

Fault Diagnosis of Air Compressor (AC) System using Local Mean Decomposition (LMD) and Logistic Regression (LR) Machine Learning Classifier

Atul Dhakar^{1,*}, Bhagat Singh², Pankaj Gupta³

Abstract

This article presents a detailed and systematic procedure for performing fault diagnosis in an air compressor (AC) system by analyzing the audio signals generated during its operation. The analysis covers both normal (healthy) conditions and seven distinct types of faults, including bearing failure, flywheel malfunction, inlet valve leakage, outlet valve leakage, non-return valve failure, piston ring defect, and rider belt issues. To acquire the acoustic signals, the researchers utilized a unidirectional microphone in combination with National Instruments hardware, specifically the NI 9234 data acquisition device and the NI 9172 USB interface. These tools ensured accurate and high-quality audio signal collection from the AC system under various operating conditions. Once the audio data were collected, they were subjected to a sophisticated, non-traditional signal processing method known as Local Mean Decomposition (LMD). This technique is especially effective for handling non-linear and non-stationary signals commonly found in mechanical systems. Following decomposition, six key statistical indicators (SIs) were computed to extract distinguishing fault features. These indicators included the mean of signals (M_S), variance (σ^2_{VS}), mean of the root square (M_{RS}), mean of the root amplitude (M_{RA}), mean of the absolute amplitude (M_{AA}), and the kurtosis index (K_U). The extracted features were classified using a Logistic Regression (LR) machine learning algorithm. The integrated use of LMD, the kurtosis index, and the LR classifier led to a classification accuracy of 90.30%, demonstrating strong effectiveness in identifying faults within the AC system

Keywords: AC system, LMD, statistical indicators, logistic regression classifier, fault diagnosis

INTRODUCTION

Online Condition Monitoring (OCM) has become a vital field of study for the diagnosis of mechanical component faults due to industrial automation and technological advancements. Fault detection is critical for system design and maintenance [1]. OCM involves continuously monitoring machine parameters such as vibration, temperature, and pressure to identify changes and detect potential issues early [2–5]. Author If faults occur in any of the components while they are in use, this proactive approach can help prevent serious damage.

Acoustic and vibration measurements are among the most widely used techniques for fault diagnosis, especially in machinery that involves rotational motion [6–8]. The AC is a type of rotary machine that requires online condition monitoring to prevent further damage and maintain safe

***Author for Correspondence**
Atul Dhakar

¹Assistant Professor, Department of Mechanical Engineering, University Institute of Technology (UIT), RGPV, Shivpuri, Madhya Pradesh, India

²Head & Associate Professor, Department of Mechanical Engineering, Jaypee University of Engineering and Technology, Guna, Madhya Pradesh, India

³Assistant Professor, Department of Mechanical Engineering, Jaypee University of Engineering and Technology, Guna, Madhya Pradesh, India

Received Date: September 13, 2025

Accepted Date: September 29, 2025

Published Date: February 14, 2026

Citation: Atul Dhakar, Bhagat Singh, Pankaj Gupta. Fault Diagnosis of Air Compressor (AC) System using Local Mean Decomposition (LMD) and Logistic Regression (LR) Machine Learning Classifier. Journal of Polymer & Composites. 2026; 14(Special Issue 1): S416–S427p.

operating conditions [9]. Fault diagnosis can be categorized into three forms: (1) Human perception, (2) Extraction of feature based on signal processing, and (3) Classification using machine learning. Data collection, the first step in the fault diagnosis process, involves recording all relevant machine parameters.

Various sensors are used to collect data from AC systems; however, obtaining data from multiple locations can sometimes be challenging and impractical. To address this issue, careful sensor placement and selection are crucial for effective online condition monitoring of AC system [10, 11]. In such applications, microphones and accelerometers are the two most commonly used types of sensors. According to various researchers, microphones are considered the most precise and appropriate tools for signal processing [12, 13]. Additionally, selecting the right signal processing technique significantly influences both the effectiveness of OCM and the accuracy of the machinery assessment [14]. Historically, researchers have employed various signal processing methods to analyze collected data for assessing the condition of rotating machinery. These methods include Local Mean Decomposition (LMD) [15, 16], Empirical Mode Decomposition (EMD) [17, 18], and the Wavelet Transform [19, 20]. These techniques can be categorized into three groups: 1) Time Domain,

2) Frequency Domain, and 3) Time-Frequency Domain [20]. Among these methods, LMD is particularly advantageous, as it effectively mitigates the drawbacks associated with the others, such as mode aliasing and end effects [21]. In this study, the local mean decomposition technique was applied to process both healthy and faulty signals from an AC system. Through demodulation, the signals were decomposed into various distinct Product Functions (PFs).

Feature extraction, a crucial step in OCM for fault diagnosis, was then performed. In previous research, researchers have suggested a range of statistical indicators (SIs) for feature extraction, including mean, variance, absolute mean amplitude, root mean amplitude, root mean square, kurtosis, and more. The selection of the most significant statistical indicators for classification is crucial after feature extraction. Rotating machinery, such as compressors, is classified based on its faulty and healthy states. Evaluating the correlation between the class data and the training samples is the first stage in the classification process. This correlation is then used to predict the class of the test data [19–22]. Logistic Regression (LR) classifier has been widely employed by re-searchers to address classification challenges [23, 24]. However, these methods are often tailored to specific problems. In this paper, defective states in an AC system are classified using the LR algorithm. Figure 1 illustrates methodology for fault diagnosis model, which involves five steps: 1) Collection of data, 2) LMD technique, 3) Extraction of feature, 4) Classification using the LR classifier, and 5) Fault diagnosis.

