Implementation of Rule based Classifiers for Wind Turbine Blade Fault Diagnosis Using Vibration Signals

A. Joshuva, R. Vishnuvardhan, G. Deenadayalan,
R. Sathishkumar, S. Sivakumar

Abstract--- The fast developing wind industry has revealed a requirement for more multifaceted fault diagnosis system in the segments of a wind turbine. "Present wind turbine researches concentrate on enhancing their dependability quality and decreasing the cost of energy production, especially when wind turbines are worked in off-shore places. Wind turbine blades are ought to be an important component among the other basic segments in the wind turbine framework since they transform dynamic energy of wind into useable power and due to environmental conditions, it get damage often and cause lack in productivity. The main objective of this study is to carry out a fault identification model for wind turbine blade using a machine learning approach through vibration data to classify the blade condition. Here five faults namely, blade bend, hub-blade loose connection, blade cracks, blade erosion and pitch angle twist have been considered. Machine learning approach has three steps namely feature extraction, feature selection and feature classification. Feature extraction was carried out by statistical analysis followed by feature selection using J48 decision tree algorithm. Feature classification was done using twelve rule based classifiers using WEKA. The results were compared with respect to the classification accuracy and the computational time of the classifier."

Keywords--- Condition Monitoring, Wind Turbine Blade, Statistical Features, Rule based Classifiers, Vibration Signals.

I. INTRODUCTION

Currently, wind is one of the renewable energy sources which has acquired enormous consideration in the energy market to moderate the frequently increasing universal necessity of fossil fuels and consequent disquiets about ecological issues. Though, bringing down the cost of wind energy production is a fundamental strategy for the development of wind energy industry in the subsequent eras. In such manner, the future of wind energy industry drives by means of bigger and more flexible wind turbines in remote areas, which are progressively offshore to benefit stronger and more unvarying wind conditions. Thus, both the size and locality factors becomes probably the most important factor and prompt to extended maintenance challenges [1]. In addition, high machine-driven stress is enforced on wind turbines because of extremely operational conditions and repeatedly varying loads. This high level of mechanical stress needs a high level of maintenance support while wind turbines are exposed to high dependability and accessibility necessities. A favourable (machine learning) approach is added to provide a fault detection and fault diagnosis system effectively [2].

Many research work have being carrying out on fault identification on wind turbine blade. To name a few, Kusiak and Verma [3] done a work on a data-driven approach for monitoring blade pitch faults in wind turbines using machine learning approach.

They conducted the study using bagging, artificial neural network (ANN), pruning rule based classification tree (PART), K-nearest neighbor (K-NN) and genetic programming (GP).

They obtained the accuracy of about GP-74.7%, Bagging-72.5%, PART-75.5%, ANN-76.2%, K-NN-73.5%. Bindi Chen *et al.,* [4] conducted an experiment on wind turbine pitch faults prognosis using a-priori knowledge based adaptive neuro-fuzzy inference system (ANFIS) using SCADA data and obtained 88.30% classification accuracy.

A comparative study on wind turbine power coefficient estimation by soft computing methodologies was carried out by Shamshirband *et al.,* [5].

In this study they used support vector regression (radial basis function), support vector regression (polynomial), ANFIS (adaptive neuro-fuzzy inference system), NN (neural network) algorithms for comparison. Correlation Coefficient of algorithms where found to be SVR (RBF)-0.997, SVR (Polynomial)-0.504, ANFIS-0.978, NN-0.922.Mark Mollineaux *et al.,* [6] have done a work on damage detection methods on wind turbine blade testing with wired and wireless accelerometer sensors using benchmark data and autoregressive moving average (ARMA) and Continuous Wavelet Transform (CWT) used as modelling techniques.

A study on wavelet transform based stress and time history editing of horizontal axis wind turbine blades was carried out by Pratumnopharat *et al.,* [7].

They used time correlated fatigue damage (TCFD), mexican hat wavelet (Mexh), meyer wavelet (Meyr), daubechies 30th order (DB30), morlet wavelet (Morl), discrete meyer wavelet (Dmey) for the classification of crack on blade.

The accuracy they found to be TCFD-89.82%, Morl-80.34%, Meyr-79.76%, Dmey-80.30%, Mexh-79.23%, DB30-80.81%.

A. Joshuva*, Department of Mechanical Engineering, Hindustan Institute of Technology and Science, Chennai, T.N, India.

R. Vishnuvardhan, Department of Mechatronics Engineering, Sri Krishna College of Engineering and Technology, Coimbatore. T.N, India.

G. Deenadayalan, Department of Mechanical Engineering, Hindustan Institute of Technology and Science, Chennai, T.N, India.

R. Sathishkumar, Department of Automobile Engineering, Hindustan Institute of Technology and Science, Chennai, T.N, India.

S. Sivakumar, Department of Mechanical Engineering, Hindustan Institute of Technology and Science, Chennai, T.N, India.

320

IMPLEMENTATION OF RULE BASED CLASSIFIERS FOR WIND TURBINE BLADE FAULT DIAGNOSIS USING VIBRATION SIGNALS

Simon Hoell and Piotr Omenzetter [8] carried out a structural damage detection in wind turbine blades based on time series representations of dynamic responses using vibration data and cross-correlations, principal component analysis (PCA), genetic programming (GP) as the diagnostic algorithm.

Damir*et. al*[9] has done a numerical models for robust shape optimization of wind turbine blades using 3D geometric modeller. A computational framework for the shape optimization of wind turbine blades is developed for variable operating conditions specified by local wind speed distributions.

A classification of operating conditions of wind turbines for a class-wise condition monitoring strategy study was done by Jong *et al.,* [10].

This paper presents a general method that can be used to classify the operating conditions of wind turbine in terms of rotor speed and power.

This study used empirical probability density functions based method and Gaussian mixture model (GMM) based method. This paper presents performance evaluation of the proposed class-wise condition monitoring strategy using vibration signals.

Numerous works were carried out using simulation analysis; however only few experimental analysis were performed for wind turbine blade condition monitoring. Machine learning technique was considered for wind turbine blade fault diagnosis; however, the usage was limited in literature [11].

A very limited set of defects were considered for analysis. This is especially true in case of fault diagnosis of wind turbine blade.

This study makes an attempt to find five different blade fault conditions by applying machine learning approach and statistical analysis. Figure 1 shows the methodology of the work done. The contribution of the present study,

- This study considers five faults (blade crack, erosion, hub-blade loose connection, pitch angle twist and blade bend) for wind turbine blade fault diagnosis.
- Statistical feature extraction tool was used to extract the required features from the vibration signals.
- J48 decision tree algorithm was used for feature selection.
- This problem is modeled as a multiclass classification problem and attempts to classify using rule based machine learning classifiers.

The rest of the paper is organized as follows. In section 2, the experimental setup and experimental procedure are explained.

Section 3 presents the feature extraction process using statistical analysis.

