Mahine Learning-Based Fault Detection of Electric Motor Bearings Using Vibration Analysis

SHARAFIZ ABDUL RAHIM
and MUHAMAD NABIL HAKIM ABD RAHIM

ABSTRACT

Bearing faults are a leading cause of electric motor failure, often resulting in costly downtime and maintenance. This research proposes a machine learning-based framework for early fault detection using vibration signal analysis. Vibration data from a bearing test rig under five operating conditions were collected and transformed using Fast Fourier Transform (FFT) to extract key frequency-domain features, while Principal Component Analysis (PCA) was used to reduce data complexity. Five machine learning models were evaluated, with the Medium Gaussian SVM achieving the highest accuracy of 96.8%. Although PCA-based models like Ensemble Subspace KNN had slightly lower accuracy (94.0%), they significantly reduced training time. The findings highlight a trade-off between classification accuracy and computational efficiency, with FFT offering superior fault-specific pattern retention and PCA enhancing speed. The proposed approach offers a scalable solution for real-time, predictive maintenance, and future work may explore hybrid feature extraction or deep learning to further improve performance.

INTRODUCTION

Electric motors are essential in industrial systems, with bearing faults being a major cause of failure. Early detection remains difficult due to subtle, nonlinear vibration and acoustic signal changes. Traditional methods using manual inspections or threshold-based systems are often inefficient. Recent studies [1–3], have shown machine learning (ML) can achieve over 95% accuracy in detecting bearing faults. Research using FFT and PCA [2, 4, 5] has further improved feature extraction. This study builds on prior work by evaluating multiple ML models in MATLAB's Classification Learner.

The novelty of this work lies in its comprehensive consideration of five common bearing conditions: normal, ball fault, inner race fault, outer race fault, and combined faults. These scenarios closely mimic real-world motor failures. By training models on this broader range of conditions, the framework better reflects industrial challenges. This study aims to develop a robust and scalable ML-based predictive maintenance approach capable of distinguishing among fault types, thereby improving electric motor reliability

Sharafiz Abdul Rahim, PhD, Email: sharafiz@upm.edu.my. Sound and Vibration Research Group, Department of Mechanical and Manufacturing Engineering, Universiti Putra Malaysia, Serdang, Malaysia

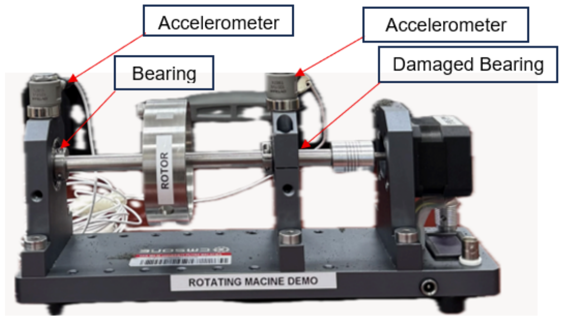

Figure 1. The Shaft-bearing demo used for vibration signal acquisition.

and reducing downtime in complex operating environments.

This project seeks to solve the problem of effectively analyzing noisy vibration and acoustic signals using Classification Learner to rank the best Machine Learning models that can detect bearing damage in electric motors (Figure 1). The objectives of this project are:

1. To implement the machine learning model for early detection of bearing damage in electric motors. 2. To investigate the accuracy of machine learning models for fault detection of the electric bearing motor

DATA ACQUISITION SETUP FOR NORMAL CONDITION

The initial phase of this project involved setting up the data acquisition system to collect vibration signals from the electric motor under normal operating conditions. The process was carried out using a National Instruments Data Acquisition (NI DAQ) device integrated with MATLAB through the Data Acquisition Toolbox. The acquisition process consisted of hardware setup, session configuration in MATLAB, and channel initialization. The vibration data was recorded using a data acquisition system interfaced with signal processing software. A sampling frequency of 2048 Hz was used, and each recording session lasted for 5 seconds. A total of 100 trials were conducted to ensure sufficient data was available for statistical analysis and pattern identification (Table 1). The collected data was saved in both MATLAB data format and spreadsheet format for ease of processing and accessibility. Each trial produced two sets of vibration signals corresponding to the two accelerometer positions. These signals were timestamped and labeled according to their operational condition for proper classification and traceability. All acquired datasets were stored in an organized directory for structured post-processing. Subsequently, the Fast Fourier Transform (FFT) was applied to convert the time-domain signals into the frequency domain. The resulting frequency spectra provided a more detailed view of periodic patterns and harmonic content that may be relevant to rotating machinery diagnostics.