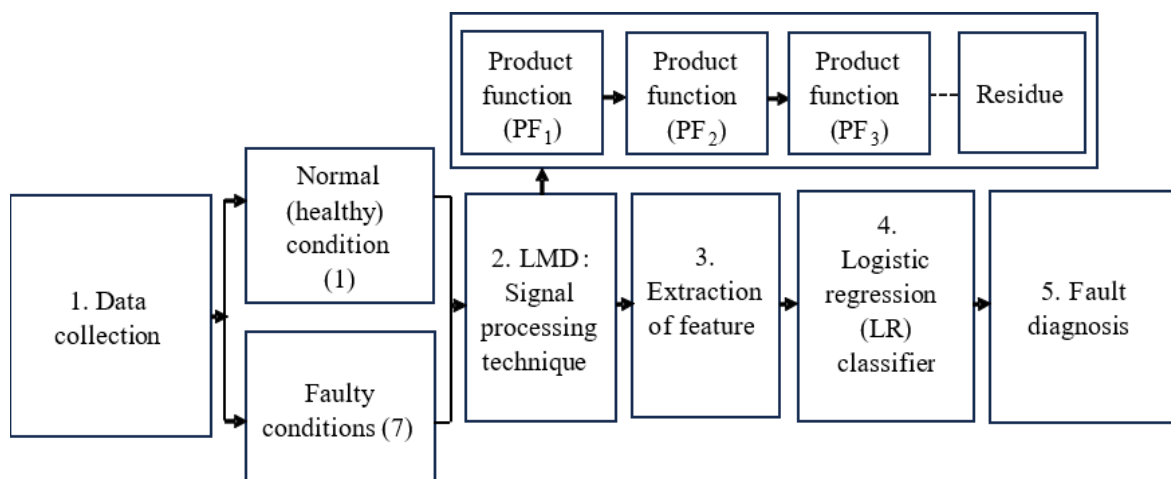


Figure 1. Complete methodology.

EXPERIMENTATION

In the last several decades, Various studies have been executed to identify and analyze issues in various rotary devices, such as rotors, air compressors, and bearings. However, there are relatively few studies specifically focused on fault diagnosis in AC system, which motivated the current research. This focus on AC system fault detection represents the novelty of this work. The authors utilized a unidirectional microphone (Model code: UTP-30) to capture both healthy and faulty signals from an AC system. By strategically positioning the microphone near problematic areas, they were able to minimize background noise and collect cleaner signals.

To capture and analyze signals during operation, a variety of sensors can be employed. Among these, micro-phones have been identified as the most effective tools for detecting faulty signals in rotary devices. Figure 2(a) presents the audio signal acquisition for 7 faulty conditions and 1 normal condition, providing an overview of the complete AC system. The placement of microphones to record faulty signals from the AC system is shown in Figure 2(b). Figure 2(c) shows the configuration used for collecting signals, featuring the integration of a microphone and DAQ hardware connected to a computer system. Several test runs revealed that positioning the micro-phone 1.5 cm from the target area significantly improved the quality of the recorded sound, resulting in clearer and more distinct signals [9, 10].

Once the audio signals from the air compressor system were successfully acquired, the next step involved processing these signals using advanced signal analysis techniques to extract important information.

PROCESSING THE ACQUIRED SIGNALS

The experimental signals obtained have been analyzed employing the most recent signal processing techniques, such as LMD. It is a method for breaking down a signal into several parts, or PFs, which represent different amplitude and frequency elements. The complete mathematics of the LMD technique is presented in Figure 3.

Figure 4 illustrates various segregated product functions (PFs) along with their respective Fast Fourier Transforms (FFT). In this study, the LMD method was used to extract the PFs from the signals acquired from the AC system.



Figure 2. Experimentation: (a) AC system, (b) microphone placements, (c) DAQ [9].

