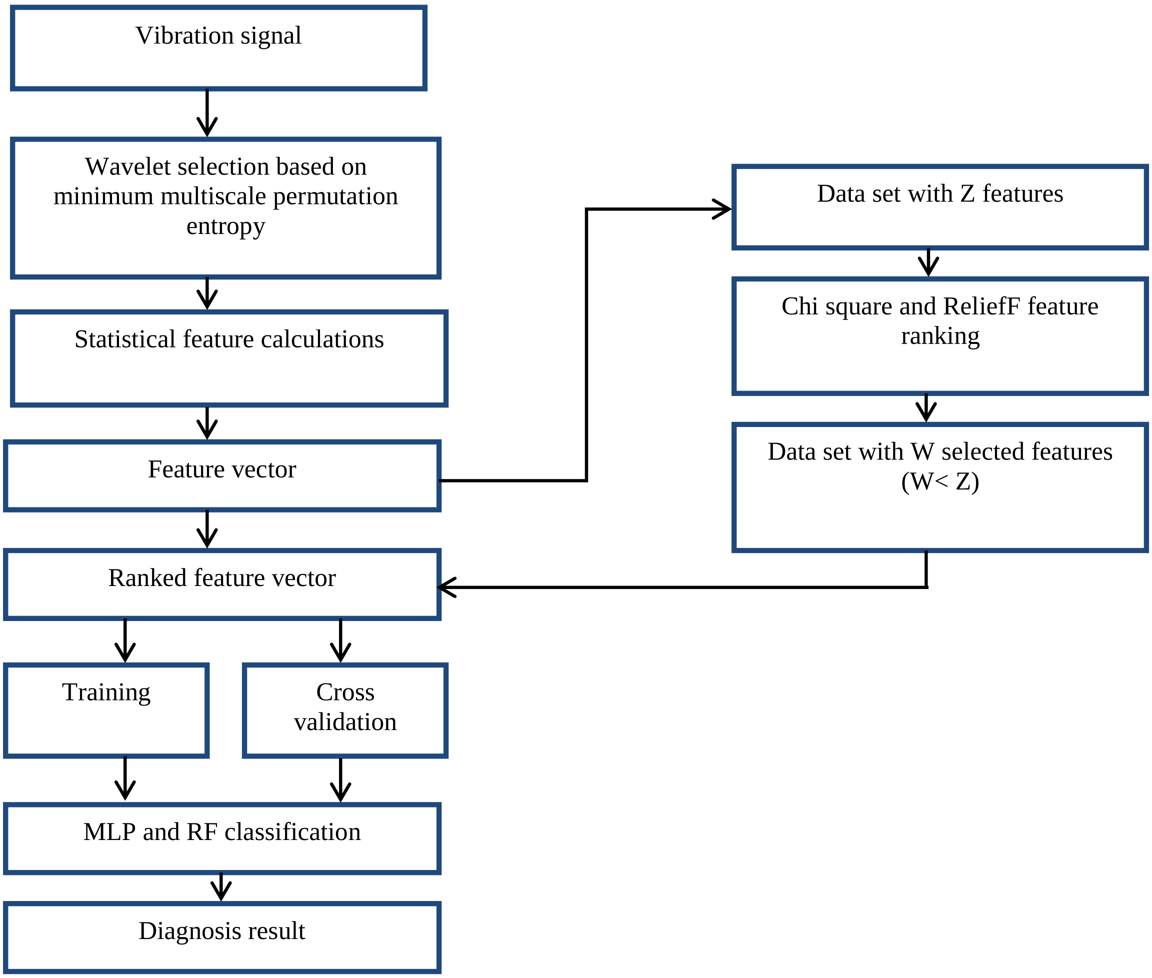



Fig.1. Proposed methodology for fault diagnosis using feature ranking method

## 3. Feature Selection

Feature selection is required to choose a small subset of features from original feature subset to reduce dimensionality without compromising information in features. To make a decision which feature has to be retain and which feature has to discard depends solely on the technique which has to be applied. In this study, two feature ranking criteria are compared namely: Chisquare and ReliefF with ANN and RF classifier.