The feature selection using J48 decision tree algorithm is presented in section 4. In section 5, the rule based classifiers used in the study were explained in detail. The results obtained from the classifiers and the discussion about their performance are presented in section 6. Conclusions are presented in the final section (section 7).
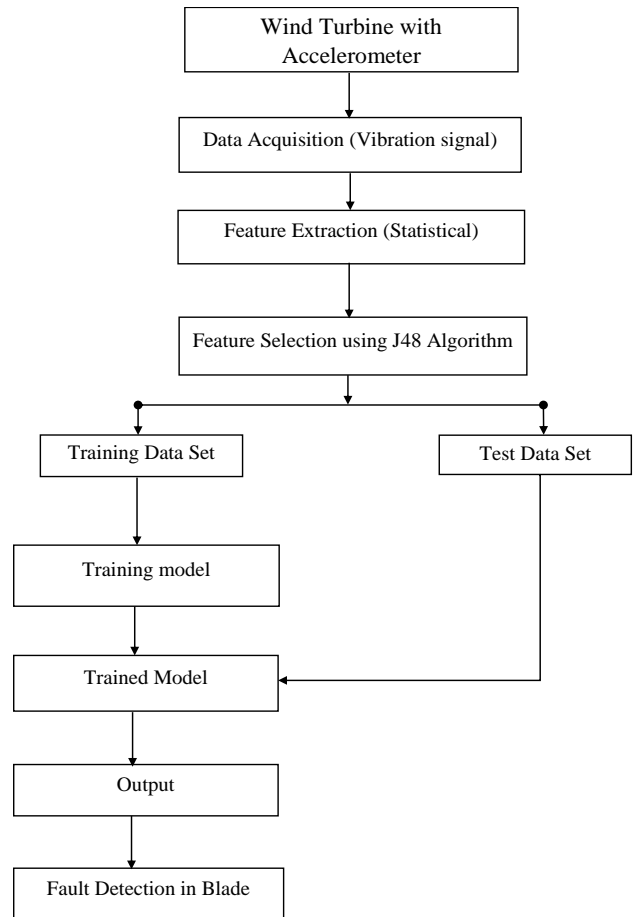

Figure 1: Methodology

II. EXPERIMENTAL STUDIES

The main aim of this study is to identify whether the blades are in good condition or in defective condition. If it is defective, then the objective is to deduce the condition of fault. Referring to Figure 1, the first two blocks are described in the following sub sections, namely experimental setup and experimental procedure. The study was conducted on a test rig resting on a stationary stand [12].

2.1. Experimental Setup

The experiment was carried out on a 50W, 12V variable wind turbine (MX- POWER, model: FP-50W-12V). The technical parameters of a wind turbine are given in Table 1.

Table 1: Technical parameters of wind turbine

Model	FP-50W-12V
Rated Power	50W
Rated Voltage	12V
Rated Rotating Rate	850r/m
Start-up Wind Speed	2.5m/s
Cut-in Wind Speed	3.5m/s
Cut-out Wind Speed	15m/s
Security Wind Speed	40m/s
Rated Wind Speed	12.5m/s
Engine	Three-phase permanent magnet generator
Rotor Diameter	1050mm
Blade Material	Carbon fiber reinforced plastics

The wind turbine was mounted on a fixed steel stand in-front of the open circuit wind tunnel outlet. The wind tunnel speed ranges from 5m/s to 15 m/s and act as a wind source to start the wind turbine. The wind speed was varied continuously in order to simulate the environmental wind condition. Experimental setup is shown in Figure 2. Piezoelectric type accelerometer was used as transducer for acquiring vibration signals. It has high-frequency sensitivity for detecting faults. Hence accelerometers are widely used in condition monitoring. In this case, a uniaxial accelerometer of 500g range, 100 mV/g sensitivity, and resonant frequency around 40 Hz was used.

The piezoelectric accelerometer (DYTRAN 3055B1) was mounted on the nacelle near to the wind turbine hub to record the vibration signals using an adhesive mounting technique. It was connected to the DAQ system through a cable. The data acquisition system (DAQ) used was NI USB 4432 model. The card has five analog input channels with a sampling rate of 102.4 kilo samples per second with 24-bit resolution.

The accelerometer is coupled to a signal conditioning unit which consists of an inbuilt charge amplifier and an analogue-to digital converter (ADC). From the ADC, the vibration signal was taken. These vibration signals were used to extract features through feature extraction technique. One end of the cable is plugged to the accelerometer and the other end to the AIO port of DAQ system. NI – LabVIEW was used to interface the transducer signal and the system (PC).
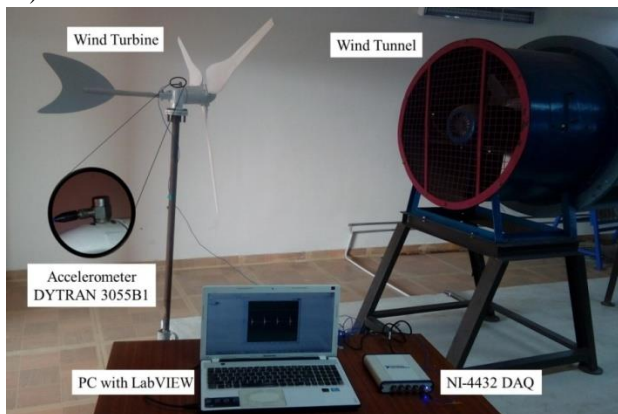

Figure 2: Experimental Setup

2.2. Experimental Procedure

In the present study, three-blade variable horizontal axis wind turbine (HAWT) was used. Initially, the wind turbine considered was in good condition (free from defects, new setup) and the signals were recorded using the accelerometer. These signals were recorded with the following specification:

1. **Sample length:** The sample length was chosen long enough to ensure data consistency; and also the following points were considered. Statistical measures are more meaningful, when the number of samples is sufficiently large. On the other hand, as the number of samples increases the computation time increases. To strike a balance, sample length of 10000 was chosen.

2. **Sampling Frequency:** The sampling frequency should be at least twice the highest frequency contained in the signal as per Nyquist sampling theorem. By using this theorem sampling frequency was calculated as 12 kHz (12000Hz).

3. **A number of signal samples:** Minimum of 100 (hundred) signal samples were taken for each condition of the wind turbine blade and the vibration signals were recorded by using NI LabVIEW.

The vibration signals are acquired using DAQ. Data acquisition (DAQ) is the process of converting analog sampling signals to digital numeric values that can be manipulated by a computer. DAQ hardware is used hereto interface between the sensor signal and a PC. The following faults were simulated one at a time on a blade while other blades remain in good condition and the corresponding vibration signals were acquired. Figure 3 shows the different blade fault conditions which are simulated on the blade.
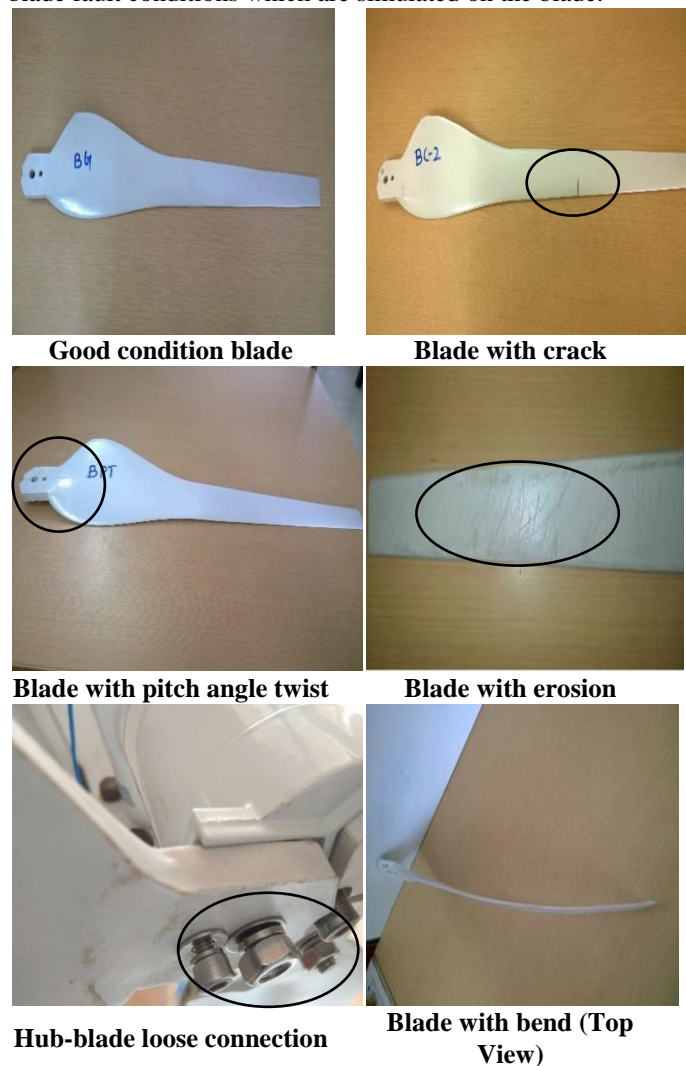

