Vibration-Based Condition Monitoring of Shaft Bearing Systems Using Machine Learning Techniques

Kishan Kumar[a], Prof. Randhvan Bhagwat M[b], Prof. Dengale Pravin B[c]
[a] PG Student, [bc] Assistant Professor, Department of Mechanical Engineering, Sahyadri Valley College of Engineering & Technology, Rajuri, Pune, Maharashtra, India-412411

***Abstract* - A shaft-bearing system is an essential part of rotating machinery. To guarantee that a shaft bearing system operates safely and reliably, the bearings' condition must be monitored on a regular basis. Bearing and shaft failures are thought to be the leading reasons of failure in various revolving machines used in the industry at highre and lower speeds. The condition of the bearing changes throughout use, so do the vibrations, and their characteristics vary depending on the reason. As a result, the bearing's unique property makes it suited for vibration monitoring and other procedures. The vibration measurement approach may reliably anticipate the upcoming failure and life of a mechanism or component based on changes in vibration signals.**

As a result, the bearing's unique property makes it suited for vibration monitoring and other procedures. The vibration measurement approach may reliably anticipate the future failure and life of a machine or component based on changes in vibration signals. As a result, the goal is to extend the machine's life by detecting faults early on, allowing for an effective maintenance program to be implemented to remedy the problem. Subsequently, this research uses machine learning methods to detect bearing problems, compare them to various faulty and standard models, and categorize the bearing type. In this research work, we use outer race fault data from the Bearing data set to extract the time domain features from the dataset using Various machine learning models, including Principal Component Analysis, K-NEAREST NEIGHBOURS (K-NN), SUPPORT VECTOR MACHINES (SVM), RANDOM FOREST CLASSIFIER, DESICION TREE, and LOGISTIC REGRESSION. As a consequence, we obtain the best model that performs optimally on the data set. Finally, the proposed methods of condition monitoring will be implemented in a real-world case study of the shaft bearing system. Thus, vibration testing is used to monitor the state of the shaft bearing system, allowing for the identification of problematic bearings and improved performance after they are replaced.

Keywords: Bearing fault diagnosis using machine learning technique , Bearing condition monitoring

I. INTRODUCTION

Previously, maintenance was exclusively referred to as breakdown maintenance (also known as run-to-failure), which occurs after a machine or component fails. This type of maintenance requires little planning. Then, preventative or periodic maintenance was formed as the following maintenance plan. Regardless of the physical asset's condition, this maintenance plan entails doing maintenance chores on a regular basis. Finally, the condition-based maintenance technique (also known as predictive maintenance) was developed. This maintenance technique uses data acquired from several condition monitoring systems to guide maintenance operations. Condition-based maintenance reduces unnecessary maintenance effort by performing maintenance only when there is evidence of aberrant behavior of the machine component or device. If the condition-based maintenance program is properly established and implemented, it is possible to minimize maintenance costs even further by avoiding unnecessary preventive maintenance procedures. Rotating machine element vibration analysis is one of the most commonly used devices in condition monitoring activities. Analyzing and measuring the extent of vibration in rotary machines allows you to find various issues such as misalignment, looseness, bent shaft, unbalance, cracked shaft, motor fault, gear fault, rubbing, and impellor or blade defects. However, the defects listed above are fairly common in high-speed machinery. Ball bearings are made up of four components: an inner race, an outer race, balls, and a cage. The health of a machine is heavily reliant on the strength and dependability of its bearings.

[1] The majority of mechanical failures are mostly caused by bearing faults. Different bearing flaws cause vibration, noise, low efficiency, and ultimately equipment failure. The typical method of bearing care involves periodic replacement, which may result in a 90% reduction in bearing life. Determining the right operational condition of the bearing and carrying out scheduled repairs for fault identification and usable operating conditions are therefore necessary in an effective maintenance method. [2]. An automated condition monitoring system for several machines was created through modifications to various types of sensors, data-gathering devices, software, and computers. Randall and Endo address the needs for various gear and bearing types for power transfer and friction reduction in transportation and rotating machinery applications. Fault detection is the process of observing system data (measured) and information about the system's state and comparing it to some standard qualities in order to discover any abnormalities in the system's health. Unfortunately, one technique alone cannot detect all defects in machines. It is indicated that only the vibration measuring approach is extensively used in the industry for condition monitoring of various machinery, and it can detect 90% of faults or failures in the machines by the changes in vibration signals [3]. The signal level reliably predicts the machine's or component's future failure and life. As a result, the goal is to extend the machine's life by detecting faults early on, allowing for an effective maintenance program to be implemented to remedy the problem.

