Vibration Analysis of Hydrodynamic Conical Journal Bearing and Fault Prediction Using Machine Learning

Harshal Bhat
Veermata Jijabai Technological Institute (VJTI), Mumbai, PG Dynamics Lab, HR Mahajani Marg, Matunga(E)

Rahul Yadav
Veermata Jijabai Technological Institute (VJTI), Mumbai, PG Dynamics Lab, HR Mahajani Marg, Matunga(E)

Sarang Bhudhar
Veermata Jijabai Technological Institute (VJTI), Mumbai, PG Dynamics Lab, HR Mahajani Marg, Matunga(E)

Avinash Mandarha
Veermata Jijabai Technological Institute (VJTI), Mumbai, PG Dynamics Lab, HR Mahajani Marg, Matunga(E)

他

https://doi.org/10.5109/7342454

出版情報：Evergreen. 12 (1), pp.278-302, 2025-03. Interdisciplinary Graduate School of Engineering Sciences, Kyushu University, Japan
バージョン：

EVERGREEN Joint Journal of Novel Carbon Resource Sciences & Green Asia Strategy, Vol. 12, Issue 01, pp287-302, March, 2025

Vibration Analysis of Hydrodynamic Conical Journal Bearing and Fault Prediction Using Machine Learning

Harshal Bhat[1], Rahul Yadav[1], Sarang Bhudhar[1], Avinash Mandarha[1], Vikas M. Phalle[1]

[1]Veermata Jijabai Technological Institute (VJTI), Mumbai, PG Dynamics Lab, HR Mahajani Marg, Matunga(E), Mumbai-400019. Maharashtra. India.

*Author to whom correspondence should be addressed:
E-mail: rkyadav2017india@gmail.com

(Received June 7, 2024; Revised January 15, 2025; Accepted February 1, 2025).

Abstract: In Industry 4.0, condition monitoring is gaining significant importance due to its capability for early diagnosis of the symptoms of the faults that result in severe damage in the rotating machinery. Hydrodynamic conical journal bearings (HCJB) are used in heavy rotating machinery due to their high radial and axial load-carrying capacity and need to be adequately monitored during the operation. Thus, an experimental study is carried out on a healthy and faulty CJB which is scratched due to frequent start and stop operation, i.e. worn-out internal surfaces of bearing. Therefore, the dynamic response of the journal changes significantly due to wear. The vibration response of bearing in the time domain was acquired and transformed into the frequency domain and a comparative analysis is carried out between the bearings. The experimental results demonstrate the unique vibration characteristic of worn-out bearing at the 4th multiple of operating frequency. Machine learning techniques have also been used as fault diagnosis approaches. A comparative study using machine learning methods i.e. SVM (Support Vector Machine), k-NN (k-Nearest Neighbour) and RF (Random Forest) has been performed. The results show that the RF classifier is more proficient in fault diagnosis with an accuracy of 93.93% compared to the other methods i.e. k-NN and SVM.

Keywords: Condition monitoring; Hydrodynamic Conical Journal Bearing (HCJB); Support Vector Machine (SVM); k-NN (k-Nearest Neighbour); RF (Random Forest)

1. Introduction

Bearings can be categorized based on their structural configuration into several types, including sliding bearings, rolling bearings, deep groove ball bearings, angular contact bearings, and thrust bearings. Additionally, from a material standpoint, bearings can be classified into metal bearings, non-metallic bearings, and porous metal journal bearings.

A journal bearing is a conventional type of hydrodynamic bearing employed in rotary machinery. A circular shaft called a journal is allowed to rotate inside a fixed sleeve called the bearing which forms its primary construction. It supports high loads due to the wedging action between the journal and bearing because of considerable relative motion[1)]. Full lubrication occurs, resulting in no contact between the bearing and the journal because of high relative motion. These bearings find application in IC engines for supporting the crankshaft, rotor assemblies, and turbochargers[2)].

The Conical hydrodynamic journal bearing can sustain high radial loads and mitigate axial load effects on rotating elements. Journal and thrust hydrodynamic bearings are being replaced by conical journal bearings due to their compact size, low cost, radial and axial load bearing capacity. These are used in supporting precision spindles of rotating equipments[3)]. When a journal bearing is operational, the surface asperities of the bearing lining and the journal interact and the mating surfaces wear out. Since bearing material is softer than journal material, its wear rate is higher and it is generally defined as abrasive wear or scratching[4)]. Abrasive wear is one of the most critical failure modes of the journal bearings[5)]. Wedge action refers to the pressure created by the lubricating film and requires a definite minimum speed to occur. During the start and shutdown sessions, this criterion is not met as the bearing slowly accelerates to the operating speed before decelerating to zero. During these instances, the journal and bearing rub against each other, causing wear on the surfaces[6)]. Eccentricity, film thickness, and load-carrying capacity are affected due to wear[7)]. The most common wear observed in lubricated machinery is abrasive wear or scratching, which occurs due to particle contamination[8)].

To allow preventive action against catastrophic failures and decrease the probable loss of production due to the breakdown of the machines, early detection of fault symptoms is important and it is detected by the condition monitoring of the bearing and thus mitigates the downtime during running, and leads to the economic operation of systems. Hence, monitoring techniques can be employed in the early wear regime such as vibration and vibro-acoustic analysis[34]. Moreover, bearing performance analysis and lubricating oil analysis (fluid contamination and wear debris analysis) can also be employed for monitoring journal bearings[9,10,11]. The vibration is a result of geometrical imperfections on the bearing components caused while manufacturing and due to wear and tear during operation[12]. The main idea behind vibration analysis is that if a system presents a change of its dynamic response, then such change can be related to an undergoing fault in the system thus making it suitable for the automatic fault diagnosis using machine learning algorithms[53]. Hence vibration analysis is one of the extensively used condition monitoring techniques that can produce useful quantitative data and be controlled remotely in real-time and it is considered as one of the best methods in determining the machine condition. Alarm limits can also be pre-set and automatically activated based on observed signals[13]. By converting vibration signals from time domain to the frequency domain, the Fast Fourier transform (FFT) can identify abnormal frequencies associated with bearing defects[38]. Each type of fault has a characteristic component of vibrational frequencies that can be filtered and defined. The amplitude of each vibration component will remain constant if the dynamics of the technical system remain unchanged[30].

Tribology has produced a significant number of multi-dimensional and multi-structural research data due to its unique system dependence, multidisciplinary coupling, and time dependence, which has become the basis of the integration of AI (Artificial Intelligence) in tribology[40-43]. Therefore, the effectiveness of data acquisition, integration, analysis, and research in tribology would be improved by the integration of AI methods and promote its development. For example, the capabilities of AI in data and pattern recognition can establish correlations between various signals (e.g., vibration, acoustics, electrical, and sound pressure) and wear, making early prediction of mechanical wear feasible. Tribological behaviors under different operational states can be analyzed and diagnosed by AI, which is crucial in designing more robust and reliable mechanical systems[44]. In 2017, Wu et al. employed RF algorithms for wear prediction and generated considerable interest among researchers[45]. In general, prediction is the most common application purpose in “AI for tribology”. Among the different components in tribology, bearings are the most complex and widely used component, and AI technology is the most widely used in bearing research[46,47]. Clustering and classification methods within AI are specifically useful in understanding tribological behavior[44].

To increase fault prediction performance, signal processing and AI technologies are combined by extracting characteristics from the vibration signals. Thus, the raw vibration signals supplement features for fault classification in the time-frequency domain and their subsequent time and frequency domains are then passed to high-fidelity artificial intelligence methods such as RF, SVM, and ANN (Artificial Neural Network)[14-18]. The GA (Genetic Algorithm) based[19] strategy was used to improve the input feature selection and the relevant classifier parameters. Samanta[20] focused on a genetic algorithm-based bearing fault detection using SVM and ANN. SVM performed substantially better than ANN with the entire feature set.

SVM is a novel computational and statistical learning theory that may be used as an expert system[21]. With finite samples and estimating (learning) dependencies, this theory incorporates essential ideas and learning principles, well-defined formulation, and self-contained mathematical theory. SVMs have been used in machine learning, computer vision and pattern recognition for their high accuracy and good generalization capabilities making them suitable for fault diagnosis in bearings[22,36]. Ziwei Wang[23] proposes a hybrid approach of random forest classifiers for fault diagnosis in rolling bearings using vibration signals and the experiment results show that the method is capable to handle high-dimensional characteristic data, and is highly efficient, robust and accurate, making it suitable to be deployed for the diagnosis in case of vibration signals.

Given the typically limited data produced from a series of tribological experiments, traditional machine learning algorithms like Random Forest (RF)[48-49] and Support Vector Machines (SVM)[50] continue to be valuable. Unlike deep learning methods, which generally require substantial training datasets to reach optimal accuracy, these classical approaches can perform effectively with smaller datasets. This is particularly advantageous in industry-oriented research settings, where the availability of large data volumes is often constrained by time and financial limitations[51]. Hence SVM and RF algorithms are selected for carrying out the binary classification.

SVM and k-NN machine learning models have been used to classify the different defects associated with ball bearing elements and results show that the k-NN classifier has a higher level of accuracy than the SVM[24]. k-NN classifier has been widely used for bearing fault diagnosis and the results enable fault detection with higher accuracy[25]. Especially in problems with abnormal and unidentified distributions, the k-NN algorithm, a nonparametric classification technique can produce excellent classification accuracy. Consequently k-NN algorithm is chosen for the classification considering its benefits over other supervised machine learning algorithms.

