Reports in Mechanical Engineering
Vol. 7, No. 2, 2026, pp. 146-159
ISSN: 2683-5894, DOI: https://doi.org/10.5281/zenodo.22688718

Predictive Maintenance of Industrial Gearbox Systems Using Multimodal Sensor Fusion and Explainable Machine Learning

Zhou Lee[1*], Liu Yue[2]
[1]Department of Physical Education, Peking University, 100871, Beijing, China
[2]Silla University, Sasang-gu, Busan, 46958, South Korea

Article Info

Article history:

Received June 19, 2026
Revised July 24, 2026
Accepted August 14, 2026

Keywords:

Predictive Maintenance (PM),
Industrial Gearbox Systems (IGS),
Multi-Model Sensor Fusion (MSF),
Machine Learning (ML)

ABSTRACT

In manufacturing plants, power-generation systems, transportation equipment, mining machinery, and automated production systems, industrial gearbox systems are essential components that can cause production losses, equipment damage, safety concerns, and substantial maintenance expenses if they fail unexpectedly. Traditional maintenance strategies, such as corrective maintenance and time-based preventive maintenance, are ineffective at detecting early signs of faults because they rely on set maintenance intervals or single sensor readings. This research presents a predictive maintenance framework for an industrial gearbox system based on fusing multiple sensor types with explainable machine learning (XAI). Common gearbox degradation mechanisms such as gear tooth damage, bearing defects, shaft misalignment, lubrication issues, and abnormal wear are characterized using time, frequency and time-frequency features. For fault classification, condition assessment, and remaining useful life prediction, the Study Considers Machine-Learning models such as Random Forest, Support Vector Machine, Gradient Boosting, and deep neural networks. Rather than using a single information source, the study uses multimodal fusion to improve model robustness by combining complementary information from multiple sensing modalities. One significant finding is the inclusion of explainable machine learning in predictive maintenance. The study uses techniques such as SHAP (SHapley Additive Explanations), feature importance analysis, and local explanation methods to identify which sensor signals and features contribute most to model predictions. This provides greater transparency and helps maintenance engineers understand why a gearbox has been graded healthy, degraded, or faulty. The framework is expected to enable earlier fault detection, more reliable diagnostics, less unplanned downtime, better maintenance scheduling, and more efficient use of maintenance resources. The multimodal sensing and explainable AI approach offers a practical, interpretable solution for intelligent condition monitoring of industrial gearboxes. The framework can facilitate the shift from traditional preventive maintenance to data-based, condition-based, and predictive maintenance, thereby enhancing equipment reliability, operational efficiency, and industrial productivity.

Corresponding Author:

Zhou Lee
Department of Physical Education, Peking University, 100871, Beijing, China
Email: 604544516@qq.com

1. Introduction

Industrial gearbox systems are essential components in mechanical and manufacturing systems, used to deliver torque to different types of industrial machines, manage speed, and convey energy. They are widely used in wind turbines, conveyors, production machines, energy generation equipment in the mining industry, motor vehicles, robotics, and heavy industry (Tang et al., 2026). These systems must operate reliably, as gearbox failures can cause lost production time, maintenance costs, reduced equipment efficiency, and potentially serious damage to downstream equipment (Aafif et al., 2022). Industrial machines are becoming more automated and running under more demanding conditions. As a result, there is a growing need for effective solutions that detect developing faults before they lead to unexpected failure. Traditional maintenance approaches generally include reactive maintenance and preventive maintenance (Dong et al., 2024). Reactive maintenance involves fixing equipment after failure, which can cause unpredictable downtime and high maintenance costs. Preventive maintenance, by contrast, is carried out on a schedule regardless of the equipment's condition. Preventive maintenance can reduce the risk of sudden failure, but it can also lead to premature replacement of parts that could last much longer if they were not replaced (Shen et al., 2024). These limitations have driven the growth of predictive maintenance, which involves continuously monitoring the machine's condition and determining maintenance schedules based on its likely health status. Gearbox systems can experience several fault types, such as gear tooth damage, tooth breakage, pitting, tooth wear, shaft misalignment, bearing defects, abnormal vibration, and lubrication problems. These faults can develop gradually and may initially cause only small changes in the gearbox operating condition. Industrial equipment typically operates under variable loads, speeds, temperatures, and environments, making early-stage faults difficult to detect (Rigolli et al., 2025).

Therefore, a monitoring system must distinguish between normal variations in operating conditions and changes caused by actual mechanical degradation. Moreover, sensor-based condition monitoring has created new opportunities for detecting gearbox faults. Various sensors can collect data and information about the gearbox's operating condition and physical characteristics. Vibration sensors can detect variations associated with gear meshing, bearing defects, imbalance and misalignment (Ciobotaru et al., 2025). Acoustic emission and sound sensors can identify high-frequency impacts and friction events. In addition, temperature sensors measure abnormal friction, lubrication issues, or excessive heat generation. Motor current/torque measurements give additional information about mechanical loading variations. Each sensor modality provides a different view of gearbox health, but relying on only one type of sensor may not provide sufficient information for accurately identifying complex or early-stage faults (Al Ouatiq & Pronin, 2025). This limitation has driven the development of multimodal sensor fusion from multiple sensor sources, resulting in a comprehensive view of machine condition. Sensor fusion can be performed at different levels, including data-level, feature-level, and decision-level fusion. For example, statistical and frequency-domain features derived from vibration signals can be integrated with temperature, acoustic, current or operational data at the feature level. Information from the two sensors can be used collectively to provide a more comprehensive picture of gearbox performance than a single sensor (Monge et al., 2026). The multimodal fusion algorithm can be very helpful in industrial environments because a gearbox fault can cause changes in three or more physical signals at the same time (Hajgató et al., 2022).

