International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 14, Issue 09, RTMSE-2026

Predictive Maintenance in Heavy Duty Vehicles Using Machine Learning

Divyam Singla[1], Sajal Rajoria[2] , Harshdeep Singh Sohal[3] , Dr Lakshmi Shankar[4]

Students, B. Tech., Mechanical Engineering, Punjabi University, Patiala

Assistant Professor, Department of Mechanical Engineering, Punjabi University, Patiala

*Corresponding Author E-mail: singladivyam09@gmail.com

ABSTRACT - Unplanned Equipment Downtime in High-Speed Automotive Manufacturing may cause a lot financial disruptions, making predictive maintenance techniques increasingly necessary in modern day production. With the help of the effectiveness of machine learning and artificial intelligence manufacturing plants can process abundant real time data collected from sensors and forecast remaining equipment lifespan, thus optimizing both maintenance costs and efficiency. This evaluates various algorithmic approaches that excel at fault detection and also focuses on data collection methods which are also one of the biggest challenges in implementing predictive maintenance. The methodology involves analyzing how embedded sensor networks monitor machinery health directly on the factory floor, by testing these models against industrial cycles and manufacturing datasets, this determines the most efficient algorithms for identifying mechanical anomalies in critical components and forecasting exact failure horizons without disrupting working of the vehicle. The successful deployment of predictive maintenance relies on prioritizing computationally efficient algorithms. Moving forward, the industry's focus must shift toward securing these expanding these methods across all tiers of manufacturing supply chain

1. INTRODUCTION

In the high-speed world of modern automotive manufacturing, just one hour of downtime can cost a lot of money. As production lines aim for faster cycles and more complexity, unplanned downtime has become a major source of financial strain and operational problems. Traditionally, manufacturers used reactive or scheduled maintenance. These methods often resulted in either major failures or replacing parts that were still functioning. However, the rise of Industry 4.0 has introduced a better solution: Predictive Maintenance (PdM). The strength of PdM is its ability to turn raw operational data into useful insights. By using Machine Learning (ML) and Artificial Intelligence (AI), manufacturing plants can process the large amounts of real-time data produced by embedded sensor networks. These technologies enable constant monitoring of machine health right on the factory floor, allowing the prediction of a component's Remaining Useful Life (RUL). This change from "fail and fix" to "predict and prevent" is crucial for reducing maintenance costs and ensuring uninterrupted production.

2. LITERATURE REVIEW:

Recent advancements in automotive engineering have shown how important predictive maintenance is for improving system reliability and reducing operational downtime. Traditional maintenance strategies, like reactive maintenance and scheduled maintenance, are being replaced by data-driven methods that use sensor technologies and machine learning algorithms.
Heavy Duty Vehicles and industrial automotive systems have multiple embedded sensors in components such as the engine, transmission, and braking systems. These sensors constantly monitor parameters like engine RPM, coolant temperature, oil pressure, vibration levels, and fuel injection rate. The data collected from these sensors offers valuable insights into the operational condition of mechanical components and helps identify early signs of system failure. Continuous data logging through control units, such as electronic control units (ECU) and transmission control modules (TCM), allows for the creation of historical datasets that can be used for predictive analytics and fault detection.

3. DATA COLLECTION:

Data Collection is already done through ecu, tcm which read rpm, coolant temperature, oil pressure, fuel rate, heavy duty vehicles we just continuously log all values over time, we may also add high frequency vibration accelerometers, brake wear sensors, etc.

So instead of maintenance is done not after breakdown but it is done before it. As Maintenance done after breakdown can be more costly and as well lead to more downtime.

Automotive predictive maintenance systems use vibration, acoustic, thermal, and electrical sensors to monitor machine health. However, the high volume of raw telemetry generated by modern production lines makes cloud-only processing impractical, necessitating edge-based data preprocessing to ensure low-latency anomaly detection and system reliability

Types of Sensors involved:

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 14, Issue 09, RTMSE-2026

Modern heavy duty vehicles are equipped with sensor systems integrated within critical subsystems such as engine transmission braking system, etc this sensors generate data essential for degradation analysis and fault prediction.