Good condition blade	Blade with crack
Blade with pitch angle twist	Blade with erosion
Hub-blade loose connection	Blade with bend (Top View)

Figure 3: Various blade fault conditions

a) **Blade bend (BB):** This fault occurs due to high-speed wind and complex forces caused by the wind. The blade was made to flap wise bend with 10^0 angle.

b) **Blade crack (BC-2):** This occurs due to foreign object damage on blade while it is in operating condition. On blade, 15mm crack was made.
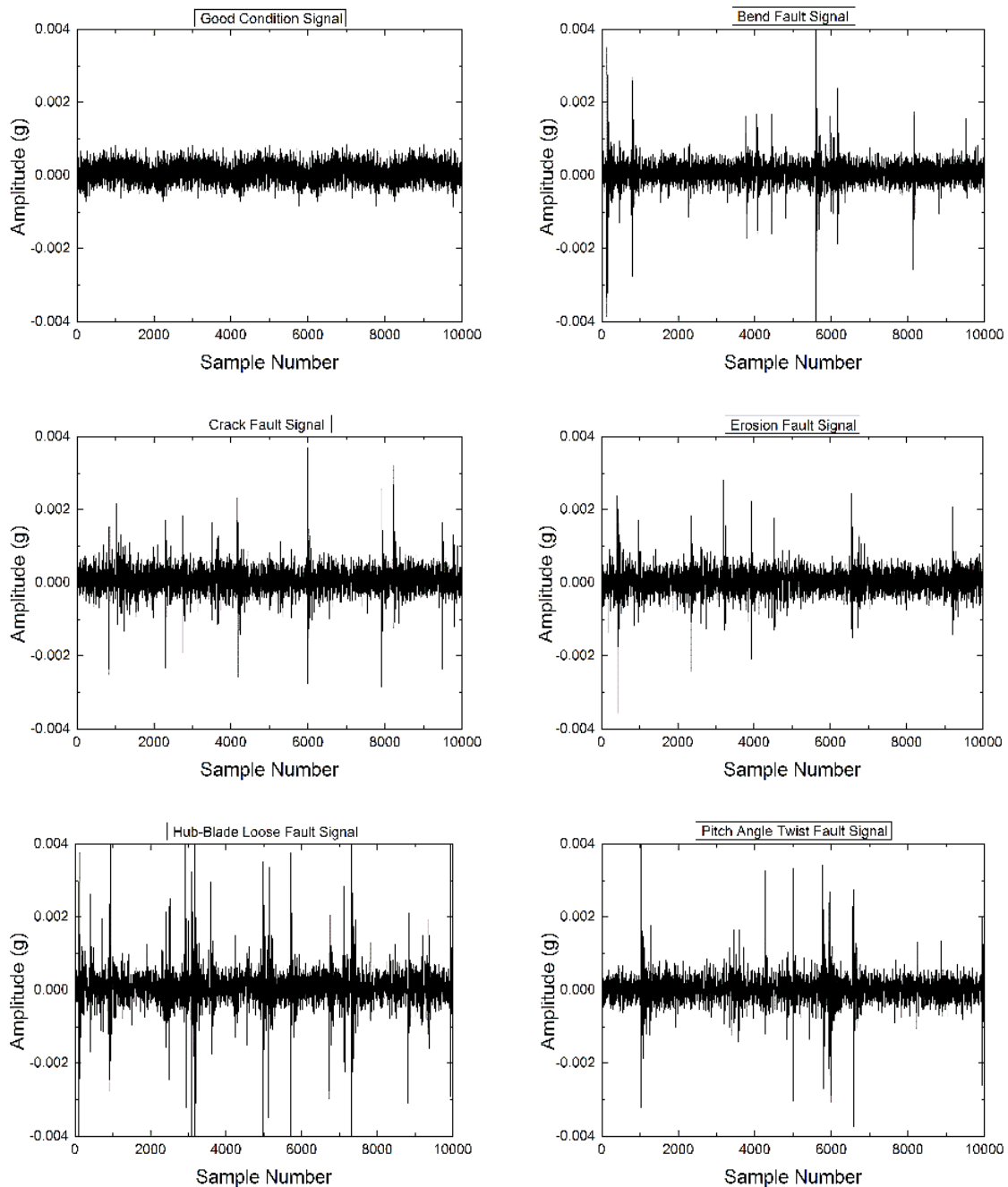
Figure 4: Time-domain signal plot

c) **Blade erosion (BE):** This fault is due to the erosion of the top layer of the blade by high-speed wind. The smooth surface of the blade was eroded using emery sheet (320Cw) to provide an erosion effect on the blade.

d) **Hub-blade loose contact:** This fault generally occurs on wind turbine blade due to over runtime. The connection between the hub and blade bolt was made loose to obtain this fault.

e) **Blade pitch angle twist (PAT):** This fault occurs due to the stress on blade caused by high-speed wind. This makes the pitch get twisted and creates a heavy vibration to the framework. To attain this fault, blade pitch was twisted about 12^0 with respect to the normal blade condition.

Figure 4 shows the rotational domain of the vibration signals of various blade conditions for one revolution of the wind turbine blade. The signal plot (Figure 4) shows the vibration acquired from good condition blade, blade bend, blade crack, pitch angle twist, hub-blade loose connection and blade erosion. This gives some basic idea about how the magnitude of the acquired vibration signal varies over time with respect to the faults that were simulated.

III. STATISTICAL FEATURE EXTRACTION

The vibration signals were obtained for good and other faulty conditions of the blades. If the time domain sampled signals are given directly as inputs to a classifier, then the number of samples should be constant.

The number of signal samples obtained is a function of rotatory motion of the blade speed. Hence, it can't be used directly as the input to the classifier. However, a few features must be extracted before the classification process. Descriptive statistical parameters [13] such as sum, mean, median, mode, minimum, maximum, range, skewness, kurtosis, standard error, standard deviation and sample variance were computed to serve as features in the feature extraction process.

- Sum: It is the sum of all feature values for each sample.
- Mean: The arithmetic average of a set of values or distribution.
- Median: Middle value sorting out the greater and lesser splits of a data set.
- Mode: Most frequent value available in the data set.
- Minimum value: It refers to the least signal point value in a given signal.
- Maximum value: It refers to the extreme signal point value in a given signal.
- Range: Difference in extreme and least signal point values for a given signal.
- Skewness: Skewness illustrates the degree of irregularity of a distribution around its mean. The following formula was used for calculation of skewness.

$$Skewness = \frac{n}{(n-1)(n-2)} \sum \left(\frac{x_i - \bar{x}}{s_d}\right)^3 \text{—(1)}$$

- Kurtosis: Kurtosis point toward the flatness or the spikiness of the signal. Its value is very low for normal condition of the blade and high for the faulty condition of the blade due to the spiky nature of the signal and 's' is the sample standard deviation

$$Kurtosis = \left\{\frac{n(n+1)}{(n-1)(n-2)(n-3)} \sum \left(\frac{x_i - \bar{x}}{s_d}\right)^4\right\} - \frac{3(n-1)^2}{(n-2)(n-3)} \text{(2)}$$

- Standard error: Standard error is a measure of the amount of error in the prediction of y for an individual x in the regression, where x and y are the sample means and 'n' is the sample size.

$$Standard\ Error\ (y) = \sqrt{\frac{1}{n-2}\left[\sum(y - \bar{y})^2 - \frac{\sum[(x-\bar{x})(y-\bar{y})]^2}{\sum(x-\bar{x})^2}\right]} \text{—(3)}$$

- Standard deviation: This is a measure of the actual energy or power content of the vibration signal. The following formula was used for calculation of standard deviation.

$$Standard\ Deviation\ (\sigma) = \sqrt{\frac{n\sum x^2 - (\sum x)^2}{n(n-1)}} \text{—(4)}$$

- Sample variance: It is the variance of the signal points and the following formula was used for calculation of sample variance.

$$Sample\ Variance = \frac{n\sum x^2 - (\sum x)^2}{n(n-1)} \text{(5)}$$

When the statistical feature extraction was completed, the features were chosen and feature selection method was carried out. The statistical features form the input to the feature selection method. With the selected feature, further classification was carried out.