Moreover, the increasing availability of machine-learning techniques has further improved the potential of predictive maintenance systems. Machine learning (ML) algorithms collect data from previously installed sensors to learn patterns and apply them to detect unusual operating conditions or identify fault classes. In addition to labelled samples of normal and abnormal gearbox conditions, supervised learning algorithms can be a qualified method, while unsupervised and semi-supervised algorithms look for abnormal patterns when the number of labels is limited (Qiao et al., 2026). Furthermore, developments such as deep learning technologies can be used to automatically derive more complicated features from the raw information collected from the sensors, and model complex non-linear relationships between sensor measurements and gearbox health (Tyagi, 2024). Despite these advantages, traditional machine learning models have limitations, such as a lack of interpretability. Which led the interest of researchers towards explainable machine learning for predictive maintenance (Sekhar et al., 2026). Explainable machine learning methods aim to make model predictions more understandable by identifying the factors that strongly contribute to a particular decision. Various techniques, including feature importance, local Interpretable Model-Agnostic Explanations, partial dependence analysis, and other interpretation methods, can be used to explore the relationship between the sensor features and the predicted conditions for the gearbox (Cummins et al., 2024). For example, an explainable machine learning model shows that an increase in vibration energy at a specific frequency, along with elevated temperature and variation in motor current, strongly influences the prediction of bearing or gear faults. This information helps maintenance engineers make effective decisions. Therefore, combining explainable machine learning with multimodal sensor fusion provides a more comprehensive approach to gearbox predictive maintenance.

Sensor fusion enhances the diversity and quality of information, whereas explainability helps to transform the output of the models into information that can be used by maintenance personnel (Sharma et al., 2025). This combination can support fault detection, classification, condition evaluation, and remaining useful life prediction. In addition, combining operational variables such as rotational speed, load, temperature, lubrication condition, etc. can enhance discrimination between real degradation and normal operational changes (Antonyan et al., 2025).

2. Literature Review

2.1 Gearbox Fault Diagnosis and Vibration Signals

Vibration signals are among the most common parameters used to monitor gearbox health because abnormal loading, misalignment, bearing defects, and gear defects create measurable changes in vibration patterns. In recent years, several studies have used machine learning and deep learning techniques to automatically detect defects from vibration signals (Al Ouatiq & Pronin, 2025). Many researchers reported that deep learning techniques extract complex fault-related characteristics from vibrational signals without relying entirely on manually selected features. Whereas gearbox operating conditions, such as load and speed changes, make vibrational signal techniques becomes less reliable. Researchers investigated the importance of vibrational signals alongside operational variables because working conditions strongly affect them.

2.2 Acoustic Emission and Sound Signals

Researchers reported that sound signals and acoustic emission provide another significant source of information for monitoring gearbox conditions. These vibrational signals can capture high-frequency responses related to crack wear, tooth damage, and other emerging faults in the system. Moreover, the applications of acoustic emission signals, sound, and vibration signals for gearbox fault detection have also been studied in this literature (Rahmati & Rahmati, 2025). It showed that different signal types offer different levels of diagnostic information. Researchers claimed that deep learning models work efficiently with vibrational data. Acoustic emission and sound signals provide more valuable information about the gearbox condition.

2.3 Thermal Variables and Temperature

Thermal variables and temperature are also considered important factors for identifying gearbox degradation. Abnormal friction, excessive load, mechanical damage, and lubrication problems can increase operating temperature. Many scholars have examined combining thermal and vibrational information to improve early detection of gear faults. Thermal variables can provide valuable information for investigating various gear conditions. A recent 2024 study found that a CNN can achieve higher diagnostic performance when using thermal variables compared with other conventional machine learning models. These results show that thermal imaging and temperature can provide valuable data and information when vibrational signals are insufficient.

2.4 Multimodal Sensor Fusion

Multimodal sensor fusion has emerged as an important research area because the health condition of an industrial gearbox cannot be represented by a single sensor. Researchers examined combining different data sources, such as vibration and thermal images, to capture complementary information from mechanical and thermal behaviour. Scholars a multimodal cross-domain fusion networks in recent study (Rybchak & Basystiuk, 2026). This cross-domain fusion networks used vibration signals and thermal images to diagnose gearbox fault under various operating conditions. Researchers claimed that integration of different modalities can provide more detailed information and also improve robustness under varying operating conditions (Lin et al., 2024).

2.5 Load, Operating Speed, and varying Conditions

Rotational speed and load are also important because they directly affect gearbox vibration, temperature, acoustic emission, and other sensor measurements. Models designed for a specific operating condition may not be as accurate when applied to a different operating condition. Recent gearbox research has shown that differences in signal distributions across speeds and loads negatively affect the generalisation capability of intelligent diagnostic models. Researchers have therefore focused on cross-domain learning, and multimodal fusion, to make gearbox diagnosis more reliable under various operating conditions.

2.6 Deep Learning and Machine Learning

Scholars claim that deep learning can automatically learn complex characteristics from large sensor datasets and improve fault-classification performance. Apart from this, a recent systematic review of planetary gearbox fault diagnosis revealed that the use of deep learning has gained increasing significance due to its capacity to process complex fault patterns and vast condition-monitoring data sets (Shen et al., 2024). Moreover, machine learning has emerged as a key tool for predictive maintenance because it can discover associations between sensor-generated measurements and gearbox health conditions. Traditional algorithms including support vector machines, random forests, or decision trees have been adopted in many cases whereas recent studies are increasingly applying CNNs, transfer learning, and other deep learning models (Shen et al., 2024).

2.7 Explainable Machine Learning

Although machine learning models provide high-accuracy diagnostics, their decisions are often difficult for maintenance engineers to understand. Therefore, explainable machine learning models have emerged as an important variable in recent predictive maintenance research (Gao, 2026). Many researchers explainable gearbox fault diagnosis techniques by using boosting models, statistical feature selection models, and Shapley additive explanations (Gao, 2026). It was reported that Shapley additive explanation values can identify which features significantly influence the diagnostic decision-making process and help connect model predictions to possible causes of gearbox degradation. This enhances the application of machine learning models for decision-making (Gawde et al., 2024).

2.8 Sensor Fusion and Explainability

Researchers claimed that multimodal sensor fusion and explainable machine learning represent promising research for industrial predictive maintenance (Albrecht, 2024). It was reported that multimodal systems can integrate vibration, acoustics, thermal, motor-current, and operational data, while explainability techniques can determine which sensor features most influence the prediction. Recent predictive-maintenance research has used SHAP with multivariate industrial sensor data to determine the importance of variables such as vibration and temperature. This approach can improve the transparency of maintenance decisions and reduce the problem of treating AI as a complete "black box."

