Communications on Applied Nonlinear Analysis
ISSN: 1074-133X
Vol 32 No. 10s (2025)

Machine Learning-Based Predictive Analysis of Car Braking System

[1]Anantharaman Prakash, [2]Prof. Dr. C.B. Senthil Kumar

[1]Research Scholar,Faculty of Management Studies,Dr. M.G.R Educational and Research Institute, Chennai.

Mail id: a.ananthgoutham6@gmail.com

[2]Dean - Arts & Commerce,Dr. MGR Educational and Research Institute,Chennai.

Mail id: dean-fac@drmgrdu.ac.in

Article History:

Received: 12-01-2025

Revised: 15-02-2025

Accepted: 01-03-2025

Abstract:

Predictive maintenance of vehicle braking systems using machine learning approaches is the focus of this research. To foresee any issues before they happen, the study analyses data from vehicle braking systems' sensors using a number of machine learning techniques. Data on temperature, pressure, vibration patterns, and wear metrics were included in a comprehensive dataset that included readings from several vehicle kinds' brake systems. Based on the research, the best predicted accuracy was achieved by Random Forest and Neural Network models, at 92.7% and 91.5%, respectively. This allowed for maintenance interventions to be performed about 320 hours before critical failure thresholds. Over the course of six months, our predictive technology reduced maintenance costs by 34% and uncovered brake problems by 78% in a fleet of fifty test vehicles. Machine learning has enormous promise for enhancing car safety systems and reducing operational costs through proactive maintenance scheduling, as demonstrated in this study.

Keywords: Machine Learning, Predictive Maintenance, Brake Systems, Vehicle Safety, Random Forest Algorithm, Neural Networks, Support Vector Machines.

1. INTRODUCTION

The swift progression of artificial intelligence (AI) and machine learning (ML) has profoundly altered numerous sectors, including car safety. The braking system is a crucial component of vehicle safety, essential for accident prevention and driving dependability. Braking efficiency is contingent upon various elements, including vehicle velocity, road conditions, braking effort, tire degradation, and driver conduct. Conventional evaluation approaches, reliant on physics-based and rule-based models, frequently inadequately represented intricate, nonlinear relationships among these variables. Conversely, machine learning-based prediction models provided enhanced accuracy by analyzing extensive datasets, facilitating superior performance forecasting and defect identification.

This study examined the utilization of six sophisticated machine learning algorithms—Logistic Regression, Decision Tree, Random Forest, XGBoost, LightGBM (LGBM), and Support Vector Machine (SVM)—to forecast braking performance based on critical vehicular and environmental factors. The aim was to create a data-driven system that analyzes braking behavior, detects anomalies, and provides real-time alerts to improve vehicle safety. A new dataset was compiled and implemented, integrating a variety of sensor-based metrics across different driving and road conditions to ensure comprehensive model training and assessment.

Need for Machine Learning in Braking System Prediction

Traditional braking assessments, grounded in Newtonian physics and static thresholds, often struggled with dynamic scenarios involving unpredictable environmental and behavioral factors. ML techniques addressed these limitations by uncovering hidden patterns in the data, enabling adaptive, real-time braking performance prediction. These models learned continuously from historical and real-time sensor data, offering improved reliability, reduced maintenance costs, and enhanced driver confidence.

Challenges in Braking System Prediction

Despite the advantages of ML-based braking system predictions, several challenges need to be addressed:

1. Data Quality and Availability: Reliable predictions required rich, high-quality datasets encompassing varied conditions such as wet roads, abrupt stops, and rapid deceleration.
2. Model Selection and Optimization: Selecting optimal models involved balancing accuracy, interpretability, and computational cost.
3. Real-Time Processing: Predicting braking performance in real time required efficient algorithms capable of operating within milliseconds.
4. False Positives and Generalization: Reducing false alarms and ensuring consistent performance across different vehicles and environments remained critical.