IV. J48 DECISION TREE FEATURE SELECTION

J48 decision tree algorithm is adapted from the C4.5 algorithm in WEKA [14]. It consists of a number of branches, one root, a number of nodes, and a number of leaves. One branch is a chain of nodes from the root to a leaf, and each node involves one attribute. The occurrence of an attribute in a tree provides information about the importance of the associated attribute [15]. A decision tree is a tree based knowledge representation methodology used to represent classification rules. J48 decision tree algorithm is a widely used one to construct decision trees [16]. The procedure of forming the decision tree and exploiting the same for feature selection is characterized by the following:

1. The set of features available at hand forms the input to the algorithm; the output is the decision tree.
2. The decision tree has leaf nodes, which represent class labels, and other nodes associated with the classes being classified.
3. The branches of the tree represent each possible value of the feature node from which they originate.
4. The decision tree can be used to classify feature vectors by starting at the root of the tree and moving through it until a leaf node, which provides a classification of the instance, is identified.
5. At each decision node in the decision tree, one can select the most useful feature for classification using appropriate estimation criteria. The criterion used to identify the best feature invokes the concepts of entropy reduction and information gain.

Information gain measures how well a given attribute separates the training examples according to their target classification. The measure is used to select the candidate features at each step while growing the tree [17]. Information gain is the expected reduction in entropy caused by portioning the samples according to this feature.

Information gain (S, A) of a feature A relative to a collection of examples S, is defined as:

$$Gain\ (S, A) =$$
$$Entropy(S) - \sum_{v \in Value(A)} \times \frac{|S_v|}{|S|} Entropy(S_v)\text{--(6)}$$

where Value (A) is the set of all possible values for attribute A, and S_v is the subset of S for which feature A has value v.

Note the first term in the equation for gain is just the entropy of the original collection S and the second term is the expected value of the entropy after S is partitioned using feature A. The expected entropy described by the second term is simply the sum of the entropies of each subset S_v, weighted by the fraction of samples $|S_v|/|S|$ that belong to S_v. Gain (S, A) is, therefore, the expected reduction in entropy caused by knowing the value of feature A. Entropy is a measure of homogeneity of the set of examples and it is given by

$$Entropy(S) = \sum_{i-1}^{c} -P_i \log_2 P_i\text{--(7)}$$

where, c is the number of classes, P_i is the proportion of S belonging to class 'i'.

The J48 decision tree algorithm has been applied to the problem of feature selection. The input to the algorithm is the set of statistical features described above and output of the decision tree shown in Figure 5. It is clearly shown that the top node is the best node for classification. The other features in the nodes of decision tree perform in descending order of significance. It is to be mentioned here that only features that contribute to the classification appear in the decision tree and other features do not contribute much.

The features which have the less discriminating capability can be consciously discarded by deciding on the threshold. This concept is made use for selecting good features. The algorithm identifies the good features for the purpose of classification of the given training data set, and thus reduces the domain knowledge required to select good features for pattern classification problem [18]. Referring from Figure 5, one can identify the most dominating feature to represent the blade conditions are the sum, range, standard deviation, and kurtosis.
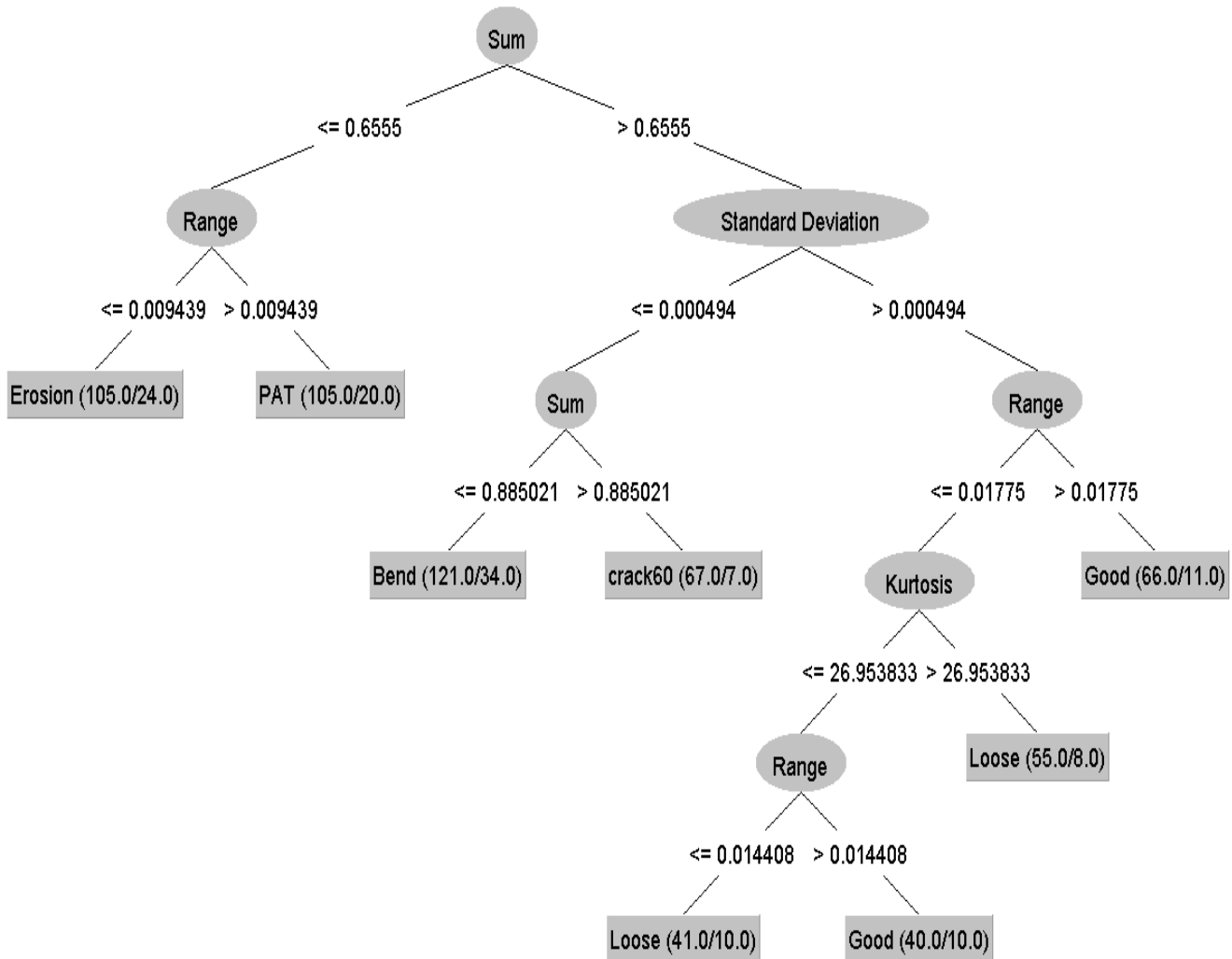

Figure 5: J48 Tree classification for feature selection

V. RULE BASED FEATURE CLASSIFICATION

The selected features are served as input to the classifiers. The wind turbine blade fault diagnosis was carried out using conjunctive rule (CR), decision table (DT), decision table and Naive Bayes hybrid classifier (DTNB), JAVA implemented repeated incremental pruning to produce error reduction (JRip), non-nested generalized exemplars (NNge), one rule (OneR), projective adaptive resonance theory (PART), ripple down rule learner (Ridor), zero rule (ZeroR), fuzzy unordered rule induction algorithm (FURIA), modified learnable evolution model(MODLEM) and ordinal learning method (OLM) classifiers.