In this workflow, vibration signals of healthy and faulty bearings were recorded when mounted on the CJBTR (Conical Journal Bearing Testing Rig) running under different loading conditions. The features derived from vibration signals were fed into the SVM, RF and k-NN models, the models were trained, tested and hyperparameters were tuned to improve the classification accuracy. Thus, the workflow proposes an ML-based condition monitoring methodology for fault prediction in the HCJBs in the field of tribology.

In the succeeding portion of the paper, section 2 discusses the principle and working of HCJBs, section 3, 4 and 5 discuss experimental setup, vibration recording and framework respectively. Section 6 and 7 elaborates on the features-hyperparameter selection and machine learning classifiers respectively. Results and conclusion are discussed in the last sections 8 and 9 respectively.

2. Conical Journal Bearings

The current section describes in detail the following: (1) Introduction of the hydrodynamic conical journal bearings, its advantages and applications (2) Principle of operation of the journal bearings

2.1 Hydrodynamic Journal Bearings

Hydrodynamic journal bearings are one of the most critical components in turbo-machineries i.e. centrifugal pumps, compressors, and turbines. These bearings are well known for sustaining high radial loads but cannot withstand significant axial loads. In ship shafting propulsion systems and other marine fields, journal bearings are definitely among the most crucial components. Marine shafts often operate at a lower speed to achieve desirable propeller characteristics (such as high efficiency). Such a situation may hurt the behavior of journal-bearing.

Hence design modifications were carried out & conical journal bearing came into the picture with two dominating advantages, it can sustain radial as well as axial loads and provides variable clearances. These bearings substituted the journal assembly and thrust bearings, and are also used in precision tool applications i.e. lathes and grinding machines.

2.2 Principle of Operation

The test bearing surface on the inner diameter is separated from the outer surface of the journal by the lubricant film generated while in rotation, the rotation of the journal causes pumping of the oil to flow around the inner surface of the bearing in the direction of rotation. If there are no forces applied on the bearing, the oil film and bearing position will remain concentric with journal circumference. However, when the load is applied on the bearing, the bearing shifts from its concentric position creating a converging gap between the bearing and the journal surface. The pumping action caused due to rotation of the journal forces the oil to squeeze through the gap thereby generating hydrodynamic pressure along the convergent zone. The pressure decreases to the cavitation pressure in the diverging gap zone. Where cavitation forms, the oil pressure creates a counter balancing force thereby separating the bearing from the journal surface. The force exerted by lubricant pressure and the hydrodynamic friction forces counterbalances the external forces applied i.e. axial and radial loads as shown in Fig. 1 below. Equilibrium between the forces decides the stable position of the bearing.

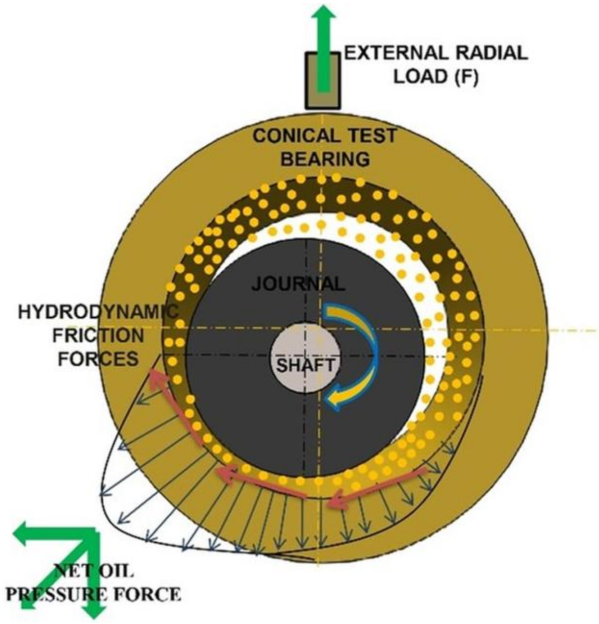

Fig. 1: Counterbalancing of the external radial load.

3. Experimental Setup and Procedure

Figure 2 shows the system arrangement of CJBTR consisting of a machine control unit for adjusting the running and loading conditions, an air supply system for actuating the cylinders for applying the radial load, a lubricant re-circulating unit having in-line water-cooled heat exchanger used during high-speed and load conditions and DAQ (Data Acquisition) system for recording the vibration signals collected by the trial-axial ICP (Integrated Circuit Piezoelectric) accelerometer mounted on the HCJB having x, y and z faces aligned along the lateral, longitudinal and radial directions of the bearing respectively. NV Gate software is used for controlling the operating parameters of the bearing and OROS DAQ unit for capturing the measured vibration signals and displaying it over the computer unit.

The machine is allowed to run and radial load is applied by pulling the test bearing upwards by a pneumatic cylinder which is mounted on a frame in the vertical direction with its piston pointing downward in line with the vertical axis of the journal. The bearing is pulled upward by a metallic chain fixed to the pneumatic piston; a load cell is connected between the chain and piston to measure the force applied as shown in Fig. 3. Load and speed conditions are controlled by the machine control unit. The radial load and speed conditions as shown in

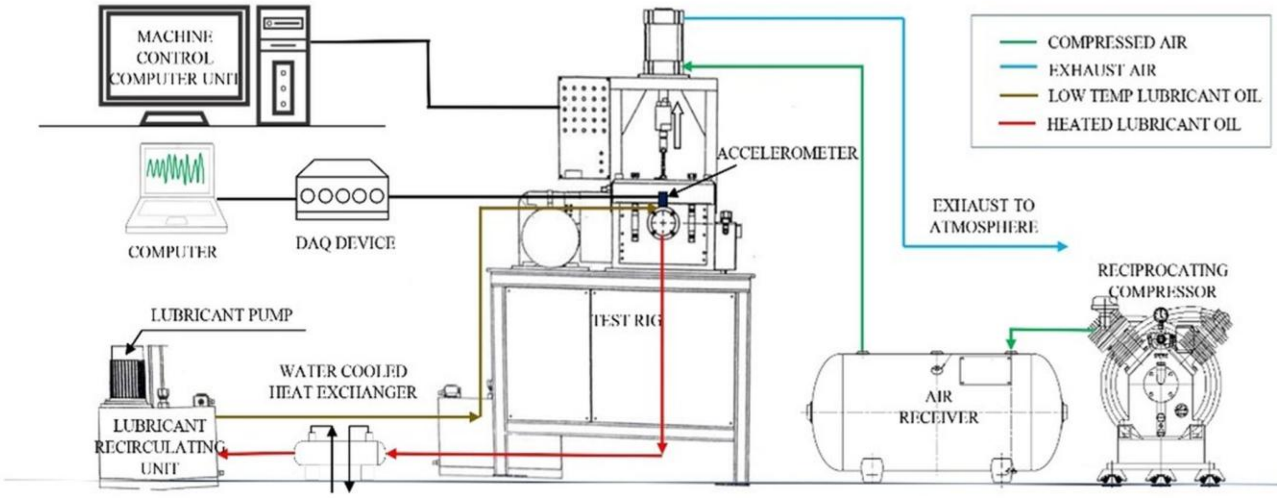

Fig. 2: Schematic diagram of CJBTR

Table 1 are followed in case of HB with perfectly finished surfaces and FB having radial and longitudinal scratches on the internal surface of the bearing as shown in Fig. 4 and Fig. 5 respectively and simultaneously the vibration signals are recorded for the conditions.

Table 1. Loading conditions

Constant Speed & Variable Radial load		Constant Radial load &Variable Speed	
Speed (RPM)	Load (N)	Speed (RPM)	Load (N)
500	0	600	350
	250	700	
	350	800	
	450	900	
	550	1000	
	650	1100	
	750	1200	
	850	1300	
	950	1400	
	1050	1500	

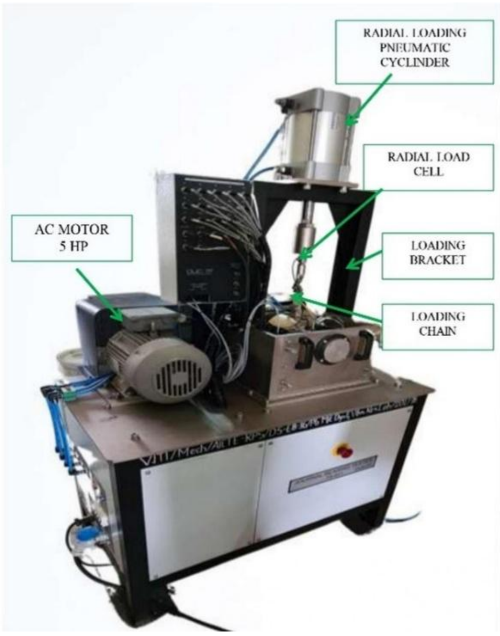

Fig. 3: Test rig of conical journal bearing.

Fig. 4: Healthy bearing

Fig. 5: Faulty bearing

4. Recording Vibration Signals

The healthy bearing shown in Fig. 4 is installed on the test rig, assembled with different pressure and temperature sensors and both actuators are linked with chains. After assembly, the procedure shown in Fig. 6 is followed for recording the vibration signals on the test rig. The vibration signals are recorded based on the two conditions indicated in Table 1. The healthy bearing is then dismantled and the scratched bearing having radial and longitudinal scratches is mounted on the test rig, and the same procedure is followed for recording the vibration signals.