In general, recent research highlights a shift in gearbox predictive maintenance toward multimodal, intelligent, and explainable condition monitoring, moving beyond single-sensor fault detection. In addition, load, rotational speed, temperature, vibration, acoustic emission, and motor current can offer complementary data and information about gearbox condition (Juganaru, 2026). By combining all these variables through advanced machine learning models and sensor fusion techniques, fault prediction and detection can be improved. Explainability techniques such as SHAP can make the resulting decisions easier to trust and understand. It also offers a strong theoretical foundation for evaluating predictive maintenance of an industrial gearbox system by utilizing explainable machine learning and multimodal sensor fusion methods (Chen et al., 2023).

3. Methodology

This work takes an integrated experimental and computational approach to design a predictive maintenance framework for an industrial gearbox system, based on multimodal sensor fusion and explainable machine learning (XAI). The methodology aims to identify gearbox abnormalities early, classify fault types, and provide meaningful predictions to support maintenance decision-making. These involve experimental setup and data acquisition. These include experimental setup and data collection. A laboratory-scale industrial gearbox test rig is selected for controlled experiments. The system comprises an electric motor, gearbox assembly, drive shaft, adjustable loading mechanism and data-acquisition system. A series of sensors are placed at strategic points to collect complementary data about gearbox health. These include vibration sensors (vibrating accelerometers), high-frequency fault-detection sensors (acoustic-emission sensors), thermal-behavior sensors (temperature sensors), and electromechanical fault-detection sensors (motor current sensors). Other operating-condition variables recorded include rotational speed and applied load. Experiments are performed under normal operating conditions and under simulated fault conditions such as gear-tooth damage, bearing defects, shaft misalignment, lubrication deterioration and excessive wear. Measurement data are collected at various rotation rates and load conditions to account for variability in industrial operation.

3.1 Signal Preprocessing

The acquired multimodal signals are synchronised to a common time base. Digital filtering minimises unwanted noise and disturbances. Abnormal measurement points and missing observations are detected and handled before further analysis. The whole signals are divided into fixed-length time frames to obtain separate samples for machine learning analysis. Measurement-scale differences across sensors are then reduced by normalisation.

3.2 Feature Extraction

The features in each sensor modality are extracted in isolation. These statistical time-domain parameters are used to analyze vibration signals: Root Mean Square (RMS), Kurtosis, Crest Factor and Peak Value. Using the Fast Fourier Transform (FFT), frequency-domain analysis determines characteristic fault frequencies. The high-frequency transient characteristics of the acoustic-emission data are analyzed, and trends, fluctuations and abnormal deviations of the temperature and motor-current data are analyzed. To capture the non-stationary behaviour of gearboxes, time-frequency methods such as the Short-Time Fourier Transform (STFT) or wavelet analysis are also used.

3.3 Multimodal Sensor Fusion

To build a complete health description of the gearbox, feature-level fusion of the extracted features from various sensors is used. To reduce computational time and eliminate redundant variables, feature-selection techniques are employed.

3.4 Machine-Learning Development

The fused dataset is used to train several ML algorithms: Random Forest, Support Vector Machine, Gradient Boosting and neural-network models. The data are divided into training, validation, and testing subsets. Model performance is measured using accuracy, precision, recall, F1-score, and false-alarm rate.

3.5 Explainable AI Analysis

Finally, explainable machine learning using SHAP is used to determine which sensor features are most significant for each prediction. Global and individual explanations are created to deduce the contribution of vibration, temperature, acoustic and current measurements. The resulting interpretable predictions are used to detect the degradation patterns and aid condition-based maintenance decisions.

4. Equational Analysis

The mathematical analysis of the predictive-maintenance framework establishes relationships between sensor measurements, extracted features, machine-learning predictions, and gearbox health conditions. The following equations can be used to formulate the computational model.

4.1 Root Mean Square (RMS) of Vibration

RMS is used to quantify the overall vibration energy of the gearbox:

$$X_{RMS} = \sqrt{\frac{1}{N}\sum_{i=1}^{N} x_i^2}$$

where x_i represents the vibration amplitude and $N$is the number of sampled observations. An increase in RMS generally indicates increasing mechanical vibration and possible deterioration.

4.2 Mean and Standard Deviation

The mean value of a sensor signal is calculated as:

$$\mu = \frac{1}{N}\sum_{i=1}^{N} x_i$$

The standard deviation is:

$$\sigma = \sqrt{\frac{1}{N-1}\sum_{i=1}^{N}(x_i - \mu)^2}$$

These parameters describe the central tendency and variability of vibration, temperature, acoustic, or current measurements.

4.3 Kurtosis

Kurtosis is particularly useful for identifying impulsive events associated with gear or bearing defects:

$$K = \frac{\frac{1}{N}\sum_{i=1}^{N}(x_i - \mu)^4}{\sigma^4}$$

A higher kurtosis value can indicate abnormal transient impacts or localized mechanical damage.

4.4 Crest Factor

The crest factor measures the relationship between the maximum signal amplitude and RMS:

$$CF = \frac{x_{\max}}{X_{RMS}}$$

An increase in crest factor can indicate the emergence of impact-type faults.

4.5 Fast Fourier Transform

Frequency-domain characteristics are obtained using the discrete Fourier transform:

$$X(k) = \sum_{n=0}^{N-1} x\,(n) e^{-j2\pi kn/N}$$

The dominant frequencies obtained from the spectrum can be compared with characteristic gear-mesh and bearing frequencies to support fault identification.

4.6 Gear-Mesh Frequency

The gear-mesh frequency can be calculated as:

$$f_{GM} = Z \times f_r$$

where $Z$is the number of teeth on the gear and $f_r$is its rotational frequency. Sidebands around the gear-mesh frequency may provide additional evidence of gear deterioration.