5.1 Conjunctive Rule (CR)

Conjunctive rule learner is one of the machine learning algorithms and is normally known as inductive learning. The goal of rule induction is generally to induce a set of rules from data that captures all generalizable knowledge within that data, and at the same time being as small as possible [19]. Classification in rule-induction classifiers is typically based on the firing of a rule on a test instance, triggered by matching feature values at the left-hand side of the rule [20]. Rules can be of various normal forms, and are typically ordered; with ordered rules, the first rule that fires determines the classification outcome and halts the classification process.

5.2 Decision Table (DT)

Decision table (DT)builds a decision table majority classifier [21]. It evaluates feature subsets using best-first search and can use cross-validation for evaluation [22].

An option uses the nearest-neighbor method to determine the class for each instance that is not covered by a decision table entry, instead of the table's global majority, based on the same set of features.

5.3 Decision Table and Naive Bayes Hybrid Classifier (DTNB)

Decision table and Naive Bayes classifier (DTNB)is a hybrid classifier that combines a decision table with Naïve Bayes proposed by Hall and Frank [23]. The model is a Bayesian Network in which the conditional probabilities are represented with Decision Tables. In the DTNB algorithm, the attributes are divided into two subsets by applying the gain function. These two subsets are used to create Decision Tables and Naïve Bayes model, respectively. The algorithm is based on a forward selection procedure, where all attributes are initially modelled by Decision Trees and the selected attributes are provided to a Naïve Bayes model. In order to generate the overall class probability, the class probabilities estimated by Decision Tables and Naïve Bayes have to be combined [24].

5.4 JAVA implemented Repeated Incremental Pruning to Produce Error Reduction (JRip)

JRip is a JAVA implemented Repeated Incremental Pruning to Produce Error Reduction (RIPPER) which was introduced by Cohen [25] to produce easily readable, fast and accurate rules from noisy and large data sets. The main idea of RIPPER approach is for seeking an initial set of rules and iteratively improving it by applying an optimization algorithm. Such modelling with determination of initial rule sets makes this approach effective and fast. The training set used in the rule induction process of this approach is split into two parts: growing set and pruning set. The instances from the growing set are used to build a rule set that starts with an empty set. Once the rule is grown using the data from the first set, the instances from the pruning g data set are applied to advance the performance of the obtained set by pruning it [26].

5.5 Non-Nested Generalized Exemplars (NNge)

Non -Nested Generalized Exemplars (NNge) is an extension of Nested Generalized Exemplars (NGE), which is also an extension to the nearest neighbor classification approach that learns incrementally from the examples. NNGE was proposed by Martin [27] with the goal to solve the overgeneralization problem in the NGE method, which leads to the poor performance. The NNge creates a new generalization each time a new in stance is added to the system by distributing it to the nearest neighbor of the same class [28].

5.6 One Rule (OneR)

One-R is one of the most widely applied rule-based classifiers due to its simplicity and agility which was proposed by Holte [29]. Based on a one level decision tree, this machine learning algorithm attempts to classify the instances by using the value of single attributes. In spite of the accepted lack in accuracy of this classifier, simplicity and speed of this approach make it as a crucial alternative to more complex rule based models [30]. Distinguished from other classifiers that use entropy measures to classify the instances, One-R classifier uses the error rate obtained from the training set. The proposed algorithm develops a rule for each individual predictor in the training set and determines the one rule with lowest error rate.

5.7 Projective Adaptive Resonance Theory (PART)

Frank and Witten [31] proposed an algorithm based on partial decision trees, PART, which differs from other alternatives in way that the rules are generated. The PART algorithm is a combination of C4.5 decision tree and RIPPER algorithms. Distinguished from other rule induction classifiers, the PART algorithm doesn't perform global optimization when inducing the rules, which makes it simple and fast. The working principle of this approach is based on separate and conquer strategy, as follows: the first rule is derived, instances covered by this rule are removed and recursively other rules are generated until there are no more instances remaining [32].

5.8 Ripple Down Rule Learner (Ridor)

Ripple down rule learner (Ridor) is a rule induction algorithm that is similar to PART and C4.5 approaches, but derives rules directly using Cendrowska's Prism algorithm in order to deal with noisy data. The Ridor approach is developed by Gaines [33]. The algorithm initially derives a rule that is followed by determination of exception to the defined rule using a least weighted error rate. For each exception the algorithm determines the most appropriate exception and this process continues recursively until all instances are covered. Derivation of exception can be also seen as tree algorithm where the exceptions are sets of rules for classification of classes [34].

5.9 Zero Rule (ZeroR)

Zero rule classifier (ZeroR) is a learner used to test the results of the other learners. ZeroR chooses the most common category all the time [35]. ZeroR learners are used to compare the results of the other learners to determine if they are useful or not, especially in the presence of one large dominating category [36].

5.10 Fuzzy Unordered Rule Induction Algorithm (FURIA)

Fuzzy Unordered Rules Induction Algorithm (FURIA) is an extension of the state-of-the-art rule learning algorithm called RIPPER [37] having its advantages such like simple and comprehensible fuzzy rule base, and introducing new features. FURIA provides three different extensions of RIPPER: (i) it takes an advantage of fuzzy rules instead of crisp ones, (ii) it applies unordered rule sets instead of rule lists, and (iii) it proposes a novel rule stretching method in order to manage uncovered examples [38-39].

5.11 Modified Learnable Evolution Model (MODLEM)

The rule induction algorithm, called Modified Learnable Evolution Model (MODLEM), has been introduced by Stefanowski [40]. It is based on the scheme of a sequential covering and it heuristically generates a minimal set of decision rules for every decision concept (decision class or its rough approximation in case of inconsistent examples).

Such a set of rules attempts to cover all (or the most significant) positive examples of the given concept and not to cover any negative examples (or as little as possible of them). The main procedure for rule induction scheme starts from creating a first rule by choosing sequentially the 'best' elementary conditions according to chosen criteria [41].

5.12 Ordinal Learning Method (OLM)

The ordinal learning model (OLM) [42] is a very simple algorithm that learns ordinal concepts by eliminating non-monotonic pairwise inconsistencies. The generated concepts are rules. During learning, each example is checked against every rule in a rule-base, which is initially empty. If an example is in consistent with a rule in the rule-base, one of them is selected at random while the other is discarded, but if the example is selected, it must be checked for consistency against all the other rules for monotonicity. If it passes this consistency test, it is added as a rule. Consequently the rule-base is kept monotonic at all times. Classification is done conservatively. All the rules are checked in decreasing order of class values against an attribute vector, and the vector is classified to the class of the first rule that covers it. If such a rule does not exist – the attribute vector is assigned the lowest possible class. It has been shown both theoretically and empirically that the OLM results in very few classification rules while learning from noisy ordinal datasets [43].

