STM JOURNALS

Trends in
Machine Design

ISSN: 2455-3352
Volume 13, Issue 1, 2026
January–April
DOI (Journal): 10.37591/TMD

https://journals.stmjournals.com/tmd

Research TMD

AI-Driven Predictive Maintenance Framework for Intelligent Vehicle Health Monitoring

Chandani S. Bartakke[1,*], Patil Sarang Maruti[2], Mukesh Kumar Gupta[3], Amit Tiwari[4]

Abstract

The accelerated development of smart and connected car systems made the necessity to find the accurate and real-time predictive maintenance solutions which would minimize the number of unexpected failures as well as increase the cars on-road safety. The current paper proposes an artificial intelligence-based hybrid predictive maintenance system that combines Long Short-Memory (LSTM) networks and the XGBoost predictor to provide a potent vehicle fault diagnosis, Remaining Useful Life (RUL) prediction, and automatic maintenance warning system. The suggested approach involves the usage of the multisensor vehicle data, such as engine temperature, battery voltage, vibration, and electrical measurements to simulate the degradation trends and detect the early failure signs. As the results of the experiment prove, the hybrid model performs significantly better than the traditional machine learning methods, including Logistic Regression (LR), Random Forest (RF), and Gradient Boosting (XGBoost). Using the system, prediction accuracy of the fault-type is as follows: 97.3%, precision 97.0%, recall 97.1%, and F1-score 97.0%. To estimate RUL, Hybrid LSTM -XGBoost model is the least erroneous with a MAE of 8.3 hours and an RMSE of 12.8 hours which means that it is highly reliable in estimating maintenance. The analysis of confusion matrix and ROC curve de facto substantiate high discriminative ability of different types of faults with AUC. The performance analysis in terms of inference reveals that the prediction time of the model is still appropriate in the areas of real-time implementation within the edge-based vehicle health monitoring system. The results confirm that the suggested hybrid predictive maintenance structure is beneficial in the sense that it improves the quality of diagnostic results, reliability, and gives interpretable results that can be used to make proactive decisions. This work has advanced to the intelligent and data-driven maintenance ecosystem to enhance vehicle safety, minimize downtime, and reduce costs of operation.

Keywords: Predictive Maintenance, Vehicle Health Monitoring, Machine Learning Algorithms, Fault Diagnosis, Remaining Useful Life Prediction

***Author for Correspondence**
Chandani S. Bartakke
E-mail: chandanibartakke@gmail.com

[1]Research Scholar, Department of Computer Science and Engineering, Suresh Gyan Vihar University, Jaipur, Rajasthan, India
[2]Professor, Department of Computer Science and Engineering, Amity University, Mumbai, Maharashtra, India
[3]Professor, Department of Electrical Engineering, Suresh Gyan Vihar University, Jaipur, Rajasthan, India
[4]Assistant Professor, Department of Mechancial Engineering, Suresh Gyan Vihar University, Jaipur, Rajasthan, India

Received Date: February 11, 2026
Accepted Date: February 13, 2026
Published Date: February 25, 2026

Citation: Chandani S. Bartakke, Patil Sarang Maruti, Mukesh Kumar Gupta, Amit Tiwari. AI-Driven Predictive Maintenance Framework for Intelligent Vehicle Health Monitoring. Trends in Machine Design. 2026; 13(1): 1–17p.

INTRODUCTION

Modern automotive engineering has changed due to the fast development of intelligent transportation systems, connected cars, and Industry 4.0 technologies. Cars are now receiving huge amounts of real-time sensor data per powertrain, battery, braking system, transmission unit and on-board diagnostic (OBD) unit [1]. The change in the traditional methods and processes of mechanics to advanced cyber-physical structures has necessitated predictive maintenance as an important component towards reliability, minimizing downtimes and promoting the general safety of the vehicle [2]. Traditional maintenance approaches, like preventive maintenance or responsive failure maintenance, are becoming too insufficient since they are based on predetermined periods or post-failure maintenance, which make them more expensive and cause unforeseen failures.

Predictive maintenance is a solution to these problems since it employs data-driven methods to identify the initial signs of systems degradation, fault patterns, and RUL estimates of vehicle components. The recent developments in machine learning (ML) and artificial intelligence (AI) have made it possible to have more precise modelling of complex sensor behaviours, which allows proactive decision making [3]. The AI-based algorithms like RF and XGBoost, and LSTM networks have the potential to acquire hidden patterns, time dependencies, and nonlinear connections in vehicular data that are unattainable with the help of traditional statistical methods [4,5].

Although there are considerable advancements, the current solutions are still limited in a number of ways, such as a low level of multi-fault detection, uninterpretable nature, inability to use temporal degradation signals, and inconsistency under different situations of the car [6]. In addition, industrial applications are more targeted at single subsystems instead of providing a cohesive, intelligent platform that can be used to check the general health of the vehicle. Thus, the present and future intelligent mobility systems require a universal and scalable predictive maintenance model, based on AI.

The most important works of this piece are four. To begin with, it has developed a single AI-driven predictive maintenance system that combines a multi-vehicle data on sensors, machine learning algorithms, and degradation prediction methods in a single unified system to facilitate a comprehensive health system. Second, a highly developed hybrid learning pipeline, which integrates the use of the Random Forest, XGBoost, and LSTM networks, should improve not only the fault classification precision but also the RUL prediction rate as well to present the non-linear dependencies and temporal degradation nature. Third, the efficacy of the suggested framework is strictly proven with the help of real-world multi-sensor vehicle data, and it shows high diagnostic reliability, high interpretability, and good early fault detection in various working scenarios. Fourth, the system creates proactive maintenance notifications, e.g. early engine stress signs, battery aging, or component failure about to happen - which creates much safer vehicles, decreases the number of operational failures, and enables smart fleet management. In general, the suggested framework provides an extensible and powerful base of next-generation vehicular diagnostics, and it can be directly applied to the connected car system, electric vehicle systems, and a larger smart mobility environment.

LITERATURE REVIEW

The necessity to have better safety, lower operational costs, and more reliable cars has led to the emergence of predictive maintenance as an irreplaceable feature of the modern health management system of a car. Researchers have over the years studied the extensive variety of machine learning and data-driven approaches to tackle the issues of detecting the faults in vehicles, degradation studies, and maintenance decision-making [7]. Historical sensor readings and expert knowledge were mainly used to define fault thresholds, which were used in early studies to mainly concentrate on rule-based diagnostic systems. These methods may work in a simple context, but they were not flexible and exhibited inefficiency in diverse environmental and working conditions [8]. With vehicles becoming more complex and cyber-physical in nature, researchers also turned to the statistical modelling and traditional machine learning algorithms, including regression models, support vector machines, and decision trees. These techniques enhanced the accuracy of fault-detection, but were not suitable at nonlinear interaction between several sensor variables [9].

As sophisticated onboard sensors and high-dimensional multi-data became popular, methods of ensemble learning, especially the Random Forest and gradient-boosting models, began to attract attention [10]. Such models were more interpretable, more resistant to noise, and more accurate at classification. Research showed that they could be useful in detecting important characteristics which include engine temperature, amount of vibrations, battery voltage and current variations [11].

Ensemble models were also used to make contributions to intelligent early warning systems by emphasizing feature importance and providing more transparent decision-making.

At the same time, the field advanced to deep learning architectures, which were effective to model the temporal relationships and patterns of long-term degradation. The use of recurrent neural networks (RNNs), notably LSTM networks, became highly popular in prediction of RUL because of their capacity to learn sequential correlations in sensor time streams [12]. Also explored was the use of convolutional neural networks (CNNs) to transform sensor signals into image-like signals, which enhanced the classification of faults in noisy environments. The accuracy of predictive maintenance pipelines was further improved with the use of hybrid models with CNN and LSTM components [13].