Firstly, the accelerometer is positioned at the appropriate location as shown in Fig.6 on the bearing. The axes X, Y and Z of the trial axial accelerometer are aligned along the lateral, longitudinal, and radial direction of the CJB respectively as shown in Fig. 6. The accelerometer is connected to the DAQ and analyzer is connected to the desktop unit with software installed on it which indicates the signals i.e. accelerometer v/s time graph of all directions with simultaneous single sided amplitude spectrum. The accelerometer measures the signals of all three directions i.e. radial (z), lateral(x) and longitudinal(y), as the process has been well illustrated by the flow chart shown in Fig. 7.

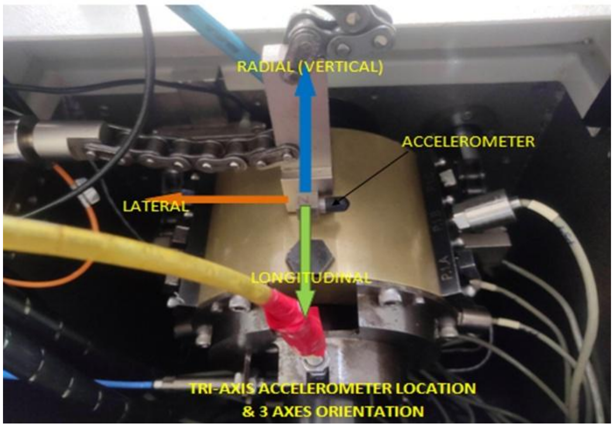

Fig. 6: Accelerometer installation and orientation.

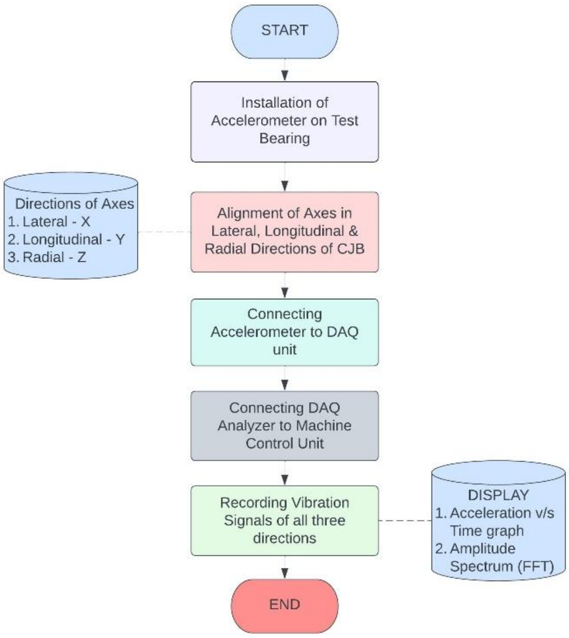

Fig.7: Vibration recording flow chart.

5. Framework

The framework in Fig. 8 shows the overview of the entire study done on the HCJB. Vibration signals are recorded and various feature vectors are extracted from the raw vibration data. Further extracted and selected features are fed to the different models i.e. SVM, RF and k-NN models followed by the training and testing of the models and bearing fault classification. The raw vibration data is also Fast Fourier Transformed (FFT) to understand the dynamic nature of HCJB, its unique vibration characteristics and it is compared with the faulty bearing.

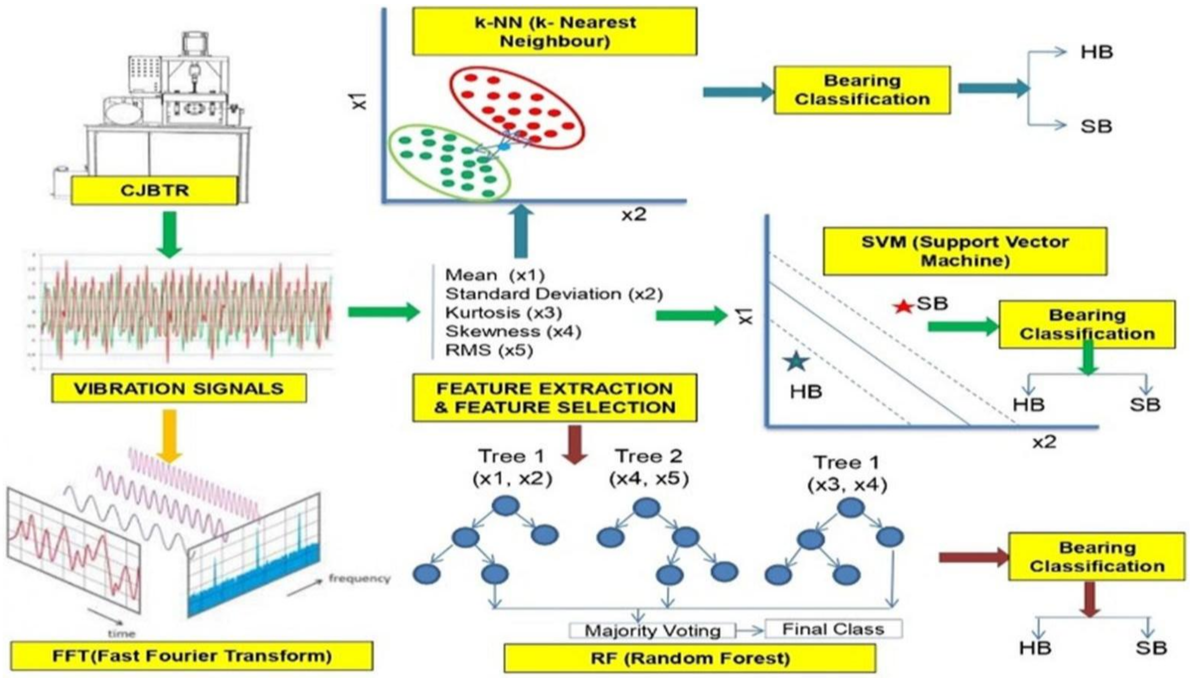

Fig. 8: Framework of the study.

6. Fault Diagnosis using Machine Learning Methods

Data-driven methodologies are frequently utilized in the problem diagnosis process in the process industries, which monitor pumps, bearings, etc. The development of sensor and computing technology is the basis for utilizing a data-driven approach. The most recent data-driven strategy includes monitoring the threshold values of the signals acquired from the sensors located on the machineries[30].

False positive alarms will occur frequently if the threshold value is set during the fault diagnostic process to a value that is too low; on the other hand, if the threshold value is set to a value that is too high, there is a probability that possible risk will not be identified in the earlier stages. The data-driven technique is also a lengthy procedure due to uncertainties in process variability and a lack of process understanding, presence of signal noise and missing data. To overcome these challenges, researchers have begun implementing ML-based defect diagnosis algorithms. The machine learning process for problem identification includes the processes of data gathering, feature extraction, feature selection and decision-making.

The machine learning algorithms are evaluated and developed using previously collected data. SVM, RF, and k-NN are a few examples of some of the most appreciated machine-learning techniques for defect diagnostics.

6.1 Feature Extraction

The technique of feature extraction retrieves meaningful data from the collected signal, illustrating the equipment's condition through signal signature. The statistical feature can be used to detect vibration signals with a pattern produced by faults in the equipment. The approach has been applied and validated by the researchers. This establishes statistical features as one of the popular feature extraction techniques with the least amount of complexity and time consumption[26–28]. The computed feature vectors are maximum value, mean, median, mode, minimum, kurtosis, range, standard error, standard deviation, RMS (Root Mean Square), variance and skewness. These are calculated using the standard formulae[39].

6.2 Feature Selection

The process of selecting the most essential features from descriptive features for classification algorithm performance is known as feature selection.

Only a few of the extracted features will have a positive impact on the categorization process. In order to prevent a decline in classification performance, certain important features are chosen during the feature selection process while the rest are eliminated.

The features having excessive variations (outliers) and error values are eliminated from the categorization process.

After multiple runs of the ML models with features on varying batch sizes were performed and following features hold the maximum significance in determination of the fault: mean, standard deviation, kurtosis, skewness and RMS value.

6.3 Hyperparameter Tuning

In case of classification problems of bearings, diagnostic accuracy is the most critical, since any error may lead to machinery failure and production losses. In these cases, tuning or optimizing the hyperparameters can significantly improve the classification accuracy and generalization ability of the models[52].

In first stage of the classification, the models were trained with the default hyperparameter settings and initial accuracies were obtained. Then the hyperparameters were optimized. C value, γ value and kernel function for SVM, number of estimators for RF and k (neighbors) value for k-NN are optimized hyperparameters. These key hyperparameters were fine-tuned using grid-search approach and optimized values are discussed in the performance assessment section. A significant improvement was observed in the accuracies after the optimization of the hyper-parameters. Confusion matrices were also obtained for the optimized values.

7. Machine Learning Classifiers

Based on the values of the given features, the class (HB or FB) is determined using the techniques of the ML classifiers. The type of classification algorithm used determines the classification process. In this study, different classifiers i.e. SVM, k-NN and RF are used for the bearing classification using the same dataset features for all the classifiers.