A. CONDITION MONITORING PRINCIPLE

It has been reported that visible signs precede 99% of mechanical failures. Condition monitoring is a technique for monitoring the equipment's indications in order to provide a warning of failure. The main premise is to choose physical parameters that show that deterioration is occurring, and then take readings at predetermined intervals. Any upward trend can thus be spotted and interpreted as an indication that a problem exists, as demonstrated by a typical trend curve and how this offers a warning that an impending failure is approaching. It also provides ample time for planning and implementation. Because failure happens in distinct components, the monitoring measurement must focus on the specific failure modes of the critical components.

The procedure for Condition monitoring consists of three steps:

- Data collection step: Collect relevant system health data.
- Data processing step: Analyse data acquired in step 1 to improve understanding and interpretation.
- Recommend efficient maintenance policies during the decision-making process.

Condition monitoring data serves a variety of applications. It could be vibration data, auditory data, oil analysis data, or temperature data. Many sensors, such as microsensors, ultrasonic sensors, and acoustic emission sensors, have been designed to collect various types of data. Bluetooth and other radio technologies now offer a low-cost alternative for data exchange. To store and manage data, maintenance information systems have been developed, including computerized maintenance management systems (CMMS) and enterprise resource planning systems. As a result, with the rapid advancement of computer and advanced sensor technologies, data-collecting facilities and technology have become more accessible and practical for CBM applications.

➢ Data Acquisition: The equipment that is used for performing the experiments (Monitor the condition of the shaft bearing system.) is as follows:

1. Single-axis accelerometer or sensor (353B33), Sensitivity- (±5%) 100 mV/g,

2. FFT analyser (OROS-3 series/NV gate, 4 channel types, Hardware-OR 34 system.)

3. Noise Level Meter (Microphone)

4. SKF Machine Condition Advisor (CMAS100-SL)

5. Laptop/computer with Vibration measurement software (NV Gate) for recording the readings,

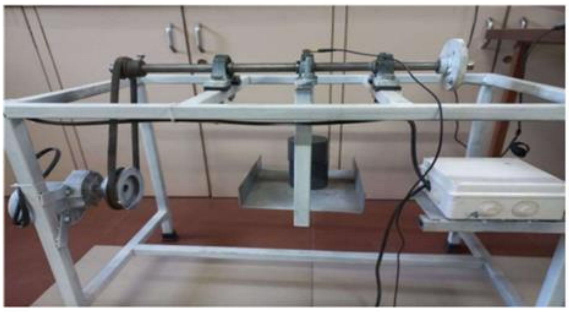

Fig. 1. Experimental setup of fault diagnosis of bearings

B. Data Structure

The data set is contained in the data container. The dataset describes a test-to-fail experiment. Each file in the data collection contains a snapshot of a 1-second vibration signal captured at predefined intervals. Separately file has a 20-kHz sampling rate and a total point count of 20,480. A data point is a record (or row) in a data file. The NI DAQ Card 6062E made data collecting easier. The experiment was continued the next working day, as evidenced by extended time stamp intervals (visible in file names). Set: Three channels are available. Bearing 1 is in Ch1, Bearing 2 is in Ch2, Bearing 3 is in Ch3, and Bearing 4 is in Ch4. Bearing 3 experienced the inside

Bearing 3 experienced inner race disappointment and bearing four experienced after the test-to-failure experiment.

C. Data Processing

This first phase will include data preparation and feature extraction from the collected data to produce meaningful features. In this situation, we retrieved time domain features from the available acceleration data. Depending on the classification requirements, features might be classified as related, redundant, or irrelevant. The purpose of feature selection is to remove as many redundant and irrelevant characteristics as possible while keeping related features, reducing the feature vector dimension, and avoiding dimensional disasters and overfitting. The primary approach of feature selection consists of four steps: creating feature subsets, evaluating feature subsets, halting criteria, and verification outcomes.