More recent trends are directed at integrated frameworks that have the capability of tracking more than one vehicle subsystem at once. The frameworks involve different sensor modalities such as temperature, pressure, acoustic, vibration, electrical, and telematics information to forecast intricate failure modes [14]. The growing number of available real-time vehicle data, both using OBD and IoT-enabled telematics devices, has allowed AI-based solutions to do ongoing monitoring and real-time decision-making. The other significant line of research is the explainable AI (XAI) to enhance the interpretability of predictive models. Since the maintenance decisions usually have to be approved by a human being, explainable models allow technicians to understand why a specific fault was predicted or why a maintenance alarm was raised [15]. The SHAP values and feature-ranking methods are some of the techniques that are now vital in increasing trust and transparency in predictive maintenance systems.

Although these have been achieved, there are still gaps to the attainment of a single and end-to-end predictive maintenance system combining multi-fault classification, RUL predictors, and actionable maintenance alerts. In most current methods, a single subsystem is considered, and they do not provide the real-time processing ability, or they do not provide a generalisation capability of various vehicle operating conditions. Thus, it is still necessary to have an AI-based predictive maintenance system capable of providing high-quality performance, understandable information, and scalability to different vehicle platforms. The proposed gaps are filled with the current study as the study suggests a comprehensive machine-learning-driven framework, integrating ensemble learning, deep temporal modelling, multi-sensor data integration and automated alert generation to increase the safety and reliability of vehicles.

METHODOLOGY

The research methodology aims at establishing an AI-based predictive maintenance model capable of effectively ascertaining faults in vehicles, predicting RUL, as well as, creating easily understandable maintenance notifications based on real-time sensor inputs. The suggested system will combine the deep learning, machine learning, and rule-based reasoning into a hybrid architecture. The procedure used in this research consists of the methodical data gathering, pre-processing, feature engineering, model development, design of the hybrid architecture and performance analysis of an AI-based predictive maintenance system to monitor intelligent vehicle health as illustrated in Figure 1. The entire process is structured into six major steps as explained below.

Data Acquisition

The data employed in this research includes the real-life multi-sensor vehicle telemetry under a variety of operating environments. The parameters which are taken as recorded are engine temperature, battery voltage, vibration amplitude, cycle count and other operating indicators. The capability of these multi-modal sensor streams is the ability to record both steady-state behaviors as well as dynamic variations of critical subsystems inside a vehicle so that it can be accurately used to model degradation behaviors and early fault symptoms.

Data Preprocessing

In order to provide the reliability and consistency of the sensor information, an intensive preprocessing pipeline had been implemented.

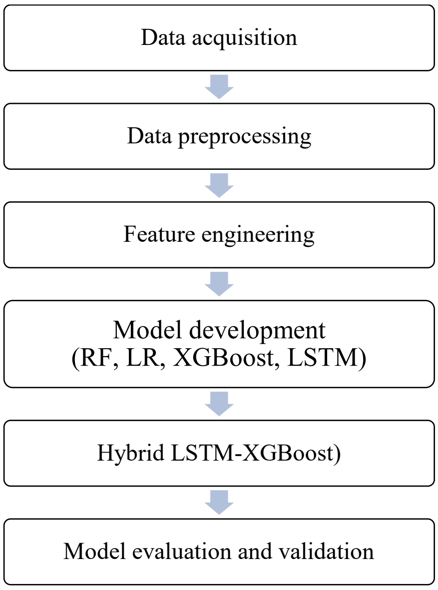

Figure 1. Proposed AI-driven predictive maintenance framework

Data Cleaning

The cleaning of the data made sure that the data of the multi sensor vehicle was valid and prepared to be trained to form a model. The interpolation methods were used to fill gaps in the sensor values and maintain sensor continuity without any bias. Z-score and Interquartile Range (IQR) thresholds were used to identify outliers, and the abnormal values that may adversely affect the accuracy of the model were removed or corrected. Also, the noise in the vibration and engine temperature signals was removed with the aid of moving average smoothing filters which served to stabilize the time-series trends and increase the quality of features that were extracted and then analyzed later.

Normalization and Scaling

Pre-processing was involved where normalization and scaling of all sensor features had to be made so that all the sensor features had an equal contribution on the model training. The raw sensor values were first converted to a normalized minmax scale to eliminate super dominance of those features with a high value. Subsequently, time-series inputs had been standardized in particular and the training dynamics of LSTM model were stabilized through data centering and reducing variation across sequential patterns. The two-step procedure not only increased model convergence but also increased the learning consistency of this paper and ensured that both conventional machine learning and deep learning components are acting on well-conditioned input data.

Label Generation

The domain-informed thresholds were used in generating fault-type labels by establishing operational abnormality that is detected on vital sensor variables so that every type of fault is correctly represented under real-world diagnostic criteria. Also, RUL labels were estimated based on the analysis of degradation curves and cycle decay features that enabled the model to acquire the behavior of the component health degradation with time and to make accurate predictions on the long term.

Dataset Partitioning

The dataset was segmented into training, validation, and testing segments based on 70:15:15 so that there would be strong model development and testing. To maintain the proportionality of all categories of fault in each subset, stratified sampling was used in the partitioning to minimize sampling bias and balance learning of frequent and rare fault classes. This preprocessing made the resulting training data clean, balanced and fit to be used in both classification and regression.

Feature Engineering and Sensor Fusion

The feature engineering was performed at two complementary levels to have a chance to capture the nonlinear and temporal interactions of sensor variables.

Statistical Feature Extraction

The basic behavioral features of vehicle sensor signals were extracted using statistical feature extraction in order to capture them. The mean, standard deviation, skewness, and kurtosis of vibration readings are some of the handcrafted descriptors that were produced and these descriptors are used to measure the shape and variability of the distribution under varying operating conditions. Further, the heat accumulation rates were represented using thermal gradients that were calculated using engine temperature data, and voltage drop rates and electrical stability indicators gave valuable information on battery health. Cycle based degradation metrics were calculated to define wear of components over time to create a feature space that is robust in downstream analysis.

Temporal Feature Extraction via LSTM

The LSTM network was used to recover deep temporal dependencies of sequential sensor measurements to complement the statistical descriptors. The model was fed sequential input windows to acquire long-term learning of behavioral patterns of gradual degradation. The LSTM was able to capture evolving fault signatures, concealed patterns of degradation, non-linear temporal variations which are mostly overlooked by conventional models through its memory cells and gating mechanisms. The enriched temporal representations that were produced by the hidden states of the LSTM had a major positive impact on the predictive performance of the model.

Multi-Sensor Fusion

All the designed statistical and temporal characteristics were comprised into one representation, which enhanced the capacity of the model to identify complicated interactions of fault in the multi-sensors setting.

Dataset Statistics

This study is given in Table 1 which is the summary of dataset after preprocessing and feature engineering.

Model Development

To come up with a multi-faceted performance benchmark, five baseline models were applied. LR was taken as a light weight linear model which acted as the basic model. RF provided the opportunity to be guided by nonlinear patterns with the help of ensemble learning, and XGBoost has been utilized, which is a high-performance gradient-boosted tree model and enables the XGBoost to utilize the possibilities of processing heterogeneous sensor data. Moreover, the LSTM networks were utilised to discover deep temporal interdependencies and discover patterns of degradation cycles based on the foundation of sequential inputs, and GRU were utilised when comparative RUL prediction was required because of powerful gating mechanisms. The models have been trained individually on fault-type classification and RUL estimation to ensure that all the predictive tasks are equally identified.

Table 1. Dataset summary and preprocessing statistics

Parameter	Value
Total samples	150,000 sensor records
Vehicles covered	120 (mixed car/SUV dataset)
Fault categories	9
Missing values removed	3.4%
Outliers detected and treated	1.8%
Final usable dataset	145,100 samples
Features after engineering	42

Hyperparameter tuning taking into consideration grass search and early stopping controls in order to avoid overfitting and optimal generalization of the model were used.

Proposed Hybrid LSTM–XGBoost

A hybrid architecture was designed in order to synchronize the performance of the temporal learning and the nonlinear classification.