VI. RESULTS AND DISCUSSION

The vibration signals were noted for good condition blade and other fault conditions of wind turbine blade using DAQ. Totally 600 signal samples were collected; 100 signal samples from each condition were collected. The statistical features were extracted as features and serves as input to the algorithm. The corresponding condition of the classified data will be the required output of the algorithm. From vibration signals, twelve descriptive statistical features were extracted. Out of theses twelve features, four best contributing features were selected using J48 decision tree algorithm. They are the sum, range, standard deviation, and kurtosis. From Figure 5, the feature 'sum' is the most contributing features when compared to other features.

The other contributing features are range, standard deviation, and kurtosis. The minimum number of instances per leaf and the number of data used for reduced-error pruning was kept at 50 for selecting 4 dominating features in J48 decision tree algorithm. The rest of the features like mean, median, mode, minimum, maximum, skewness, sample variance and standard error were eliminated as they contribute very less in fault classification. In Figure 6, the number of features vs classification accuracy is presented. The classification accuracy during the feature selection process using J48 decision tree algorithm is 86.67%. Other feature combinations did not perform well (Figure 6). Hence, sum, range, standard deviation, and kurtosis were chosen. Then, these selected features were given as input to the classifier to determine the classification accuracy.

Figure 6: Classification accuracy for number of features

In Figure 6, the number of features vs classification accuracy is presented. The classification accuracy during the feature selection process using J48 decision tree algorithm is 86.67%. Other feature combinations did not perform well (Figure 6). Hence, sum, range, standard deviation, and kurtosis were chosen. Then, these selected features were given as input to the classifier to deter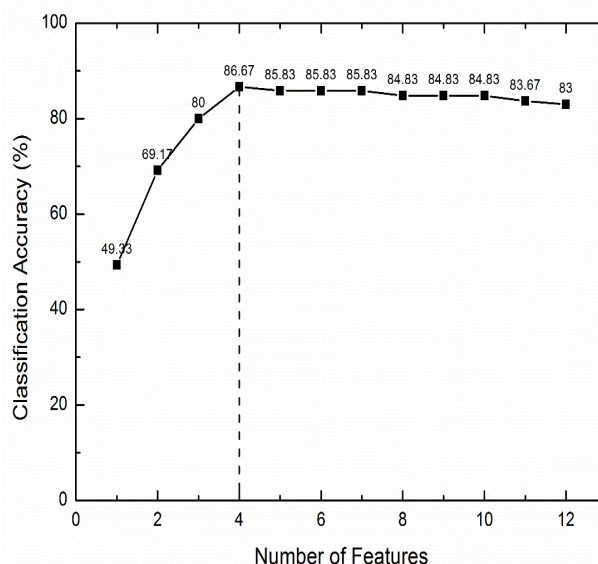mine the classification accuracy. From Figure 6, the selected features were given as the input to rule based classifiers like conjunctive rule (CR), decision table (DT), decision table and Naive Bayes hybrid classifier (DTNB), JAVA implemented repeated incremental pruning to produce error reduction (JRip), non-nested generalized exemplars (NNge), one rule (OneR), projective adaptive resonance theory (PART), ripple down rule learner (Ridor), zero rule (ZeroR), fuzzy unordered rule induction algorithm (FURIA), modified learnable evolution model (MODLEM) and ordinal learning method (OLM) classifiers.

When comparing these twelve algorithm (from Table 2), one can find that the FURIA gives the maximum classification accuracy (87.50%) when compared to other classifiers.

In FURIA algorithm, the T-Norm that is used with fuzzy AND operator has been chosen as standard condition for the classifier to perform.

The error rate of $\geq 1/2$ is included as the stopping criteria with the batch size of 100. The amount of data used as pruning by the classifier is 3 along with the minimum total number of weight in the rule was set to be 2. For the uncovered instances, the rule stretching sub classifier has been assigned with respect to the decision carried out by the FURIA classifier.

The computational time taken by the FURIA classifier for creating a model for fault identification and detection is about 0.12s.

Table 2: Classification Accuracy of Rule based Classifiers

Classifiers	Classification Accuracy (%)	Computational Time (s)
Conjunctive Rule	33.17	0.02
Decision Table	84.00	0.05
DTNB	83.50	0.17
JRip	84.17	0.11
NNge	85.33	0.06
OneR	60.33	0.01
PART	85.00	0.02
Ridor	84.67	0.13
ZeroR	16.67	0.01
FURIA	**87.50**	**0.12**
MODLEM	87.17	0.21
OLM	21.83	0.02

The first row of the confusion matrix, Table 3, represents good condition. The first element (location (1, 1)) represents the number of correctly classified instances belonging to the same. The second element (location (1, 2)) represents the number of good instances that were incorrectly classified as bend fault condition (bend). The third element (location (1, 3)) represents the number of good instances that were incorrectly classified as crack fault condition (crack). The fourth element (location (1, 4)) represents the number of good instances that were incorrectly classified as hub-blade loose fault condition (loose). The fifth element (location (1, 5)) represents the number of good instances that were incorrectly classified as pitch angle twist fault condition (pitch twist). The sixth element (location (1, 6)) represents the number of good instances that were incorrectly classified as erosion fault condition (erosion). Similarly, the second row represents the second condition i.e bend fault condition. The third row represents the data points for third condition, i.e. crack fault condition. The fourth row represents the data points for fourth condition, i.e. hub-blade loose fault condition. The fifth row represents the data points for fourth condition, i.e. pitch angle twist fault condition. The sixth row represents the data points for fourth condition, i.e. erosion fault condition.

Table 3: Confusion matrix for FURIA algorithm

Blade conditions	Good	Bend	Crack	Loose	Pitch twist	Erosion
Good	**82**	0	1	17	0	0
Bend	0	**90**	5	0	0	5
Crack	0	9	**86**	5	0	0
Loose	14	0	6	**80**	0	0
Pitch twist	0	0	0	0	**98**	2
Erosion	0	7	1	0	3	**89**

The confusion matrix of FURIA is shown in Table 3. In confusion matrix, the diagonal elements represent the correctly classified instances and the others are misclassified ones. In FURIA algorithm, out of 600 samples, 525 samples were correctly classified (87.50%) and remaining 75 were misclassified (12.50%). Also one can observe more misclassifications between good and loose conditions. For the loose condition, the bolts between the hub and the blade were made loose (please note that the blade was in good condition). However, at high wind speed, the blade can stick to the hub and behave like a good condition during operation. Because of this, the signature of the loose condition sometimes resembles good condition and the classifier finds difficult to distinguish between them; hence, more misclassifications.

Table 4: Class-wise accuracy of FURIA

Class	TP Rate	FP Rate	Precision	Recall	F-Measure	ROC area
Good	0.820	0.028	0.854	0.820	0.837	0.930
Bend	0.900	0.032	0.849	0.900	0.874	0.957
Crack	0.860	0.026	0.869	0.860	0.864	0.931
Loose	0.800	0.044	0.784	0.800	0.792	0.899
Pitch twist	0.980	0.006	0.970	0.980	0.975	0.987
Erosion	0.890	0.014	0.927	0.890	0.908	0.963

From FURIA, the kappa statistics were found to be 0.85. It is used to measure the arrangement of likelihood with the true class. The mean absolute error was found to be 0.0469. It is a measure used to measure how close forecasts or prediction are with the ultimate result [44]. The root mean square error was found to be 0.1951. It is a quadratic scoring rule which processes the average size of the error. The relative absolute error was found to be 16.8987 % and the root relative squared error was 52.3636 %. The time taken to build the model is about 0.12s; hence, this can be used in real time for the fault detection on the wind turbine blade. The detailed class-wise accuracy is shown in Table 4. The class-wise accuracy is expressed in terms of the true positive rate (TP), false positive rate (FP), precision, recall and F-Measure [45].

