Technical Journal, University of Engineering and Technology (UET) Taxila, Pakistan Vol. 23 No. 3-2018
ISSN:1813-1786 (Print) 2313-7770 (Online)

Classification of Rigid Rotor Faults Using Time Domain Features Extracted from Multiple Vibration Sensors

S. Hassan[1], M. M. Tahir[2], S. Badshah[3], A. Hussain[4], N. A. Anjum[5]

[1]*Air University, Islamabad, Pakistan*
[2,3,4]*International Islamic University, Islamabad, Pakistan*
[5]*University of Engineering and Technology, Taxila, Pakistan*
[1]shahwazhassan@gmail.com

Abstract-Rotating machinery is the most common in industry. To avoid financial losses and catastrophic failures, accurate identification of rotor faults is crucial. Common rotor faults include unbalance (UF), misalignment (MF), bent (BF), eccentricity (EF) and cocked rotor (CF). Each fault is addressed through distinctive maintenance technique, and thus inaccurate identification of these faults may introduce additional problems in the machinery. Vibration-based predictive maintenance is very effective method to monitor the condition of machinery. Problem arises when traditional vibration analysis methods do not provide clear picture of the rotor faults. To address the issue, this research presents a predictive maintenance-based fault diagnostic model, which employs supervised learning-based pattern recognition (PR) method using time domain statistical time domain features. The TD features are extracted from vibration signals acquired from multiple accelerometers to capture radial and axial vibrations simultaneously. Difference of mechanical forces, exhibited by these faults on the multiple axes, provides very informative fault related TD features. Salient features are selected with the help of decision tree (DT) to be utilized by support vector machine (SVM). The proposed model provides very accurate classification of the faults, and model identifies maximum number of rotor faults reported so far. The model provides classification accuracy of 98% and outperforms the previously presented methods for the problem at hand.

Keywords-Rotor Fault Diagnosis, Supervised Learning-Based Classification, Vibration Analysis, Multi-Axes Time Domain Features, Feature Selection, Support Vector Machine.

I. INTRODUCTION

Increase in production is the main challenge for modern industry. New techniques are being introduced rapidly to enhance the productivity. To meet larger production targets, machines run continuously, and thus increase the chance of machine failures [i, ii]. In order to avoid plant's downtime, machine condition monitoring (MCM) is implemented. That is a process of monitoring one or more parameters of machine like temperature, noise or vibration to predict its potential faults early [iii, iv]. Variety of MCM methods have been used such as mechanical measurements, electrical measurements and electro-mechanical measurements [v-viii]. Different kinds of sensors are mounted on machine at prime locations [ii] to observe any unusual behavior. Most of the problems in rotating machinery arise due to rotor faults, which may lead to secondary faults such as gear faults, bearing faults, belt faults, coupling faults and motor faults [ix-xii]. Vibration based techniques are the most common for condition monitoring of rotating machinery because vibration is the earliest indicator [xiii-xv]. The most common methods include frequency domain analysis, time domain analysis, time-frequency domain analysis and current based analysis [xvi-xix]. In case of rotor faults, conventional techniques such as spectral analysis alone is not sufficient as most of the rotor faults exhibit similar sort of spectra making the faults detection process difficult [xx-xxii]. Common rotor faults are Unbalance (UF), Misalignment (MF), Bent rotor (BF), Eccentricity (EF) and Cocked rotor (CF) [xxiii-xxv].

Artificial intelligence (AI) is a very popular and important domain for computer-based rotor fault diagnosis [iv], main methods includes neural networks, data mining [xxv-xxviii], entropy and optimization techniques and fuzzy logic [xii]. The AI techniques have also been employed to identify rotor faults using TD features extracted from vibration signals [xxix, xxxix]. However, the researchers mostly used single sensor-based features extraction on various machinery faults [xl-xliv].

[xxviii] presented a method using multiple vibration sensors to extract TD features for rotor faults but the authors worked only on UF and MF. Present study not only includes common rotor faults but also enhances overall rotor fault classification capability of the multi-axes TD features.

The features include skewness, RMS, mean, kurtosis, impulse factor, crest factor, shape factor, range, standard deviation, variance and median. These statistical features are selected using DT for efficient data processing. Selected features are then fed to the multi-class SVM, which provides very accurate classification of these faults. Contributions of the research are summarized as,

- Employing multi-axes TD features such as axial and radial axis, for common rotor faults.
- The DT is used to select salient features from both axes.
- 98% is achieved using multi-class SVM.

The paper is organized in the following way. Section II briefly explains the rotor faults. Proposed methodology is elaborated in section III. Results and outcomes are discussed in section IV. Conclusions are made in section V.

II. ROTOR FAULTS

Rotor faults are often experienced by a machine due to thermal and mechanical stresses because of continuous running. Rotor faults are one of the main reason of machine failure. Misalignment generally occurs due to bent shaft, manufacturing defects, thermal growth, poor alignment and shifted foundations. In some estimate almost 50% to 70% of machine vibration problems are the result of misalignment [xxiii]. Power wastage, wear tear or permanent damage to the machine can occur due to misalignment. Angular misalignment and parallel misalignment are the common types of misalignment but this study only deals with parallel misalignment. which is shown in Fig. 1.

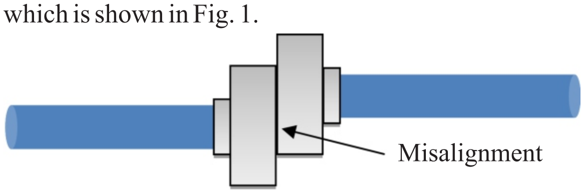

Fig. 1. Misaligned Rotor

Arc or bent present at the middle of a shaft is known as shaft bending. Major reasons of shaft bending are damage during shipment rigging, excessive heat and stress beyond elastic limit. A typical bent rotor can be seen in Fig. 2 with two loaders and fixed bearings at both ends.