1. Engine and Power train sensors :

These sensors would monitor parameters like rpm, coolant temp, oil pressure, fuel injection rate, etc. These measurements would enable fault detection of abnormal combustion behaviour and thermal stress conditions.

2. Drive train and Vibration Monitoring:

High frequency accelerometers can be used to capture vibration signatures which would be helpful in faults like wear, misalignment and imbalance. These types of sensors could give useful data to reveal early stage degradation before any vibration. This also becomes necessary as vibration can lead to catastrophic, fatigue and sudden failure of a part which becomes a threat to both the machine and the operator.

3. Load Sensors:

Heavy duty vehicles frequently operate under high loads, as the name suggests and also faces variable road conditions

4. SHIFT TOWARDS AI

In modern automotive systems and production lines, machines are equipped with many sensors. These sensors continuously monitor important parameters such as vibration, temperature, pressure, and energy consumption. Traditionally, all this sensor data was sent to a centralized cloud server for analysis. The cloud would process the data, detect problems, and then send alerts or instructions back to the system.

However, this cloud-only approach creates several problems.

First, it introduces latency (delay). Data must travel from the machine to the cloud and then back again. In high-speed automotive production lines, even a small delay can be dangerous because machines operate very quickly. A few milliseconds of delay can lead to serious mechanical damage.

Second, sending huge amounts of raw sensor data to the cloud consumes massive network bandwidth. Modern systems generate enormous amounts of high-frequency data. Transmitting all of this data is inefficient and expensive.

Third, there are cybersecurity concerns. Sending sensitive operational data over networks increases the risk of cyber attacks or data breaches.

To solve these problems, the automotive industry is shifting toward Edge AI.

Edge AI means that data is processed locally, close to where it is generated. Instead of sending all raw data to the cloud, machine learning algorithms run directly on local devices such as robotic controllers, motor gateways, embedded processors, or programmable logic controllers (PLCs).

5. IMPLEMENTATION OF MACHINE LEARNING MODELS:

Machine Learning Models: They play a critical role in enabling predictive maintenance for heavy-duty vehicles. These models analyze historical and real-time sensor data to identify patterns associated with component degradation and failure. Depending on the prediction objective—classification, regression, or anomaly detection—different machine learning algorithms are applied.

5.1 Classification Models:

Classification models are used when the objective is to predict whether a component will fail within a certain time window. In heavy-duty vehicles, this can include predicting engine breakdown, gearbox malfunction, or brake failure.

Logistic Regression is one of the simplest classification techniques. It estimates the probability of failure using a sigmoid function. Although computationally efficient, it may not perform well with complex nonlinear relationships.

Random Forest is widely used in predictive maintenance due to its robustness and ability to handle nonlinear data. It consists of multiple decision trees, and the final prediction is determined by majority voting. Random Forest performs well with noisy sensor data and provides feature importance analysis, which is useful for interpretability.

Support Vector Machine (SVM) is effective in high-dimensional spaces and works well when there is a clear margin of separation between failure and normal states. However, it may require careful parameter tuning.

Gradient Boosting Algorithms, such as XGBoost, are powerful ensemble methods that iteratively improve model accuracy by correcting previous errors. These models often achieve high predictive performance in industrial applications.

5.2 Regression Models

Regression models are used when the goal is to estimate Remaining Useful Life (RUL) of a vehicle component.

Linear Regression models the relationship between sensor inputs and RUL as a linear function. While simple and

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 14, Issue 09, RTMSE-2026

interpretable, it may not capture complex degradation patterns.

Support Vector Regression (SVR) extends SVM principles to regression problems and is effective in handling nonlinear degradation trends.

Artificial Neural Networks (ANNs) are capable of modeling complex nonlinear relationships between sensor data and component health. They are particularly useful when large amounts of data are available.