7.1 SVM (Support Vector Machine)

Figure 9 shows the flowchart of SVM algorithm. The recorded dataset (132x5x2) consists of 5 normalized features with 132 sample cases and 2 classes (healthy & faulty bearing) which are split into two subsets for training and testing the SVM algorithm. Statistical feature vectors i.e. mean, skewness, kurtosis, standard deviation, and RMS are calculated from the vibration signals data sets. For this algorithm, the Gaussian Radial Basis Function (RBF) kernel is used since it is a general-purpose kernel. Estimation and approximation errors are minimized when it is used with appropriate regularization parameters. While using SVM with Gaussian RBF, a trade-off has to be done between the parameters i.e. C and γ, where C is the inverse of regularization parameter λ, where λ is used in linear regression to address the problem of over-fitting. Increasing or decreasing the value of C and γ results in over-fitting or under-fitting of the dataset. A set of combinations of these hyper-parameters are fed to the algorithm and an average score is calculated to select the best hyper-parameters. 500 and 5 were found to be the best hyper-parameters for C and γ, giving the highest average score of 87.88%.

7.2 RF (Random Forest)

The recorded faulty datasets with five features are re-sampled using the RFs technique and then they are split into training and testing sets.

The bootstrapping re-sampling technique is used to create the corresponding new training datasets randomly using the original datasets followed by the random feature selection for the bootstrapped datasets i.e. two features for the datasets (square root of the total number of the features). Many decision trees are produced using the training dataset, and the trees then result in RFs as shown in Fig.10. The trees generally grow together without being pruned. Internal nodes represent the features of a dataset, branches reflect the decision-making process, and each leaf node represents the classification outcome in decision trees. It is a visual depiction of finding every solution that is possible under the circumstances of a choice or problem. This is integrated to create the RFs classifier. The test datasets are passed through the different decision trees of the RF. The corresponding categories are determined by analyzing the test dataset. Multiple results are predicted from the decision trees and aggregated. Further majority voting decides the final classification results. The category with the largest output class is selected as the testing set category and the fault diagnosis is finished[23].

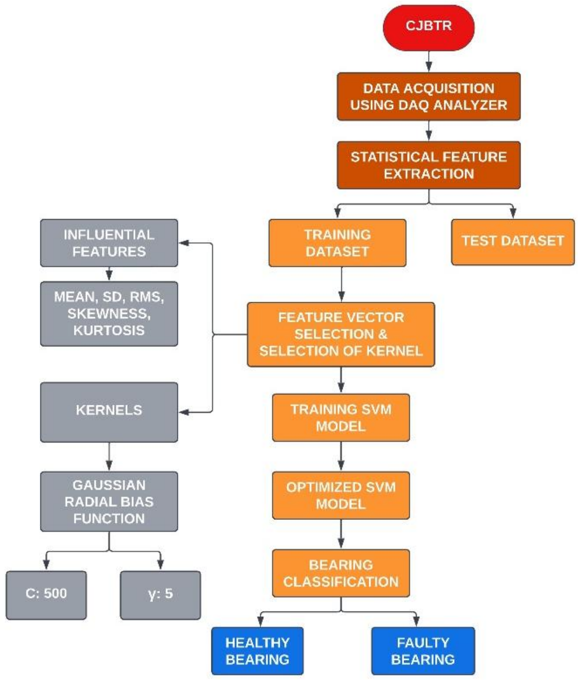

Fig. 9: SVM fault detection flow.

7.3 k-NN (k-Nearest Neighbour)

An N X M data matrix serves as the foundation for the k-NN method. N and M stands for the total number of data points and the size of each data point respectively. The data matrix's k closest points to the query point are searched for. The Euclidean distance between the query and the remaining points of the data matrix is calculated. The results of this procedure are N Euclidean distances which indicate the distances between each relevant point in the data set and the query as shown in Fig. 11. The k points (k=5, default value) with the least distance between the data set and the query may then be easily found by sorting the distances in ascending order and retrieving the k points[29].

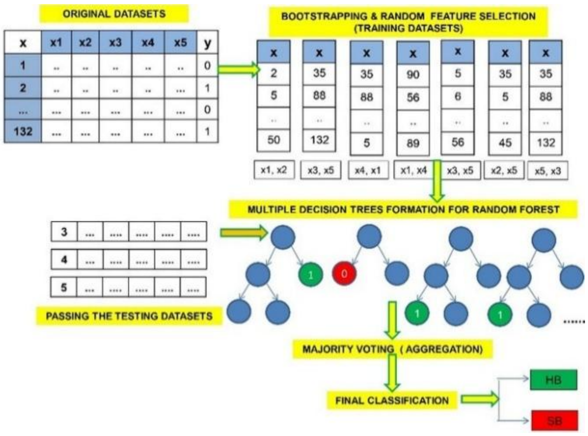

Fig. 10: Random forest fault detection flow.

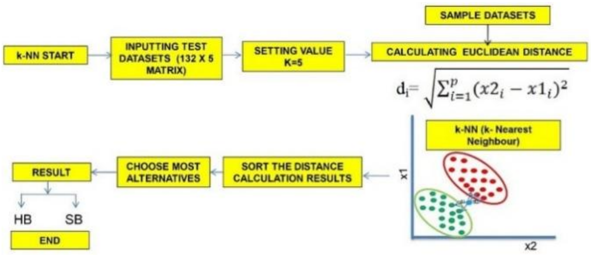

Fig. 11: k-NN algorithm fault detection flow.

8. Results & Discussions

This section describes the results as follows:

(1) FFT- Amplitude Spectrum Comparison

(2) ML classifiers SVM, RF and k-NN with their confusion matrices and feature vectors comparison plots

8.1 Comparison between HB and FB

Figure 12 shows the comparison of the acceleration amplitudes in the Y-direction (longitudinal), the direction across which dominating vibration signals are observed. This is the plot of the recorded vibration signals. Amplitude values range from 3.29 to -2.67 m/s2 in case of scratched bearing and 1.83 to -1.72 m/s2 in case of healthy bearing are observed at the same operating condition. The maximum values help in setting up the alarming values for the system35). The operating frequency (OF) is 1500 RPM (25 Hz).

Characteristic spectral components pointing to the defect of the bearing can be detected by the FFT (Fast Fourier Transform) spectrum of the vibration signals[30,34,35]. Hence the signals are Fast Fourier transformed (FFT) and frequency-amplitude spectrums are obtained, following are the observations:

Table 2. Observations of amplitude spectrums.

Sr. No.	Bearing	Peak Value (m/s^2)	Peak Frequency (Hz)
1.	FB/SB	0.3441	49.95 (2x Multiple of OF)
2.	HB	0.8773	99.99 (4x Multiple of OF)

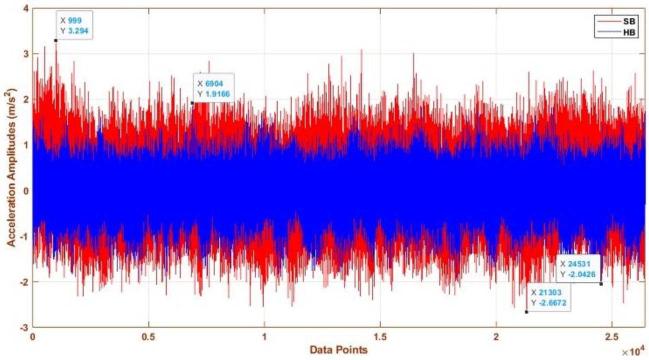

Fig. 12: Acceleration amplitudes comparison between HB & SB at 1500 RPM & 350 N.

From the observations shown in Table 2 observed from the Fig. 12 and 13, the peak frequency shifts at the higher values in case of SB (Scratched Bearing).

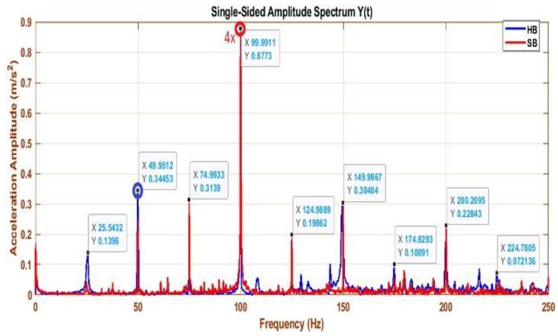

Fig. 13: Amplitude spectrum comparison between HB & SB at 1500 RPM & 350 N.

8.2 Performance Assessment of the Classifiers

This section reviews the performance of three different classifiers. The extracted optimum features of the acquired signals are fed to the different ML classifiers and their performances are compared based on the different parameters and accuracies[33].

Table 3. Optimal hyper-parameters for the ML classifiers.

Machine Learning Algorithm	Optimum Hyper-parameter
SVM	C: 500, γ: 5, kernel: RBF (radial basis function)
RF	No. of estimators= 100
k-NN	k= 5 (neighbors)

Table 3 shows the description of the optimum hyper-parameters used for executing the classification for the respective models, these parameters show a significant role in deciding the cross-validation accuracies of the models.

8.2.1 Features Comparison

The extracted features are compared to determine their behavior. The extracted features for healthy and faulty bearings at the above-defined loading conditions are compared as shown in Fig. 14, 15, 16, 17 and 18.

Table 4. Fluctuating ranges of the feature vectors.

