Hindawi
International Journal of Photoenergy
Volume 2022, Article ID 5389574, 14 pages
https://doi.org/10.1155/2022/5389574

Research Article

Investigating Rotor Conditions on Wind Turbines Using Integrating Tree Classifiers

Bikash Chandra Saha,[1] Joshuva Arockia Dhanraj,[2] M. Sujatha,[3] R. Vallikannu,[4] Mohana Alanazi,[5] Ahmad Almadhor,[6] Ravishankar Sathyamurthy,[7] Kuma Gowwomsa Erko,[8] and V. Sugumaran[9]

[1]*Department of Electrical and Electronics Engineering, Cambridge Institute of Technology, Ranchi, Jharkhand 835103, India*
[2]*Centre for Automation and Robotics (ANRO), Department of Mechatronics Engineering, Hindustan Institute of Technology and Science, Padur, Chennai, 603103 Tamil Nadu, India*
[3]*Department of Electronics and Communication Engineering, Koneru Lakshmaiah Education Foundation Vijayawada, Andhra Pradesh 522502, India*
[4]*Department of Electronics and Communication Engineering, Hindustan Institute of Technology and Science, Padur, Kelambakkam, Tamil Nadu 603103, India*
[5]*Department of Electrical Engineering, College of Engineering, Jouf University, Sakaka 42421, Saudi Arabia*
[6]*Department of Computer Engineering and Networks, Jouf University, Sakaka 42421, Saudi Arabia*
[7]*Department of Mechanical Engineering, KPR Institute of Engineering and Technology, Arasur, Coimbatore, 641407 Tamil Nadu, India*
[8]*Department of Mechanical Engineering, Ambo University, Ethiopia*
[9]*School of Mechanical Engineering, Vellore Institute of Technology, Chennai, Tamil Nadu 600127, India*

Correspondence should be addressed to Kuma Gowwomsa Erko; kuma.gowwomsa@ambou.edu.et

Received 31 January 2022; Accepted 26 April 2022; Published 9 June 2022

Academic Editor: V. Mohanavel

Renewable wind power is productive and feasible to manage the energy crisis and global warming. The wind turbine's blades are the essential components. The dimension of wind turbine blades has been increased with blade sizes varying from approx. 25 m up to approx. 100 m or even greater with a specific purpose to increase energy efficiency. While wind turbine blades tend to be highly stressed by environmental conditions, the wind turbine blade must be constantly tested, inspected, and monitored for wind turbine blades safety monitoring. This research presents a methodology adaptation on machine learning technique for appropriate classification of different failure conditions on blade during turbine operation. Five defects were reported for the diagnosis study of defective wind turbine rotor blades, and the considered defects are blade crack, erosion, loose hub blade contact, angle twist, and blade bend. The statistical features have been drawn from the recorded vibration signals, and the important features was selected through J48 classifier. Eight tree-dependent classifiers were used to categorize the state of the rotor blades. Among the classifiers, the least absolute deviation tree performed better with the classification percentage of 90% (Kappa statistics = 0.88, MAE = 0.0362, and RMSE = 0.1704) with a computational time of 0.06 s.

1. Introduction

The increasing impact of fossil fuel energy on global processes such as global warming and the availability of electricity has prompted various countries to explore clean options, like green solar and wind energy, to meet expanding energy demand [1]. The majority of wind turbines have the same essential components: shafts, gearboxes, blades, and generators, both of which giant and small. These sectors work together to turn wind power into electricity. The well-

known fact is that the energy of the wind turbines relies on its propellers [2]. The pivotal parts of the turbine blades and the costs for the blade are approximately 15–25 percent of the cost of generation of wind turbines. The scale of wind turbine blades with blade diameters varies from approx. 25 m to approx. 100 m, or even greater, to improve energy efficiency. The blades may suffer from absorption of moisture, ultraviolet radiation, degradation of the environment, exhaustion, wind or lightning, etc. The use of automated systems becomes troublesome during a traditional approach, as wear and tear vary in the pace for the parts. In the machine-learning method, algorithms will continuously adapt and get to know the changing circumstances [3]. Researchers in the machine learning approach routinely take follow-up on fault diagnosis of mechanical systems.

In specific, economic, social, and environmental influence are the criteria for choosing a maintenance strategy. Even though wind turbines on-shore and off-shore are located remotely from populated areas, there would be little social and environmental concern. The economic dimension is also the key consideration for assessing the annual maintenance strategies [4]. The model proposed in this study can provide a typical approach for the determination of the optimum cost of on-shore and off-shore turbine preventive maintenance. Two key maintenance types are available, namely, corrective maintenance and proactive maintenance. Corrective maintenance measures are based on a strategy of failure [5]. Two common categories of proactive maintenance are condition-based maintenance and time-based maintenance [6]. Predictive maintenance, which can be called a functional method of condition-based maintenance in some fields of research like artificial intelligence and machine learning, is closely related to predictive maintenance [7]. The emphasis of this paper is on condition-based maintenance [8].This study makes an attempt to find five different blade fault conditions by applying an approach through tree-based machine learning classifiers and statistical features. Figure 1 shows the methodology of the work done.

The present study's contributions are as follows:

(i) Five defects were reported for the diagnosis study of defective wind turbine rotor blades, and the considered defects are blade crack, erosion, loose hub blade contact, angle twist, and blade bend

(ii) Statistical features have been drawn from the recorded vibration signals

(iii) For the feature selection, J48 decision algorithm was used

(iv) It is modeled as multiclass problem and aims to examine the state of blade using a tree-dependent classifiers

The organization of the document is prepared accordingly. Section 2 provides the literature review of previous research carried out for the assessment of blade condition via simulation and the machine learning process. Section 3 discusses the experimental setup and procedure. Section 4 includes the procedures of statistical feature extraction. Section 5 illustrates the selection of features through the J48 decision tree algorithm. Section 6 outlines in depth the concept of tree-based machine learning in this research. Section 7 presents the results of the tree-based machine learning classifiers and discussion with their performance. The last section (Section 8) contains conclusions.

2. Literature Survey

Machine learning experiments and wind turbine failure prediction modeling testing were carried out for the most part; they [9] worked on wavelet-based and neural network fault diagnosis by the use of a wavelet transformation process to extract data from the blades collected. The wavelet principle is used to suppress device distortion and to minimize vibration interfering with wind turbine blades defect diagnoses. The processing method shall collect data from the various vibration frequencies of wind turbine blades and insert data into the neural system. The neural networks are used for data analysis and wind turbine blade condition recognition.

This work [10] conducted a work on sparse Bayesian learning through a short-term Fourier transformation to detect errors in wind turbine blades. In this analysis, sparse Bayesian learning (SBL) beam forming (BF) for the identification of blade defects is applied in order to improve signal acoustic information received by a microphone array on the ground. With high-resolution interference, such as noise produced by cooling fans, the path of arrival of the irregular sound can be determined by the low side lobes of the SBL-BF.

Researchers [11] conducted a report on the modeling of failures in large-scale wind turbines and electronic monitoring of blade damage. The study proposes a new approach of the analysis of the wavelet energy spectrum packet and modal operational analytics in the area of wind turbine blades. The first move is to determine the change point of the blades to achieve energy spectrum using the wavelet packet transform. Thanks to the energy change in multiple frequency bands, the harm is observed provisional.