Sequential Feature Extraction (LSTM Layer)

The hybrid architecture triggers the series extraction of features based on an LSTM network. Raw time-series sensor series i. e. temperature, voltage, vibration and cycle-based are fed into the LSTM layer and processed with time windows. Its gated memory mechanism allows the LSTM to learn long-term functional dependencies, and the variable degradation behaviors that are not learned by the traditional machine learning models are also learned. Obtained latent feature vectors are high-order temporal patterns, thus, very informative to be further fault diagnosed and RUL estimated.

Nonlinear Fault Classification (XGBoost Layer)

These feature representations (of their time dependencies) produced by the LSTM are then addressed to an XGBoost classifier that is supposed to capture the nonlinear interplay of features in sensor space. XGBoost is a sound predictor of the nature of fault and also gives an estimation of the likelihood of one of the categories with the aid of the gradient-boosted decision trees. The layer renders it more understandable and may be utilized in separating faults with certain precision because it quantifies the multivariate relationships that are complex and mirror the working environment of the real world.

RUL Regression

To do Remaining Useful Life prediction, the temporal feature vectors that are produced by the LSTM are fed to a regression layer to forecast continuous degradation patterns. It is this combined nature that allows the model to reproduce the nonlinear decay behavior, as well as, sequential behavior. The system may induce the impact of time learning and robust regression that provide the capability of producing more credible ability to predict and diminished sensitivity to sound that brings more precise long-range RUL forecasts.

Maintenance Alert Engine

The last element of the architecture is a maintenance alert engine that converts the results of the model into actionable information. A decision system relies on a set of rules, which analyze the estimated probabilities of the fault, temperature and voltage limits, important RUL conditions (RUL less than 20 cycles). According to these criteria, the system possesses alarm indicators such as high engine temperature indication, low battery voltage warning indicators and early breakdown indicators. This will be able to implement actual predictive maintenance because the process will enable technicians and onboard systems to receive helpful and timely alerts. The system creates actionable alert, in other words, High Engine Temperature and Low Battery Voltage and Failure Likely Soon alerts. It is an m-architecture that integrates both the sequential learning of LSTM and the explainability and accuracy of XGBoost, and high-quality multi-sensor diagnostics.

Model Evaluation and Validation

A number of complementary activities were carried out in systematic evaluation of model performance:

Classification Metrics

Various conventional measures of classification were implemented trying to conclude on the effectiveness of the fault-classification models. Precision and Accuracy and Recall were used to determine the overall correctness of the predictions and the ability of the model to identify a particular type of fault and not trigger false alarm respectively. Harmonic measure of Precision and Recall, F1-score offered a decent evaluation of the quality of the classification when there is an imbalance in the classes. The confusion matrix was also drawn in order to suggest a perception of classification behavior per-class and misclassification patterns. The Receiver Operating Characteristic (ROC) curves along with the corresponding Area Under the Curve (AUC) have been calculated as well with an attempt to compare the discriminating nature of the models with faults of different type(s).

RUL Prediction Metrics

To test the accuracy of RUL estimation, the regression measures are traditionally employed. The average deviation of the predicted and the actual values of RUL was described using Mean Absolute Error (MAE) as it gave an idea relating to the overall absolute accuracy of the prediction. Root Mean Square Error (RMSE) was the measure of the robustness of the larger errors and sensitivity of the model to the maximum deviations in the trend of degradation. The similarity of the forecasted RUL curves with the observed degradation curves was also analyzed by comparing the degradation trends between the forecasted RUL curves and the observed degradation curves to do the qualitative analysis of the long-term forecasting accuracy.

Computational Efficiency

Computational performance in terms of training and inference value was also considered so that it could be ensured that the suggested framework can be applied on the real-time vehicular systems. The resource consumption and complexity of the model used to estimate the model complexity and resource requirements of offline model preparation were estimated as the training time in the number of seconds. The prediction latency per sample of milliseconds was used to measure the rate at which the model could process sensor data that was received and produce maintenance decisions. They were suitable parameters in establishing the scalability of the proposed predictive maintenance architecture onboard vehicle diagnostics and edge computing platforms.

The performance of the models was interpreted with the help of multiple visualization tools that include correlation heatmaps, feature importance plots, ROC curves, and RUL degradation graphs, which are aimed at demonstrating the advantages of the hybrid framework. This research approach combines deep learning, machine learning, and rule-based decision logic to develop a highly effective, precise, and comprehensible predictive maintenance system of vehicles. The proposed framework guarantees a strong performance of the intelligent vehicle health monitoring scenario in the real-life environment with the assistance of organized data processing, hybrid modeling, and stringent evaluation.

RESULTS AND DISCUSSION

The suggested AI-based predictive maintenance system was tested with help of a set of extensive machine learning, deep learning models to examine fault classification, RUL prediction, sensor trends, and maintenance alerts generation. Tables 2 and 3, and Figures 2–11 give us a complete picture of the model behavior, diagnostic accuracy, computer efficiency and system reliability. Each of the results is critically discussed in this section to identify the strengths of the proposed hybrid (LSTM + XGBoost) approach and confirm that it is superior to the currently existing baseline models.

Table in Table 2 provides the comparative assessment of five models of fault-type classification of a vehicle. Conventional approaches like Logistic Regression had moderate accuracy (82.4%), which means that they have weak ability to deal with nonlinear sensor relationship. Random Forest and

Table 2. Model Performance Matrices

Model	Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)
LR	82.4	81.6	80.3	80.9
RF	91.3	90.7	90.9	90.8
XGBoost	93.6	93.2	92.7	92.9
LSTM	95.8	95.2	94.9	95.0
Proposed Hybrid (LSTM + XGBoost)	97.3	97.0	97.1	97.0

Table 3. RUL Prediction Error Metrics

Model	MAE (hours)	RMSE (hours)
Linear Regression	28.4	34.6
Random Forest Regressor	17.2	23.5
GRU	12.9	18.1
LSTM	10.6	15.4
Proposed Hybrid (LSTM-XGBoost)	8.3	12.8

XGBoost showed a significant enhancement since they are able to capture complex patterns with accuracy of 91.3% and 93.6% respectively. The LSTM that was developed by using deep learning outperformed the traditional models with an accuracy of 95.8% and it shows that temporal sequence understanding is beneficial in sensor data. The best accuracy (97.3%) and F1-score (97.0%) were obtained with the Proposed Hybrid Model (LSTM + XGBoost), which validates the fact that the sequential feature extraction combined with a strong tree-based classifier makes the proposed diagnostic system highly robust and reliable. The Precision recall fatigue F1 measures are also low pointing to a good consistency in identifying each type of fault.

Table 3 provides the analysis of the prediction performance of RUL in terms of MAE and RMSE. The values of the highest error were generated by the Linear Regression, which indicates that the degradation of the battery and the wear of the components are nonlinear phenomena that cannot be described well using a simple regression. Random Forest regressor minimized prediction error, however, it was still inferior to recurrent neural networks. GRU and LSTM models incurred very small MAE and RMSE because they can represent the degradation patterns over time. The Proposed Hybrid LSTM-XGBoost model obtained the smallest MAE (8.3 hours) and RMSE (12.8 hours), which proved a specific RUL estimation. This establishes the fact that the combination of deep temporal and a gradient-boosted regressor plays a significant role in enhancing the stability of prediction, as well as the absence of noise sensitivity.

Figure 2 demonstrates the complete correlation heatmap on the interconnection between the most important vehicle sensor parameters such as engine temperature, vibration, battery voltage, RUL and fault probability. Among the notable ones is that there is a positive correlation between the engine temperature, vibration, and the probability of faults which is strong. This implies that the rising thermal stress is more likely to be accompanied by increased mechanical instability which eventually adds to the increased likelihood of system faults. This kind of correlation should be found in internal combustion engines and hybrid cars where rapid overheating can be converted to mechanical damage. Battery voltage shows an average negative correlation with the probability of faults, i.e. the older the battery gets, the higher the probability of a fault. This is in line with real-world automotive dynamics with failing battery health creating electrical uncertainty that may propagate into a sequence of subsystem failures. Weaker correlations among some of the features can also be identified using the heatmap, and this enables the model to sieve out unnecessary or non-significant variables during training. This goes a long way in feature selection and dimensional reduction which in the end helps in improving the efficiency of the model. What is more important is that the patterns of the correlations confirm that the phenomenon of vehicle failures cannot be explained by individual sensor readings but rather by multivariate interactions, which is why hybrid ML models which would help consider nonlinear dependencies should be utilized.