Sr. No.	Feature Vector	Fluctuating Range (m/s^2)	
		HB	FB
1	Mean	(-0.0053) to 0.0104	(-0.0167) to 0.0192
2	Standard Deviation	0.4121 to 0.9076	0.7236 to 0.9150
3	Kurtosis	(-0.1184) to (-1.1704)	(-1.036) to 2.1115
4	Skewness	(-0.1388) to 0.2293	(-0.2052) to 0.2654
5	RMS	0.4121 to 0.9076	0.7237 to 0.9152

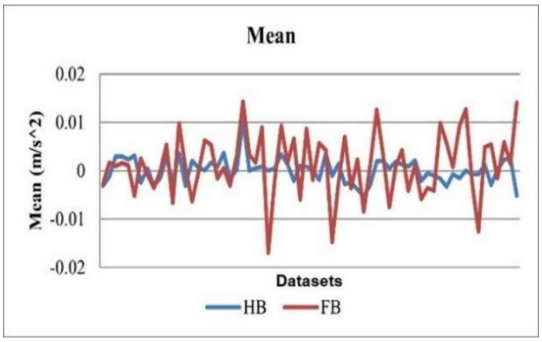

Fig. 14: Comparison between mean values.

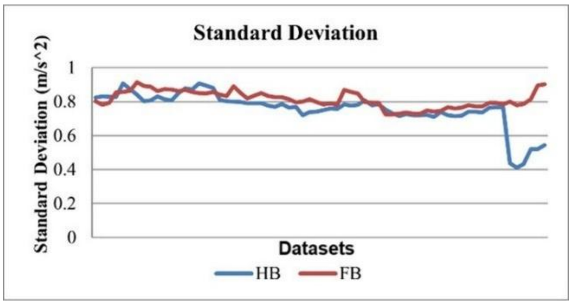

Fig. 15: Comparison of standard deviation values.

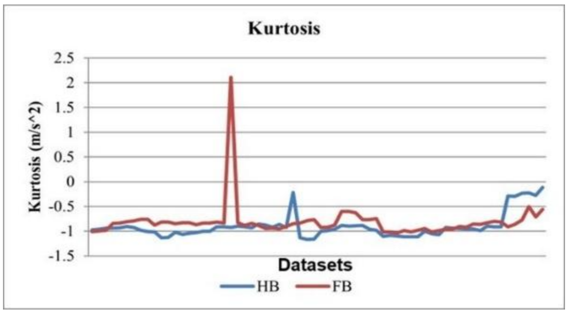

Fig. 16: Comparison of kurtosis values.

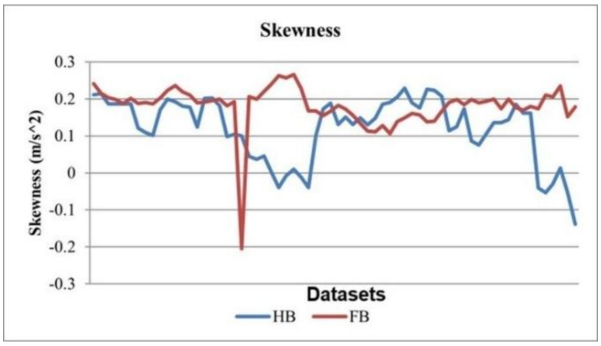

Fig. 17: Comparison between skewness values.

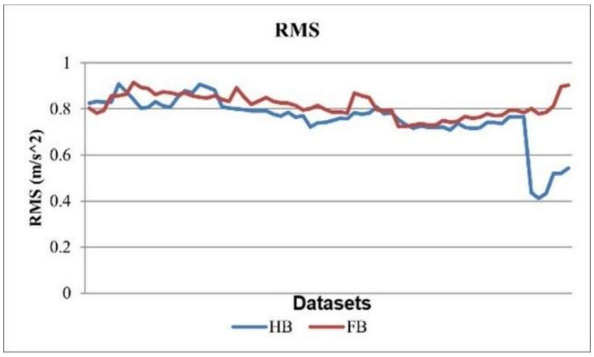

Fig. 18: Comparison between RMS values.

From the comparison plots (Fig. 14-18) of the features, the fluctuating ranges of the features have been determined as mentioned in Table 4. It can be inferred that mean, kurtosis, and skewness have predominating variations in case of FB compared to HB whereas other features have lesser variations except the abrupt changes caused due to noises involved in the signals. These features have dominating effects in the classification process.

8.2.2 ML Algorithm Results

The Results of the ML models are obtained in the form of a confusion matrix[37)]. It allows us to visualize the performance of the algorithm especially in case of classification problems. Any confusion matrix consists of TP, FP, TN, and FN arranged as shown in the Fig. 19, several other performance parameters can also be calculated using the below standard formulae as shown in figure below.

	1	0
1	TP (True Positives)	FN (False Negatives)
0	FP (False Positives)	TN (True Negatives)

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN}\ ;\ Recall\ Value = \frac{TP}{TP+FN}$$

$$Precision\ Value = \frac{TP}{TP+FP}\ ; F1\ Score = 2\ X\ \frac{Precision\ X\ Recall}{Precision + Recall}$$

Fig. 19: Confusion matrix terminology.

A 2x2 confusion matrix as shown in Fig. 20 and 21 below is obtained which predicts the classification between two classes i.e. 'Healthy Bearing' and 'Faulty Bearing'. For the training set as shown in Fig. 20, out of 99 cases, the algorithm predicts the right decision in 92(46+46) cases which gives an accuracy of 92.93% whereas in the test set as shown in Fig. 21, out of 33 examples, the right prediction is made for 29 examples which provide a validation accuracy value of 87.88%. Similarly, Fig. 22, 23, 24 and 25 depict the training and testing confusion matrices of RF and k-NN respectively.

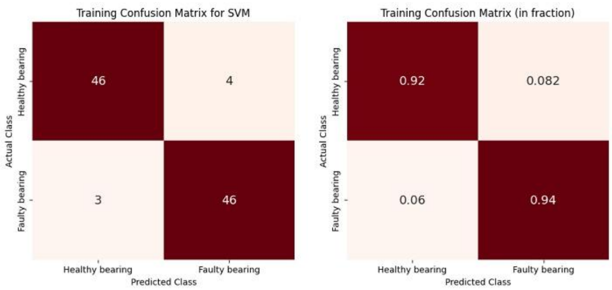

Fig. 20: Training confusion matrices for SVM.

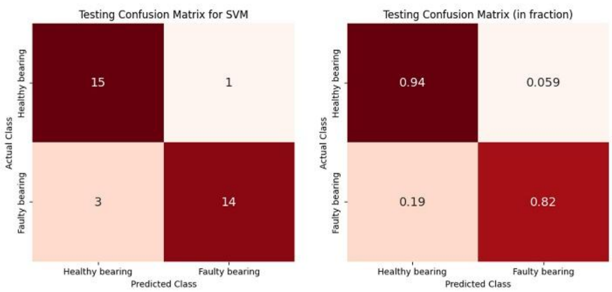

Fig. 21: Testing confusion matrices for SVM.

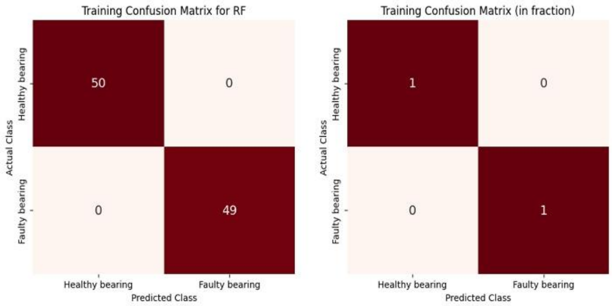

Fig. 22: Training confusion matrices for RF.

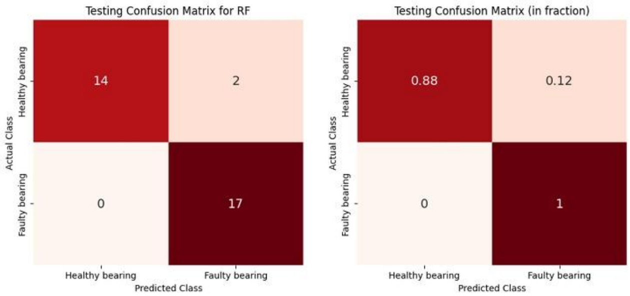

Fig. 23: Testing confusion matrices for RF.

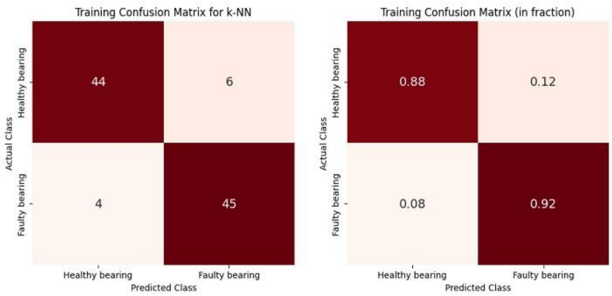

Fig. 24: Training confusion matrices for k-NN.