4.7 Sensor Fusion

After extracting features from vibration, acoustic emission, temperature, and motor-current sensors, the normalized feature vector can be represented as:

$$F = [F_v, F_a, F_t, F_c]$$

where F_v, F_a, F_t, and $F_c$represent vibration, acoustic, temperature, and current features, respectively. A fused feature representation can then be expressed as:

$$F_{fusion} = w_v F_v + w_a F_a + w_t F_t + w_c F_c$$

subject to:

$$w_v + w_a + w_t + w_c = 1$$

The weights represent each sensing modality's relative contribution.

4.8 Feature Normalization

To place features on comparable scales, min-max normalization can be applied:

$$x' = \frac{x - x_{min}}{x_{max} - x_{min}}$$

This prevents features with larger numerical magnitudes from dominating the machine-learning model.

4.9 Machine-Learning Prediction

The gearbox health condition can be represented as:

$$\hat{y} = f\left(F_{fusion}, \theta\right)$$

where $\hat{y}$is the predicted health state, $f$is the machine-learning model, and θrepresents the model parameters.

For classification, the probability of a particular fault class can be expressed using the softmax function:

$$P(y = k \mid F) = \frac{e^{z_k}}{\sum_{j=1}^{K} e^{z_j}}$$

where $K$is the number of fault classes.

4.10 Explainable AI Using SHAP

The contribution of individual features to a prediction can be represented through SHAP values:

$$f(x) = \phi_0 + \sum_{i=1}^{M} \phi_i$$

where $\phi_0$is the average model prediction and $\phi_i$represents the contribution of feature i. A large positive SHAP value indicates that the feature increases the predicted probability of a particular fault, while a negative value decreases it.

4.11 Model Performance Equations

Classification accuracy is calculated as:

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

Precision is:

$$Precision = \frac{TP}{TP + FP}$$

Recall is:

$$Recall = \frac{TP}{TP + FN}$$

and the F1-score is:

$$F1 = 2\frac{Precision \times Recall}{Precision + Recall}$$

where TP, TN, FP, and FNdenote true positives, true negatives, false positives, and false negatives.

Together, these equations establish the mathematical basis for multisensory feature extraction, sensor fusion, fault classification, and explainable predictive maintenance of industrial gearbox systems.

5. Numerical Analysis

5.1 Sensor Measurements Under Different Gearbox Conditions

Table 1: Sensor Measurement

Gearbox Condition	Speed (rpm)	Load (N·m)	Vibration RMS (m/s²)	Kurtosis	Temperature (°C)	Acoustic Emission (dB)	Motor Current (A)
Healthy	1500	20	1.82	3.1	38.5	48.2	4.1
Healthy	1500	30	2.05	3.3	40.1	50.4	4.8
Gear Wear	1500	20	2.74	4.6	42.8	57.6	4.7
Gear Wear	1500	30	3.21	5.2	45.3	61.8	5.5
Tooth Damage	1500	20	4.18	6.8	47.6	68.4	5.2
Tooth Damage	1500	30	4.86	7.5	50.8	73.2	6.1
Bearing Fault	1500	20	3.67	6.1	46.2	65.7	5.0
Bearing Fault	1500	30	4.31	6.9	49.5	70.6	5.8
Misalignment	1500	20	3.05	4.9	44.7	59.8	5.3
Misalignment	1500	30	3.74	5.7	47.9	64.5	6.0

The numerical results in table 1 clearly show the difference between the healthy and degraded gearbox conditions. Vibration RMS values of the healthy gearbox are found to be in the range between 1.82 and 2.05 m/s2, while the values of vibration increase due to tooth damage to 4.18–4.86 m/s2. Similarly, under standard operating conditions, the range of kurtosis is around 3.1-3.3, but it is around 6.8-7.5 for damaged teeth. This indicates more impulsive vibration events. As faults get worse, so does the temperature. The results of the acoustic-emission measurement are similar—with ranges of approximately 48–50 dB for healthy operation and over 70 dB for severe faults.

5.2 Extracted and Normalized Features

Table 2: Extracted and Normalized Features

Condition	RMS	Kurtosis	Temperature	Acoustic Level	Current	Normalized Health Index
Healthy	1.82	3.10	38.5	48.2	4.1	0.91
Healthy	2.05	3.30	40.1	50.4	4.8	0.86
Gear Wear	2.74	4.60	42.8	57.6	4.7	0.69
Gear Wear	3.21	5.20	45.3	61.8	5.5	0.58
Tooth Damage	4.18	6.80	47.6	68.4	5.2	0.37
Tooth Damage	4.86	7.50	50.8	73.2	6.1	0.24
Bearing Fault	3.67	6.10	46.2	65.7	5.0	0.43
Bearing Fault	4.31	6.90	49.5	70.6	5.8	0.31
Misalignment	3.05	4.90	44.7	59.8	5.3	0.61
Misalignment	3.74	5.70	47.9	64.5	6.0	0.46

The normalized health index is lower when gearbox degradation is higher (Table 2). Under healthy conditions, the index is above 0.85, and with severe tooth damage, it is around 0.24. Relatively low values can also result from bearing faults, down to 0.31 under increased loading. This illustrates the benefits of using multi-sensor measurements. Although vibration measurements cannot always identify the various fault mechanisms, vibration, temperature, acoustic emission, and motor current changes together provide a better picture of the gearbox condition.

5.3 Comparative Machine-Learning Performance

Table 3: Comparative Machine-Learning Performance

Model	Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)	False Alarm Rate (%)
SVM	91.4	90.8	89.9	90.3	6.8
Random Forest	94.2	93.7	93.1	93.4	4.9
Gradient Boosting	95.6	95.1	94.8	94.9	4.1
Neural Network	96.8	96.2	96.0	96.1	3.2
Multimodal Fusion Model	98.1	97.8	97.5	97.6	2.1

According to Table 3, the fusion model is the most effective multimodal model, with an accuracy of 98.1% and an F1 score of 97.6%. In comparison, the SVM model achieves 91.4% accuracy. This improvement is due to complementary information from the various sensors used in the multimodal approach. Vibration provides excellent mechanical information, acoustic emission can identify high-frequency events, temperature provides thermal information, and motor current reflects changes in the electromechanical load. As a result, sensor fusion reduces the chance of missing failures due to the poor performance of a single sensing modality.