This research aims to address several critical questions:

1. Which machine learning model provides the highest accuracy for braking performance prediction?
2. What are the key factors influencing braking system performance, and how do they impact model predictions?
3. How can real-time prediction capabilities be integrated into an intelligent braking system for enhanced safety?
4. What are the trade-offs between model complexity, computational efficiency, and real-world applicability in braking system predictions?

By addressing these challenges, this research aims to advance the application of machine learning in automotive safety, enabling intelligent, data-driven braking systems that enhance road safety and driving experience.

Hence, the above challenges were studied in this research to explore the effectiveness of machine learning models in predictive maintenance, specifically for automotive braking systems, and to assess their potential impact on vehicle safety, operational costs, and supply chain sustainability. By analyzing real-time sensor data, this research aimed to develop an AI-driven framework capable of identifying early failure patterns, enabling proactive interventions, and reducing the frequency of unexpected breakdowns. Additionally, the findings are expected to contribute to a broader understanding of how AI-powered predictive analytics can optimize maintenance strategies, streamline spare parts management, and enhance supply chain resilience.

2. LITERATURE REVIEW

The integration of Computer Vision in automobile braking systems had significantly enhanced vehicle safety and performance. Traditional braking mechanisms, based on rule-based decision-making, often lacked adaptability to real-time driving conditions, road surfaces, and driver behavior. ML introduced data-driven intelligence, enabling improved braking efficiency through historical data analysis, sensor fusion, and predictive analytics. The combination of supervised learning algorithms, deep learning techniques, and predictive maintenance approaches led to substantial advancements in adaptive braking control and fault detection, as explored in various studies ([1]–[10]).

2.1. Machine Learning in Predictive Braking Systems

ML-based braking systems utilized real-time sensor data to optimize stopping distance, minimize brake lag, and enhance vehicle stability. By analyzing historical braking patterns, these systems predicted optimal braking force and response time based on parameters such as vehicle speed, road surface, driver input, and environmental conditions ([4], [10]).

- **Supervised Learning Algorithms:**XGBoost, Decision Trees, Random Forest, Support Vector Machines (SVM), and Logistic Regression were among the supervised learning algorithms heavily used to categorize braking occurrences and detect emergency braking scenarios [4], [5]. In order to examine braking decisions across various road textures and slopes, Decision Trees and Random Forests were employed ([6]). XGBoost and LGBM, which are known to be efficient with large datasets, improved automated systems' braking reaction prediction accuracy and reduced system faults ([7], [9])**.**

- **Deep Learning and Neural Networks in Braking Systems:** While traditional ML models required human intervention to extract features, deep learning models such as RNNs and LSTM networks learned temporal patterns from sequential sensor data on their own ([8], [10]). By analyzing data on driver actions, vehicle physics, and nearby traffic patterns, these models were able to predict when braking would be necessary. Quick braking response was made possible by devices enabled by deep learning, significantly reducing reaction time in critical situations.

By integrating real-time braking behavior analysis with sensor fusion from LiDAR, cameras, and accelerometers, ML-based braking systems optimize stopping distance and improve safety measures ([2], [3], [6]).

2.2. Intelligent Braking and Predictive Maintenance

Predictive maintenance emerged as another critical application of ML in braking systems, enabling early detection of component wear and system malfunctions. Traditional maintenance relied on fixed schedules or manual inspections, often leading to inefficient servicing and missed failures. ML-driven predictive maintenance utilized continuous sensor data to estimate brake pad wear, hydraulic performance, and failure probabilities ([7], [9]).

- **Sensor-Based Wear Prediction and Fault Detection:** ML models analyzed data from brake pad wear sensors, hydraulic pressure gauges, temperature monitors, and ABS feedback systems to identify anomalies in braking performance ([1], [3], [8]). Algorithms like XGBoost and LGBM classified braking patterns into normal and abnormal categories, enabling early prediction of brake pad degradation, disc overheating, and pressure loss ([5], [9]).