The observed fault tree model for offshore floating wind turbines was suggested by [12]. In this study, the approach for the qualitative and quantitative assessment of semisubmersible offshore wind turbine malfunction components is used in the failure tree study technique. The multiple components include supports, pitch and hydraulics, gearboxes, engines, and other systems. Failure rates are gathered from the previous studies, publications, and reliability records for the related offshore systems.

They [13] have been focusing on the complex reaction of the V-shaped offshore wind turbine to cope with fault states. The dynamic response to different possible fault conditions of the semisubmersible V-form is examined in this paper with different behaviors of the floating wind turbine which were related in the instance of operating and fault conditions of the system. In the different fault situations, a numerical model was developed to provide a fully coupled dynamic analysis of the offshore wind turbine.

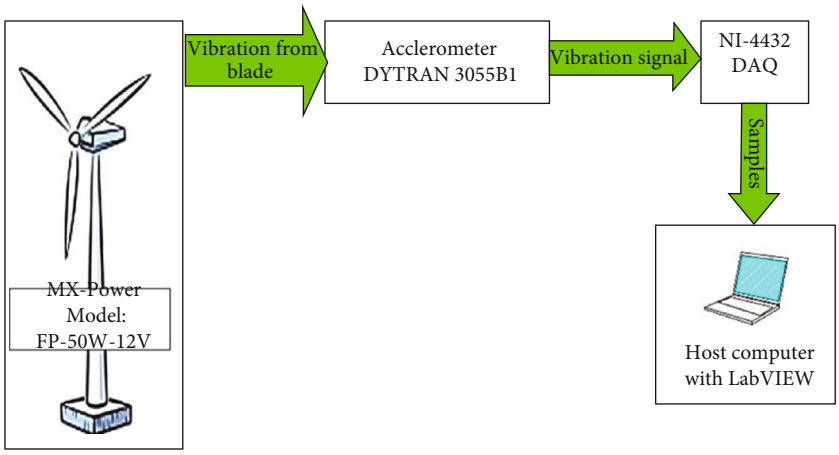


Figure 1: Methodology of the present work.

They have carried out maintenance control studies on artificial intelligence and nonlinear characteristics in wind turbines [14]. This article presents a new scientific solution to delamination research for reliability systems for wind turbine blades. It relies on the abstraction of an exogenous input (NARX) and linear autoregressive process (AR) from a nonlinear self-extraction. Six scenarios were performed using supervised machine learning techniques with various delamination measurements. This involve discriminatory analysis, nearest neighbors, quadratic SVM, ensemble, and tree classifiers.

They [15] have developed a study that uses atmospheric inputs for artificial neural systems to improve the prediction of wind turbine power. The research has a major effect on the calculation of wind farms' financial feasibility by enhancing the precision of monthly energy forecasts. The importance of the analysis is that the wind turbine's performance forecasts take account of ambient equilibrium and air density. A multiparameter input model is used to approximate the power produced by the wind turbine in artificial neural networks (ANN) machine learning method. ANN concept uses the feed forward back propagation (FFBP) algorithm. A 40 percent increase in the mean absolute error (MAE) associated with the density correction indicated an enhancement of the FFBP-ANN model with the density correction techniques.

They [16] have done a research on the rigorous process of detection through acoustic emission and machine learning of wind turbine blade adhesive composite joints. This article provides a framework for clustering method by fast search and find of density peaks (CFSFDP), in an effective detection of various types of damage which can be handled using similarities and differences of AE signals. Based on the clustering study, fatigue cracks and adhesive layer shear failure are, respectively, shown to be critical and characteristic measures of damage. Fiber problems are seen quite close to delamination as a unique damage feature, in comparison to other damage modes in the subspace. In comparison, the effects of the cluster number selection, the spatial similarity metric, and the importance of the cutoff index are seen to be negligible for the cluster performance.

This work [17] carried out adaptive selection for enhancement of blade damage for an operating wind turbine. The proposed works incorporate a structural health monitoring method focused on a semisupervised algorithm for damage detection, using preliminary results to improve the collection of the features of damage detection. The approach suggests a new method of extracting functions by sorting the acceleration values in increasing vibration response. An evolutionary algorithm for choosing applications should be used in order to determine the most appropriate damage prevention. All these technique strengthens the connection of same blade measures and thus the robustness of the proposed structural health monitoring technique.

They [18] conducted an experimental research on passive acoustic detection of damage to structural wind turbine blades health monitoring. The approach uses intrinsic microphones with blades to detect trends, shifts, or spikes in the cavity of the blade by means of a limited network of airborne, natural passive structure amplitude, and recurrent quantification windows. K-means clustering is used to classify the outliers in the calculations and include the feature-spatial representation of the data collection. The system's efficiency is measured on the basis of its capabilities to identify structural events inside the blade detected by manual measurement observation. Novelty of this research work could be identified by Table 1 as discriminant analysis using tree classifiers on wind turbines.

3. Experimental Studies

This research is primarily aimed at assessing whether the blades are in decent or poor shape. The aim is to determine if it is faulty, the form of fault. The following sections describe the experimental and experimental configuration. These were achieved on a wind turbine with varying speeds. Table 2 specifies the basic specifications for the wind turbine. The wind turbine has been fitted opposite to the open-circuit wind tunnel on a stable stainless steel stand. The wind tunnel speed ranges between 5 and 15 m/s and it provides the wind turbine to operate. To mimic the local wind conditions, the wind speed was continuously varying. The diagnostic

Table 1: Comparison of related studies on wind turbine using machine learning algorithms.

Reference number	Year	Literature survey (methods used)																				
		Wavelet	Neural network	Sparse Bayesian learning	Spectrum analysis	Tree classifiers	Dynamic analysis	Autoregressive model	Discriminant analysis	Nearest neighbors	Support vector machines	Ensemble classification	Clustering	Evolutionary algorithm	K-means	Genetic algorithm	Fuzzy logic	Interval observer approach	Blade element momentum	Kriging model	Simulator	Damage detection algorithm	Whale optimization algorithm
Bi et al. [9]	2019	*	*																				
Pan et al. [10]	2019			*																			
Gao et al. [11]	2019	*			*																		
Kang et al. [12]	2019					*																	
Karimirad and Michailides [13]	2019						*																
Jiménez et al. [14]	2020					*		*	*	*	*	*											
Nielson et al. [15]	2020		*																				
Xu et al. [16]	2020												*										
Movsessian et al. [17]	2020													*									
Solimine et al. [18]	2020														*								
Kim et al. [19]	2020															*							
Civelek [20]	2020															*	*						
Beale et al. [21]	2020	*																					
Pujol-Vazquez et al. [22]	2020																	*					
Emami et al. [23]	2020										*												
Elgammi et al. [24]	2020																		*				
Awadallah and El-Sinawi. [25]	2020				*															*			
Saleem and Kim [26]	2020															*							
Philippidis and Eliopoulos [27]	2020																				*		
Krause and Ostermann [28]	2020				*																	*	
Liu et al. [29]	2020															*							
Qais et al. [30]	2020																*						*
Current research	2020					*			*														

Table 2: Technical parameters of wind turbine [31].