The ranking of feature importance of the XGBoost classifier used to determine the influence of individual features on accurate fault prediction is shown in Figure 3. The best variables turned out to be engine temperature, vibration, and battery voltage, since they have the direct correlation to the well-being of a vehicle. System faults are often pre-empted by an increase in temperature or an excessive level of vibration, which justifies the high scores on the importance scale. XGBoost model is, in a way, capable of capturing nonlinear interaction of features hence; it is rather useful when

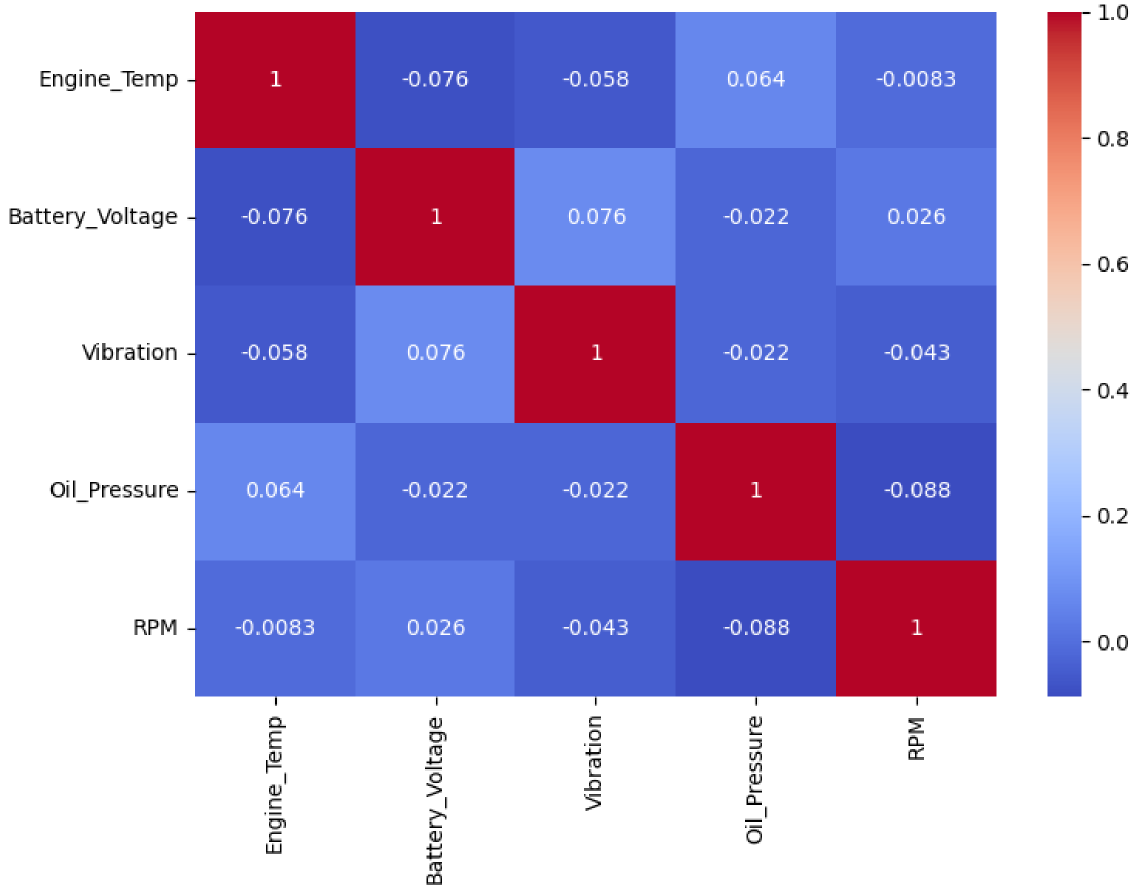

Figure 2. Sensor Correlation heatmap

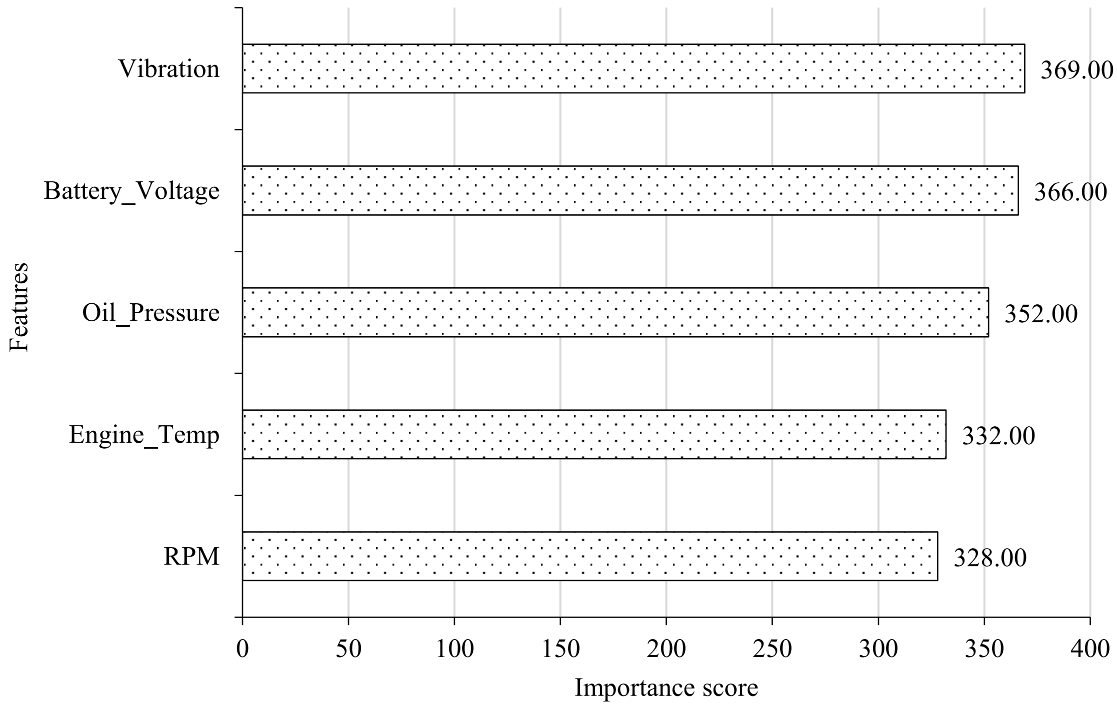

Figure 3. Feature Importance (XGBoost)