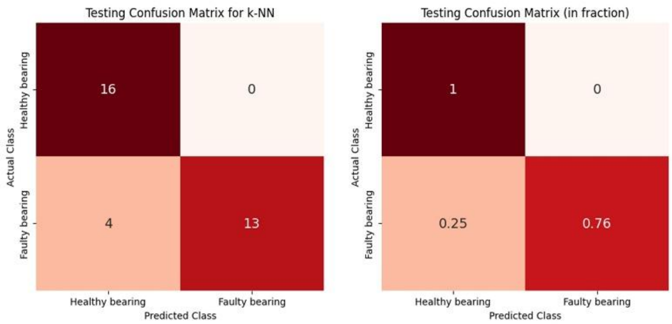

Fig. 25: Testing confusion matrices for k-NN.

8.2.3 Illustration of RF Results: Decision Tree

Each internal node in the decision tree generated has a decision rule that divides the data. Each node in Fig. 26 has five important pieces of information about it.

In reference to the decision tree, x0, x1, x2, x3 and x4 refer to as mean, standard deviation, kurtosis, skewness & RMS value respectively. The primary criterion for the division is the first item of data.

As an illustration, the first node's condition is "x0≤0.004". At this root node of the prediction processing, the sample is split into two groups depending on whether its x0 value is less than or equal to 0.004 or not.

The second piece of data in the node represents the Gini index value generated by the Gini method to generate split points. The final piece of information is the number of examples (from the training set) that contain the node (99 for the first node). The node's sample count (from the training set), which has been separated into 50 healthy bearing and 49 faulty bearing samples for the first node and a majority HB group for the second node, is the node's fourth and final piece of information. The sample tree only had 99 samples, despite the fact that we used 132 HB and FB samples in our research. The remainder is used as test cases to verify the tree's correctness[31].

The split samples must then be evaluated by the second layer leaf node. The values x3 and x1 serve as the classification criteria in second-layer leaf nodes. It will then be processed into the third layer of leaf nodes until there are no more leaf nodes. The healthy bearing probability is calculated when the decision reaches the final leaf node by dividing the total number of samples in the node by the number of class samples. For instance, the likelihood of HB and FB at the leftmost leaf node is 1/1& 0/1 respectively[32,37].

8.2.4 Accuracy Comparison

Measurement of accuracy can be accessed through cross-validation to see how various test methods perform[37].

Hence, we compared the accuracies of the ML models as shown in Fig. 27, in which RF has the highest cross-validation accuracy of 93.93 % with an accuracy difference of the 6.05 % compared to SVM and k-NN, whereas k-NN and SVM have the same validation accuracy of 87.88%.

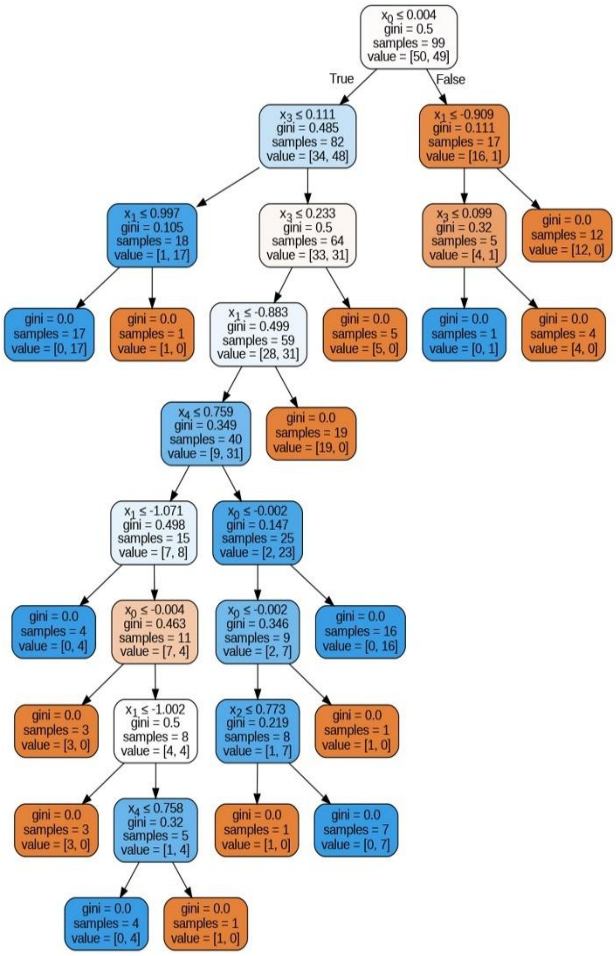

Fig. 26: Decision tree

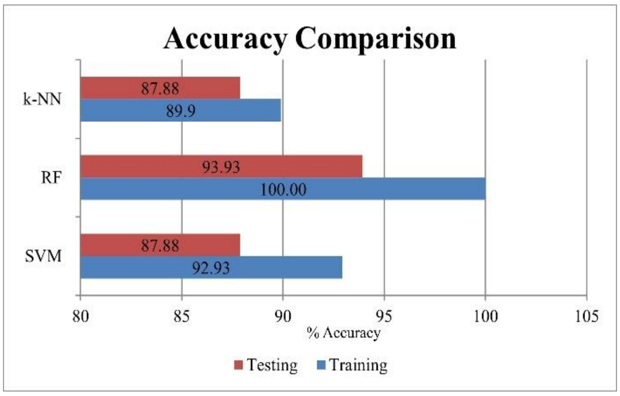

Fig. 27: Accuracy comparison

8.2.5 10-Fold Cross Validation (CV) Accuracy

It may not be accurate to evaluate a model's dependability only on how well it works during testing and training. This is because a model's performance depends on the use of accurate data during both the training and testing stages of model's lifecycle. These data might not accurately represent the dataset or might not be broadly applicable to newly discovered, unidentified data. Predictive studies frequently employ the well-known 10-fold cross-validation approach to get across this restriction. It helps in attaining unbiased and reliable prediction results[54].

Ten equivalent subsets, or "folds" are generated from the dataset using 10-fold cross-validation. Each of the ten model iterations that are trained and evaluated uses a different fold (1 in number) as the testing set, while the remaining folds (9 in numbers) serve as the training set. It evaluates the accuracies of all the iterations using the different folds for all the classifiers[54].

Figure 28 shows the different accuracies obtained for the 10-fold CV iterations and the mean accuracy line for all the classifiers. Figure 28 shows the variation of the accuracies for the different folds for the classifiers and the mean lines. As shown in Fig. 29, the mean accuracies for k-NN, RF and SVM are 85.55, 83.24 and 72.53 % respectively.

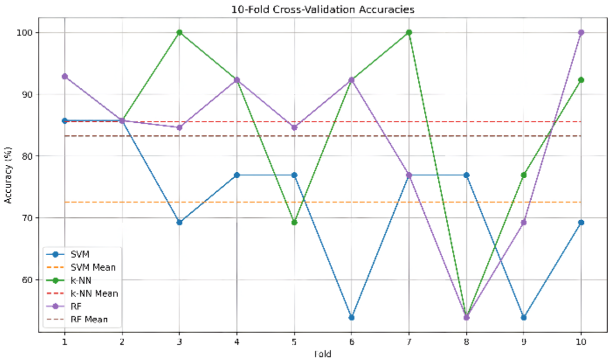

Fig. 28: Accuracy comparison of machine learning models.

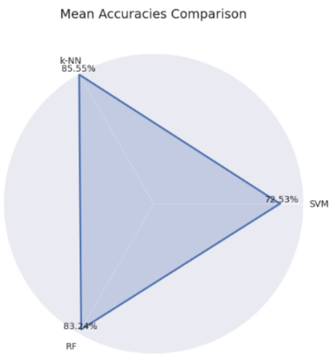

Fig. 29: Mean accuracies comparison

8.2.6 Parameters Comparison

Figure 30 shows the comparison of the different parameters i.e. F1 score, recall value and precision value[37]. The definitions of the parameters are as follows:

(1) How many of the made positive predictions (true positives) are accurate is a measure of precision.

(2) Recall also known as sensitivity is a metric for how many out of all the positive cases in the data that the classifier correctly predicted.

(3) The F1-Score is a measurement that combines recall and precision. Generally speaking, it is referred to as the harmonic mean of the two.

The above parameters are calculated using the standard formulae as shown in Fig. 19 and compared. The comparison shows that the RF has higher parameter values compared to the other methods, it outperformed compared to the others.

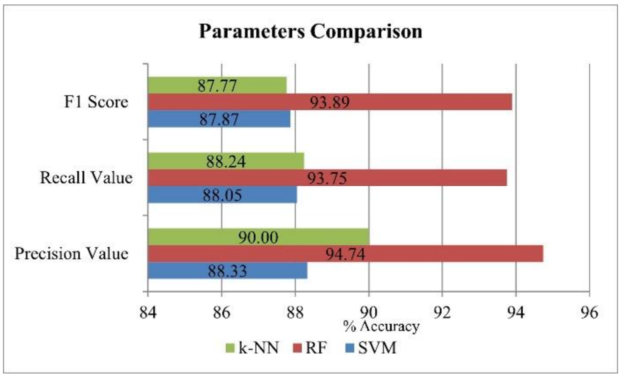

Fig. 30: Parameters comparison

8.2.7 ROC Curves Comparison

An evaluation metric for binary classification problems is the Receiver Operator Characteristic (ROC) curve. In essence, it separates the "signal" from the "noise" by plotting the TPR (True Positive Rate) against the FPR (False Positive Rate) at various threshold values. In other words, it displays how well a categorization model performs across the board.