Model	FP-50W-12V
Power (W)	50
Voltage (V)	12
Current (A)	8
Rotating rate (rpm)	850
Maximum-power (W)	150
Wind velocity during start up (m/s)	2.5
Wind velocity (m/s) (cut in)	3.5
Wind velocity (m/s) (cut out)	15
Wind velocity (m/s) (security)	40
Wind velocity (m/s) (rated)	12.5
Turbine	3Φ PM generator
Diameter of rotor (mm)	1050
Material of the blade	Reinforced plastics (carbon fiber)

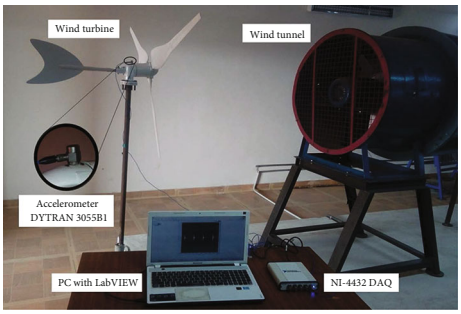


Figure 2: Experimental setup.

test setup is illustrated in Figure 2. As a transducer for the processing of vibration signal, a piezoelectric accelerometer was used. The sensitivity to detect faults is high-frequency. Accelerometers are therefore commonly used for measuring environments. A 500 g range with the sensitivity of 100 mV/g and a 40 Hz accelerometer (resonant frequency) have been used in this case (DYTRAN 3055B1). The DAQ network linked via a wire used was the NI-USB-DAQ-4432 model as data acquisition system. The sample rate of the card is 102.4 kilograms per second, and the resolution of a sample is 24-bit. A signal conditioning device that includes a built-in charge amp and an ADC is attached to the accelerometer. The vibration signal had been received from the ADC. Such vibration signals have been used to disable functionality through the isolation of information. The accelerometer is attached to one end of the cable and the AIO port of the device. The transducer and the computer have been interfaced via NI-LabVIEW.

The three-blade HAWT has been used in this study. At first, the wind turbine (free from faults, new installation) was reasoned to be in good state, and the signals were measured with an accelerometer. Figures 3(a)–3(f) show the faults fabricated on the blade, and Figures 4(a)–4(f) show its corresponding signal, and the following specifications were recorded for these signals:

(1) Sample length: sampling period was chosen appropriately for data accuracy, and the following points were also considered. If the numbers of the sample are sufficiently large, statistical measures are more significant. On the other side, with the amounts of measurements, measurement time decreases. To achieve it, the length of the sample was preferred to be 10000

(2) Sampling frequency: It should be at least twice the maximum signal frequency as set out in the Nyquist theorem. The sampling frequency was calculated using this principle for 12 kHz

(3) Number of trials: 100 measurements (samples) were obtained for each wind turbines blade state

(i) Blade bend (BB): this defect is caused by high velocity wind and storm-induced dynamic movements. With 10° angles, the blade was made to bend

(ii) Blade crack (BC-2): it happens during service as a consequence of foreign object disruption to the sword. A 15-mm crack was made on the trailing edge

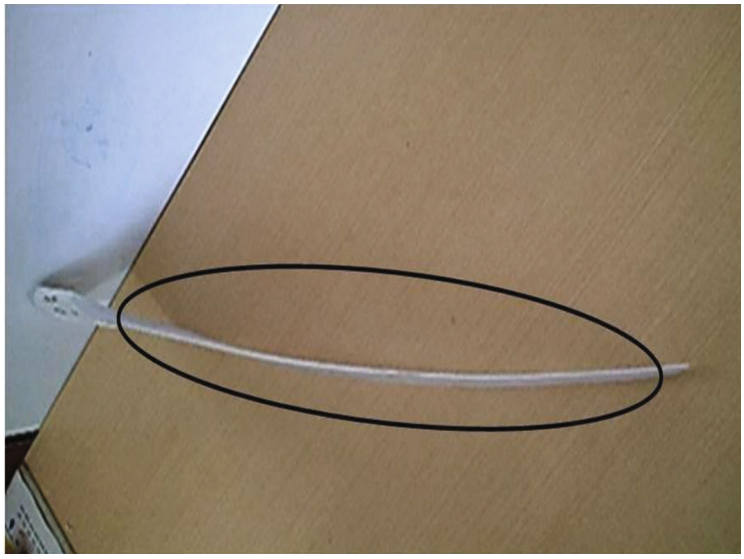
(iii) Blade erosion (BE): this failure is caused by high-speed wind on blade's top surface. The smooth blade surface has been eroded with a 320 Cw sheet to erode the blade

(iv) Hub-blade loose contact: this fault is usually caused by excessive runtime or usage time of a wind turbine. This fault was fabricated by the releasing bolt between the hub and blade

(v) Blade pitch angle twist (PAT): this failure occurs because of high-speed wind pressures on the surface. It shifts the sound, producing a strong frame vibration. The pitch was twisted around 12° angles in relation to the normal condition to achieve this failure

The turbine was designed to operate at 850 revolutions per minute. The vibration signals are captured at a rate of 12000 Hz with a sample size of 10000. In terms of rotations per second ($850/60 \approx 14$ revolutions per second), this equals 14 rotations per second. Besides considering the following factors, the sample length was selected to be long enough to assure consistency in results. When the number of samples is sufficiently big, statistical measures become more relevant. When the number of samples is increased, however, the amount of time spent computing grows. Sample length of around 10000 was selected in order to achieve a good balance. The signal contains data for 14 revolutions, which is good enough to capture the fault information. In this study 100 signal (vibration signatures) were taken for each condition (totally 600 vibration signatures). Each signal, it contain 10000 data points which is very much sufficient in this study.

(a)



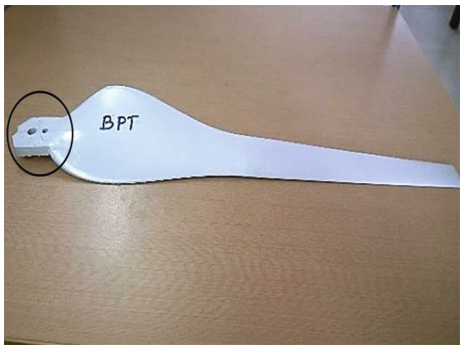
(b)

(c)

Figure 3: Continued.

(d)

(e)



(f)

Figure 3: (a) Good blade. (b) Bend fault. (c) Crack fault. (d) Erosion fault. (e) Hub-blade loose fault. (f) Pitch fault.