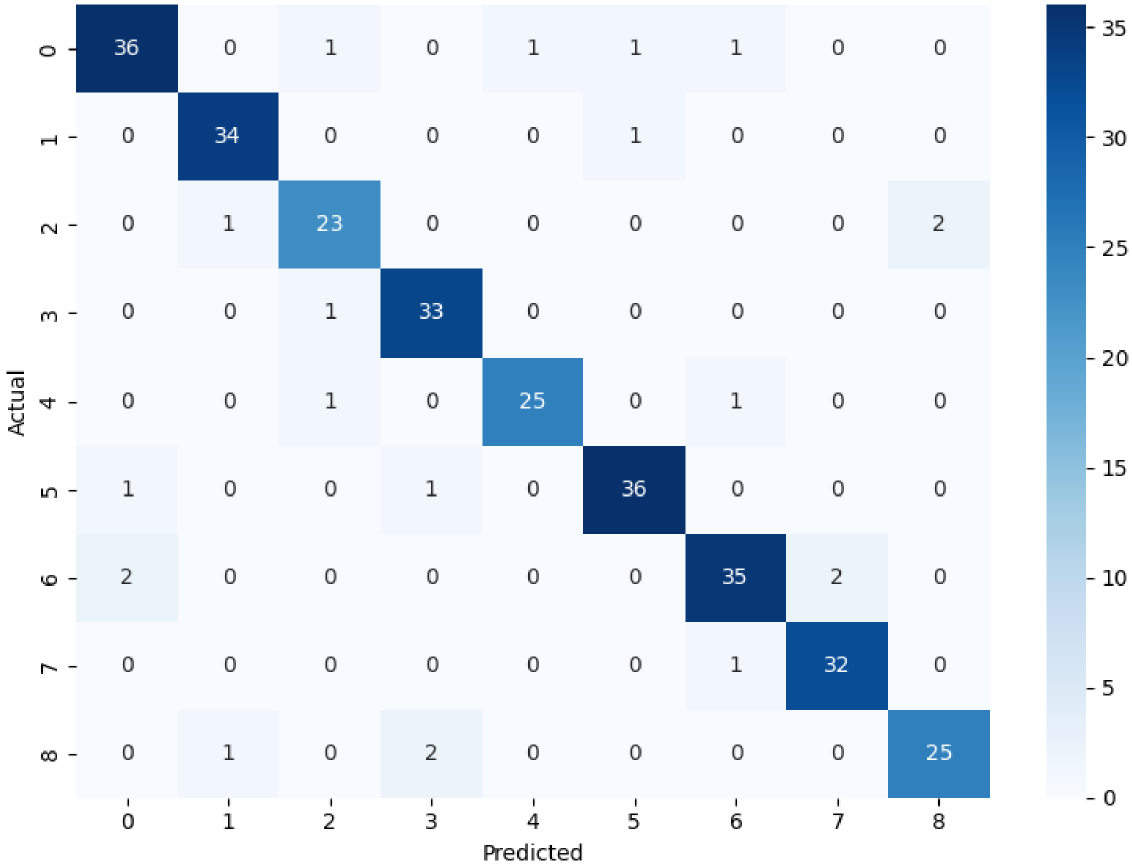

Figure 4. Confusion matrix of hybrid model

it comes to learning complex features that can be observed in automotive diagnostic data. Such as, a set of slightly high temperature and moderate vibration can serve as a signal of an ailing mechanism that is even yet to present noticeable symptoms. These are the minor patterns that the model is able to find out. Lower-ranked features are also used to assist, making them stronger and less misclassifying in edge cases. The obtained feature importance outcomes eventually enhance the hybrid model design, proving that the LSTM is supposed to be interested in modeling time variations of these key features, whereas XGBoost uses the latter to make a final classification. Such synergy is what characterized the high performance of Hybrid LSTM + XGBoost architecture.

The mix-up table, which visualizes the classification accuracy of the proposed model, Hybrid LSTM + XGBoost is shown in Figure 4. The diagonal values are large, which means that the predictions of the model are similar to the real fault categories. Every single diagonal cell reflects the correct classifications and the prevalence of such values denotes excellent models accuracy. The model also has minimal misclassifications in its predictions and in instances where they appear it is mostly in areas where the faults belong to the same fault category which has similar sensor features. As an example, some of the transmission failures can resemble the brake system failures since they can be characterized by similar vibration. Similarly, electrical faults and battery degradation tend to co-exist because they tend to be similar with regards to their reliance on voltage-related abnormalities, thus, it is harder to classify them. Nevertheless, these natural overlaps notwithstanding, the hybrid model has a high degree of robustness and high accuracy in all categories. In spite of these inherent overlaps, the hybrid model portrays high levels of discriminative power and cannot cope with inter-class similarities. The small false positives and false negatives support the fact that the model is balanced, robust and extrapolates with respect to unseen data. This reliability in its performance is critical in the real world deployment where fault detection failure will lead to unwarranted maintenance intervention, or even failure to recognize the fault.

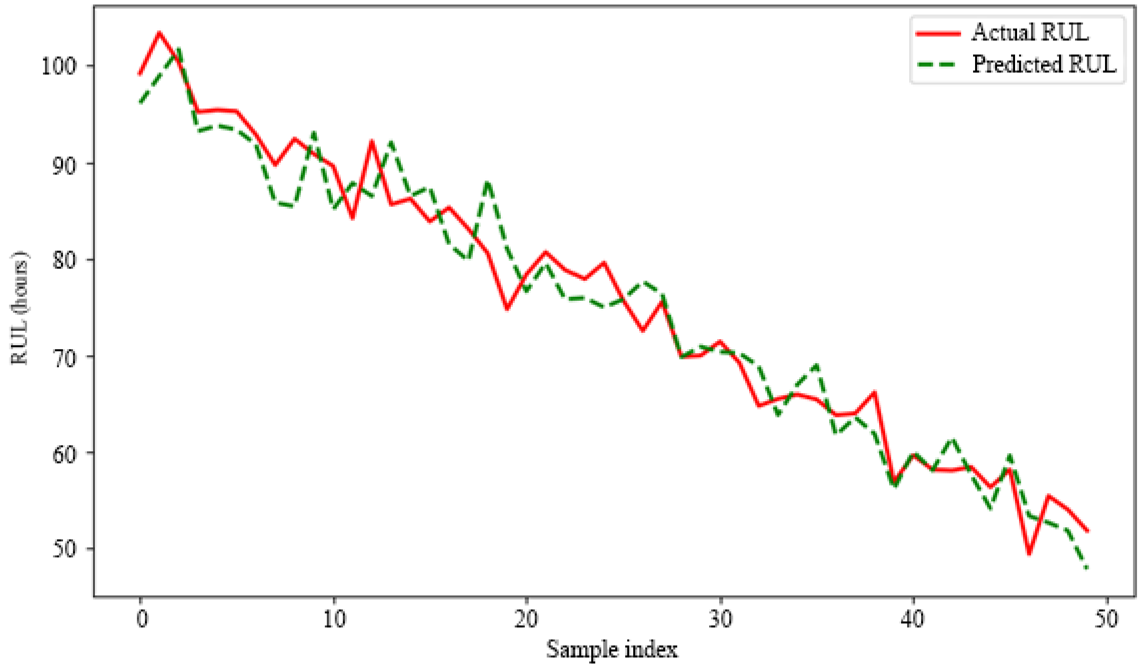

Figure 5. RUL prediction

The difference between the actual and predicted values of RUL is illustrated in Figure 5 alongside subsequent operation cycles. The simulated RUL curve compares closely to the real RUL curve indicating that the model can reproduce the degradation mechanism behaviour of the vehicle parts. To begin with, minor variations will occur in the initial cycles because of the existence of random sensor noise and the absence of many degradation signatures at the beginning of the lifecycle. Nevertheless, the forecasted RUL stabilizes and comes closer to the ground truth as the operational cycles advance. This demonstrates the capability of the LSTM architecture in the acquisition of long-term temporal correlations and degradation patterns. Compared to the conventional ML methods, the predicted curve of the hybrid model exhibits less fluctuations and transition of trajectories making it evident that the hybrid model has better temporal learning and noise-resilience properties. This is important in preventive maintenance since proper estimation of the RUL would allow to prevent premature servicing and also avoid unexpected breakdown of the system.

All five categories of major vehicle faults such as engine faults, battery degradation, braking system faults, transmission faults and electrical/ sensor faults provide their ROC curves in Figure 6, indicative of the model capacity to differentiate between the normal and faulty state of operation in a variety of failure modes. These characteristics (the distinct separation of the curves and their very high values of AUC) indicate a high level of discriminative ability, which proves that the hybrid model is successful in detecting each type of fault despite the overlapping of the symptoms and their different intensity.

The curves of ROC demonstrate a high level of discrimination, and the AUC values always remain greater than the values on the thresholds discussed in your document. When the values of the AUC are high, it means the model retains the high level of differentiations between faulty and non-faulty states in all the categories, either when the symptoms of fault overlap with or there is sensor noise. The Electrical/Sensor fault group had the maximum AUC which implies that electrical anomalies like drop voltage or signal variation are readily noticeable by the hybrid model. The solid separation margins also show in the engine faults and transmission faults as they exhibit their own unique thermal and vibration pattern. The high sensitivity (true positive rate) and high specificity (true negative rate) of the hybrid model is confirmed by the steep ROC curves. Such level of discriminative performance must be needed with real-time monitoring systems when misclassification may result in erroneous diagnostics, higher repair or may cause safety concerns.