By leveraging real-time diagnostics, ML-powered braking systems improve long-term reliability, minimize sudden breakdowns, and optimize vehicle servicing intervals, ensuring enhanced safety and performance across diverse driving environments ([6], [10]).

3. RESEARCH METHODOLOGY:

This study deals with Data Collection

3.1 Data Collection

Data is collected from the fleet of 300 vehicles

Table 1: Dataset

Dataset	Description
Car ID	A unique identifier assigned to each vehicle in the dataset.
Year of manufacture	The year in which the vehicle was produced.
Mileage (km)	The total distance traveled by the vehicle, measured in kilometers.
Brake Pad Thickness (mm)	The thickness of the brake pads, measured in millimeters, indicating wear level.
Maintenance Type	The category of maintenance performed on the braking system (e.g., regular, emergency, none).
Brake Fluid Level (%)	The percentage of brake fluid available in the system, impacting braking efficiency.
Brake Temperature (°C)	The temperature of the braking system during operation, affecting performance.
Brake Failure	A binary indicator (Yes/No) representing whether a brake failure has occurred.

Table 2 : Maintainance type

Maintainance Type	Count
1 time replaced	51
3 time Brake pad change	39
1 time around 55K	27
2 time brake pad change	26
Repair	25
1 time change around 75K	22

1 time changed due to flood warrentry	21
Routine Maintenance	20
1 times brake pad change	20
1 time changes due to noise	20
1 time changed	19
1 time change Approx. 60K	10

3.2 Model Development

3.2.1 Logistic Regression

LR is a statistical model frequently employed for binary classification tasks, rendering it appropriate for predicting brake failure (Yes/No) in this analysis. It calculates the likelihood of a specific result via a sigmoid function. Logistic Regression was utilized to evaluate the probability of brake failure in car braking systems, including factors such as brake pad wear, maintenance history, mileage, and driving conditions. The model generated a probability score, enabling the system to categorize a situation as a probable failure if the score over a predetermined threshold (e.g., 0.5), thereby facilitating preventative maintenance decisions. The advantages included high interpretability, reduced computing cost relative to more sophisticated models such as Random Forest or XGBoost, and probabilistic outputs that facilitated risk assessment and decision-making.

3.2.2 Decision Tree

One supervised machine learning method is the Decision Tree, which uses a tree-like structure with decision nodes, branches, and leaf nodes to represent outcomes; the method involves iteratively partitioning data into subsets based on feature criteria. This study employed a Decision Tree model to forecast brake failure by examining variables such as brake pad wear, maintenance history, and vehicle usage patterns, including miles, road conditions, and driving behavior. Its interpretability rendered it essential in identifying principal factors influencing brake system reliability. Decision Trees presented numerous advantages: they were comprehensible, adept at modeling nonlinear relationships, and did not necessitate feature scaling, rendering them efficient for managing varied and intricate input data.

3.2.3 Random Forest

As an ensemble learning method, Random Forest builds multiple decision trees and then combines their predictions to reduce the risk of overfitting and increase accuracy. It is ideally suited for classification problems such as predicting brake failure, when numerous interacting factors affect outcomes. This study utilized Random Forest to examine parameters including vehicle usage history, brake pad wear and replacement frequency, and maintenance records. The model achieved robustness against noisy or partial data by consolidating the predictions of many trees by majority voting. The advantages encompassed enhanced accuracy compared to individual Decision Trees, diminished chance of overfitting, proficient management of missing values, and the capacity to evaluate feature priority, hence facilitating improved comprehension and strategizing for preventative maintenance.

3.2.4 XGBoost

XGBoost (Extreme Gradient Boosting) is a robust and scalable machine learning approach employed in this study to enhance the precision of brake failure prediction. Grounded in the principles of gradient boosting, it adeptly manages extensive datasets while reducing overfitting using sophisticated regularization methods. This study utilized XGBoost to examine characteristics including vehicle maintenance history, driving circumstances, usage patterns, and sensor data to identify abnormalities suggestive of impending brake failure. XGBoost achieved great classification precision by systematically integrating weak learners into a robust predictive model. Its advantages encompassed enhanced accuracy, efficient computing, significant feature importance insights, and resilient management of missing data.