4. Feature Extraction

Vibration patterns were acquired from the DAQ for different blade fault conditions in this analysis. The feedback to the classifier cannot be explicitly used by vibration signals [31, 32]. Attributes must therefore be determined by statistical methods. In order to explore issues such as fault classification from an objective standpoint, statistical analysis is used by academics to quantify a wide variety of events. It has a propensity to generate overly simplistic solutions to complicated problems when given a complex question. Each of the statistical approaches is categorized according to a set of factors that allows us to do the best analysis possible for the issue and forecast how much error or deviation has occurred for a certain problem. The calculation process is called feature extraction. Some measures represent the

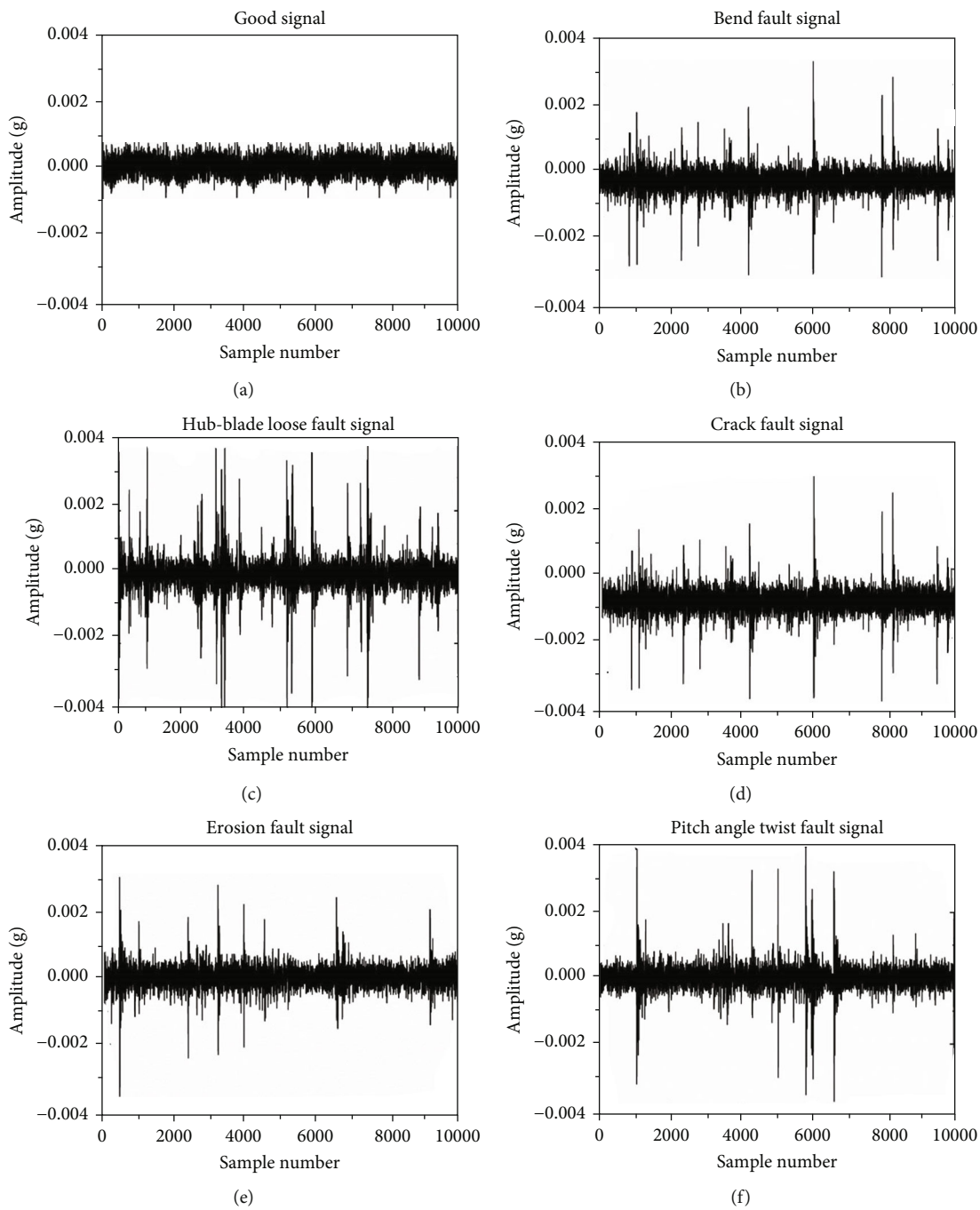


Figure 4: (a) Good condition signal plot. (b) Signal plot for bend condition. (c) Signal plot for crack condition. (d) Signal plot for erosion condition. (e) Signal plot for hub-blade loose condition. (f) Signal plot for pitch angle twist condition.

signal. Vibration signals statistical information provides various parameters [33]. The functions have been taken and the feature collection procedure has been applied once the computational extraction phase has been completed. The most significant characteristics are chosen from the descriptive statistical parameters obtained via the J48 algorithm.

5. Feature Selection

The practice of data mining is commonly used to retrieve useful information structures from libraries utilizing vibration knowledge. An important information framework, which might be produced through the processes of data mining. Decision trees (DT) are recurrently developed after a top-down approach. The regular mediated trees of C4.5 contain a number of branches, roots, many nodes, and leaves. Single branch is a root to a leaf chain of nodes, with one attribute per node. The inclusion of a characteristic in a tree gives information on the importance of the associated feature. Algorithm J48 is a commonly used algorithm to build decision trees (WEKA Application of the C4.5 Algorithm) [34].

J48 decision tree is widely used as a feature selection technique where it uses information gain to select required attributes. This metric assesses how effectively a specific characteristic distinguishes the training instances based on their categorization in the target class. While the tree is being grown, the measure is being used to choose the candidate features at each stage. The projected decrease in entropy as a result of portioning the samples according to this characteristic is referred to as the information gain. The information gain (S, A) of a feature A when compared to a collection of samples S is defined as follows:

$$Gain\,(S, A) = Entropy(S) - \sum_{v \in Value(A)} \times \frac{|S_v|}{|S|} Entropy(S_v), \tag{1}$$

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. It is important to note that the first component in the equation for gain is just the entropy of the original collection S, and the second term is the predicted value of the entropy after S has been 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. As a result, the predicted decrease in entropy induced by knowing the value of feature A is denoted by the symbol Gain (S, A). Entropy is a measure of the homogeneity of a collection of instances, and it is represented by the expression as follows:

$$Entropy(S) = \sum_{i-1}^{c} - P_i \log_2 P_i, \tag{2}$$

where c is the number of classes and P_i is the proportion of S belonging to class i. The method by which the decision tree is shaped and used for vibration analysis is defined by the following:

(1) The mathematical features from analyses of the blade vibration of the wind turbine are the contribution to the algorithm

(2) DT contains leaf nodes containing class marks and other class nodes (in this case, magnitude level)

(3) The tree branches reflect the parameter node from which they derive with each possible value

(4) One can start from the tree root (top node) and go through a branch to a leaf node by using the decision tree to express structural information

(5) A statistical step is taken within the parentheses in the decision tree to provide the contribution level for each individual parameter (Figure 5). The first parenthesis indicates the amount of data points that can be defined by this set parameter. There are decreasing parameters in the decision tree nodes

(6) Using the appropriate estimation criteria at each decision node of the decision tree, the most useful parameter for classification can be used

The entropy and information gain concept is the principle used to categorize the superlative parameter. Two phases have been developed for the decision tree algorithm (C4.5). The construction phase is also referred to as the "development phase" [35]. The main features of the wind turbine blade descriptors are usually available [36]. With regard to Table 3 (selected features with their accuracy) and Figure 5, four of these leading features can be predictable, (a) sum, (b) range, (c) kurtosis, and (d) standard deviation.