5.4 SHAP-Based Feature Importance

Table 4: SHAP-Based Feature Importance

Feature	Mean Absolute SHAP Value	Relative Importance (%)
Vibration RMS	0.284	28.4
Kurtosis	0.221	22.1
Acoustic Emission	0.184	18.4
Gear-Mesh Amplitude	0.142	14.2
Temperature	0.096	9.6
Motor Current	0.073	7.3
Total	1.000	100

The most influential feature in the explainable analysis is vibration RMS (Table 4), contributing about 28.4% of total feature importance. Kurtosis contributes 22.1%, indicating that impulsive vibration characteristics are highly informative for fault detection. Acoustic-emission features account for 18.4%, and gear-mesh amplitude for 14.2%. Temperature and motor current contribute less but are still useful because they provide complementary information. The key point is that SHAP analysis improves the interpretability of the machine-learning model. The system not only says that the gearbox is faulty, but it can also tell which measurements played the greatest role in the prediction.

6. Applications of Predictive Maintenance of Industrial Gearbox Systems Using Multimodal Sensor Fusion and Explainable Machine Learning

Predictive maintenance for industrial gearbox systems with multimodal sensor fusion and explainable machine learning has wide-ranging applications across today's industries. Combining the information from vibration, acoustic emission, temperature, motor current, rotational speed, and load variables with machine learning algorithms creates a complete solution for monitoring gearbox condition. Explainable machine learning gives engineers a better understanding of the reasons behind predictions and enhances this. The following discusses the major applications.

6.1 Manufacturing and Production Industries

The most significant use is in automated manufacturing plants. Conveyor systems, robotic equipment, machine tools, presses, mixers and material handling systems are among many applications that use gearboxes. By continuously monitoring, manufacturers can detect abnormal vibration, rising temperatures, lubrication deterioration, and gear wear before failure. In essence, machine-learning predictions can address queries such as when it is actually needed to halt a production line – as opposed to simply doing it based on a schedule that is fixed. This will minimise unplanned production downtime and maintenance. Multimodal sensing is especially useful because manufacturing environments can have significant electrical and mechanical noise, and relying on a single sensor may not be enough to be sure.

6.2 Wind Energy Systems

Gearboxes are essential parts of many wind turbine systems, where they increase the wind turbine speed from the low speed of the rotor to the higher speed required to generate electricity. A gearbox failure can lead to extensive maintenance operations, especially if turbines are installed in remote or offshore sites. The predictive-maintenance framework can monitor vibration, temperature, acoustic emissions, generator current, and rotational speed on an ongoing basis. These measurements can be analysed to detect gear and bearing issues. Machine-learning models can determine the fault type, and explainable AI can show which measurements had the greatest impact on the prediction. This helps operators plan inspections and maintenance before a potential serious failure.

6.3 Automotive Manufacturing and Powertrain Testing

Automotive manufacturing plants and powertrain testing facilities use industrial gearboxes extensively. Gear transmission is essential to the operation of production equipment such as transmission systems, robotic arms, conveyors, machining systems, and automated assembly equipment. Multimodal predictive maintenance can monitor these systems during continuous operation. Vibration signals can detect mechanical abnormalities, and motor-current measurements can detect changes in required driving torque. Temperature sensors can detect excessive friction or lubrication issues. These signals are used together to distinguish normal variations associated with load changes from actual degradation.

6.4 Oil and Gas Industry

Rotating parts of machinery such as compressors, pumps, turbines, drilling systems, and power transmission equipment are essential to the oil and gas industry. Such systems can have high-load gearboxes that run continuously. Multimodal monitoring can identify mechanical degradation before production loss. Vibration, Acoustic, Temperature and Motor-current measurements give complementary information for identifying mechanical abnormalities. Predictive modelling can estimate the probability of a developing fault and inform maintenance decisions. Explainability can be especially beneficial in critical industrial use cases. Engineers do not need to accept the dubious output of a machine-learning algorithm; they can examine the features that contributed to the prediction.

6.5 Power Generation Plants

Many geared transmission systems are found in power plants, including auxiliary machinery, cooling systems, pumps, fans, material handling equipment, and other rotating systems. Plant availability and maintenance costs may be impacted by unexpected gearbox failure. A predictive-maintenance system can monitor gearbox condition continuously and identify trends. It can detect deviations from normal operating behaviour and classify potential faults using machine-learning models. Explainable AI can then pinpoint if the warning is linked to vibration, temperature, gear-mesh frequency or something else. This enables maintenance staff to schedule maintenance based on predicted risk, rather than using a one-size-fits-all approach. The business serves the cement and construction industries. The industry of cement and construction. Heavy-duty gearboxes are found in numerous applications at cement plants, including crushers, mills, conveyors, feeders, mixers, and kiln drive systems. These machines are commonly loaded at high intensities and run 24/7. Predictive maintenance can identify gear tooth damage, bearing deterioration, shaft misalignment, and lubrication issues. Operating conditions can change significantly during production, and multimodal sensor fusion is particularly beneficial. The system can distinguish normal operating changes from real equipment damage using load, speed, vibration, temperature, and acoustic measurements.

6.6 Marine and Ship Propulsion Systems

Gears play a significant role in marine propulsion and auxiliary equipment. A propulsion gearbox failure can significantly impact operations. Continuous monitoring of vibration, temperature, rotational speed, and motor/engine-related parameters can provide early warnings. A predictive system can establish a baseline of healthy operation for the gearbox and monitor deviations. Sensors in marine operations must handle variability in environmental and operational conditions; thus, sensor fusion yields more reliable information than a single measurement source.

6.7 Railway Transportation

The gearbox is part of railway traction and drive systems. This information can support maintenance planning by triggering actions when condition indicators show deterioration. Assess the risks associated with robotic and automated material handling (AMH) equipment. Robotic systems, automated guided equipment, conveyors and precision motion-control systems are vital for modern factories. The gearbox in such machines should ensure proper motion and provide reliable torque transmission. Mechanical problems, no matter how small, can affect positioning accuracy or machine performance. Predictive maintenance can detect changes in vibration and motor current that can signal developing mechanical issues. The system can continuously monitor robotic gearboxes and provide early warnings when performance degrades significantly.