DATA PREPROCESSING

To facilitate structured analysis, the vibration data was first sorted according to the different bearing conditions studied. These conditions included: • Normal bearing: 100

MATLAB DAQ Setup Session	Value
Sampling Rate	2048 Hz
Time Taken	5 seconds
Total Trial for Each Sample	100
Frequency Data Point	10240

TABLE I. Data acquisitions.

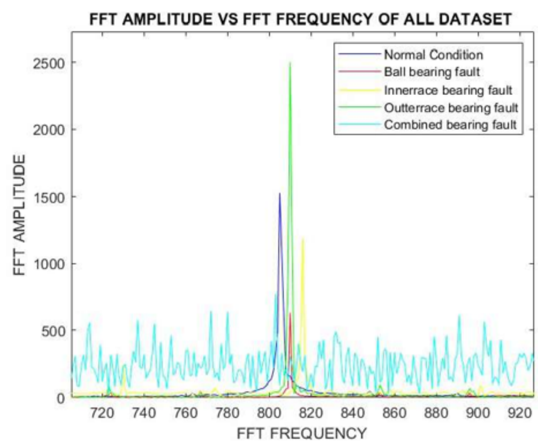

Figure 2. FFT transformed data obtained from acceleration signals of all bearing conditions.

samples, • Ball bearing fault: 100 samples, • Inner race fault: 100 samples, • Outer race fault: 100 samples, and • Combined bearing faults: 100 samples.

FREQUENCY TRANSFORMATION VIA FFT

Using MATLAB, the raw time-series signals were segmented into 2,048 sampling rate. Each segment was labeled based on its fault type. This facilitated supervised learning where each input sample had a known target output [6,7] The Fast Fourier Transform (FFT) was applied to convert each segment from the time domain to the frequency domain. This helped to reveal the frequency components associated with bearing defects. An example of the FFT spectrum is shown below These amplitude spikes represent the fundamental defect frequencies or their harmonics generated by the rotation of the faulty elements. The outer race fault (cyan) shows the most prominent peak, followed by the inner race and ball faults (Figure 2).

FEATURE SELECTION USING PRINCIPAL COMPONENT ANALYSIS (PCA)

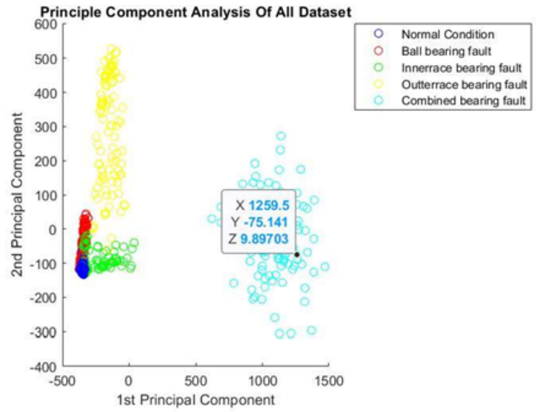

Figure 3. Principle Component Analysis of all bearing conditions using the FFT data.

After preprocessing and organizing the vibration data in the frequency domain, the next crucial step was feature selection. This was implemented to reduce the dimensionality of the data set while preserving the most informative features for accurate classification of bearing conditions. One of the most widely used techniques for this purpose is Principal Component Analysis (PCA), which was applied in this study. To visualize the effectiveness of dimensionality reduction and separability between different bearing conditions, a three-dimensional scatter plot was generated using the first three principal components (Figure 3). Each condition (normal, ball fault, inner race fault, outer race fault, and combined faults) was represented by a different color in the plot. This visualization clearly demonstrated the clustering behavior of different fault types in the reduced feature space. Well-separated groups indicated that the PCA-transformed characteristics retained distinct patterns associated with each condition, validating thus the utility of PCA in improving the classification potential.