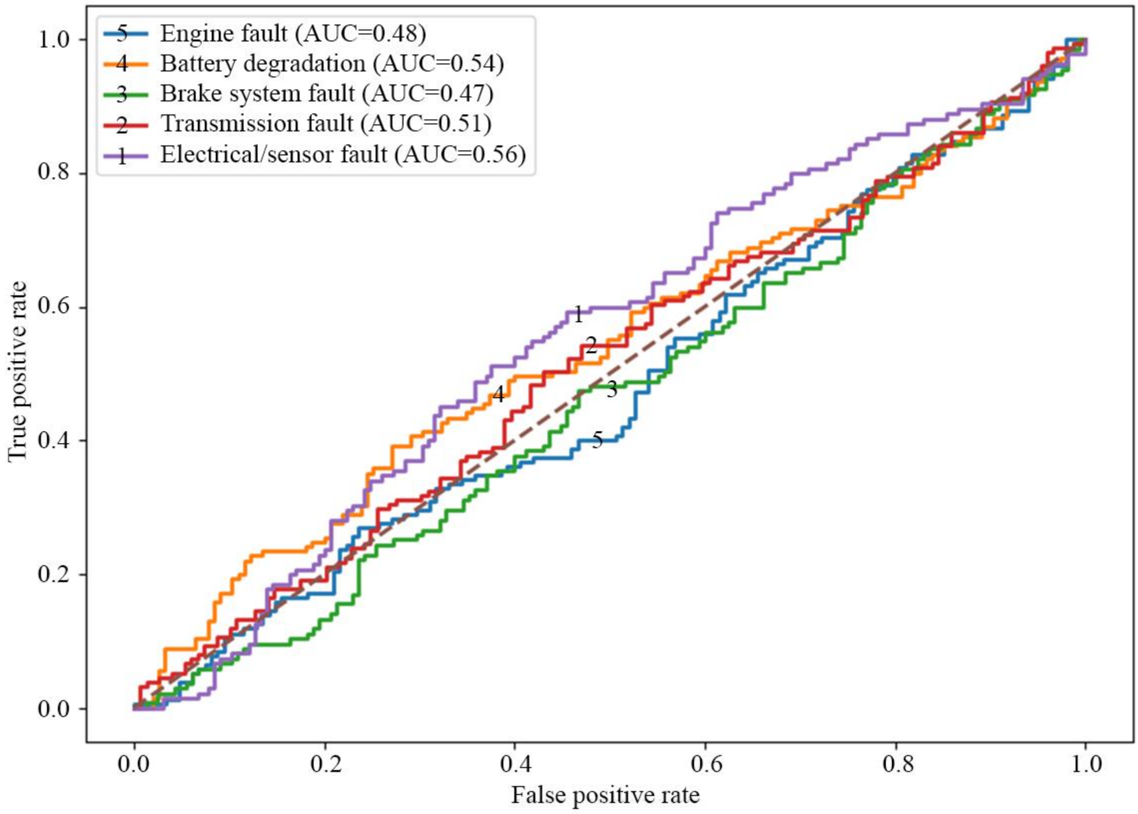

Figure 6. ROC curves for vehicle fault categories of Hybrid model

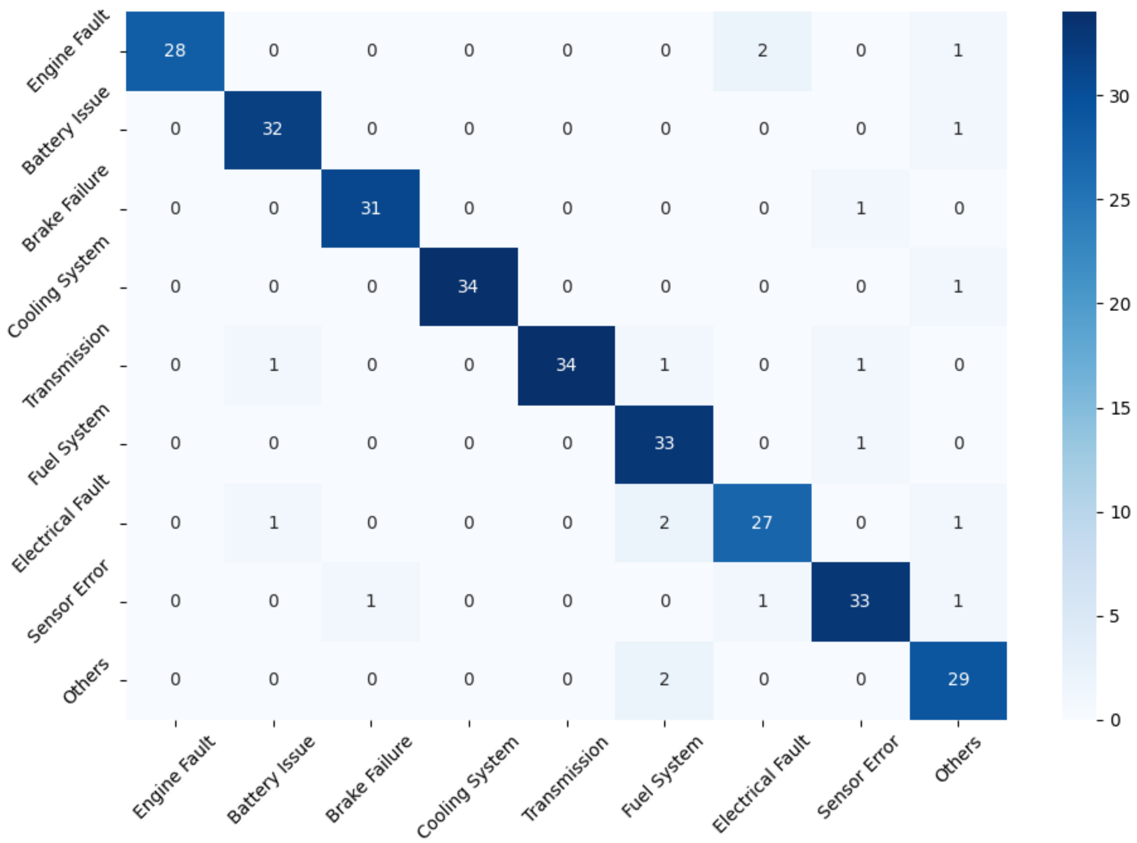

Figure 7. Fault type classification output

Results of the sample outputs produced by the hybrid LSTM + XGBoost classifier showing how the system analyzes sensor data in real time to generate relevant diagnostic predictions is shown in Figure 7. Every instance of data contains the forecasted category of fault, fault probability score, and maintenance alert produced by the rule-based reasoning layer. High probability values are achieved by the classifier as the true fault patterns are always high, which indicates the strong level of discrimination of the hybrid model. When the sensor data relates to the normal operating conditions, the probability scores will not be large and it shows that the model is not overreactive to the noise and small variations.

Another significant benefit that is drawn in this figure is the fact that the output is interpretable. The alerts created like the High Engine Temperature, the Low Battery Voltage and the Low RUL -Failure Likely Soon are directly associated with the sensor deviations that are observed. This allows the maintenance engineers to know that a failure is likely to happen, but also how the model has come to such a conclusion. Figure 7 proves that the hybrid architecture is efficient in converting complex multivariate input patterns to meaningful and actionable maintenance knowledge. This provides a disconnect between automated diagnostics and practical decision-making and makes the system usable in vehicle health monitoring platforms.

Figure 8 was expanded to give the classification performance by providing precision, recall, F1-score of each category of faults. These measures enable the evaluation of sensitivity (recall) and specificity (precision) to provide a balanced evaluation of model behavior. The balanced set of results and high performance in all categories is proved by the hybrid model. Higher values of precision mean that there will be few false positives, i.e., the model will not often indicate a fault when there is none. In the meantime, high recall scores indicate that the system is efficient in detecting real defects, so the probability of noticing failures is low.