6. Feature Classification

It is a process in which data are grouped into a number of classes. The primary objective of a classification is to define the class of new data. The various tree classifiers are as follows [37, 38].

6.1. Decision Stump (DS). The decision stumbling block is a one-level decision tree machine learning model [39]. It is also known as weak students or basic students. It is an internal hub decision tree linked immediately with the hubs (the leaves) of the terminal. A judgment stump makes a prediction on the basis of the value of a single feature [40]. They are also referred to as 1-rules. The tree may be more complex if the function is numeric and the teaching and boosting methods of machine-learning ensemble [41].

6.2. Extra Tree (ET). The ET [42] is a randomized decision-making technique. It is different from other random decision-making bodies. The split and the chosen attribute for each judgment node would be fully randomized with the last parameter setup, and the tree would be entirely independent from the training data. The additional tree is also referred to as a very random tree [43].

6.3. J48 Consolidated (J48C). Its is also known as the consolidated tree construction (CTC) algorithm (J48C) J48 consolidated (J48C) algorithm [44]. The decision on each split shall be voted on by all the trees rather than building each tree independently. All trees conform to the majority voting and are split equally irrespective of their vote [45].

6.4. J48 Graft (J48G). In view of the grafting of further branches into the tree during a postprocessing phase, the J48 decision tree algorithm is extended [46]. The grafting cycle aims to accomplish some of the strength and stability of ensemble approaches like sacking and boosting of trees. It identifies parts of an instance which include either wrong instances or which are either null or explores those classifications for various tests that could be picked at nodes above the regional leaf [47].

6.5. LAD Tree (LADT). The least absolute deviation tree (LADT) is a multiclass problem-solving alternate algorithm, based on the LogitBoost algorithm [48]. LADT can adapt the

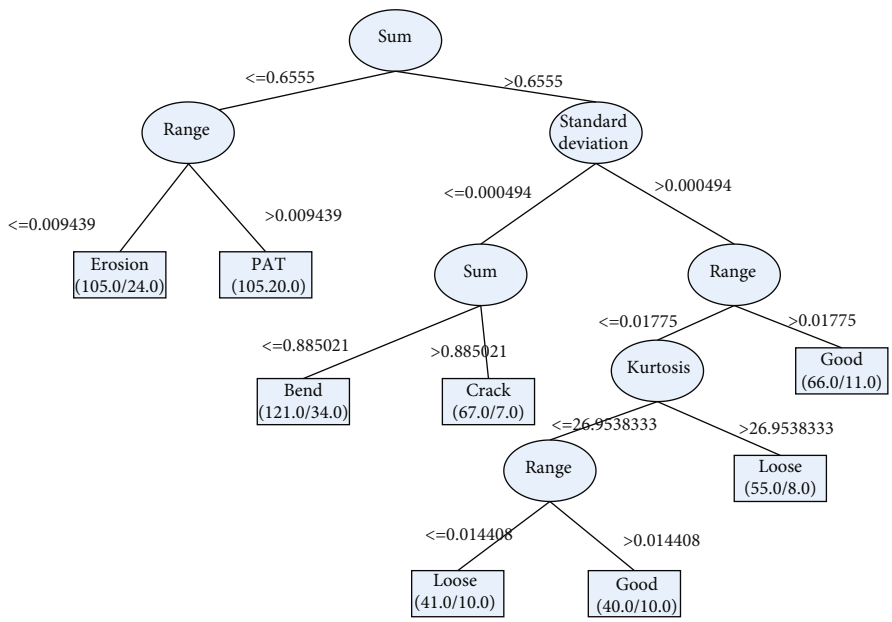


Figure 5: J48 Tree classification for feature selection.

Table 3: Features selected.

No. of features	Name of the features	Classification accuracy (%)
1	Sum	49.33
2	Range and sum	69.17
3	Standard deviation, range, and sum	80.00
4	Standard deviation, range, sum, and kurtosis	86.67
5	Kurtosis, standard deviation, range, sum, and mean	85.83
6	Median, kurtosis, standard deviation, range, sum, and mean	85.83
7	Mode, median, kurtosis, standard deviation, range, sum, and mean	85.83
8	Minimum, mode, median, kurtosis, standard deviation, range, sum, and mean	84.83
9	Maximum, minimum, mode, median, kurtosis, standard deviation, range, sum, and mean	84.83
10	Standard error, maximum, minimum, mode, median, kurtosis, standard deviation, range, sum, and mean	84.83
11	Skewness, standard error, maximum, minimum, mode, median, kurtosis, standard deviation, range, sum, and mean	83.67
12	Sample variance, skewness, standard error, maximum, minimum, mode, median, kurtosis, standard deviation, range, sum, and mean	83.00

amount of boost iterations to the provided data and decide the size of the tree.

6.6. NB-Tree (NBT). The NB-Tree (NBT) hybrid is a hybrid of Naïve Bayes and decision trees. This produces a tree of leaves which are marked as Naïve Bayes for entering the top. Cross-validation is used when constructing a tree to determine if a split node is required or the model of the Naïve Bayes is used in place [49].

6.7. REP Tree (REPT). REP tree (REPT) or REPT builds a decision-regression tree for knowledge gain or variance reduction that is plumed by decreased error pruning. It only speeds up numerical characteristics. This deals, as in C4.5 [50], with missed values by splitting instances into bits. Minimum number of instances can be determined with the root, maximum depth of a tree (useful for boosting trees), and maximum split time (only number groups) as well as several slicing folds.

6.8. Simple Cart Classifiers (SCC). The simple cart classifier (SCC) is a member of a classification decision tree that uses a minimum cost difficulty approach. Similarities do however come to an end here when none of CART's other characteristics are discussed, namely, a CART specialist, a founder of this technique [51]. The overall number of cases per paper can be estimated, the amount of the preparation details used

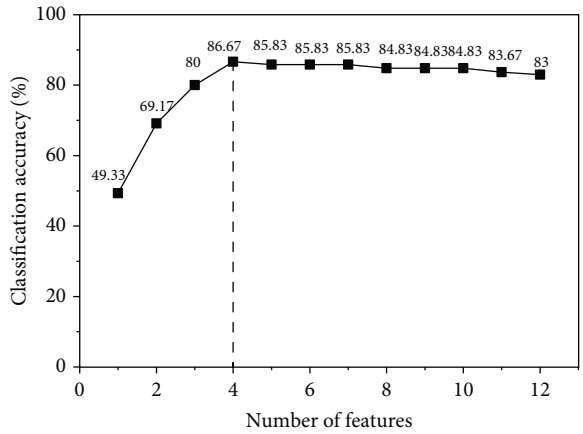


Figure 6: Number of features vs. classification accuracy.

for tree construction and the number of folds used during the cutting process.

7. Result and Discussion

Vibration patterns was obtained with DAQ (total of 600 samples; 100 samples were collected in each condition), in good condition or in bad bladder conditions. 12 concise statistical features were derived from the vibration patterns. Of them, the J48 has selected four best contributing characteristics. In order to select 4 primary features from the J48 method, the minimum number of instances per sheet was set at 50 for reduced-error pruning, and Figure 6 illustrates the number of characteristics vs. precision in classification. The classification performance of the J48 algorithm is 86.67% during the selection process. Other combinations of features were not successful as in Figure 6. The chosen features were then given to the classifier as an input to measure the accuracy of the classification [52].