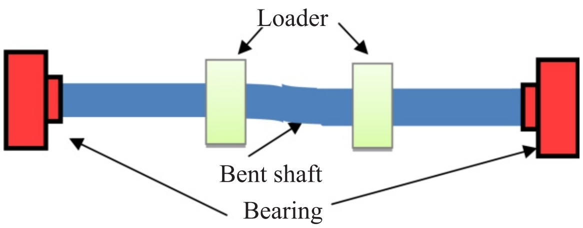

Fig. 2. Bent rotor

Unbalance in a rotating body occurs due to uneven distribution of mass around its axis of rotation. This occurs due to blow holes in casting, uneven number and position of bolt holes and parts fitted off-center etc. Unbalance occurs when center of mass (center of gravity) moves away from the center of rotation that creates centrifugal forces on the rotor. Unbalance is shown in Fig. 3.

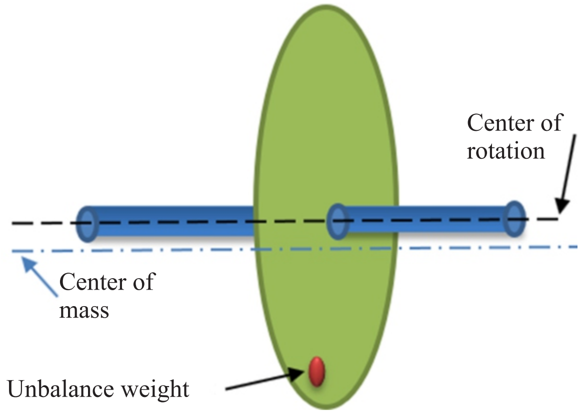

Fig. 3. Unbalanced rotor

When the stator and the rotor are not centrally aligned i.e. geometric center of stator is not same as rotor's center of rotation, the rotor is said to be in eccentric state and called eccentric rotor. Main reason of eccentricity rotor is uneven gap between rotor and stator. Fig. 4 shows that center of rotation is different from geometric center.

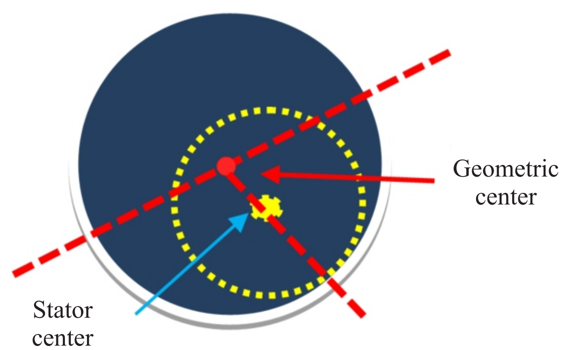

Fig. 4. Eccentric rotor

III. PROPOSED METHODOLOGY

The proposed methodology is summed up with four steps. Fig. 5 shows at first step data is acquired from vibration sensors fixed on machinery fault simulator at multiple axis. TD features are extracted from the acquired vibration data of rotor faults in the second step. Feature selection is done in the third step. At final step, fault classification is done using SVM. Details of the presented methodology are below.

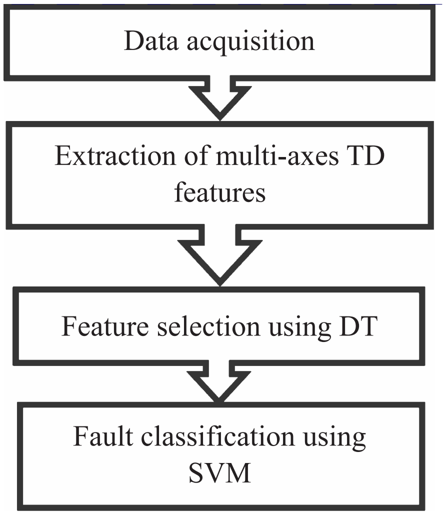

Fig. 5. Proposed methodology

A. Data acquisition

Machine fault simulator (MFS) was used to generate multiple rotor faults using its additional faulty kits [xxxii-xxxiv]. Radial and axial accelerometers were mounted on the bearings to acquire vibration signals. Piezoelectric accelerometers having sensitivity of 10.2 mv/ms-2 were mounted at outboard bearing. ER-12K bearing model was used with ¾ inch diameter shaft having different faults. One 5 kg disk was placed as loader at the middle of shaft to enhance the vibration signals [xxxv-xxxvi]. Vertical and horizontal sensors vibration data respectively. Details are shown in Fig. 6. Data acquisition is done at a sampling rate of 40K samples/sec. 60 second data at steady state motor speed of 800 rpm was acquired for each sample. Set of six rotors were used having different localized faults i.e. Healthy (HY), Bent, Eccentric, Unbalance, Misalignment and Cocked. Accelerometers were mounted on out-board bearing at axial and radial positions and tachometer was placed at motor to show RPM.

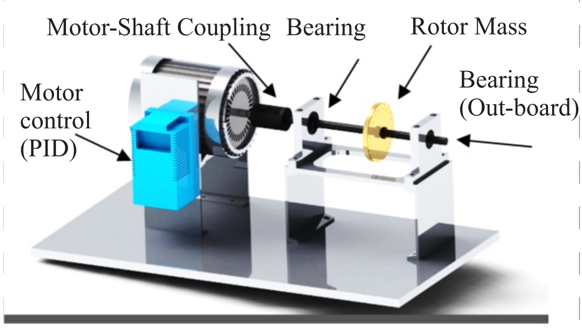

Fig. 6. Experimental setup (MFS)

In Fig. 7(a)-Fig. 7(e) frequency spectrums of all rotor faults are shown. It is shown that unbalance and eccentric rotor faults spectrums show harmonics at 1X whereas misaligned rotor, bent rotor and cocked rotor faults shows harmonics at 1X and 2X. These frequency spectrums are quiet confusing because of similar sort of pattern and it is almost impossible to classify the rotor faults with the help of these frequency spectrums.