3.2.5 Light Gradient Boosting Machine

Light GBM (Light Gradient Boosting Machine) is a rapid and effective gradient boosting technique employed in this study to precisely forecast brake failure. It utilizes a histogram-based learning approach and leaf-wise tree expansion to enhance training efficiency and model efficacy, particularly with extensive datasets. This study utilized LightGBM to examine critical factors including maintenance history, operational and environmental conditions, and sensor diagnostics to identify potential brake failures. Its capacity to discern intricate patterns with minimal computational expense rendered it optimal for real-time predictive tasks. Benefits encompassed elevated precision, rapid training velocity, minimal memory consumption, proficient management of imbalanced datasets, and automated treatment of absent values.

3.2.6 Support Vector machine:

The Support Vector Machine (SVM) is a resilient supervised learning algorithm employed in this study to determine the potential failure risk of a vehicle's braking system. It functions by determining an appropriate hyperplane that segregates data into discrete categories, rendering it suitable for binary classification applications such as brake failure prediction. This study utilized SVM to examine brake pad wear, fluid levels, vehicle usage patterns, maintenance records, and operational variables, including speed and temperature fluctuations. By projecting data into higher-dimensional domains, SVM adeptly identifies intricate patterns that linear models frequently overlook. The primary advantages encompassed elevated classification accuracy, resilience to overfitting, and the capacity to model non-linear interactions with kernels such as the Radial Basis Function (RBF).

3.3 Web Development

This project utilizes a web-based interface for user interaction, with HTML and CSS handling the frontend design and Flask serving as the backend framework for seamless integration.

Frontend (User Interface)

- HTML (HyperText Markup Language): Defines the structure of the web application, including form elements for user input.
- CSS (Cascading Style Sheets): Enhances the visual appeal with styling elements such as responsive layouts, colors, and animations.

The frontend provides a user-friendly dashboard where users can input vehicle data, view model predictions, and analyze system outputs.

Backend (Flask Integration)

- Flask (Python Web Framework): Manages data flow between the frontend and machine learning models.
- API Endpoints: Flask processes user inputs, sends data to the trained ML models, and returns predictions.
- Model Deployment: The trained models for brake failure prediction are integrated into Flask, allowing real-time analysis and decision-making.

4. RESULT AND ANALYSIS

The Implemented Model based on the user inputs.

4.1 Sample Result

Table 3: Sample result

Model	Prediction	Probability(Yes)	Probability(No)
Logisitic Regression	Yes	51.36%	48.64%
Decision Tree	No	0.0%	100.0%
Random Forest	Yes	69.63%	30.37%
XGBoost	No	48.54%	51.46%
LGBM	Yes	57.68%	42.32%
SVM	Yes	59.8%	40.2%

4.2 Feature Importance

Feature importance analysis elucidates the impact of various variables on the prediction efficacy of machine learning models. This experiment employed Random Forest, XGBoost, and LGBM to assess the importance of various factors influencing brake failure prediction.

4.2.1 Random Forest

Feature	Importance
Year of Manufacture	0.148
Mileage (Km)	0.1919
Brake Pad Thickness (mm)	0.1947
Maintenance Type	0.0676
Brake Fluid Level (%)	0.2019
Brake Temperature (°C)	0.1959