### *3.1 Chisquare Feature Selection*

Chisquare test is the method used for feature ranking and it was introduced by British statistician Karl Pearson introduced initially for measure of goodness of fit. The Chisquare test is generally used in statistics to check the independence between two events [12].In feature selection we use it to test whether the occurrence of a specific term and the occurrence of a specific class are independent or not. Thus we estimated the following quantity for each term and we rank them by their score [13].Let Yj denote the number of observations in category j and $\mu_j$ represents the expected value of Yj and are given by $\mu_j = np_{j0}$.Here $p_{j0}$ represents the known probability of occurrence.

$$x^2 = \sum_{j=1}^{k} \frac{(y_j - \mu_j)^2}{\mu_j} \tag{1}$$

### *3.2 ReliefF*

ReliefF is a supervised algorithm for feature ranking [15]. It is usually applied in data pre-processing as a feature subset selection method. The basic idea behind use of ReliefF is to compute instances at random, compute their nearest neighbors and adjust a feature weighting vector to give more weight to features which discriminate the instance from neighbors of different classes. Let the total number of instances be N. For a two class problem the evaluation criteria of ReliefF is

$$\left(f_i\right) = \frac{1}{2}\sum_{i=1}^{N} P\left(f_{t,i} - f_{dc(xi)}\right) - P\left(f_{t,i} - f_{sc(xi)}\right) \tag{2}$$

where $f_{t,i}$ denotes value of instance $x_i$ on feature $f_i$ ,P denotes distance measurement and $f_{dc(xi)}$,$f_{sc(xi)}$ represents value of $i^{th}$ feature of nearest points to $x_i$ with different and same class label respectively.

Table 1 Specification of bearing 6205 (Drive end)

| Bearing Type | Inside diameter (mm) | Outside diameter (mm) | Ball diameter (mm) | Pitch diameter (mm) |
|---|---|---|---|---|
| 6205 | 25 | 52 | 7.94 | 39 |

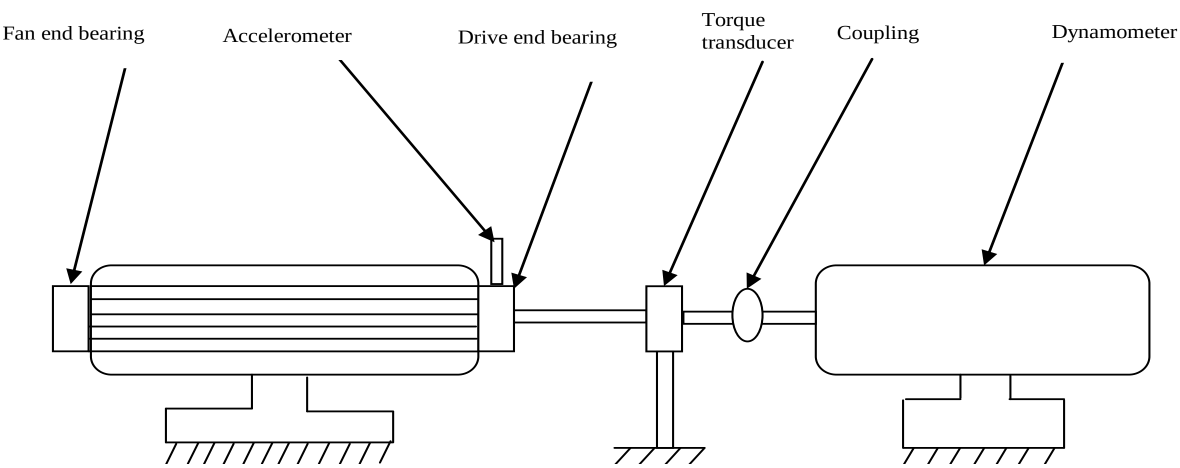
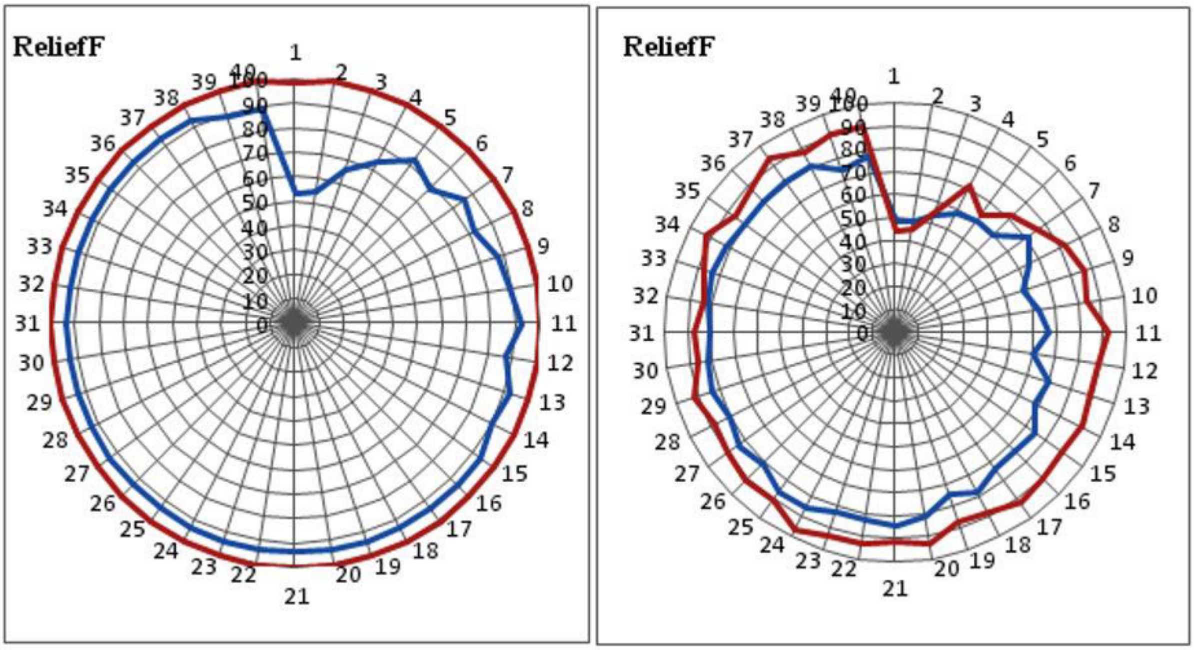