B. Extraction of multi-axes TD features

After data acquisition, TD features were extracted. Eleven features were extracted from each signal making total of twenty-two combined features. These features are described below.

$$Mean = \frac{1}{Z}\sum_{i=1}^{Z} Y(i) \tag{1}$$

$$RMS = \left(\frac{1}{Z}\sum_{i=1}^{Z}[Y(i)]^2\right)^{\frac{1}{2}} \tag{2}$$

$$Skewness = \frac{1}{Z}\sum_{i-1}^{Z}\left(\frac{Y(i)-\mu}{\sigma}\right)^3 \tag{3}$$

$$Kurtosis = \frac{1}{Z}\sum_{i-1}^{Z}\left(\frac{Y(i)-\mu}{\sigma}\right)^4 \tag{4}$$

$$Standard\ Deviation = \left(\frac{1}{Z}\sum_{i=1}^{Z}(Y(i)-\mu)^2\right)^{\frac{1}{2}} \tag{5}$$

$$Variance = \frac{1}{Z}\sum_{i=1}^{Z}(Y(i)-\mu)^2 \tag{6}$$

$$\text{Crest factor} = \frac{\max(|Y|)}{RMS} \tag{7}$$

$$Shape\ Factor = \frac{RMS}{\frac{1}{Z}\sum_{i=1}^{Z}|Y(i)|} \tag{8}$$

$$\text{Range} = \max(Y) - \min(Y) \tag{9}$$

$$Median = magnitude\left(\frac{Z+1}{2}\right) \tag{10}$$

$$Impulse\ Factor = \frac{\max(|Y|)}{\frac{1}{Z}\sum_{i=1}^{Z}|Y(i)|} \tag{11}$$

Here Y represents the data collected from each segment and amplitude of ith sample is shown by Y(i). Z represents the total number of samples

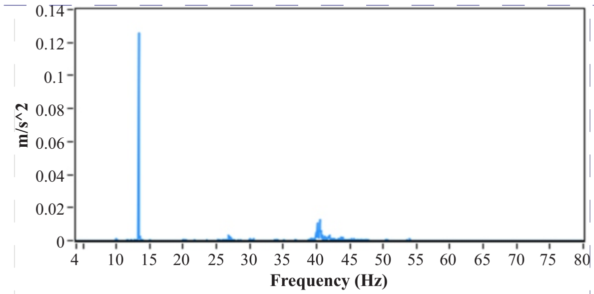

Fig. 7(a). Frequency spectrum of unbalance rotor

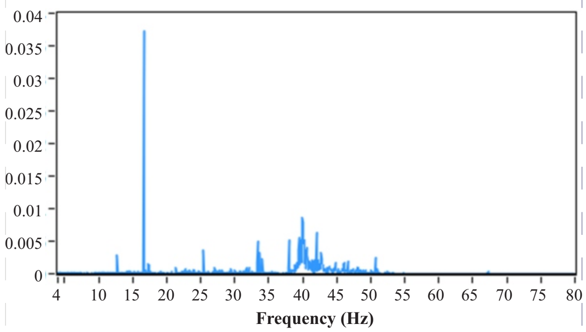

Fig. 7(b). Frequency spectrum of eccentric rotor

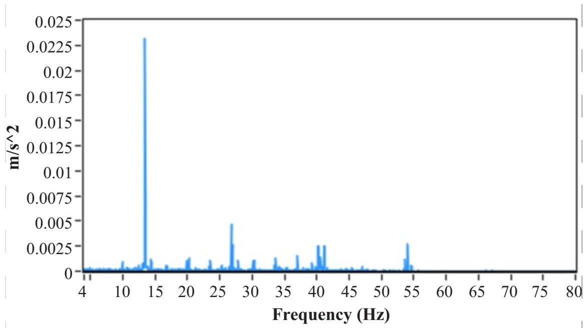

Fig. 7(c). Frequency spectrum of misaligned rotor

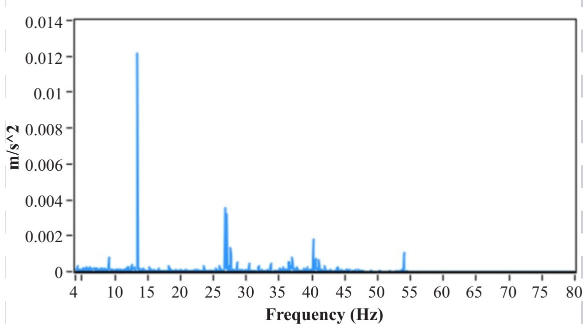

Fig. 7(d). Frequency spectrum of bent rotor

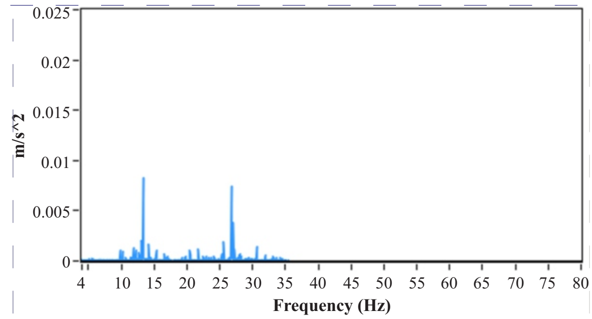

Fig. 7(e). Frequency spectrum of cocked rotor

C. Feature selection using DT

The DT is a tree-based knowledge representation methodology used to represent classification rules [xxxvii- xxxix].

It was first introduced by J. Ross Qainlan using ID3 (Iterative Dichotomizer 3) algorithm. Ross Quintan developed a simple erudition algorithm called ID3 algorithm. The ID3 algorithm builds a DT from a fixed set of data [xxxviii]. Data is described by set of attributes. Each attribute contains fixed set of values. "The advantage of DT is that a program instead of knowledge person elicits knowledge from an expert" [xxxviii, xxxix].