Table 4 shows the classification precision and measurement time of various classifiers. The purpose of the study is to develop the classification system. It is easy to notice (Table 4) that among classifiers considered, the LADT classification is stronger than other classifiers. This rule is based on high precision (90%) in low time (0.06 s) grouping. The number of iterations increased ranged from 10 to 100 in order to achieve this classification accuracy in the LADT and showed that the optimum classification accuracy was reached in a 90% increase in the number of iterations (Figure 7) [53].

Table 5 shows the LADT confusion matrix. The diagonal components signify the properly categorized occurrences in the confusion matrix, and the other instances are misclassified. For estimating the error rate of a learning process given a single, fixed sample of data, stratified ten-fold cross-validation is the most used approach. At random, the data is split into 10 parts, with each section representing the class in about the same proportion as the overall sample population. It is necessary to train the learning scheme on the remaining nine tenths of each component before calculating its error rate on the holdout set. Because of this, the learning

Table 4: Classifiers accuracy.

Classifiers	Classification accuracy (%)	Computational time (s)
Decision stump (DS)	33.17	0.04
Extra tree (ET)	81.50	0.03
J48 consolidated (J48C)	85.17	0.03
J48 graft (J48G)	79.67	0.09
LAD tree (LADT)	90.00	0.06
NB tree (NBT)	82.33	0.90
REP tree (REPT)	83.83	0.04
Simple cart classifiers (SCC)	85.00	0.32

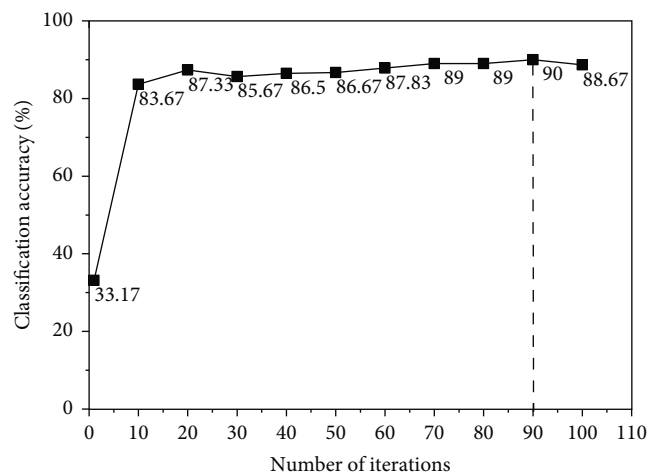


Figure 7: Number of iterations considered for LAD tree.

Table 5: Confusion matrix for LAD tree.

Blade class	Class A	Class B	Class C	Class D	Class E	Class F
Class A	*82*	0	1	0	17	0
Class B	0	*95*	4	0	0	1
Class C	0	5	*91*	4	0	0
Class D	0	4	0	*91*	0	5
Class E	13	0	3	0	*84*	0
Class F	0	0	0	3	0	*97*

Table 6: Class-wise accuracy of LAD tree.

Blade class	TP	FP	PRE	REC	F-M
Class A	0.820	0.026	0.863	0.820	0.841
Class B	0.950	0.018	0.913	0.950	0.931
Class C	0.910	0.016	0.919	0.910	0.915
Class D	0.910	0.008	0.958	0.910	0.933
Class E	0.840	0.042	0.800	0.840	0.820
Class F	0.970	0.010	0.951	0.970	0.960

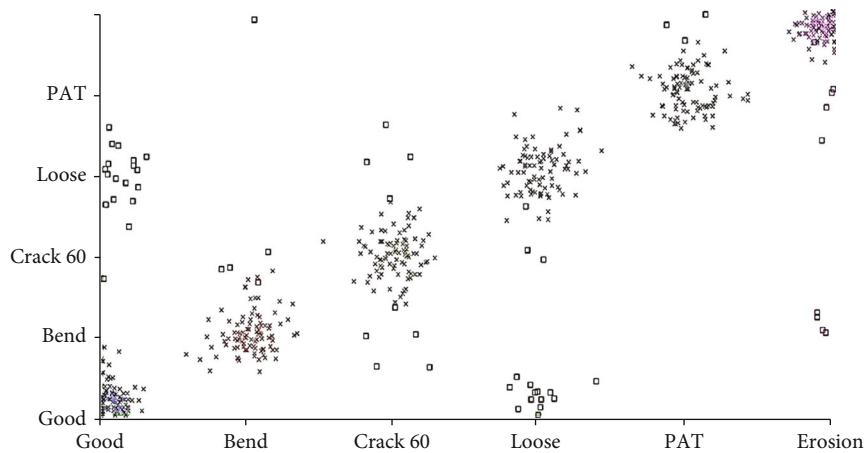


Figure 8: Errors chart.

procedure is repeated 10 times on a variety of practice sets. Finally, an overall error estimate is calculated by averaging the ten error estimates. In this approach, the error rate may be calculated quickly and accurately. The major component (area 1,1) in the first row of the confusion matrix refers to the number of properly categorized items that have a position with the equivalent that reflects good condition (class A). The next component (section 1,2) discusses the number of excellent examples of bent fault that were incorrectly allocated (bend) (class B). The third component (area 1,3) describes the number of excellent crack loss (crack) situations that have been incorrectly delegated (class C). The fourth component (region 1,4) refers to the number of positive occurrences incorrectly referred to as hub-blade (loose) fault situations (class D). The fifth component (area 1,5) refers to the number of good occurrences incorrectly referred to as pitch angle twist fault condition (PAT) (class E). The sixth component (area 1,6) refers to the number of good samples that were incorrectly assigned to the erosion fault condition (erosion) (class F). Kappa was observed to be 0.88 from the LADT. It was observed that the mean absolute error (MAE) was 0.0362. The root mean square error (RMSE) implies that 0.1704 is a square defect. The average error size is processed by a quadratic scoring rule. Table 6 displays the comprehensive class accuracy. The accuracy of the class is shown in terms of the true positive rate (TP), false positive rate (FP), precision (PRE), recall (REC), and F-measure (F-M). [54].The percentage of positive which is accurately labeled as defects is estimated using TP. In a stronger classifier, the TP probability should be about 1 and the FP number similar to 0. Table 6 reveals that the TP average in most groups is below 1 and that the FP rate was below 0 [55]. It ensures the uncertainty matrix provided in Table 5. Figure 8 shows the error classification chart. The squared dots here represent errors, and the "x" corresponds to the right classification. Of the 600 test samples, 540 (90 percent) and the remaining 60 (10 percent) are misclassified. It takes approximately 0.06 seconds to construct the pattern.