Fig.2. Schematic diagram of test rig




(a) (b)

Fig.3. (a-b) Training and cross validation efficiency of ANN and RF based on ReliefF ranking method.

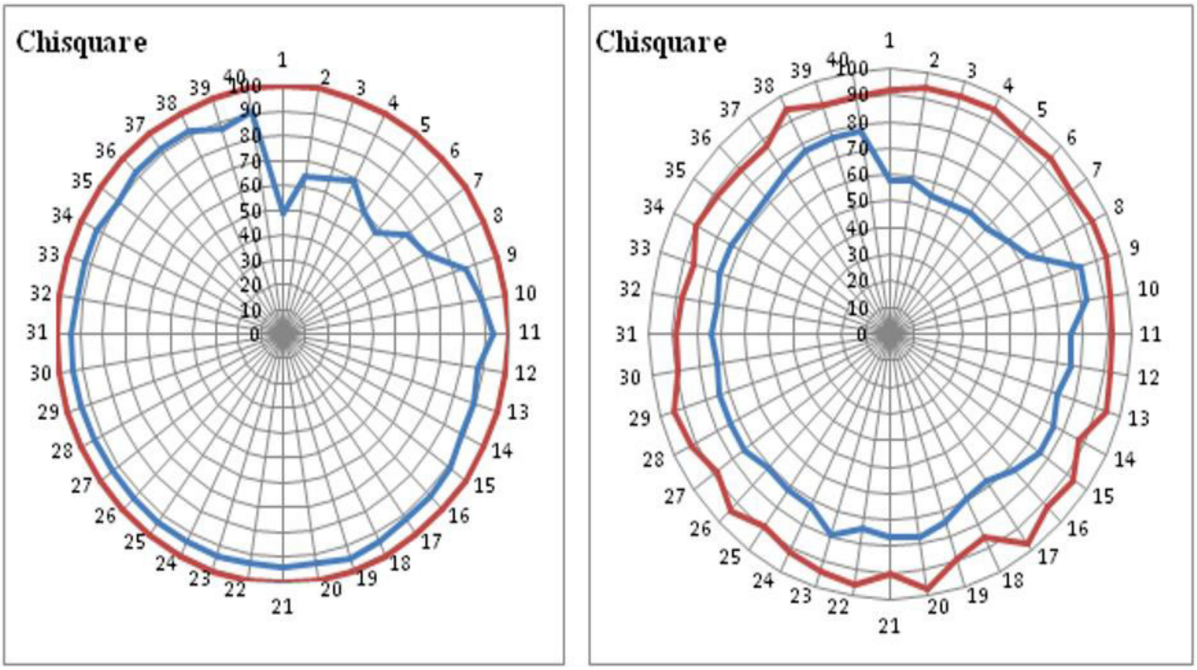



Fig.4. (a-b) Training and cross validation efficiency of ANN and RF based on Chisquare ranking method.

## 4. Experimental Procedure

Raw vibration signals from healthy bearing and bearing with defect has been utilized to calculate statistical features such as kurtosis, skewness, RMS, mean, Shannon entropy from time domain, frequency domain and time frequency (DWT, detail and approximate) domain. In total 40 statistical features are calculated from bearing test rig which belongs to Case Western Reserve University Bearing Data Center (CWRU). Schematic diagram of rotor bearing system is shown in Fig.2 [16].The data set belongs to 12K drive end bearing due to broader variety of fault sets. Dimensions of ball bearing utilized to conduct experiment are shown in Table 1. Vibration signals are recorded at various rotational speed 1725, 1748, 1772 and 1796 rpm with following bearing fault classes considered for study: Healthy bearing (HB), Inner race defect (IRD), Outer race defect (ORD) and Ball defect (BD).Sampling frequency was set to 12 kHz.

## 5. Results and Discussion

In the present study statistical features are calculated from time domain, frequency domain and discrete wavelet transform respectively. Features calculated are used to form feature vector and are fed as an input to machine learning techniques such as ANN and RF for diagnosis of bearing faults [13]. To reduce the size of feature set two feature ranking algorithm viz. ReliefF and Chisquare are used. Training and tenfold cross validation results are shown in Fig.3 and Fig.4.

From figure 3 (a) it is clear that maximum training efficiency of 100 % is obtained with RF classifier with only two ReliefF ranked features while maximum training efficiency of 93.54 % is obtained with top fifteen ranked features when ANN is used as classifier. Tenfold cross validation efficiency obtained through ANN and RF is shown in Fig. 3 (b). With ANN as classifier and ReliefF as feature ranking method maximum tenfold cross validation efficiency obtained is 82.25 % with twenty two ranked features. Similarly with RF as classifier and ReliefF as feature ranking method maximum tenfold cross validation efficiency obtained is 93.54 % with twenty ranked features. Thus it can be infer from Fig. 3(a) and Fig. 3(b), RF is the best classifier as compare to ANN since it is giving better training and tenfold cross validation efficiency.

Further for effective fault diagnosis and to evaluate the effect of feature ranking method on classifiers authors have used Chisquare feature ranking method. Training and tenfold cross validation efficiency obtained through Chisquare feature ranking is shown in Fig. 4(a) and Fig.4 (b).It is observed that 100 % training efficiency is obtained when only one feature is used when RF is used as a classifier. When ANN is used as classifier then maximum training efficiency obtained is 95 % with fifteen ranked features. For tenfold cross validation maximum fault identification accuracy 93.54 % is obtained through RF classifier with only top two ranked features and 77.41 % fault identification accuracy is obtained with ANN as a classifier with twenty ranked features which is shown in Fig. 4 (a) and Fig. 4 (b).Based on results obtain it can be conclude that RF gives better fault identification accuracy for both training and tenfold cross validation and Chisquare is the better ranking method for obtaining informative features.