Id3 consists of one root with multiple branches, different number of nodes and several leaves [xxxiii-xxxvi]. Every branch (chain of nodes) starts from root and ends on a leaf and every node contains one attribute.

The importance of the associated attribute is calculated by the occurrence of any attribute in the tree [xxxv]. DT is drawn by calculating entropy and information gain for each feature and the feature. Purity and impurity of a subset is calculated by its entropy which calculates its homogeneity in the set of given examples. "Entropy is a measure of disorder in a data"or "Entropy is an indicator that indicates how messy is our data."

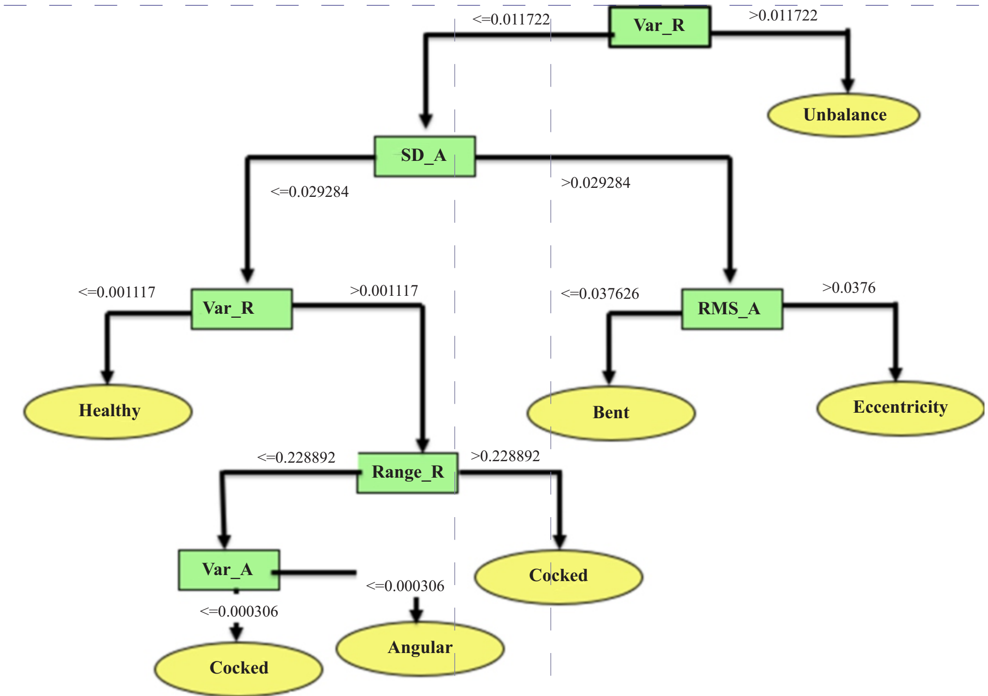

Fig. 8. Decision tree of selected features

Main focus or goal of decision tree is to tidy the given data. Randomly distributed data is given to DT and what DT tries is to separate data based on different classes. DT maximizes the purity of given data as much as possible. When given data is randomly mixed it will have maximum entropy. When DT is applied on data it tries to minimize its Entropy or its randomness.

Entropy tells us how pure or impure one subset is. Suppose we have two classes i.e. Class(A) and Class(B).The entropy will be calculated by

$$H(S) = -P(A)log_2P(A) - P(B)log_2P(B) \qquad (12)$$

Where
H(S) = Entropy
P(A) = Probability of Class (A) occurring in a subset
P(B) = Probability of Class (B) occuring in a subset

Information gain measures the expected reduction in entropy, or uncertainty. Information gain is given by

$$Gain\ (S, A) = Entropy(S) - \sum \frac{|S_W|}{S} Entropy(S) \qquad (13)$$

Set of all possible values for attribute A are given by Values(A), and subset of S for which attribute A has value v S_W = {s in S | A(s) = w} is given by S_w. The first term *Entropy(S)* is the entropy of original collection. Expected value of the entropy is derived by second value after S is portioned using first attribute [xxxiii].

The same attribute that is already been included in the path of tree or in the training examples that are associated is the leaf node all will have a same target attribute having zero entropy [xxxviii].

Decision tree algorithm is based on information gain theory that is calculated by entropy. Test features at each node are selected on the basis of information gain ratio. This is known as feature selection measure. Attribute having highest information gain is chosen as test feature for that node.

We have a data set "D" having data (D_1...D_l). If class label attribute has n different classes, C_i(for i = 1...n).

D_j is number of samples of D for class C_i

$$Splitinfo_A(D) = -\sum_{i=1}^{n} \frac{|D_j|}{|D|} \log \frac{|D_j|}{|D|} \qquad (14)$$

$$Gain\ ratio\ (A) = \frac{Gain(A)}{Splitinfo_A(D)} \qquad (15)$$

Where Gain is given by

$$Gain = Info(D) - Info_A(D) \qquad (16)$$

$$Info(D) = -\sum_{i=1}^{n} P_i \log_2(P_i) \qquad (17)$$

And

$$Info_A(D) = -\sum_{i=1}^{n} \frac{|D_j|}{|D|} Info(D_j) \qquad (18)$$

Where P_i is probability if distinct class C_i

D is data set
A is sub attribute from main attribute

$\frac{|D_j|}{|D|}$ is weight if jth partition

We can say that Gain (A) is the reduction of entropy that is reduced with the value of A feature

Algorithm

i. Data set is given as input.
ii. If it contains more than one feature apply feature selection (re-processing).
iii. Calculate entropy and information gain.
iv. Construct models
v. Find accuracy and execution time
vi. Pick model with maximum accuracy
vii. If two features have same accuracy find minimum execution time
viii. End