- ➢ Maximum and minimum values.
- ➢ Mean value.
- ➢ Standard deviation (unbiased std).
- ➢ Root mean square value
- ➢ Skewness
- ➢ Kurtosis
- ➢ Crest factor = Max/RMS
- ➢ Form factor = RMS/Mean.

D. Bearing Dataset Visualization

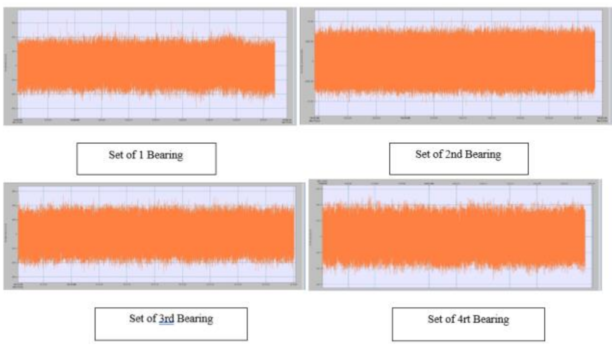

Fig. 2. Acceleration vs Number of points present in one Second

The graph you're referring to represents the relationship between acceleration and the number of points recorded per second. It seems to be a graphical representation of experimental data.

Acceleration refers to how quickly an object's velocity varies over time. It is often represented in units like meters per second squared (m/s²).The acceleration can be either positive (showing a rise in velocity) or negative (representing a decrease in velocity). The number of points per second, in this context, likely refers to the frequency at which data points are recorded or sampled. For instance, in a scientific experiment or data collection process, measurements of acceleration taken at regular intervals, and the number of data points obtained per second determines the sampling frequency.

The graph illustrates how acceleration values vary with different sampling frequencies. Each point on the graph represents a combination of acceleration and the corresponding number of points recorded per second. The x-axis represents the number of points per second, and the y-axis represents the acceleration.

By examining the graph, you can analyse the relationship between acceleration and the sampling frequency. You might observe patterns or trends, such as whether higher sampling frequencies lead to more accurate or precise acceleration measurements. Additionally, you could investigate whether there are any limits or thresholds in the data, where increasing the sampling frequency no longer yields noticeable changes in the acceleration values.

II. IDENTIFICATION AND DIAGNOSIS OF FAULT

To automate the detection and diagnosis of defects in rolling element bearings using short data sets. Used the Artificial Neural Networks (ANNs) and SVM.

Test-1

The test-to-failure experiment ended with an inner race flaw in bearing 3 and a roller element defect in bearing 4.

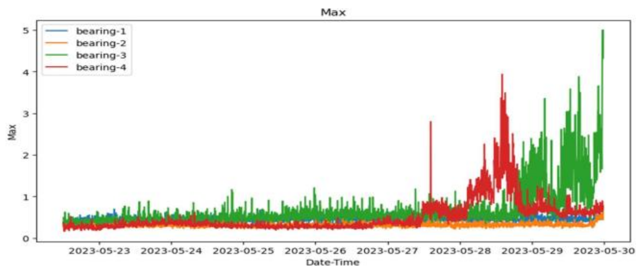

Fig. 3. the maximum acceleration recorded over time for four different types of bearings

The above graph referring to displays the maximum acceleration recorded over time for four different types of bearings, one of which is faulty and shows a maximum peak. This graph likely represents a comparison of the acceleration performance of different bearings in a system.

➢ **Maximum Acceleration:** The y-axis represents the maximum acceleration recorded during each time interval. It is commonly measured in figures like meters per second squared (m/s²).

Three Types of Bearings: The graph compares the performance of four different types of bearings. Each type of bearing is represented by a different line or data series on the graph. These lines show how the maximum acceleration changes over time for each bearing type.

Faulty Bearing: Two bearings are indicated as faulty and exhibit a maximum peak. This peak likely represents an anomalous or unexpected spike in acceleration for the faulty bearing compared to the other types. This suggests that the faulty bearing is experiencing some issue that is causing unusually high acceleration values.