The Area under the Curve (AUC), which provides a summary of the ROC curve, is a measurement of a binary classifier's capability to distinguish between classes. The higher the AUC, the better the model performs at differentiating between the positive and negative classes.

Figure 31 shows the comparison of the best fit curves of the three methods, it can be inferred that the AUC of RF (0.98) is more than the other methods SVM and k-NN. It excelled at differentiating between positive and negative class values and was able to identify more true positives and true negatives than false positives and false negatives.

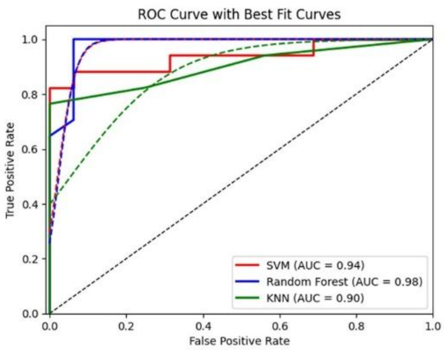

Fig. 31: ROC curves comparison.

9. Conclusion

The condition of the HCJB in the machineries is of utmost importance in tribology research and plant maintenance. Hence the ML model for condition monitoring of the HCJB is developed which will observe the vibration signals continuously, predict the fault, and shut-off the system when reaching critical values. This paper advances the state of art in tribology by ML-based fault prediction especially in the evolving field of conical journal bearings. The analysis compares healthy and scratched HCJBs based on acceleration amplitudes spectrum and ML-based fault prediction using vibration signals. Here are the findings of the work:

(1) Higher acceleration amplitudes are observed in case of the internally scratched bearing (SB/FB) during constant load and variable speed conditions along the longitudinal direction. The observed peak amplitude is 3.29 m/s^2 when faulty bearing running at 1500 RPM with 350N radial load (Fig. 12), thus giving us the pre-set alarming value in the operating region with wear for HCJB, which is useful for condition monitoring of the bearing.

(2) Peak frequency occurs at 2nd multiple of operating frequency (2x OF) as shown in the FFT spectrum in case of HB. Peak frequencies have consequently occurred at consequent multiples showing a definite pattern in case of FB.

(3) Peak frequency occurs at 4th multiple (4x OF) in case of FB and shifts from 49.95 Hz to 99.99 Hz due to scratches present on the bearing surfaces as shown in Fig. 13. The observation depicts the unique characteristic of the HCJB when wear is present on the internal surfaces.

(4) Mean, standard deviation, kurtosis, skewness and RMS are the most influential features in the classification process out of which mean, kurtosis, and skewness have greater variations in case of FB compared to HB.

(5) From the performance assessment of the different classifiers, all three methods have accuracies with a difference of 6.05% and have an overall average validation accuracy of 88.89%. RF has the highest validation accuracy of 93.93% compared to SVM and k-NN with higher evaluation metrics values.

(6) The 10-fold CV results further empower the classification accuracy and reliability of the RF and k-NN algorithms over SVM.

The findings of the work will assist in the predictive maintenance of the HCJBs, which can be operational real time in the condition monitoring tool. Impeding catastrophic failure of the HCJB due to abrasive wear propagation will be predicted before it occurs thus avoiding the breakdown of the machineries consequently eliminating production losses. The workflow can be applied to wider range of use-cases especially the journal bearings with other types of faults. Hence the presented methodology can be applied to the condition monitoring of the hydrodynamic conical journal bearing when used in machineries (centrifugal pumps and turbines, etc.).

Acknowledgements

The financial support from the All India Council for Technical Education (AICTE), New Delhi, India, under Research Promotion Scheme (RPS) Ref. No. 8-221/RIFD/RPS/Policy-1/2014- 15 is gratefully acknowledged. The author is grateful to the Director, Veermata Jijabai Technological Institute (V.J.T.I), Mumbai.

Nomenclature

List of Symbols

Symbols	Definition
λ	Regularization parameter in linear regression
C	Inverse of the strength of regularization in SVM Model
γ	RBF Kernel parameter to determine the smoothness or complexity of the decision boundary
k	Number of nearest neighbors to a data point in the k-NN algorithm

References

1) Tiwari, P., and Kumar, V. “Analysis of hydrodynamic journal bearing: A review.” *International Journal of Engineering Research & Technology*, 1-7 (2012). https://doi.org/10.17577/IJERTV1IS7470
2) Childs, P. R. “Mechanical design engineering handbook.” Butterworth-Heinemann (2013).
3) Gangrade, A. K., Phalle, V. M., and Mantha, S. S. (2019). “Performance analysis of a conical hydrodynamic journal bearing. Iranian Journal of Science and Technology,” *Transactions of Mechanical Engineering,* 43(3), 559-573 (2019). https://doi.org/10.1007/s40997-018-0217-2
4) Singh, H., and Sen, S. S. “Wear in hydrodynamic journal bearings: a review.” *International Journal Of Modern Engineering Research (IJMER),* 4, 45-57 (2014). https://api.semanticscholar.org/CorpusID:212464813
5) S. M., Muzakkir and P., Lijesh and Hirani, Harish. “Failure mode and effect analysis of journal bearing.” *International Journal of Applied Engineering Research*. 10. 37752-37759 (2015). https://doi.org/10.1016/j.engfailanal.2015.11.033
6) Harnoy, A. “Bearing design in machinery: engineering tribology and lubrication.” *CRC press* (2002). https://doi.org/10.1201/9780203909072
7) Rozeanu, L., and Kennedy, F. E. “Wear hydrodynamic journal bearings.” *In Tribology Series 39 161-166 Elsevier* (2001). https://doi.org/10.1016/S0167-8922(01)80103-6
8) Scott, R. “Basic wear modes in lubricated systems.” *Mach. Lubr*, 7 (2008).
9) Poley, J. A. C. K., and Murphy, M. I. C. H. E. L. “The history of oil analysis”. *Machinery Lubrication* (2007).
10) Kumar, S., Lokesha, M., Kumar, K., and Srinivas, K. “Vibration based Fault Diagnosis Techniques for Rotating Mechanical Components: Review

Paper." *IOP Conference Series: Materials Science and Engineering, 376* (2018). https://doi.org/10.1088/1757-899X/376/1/012109.

11) Ranjan R, Ghosh SK, Kumar M. "Fault diagnosis of journal bearing in a hydropower plant using wear debris, vibration and temperature analysis: A case study." *Proceedings of the Institution of Mechanical Engineers, Part E: Journal of Process Mechanical Engineering* 234 (3) 235-242 (2020). https://doi.org/10.1177/0954408920910290
12) Kankar, P. K., Sharma, S. C., Harsha, S. P. "Fault diagnosis of ball bearings using machine learning methods." *Expert Systems with Applications*, *38(3),* 1876-1886 (2011). https://doi.org/10.1016/j.eswa.2010.07.119
13) Roylance, B.J. "Machine failure and its avoidance—what is tribology's contribution to the effective maintenance of critical machinery?" *Proceedings of the Institution of Mechanical Engineers, Part J: Journal of Engineering Tribology,* 217 349-364 (2003).
14) Cerrada, M., Sánchez, R. V., Li, C., Pacheco, F., Cabrera, D., de Oliveira, J. V., and Vásquez, R. E. "A review on data-driven fault severity assessment in rolling bearings." *Mechanical Systems and Signal Processing, 99,* 169-196 (2018). https://doi.org/10.1016/j.ymssp.2017.06.012
15) El-Thalji, I., and Jantunen, E. "A summary of fault modeling and predictive health monitoring of rolling element bearings." *Mechanical systems and signal processing, 60*, 252-272 (2015). https://doi.org/10.1016/j.ymssp.2015.02.008.
16) Bellini, A., Filippetti, F., Tassoni, C., and Capolino, G. A. "Advances in diagnostic techniques for induction machines." *IEEE Transactions on industrial electronics,55(12),* 4109-4126 (2008). http://dx.doi.org/10.1109/TIE.2008.2007527.
17) Dai, X., and Gao, Z. "Fromthemodel,signal to knowledge:A data-driven perspective of fault detection and diagnosis." *IEEE Transactions on Industrial Informatics,9(4)* 2226-2238 (2013). http://dx.doi.org/10.1109/TII.2013.2243743.
18) Henriquez, P., Alonso, J. B., Ferrer, M. A., and Travieso, C. M. "Review of automatic fault diagnosis systems using audio and vibration signals." *IEEE Transactions on systems, man, and cybernetics: Systems, 44(5)* 642-652 (2013). http://dx.doi.org/10.1109/TSMCC.2013.2257752
19) Chen, W., and Yuan, H. M. "An improved GA-SVM algorithm." *9th IEEE Conference on Industrial Electronics and Applications, IEEE 2137-2141* (2014).
20) Samanta, B. I. S. W. A. J. I. T., Al-Balushi, K. R., and Al-Araimi, S. A. "Artificial neural networks and support vector machines with genetic algorithm for bearing fault detection." *Engineering applications of artificial intelligence, 16(7-8)* 657-665 (2003). http://dx.doi.org/10.1016/j.engappai.2003.09.006.
21) Vapnik, V. N. "The nature of statistical learning. Theory." (1995).
22) Gunn, S. R. "Support vector machines for classification and regression." *ISIS technical report, 14(1) 5-16* (1998). http://dx.doi.org/10.1039/b918972f.
23) Wang, Z., Zhang, Q., Xiong, J., Xiao, M., Sun, G., and He, J. "Fault Diagnosis of a Rolling Bearing Using Wavelet Packet Denoising and Random Forests." *IEEE Sensors Journal, 17(17)* 5581–5588 (2017). http://dx.doi.org/10.1109/JSEN.2017.2726011.
24) Jamil, M. A., Khan, M. A. A., and Khanam, S. "Erratum: Feature-based performance of SVM and KNN classifiers for diagnosis of rolling element bearing faults." *Vibroengineering PROCEDIA, 40* (2021). http://dx.doi.org/10.21595/vp.2021.22307.
25) Wang, Q., Liu, Y. B., He, X., Liu, S. Y., and Liu, J. H. "Fault Diagnosis of Bearing Based on KPCA and KNN Method." *Advanced Materials Research, 986–987* 1491–1496 (2014). http://dx.doi.org/10.4028/www.scientific.net/AMR.986-987.1491.
26) Sharma A, Sugumaran V, Babu Devasenapati S "Misfire detection in an IC engine using vibration signal and decision tree algorithms." *Meas J IntMeas Confed 50 370–380* (2014). https://doi.org/10.1016/j.measurement.2014.01.018.
27) Anoop PS, Sugumaran V "Classifying machine learning features extracted from vibration signal with logistic model tree to monitor automobile tyre pressure." *SDHM StructDurab Heal Monit 11 191–208* (2017). http://dx.doi.org/10.3970/sdhm.2017.011.191.
28) Joshuva A, Sugumaran V "Comparative study on tree classifiers for application to condition monitoring of wind turbine blade through histogram features using vibration signals: a data-mining approach." *SDHM StructDurab Heal Monit 13 399–416* (2019). http://dx.doi.org/10.32604/sdhm.2019.03014.
29) AsadiMajd, A., Samet, H., and Ghanbari, T. "k-NN based fault detection and classification methods for power transmission systems." *Protection and Control of Modern Power Systems, 2(1)* (2017). http://dx.doi.org/10.1186/s41601-017-0063-z.
30) Saravanja, Davorka and Grbesic, Marko. "Application of Vibration Analysis in Journal Bearing Problems Diagnostics," *Proceedings of the 30th DAAAM International Symposium, pp.0100- 0106, B. Katalinic (Ed.), Published by DAAAM International, ISBN 978-3-902734-22-8, ISSN 1726-9679,* Vienna, Austria (2019). https://doi.org/10.2507/30th.daaam.proceedings.013.
31) Kapusta, J., and Obonya, J. "Improvement of Misleading and Fake News Classification for Flective Languages by Morphological Group Analysis."