Entropy is calculated by different methods

Quadratic entropy was developed by T. Vajdel and it was first used by Fermi in theoretical physics

If we have discrete random variable "m" with probability P_i

$P_i \geq 0$(i=1…n)

Entropy H(X) will be given by

$$H(X) = \sum_{i=1}^{n} P_i(1 - P_i) \quad (19)$$

If we have random variable X with discrete probability

P(i)= $P_1, P_2 \ldots P_m$

$$H(X) = -\sum_{i=1}^{n} P_i \log_2 P_i \quad (20)$$

DT is drawn until a leaf is reached. It is used to generate DT from datasets. C45 is the next version of ID3 algorithm. C45 includes acceptance of both continuous and discrete values, handles incomplete data points and solves over fitting problem by bottom up technique also known as "Pruning". J48 is the open source JAVA implementation of C4.5 algorithm that is implemented using WEKA (tool used for construction of DT).

The formation of DT and using it for feature selection is based on the following.

i. The set of features is fed as an input and result is a DT
ii. Different classes are labeled by DT leaf nodes and nodes related with other class are classified
iii. Tree branches are associated with possible values of feature nodes from which they are originated
iv. Classification is done by moving from the root through nodes until the leaf node is reached. Leaf node is the final classification of the fault.
v. Most useful feature is selected by entropy calculation and information gain. The feature that reduces the entropy is most useful.

The DT is a graphical representation of features showing their importance for fault classification. In Fig. 8 DT is shown, providing the details of selected features and their effect on fault classification.

D. *Fault classification*

Fault classification was done by using SVM, which was given labeled inputs of feature vector. The faults classes are separated with the help of a hyper-plane. The hyper-plane is constructed to separate the two classes and the binary SVM can be further enhanced for multiple class problems. The labeled set is given as input and the SVM creates an optimal hyper-plane using support vectors. Separation margin is maximized between the classes. The process of maximizing separation plane creates two hyper planes parallel to separation plane on both sides. The generalization error (GE) is minimized. With low value of GE machine can correctly classify [xxvi].

These separation planes are created with single or multiple points. Separation planes are also known as boundary planes which are separated by distance called margin. Main purpose of SVM is to find a hyper-plane that can maximize the margin and minimize the GE [xxvi]. Standard SVM classifier is shown in the Fig. 9.

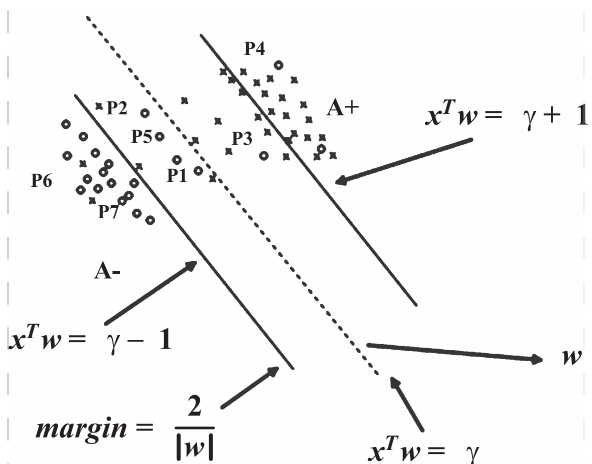

Fig. 9. SVM classifier

There are two classes A- and A+. that are to be classified. Different data points are given i.e. P1, P2, P3, P4, P5, P6 and P7. The data set P1-P5 are support vectors that belong to A- but P6 and P7 are not. Same is for the case of A+ class. These support vectors are very important for SVM. In case of formulation and D is a (M X 1) matrix that represents +1 And -1 class labels. Whereas

e is vectors of ones
A is a (m X n) matrix
x is given training vector
v is controlling parameter

w is orientation parameter

Artificial neural networks (ANN) are widely used for classification purposes and they have proven to be good classifiers. Only drawback of ANN is that they require large number of samples for training. That is not possible in every case. Support vector machine (SVM) works on statistical learning theory and it can also work on small number of samples. SVM gives improved generalization results as compared to ANN and its global optimization solution is same as obtain by ANN. In past few years SVM is efficiently working in real world application, especially in fault diagnosis [xl-xliv]. SVM have been used by many researchers for fault diagnosis in recent times [xl-xliv].

SVM is developed by the optimal separation plane by using linearly separable conditions. Basic SVM principle is shown in Fig. 9. Two different classes of data class A (circle) and class B (cross) are given to SVM. SVM tries to create a linear boundary H between the two classes in such a way that the margin between the classes is maximized, namely, the distance between the boundary and the nearest data point in each class is maximized. The points near to data define the margin and they are known as support vectors. Given training data sets for training $G = \{(x_i, y_i)\}(i = 1...n)$ each sample $x_i \epsilon R^n$ denotes the input vector $y_i \epsilon \{+1, -1\}$ Here n denotes the number of training data sets. The boundary is expressed as $w.x + b = 0$

Where ω denotes weight vector and b denote biased term. So the regression function to classify any data point in either class A or class B is given by

$$f(x) = w.x + b \tag{21}$$

ω and b can be gained by minimizing the regularized risk function.

$$R(C) = C\frac{1}{n}\sum_{i=1}^{n} L_c(y) + \frac{1}{2}\|w\|^2 \tag{22}$$

$$L_c(y) = \begin{bmatrix} |f(x) - y| - \epsilon & |f(x) - y| \geq \epsilon \\ 0 & |f(x) - y| < \epsilon \end{bmatrix}$$

C is the cost function measuring empirical risk $\frac{\|w\|^2}{2}$ denotes Euclidian norm. The ϵ- insensitive loss function is implemented to stabilized the estimation.

a_i and a_i* are introduced as lagrangian multipliers that satisfy the equation.