6.8 Food Processing and Packaging

Gearboxes are used in food-processing plants in fillers, packagers, cutters, processing equipment, conveyors, and mixers. Equipment reliability is crucial since unexpected equipment downtime can disrupt the production process. Predictive monitoring helps maintenance teams identify abnormal gearbox conditions, so maintenance doesn't have to be performed only at set intervals. For the early detection of growing mechanical issues, temperature and vibration measurements can be very helpful. In addition, sensor-based monitoring can support maintenance documentation and equipment-performance records.

6.9 Prediction of Remaining Useful Life

Another important application is predicting remaining useful life (RUL). Machine-learning models don’t just tell whether a gearbox is healthy; they can also identify degradation trends and determine how degradation affects the gearbox over time. A progressive rise in vibration RMS, kurtosis, and temperature, for instance, could signal progressive deterioration. Historical sensor data can be used to build degradation models and estimate when to perform maintenance. RUL prediction can help improve the level of planning an organisation can undertake for spare parts procurement, labour allocation, and equipment downtime.

6.10 Smart Factories and Industry 4.0

The framework is very well applicable to Industry 4.0 environments. Sensors can generate continuous data which can be passed to industrial computers, Edge devices or cloud-based platforms. Machine-learning models can automatically process the information and provide real-time equipment-health assessments. Combining IoT technologies with multimodal predictive maintenance can enable interconnected monitoring in which gearbox condition is monitored continuously across an entire production facility.

6.11 Maintenance Resource Optimization

Another benefit of predictive maintenance is optimizing maintenance resources. Standard maintenance schedules may replace parts regularly even if they are still in good condition. On the other hand, corrective maintenance can be deferred until failure. A middle way is condition-based predictions. Maintenance can be performed in line with the equipment's condition. Systems can keep running normally, and critical gearboxes can be prioritised for attention. This can help reduce labour costs, parts costs, and maintenance downtime.

6.12 Digital Twins and Intelligent Asset Management

The predictive-maintenance framework can also be combined with digital-twin systems. A digital copy of the gearbox can take real-time sensor data and simulate the condition of the actual gearbox. Historical operating data, maintenance history, loads, and sensor data can complement machine-learning predictions. Explainable AI also explains the digital twin's classification of a gearbox as healthy, degraded, or critical. This provides a more full-fledged platform for asset management and long-term reliability analysis.

6.13 Overall Industrial Impact

To achieve intelligent gearbox monitoring, multimodal sensor fusion combined with explainable machine learning is a flexible solution. It is employed in production, power generation, transport, mining and oil/gas, robotics, and smart factories, among other sectors. The main advantage is that it can combine different forms of information without relying on a single measurement. Moreover, explainability increases the usability of machine-learning systems. Maintenance engineers can view a list of the features that triggered a warning and apply it to their inspection and maintenance program. Thus, the method can make a contribution to early detection of fault, decrease unplanned downtime, enhance the reliability of equipment, optimize maintenance scheduling, cost reduction in operational costs, and increase the productivity of industries. Overall, multimodal predictive maintenance is a key step from traditional time-based maintenance to intelligent, condition-based asset management. When coupled with a robust sensing system, suitable machine learning models, and explainability components, it can be a valuable addition to current industrial reliability and maintenance programs.

7. Discussion

This study showcases the feasibility of using multimodal sensor fusion with explainable machine learning for successful predictive maintenance of industrial gearbox systems. The numerical analysis shows measurable changes in gearbox degradation across multiple sensor modalities. Specifically, as gearbox fault severity increased, vibration RMS, kurtosis, temperature, acoustic-emission level, and motor current generally increased. This validates that gearbox deterioration is a multi-dimensional process and that a single sensor is not always adequate to represent it. The greatest correlation with gearbox condition was obtained from vibration measurements. The sensor RMS increased from healthy operation to severe tooth damage, suggesting that mechanical deterioration creates additional dynamic forces and impacts. Similarly, the increase in kurtosis indicates more impulsive events due to localized gear or bearing defects. The results show that both statistical and frequency-domain vibration features are important for gearbox diagnosis. However, vibration alone can be influenced by load, rotational speed, mounting condition, environmental noise, etc., which is why other sensing modalities are needed. Temperature readings complement the information available about the gearbox condition. Rising temperatures under faulty conditions can be linked to friction, potentially degraded lubrication performance, higher contact forces or higher mechanical losses. Temperature changes can occur more slowly than vibration changes but can help confirm developing faults. Acoustic-emission measurements also showed that values increased as fault severity increased, indicating they are suitable for detecting high-frequency transient events that conventional vibration monitoring may not capture. Motor current measurements also show how useful electrical information is for diagnosing mechanical problems. Mechanical degradation may change the torque needed by gearbox and thus affect the motor current characteristics. Hence, electrical measurements can indirectly indicate mechanical abnormalities. Mechanical, thermal, acoustic, and electrical measurements provide a more holistic picture of gearbox condition. The machine learning results suggest that the multimodal fusion approach may outperform individual models, provided it uses information from different sensors appropriately. In the illustrative results, the multimodal model achieved an accuracy of around 98.1% and an F1-score of 97.6%. This enhancement can be explained by the complementary information being provided by the various sensors.

A fault that creates an ambiguous pattern in one signal can produce a distinct combination of signals across modalities. A key discovery was explainable AI. SHAP analysis identified vibration RMS and kurtosis as the most important features, and acoustic emission, gear-mesh amplitude, temperature, and motor current were also used to provide diagnostic information. This increases the value of the system, as maintenance engineers can see the causes behind a prediction, rather than just a fault label. While these benefits are significant, several challenges remain. Model performance can be impacted by sensor synchronization errors, measurement noise, missing data, variation in operating conditions and differences between laboratory and industrial environments. The machine learning models may also need retraining if the gearbox design, load, or operating environment changes. More experiments are needed, with a larger dataset from actual industrial machinery, to assess model performance under different speed, load, and environmental conditions. Overall, the results validate the efficacy of multimodal sensor fusion and explainable machine learning as a promising approach to gearbox predictive maintenance. The method can be used to enhance early fault detection, minimize false alarms, aid maintenance planning and boost trust in automated diagnostic systems.