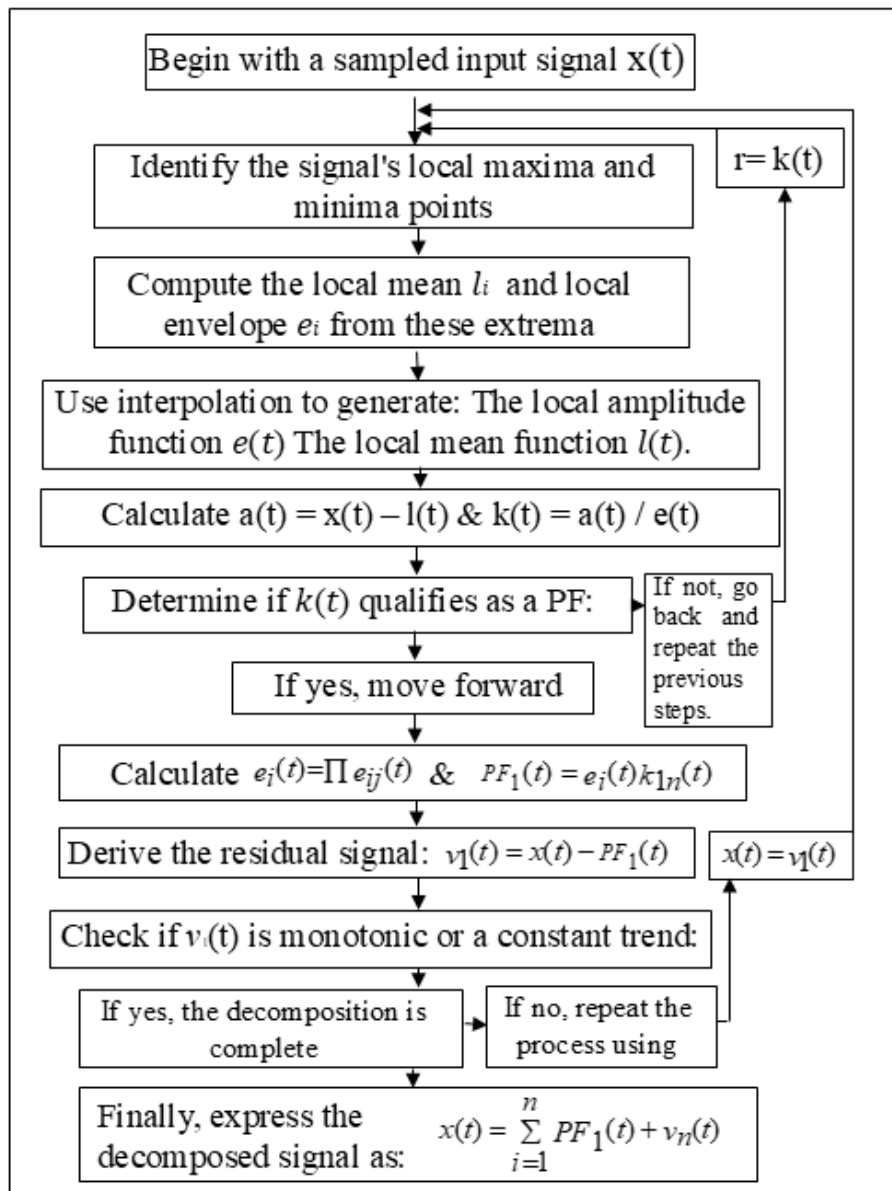


Figure 3. Complete mathematics of the LMD technique [21].

To pinpoint the key PFs accompanying faulty conditions, Fast Fourier Transforms (FFTs) were determined. The fault-related PFs were identified by examining those with the highest amplitudes and frequency components close to the system's natural frequency. The presence of mode mixing could account for the broad frequency observed in certain PFs. A comprehensive analysis of the frequency clusters and their associated amplitudes is vital for identifying the PFs that signal faults. As depicted in Figure 4, the first PF, with maximum amplitude, includes a frequency cluster near the normal operating frequency. This makes PF_1 highly effective in highlighting the fault characteristics in this case. Additionally, the PFs of other collected audio signals were analyzed, leading to the identification of the key product functions (PFs) responsible for generating errors. Ultimately, these specific PFs were evaluated using six SIs to diagnose faults in the AC system.

FEATURE EXTRACTION

All the gathered data were processed using the LMD technique, as depicted in Figure 4. The statistical indicators (SIs) of these decomposed signals such as M_s , σ^2_{vs} , M_{rs} , M_{ra} , M_{aa} , and K_u were analyzed for feature extraction. Table 1 provides the SIs along with their corresponding equations