$a_i . a_i* = 0$

$a_i \geq 0$

$a_i* \geq 0$

Optimization problem is solved using following lagrangian form

$$Min - \sum_{i-1}^{n} y_i(a_i - a_i^*) - \epsilon\sum_{i-1}^{n}(a_i + a_i^*) + \frac{1}{2}\sum_{i,j}^{n} y_i(a_j - a_j^*)(a_i - a_i^*)K(x_i, x_i^*) \tag{23}$$

$$s.t \sum_{i=1}^{n}(a_i - a_i^*) = 0$$

Where $a_i, a_i^* \epsilon[0, C]$

$$f(x) = \sum_{i=1}^{n} y_i(a_i - a_i^*)(x_i, x_i^*) + b \tag{24}$$

When the linear boundary in space "**S**" is not enough to separate the planes, a hyper plane is created which allows line separation in hyper plane dimension. In case if SVM we achieve it by using a transformation ø(*x*) which is used to map data from input space to feature space. A kernel function is introduced

$$K(x,y) = ø(x) - ø(y) \tag{25}$$

This kernel function is used to obtain regression function.

$$f(x) = \sum_{i=1}^{n} y_i(a_i - a_i^*)K(x_i, x_i^*) + b \tag{26}$$

The value of error minimization is defined by *v.* γ is location parameter and it is relative to origin.

$$ve'y + \frac{1}{2}w'w \tag{27}$$

$$s.t. D\ (Aw - e\gamma) + y\ \geq e$$

$$y \geq o$$

Where, A∈$R^{m \times n}$, D ∈ {-1,+1}$^{m\times l}$, e=1$^{m\times l}$

The value of error minimization is defined by *v.* γ is location parameter and it is relative to origin.

$$ve'y + \frac{1}{2}w'w \tag{28}$$

$$s.t. D\ (Aw - e\gamma) + y\ \geq e$$

$$y \geq o$$

Where, A∈$R^{m \times n}$, D ∈ {-1,+1}$^{m\times l}$, e=1$^{m\times l}$

Once the training is done and the machine is ready for classification. Classification is done by using the decision function

$$f(x) = sign\ (w^T x - \gamma) \tag{29}$$

If f(x) have a positive value the new features will belong to A+ class otherwise it will belong to A- Class. The interested reader can find the details of SVM mathematics and algorithm at [xxvi].

After the supervised learning process 240 instances were fed to the machine 40 for each fault. Fault classification process is explained in Fig. 10.

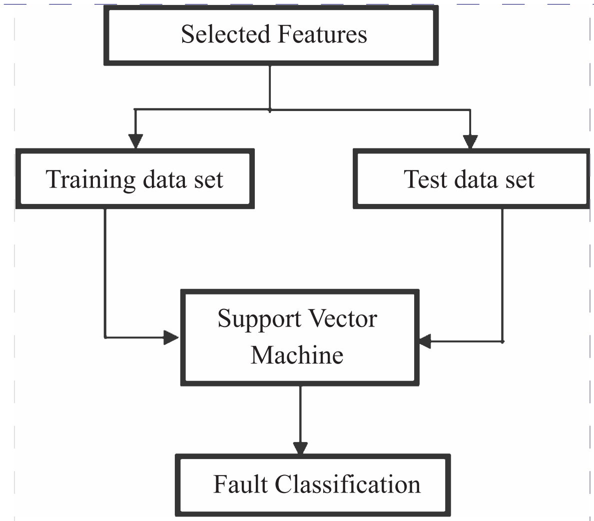

Fig. 10. Fault classification process

IV. RESULTS AND DISCUSSION

Due to alike frequency spectra produce by rotor faults, their identification becomes very hard. Each fault has its own treatment method. For example, in case of unbalance a weight is attached or removed from the rotor and bent rotor is fixed by applying a force at the bent area of the shaft. Whereas in case of misalignment shaft coupling or structure needs to be adjusted. Same goes with other faults. So, if the fault is wrongly classified and wrong procedure is implemented on the rotor then it can lead to catastrophic failures. To avoid all these problems, this study not only enhanced the rotor fault diagnosis but also included the fault diagnosis of all common rotor faults with the help of selected features. Whereas previous studies only gave us classification of Unbalance and misalignment [xxviii].

Reason for placing sensors at multi-axes is that every rotor fault generates almost distinct forces. In case of EF and BF force acts in axial direction and BF create force in Radial direction while MF creates two forces one in axial and other in radial directions. Total 11 features were extracted from vibration data acquired from sensors placed at each axis simultaneously. Making total of 22 features. These 22 features were fed into the DT for feature selection. Firstly, Training data sets with labels are given to the DT with cross validation of 10 folds. Each fault is separately given to DT with labels for machine learning. After that unlabeled data is given to the DT for accurate classification and DT selected only 5 features from 22 features that that have the major impact on fault classification. The selected features are Variance (Radial), RMS(Axial), Standard deviation (Axial), Variance (Axial) and Range(Radial). Table I shows the results.11 features of Radial axis sensor were taken it gave 71% classification accuracy and Axial sensor gave 80% accuracy. When machine is fed with 22 combined features from both sensors it gave 97% accuracy and using 5 salient features from SVM outcome accuracy reached 98%.40 randomly mixed trial sets for each fault are fed in to the machine making total of 240 trial sets. SVM gave 235 correct classification results and 5 wrong classifications. As shown in Table II.