Informatics, 7(1) 4 (2020). http://dx.doi.org/10.3390/informatics7010004.

32) Feng, Y., Wang, Z., Cui, R., Xiao, M., Gao, H., Bai, H., Delvoux, B., Zhang, Z., Dekker, A., Romano, A., Wang, S., Traverso, A., Liu, C., and Zhang, Z. "Clinical analysis and artificial intelligence survival prediction of serous ovarian cancer based on preoperative circulating leukocytes." *Journal of Ovarian Research, 15(1)* (2022). https://doi.org/dx.doi.org/10.1186/s13048-022-00994-2.
33) Rayjade G, Bhagure A, Kushare PB, Bhandare R, Matsagar V, Chaudhari A. "Performance evaluation of machine learning algorithms and impact of activation functions in artificial neural network classifier for bearing fault diagnosis." *Journal of Vibration and Control. 0(0)* (2024). https://doi.org/10.1177/10775463241235778.
34) Surojit Poddar, N. Tandon, "Detection of particle contamination in journal bearing using acoustic emission and vibration monitoring techniques", *Tribology International, Volume 134, 154-164* (2019) *ISSN 0301-679X.* https://doi.org/10.1016/j.triboint.2019.01.050.
35) Lin H-C, Ye Y-C, Huang B-J, Su J-L. "Bearing vibration detection and analysis using enhanced fast Fourier transform algorithm." *Advances in Mechanical Engineering. 8(10)* (2016). https://doi.org/10.1177/1687814016675080.
36) Pinheiro, A.A., Brandao, I.M., and da Costa, C. (2019). "Vibration Analysis in Turbomachines Using Machine Learning Techniques." *European Journal of Engineering Research and Science* (2019). https://api.semanticscholar.org/CorpusID:86719785
37) KaliappanSivakumar, K., Ganesan, G., Chermakani, G.S., and Muthukumar, D. "Identification of bearing fault in induction motor using random forest algorithm." *E3S Web of Conferences* (2023). https://doi.org/10.1051/e3sconf/202338701005.
38) Wang, D.; Guo, W.; Wang, X. A joint sparse wavelet coefficient extraction and adaptive noise reduction method in recovery of weak bearing fault features from a multi-component signal mixture. *Appl. Soft* Comput. 13, 4097–4104 (2013). https://doi.org/10.1016/j.asoc.2013.05.015
39) Dave, V., Borade, H., Agrawal, H., Purohit, A., Padia, N., & Vakharia, V. Deep Learning-Enhanced Small-Sample Bearing Fault Analysis Using Q-Transform and HOG Image Features in a GRU-XAI Framework. *Machines*, *12*(6), 373 (2024). https://doi.org/10.3390/machines12060373.
40) Mozaffar, M., Liao, S., Xie, X., Saha, S., Park, C., Cao, J., Liu, W., & Gan, Z. (2021). Mechanistic artificial intelligence (Mechanistic-AI) for modeling, design, and control of advanced manufacturing processes: Current state and perspectives. *Journal of Materials Processing Technology.* 302 117485 (2022). https://doi.org/10.1016/j.jmatprotec.2021.117485
41) Tremmel S, Marian M. Machine learning in tribology—More than buzzwords? Lubricants 10(4) 68 (2022). https://doi.org/10.3390/lubricants10040068
42) Voukantsis D, Karatzas K, Mihailidis A, Gatsios S, Sahanas C, Bakolas V, Hoffinger C. Application of computational intelligence to the analysis of friction measurements. TribolTrans 53(5) 748–754 (2010). https://doi.org/10.1080/10402001003760221
43) Marian M, Tremmel S. Current trends and applications of machine learning in tribology—A review. Lubricants 9(9) 86 (2021). https://doi.org/10.3390/lubricants9090086
44) Nian Yin, Pufan Yang, Songkai Liu, Shuaihang Pan, Zhinan Zhang. AI in tribology: Present and future. Volume 12 1060-1097 (2024). https://doi.org/10.1007/s40544-024-0879-2
45) Wu D Z, Jennings C, Terpenny J, Gao R X, Kumara S. A comparative study on machine learning algorithms for smart manufacturing: Tool wear prediction using random forests. J Manuf Sci Eng 139(7) 071018 (2017). https://doi.org/10.1115/1.4036350.
46) Wierzcholski K. Fuzzy logic tools in intelligent micro- bearing systems. J Appl Math Comput 17(2): 123–131 (2009). https://it.p.lodz.pl/file.php/12/2009-2/JACS-2009-2-Wierzcholski.pdf.
47) Walker J, Questa H, Raman A, Ahmed M, Mohammadpour M, Bewsher S R, Offner G. Application of tribological artificial neural networks in machine elements. Tribol Lett 71(1) 3 (2022) https://doi.org/10.1007/s11249-022-01673-5.
48) Shevchik S, Saeidi F, Meylan B, Wasmer K. Prediction of failure in lubricated surfaces using acoustic time-frequency features and random forest algorithm. IEEE Trans Ind Inform 13(4):1541–53 (2017). https://doi.org/10.1109/TII.2016.2635082.
49) Bienefeld C, Kirchner E, Vogt A, Kacmar M. On the importance of temporal information for remaining useful life prediction of rolling bearings using a random forest regressor. Lubricants 10 67 (2022) https://doi.org/10.3390/lubricants10040067.
50) Pandiyan V, Caesarendra W, Thajhowidodo T, Tan H. In-process tool condition monitoring in compliant abrasive belt grinding process using support vector machine and genetic algorithm. J Manuf Process 31 199–213 (2018). https://doi.org/10.1016/j.jmapro.2017.11.014.
51) J. Prost, G. Boidi, A.M. Puhwein, M. Varga, G. Vorlaufer, Classification of operational states in porous journal bearings using a semi-supervised multi-sensor Machine Learning approach, Tribology International,Volume 184,2023,108464, ISSN 0301-679X (2023) https://doi.org/10.1016/j.triboint 108464
52) Turan, Tülay. Effect of Hyperparameter Settings on

Classification Performance of SVM and KNN Algorithms in Breast Cancer Diagnosis. *Bilsel International Korykos Scientific Researches And Innovation Congress*, 07-08 (2024).

53) Almeida, F & Coletta, Luiz & Campos, M & Gonçalves, A & Silva, E. Tribology and Vibration Analysis Enhanced by a Machine Learning Technique: A Short Review (2017). https://doi.org/10.13140/RG.2.2.10187.69923.

54) Shah M, Borade H, Dave V, Agrawal H, Nair P, Vakharia V. Utilizing TGAN and ConSinGAN for Improved Tool Wear Prediction: A Comparative Study with ED-LSTM, GRU, and CNN Models. *Electronics*. 13(17) 3484 (2024). https://doi.org/10.3390/electronics13173484.