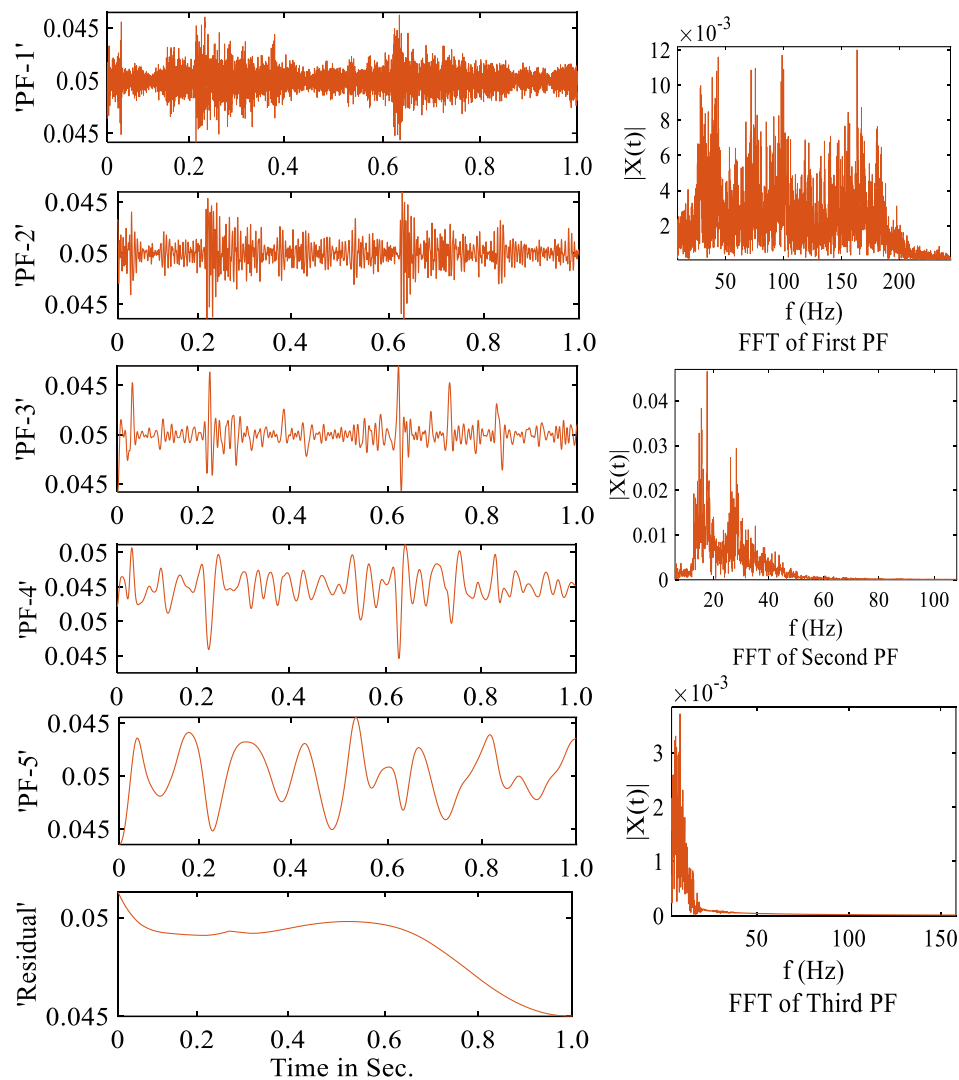


Figure 4. PFs and FFTs of signals.

Table 1. SIs and their equations [21].

SIs	Equations	
$MOS (M_s)$	$M_s = \frac{1}{L} \sum_{i=1}^N S_i(t)$	Where, MOS = Mean of signals, VOS = Variance of signals, MRS = Mean of root square, MRA = Mean of root amplitude, MAA = Mean of absolute amplitude, Si(t) = Signal amplitude, Ms = Signals mean, L = Number of samples.
$VOS (\sigma_{VS}^2)$	$\sigma_{VS}^2 = \frac{1}{L-1} \sum_{i=1}^N (S_i(t) - M_s)^2$	
$MRS (M_{RS})$	$M_{RS} = \sqrt{\frac{1}{L} \sum_{i=1}^N S_i(t)^2}$	
$MRA (M_{RA})$	$M_{RA} = \sqrt{\frac{1}{L} S_i(t)_1^2 + S_i(t)_2^2 + \dots + S_n(t)_n^2}$	
$MAA (M_{AA})$	$M_{AA} = \frac{1}{L} \sum_{i=1}^N S_i(t) $	
$Kurtosis (K_U)$	$K_U = \frac{\sum_{i=1}^N S_i(t) - (M_s) ^4}{(N-1)(\sigma_{VS})^4}$	

The frequency components of the 1,800 recorded signals were analyzed using FFTs. After a detailed examination of the prominent peaks in the corresponding frequency spectra, significant fault frequencies (PFs) were identified. The dataset consists of 1 healthy case and 7 faulty cases (as presented in Table 2), containing all the required signal indicators (SIs).

The primary object of this research is to classify the most critical fault among various unhealthy conditions. Figure 5 presents the signal plots for 1 healthy state and 7 faulty scenarios, highlighting differences in amplitude and sample count. Notably, the bearing and flywheel faults showed prominent signal peaks, with the bearing fault exhibiting the highest amplitude. Consequently, the bearing defect was deemed the most severe and chosen for in-depth analysis.

CLASSIFICATION USING LOGISTIC REGRESSION (LR) CLASSIFIER

LR is a statistical model primarily used for binary classification problems, where the goal is to predict one of two possible outcomes, typically represented as 0 or 1. De-spote its name, it is a classification algorithm rather than a regression model. The complete mathematics of LR classifier are as follows:

Step 1: Model Representation

In In Logistic Regression (LR), the likelihood that the response variable Y belongs to the positive class (label = 1) is estimated using the sigmoid function:

Table 2. PFs of SIs.

Exp. No.	Ms	σ^2_{VS}	M _{RS}	M _{RA}	M _{AA}	K _U	States
1.	0.0008	0.03	0.18	0.10	0.13	6.9	Faulty Bearing (FB)
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
225.	0.004	0.05	0.24	0.15	0.18	4.1	FB
226.	0.001	0.05	0.22	0.14	0.17	4.1	Faulty Flywheel (FF)
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
450.	0.0032	0.039	0.19	0.12	0.14	5.0	FF
451.	0.0006	0.073	0.27	0.18	0.21	3.1	Healthy (H)
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
675.	-0.0002	0.05	0.24	0.15	0.18	3.7	H
676.	-0.0005	0.02	0.16	0.09	0.11	7.1	Faulty Inlet Valve (FIV)
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
900.	0.0018	0.03	0.18	0.11	0.13	5.2	FIV
901.	0.0003	0.03	0.19	0.12	0.14	6.2	Faulty Outlet Valve (FOV)
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
1125.	0.0016	0.05	0.22	0.15	0.17	3.5	FOV
1126.	-0.0004	0.03	0.19	0.11	0.13	10.9	Faulty non-return Valve (FNRV)
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
1350.	0.0008	0.04	0.20	0.12	0.15	4.5	FNRV
1351.	-0.0004	0.07	0.27	0.17	0.21	3.6	Faulty Piston Ring (FPR)
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
1575.	-0.0004	0.05	0.22	0.14	0.17	4.2	FPR
1576.	0.0014	0.02	0.15	0.08	0.10	10.8	Faulty Rider Belt (FRB)
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮

1800.	0.0030	0.03	0.18	0.11	0.13	5.3	FRB
-------	--------	------	------	------	------	-----	-----

$$P(Y=1|X) = \sigma(S) \quad (1)$$

Where S is defined as a weighted sum of input features:

$$S = \omega_0 + \omega_1 x_1 + \omega_2 x_2 + \dots + \omega_n x_n = \omega x \quad (2)$$

$P(Y = 1|X)$ = Probability of the positive class, $\sigma(S)$ = The logistic function.