By examining the graph, you can compare the performance of the different bearings over time and identify any significant differences, particularly the maximum acceleration values. You can observe how the acceleration varies for each bearing type and determine whether the faulty bearing consistently exhibits higher acceleration peaks compared to the other bearings. The purpose of this analysis might be to identify the faulty bearing by its distinct acceleration behavior and investigate the underlying cause of the issue. It could help in troubleshooting or making informed decisions about the maintenance or replacement of the faulty component.

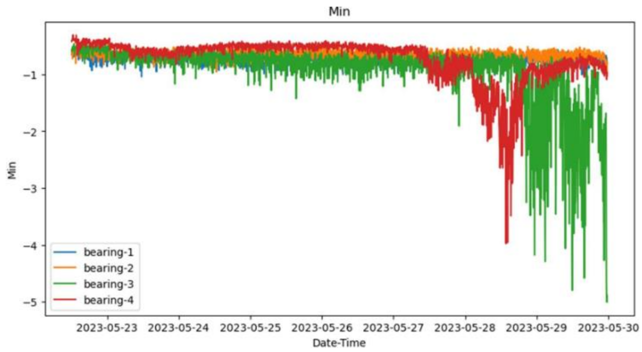

Fig. 4. the minimum acceleration recorded over time for four different types of bearings

➢ **Minimum Acceleration:** The y-axis represents the minimum acceleration recorded during each time interval. It is commonly measured in figures like meters per second squared (m/s²).

Faulty Bearing: One of the bearing types is indicated as faulty and exhibits a maximum peak. This peak likely represents an anomalous or unexpected spike in minimum acceleration for the faulty bearing compared to the other types. This suggests that the faulty bearing is experiencing some issue that is causing unusually low acceleration values.

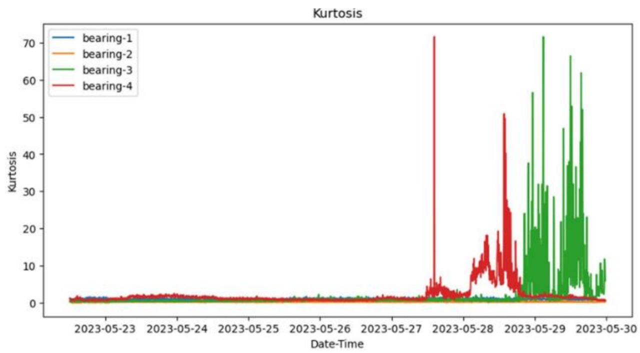

Fig. 5. Kurtosis measures acceleration recorded over time for four different types of bearings

Kurtosis indicates if data is heavily left- or right-tailed. Data sets with high kurtosis have larger tails and more outliers, whereas data sets with low kurtosis have smaller tails and fewer outliers.

The purpose of this analysis might be to detect and diagnose the faulty bearing based on its distinctive kurtosis profile. It could help in identifying potential issues such as irregular vibrations, unusual wear patterns, or other abnormalities that could be causing the observed extreme values.

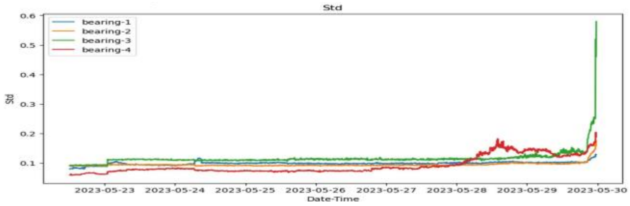

Fig. 6. relationship between the standard deviation (std) parameter and time.

The graph you're describing represents the relationship between the standard deviation (std) parameter and time for four different types of bearings, one of which is faulty and showing a maximum peak.

By analyzing the graph, you can compare the standard deviation values for each bearing type over time. You can observe how the variability changes and identify any distinct patterns or differences, particularly the maximum peak associated with the faulty bearing. This peak indicates that the faulty bearing's performance is less consistent or predictable, exhibiting more fluctuations or variations.