8. Conclusion

For harvesting wind energy from the open wind, the wind turbine is a very critical device. The paper presented an algorithmic vibration signal classification for wind turbine blades estimation. Five defects were reported for the diagnosis study of defective wind turbine rotor blades, and the considered defects are blade crack, erosion, loose hub blade contact, angle twist, and blade bend. The collected vibration data has been used to create eight models utilizing tree dependent classifiers. The algorithms were evaluated using the 10-fold cross-validation technique. The following where the findings from this study:

(1) Least absolute deviation tree (LADT) provided the cumulative accuracy of 90% when compared to other tree dependent classifiers with the computational time of 0.06 s (Table 4)

(2) The obtained Kappa statistics is 0.88 for LADT which shows a better fit model

(3) Mean absolute error was found to be 0.0362

(4) Root mean square error was found to be 0.1704

(5) The error rate of 10% is merely less when compared to other classifiers

From the above findings, it could be said with assurance that the least absolute deviation tree (LADT) can be used in real-time to identify blade defects in operating wind turbines. This study can reduce the framework downtime economically, leading to high productivity of wind energy.

Data Availability

The measured data used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

The authors declare that there is no conflict of interest regarding the publication of this paper.

Acknowledgments

The authors thank Ambo University, Ethiopia, for providing help during the research and preparation of the manuscript. The authors thank Vellore Institute of Technology, Chennai; Hindustan Institute of Technology and Science, Chennai; and KPR Institute of Engineering and Technology for providing assistance in completing the work.

References

[1] R. McKenna, S. Pfenninger, H. Heinrichs et al., "High-resolution large-scale onshore wind energy assessments: a review of potential definitions, methodologies and future research needs," *Renewable Energy*, vol. 182, pp. 659–684, 2022.

[2] A. L. Pires, P. Rotella Junior, S. N. Morioka, L. C. Rocha, and I. Bolis, "Main trends and criteria adopted in economic feasibility studies of offshore wind energy: a systematic literature review," *Energies*, vol. 15, no. 1, p. 12, 2022.

[3] A. Chaudhuri, R. Datta, M. P. Kumar, J. P. Davim, and S. Pramanik, "Energy conversion strategies for wind energy system: electrical, mechanical and material aspects," *Materials*, vol. 15, no. 3, p. 1232, 2022.

[4] Z. Zhao, D. Wang, T. Wang, W. Shen, H. Liu, and M. Chen, "A review: approaches for aerodynamic performance improvement of lift-type vertical axis wind turbine," *Sustainable Energy Technologies and Assessments*, vol. 49, no. 49, article 101789, 2022.

[5] T. Asim, S. Z. Islam, A. Hemmati, and M. S. Khalid, "A review of recent advancements in offshore wind turbine technology," *Energies*, vol. 15, no. 2, p. 579, 2022.

[6] Q. S. Ali and M. H. Kim, "Power conversion performance of airborne wind turbine under unsteady loads," *Renewable and Sustainable Energy Reviews*, vol. 153, article 111798, 2022.

[7] N. Hoksbergen, R. Akkerman, and I. Baran, "The springer model for lifetime prediction of wind turbine blade leading edge protection systems: a review and sensitivity study," *Materials*, vol. 15, no. 3, p. 1170, 2022.

[8] C. Jung and D. Schindler, "On the influence of wind speed model resolution on the global technical wind energy potential," *Renewable and Sustainable Energy Reviews*, vol. 156, article 112001, 2022.

[9] J. Bi, C. Zheng, H. Huang, X. Song, and J. Li, "Fault diagnosis of wind turbine blades based on wavelet theory and neural network," *International Journal of Performability Engineering*, vol. 15, no. 7, 2019.

[10] X. Pan, Z. Cheng, Z. Zheng, and Y. Zhang, "Sparse Bayesian learning beamforming combined with short-time Fourier transform for fault detection of wind turbine blades," *The Journal of the Acoustical Society of America*, vol. 145, no. 3, p. 1802, 2019.

[11] F. Gao, X. Wu, Q. Liu, J. Liu, and X. Yang, "Fault simulation and online diagnosis of blade damage of large-scale wind turbines," *Energies*, vol. 12, no. 3, p. 522, 2019.

[12] J. Kang, L. Sun, and C. G. Soares, "Fault tree analysis of floating offshore wind turbines," *Renewable Energy*, vol. 133, pp. 1455–1467, 2019.

[13] M. Karimirad and C. Michailides, "Fault condition effects on the dynamic response of V-shaped offshore wind turbine," *Journal of Marine Science and Technology*, vol. 24, no. 1, pp. 34–45, 2019.

[14] A. A. Jiménez, L. Zhang, C. Q. Muñoz, and F. P. Márquez, "Maintenance management based on machine learning and nonlinear features in wind turbines," *Renewable Energy*, vol. 146, pp. 316–328, 2020.

[15] J. Nielson, K. Bhaganagar, R. Meka, and A. Alaeddini, "Using atmospheric inputs for artificial neural networks to improve wind turbine power prediction," *Energy*, vol. 190, article 116273, 2020.

[16] D. Xu, P. F. Liu, Z. P. Chen, J. X. Leng, and L. Jiao, "Achieving robust damage mode identification of adhesive composite joints for wind turbine blade using acoustic emission and machine learning," *Composite Structures*, vol. 236, article 111840, 2020.

[17] A. Movsessian, D. Garcia, and D. Tcherniak, *Proceedings of the 13th International Conference on Damage Assessment of Structures*, Adaptive feature selection for enhancing blade damage diagnosis on an operational wind turbine, Springer, Singapore, 2020.

[18] J. Solimine, C. Niezrecki, and M. Inalpolat, "An experimental investigation into passive acoustic damage detection for structural health monitoring of wind turbine blades," *Structural Health Monitoring*, vol. 19, no. 6, pp. 1711–1725, 2020.

[19] Y. Kim, G. Bangga, and A. Delgado, "Wind turbine airfoil boundary layer optimization using genetic algorithm with 3D rotational augmentation," in *Design Optimization of Wind Energy Conversion Systems with Applications*, IntechOpen, 2020.

[20] Z. Civelek, "Optimization of fuzzy logic (Takagi-Sugeno) blade pitch angle controller in wind turbines by genetic algorithm," *Engineering Science and Technology, an International Journal*, vol. 23, no. 1, pp. 1–9, 2020.

[21] C. Beale, C. Niezrecki, and M. Inalpolat, "An adaptive wavelet packet denoising algorithm for enhanced active acoustic damage detection from wind turbine blades," *Mechanical Systems and Signal Processing*, vol. 142, article 106754, 2020.

[22] G. Pujol-Vazquez, L. Acho, and J. Gibergans-Báguena, "Fault detection algorithm for wind turbines' pitch actuator systems," *Energies*, vol. 13, no. 11, p. 2861, 2020.

[23] A. Emami, A. Tavana, M. Mahnama, and A. Sadighi, "Model-based data-driven structural health monitoring of a wind turbine blade," *AUT Journal of Mechanical Engineering*, vol. 5, no. 1, pp. 97–108, 2021.

[24] M. Elgammi, T. Sant, and M. Alshaikh, "Predicting the stochastic aerodynamic loads on blades of two yawed downwind hawts in uncontrolled conditions using a bem algorithm," *Renewable Energy*, vol. 146, pp. 371–383, 2020.