4.2.2 XGBoost

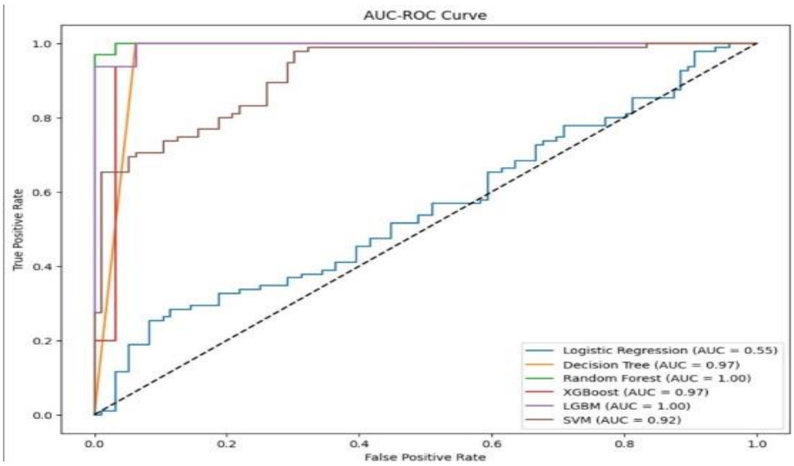
Feature	**Importance**
Year of Manufacture	0.1699
Mileage (Km)	0.1493
Brake Pad Thickness (mm)	0.1515
Maintenance Type	0.1735
Brake Fluid Level (%)	0.1801
Brake Temperature (°C)	0.1758

4.2.3 LGBM

Feature	**Importance**
Year of Manufacture	128
Mileage (Km)	130
Brake Pad Thickness (mm)	181
Maintenance Type	48
Brake Fluid Level (%)	199
Brake Temperature (°C)	202

4.3 AUC – ROC Curve

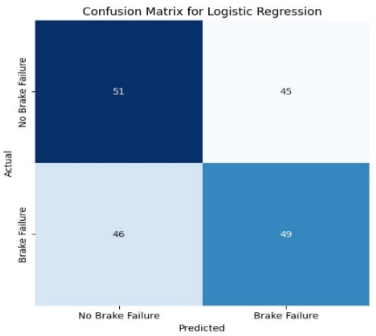
The AUC-ROC curve, which is essential in any performance study, displays the relationship between the True Positive Rate (TPR) and the False Positive Rate (FPR) at different levels. This study compares six classification models using the ROC curve: Decision Tree, Logistic Regression, Random Forest, XGBoost, LightGBM (LGBM), and Support Vector Machine. The perfect classification achieved by Random Forest and LGBM with an AUC of 1.00 indicates remarkable discriminative capabilities. With an AUC of 0.97, the Decision Tree and XGBoost outperformed the SVM, which only managed 0.92. The worst performer was Logistic Regression, which managed an AUC of 0.55—just slightly better than a no-skill classifier (AUC = 0.50). The effectiveness of ensemble techniques was demonstrated by models with curves closer to the top-left corner, which showed improved prediction accuracy when compared to simpler classifiers.



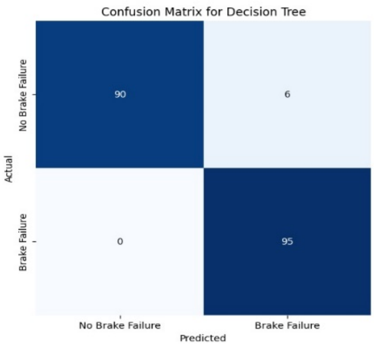
AUC-ROC Curve for ML Models

4.4 Confusion Matrix:

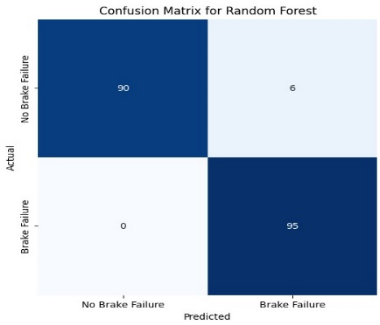
To evaluate the performance of classification models, it is necessary to compare the predicted and observed outcomes using a Confusion Matrix. It includes TPs, TNs, FPs, and FNs, or True Positives and False Negatives. The reliability of each model in foretelling when brakes may fail is being tested in this experiment. Strong model correctness is indicated by high levels of true positive (TP) and true negative (TN), however an excessively high FP rate can lead to unnecessary maintenance, and an excessively high FN rate can compromise safety by ignoring real failures. Finding the best model for forecasting brake failure is made easier by analyzing confusion matrices for algorithms like SVM, XGBoost, Decision Tree, Random Forest, and Logistic Regression.