TABLE I
SVM CLASSIFICATION ACCURACY WITH RADIAL, AXIAL, MULTI-AXES AND MULTI-AXIS SELECTED TD FEATURES

	Radial	Axial	Multi-axis all features	Multi-axis Selected features
Features	11	11	22	5
Accuracy	71%	80%	97%	98%

In the table Unbalance fault is only once confused with cocked fault, Eccentric fault is only three times confused with unbalance, healthy and misalignment faults. Whereas cocked fault is only once confused with healthy rotor. Misalignment fault is only once confused with unbalance fault. Selected 5 features gave 98% accuracy. So, the accuracy of SVM is enhanced by using multi-axis TD features In the previous research [xl-xliv] used statistical time domain features for processing of bearing and rotor faults. The author used feature processing for accurate rotor fault classification whereas this research uses feature selection instead of feature processing with the help of DT. With the help of feature processing [xxviii] reduced to 5 features for two rotor faults. This study utilized only 5 features to classify 6 rotor faults with 98% accuracy.

TABLE II
CONFUSION MATRIX

Fault	HY	MF	EF	CF	UF	BF
BF	0	0	0	0	0	40
UF	0	0	0	1	39	0
CF	1	0	0	39	0	0
EF	0	1	38	0	1	0
MF	0	39	0	0	1	0
HY	40	0	0	0	0	0

V. CONCLUSION

In rotor fault diagnosis, it is crucial to confirm the fault before applying appropriate predictive maintenance method. Wrong prediction can lead to more machine problems. Rotor fault problem is not easy to handle using conventional spectral analysis methods. The present study explores that most of the introduced rotor faults show similar sort of frequency behavior. In order to solve the problem, this research used the dynamic behavior of these faults i.e. every rotor fault creates its own forces that may be in axial, radial or in both directions. Combining the multi-axes TD features improves the fault classification accuracy of the multi-class SVM. The SVM provides excellent identification of the faults. Six rotor faults are identified with the accuracy of 98% with the help of only five TD features. This research can enhance the accuracy of predictive maintenance for rotating machinery industries, by applying this technique catastrophic machine failure can be avoided, that will reduce machine down time.

REFERENCES

[i] T. Renwick and P. E. Babson, "Vibration analysis-a proven techniques as a productions maintenance tool," IEEE Transactions on Industry Applications, vol. 21, no. 2, pp. 324–332, 1985.

[ii] Loparo, K. A., Adams, M., Lin, W., Abdel-Magied, M. F., and Afshari, N.: Fault detection and diagnosis of rotating machinery, IEEE Transactions on Industrial Electronics, vol. 47, pp. 1005-1014, 2002.

[iii] P. Jayaswal, A. Wadhwani, and K. Mulchandani, "Machine fault signature analysis," International Journal of Rotating Machinery, vol. 2008,2008.

[iv] R. B. Randall, Vibration-based condition monitoring : industrial, aerospace and automotive applications, 2011.

[v] S. Edwards, A. W. Lees, and M. I. Friswell, "Fault diagnosis of rotating machinery," pp. 4–13, 1988.

[vi] S. Singh and N. Kumar, "Rotor faults diagnosis using artificial neural networks and support vector machines," International Journal of Acoustics and Vibration (Accepted for publication), 2014.

[vii] Darpe, K. Ashish, "A novel way to detect transverse surface crack in a rotating shaft", Journal of sound and vibration, 305(1-2), p. 151-171, Aug. 2007.

[viii] Al-Ghamd, Mba, "A comparative experimental study on the use of acoustic emission and vibration analysis for bearing defect identification and estimation of defect size", Mechanical Systems and Signal Processing, vol. 20, pp. 1537-1571, 2006.

[ix] R. Isermann, "Model-based fault-detection and diagnosis-status and applications", Annual Reviews in control, vol. 29, pp. 71-85, 2005.

[x] J. Bangura, and N. Demerdash, "Diagnosis and characterization of effects of broken bars and connectors in squirrel-cage induction motors by a time-stepping coupled finite element-state space modeling approach", IEEE Transactions on Energy Conversion, vol. 14, pp. 1167-1176, 1999.

[xi] R. Isermann, and P. Balle, "Trends in the application of model-based fault detection and diagnosis of technical processes", Control Engineering Practice, vol. 5, pp. 709-719, 1997.

[xii] M. Saimurugan, K. I. Ramachandran,, "Automatic rule learning using decision tree for fuzzy classifier in fault diagnosis of roller bearing" , Mechanical Systems and Signal Processing 21 (2007) 2237–2247 science direct.

[xiii] P. Jayaswal, A. K. Wadhwani, and K. B. Mulchandani, "Machine Fault Signature Analysis: Review Article", International Journal of Rotating Machinery, vol. 2008, pp. 1155-1165, 2008.

[xiv] J. J. Carbajal-Hern´andez, L. P. S´anchez-Fern´andez, I. Hern´andez- Bautista, J. d. J. Medel-Ju´arez, and L. A. S´anchez-P´erez, "Classification of unbalance and misalignment in induction motors using orbital analysis and associative memories," Neurocomputing, vol. 175, pp. 838–850, 2016.

[xv] C. Alarcon, J. A. Daviu, M. R. Guasp, R. Puche Panadero, and L. Escobar, "Application of the wigner–ville distribution for the detection of rotor asymmetries and eccentricity through high order harmonics", Electric Power Systems Research, vol. 91, pp. 28–36, 2012.

[xvi] R. Kechida, A. Menacer, and H. Talhaoui, "Approach signal for rotor fault detection in induction motors," Journal of failure analysis and prevention, vol. 13, no. 3, pp. 346–352, 2013.

[xvii] K. Kim and A. G. Parlos, "Induction motor fault diagnosis based on neuropredictors and wavelet signal processing," IEEE/ASME transactions on mechatronics, vol. 7, no. 2, pp. 201–219, 2002.

[xviii] D. Liu, Y. Zhao, B. Yang, and J. Sun, "A new motor fault detection method using multiple window s-method time-frequency analysis," in Systems and Informatics (ICSAI), 2012 International Conference on. IEEE, 2012, pp. 2563–2566.

[xix] R. R. Obaid, T. G. Habetler, and R. M. Tallam, "Detecting load unbalance and shaft misalignment using stator current in inverter-driven induction motors," in Electric Machines and Drives Conference, 2003. IEMDC'03. IEEE International, vol. 3. IEEE, 2003, pp. 1454–1458.