The purpose of this analysis might be to detect and diagnose the faulty bearing based on its distinctive standard deviation profile. The maximum peak suggests an abnormal level of variability that could be caused by factors such as looseness, misalignment, or other mechanical issues affecting the bearing's stability.

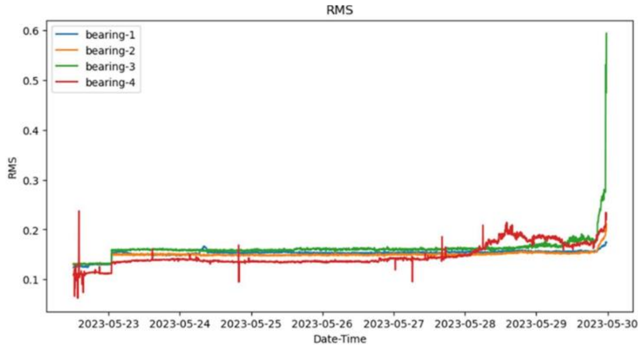

Fig. 7. relationship between the Root Mean Square (RMS) parameter and time.

Graph represents the relationship between the RMS Vs time for four different types of bearings, with one of them being faulty and showing a maximum peak.

RMS is an arithmetic measure that quantifies the average size or amplitude of a indication or dataset. In this context, it likely represents a characteristic related to the vibration or oscillation of the bearings.

By examining the graph, you can compare the RMS values for each bearing type over time. You can observe how the vibration or oscillation characteristics vary and identify any distinct patterns or differences, particularly the maximum peak associated with the faulty bearing. This peak indicates that the faulty bearing has the highest amplitude of vibration, suggesting a potential issue or malfunction.

TABLE I. SUMMARY OF FEATURE EXTRACTION:

Date & Time	Max	Min	Mean	Std	RMS	Skewness	Kurtosis	Crest Factor	Form Factor
04-04-2023 09:27	0.361	-0.31	-0.00515	0.066192	0.06639	0.083449	0.591902	5.437565	-12.8964
04-04-2023 09:32	0.295	-0.295	-0.00631	0.065986	0.066285	0.011828	0.359751	4.450449	-10.5056
04-04-2023 09:42	0.286	-0.286	-0.00129	0.069176	0.069186	0.049056	0.213724	4.133782	-53.5195
04-04-2023 09:52	0.349	-0.31	-0.00228	0.0686	0.068637	0.002209	0.299035	5.084751	-30.0995
04-04-2023 10:02	0.388	-0.374	-0.00173	0.067449	0.067469	0.031884	0.395754	5.750781	-39.0583
04-04-2023 10:12	0.383	-0.344	-0.00265	0.069782	0.06983	0.044135	0.423808	5.484716	-26.3686
04-04-2023 10:22	0.605	-0.381	-0.00177	0.069478	0.069499	0.010153	0.666443	8.705145	-39.296
04-04-2023 10:32	0.359	-0.354	-0.00189	0.067005	0.06703	0.073132	0.430821	5.355775	-35.4239
04-04-2023 10:42	0.354	-0.3	-0.00166	0.070047	0.070065	-0.03024	0.338647	5.052455	-42.2076
04-04-2023 10:52	0.359	-0.4	-0.0023	0.069329	0.069366	0.025425	0.330919	5.175468	-30.2007
04-04-2023 11:02	0.32	-0.312	-0.0022	0.068443	0.068476	0.04173	0.263364	4.673156	-31.1795
04-04-2023 11:12	0.386	-0.342	-0.00208	0.069379	0.069409	0.055657	0.344013	5.561277	-33.3174
04-04-2023 11:22	0.349	-0.361	-0.00191	0.069574	0.069598	0.003434	0.346268	5.0145	-36.3643
04-04-2023 11:32	0.398	-0.337	-0.00218	0.068573	0.068606	0.002784	0.466325	5.801211	-31.4493
04-04-2023 11:42	0.344	-0.352	-0.00173	0.069439	0.069459	0.013238	0.401559	4.952577	-40.242

The analysis of the graph aims to detect and diagnose the faulty bearing based on its distinctive RMS profile. The maximum peak represents an abnormal level of vibration, which could be caused by factors such as misalignment, imbalance, or other mechanical problems affecting the bearing's performance.