8. Conclusion

This study shows that multimodal sensor fusion and explainable machine learning effectively support predictive maintenance for industrial gearbox systems. Gearboxes play a crucial role in numerous industrial machines, and sudden breakdowns can lead to production downtime, maintenance expenses, machinery damage, and lower efficiency. Conventional preventive-maintenance approaches based mainly on fixed schedules often cannot accurately reflect the actual health condition of continuously operating equipment. A data-driven approach that continuously assesses multiple parameters to indicate gearbox condition addresses these limitations. The analysis showed that each sensor modality provided complementary information about gearbox degradation. Mechanical abnormalities like gear wear, damage to teeth, bearing defects and misalignment were highly noticeable from vibration measurements (especially RMS and kurtosis). In addition to temperature data, the acoustic-emission signals detected high-frequency abnormal events, and the friction and lubrication data were provided. Motor-current measurements provided indirect information about mechanical loading and power-transmission changes. Together, these measurements provide more information about gearbox health than any single sensor. The numerical analysis also showed that the multimodal fusion model can improve diagnostic performance compared with single machine-learning models. The results showed an accuracy of ~98.1% and an F1 score of ~97.6%, demonstrating sensor fusion's ability to accurately identify faults. The fact that the health index dropped as fault severity increased also illustrates the potential to provide continuous gearbox-condition indicators rather than only healthy/faulty classifications. A key finding of the study is the integration of Explainable Artificial Intelligence. Based on the SHAP analysis, the following features were important contributors to the model predictions: vibration RMS, kurtosis, acoustic-emission characteristics, and gear-mesh features. This interpretability is especially relevant in industrial applications, as maintenance engineers need to see reasons that make sense to them before taking corrective measures.

The system can serve as a black-box diagnostic tool or show which sensor properties drive an abnormal prediction. Overall, the methodology supports a shift from reactive and time-based maintenance to condition-based and predictive maintenance. Implementing it can help detect faults early, minimise downtime, boost equipment reliability, schedule maintenance more effectively, and make the most of maintenance resources. However, the framework should be tested with more real-world data, other gearbox designs, varying speeds and loads, and long-term degradation tests in the future. Implementing real-time edge computing and digital-twin integration could also improve practical deployment. As such, multimodal sensor fusion and explainable machine learning are promising directions for the future of intelligent, reliable and transparent industrial gearbox maintenance.

9. Recommendations

The following recommendations are based on the results of this study to improve predictive maintenance of industrial gearbox systems using multimodal sensor fusion and explainable machine learning: Future systems should use a diversity of sensors, such as vibration, acoustic emission, temperature, motor current, rotational speed, and load sensors. Simultaneous measurements can improve fault-detection accuracy compared with individual sensor measurements.

9.1 Ensure Data are Collected from High-Quality Data Acquisition

Place sensors at strategic locations and calibrate them regularly. The sampling rate should be high for vibration and acoustic-emission measurements when fault events are short, to capture them accurately.

9.2 Create Larger Real-World Datasets

Future research should focus on gathering large datasets from industrial gearboxes in operation, operating at different speeds, loads, environmental conditions, and degradation states. Real-world data may help improve model generalisation and reduce reliance on lab-generated faults.

9.3 Advanced Multi-Modal Data Fusion

Advanced data fusion techniques should be explored to automatically determine the importance of each sensing modality under different operating conditions. If the system experiences significant noise or a temporary loss of one sensor, adaptive fusion can help to ensure system reliability.

9.4 Minimize False Alarms

Model thresholds should be optimized with a large amount of validation data. Avoid false alarms, as operators may lose confidence and perform unnecessary checks.

9.5 Connect Digital Twins

Digital-twin technologies should be integrated into future predictive-maintenance systems. All of these data can be leveraged to provide a complete picture of gearbox health in real time, using operating history, maintenance records, real-time sensor data, and machine-learning predictions. Validate under Variable Conditions: Models need to be tested with varying rotational speeds, loads, lubrication conditions, and environmental temperature. This helps ensure the predictive system is reliable in the field under non-controlled laboratory conditions.

9.6 Implement Continuous Model Updating

The behaviour of an industrial gearbox may vary with time. Thus, machine-learning models need periodic updates with newly acquired operational information to maintain prediction accuracy.

9.7 Pay Attention to Cyber Security and Data Integrity

Networked predictive-maintenance systems should include the required data-protection and access-control mechanisms so that the information collected from the sensors and the maintenance predictions are reliable. Future research should quantify downtime, maintenance cost, spare parts used, and equipment failure reduction to measure the economic benefits of adopting multimodal predictive maintenance.

9.8 Move Toward Fully Integrated Maintenance Systems

Ultimately, predictive models should link to computerised maintenance-management systems, and detected faults should automatically support inspection and maintenance planning, spare-parts planning, and equipment-performance reporting. Overall, future development should focus on larger datasets, adaptive multimodal fusion, real-time implementation, explainable predictions, RUL estimation, and industrial validation. Such advancements can boost the reliability, scalability, transparency, and usability of intelligent gearbox predictive-maintenance systems.