Step 2: Logistic (Sigmoid) Function

The logistic function transforms any real-valued input into the interval (0, 1):

$$\sigma(S) = \frac{1}{1 + e^{-S}} \quad (3)$$

Step 3: Classification Rule

An instance is classified into the positive class if the predicted probability exceeds a certain threshold (commonly 0.5):

$$\hat{Y} = \begin{cases} 1, & \text{if } \sigma(S) \geq 0.5 \\ 0, & \text{otherwise} \end{cases} \quad (4)$$

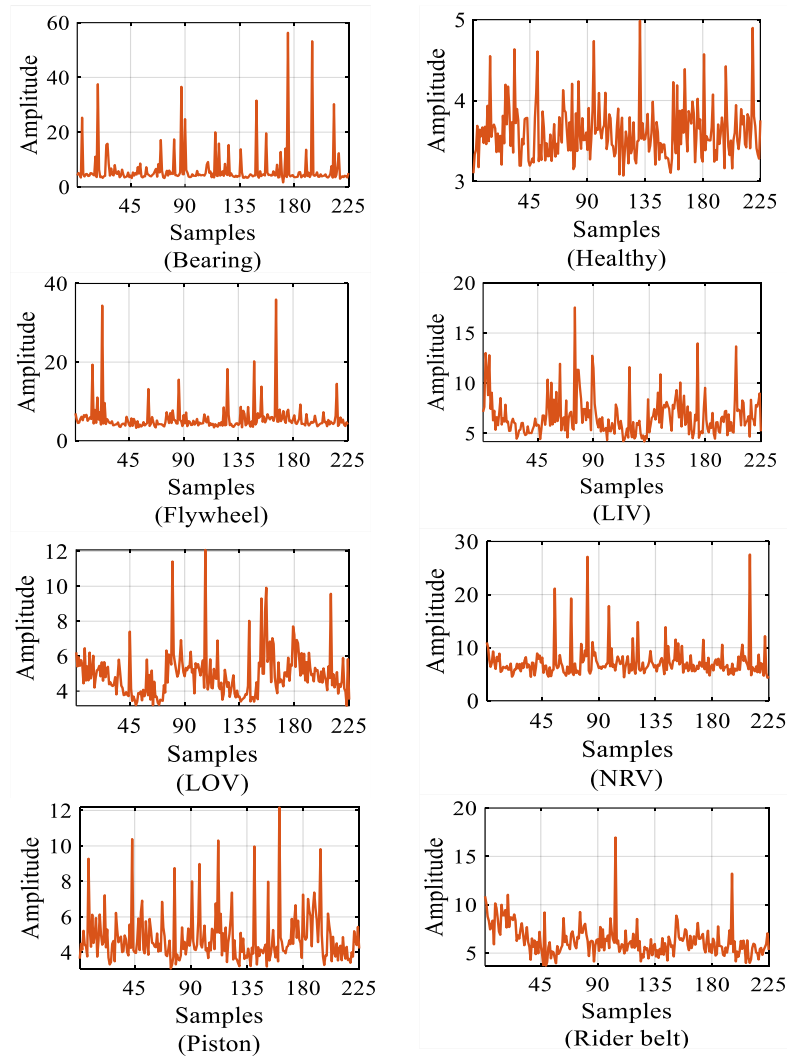


Figure 5. Amplitude vs samples graphs.

Step 4: Cost Function (Loss Function)

The model assesses prediction error using the binary cross-entropy (log-loss),

$$L = -\frac{1}{K} \sum_{i=1}^K \left[y^{(a)} \log(\hat{y}^{(a)}) + (1 - \hat{y}^{(a)}) \log(1 - \hat{y}^{(a)}) \right] \quad (5)$$

Where K = Number of samples, $y^{(A)}$ = Actual label (0 or 1) for sample a .

Step 5: Gradient Descent

To minimize the loss function, weights are iteratively updated using gradient descent:

$$\omega_j := \omega_j - \eta \cdot \frac{\partial L}{\partial \omega_j} \quad (6)$$

Step 6: Gradient of the Cost Function

The gradient for each weight is computed as:

$$\frac{\partial L}{\partial \omega_j} = \frac{1}{K} \sum_{i=1}^K (\hat{y}^{(a)} - y^{(a)}) x_j^{(a)} \quad (7)$$

Step 7: Final Model Representation

After training, the model prediction is represented as:

$$\hat{y} = \sigma(\omega x) \quad (8)$$

With the classification process completed, the next step was to analyze how effectively the model diagnosed faults were using various evaluation techniques, as discussed in the fault diagnosis section.

DIAGNOSING OF FAULTS

This study utilizes the LR classifier for fault diagnosis. The analytical results are presented using three methods: Confusion Matrix (C-Matrix), Scatter Plots (SP) and Receiver Operating Characteristic (ROC) curves.

Fault Diagnosis Using SP

The distinction between the two classes in the dataset is clearly depicted using a scatter plot. Figure 6 illustrates the SP of normal and FB signals based on 6 statistical features across various samples. Notably, the kurtosis plot demonstrates a distinct separation between H and FB conditions.

In this plot, blue circles represent the normal (healthy) state, while red circles indicate the faulty state. The red circles display a significantly disrupted position referred to as Leptokurtic (kurtosis > 3), which indicates a signal with a relatively high peak, suggesting a faulty bearing case. Conversely, the blue samples represent a stable state known as Platykurtic (kurtosis < 3), characterized by a flat-topped signal that implies the absence of a fault. Thus, it can be concluded from this observation that kurtosis serves as an effective predictor of fault characteristics in the AC system.