III. CLASSICAL AIML BASED APPROACHES

A thorough exploratory data analysis is typically performed on the dataset first, followed by feature extraction and dimension reduction techniques such as principal component analysis (PCA). Finally, the most representative features are supplied to the ML algorithm. The knowledge base of different domains and applications can be quite different and often requires extensive specialised expertise within each field, making it difficult to perform appropriate feature extraction and maintain a good level of transferability for ML models trained in one domain to be generalised or transferred to other contexts or settings.

A. Principle Component Analysis (PCA):-

A approach known as PCA reveals the core structure of data in the most effective way to account for its variance. If a multivariate dataset is represented as a set of coordinates in a high-dimensional data space (one axis per variable), PCA can provide the user with a lower-dimensional projection of the object as seen from its most informative perspective. Because the sensitivity of various bearing defect characteristics can vary significantly across operating conditions, PCA has proven to be an effective and systematic feature selection scheme that provides guidance on manually selecting the most representative features for classification purposes. One of the early applications of PCA for bearing problem detection is documented in [11]. Experimental data show that the fault diagnostic accuracy ranges from 88% to 98%. The study found that the proposed PCA technique is more effective for classifying bearing problems with a better degree of precision and fewer input characteristics than using all of the original features.

1) PCA with two main component

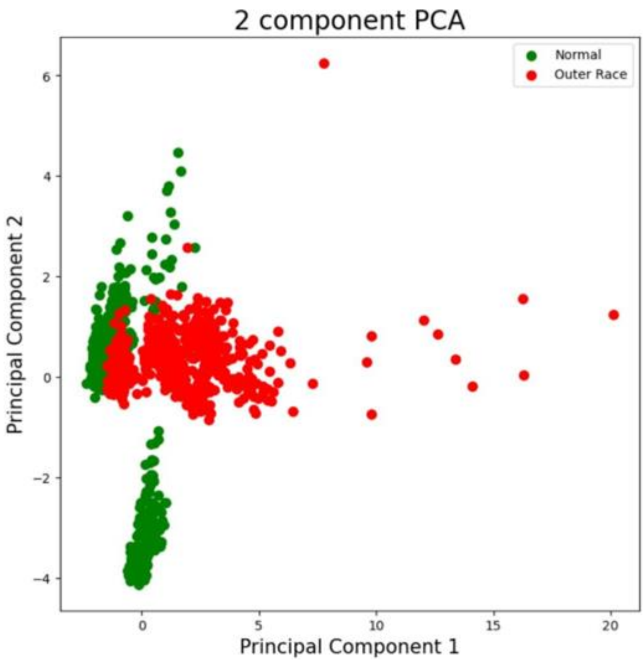

Fig. 8. PCA with two main component

2) PCA with three main component

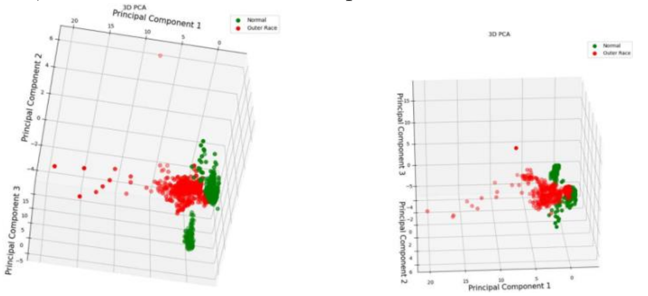

Fig. 9. PCA with three main component

B. K-Nearest Neighbors (K-NN)

The k-NN algorithm is a nonparametric approach for classification or regression. In k-NN classification, the output is the object's class, which is determined by a majority vote of its k nearest neighbours. k-NN is the central method for a data mining-based ceramic bearing defect classifier that uses acoustic signals. The experimental results indicate that the defect diagnosis accuracy is 77%.

C. Support Vector Machines (SVM)

SVMs are supervised learning models that process data for non-probabilistic classification or regression. One famous study on the use of SVM to find bearing flaws can be found in [12], where the classification results achieved by the SVM are optimal in all circumstances, with an overall improvement over the performance of ANN. Experimental results that the fault diagnosis accuracy is 84%.