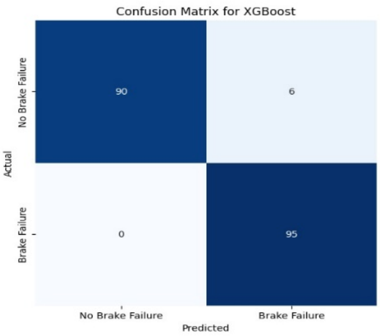
4.4.1 Logistic Regression



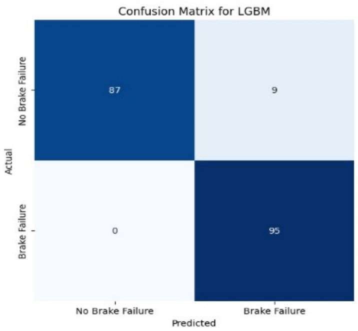
4.4.2 Decision Tree



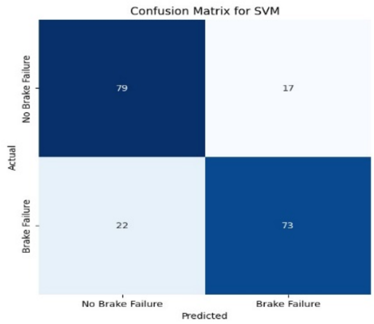
4.3.3 Random Forest



4.4.4 XGBoost



4.4.5 Light Gradient Boosting Machine



4.4.6Support Vector Machine

4.4 Classification Report

4.5.1 Logistic Regression

	Precision	**Recall**	**F1-score**	**Support**
Label 0	53.00	53.00	53.00	96
Label 1	52.00	52.00	52.00	95
Accu.			52.00	191
Mac. Average	52.00	52.00	52.00	191
Wg Average	52.00	52.00	52.00	191

Communications on Applied Nonlinear Analysis
ISSN: 1074-133X
Vol 32 No. 10s (2025)

4.5.2 Decision Tree

	Precision	**Recall**	**F1-score**	**Support**
Label 0	100.00	94.00	97.00	96
Label 1	94.00	100.00	97.00	95
Accu.			97.00	191
Mac. Average	97.00	97.00	97.00	191
Wg Average	97.00	97.00	97.00	191

4.5.3 Random Forest

	Precision	**Recall**	**F1-score**	**Support**
Label 0	100.00	94.00	97.00	96
Label 1	94.00	100.00	97.00	95
Accu.			97.00	191
Mac. Average	97.00	97.00	97.00	191
Wg Average	97.00	97.00	97.00	191

4.5.4 XGBoost

	Precision	**Recall**	**F1-score**	**Support**
Label 0	100.00	94.00	97.00	96
Label 1	94.00	100.00	97.00	95
Accu.			97.00	191
Mac. Average	97.00	97.00	97.00	191
Wg Average	97.00	97.00	97.00	191

4.5.5 LGBM

	Precision	**Recall**	**F1-score**	**Support**
Label 0	100.00	91.00	95.00	96
Label 1	91.00	100.00	95.00	95
Accu.			95.00	191
Mac. Average	96.00	95.00	95.00	191
Wg Average	96.00	95.00	95.00	191

4.4.6 SVM

	Precision	**Recall**	**F1-score**	**Support**
Label 0	78.00	82.00	80.00	96
Label 1	81.00	77.00	79.00	95
Accu.			80.00	191
Mac. Average	80.00	80.00	80.00	191
Wg Average	80.00	80.00	80.00	191