TP is used to predict the ratio of positives which are correctly classified as faults. FP is commonly described as a false alarm in which the result that shows a given fault condition has been achieved when it really has not been achieved [46]. The true positive (TP) rate should be close to 1 and the false positive (FP) rate should be close to 0 to propose the classifier is a better classifier for the problem classification [47]. In FURIA, it shows that the TP near to 1 and FP close to 0, hence one can conclude that the classifier built for the specific problem is effective for the fault diagnosis problem. Precision is the probability of retrieved instances that are relevant for the class. That is, it is the ratio of true positive (TP) to the retrieved instances (TP+FP). It is stated as $\frac{TP}{TP+FP}$. Precision is also called as the positive predictive value and can be defined as a measure of exactness or quality [48].
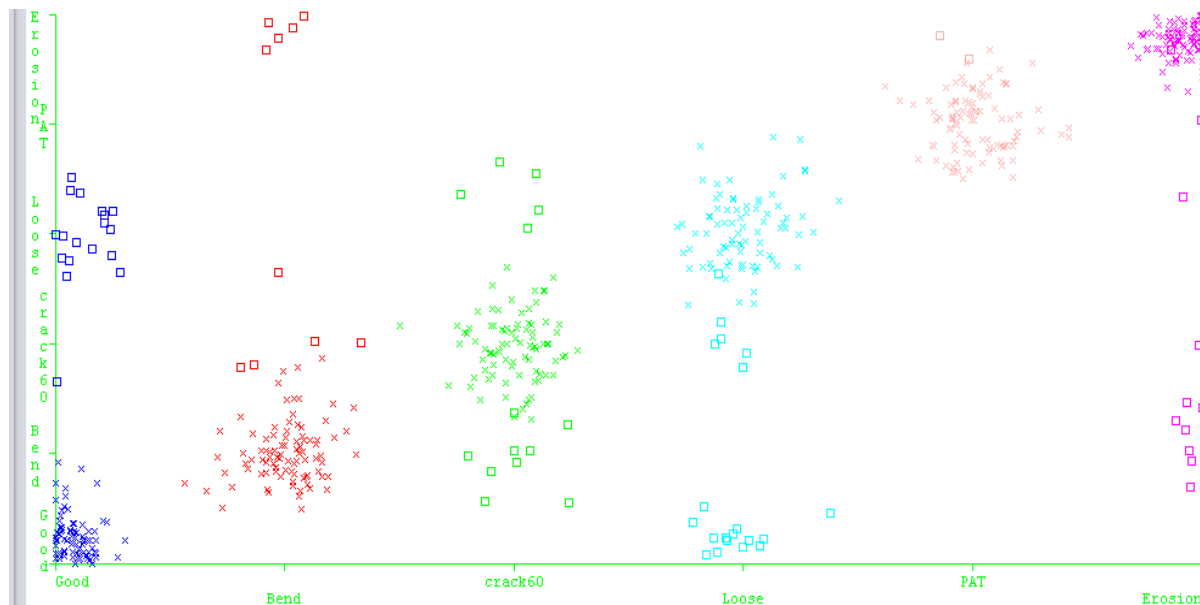
Figure 7: Classifier Errors (Classification vs Misclassification)

The recall is the information retrieval which shows the probability of the faults that are relevant to the classification that is successfully retrieved. That is the ratio of true positive (TP) to the overall instances (TP+FN) [49]. False negative (FN) is considered as type 2 error in which the instances indicates the misclassification but it is actually correctly classified. It is stated as $\frac{TP}{TP+FN}$. Recall is also called as the measure of completeness or quantity. F-measure is defined as the harmonic mean of both recall and precision [50]. That is, this measure is approximately the average of the two (recall and precision) when they are close, and is more generally the square of the geometric mean divided by the arithmetic mean. The f-measure is expressed as $2 * \frac{Recall*Precision}{Recall+Precision}$. The classifier error chart is shown in Figure 7. Here the squared dots represent the misclassification and the 'x' denotes the correct classification.

VII. CONCLUSION

The wind turbine is very much essential in the production of wind energy in our day-to- day life. This paper represents an algorithmic based interpretation of vibration signals for the valuation of wind turbine blade conditions. From the acquired vibration data, twelve rule based model have been developed using data modelling technique. From twelve, fuzzy unordered rules induction algorithm (FURIA)provide the maximum classification accuracy of 87.50% for the prediction of blade fault condition in wind turbine. The model is tested under 10-fold cross validation. The error rate is relatively less and FURIA may be considered for the blade fault diagnosis. Hence, fuzzy unordered rules induction algorithm (FURIA) can be practically used for the condition monitoring of wind turbine blade to reduce the downtime and to maximize the usage of wind energy. The methodology and algorithm suggested in this paper can be potentially used for any kind of wind turbine blade to diagnose the blade fault with minimal modification.

REFERENCES

1. Badihi H, Zhang Y, Hong H. Fuzzy gain-scheduled active fault-tolerant control of a wind turbine. *Journal of the Franklin Institute.* 2014 Jul 31;351(7):3677-706.
2. Chehouri A, Younes R, Ilinca A, Perron J. Review of performance optimization techniques applied to wind turbines. *Applied Energy.* 2015 Mar 15;142:361-88.
3. Kusiak A, Verma A. A data-driven approach for monitoring blade pitch faults in wind turbines. Sustainable Energy, *IEEE Transactions on.* 2011 Jan;2(1):87-96.
4. Chen B, Matthews PC, Tavner PJ. Wind turbine pitch faults prognosis using a-priori knowledge-based ANFIS. *Expert Systems with Applications.* 2013 Dec 1;40(17):6863-76.
5. Shamshirband S, Petković D, Saboohi H, Anuar NB, Inayat I, Akib S, Ćojbašić Ž, Nikolić V, Kiah ML, Gani A. Wind turbine power coefficient estimation by soft computing methodologies: comparative study. *Energy Conversion and Management.* 2014 May 31;81:520-6.
6. Mollineaux M, Balafas K, Branner K, Nielsen P, Tesauro A, Kiremidjian A, Rajagopal R. Damage detection methods on wind turbine blade testing with wired and wireless accelerometer sensors. *In EWSHM-7th European Workshop on Structural Health Monitoring* 2014 Jul 8.
7. Pratumnopharat P, Leung PS, Court RS. Wavelet transform-based stress-time history editing of horizontal axis wind turbine blades. *Renewable Energy.* 2014 Mar 31;63:558-75.
8. Hoell S, Omenzetter P. Damage Detection in a Wind Turbine Blade Based on Time Series Methods. *In EWSHM-7th European Workshop on Structural Health Monitoring* 2014 Jul 8.
9. Vučina D, Marinić-Kragić I, Milas Z. Numerical models for robust shape optimization of wind turbine blades. *Renewable Energy.* 2016 Mar 31;87:849-62.
10. Ha JM, Oh H, Park J, Youn BD. *Classification of operating conditions of wind turbines for a class-wise condition monitoring strategy. Renewable Energy.* 2017 Apr 30;103:594-605.
11. Joshuva A, Sugumaran V. Fault diagnostic methods for

wind turbine: A review. *Asian Research Publishing Network (ARPN) Journal of Engineering and Applied Sciences.* 2016 Apr;11(7):4654-68.

12. Joshuva A, Sugumaran V. A data driven approach for condition monitoring of wind turbine blade using vibration signals through best-first tree algorithm and functional trees algorithm: A comparative study. *ISA transactions.* 2017 Mar 31;67:160-72.

13. Joshuva A, Sugumaran V. Wind Turbine Blade Fault Diagnosis Using Vibration Signals through Decision Tree Algorithm. *Indian Journal of Science and Technology.* 2016 Dec 29;9(48).