D. Random Forest Classifier

Random Forest is a common machine learning approach that can be used to identify faults. It is a form of ensemble learning technique that uses numerous decision trees to improve accuracy and prevent overfitting. Each decision tree in the forest is formed on a random portion of training data and a random subset of input features, so they are less likely to overfit to the data. When applied to fault diagnosis, a random forest model can take sensor data or other measurements from a system and predict whether a fault has occurred or is likely to occur. The model can also provide insights into which features are most important for predicting faults and which factors are contributing the most to the faults.Accuracy, precision, recall, and F1 score are popular measures used to evaluate the effectiveness of a random forest model for fault identification. These measures can help you determine how effectively the model performs & whether it is correctly identifying faults and non-fault conditions. Experimental results that the fault diagnosis accuracy is 84%.

E. Decision Tree

The primary categorization and prediction technology is the decision tree. Decision tree learning is a common instance-based inductive approach that focuses on classification rules and displays them as decision trees inferred from a group of disordered and irregular instances. It analyses attributes between internal nodes of the decision tree in a top-down recursive manner, judges the downward branches based on the node's distinct attributes, and draws conclusions from the decision tree's leaf nodes. Experimental results show that 97% of faults can be diagnosed correctly.

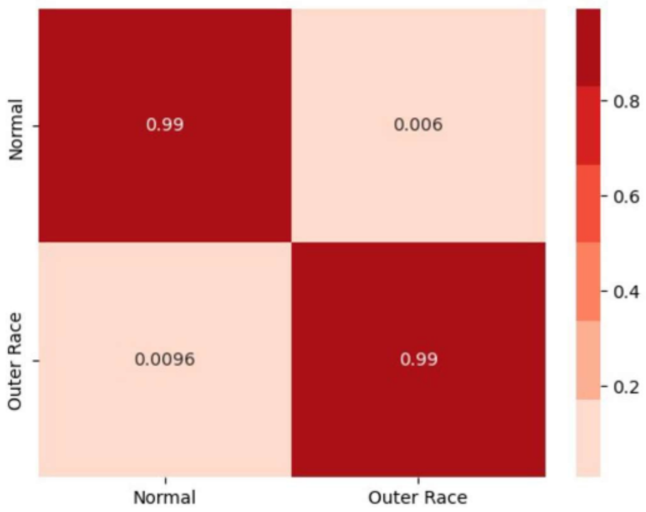

Fig. 10. Random forest.

F. Logistic Regression

The natural logarithm of an odds ratio is the fundamental mathematical principle behind logistic regression. The simplest example of a logistic regression is a 2 × 2 contingency table. Experimental results that the fault diagnosis accuracy is around 82%.

IV. CONCLUSION

In this study, the Decision Tree model and Principal Component Analysis (PCA) performed particularly well. The Decision Tree successfully captures complicated correlations in data using a hierarchical structure and provides easy interpretability, making it useful for understanding crucial predictive aspects. PCA, a dimensionality reduction technique, reduces high-dimensional data while keeping crucial information, hence improving the performance of future machine learning models.

The Decision Tree is best suited for tasks that require transparency and detailed pattern identification, but PCA excels at optimizing data for other algorithms. Both approaches have proven effective, and the choice between them is determined by the dataset and the problem's specific requirements. They complement each other's strengths, resulting in powerful machine-learning solutions.

REFERENCES

[1] Slavic, J. et al. “Bearing-Fault Rating Using Force Measurements-Application to Real data‖”, Journal of Vibration & Control, vol. 17 no. 14, pp. 2164-2174, 2011.

[2] Kunli, M.et al. “Fault Diagnosis of Rolling Element Bearing Based on Vibration Frequency Analysis”, IEEE Third International Conference on Measuring Technology and Mechatronics Automation (ICMTMA), pp. 337-341, 2011.

[3] Tandon, N.et al. “A Review of Vibration and Acoustic Measurement Methods for the Detection of Defects in Rolling Element Bearings”, Tribology International, Vol. 32, pp.469–480, 1999.