[xx] D. Yang, "Induction motor bearing fault diagnosis using hilbert-based bispectral analysis," in Computer, Consumer and Control

(IS3C), 2012 International Symposium on. IEEE, 2012, pp. 385–388.

[xxi] D. Zhen, T. Wang, F. Gu, and A. Ball, "Fault diagnosis of motor drives using stator current signal analysis based on dynamic time warping," Mechanical Systems and Signal Processing, vol. 34, no. 1, pp. 191–202, 2013.

[xxii] C. E. Kim, Y. B. Jung, S. B. Yoon, and D. H. Im, "The fault diagnosis of rotor bars in squirrel cage induction motors by time stepping finite element method", IEEE Transactions on Magnetics, vol. 33, pp. 2131-2134, 1997.

[xxiii] Machinery fault diagnosis, A basic guide to understand vibration analysis for machine diagnosis

[xxiv] I. Howard, S. Jia, and J. Wang, "The dynamic modelling of a spur gear in mesh including friction and a crack", Mechanical Systems and Signal Processing, vol. 15, pp. 831-853, 2001.

[xxv] F. Zidani, M. E. H. Benbouzid, D. Diallo, and M. S. Na¨ıt-Sa¨ıd, "Induction motor stator faults diagnosis by a current concordia pattern based fuzzy decision system," IEEE Transactions on Energy Conversion, vol. 18, no. 4, pp. 469–475, 2003.

[xxvi] M. Saimurugan, K. I. Ramachandran, V. Sugumaran, and N. R. Sakthivel, "Multi component fault diagnosis of rotational mechanical system based on decision tree and support vector machine," Expert Systems with Applications, vol. 38, no. 4, pp. 3819–3826, 2011.

[xxvii]A. K. Jalan and A. Mohanty, "Model based fault diagnosis of a rotor–bearing system for misalignment and unbalance under steady-state condition," Journal of Sound and Vibration, vol. 327, no. 3, pp. 604–622,2009.

[xxviii]M. MasoodTahir and Ayyaz Hussain, Saeed Badsha "Classification of Unbalance and Misalignment Faults in Rotor using Multi-Axis Time Domain Features", 978-1-5090-3552-6/16/2016 IEEE.

[xxix] H. P. Bloch and F. K. Geitner, in Machinery Failure Analysis and Troubleshooting, chapter 5, Gulf Publishing Company, Houston, Tex, USA, 1983.

[xxx] S. Nandi and H. A. Toliyat, "Condition monitoring and fault diagnosis of electrical machines-a review," in Proceedings of the 34th IEEE IAS Annual Meeting on Industry Applications Conference, vol. 1, pp. 197–204, Phoenix, Ariz, USA, October 1999.

[xxxi] Z. Li, J. Zhu, X. Shen, C. Zhang, and J. Guo, "Fault diagnosis of motor bearing based on the bayesian network," Procedia Engineering, vol. 16, pp. 18–26, 2011.

[xxxii]http://spectraquest.com/machinery-fault-simulator/details/mfs/

[xxxiii]R. Olivas, "Decision Trees – A primer for Decision making Professionals", Rev. 5, 2007.

[xxxiv]O. Maimon, L. Rokach, "Data Mining and Kowledge Discovery Handbook", Second Edition, ISBN 978-0-387-09822-7, Springer Science+Business Media, pp. 149-174, 2009.

[xxxv]J. R. QUINLAN, "Induction of Decision Trees", Centre for Advanced Computing Sciences, New South Wales Institute of Technology, Sydney 2007, Australia.

[xxxvi]K. Madadipouya, "A new decision tree method for data mining in medicine", Advanced Computational Intelligence An International Journal (ACII), Vol.2, No.3, pp. 31-37, July 2015.

[xxxvii]I. Guyon, & A. Elisseeff, "An introduction to variable and feature selection" , Journal of Machine Learning Research, vol. 3, pp. 1157-1182, 2003.

[xxxviii]M. Kudo, J. Sklansky, "Comparison of algorithms that select features for pattern classifiers", Pattern Recognition, vol. 33, Issue 1, pp. 25-41, 2000.

[xxxix]Correlation-based Feature Selection for Machine Learning Mark A. Hall This thesis is submitted in partial fulfilment of the requirements for the degree of Doctor of Philosophy at The University of Waikato. April 1999 Mark A. Hall (thesis)

[xl] M. M. Tahir, A. Q. Khan, N. Iqbal, A. Hussain, S. Badshah, "Enhancing Fault Classifcation Acuuracy of Ball Bearing using Central Tendency based Time Domain Features", IEEE Access, Vol.5, PP. 72 - 83, 2016.

[xli] M. M. Tahir, A. Hussain, S. Badshah, Q. Javaid "Rule-based Identifcation of Bearing Faults using Central Tendency of Time Domain Features",Journal of Engineering and Applied Sciences, Vol.35. No. 2, 2016.

[xlii] M. M. Tahir, A. Hussain, S. Badshah, "Enhancing Classifcation Accuracy of Ball Bearing Faults using Statistically Processed Features", IEEE InternationalConference on Intelligent Systems Engineering (ICISE), Islamabad, January 2016.

[xliii] M. M. Tahir, A. Hussain, S. Badshah, A. Q. Khan, N. Iqbal,"Classifcation of Unbalance and Misalignment Faults in Rotor using Multi-Axis Time Domain Features", IEEE International Conference on Emerging Technologies (ICET), Islamabad, October 2016.

[xliv] M. M. Tahir, A. Hussain, S. Badshah, "Accurate Extraction of Time Domain Features for Reliable Classification of Ball Bearing Faults", International Journal of Advanced and Applied Sciences, Vol. 5. No. 1, PP. 156-163, 2018.