Fault Diagnosis Using Confusion Matrix (CM)

The CM is a table that shows how well an algorithm can classify data by predicting a categorical label for every instance of an input. As shown in Figure 7, it displays the positive values that are true (TP), negative values that are true (TN), positive values that are false (FP), and negative values that are false (FN) values determined by the test data and model. This matrix is used to assess the predictive effectiveness of a particular classification method in the current study.

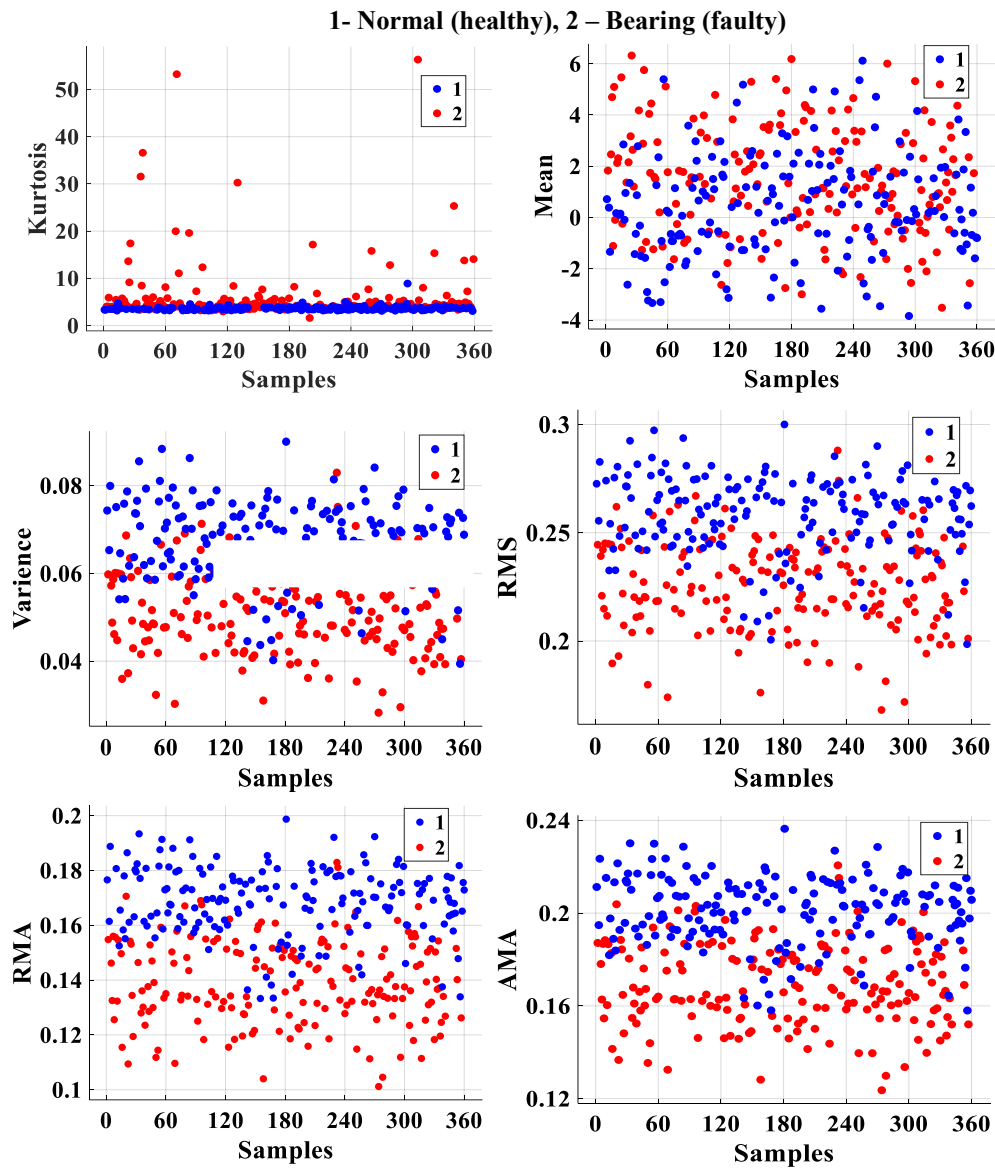


Figure 6. Amplitude vs samples graphs.

		Predicated Values	
		Positive (+)	Negative(-)
Actual Values	Negative(-)	T_P	F_P
	Positive (+)	F_N	T_N

Figure 7. C-Matrix.

In the current study, the LR classifier was employed to diagnose faults. The confusion matrix for the ensemble-based classifiers takes into account the normal and faulty bearing conditions in the AC system. The CM for the LR classifier is displayed in Figure 8 and has been used for assessment. These statistics reflect the T_P , T_N , F_P , and F_N for both normal and faulty bearings in training and testing cases. These parameters contribute to the overall classification accuracy of the LR classifier, and the performance of each LR classifier is described in detail below.

Fault Diagnosis using ROC Curves

The ROC curves used to assess how well the LR classifier performs in identifying bearing problems in an air compressor system are shown in Figure 9. For both training and testing datasets, these graphs show the correlation between the True Positive Rate (TPR) and False Positive Rate (FPR). A red dot on each curve marks the classifier's specific performance, representing the corresponding T_{PR} and F_{PR} values.

For the training data, Figures 9(a) and 9(b) show that the LR classifier resulted in FPR of 0.11 and 0.08, indicating that 11% and 8% of instances were incorrectly classified as positive. In the testing phase, Figures 9(c) and 9(d) display FPR of 0.00 and 0.04, meaning no false positives in one case and a 4% misclassification rate in the other.