To get an insight about class wise fault identification accuracy confusion matrix is used. Table 2-Table 4 gives detail about prediction accuracy of fault cases consider such as IRD, BD, ORD and HB. Table 2 shows confusion matrix which is obtained with ReliefF ranking method and top fifteen selected features based on training of feature set. When ANN is used then fault identification rate of IRD, BD and HB is 100 % while for ORD 22 out of 26 cases are predicted correctly while 4 cases are predicted incorrectly as BD. For RF 100 % training accuracy achieved so incorrectly predicted instances are not seen in confusion matrix. Table 3 shows confusion matrix with maximum cross validation efficiency using ReliefF ranking method. HB i.e. bearing with no fault is identified correctly by both ANN and RF. In Table 4 when ANN is used as a classifier then prediction rate of IRD, BD and HB is 100 % whereas 23 out of 26 instances are predicted correctly giving 95.1 % fault identification accuracy. Table 5 shows confusion matrix based on tenfold cross validation result. When one top ranked feature is used then prediction rate of HB is 100 % where as other fault cases such as IRD, BD and ORD are misclassified giving 93.54 % tenfold cross validation efficiency.

Table 2 Confusion matrix showing maximum training efficiency with ReliefF ranking method

| Fifteen selected features (ANN)-training | | | | | One selected features (ReliefF and RF)-training | | | | |
|---|---|---|---|---|---|---|---|---|---|
| IRD | BD | ORD | HB | Identification Result (%) | IRD | BD | ORD | HB | Identification Result (%) |
| 16 | 0 | 0 | 0 | IRD | 16 | 0 | 0 | 0 | IRD |
| 0 | 16 | 0 | 0 | BD | 0 | 16 | 0 | 0 | BD |
| 0 | 4 | 22 | 0 | ORD | 0 | 0 | 26 | 0 | ORD |
| 0 | 0 | 0 | 4 | HB | 0 | 0 | 0 | 4 | HB |
| Fault identification accuracy –93.54 % | | | | | Fault identification accuracy – 100 % | | | | |

Table 3 Confusion matrix showing maximum cross validation efficiency with ReliefF ranking method

| Twenty Two selected features (ReliefF and ANN)-cross validation | | | | | Twenty selected features (ReliefF and RF)-cross validation | | | | |
|---|---|---|---|---|---|---|---|---|---|
| IRD | BD | ORD | HB | Identification Result (%) | IRD | BD | ORD | HB | Identification Result (%) |
| 14 | 0 | 2 | 0 | IRD | 14 | 0 | 2 | 0 | IRD |
| 1 | 12 | 3 | 0 | BD | 0 | 15 | 1 | 0 | BD |
| 1 | 4 | 21 | 0 | ORD | 1 | 0 | 25 | 0 | ORD |
| 0 | 0 | 0 | 4 | HB | 0 | 0 | 0 | 4 | HB |
| Fault identification accuracy – 82.25 % | | | | | Fault identification accuracy – 93.54 % | | | | |

Table 4 Confusion matrix showing maximum training efficiency with Chisquare ranking method

| Ninteen selected features (Chisquare and ANN)-training | | | | | One selected features (Chisquare and RF)-training | | | | |
|---|---|---|---|---|---|---|---|---|---|
| IRD | BD | ORD | HB | Identification Result (%) | IRD | BD | ORD | HB | Identification Result (%) |
| 16 | 0 | 0 | 0 | IRD | 16 | 0 | 0 | 0 | IRD |
| 0 | 16 | 0 | 0 | BD | 0 | 16 | 0 | 0 | BD |
| 0 | 3 | 23 | 0 | ORD | 0 | 0 | 26 | 0 | ORD |
| 0 | 0 | 0 | 4 | HB | 0 | 0 | 0 | 4 | HB |
| Fault identification accuracy – 95.16 % | | | | | Fault identification accuracy – 100 % | | | | |

Table 5 Confusion matrix showing maximum cross validation efficiency with Chisquare ranking method

| Twenty selected features (Chisquare and ANN)-cross validation | | | | | One selected features (Chisquare and RF)-cross validation | | | | |
|---|---|---|---|---|---|---|---|---|---|
| IRD | BD | ORD | HB | Identification Result (%) | IRD | BD | ORD | HB | Identification Result (%) |
| 15 | 0 | 1 | 0 | IRD | 14 | 0 | 2 | 0 | IRD |
| 1 | 10 | 5 | 0 | BD | 0 | 15 | 1 | 0 | BD |
| 2 | 5 | 19 | 0 | ORD | 2 | 0 | 24 | 0 | ORD |
| 0 | 0 | 0 | 4 | HB | 0 | 0 | 0 | 4 | HB |
| Fault identification accuracy – 77.41 % | | | | | Fault identification accuracy – 93.54 % | | | | |