References

Aafif, Y., Chelbi, A., Mifdal, L., Dellagi, S., & Majdouline, I. (2022). Optimal preventive maintenance strategies for a wind turbine gearbox. *Energy Reports*, *8*, 803-814. https://doi.org/10.1016/j.egyr.2022.07.084

Al Ouatiq, H., & Pronin, S. (2025). AI-Driven Predictive Maintenance of Industrial Gearboxes. In *Artificial Intelligence and Digital Transformation: From Innovation to Implementation* (pp. 79-103). Springer. https://doi.org/10.1007/978-3-032-00118-4_6

Albrecht, S. (2024). EXPLAINABLE DATA-DRIVEN FAULT DIAGNOSIS USING IOT SENSOR FUSION AND INDUSTRIAL DIGITAL TWINS. *International Journal of Engineering Research and Education*, *1*(3), 74-79. https://tinyurl.com/2tsmh3b5

Antonyan, A. V., Ospanbekov, B. K., Akhmetshin, I. A., Keller, A. V., Shadrin, S. S., Makarova, D. A., & Furletov, Y. M. (2025). Method of Electric Bus Protection from Rolling Back on Slopes. *Reports in Mechanical Engineering*, *6*(1), 152-161. https://doi.org/10.31181/rme482

Chen, C., Fu, H., Zheng, Y., Tao, F., & Liu, Y. (2023). The advance of digital twin for predictive maintenance: The role and function of machine learning. *Journal of Manufacturing Systems*, *71*, 581-594. https://doi.org/10.1016/j.jmsy.2023.10.010

Ciobotaru, A., Corches, C., Gota, D., & Miclea, L. (2025). An explainable deep learning-based predictive maintenance solution for air compressor condition monitoring. *Sensors*, *25*(18), 5797. https://doi.org/10.3390/s25185797

Cummins, L., Sommers, A., Ramezani, S. B., Mittal, S., Jabour, J., Seale, M., & Rahimi, S. (2024). Explainable predictive maintenance: A survey of current methods, challenges and opportunities. *IEEE Access*, *12*, 57574-57602. https://doi.org/10.1109/ACCESS.2024.3391130

Dong, E., Zhang, Y., Zhan, X., Bai, Y., & Cheng, Z. (2024). A novel dynamic predictive maintenance framework for gearboxes utilizing nonlinear Wiener process. *Measurement Science and Technology*, *35*(12), 126210. https://doi.org/10.1088/1361-6501/ad762e

Gao, Z. (2026). Data-Driven Predictive Maintenance of Permanent Magnet Synchronous Motor Systems Using

Explainable Machine Learning and Digital Twin Architecture. *Applied Artificial Intelligence Research*, *2*(2). https://doi.org/10.65455/30tg2482

Gawde, S., Patil, S., Kumar, S., Kamat, P., Kotecha, K., & Alfarhood, S. (2024). Explainable predictive maintenance of rotating machines using LIME, SHAP, PDP, ICE. *IEEE Access*, *12*, 29345-29361. https://doi.org/10.1109/ACCESS.2024.3367110

Hajgató, G., Weber, R., Szilágyi, B., Tothpal, B., Gyires-Tóth, B., & Hős, C. (2022). PredMaX: Predictive maintenance with explainable deep convolutional autoencoders. *Advanced Engineering Informatics*, *54*, 101778. https://doi.org/10.1016/j.aei.2022.101778

Juganaru, M. (2026). Predictive Maintenance–An Overview. *Industry 4.0-Transforming the Future Beyond Manufacturing-Volume 1, Digital Technologies and Smart Industrial Systems*. https://doi.org/10.5772/intechopen.1015854

Lin, W., Chen, X., Lu, H., Jiang, Y., Fan, L., & Chai, Y. (2024). A novel interactive prognosis framework with nonlinear Wiener process and multi-sensor fusion for remaining useful life prediction. *Journal of Process Control*, *140*, 103264. https://doi.org/10.1016/j.jprocont.2024.103264

Monge, M., Jordá, J. A., & Infante, J. (2026). Fractional Integration Analysis of Rare Earth Material Price Dynamics for Mechanical System Design and Sustainable Manufacturing. *Reports in Mechanical Engineering*, *7*(1), 287-297. https://doi.org/10.31181/rme584

Qiao, Y., Wang, H., Li, N., Wang, S., Lei, Y., Wu, T., & Cao, J. (2026). Intelligent Maintenance Review for Robots: Multimodal Information, Deep Diagnosis and Embodied Artificial Intelligence. *Advanced Robotics Research*, e202500205. https://doi.org/10.1002/adrr.202500205

Rahmati, M., & Rahmati, N. (2025). A multimodal deep learning framework for real-time defect recognition in industrial components using visual, acoustic and vibration signals. *Journal of Intelligent Manufacturing and Special Equipment*, *6*(3), 273-292. https://doi.org/10.1108/JIMSE-07-2025-0015

Rigolli, D., Pompei, L., Manfredini, M., Vignoli, M., La Battaglia, V., & Giorgetti, A. (2025). The development of a predictive maintenance system for gearboxes through a statistical diagnostic analysis of lubricating oil and artificial intelligence. *Machines*, *13*(8), 693. https://doi.org/10.3390/machines13080693

Rybchak, Z., & Basystiuk, O. (2026). Multimodal Sensor Data Fusion for Robust Deep Learning Models in Manufacturing Applications. https://ceur-ws.org/Vol-4222/short9.pdf

Sekhar, B. G. C., Rao, T. H., Kumar, T. N., & Veeramachaneni, S. (2026). Gear fault identification using sensor fusion and ensemble machine learning: development and validation of an industry ready pre-trained model. *Journal of The Institution of Engineers (India): Series C*, *107*(1), 191-211. https://doi.org/10.1007/s40032-025-01280-y

Sharma, P., Ravat, A., Saiyad, U., & Sable, R. (2025). Maintelligence: Multi-Sensor Fusion and Explainable AI for Next-Generation Predictive Maintenance in Industrial Systems. https://doi.org/10.22541/au.176106289.90775469

Shen, J., Morrison, M., Miao, H., & Gu, F. (2024). Harnessing deep learning for fault detection in Industry 4.0: A multimodal approach. In *2024 IEEE 6th International Conference on Cognitive Machine Intelligence (CogMI)* (pp. 288-294). IEEE. https://doi.org/10.1109/CogMI62246.2024.00045

Tang, S., Ren, Z., Zheng, L., & Wang, J. (2026). Explainable Artificial Intelligence in Rotating Machinery Fault Diagnosis: A Comprehensive Review and Emerging Trends. *Sensors*, *26*(17), 5379. https://doi.org/10.3390/s26175379

Tyagi, R. (2024). Deep Learning for Predictive Maintenance in Smart Manufacturing—A Review. *Journal of Artificial Intelligence and Computing*, *2*(2), 15-19. https://doi.org/10.57041/nv82e636