[4] Malla, C.et al. “Review of Condition Monitoring of Rolling Element Bearing using Vibration Analysis and Other Techniques”, Journal of Vibration Engineering and Technology, Vol. 7 (4), pp. 407-414, 2019.

[5] Caesarendra, W.et al. “An application of nonlinear feature extraction-a case study for low speed slewing bearing condition monitoring and prognosis”, ASME International Conference on Advanced Intelligent Mechatronics (AIM), Wollongong, Australia, July 9-12, 2013.

[6] Wang, J.et al. “Fractional envelope analysis for rolling element bearing weak fault feature extraction”, IEEE/CAA Journal of Automatica Sinica, Vol. 4, No. 2, pp. 353-360, 2017.

[7] Tandon, N.et al. "Vibration and acoustic monitoring techniques for the detection of defects in rolling element bearings - a review," Shock and Vibration Digest, vol. 24, pp. 3-11, 1992.

[8] Braun, S.et al. "Analysis of Roller/Ball Bearing Vibrations," Transactions of the ASME, vol. 101, pp. 118-125, 1979.

[9] Alfredson, R. J.et al. "Time Domain Methods for Monitoring the Condition of Rolling Element Bearings," presented at Mechanical Engineering Transactions, 1985.

[10] Dyer, D.et al. "Detection of Rolling Element Bearing Damage by Statistical Vibration Analysis," Journal of Mechanical Design, Transaction of the ASME, vol. 100, pp. 229-235, 1978.

[11] Mathew, J.et al. "The Condition Monitoring of Rolling Element Bearings Using Vibration Analysis,"Journal of Vibration, Acoustics, Stress, and Reliability in Design, Transaction of ASME, vol. 106, 1984.

[12] Tandon, N.et al. "Detection of Defects in Rolling Element Bearings by Vibration Monitoring," IE(I) Journal MC, vol. 73, pp. 271-282, 1993.

[13] Liu, T.I.et al. "Intelligent Monitoring of Ball Bearing Conditions," Mechanical Systems and Signal Processing, vol. 6, pp. 419-431, 1992.

[14] Mathew, J.et al. "Frequency Domain Methods for Monitoring the Condition of Rolling Element Bearings," presented at Mechanical Engineering Transactions, 1985.

[15] Taylor, J.I.et al. "Identification of Bearing Defects by Spectral Analysis," Journal of Mechanical Design, Transaction of the ASME, vol. 102, pp. 199-204, 1980.

[16] Smith, J.D.et al. "Vibration monitoring of rolling element bearings by the high frequency resonance technique - a review," Tribology International, vol. 17, pp. 3-10, 1984.

[17] Yudin, L.A.et al. "Automatic Diagnostics and Condition Prediction of Rolling Element Bearings Using Enveloping Methods," presented at Vibration Institute Proceedings, 18th Annual Meeting, Hershey, Pennsylvania, June 21-23, 1994.

[18] Mitchell, J.S.et al. "Assessing the Condition and Lifetime of Rolling Element Bearings from a Single Measurement," presented at Vibration Institute Proceedings, 19th Annual Meeting, Indianapolis, Indiana, June 20-22, 1995.

[19] Fnaiech, F.et al. “Application of empirical mode decomposition and artificial neural network for automatic bearing fault diagnosis based on vibration signals” Applied Acoustics, vol. 89, pp. 16-27, 2015.

[20] Harsha, S. P.et al. “Rolling element bearing fault diagnosis using wavelet transform” Neurocomputing, vol. 74, pp. 1638-1645, 2011.

[21] Hemmati, F.et al. “Roller bearing acoustic signature extraction by wavelet packet transform, applications in fault detection and size estimation” Applied Acoustics, vol. 104, pp.101-118, 2016.

[22] Harsha, S. P.et al. “Fault diagnosis of rolling element bearing using cyclic autocorrelation and wavelet transform” Neurocomputing, vol. 110, pp. 9-17, 2013.

[23] Kumar, N. P.et al. “Journal Bearing Fault Detection Based on Daubechies Wavelet” Archives of Acoustics, vol. 42(3), pp. 401-414, 2017.