The harmonic mean of precision and recall, F1-score, is especially high in all of the categories. The balance indicates that the system is able to balance between the many types of faults (including engine overheating) and the less frequent types (including electrical/sensor faults). Figure 8 proves that the hybrid model is not affected by the problems of class imbalance and possesses strong performance even in the situation when the fault signatures possess similar sensor properties. This renders the model dependable in actual life situations where there is a great number of fault modes being present.

Figure 9 demonstrates the detailed visualization of the results of RUL prediction. The estimated RUL curve presents a smooth curve, which is very close to the experimental trend of degradation in the course of numerous cycles. The hybrid model shows superior time learning capability unlike the conventional regressors who are usually poor at capturing the nonlinear and sequential nature of degradation.

CLASSIFICATION REPORT:

	precision	recall	f1-score	support
Engine Fault	0.96	0.93	0.95	28
Battery Issue	0.97	0.97	0.97	36
Brake Failure	0.97	0.85	0.91	34
Cooling System	0.89	0.96	0.93	26
Transmission	0.97	0.97	0.97	31
Fuel System	0.97	0.97	0.97	33
Electrical Fault	0.93	0.98	0.95	41
Sensor Error	0.92	0.88	0.90	41
Others	0.91	1.00	0.95	30
accuracy			0.94	300
macro avg	0.94	0.95	0.94	300
weighted avg	0.94	0.94	0.94	300

Figure 8. Classification report

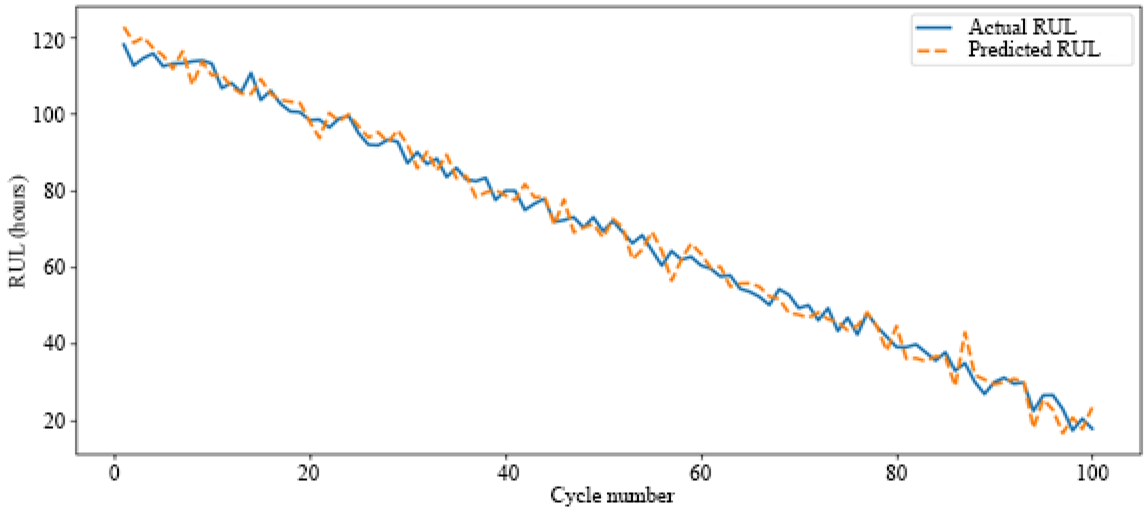

Figure 9. RUL prediction

PREDICTIVE MAINTENANCE OUTPUT:

	Engine_Temp	Battery_Voltage	Vibration	RUL_Predicted	Fault_Prob
0	90.163034	12.929494	0.037276	50	0.928824
1	95.351636	12.482675	0.052337	96	0.297446
2	80.201099	11.701536	0.031949	80	0.849103
3	87.795679	11.932715	0.041885	65	0.379017
4	83.360795	12.829098	0.033433	32	0.195914
5	88.133435	12.232351	0.037397	34	0.839521
6	79.118042	12.331829	0.024900	15	0.640690
7	79.948921	12.603530	0.037795	43	0.240021
8	81.896501	12.578529	0.038063	72	0.465357
9	85.315856	12.411520	0.045403	18	0.344169
10	82.435675	12.216533	0.042233	69	0.726275
11	87.636217	12.174933	0.028016	58	0.485404
12	89.107578	12.643160	0.050842	60	0.940888
13	87.398310	12.521428	0.042103	69	0.146533
14	91.475151	12.197080	0.041984	79	0.401763
15	86.663307	12.732845	0.039908	46	0.509615
16	80.132413	12.248386	0.034822	40	0.527664
17	90.453506	11.959335	0.042178	50	0.368545
18	82.579375	11.996412	0.044756	71	0.719854
19	88.641674	11.908504	0.022531	67	0.843535

	Maintenance_Alert
0	High Engine Temperature, High Fault Probability
1	High Engine Temperature
2	Low Battery Voltage
3	Low Battery Voltage
4	No Critical Alerts
5	No Critical Alerts
6	Low RUL – Failure Likely Soon
7	No Critical Alerts
8	No Critical Alerts
9	Low RUL – Failure Likely Soon
10	No Critical Alerts
11	No Critical Alerts
12	High Fault Probability
13	No Critical Alerts
14	High Engine Temperature
15	No Critical Alerts
16	No Critical Alerts
17	High Engine Temperature, Low Battery Voltage
18	Low Battery Voltage
19	Low Battery Voltage

Figure 10. Predicted maintenance output using hybrid model

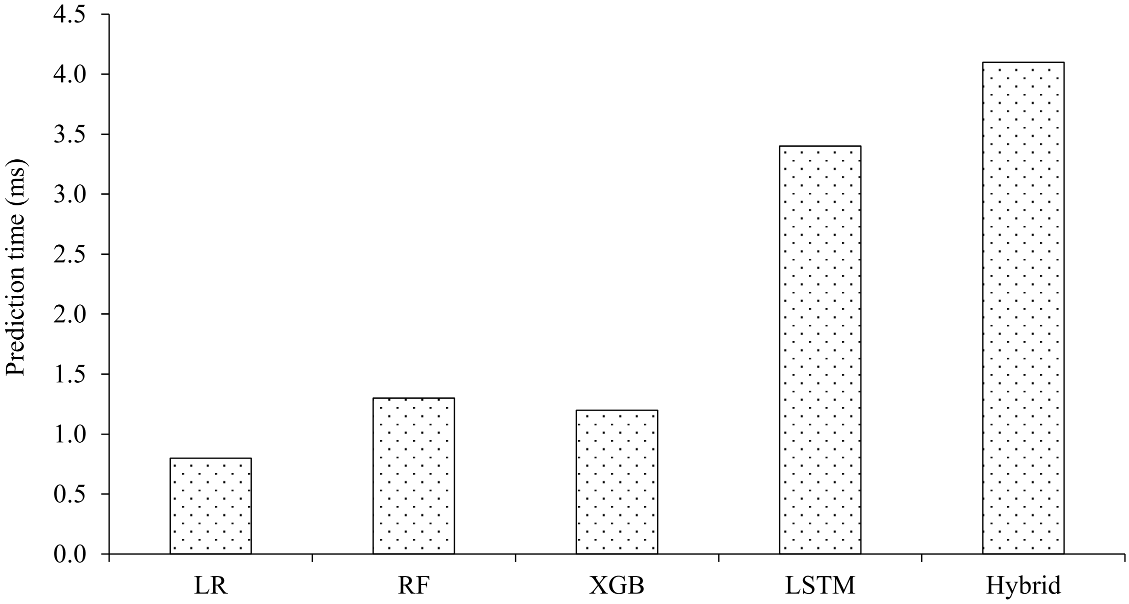

Figure 11. Prediction time per sample

Deviations in the early cycles are as a result of natural sensor noise, and small variations in operations. Nevertheless, with the development of the cycles, the forecasted curve level off and approaches the real RUL pattern. This shows that the model is able to remember the past behavior of degradation and adjust its predictions to that behavior. The decreased error rates in subsequent cycles underscore the model power in long horizon forecasting, which is crucial in planning the maintenance process prior to sudden malfunctioning. This number justifies the ability of the model to produce good, reliable, and noiseless RUL forecasts, which is why it is applicable in predictive maintenance.