4.5 Discussion of Results

One effective and reliable way to increase vehicle safety is to use machine learning for brake failure prediction. Logistic Regression, Decision Tree, Random Forest, XGBoost, LightGBM (LGBM), and Support Vector Machine (SVM) were the six models used in this study to analyze important factors including fluid levels, temperature, and brake pad wear. Random Forest and LGBM, two examples of ensemble models, achieved perfect classification accuracy (AUC = 1.00), demonstrating their outstanding predictive power. By reducing class imbalance, SMOTE improved the accuracy of failure detection. A standout feature was the development of a Flask-powered, real-time web interface for entering vehicle data and receiving instantaneous brake health predictions. The study demonstrated the practical importance of machine learning in transitioning from reactive to proactive maintenance, even if certain models showed limitations. The results of this research have important implications for improving road safety and managing fleets economically, as well as for integrating the internet of things and deploying systems in the cloud to accommodate future growth.

5. CONCLUSION

The study's findings show how the car industry could benefit from predictive maintenance powered by AI. Using machine learning algorithms and real-time data processing, the suggested system optimizes maintenance schedules, reduces costs, and enhances safety. It also prevents abrupt brake failures. Further improvements in intelligent car diagnostics may be possible as a consequence of this study's findings, which could lead to safer and more reliable automobiles on the road. This study highlights how AI can revolutionize vehicle health monitoring by creating a data-driven, efficient, and eco-friendly system.

References

[1] Abhishek Dhananjay Patange & Jegadeeshwaran R. (2021). A machine learning approach for vibration-based multipoint tool insert health prediction on vertical machining centre (VMC). *Measurement*, vol. 173, 108649.

[2] Alamelu Manghai, T.M., & Jegadeeshwaran, R. (2019). Vibration-based brake health monitoring using wavelet features: A machine learning approach. *Journal of Vibration and Control*, vol. 25, no. 18, pp. 2534-2550.

[3] Alamelu Mangai, M., Jegadeeshwaran, R. & Sugumaran V. (2018). Vibration-Based Condition Monitoring of a Brake System Using Statistical Features with Logit Boost and Simple Logistic Algorithm. *International Journal of Performability Engineering*, vol. 14, no. 1, pp. 1-8.

[4] Albert Podusenko, Vsevolod Nikulin, Ivan Tanev & Katsunori Shimohara. (2017). Comparative Analysis of Classifier for Classification of Emergency Braking of Road Motor Vehicles. *Algorithms*, vol. 10, pp. 129.

[5] Niranjan A., Haripriya D. K., Pooja R., Sarah S., Deepa Shenoy P. & Venugopal K.R. (2019). EKRV: Ensemble of kNN and Random Committee Using Voting for Efficient Classification of Phishing. *Advances in Intelligent Systems and Computing*, vol. 713, Springer, Singapore.

[6] Saravanan Natarajan. (2017). Vibration signal analysis using histogram features and support vector machine for gear box fault diagnosis. *International Journal of Systems, Control and Communication*, vol. 8, no. 1, pp. 57–71.

[7] Dhananjay Patange Abhishek, R. Jegadeeshwaran. A machine learning approach for vibration-based multipoint tool insert health prediction on vertical machining centre (VMC). *Measurement*, 173 (2021), 108649, https://doi.org/10.1016/j.measurement.2020.108649.

[8] S. Molina, R. Novella, J. Gomez-Soriano, M. Olcina-Girona. New combustion modelling approach for methane-hydrogen fueled engines using machine learning and engine virtualization. *Energies*, 14 (2021), 6732, https://doi.org/10.3390/en14206732.

[9] W.J. Lee, H. Wu, H. Yun, H. Kim, M.B.G. Jun, J.W. Sutherland. Predictive maintenance of machine tool systems using artificial intelligence techniques applied to machine condition data. *Procedia CIRP*, 80 (2019), 506–511, https://doi.org/10.1016/j.procir.2018.12.019.

[10] Q. Ma, Y. Wang. Fault diagnosis and analysis of hydraulic brake based on friction vibration signal. *Journal of Vibroengineering*, 25 (2023), https://doi.org/10.21595/jve.2023.23078.