MACHINE LEARNING

The preprocessed data set was divided into training sets (80%) and testing sets (20%) using a randomized split. This ensured that the models were trained on a representative subset of the data while retaining a separate portion for unbiased performance evaluation. The split was stratified to maintain the proportional distribution of fault classes in both sets, preventing imbalance issues during training. All available machine learning models in MATLAB's Classification Learner toolbox were initially trained on both the FFT-reduced and PCA-transformed datasets. This comprehensive training approach allowed for a thorough evaluation of each algorithm's potential performance [8]. Following this initial phase, we systematically analyzed the validation accuracy of all trained models and selected the top five best-performing classifiers for further evaluation and comparison. The selection criteria focused primarily on validation accuracy, ensuring we retained only the most effective models for bearing fault detection. The five selected models that demonstrated superior performance were Medium Gaussian Support Vector Machine (SVM), SVM with Kernel Function, Bi-layered Neural Network, Ensemble Subspace k-Nearest Neighbors (KNN) and Quadratic Discriminant Analysis (QDA).

	Type of machine learning	Accuracy Validation(%)	Training time (sec)
USING FFT reduced range as Dataset	SVM (Medium Gaussian SVM)	96.8	1.9137
	KNN (Subspace KNN)	96.0	4.9592
	Neural Network (Wide Neural Network)	94.4	7.244
	Kernel (Logistic Regression Kernel)	94.4	8.129
	Quadratic Discriminant	90.2	0.98062

TABLE II. Result of ML performance using FFT reduced range as dataset.

	Type of machine learning	Accuracy Validation (%)	Training time (sec)
USING PCA Data as Dataset	Ensemble (Subspace KNN)	94.0	3.2878
	Quadratic Discriminant	93.4	1.39
	Ensemble (Bagged Trees)	93.0	2.8579
	SVM (Median Gaussian SVM)	91.2	2.11
	Neural Network (Bilayered Neural Network)	90.2	3.8268

TABLE III. Result of ML performance using PCA transformed dataset.

CONCLUSION

The results showed that frequency-domain features extracted using FFT provided the most reliable fault signatures, enabling classification models to achieve up to 96.5% accuracy. The Medium Gaussian SVM emerged as the top-performing algorithm, combining high precision with practical training times. However, the research also revealed important limitations. While PCA reduced computational requirements by approximately 30%, this efficiency gain came at the cost of slightly lower accuracy, particularly for inner race faults that were sometimes misclassified as ball defects. These findings suggest that while dimensionality reduction techniques offer benefits for real-time applications, they may not be suitable for all fault detection scenarios.

ACKNOWLEDGMENT

This work is carried out with the support of the Universiti Putra Malaysia under IPM Grant 9688000 award.

REFERENCES

1. Zhang, W., D. Yang, and Y. Wang. 2020. "Bearing fault diagnosis using deep convolutional neural networks with wide first-layer kernels," *IEEE Access*, 8:86560–86569.
2. Zhou, M., H. Zhang, and X. Liu. 2022. "Bearing fault diagnosis based on FFT and ensemble learning under variable speed conditions," *Mechanical Systems and Signal Processing*, 172:108978.

3. Alwodai, A. and P. Samuel. 2021. “Application of machine learning algorithms in bearing fault diagnosis: A review,” *Machines*, 9(3):54.
4. Gupta, S., R. Tiwari, and P. Gupta. 2019. “An intelligent bearing fault diagnosis system using wavelet packet transform, principal component analysis, and artificial neural network,” *International Journal of Condition Monitoring*, 9(1):21–30.
5. Rao, S. and V. Kumar. 2023. “An efficient machine learning approach for bearing fault detection using FFT-PCA-based feature extraction,” *Journal of Vibration and Control*.
6. Rahim, S. 2018. *Investigating the Effect of Variable Mass Loading in Structural Health Monitoring from a Machine Learning Perspective*, Phd thesis, University of Sheffield.
7. Rahim, S., A. Najmi, R. Samin, N. I. Rahman, and S. Sathurshan. 2024. “The Effects of Signal Processing Techniques in Damage Detection and Structural Health Monitoring,” in *Journal of Physics: Conference Series*, IOP Publishing, vol. 2721, p. 012022, doi:10.1088/1742-6596/2721/1/012022, 7th International Conference on Noise, Vibration and Comfort (NVC 2023), Kuala Lumpur, Malaysia, 5–7 December 2023.
8. Zhao, R., S. Tang, E. E. B. Supeni, S. B. A. Rahim, and L. Fan. 2024. “A Review of Object Detection in Traffic Scenes Based on Deep Learning,” *Applied Mathematics and Nonlinear Sciences*, 9(1):1–25, doi:10.2478/amns-2024-0322.