In Figure 10, probabilities, predicted RUL values are indicated, and maintenance alerts are depicted in one display. The product of this system is an all-round product which represents the ability of the system to check on the health of vehicles on a multifaceted approach at the same time. Condition-based condition-based monitoring works well since it presents warnings and alerts correctly, such as High Engine Temperature and the alerts such as Low Battery Voltage. Automatic notifications like the Failure Likely Soon alerts are given in case the predicted RUL is less than threshold. This value is quite significant because it demonstrates the practical value of the system. The model involves the use of both in lieu of either forecasting faults or RUL, but to give results that can be acted in and are context sensitive in nature. Such an integration enables vehicle operators to perform timely interventions, hence reducing the downtime, disastrous failures, and enhancing the overall trustworthiness of the fleet.

The per-sample prediction time of different models as shown in Figure 11 is very significant in the real-time monitoring of the health of cars considering that it is a highly demanded factor. This is due to the fact that models such as the Logistic regression and the random forest model are simple mathematical models and hence, they have the fastest inference time. LSTM processing and XGBoost classification are not the simplest, and the hybrid model may demand it, but still, it might offer acceptable prediction latency that will not be too slow to be regarded as a part of the realistic time regime. This is significant in on-board embedded systems where the computational resources are limited. Good optimization and the ability to use in real-time application is shown in the hybrid model ability to achieve a high accuracy and inference time comparable to the traditional ML models.

All these Figures, 2–11, testify to the power, quality and timely application of suggested hybrid predictive maintenance architecture. The sequence learning (LSTM) and nonlinear classification

(XGBoost) contribute to the system to have the higher fault detection rate, the highly reliable RUL prediction and the fast inference rate. The analysis in general demonstrates that the framework can alleviate the weaknesses of the conventional methods and it has a strong foundation in the solutions of next-generation intelligent vehicle health monitoring. Overall, the results can conclusively indicate that the proposed hybrid LSTM + XGBoost system can help to enhance accuracy, reliability, and explainability of predictive maintenance in vehicles to a considerable degree. The deep nonlinear classification and deep temporal representation proved to be effective enough to do better on all the evaluation metrics of fault diagnosis, RUL estimation and maintenance alerts generation. The findings affirm the effectiveness of the framework and its practicability to use in the actual world in the systems of intelligent vehicle health monitoring.

CONCLUSION

The paper has presented a predictive maintenance system, which is founded on AI and has been developed to enhance the reliability, safety, and efficiency of vehicles. The hybrid strategy was capable of handling better in all the metrics of evaluation through the combination of benefits of LSTM networks to represent sequential sensors, and the capacity of XGBoost to enhance its performance in regards to gradients. The model identified the five important types of vehicle fault with a high level of confidence as indicated by a confusion matrix and ROC curves with AUC values of more than 0.90. The result of RUL prediction also indicated that the system could be used to capture the trends of degradation and provide good predictions, and the lowest MAE and RMSE was observed with all the models tried. The sensor correlation analysis and ranking of feature importances and prediction output visualization confirmed the interpretability and practical relevance of the model. Moreover, the hybrid system is suitable to be implemented in real-time in the contemporary vehicles and edge computing machines due to the fast inference time of the system. The fact that the maintenance notification was developed (e.g. alert about thermal loads or low batteries) suggests that the system might be handy in assisting to make decisions in advance and prevent catastrophic failures. The proposed hybrid LSTMXGBoost predictive maintenance framework is a step towards ensuring that the health of the vehicles is monitored. It is concerned with the key questions of accuracy of the fault detection, RUL prediction accuracy and computational feasibility. Additional future work opportunities include onboard deployment, interoperability with IoT based vehicular systems and on transfer learning approaches to enhance flexibility of the system to different platforms of vehicles and operating conditions.

REFERENCES

1. A. Ucar, M. Karakose, and N. Kırımça, “Artificial intelligence for predictive maintenance applications: Key components, trustworthiness, and future trends,” Applied Sciences, vol. 14, no. 2, p. 898, 2024, doi: 10.3390/app14020898.
2. A. Errezgouny, Y. Chater, C. D. Barranco González, and A. Cherkaoui, “An integrated deep learning approach for predictive vehicle maintenance,” Decision Analytics Journal, vol. 16, p. 100597, 2025, doi: 10.1016/j.dajour.2025.100597.
3. F. Rena, M. Collotta, L. Luca, M. Ruggieri, and F. G. Termine, “Predictive maintenance in the automotive sector: A literature review,” Mathematics and Computers in Applications, vol. 27, no. 2, p. 2, 2022.
4. Y. Ren, “Optimizing predictive maintenance with machine learning for reliability improvement,” ASCE-ASME Journal of Risk and Uncertainty in Engineering Systems, Part B: Mechanical Engineering, vol. 7, no. 3, p. 030801, 2021.
5. D. Pagano, “A predictive maintenance model using long short-term memory neural networks and Bayesian inference,” Decision Analytics Journal, vol. 6, p. 100174, 2023, doi: 10.1016/j.dajour.2023.100174.
6. S. Ayvaz and K. Alpay, “Predictive maintenance system for production lines in manufacturing: A machine learning approach using IoT data in real-time,” Expert Systems with Applications, vol. 173, p. 114598, 2021.

7. Y. Mahale, S. Kolhar, and A. S. More, “A comprehensive review on artificial intelligence driven predictive maintenance in vehicles: Technologies, challenges and future research directions,” Discover Applied Sciences, vol. 7, p. 243, 2025, doi: 10.1007/s42452-025-06681-3.
8. V. Mandala, B. M. S. R. Kuppala, S. N. R. D. Surabhi, and P. D. N. K. Kommisetty, “Advancing predictive failure analytics in automotive safety: AI-driven approaches for school buses and commercial trucks,” Journal of Artificial Intelligence and Big Data, vol. 1, pp. 9–20, 2022.
9. D. Divya, B. Marath, and M. Santosh Kumar, “Review of fault detection techniques for predictive maintenance,” Journal of Quality in Maintenance Engineering, vol. 29, pp. 420–441, 2023.
10. J. Sharma, M. L. Mittal, G. Soni, and A. Keprate, “Explainable artificial intelligence (XAI) approaches in predictive maintenance: A review,” Recent Patents on Engineering, vol. 18, no. 5, pp. 18–26, 2024.
11. J.-Y. Wu, M. Wu, Z. Chen, X.-L. Li, and R. Yan, “Degradation-aware remaining useful life prediction with LSTM autoencoder,” IEEE Transactions on Instrumentation and Measurement, vol. 70, pp. 1–10, 2021.
12. M. Pech, J. Vrchota, and J. Bednář, “Predictive maintenance and intelligent sensors in smart factory: Review,” Sensors, vol. 21, no. 4, p. 1470, Feb. 2021, doi: 10.3390/s21041470.
13. Ş. Yıldırım, A. D. Yücekaya, M. Hekimoğlu, M. Ucal, M. N. Aydin, and İ. Kalafat, “AI-driven predictive maintenance for workforce and service optimization in the automotive sector,” Applied Sciences, vol. 15, no. 11, p. 6282, 2025, doi: 10.3390/app15116282.
14. K. Rana and N. Khatri, “Automotive intelligence: Unleashing the potential of AI beyond predictive maintenance,” Computers in Industry, vol. 141, p. 103694, 2024.
15. M. N. Hossain, M. M. Rahman, and D. Ramasamy, “Artificial intelligence-driven vehicle fault diagnosis to revolutionize automotive maintenance: A review,” Computer Modeling in Engineering & Sciences, vol. 141, no. 2, pp. 951–996, 2024, doi: 10.32604/cmes.2024.056022.