## 6. Conclusion

In this paper we proposed methodology to identify bearing faults using ANN and RF fault identification algorithm. Forty statistical features are calculated from time domain, frequency domain and discrete wavelet transform. To select most informative feature and simultaneously to reduce size of feature vector feature ranking method Chisquare and ReliefF is utilized. Effect of ranked feature on the performance of ANN and RF are investigated in detail. It is observed from experimental study that 93.54 % tenfold cross validation accuracy is obtain when Chisquare feature ranking method is used along with RF. The methodology proposed is useful for developing online fault diagnosis.

## Acknowledgments

The authors would like to thank Prof. K.A. Loparo of Case Western Reserve University for providing bearing vibration data sets which is used in present study.

## References

[1] N.Tandon, and A. Choudhury, A review of vibration and acoustic measurement methods for the detection of defects in rolling element bearings, Tribology International.32 (1999) 469-480.
[2] R.J. Alfredson and J. Mathew, Time domain methods for monitoring the condition of rolling element bearings,Mechanical. Engineering Transaction.10 (2) (1985) 102-107.
[3] A. Jardine, D. Lin and D. Banjevic, A review on machinery diagnostics and prognostics implementing condition-based maintenance, Mechanical Systems and Signal Processing.20(2006) 1483-1510.
[4] R. Upadhyaya, A.Manglick, D.K.Reddy, P.K.Padhy, P.K.Kankar, Channel optimization and nonlinear feature extraction for electroencephalogram signals classification.,Computers & Electrical Engineering. 45 (2015) 222-234.
[5] V.Vakharia,V.K Gupta, P.K. Kankar, A multiscale entropy based approach to select wavelet for fault diagnosis of ball bearings, Journal of Vibration and Control. 21 (16) (2015) 3123-3131.
[6] P.K. Kankar, S.C. Sharma and S.P. Harsha, Rolling element bearing fault diagnosis using wavelet transform, Neurocomputing.74 (2011) 1638-1645.
[7] W. Du, J. Yuan, Y. Li, and C. Liu, Adaptive wavelet filtering for bearing monitoring based on block bootstrapping and white noise test,Proceedings of the Institution of Mechanical Engineers. Part C: Journal of Mechanical Engineering Science.226 (9) (2012) 2345-2360.
[8] R. Yan, R.X. Gao, Base wavelet selection for bearing vibration signal analysis, International Journal of Wavelets Multiresolution and Information Processing.7(4) (2009) 411-426.
[9] M.H. Hassoun, Fundamentals of artificial neural networks, MIT Press Cambridge, Massachusetts London, England,1995.
[10] V. Vakharia, V.K Gupta and P.K. Kankar, A comparison of feature ranking techniques for fault diagnosis of ball bearing, Soft Computing. Springer.DOI 10.1007/s00500-015-1608-6 (2015a)
[11] L. Breiman, Random Forests, Machine Learning, 45 (1) (2001) 5-32.
[12] V. Vakharia, V.K Gupta and P.K. Kankar, Application of chi square feature ranking technique and random forest classifier for fault classification of bearing faults,22nd ICSV, Florence ,Italy, July 12-16,(2015b)
[13] D. S. Moore, W. I. Notz, , M. A., Fligner, R. S. Linder, The basic practice of statistics, W. F. Freeman and Company, New York, 2013.
[14] M.Hall, E.Frank, G.Holmes, B. Pfahringer, P. Reutemann. and I.H.Witten, The WEKA data mining software: An Update, SIGKDD Explorations.11 (1) (2009) 10-18.
[15] M.R. Sikonja and I. Kononenko, Theoretical and empirical analysis of ReliefF and RReliefF, Machine Learning Journal. 53 (2003) 23-69.
[16] K.A. Loparo, Bearing vibration dataset, Case Western Reserve University. Available at: www.eecs.cwru.edu/laboratory/bearing (accessed May 2015)