[25] M. Awadallah and A. El-Sinawi, "Effect and detection of cracks on small wind turbine blade vibration using special Kriging analysis of spectral shifts," *Measurement*, vol. 151, article 107076, 2020.

[26] A. Saleem and M. H. Kim, "Aerodynamic performance optimization of an airfoil-based airborne wind turbine using genetic algorithm," *Energy*, vol. 203, article 117841, 2020.

[27] T. P. Philippidis and E. N. Eliopoulos, "A progressive damage mechanics algorithm for life prediction of composite materials under cyclic complex stress," in *Fatigue life prediction of composites and composite structures*, pp. 495–537, Woodhead Publishing, 2020.

[28] T. Krause and J. Ostermann, "Damage detection for wind turbine rotor blades using airborne sound," *Structural Control and Health Monitoring*, vol. 27, no. 5, article e2520, 2020.

[29] F. Liu, X. Ju, N. Wang, L. Wang, and W. J. Lee, "Wind farm macro-siting optimization with insightful bi-criteria identification and relocation mechanism in genetic algorithm," *Energy Conversion and Management*, vol. 217, article 112964, 2020.

[30] M. H. Qais, H. M. Hasanien, and S. Alghuwainem, "Whale optimization algorithm-based Sugeno fuzzy logic controller for fault ride-through improvement of grid-connected variable speed wind generators," *Engineering Applications of Artificial Intelligence*, vol. 87, article 103328, 2020.

[31] A. Joshuva and V. Sugumaran, "A lazy learning approach for condition monitoring of wind turbine blade using vibration signals and histogram features," *Measurement*, vol. 152, article 107295, 2020.

[32] X. Meng, P. Zhang, Y. Xu, and H. Xie, "Construction of decision tree based on C4.5 algorithm for online voltage stability assessment," *International Journal of Electrical Power & Energy Systems*, vol. 118, article 105793, 2020.

[33] Y. Cuan, Z. Wang, and J. Han, "Research on TV imaging casing damage detection and classification method based on C4. 5 decision tree," *Journal of Physics: Conference Series*, vol. 1437, no. 1, article 012132, 2020.

[34] R. Hariharan, I. S. Thaseen, and G. U. Devi, "Performance analysis of single- and ensemble-based classifiers for intrusion detection," in *Soft Computing for Problem Solving*, pp. 759–770, Springer, Singapore, 2020.

[35] N. Bhatt, A. Thakkar, N. Bhatt, and P. Prajapati, "Algorithm selection via meta-learning and active meta-learning," in *Smart Systems and IoT: Innovations in Computing*, pp. 169–178, Springer, Singapore, 2020.

[36] D. Lou, M. Yang, D. Shi et al., "K-means and C4. 5 decision tree based prediction of long-term precipitation variability in the Poyang Lake Basin, China," *Atmosphere*, vol. 12, no. 7, p. 834, 2021.

[37] I. H. Witten and E. Frank, "Data mining: practical machine learning tools and techniques with Java implementations," *ACM SIGMOD Record*, vol. 31, no. 1, pp. 76-77, 2002.

[38] Y. J. Chew, S. Y. Ooi, K. S. Wong, and Y. H. Pang, "Decision tree with sensitive pruning in network-based intrusion detection system," in *Computational Science and Technology*, pp. 1–10, Springer, Singapore, 2020.

[39] Y. Qu, B. L. Adam, Y. Yasui et al., "Boosted decision tree analysis of surface-enhanced laser desorption/ionization mass spectral serum profiles discriminates prostate cancer from noncancer patients," *Clinical Chemistry*, vol. 48, no. 10, pp. 1835–1843, 2002.

[40] Y. Zhao and Y. Zhang, "Comparison of decision tree methods for finding active objects," *Advances in Space Research*, vol. 41, no. 12, pp. 1955–1959, 2008.

[41] R. C. Holte, "Very simple classification rules perform well on most commonly used datasets," *Machine Learning*, vol. 11, no. 1, pp. 63–90, 1993.

[42] P. Geurts, D. Ernst, and L. Wehenkel, "Extremely randomized trees," *Machine Learning*, vol. 63, no. 1, pp. 3–42, 2006.

[43] D. Ernst, P. Geurts, and L. Wehenkel, "Tree-based batch mode reinforcement learning," *Journal of Machine Learning Research*, vol. 6, pp. 503–556, 2005.

[44] I. Ibarguren, J. M. Pérez, J. Muguerza, I. Gurrutxaga, and O. Arbelaitz, "Coverage-based resampling: building robust consolidated decision trees," *Knowledge-Based Systems*, vol. 79, pp. 51–67, 2015.

[45] J. M. Pérez, J. Muguerza, O. Arbelaitz, I. Gurrutxaga, and J. I. Martín, "Combining multiple class distribution modified subsamples in a single tree," *Pattern Recognition Letters*, vol. 28, no. 4, pp. 414–422, 2007.

[46] G. I. Webb, "Decision tree grafting from the all-tests-but-one partition," in *IJCAI'99: Proceedings of the 16th international joint conference on Artificial intelligence - Volume 2*, pp. 702–707, Stockholm Sweden, 1999.

[47] V. K. Pachghare and P. Kulkarni, "Pattern based network security using decision trees and support vector machine," in *2011 3rd International Conference on Electronics Computer Technology*, Kanyakumari, India, 2011.

[48] G. Holmes, B. Pfahringer, R. Kirkby, E. Frank, and M. Hall, "Multiclass alternating decision trees," in *European Conference on Machine Learning*, pp. 161–172, Springer, Berlin Heidelberg, 2002.

[49] R. Kohavi, "Scaling up the accuracy of Naive-Bayes classifiers: a decision-tree hybrid," in *KDD'96: Proceedings of the Second International Conference on Knowledge Discovery and Data Mining*, pp. 202–207, Portland Oregon, 1996.

[50] C. Thornton, F. Hutter, H. H. Hoos, and K. Leyton-Brown, "Auto-WEKA: combined selection and hyperparameter optimization of classification algorithms," in *Proceedings of the 19th ACM SIGKDD international conference on Knowledge discovery and data mining*, pp. 847–855, Chicago Illinois USA, 2013.

[51] L. Breiman, J. Friedman, C. J. Stone, and R. A. Olshen, *Classification and Regression Trees*, CRC press, 1984.

[52] C. N. Villavicencio, J. J. Macrohon, X. A. Inbaraj, J. H. Jeng, and J. G. Hsieh, "COVID-19 prediction applying supervised machine learning algorithms with comparative analysis using WEKA," *Algorithms*, vol. 14, no. 7, p. 201, 2021.

[53] G. Zeng, "On the confusion matrix in credit scoring and its analytical properties," *Communications in Statistics-Theory and Methods*, vol. 49, no. 9, pp. 2080–2093, 2020.

[54] C. H. Han, E. Kim, and C. H. Im, "Development of a brain–computer interface toggle switch with low false-positive rate using respiration-modulated photoplethysmography," *Sensors*, vol. 20, no. 2, p. 348, 2020.

[55] S. Suresh and S. Mohan, "ROI-based feature learning for efficient true positive prediction using convolutional neural network for lung cancer diagnosis," *Neural Computing and Applications*, vol. 32, no. 20, pp. 15989–16009, 2020.