The TPR values of 0.92 and 0.89 indicate that the LR classifier successfully identifies 92% of the normal class and 89% of the faulty bearing class as positive during the training phase. In the testing phase, TPR values of 0.96 and 1.00 show that the classifier accurately classifies 96% of the normal class and 100% of the faulty bearing class as positive.

Result and accuracy of LR classifier

The accuracy of the Logistic Regression (LR) classifier, which is determined by dividing the number of properly predicted instances by the total number of input samples, is used to assess the classifier's efficacy in this study.

$$Accuracy = \left[\frac{TP + TN}{TP + FP + FN + TN} \times 100 \right] \quad (9)$$

The values for TP, TN, FP, and FN for the LR classifier are 163, 162, 14, and 21, respectively, as illustrated in Figure 8(a).

$$Accuracy = \frac{163 + 162}{163 + 14 + 21 + 162} \times 100$$

$$Accuracy = 90.30\%$$

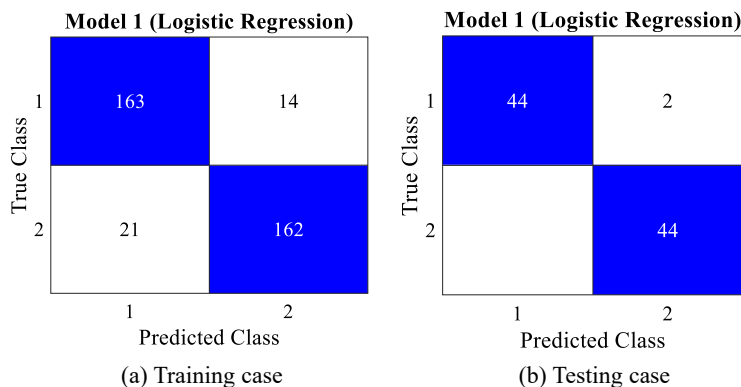
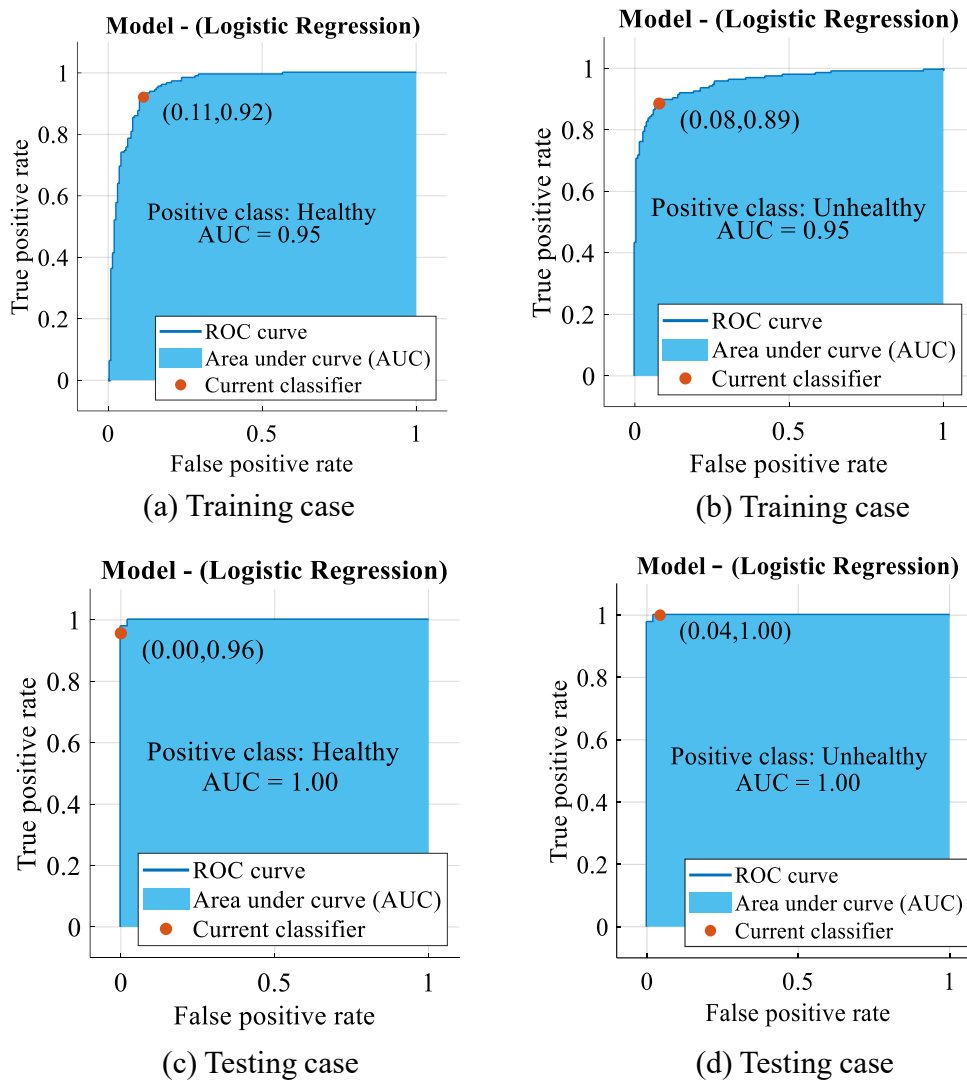


Figure 8. CM of LR classifier.**Figure 9.** ROC of LR classifier.