14. Mitchell TM. Machine learning. 1997. *Burr Ridge, IL: McGraw Hill.* 1997;45:37.

15. Witten IH, Frank E. Data Mining: Practical machine learning tools and techniques. *Morgan Kaufmann;* 2005.

16. Joshuva, A., Sugumaran, V. Classification of Various Wind Turbine Blade Faults through Vibration Signals Using Hyperpipes and Voting Feature Intervals Algorithm. *International Journal of Performability Engineering.* 2017 May;13:247-258.

17. Joshuva A, Sugumaran V, Amarnath M. Selecting kernel function of Support Vector Machine for fault diagnosis of roller bearings using sound signals through histogram features. *International Journal of Applied Engineering Research.* 2015;10(68):482-7.

18. Joshuva A, Sugumaran V. A machine learning approach for condition monitoring of wind turbine blade using autoregressive moving average (ARMA) features through vibration signals: a comparative study. *Progress in Industrial Ecology, an International Journal.* 2018;12(1-2):14-34.

19. Bin Othman MF, Yau TM. Comparison of different classification techniques using WEKA for breast cancer. *In3rd Kuala Lumpur International Conference on Biomedical Engineering 2006* 2007 (pp. 520-523). Springer Berlin Heidelberg.

20. Montgomery H, Svenson O. On decision rules and information processing strategies for choices among multiattribute alternatives. *Scandinavian Journal of Psychology.* 1976 Sep 1;17(1):283-91.

21. Wang GY, Yu H, Yang DC. Decision table reduction based on conditional information entropy. *Chinese Journal Of Computers-Chinese Edition-.* 2002 Jul;25(7):759-66.

22. Kohavi R. The power of decision tables. InEuropean Conference on Machine Learning 1995 Apr 25 (pp. 174-189). *Springer Berlin Heidelberg.*

23. Hall MA, Frank E. Combining Naive Bayes and Decision Tables. *In FLAIRS Conference* 2008 May 15 (Vol. 2118, pp. 318-319).

24. Chen C, Zhang G, Yang J, Milton JC. An explanatory analysis of driver injury severity in rear-end crashes using a decision table/Naïve Bayes (DTNB) hybrid classifier. *Accident Analysis & Prevention.* 2016 May 31;90:95-107.

25. Cohen WW, Singer Y. Context-sensitive learning methods for text categorization. *ACM Transactions on Information Systems (TOIS).* 1999 Apr 1;17(2):141-73.

26. Nguyen HA, Choi D. Application of data mining to network intrusion detection: classifier selection model. *In Asia-Pacific Network Operations and Management Symposium* 2008 Oct 22 (pp. 399-408). Springer Berlin Heidelberg.

27. Martin B. Instance-based learning: nearest neighbour with generalisation, 1995.

28. Weber BG, Mateas M. A data mining approach to strategy prediction. *In Computational Intelligence and Games, 2009. CIG 2009. IEEE Symposium on 2009* Sep 7 (pp. 140-147). IEEE.

29. Holte RC. Very simple classification rules perform well on most commonly used datasets. Machine learning. 1993 Apr 1;11(1):63-90.

30. Menzies T, Greenwald J, Frank A. Data mining static code attributes to learn defect predictors. *IEEE transactions on software engineering.* 2007 Jan;33(1).

31. Frank E, Witten IH. Generating accurate rule sets without global optimization, 1998.

32. Cao Y, Wu J. Dynamics of projective adaptive resonance theory model: the foundation of PART algorithm. *IEEE Transactions on Neural Networks.* 2004 Mar;15(2):245-60.

33. Hall MA, Frank E. Combining Naive Bayes and Decision Tables. *In FLAIRS Conference* 2008 May 15 (Vol. 2118, pp. 318-319).

34. Daud NR, Corne DW. Human readable rule induction in medical data mining. *In Proceedings of the European Computing Conference* 2009 (pp. 787-798). Springer US.

35. Aher SB, Lobo LM. Data mining in educational system using Weka. *In IJCA Proceedings on International Conference on Emerging Technology Trends (ICETT)* 2011 (Vol. 3, pp. 20-25).

36. Kalapanidas E, Avouris N, Craciun M, Neagu D. Machine Learning Algorithms: A study on noise sensitivity. *In Proc. 1st Balcan Conference in Informatics* 2003 Oct (pp. 356-365).

37. Hühn J, Hüllermeier E. FURIA: an algorithm for unordered fuzzy rule induction. *Data Mining and Knowledge Discovery.* 2009 Dec 1;19(3):293-319.

38. Hühn JC, Hüllermeier E. An analysis of the FURIA algorithm for fuzzy rule induction. *In Advances in machine learning* I 2010 (pp. 321-344). Springer Berlin Heidelberg.

39. Trawiński K, Cordón O, Quirin A. On designing fuzzy rule-based multi classification systems by combining FURIA with bagging and feature selection. *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems.* 2011 Aug;19(04):589-633.

40. Stefanowski J. On combined classifiers, rule induction and rough sets. *In Transactions on rough sets* VI 2007 (pp. 329-350). Springer Berlin Heidelberg.

41. Stefanowski J. The bagging and n 2-classifiers based on rules induced by MODLEM. *In International Conference on Rough Sets and Current Trends in Computing* 2004 Jun 1 (pp. 488-497). Springer Berlin Heidelberg.

42. Ben-David A, Sterling L, Pao YH. Learning and classification of monotonic ordinal concepts. *Computational Intelligence.* 1989 Jan 1;5(1):45-9.

43. Ben-David A, Jagerman DL. Evaluation of the number of consistent multiattribute classification rules. *Engineering Applications of Artificial Intelligence.* 1997 Apr 1;10(2):205-11.

44. Joshuva A, Sugumaran V. A Comparative Study for Condition Monitoring on Wind Turbine Blade using Vibration Signals through Statistical Features: a Lazy Learning Approach. *International Journal of Engineering & Technology.* 2018;7(4.10):190-6.

45. Peck R, Devore JL. Statistics: The exploration & analysis of data. *Cengage Learning*; 2011.

46. Joshuva A, Sugumaran V, Amarnath M, Lee SK. Remaining Life-Time Assessment of Gear Box Using Regression Model. *Indian Journal of Science and Technology.* 2016 Dec 28;9(47):1-8.

47. Powers DM. Evaluation: from precision, recall and F-measure to ROC, informedness, markedness and correlation.

48. Joshuva A, Sugumaran V. A Study of Various Blade Fault Conditions on a Wind Turbine Using Vibration Signals through Histogram Features. *Journal of Engineering Science and Technology.* 2018 Jan;13(1):102-121.

49. Manju BR, Joshuva A, Sugumaran V. A Data Mining Study for Condition Monitoring on Wind Turbine Blades Using Hoeffding Tree Algorithm Through Statistical and Histogram Features. *International Journal of Mechanical Engineering and Technology.* 2018 ;9(1):1061-1079.

50. Joshuva A, Sugumaran V. A comparative study of Bayes classifiers for blade fault diagnosis in wind turbines through vibration signals. *StructDurab Health Monit (SDHM).* 2017;12(1):69-90.

51. Deenadayalan, Sivakumar, S**.**, Vishnuvardhan R., Sathish Kumar R, "Fabrication and Characterisation of B-H-G Fiber with Teak Wood Particles Reinforced Hybrid Composite, *International Journal of Engineering and Technology,* Vol 7(2.31), 208 -211, 2018

52. Sivakumar S, Dhanalakshmi V and Vishuvardhan R, **"**Extraction of Subtractive Features of Prismatic Parts from STEP File for CAD/CAM Integration" *Journal of Science and Technology, Malaysia*, Vol. 27 (1), 343-356, Jan. 2019.

53. Vishnuvardhan R, Pooranam N, Sivakumar S, Vignesh T., "Producing Electrical Energy from Light Intensity, Design Traking System by Dual Axis Method", *International Journal of Mechanical and Production Engineering Research and Development (IJMPERD)*Vol. 9, Jan 2019, 503-509.