5.3 Model Selection Considerations

The choice of machine learning model depends on:

- Availability of labelled data
- Type of prediction (classification or regression)
- Computational resources
- Requirement for interpretability
- Real-time deployment constraints

For most heavy-duty vehicle applications, ensemble models such as Random Forest and Gradient Boosting provide a good balance between accuracy and interpretability. For time-series RUL prediction, LSTM models are generally preferred.

.5.4 Making Model working:

Step1: Machine learning models do not work without proper data collection. The proper data collection has been discussed above. For example, our data collection must looks like in a table format having all the required sensing data such as Temperature, Vibration, RPM, Oil pressure.

Step2: The required library must be installed on our systems such as panda, numpy, scikit-learn, Matplotlib, Tensor flow (for deep learning, if using).

Step3: For research and simulation purposes our data must be stored in a particular manner.

- CSV file.
- SQL database
- Real Time streaming data.

This above process is called Data Loading and preprocessing. For this, we also have a line of code.

6. CONCLUSIONS:

Unplanned downtime is one of the biggest financial and operational challenges in high-speed automotive manufacturing. This paper shows that shifting from traditional reactive maintenance to data-driven Predictive Maintenance (PdM) is essential in the Industry 4.0 era. By using advanced sensor networks that capture important factors like vibration, thermal signatures, and powertrain metrics, manufacturers can gain real-time insights into machine health. The large amount and fast pace of telemetry data generated by modern automotive production lines make relying solely on centralized cloud processing ineffective. Moving to Edge AI allows for local data processing. This change reduces delays, saves bandwidth, and lessens cyber security risks. It helps ensure that the quick response times needed for high-speed manufacturing are achieved without compromising system integrity. Successful PdM implementation depends on choosing the right Machine Learning models. While simpler algorithms like Logistic Regression provide a starting point, more complex methods like Random Forest offer the strength needed to manage messy, non-linear industrial data. For accurately predicting a component's Remaining Useful Life (RUL), regression models and advanced neural networks are crucial. Looking ahead, the main challenge for the automotive industry will be not only detecting faults with algorithms but also securely and effectively deploying these Edge AI and Machine Learning systems across all levels of the manufacturing supply chain. Ultimately, combining continuous sensor monitoring and local data

REFERENCES:

[1] M. A. A. et al., "Secure Explainable-AI Approach for Brake Faults Prediction in Heavy Transport," IEEE Access, vol. 12, pp. 10638-10650, 2024.

[2] A. R. M. N. et al., "Cost-Optimised Machine Learning Model Comparison for Predictive Maintenance," MDPI Electronics, vol. 14, no. 12, 2025. (Note: This study specifically utilizes the SCANIA Component X heavy-duty truck dataset).

[3] R. Aravind, "Machine Learning Applications in Predictive Maintenance for Vehicles: Case Studies," International Journal of Engineering and Computer Science, vol. 10, no. 11, pp. 25628-25640, 2022.

[4] S. K. et al., "Predictive Maintenance of Lead-Acid Batteries with Sparse Vehicle Operational Data," International Journal of Prognostics and Health Management, vol. 12, no. 1, 2021.

[5] M. Lei, Z. Jiang, X. Pan, and C. Feng, "A review of edge computing in predictive maintenance," IEEE Access, vol. 8, pp. 161825-161838, 2020.

[6] R. Zhao, R. Yan, Z. Chen, K. Mao, P. Wang, and R. X. Gao, "Deep learning and its applications to machine health monitoring," Mechanical Systems and Signal Processing, vol. 115, pp. 213-237, 2019.

[7] A. K. S. Jardine, D. Lin, and D. Banjevic, "A review on machinery diagnostics and prognostics implementing condition-based maintenance," Mechanical Systems and Signal Processing, vol. 20, no. 7, pp. 1483-1510, 2006

[8] Theissler, Andreas, et al. "Predictive maintenance enabled by machine learning: Use cases and challenges in the automotive industry." *Reliability engineering & system safety* 215 (2021): 107864.