CONCLUSIONS

This research introduces a methodology for detecting defects in AC systems, encompassing data gathering, microphone positioning, signal processing, extraction of feature, and fault classification. The following salient features have been inferred from the analysis:

- LMD has been identified as a powerful tool, successfully overcoming the limitations associated with traditional diagnostic techniques in air compressors.
- Six statistical indicators such as MS, σ^2_{VS} , M_{RS} , M_{RA} , M_{AA} , and K_U were assessed. Kurtosis (K_U) emerged as the most effective feature for fault prediction.
- LR classifier achieved an impressive prediction accuracy of 90.30%, validating its efficacy for real-time fault diagnosis in AC systems.
- The proposed approach is highly effective for real-time diagnostics of air compressor issues, enabling timely maintenance and reducing operational downtime.

ACKNOWLEDGMENTS

- *Funding:* The present work has received no funds in any manner from any organization.
- *Conflicts of interest:* The authors declare that they have no involvement in any organization or entity with any financial interest.

REFERENCES

1. Ruonan L, Boyuan Y, Enrico Z, Xuefeng C. Artificial intelligence for fault diagnosis of rotating machinery: A review. *Mech Syst Signal Process*. 2018; 108:33–47.
2. Blodt M, Grabjon P, Raison B, Rostaing G. Models for bearing damage detection in induction motors using stator current monitoring. *IEEE Trans Ind Electron*. 2008;55(4):1813–1822.
3. Haji M, Toliyat HA. Pattern recognition – A technique for induction machines rotor broken bar detection. *IEEE Trans Energy Convers*. 2001; 16(4):312–317.
4. Nandi S, Toliyat HA, Xiaodong L. Condition monitoring and fault diagnosis of electrical motors – A review. *IEEE Trans Energy Convers*. 2005; 20(4):719–729.
5. Yan R, Gao R. Hilbert-Huang transform-based vibration signal analysis for machine health monitoring. *IEEE Trans Instrum Meas*. 2006; 55(6):2320–2329.
6. Scanlon P, Kavanagh DF, Boland FM. Residual life prediction of rotating machines using acoustic noise signals. *IEEE Trans Instrum Meas*. 2013; 62(1):95–108.
7. McInerny SA, Dai Y. Basic vibration signal processing for bearing fault detection. *IEEE Trans Educ*. 2003; 46(1):149–156.
8. Li R, He D. Rotational health machine monitoring and fault detection using EMD-based acoustic emission feature quantification. *IEEE Trans Instrum Meas*. 2012; 61(4):990–1001.
9. Verma NK, Salour AL. *Intelligent Condition Based Monitoring*. Springer. 2020; 256.
10. Verma NK, Sevakula RK, Dixit S, Salour AL. Intelligent condition-based monitoring using acoustic signals for air compressors. *IEEE Trans Reliab*. 2015; 99:1–19.
11. Delio T, Thlusty J, Smith S. Use of audio signals for chatter detection and control. *J Manuf Sci Eng Trans ASME*. 1992; 114(2):146–157.
12. Verma NK, Kadambari J, Abhijit B, Tanu S, Salour AL. Finding sensitive sensor positions under faulty condition of reciprocating air compressors. *IEEE Recent Adv Intell Comput Syst (RAICS)*. 2011: 242–246.
13. Shrivastava Y, Singh B, Sharma A. Identification of chatter in turning operation using WD and EMD. *Mater Today Proc*. 2018; 5(11):23917–23926.
14. Jun W, Guifu D, Zhongkui Z, Changqing S, Qingbo H. Fault diagnosis of rotating machines based on the EMD manifold. *Mech Syst Signal Process*. 2020; 135:1–21.
15. Yongbo L, Shubin S, Zhiliang L, Xihui L. Review of local mean decomposition and its application in fault diagnosis of rotating machinery. *J Syst Eng Electron*. 2019; 30(4):799–814.
16. Estrada A. Optimized LMD method and its applications in rolling bearing fault diagnosis. *Bioorg Med Chem Lett*. 2012; 23(6):149–152.
17. Maurya S, Singh V, Dhar NK, Verma NK. Improved EMD local energy with SVM for fault diagnosis in air compressor. *Adv Intell Syst Comput*. 2019; 799:81–92.
18. Yaguo L. Fault diagnosis of rotating machinery based on empirical mode decomposition. *Smart Sens Meas Instrum*. 2017; 26:259–292.
19. Peng ZK, Tse PW, Chu FL. A comparison study of improved Hilbert-Huang transform, and wavelet transform: Application to fault diagnosis for rolling bearing. *Mech Syst Signal Process*. 2005; 19(5):974–988.
20. Feng Z, Liang M, Chu F. Recent advances in time-frequency analysis methods for machinery fault diagnosis: A review with application examples. *Mech Syst Signal Process*. 2013; 38(1):165–205.
21. Gupta P, Singh B. Local mean decomposition and artificial neural network approach to mitigate tool chatter and improve material removal rate in turning operation. *Appl Soft Comput J*. 2020;96.
22. Bin GF. Early fault diagnosis of rotating machinery based on wavelet packets – Empirical Mode Decomposition feature extraction and neural network. *Mech Syst Signal Process*. 2012; 27(1):696–711.
23. Pandya DH, Upadhyay SH, Harsha SP. Fault diagnosis of rolling element bearing by using multinomial logistic regression and wavelet packet transform. *Soft Comput*. 2014; 18:255–266.
24. Yan J, Lee J. Degradation assessment and fault modes classification using logistic regression. *J Manuf Sci Eng*. 2004; 124(4):912–914.