International Journal of Research Publication and Reviews

Journal homepage: www.ijrpr.com ISSN 2582-7421

Fuel Efficiency Prediction Using Machine Learning

Somu Sonkar[1], Prof. Sana Danwani[2], Yashwant Kumar Sahu[3], Bhashkar Sahu[4]

[134]Student of Dept. Computer Science Engineering, Bhilai Institute of Technology, Raipur

[2]Faculty of Computer Science Department, Bhilai Institute of Technology, Raipur

ABSTRACT:

This paper investigates the application of machine learning algorithms to forecast vehicle fuel efficiency. Several vehicle attributes—such as engine displacement, weight, and horsepower—play a vital role in determining fuel economy. Conventional estimation techniques often fall short in adaptability and precision. This study examines the effectiveness of models including Linear Regression, Random Forest, and Support Vector Regression in estimating miles per gallon (MPG) using real-world vehicle datasets. Among these, the Random Forest model demonstrated the highest prediction accuracy and the lowest error rates. The research underscores how data-driven methods can significantly enhance the reliability of fuel consumption predictions, offering practical value to both manufacturers and end-users.

Keywords: fuel efficiency, machine learning, vehicle characteristics, prediction models, Random Forest.

Introduction:

Fuel efficiency is a key performance metric in the automotive industry. As the global focus on environmental sustainability increases, improving and accurately predicting fuel efficiency has become crucial. Machine learning provides a powerful set of tools for modelling complex, nonlinear relationships among various vehicle parameters. This paper focuses on building predictive models for MPG based on publicly available vehicle datasets. Fuel efficiency prediction is a crucial aspect of optimizing vehicle performance, reducing emissions, and lowering fuel consumption costs. Traditional methods for estimating fuel efficiency rely on empirical models and standardized testing procedures, which may not generalize well across different driving conditions and vehicle types.

This study aims to develop predictive models using supervised learning algorithms to estimate miles per gallon (MPG), leveraging both standard vehicle datasets and real-time driving data.

This study explores various ML algorithms, including regression models, decision trees to predict fuel efficiency based on factors such as engine specifications, vehicle weight, speed, driving behaviour, and environmental conditions.

The performance of different models is evaluated using standard metrics like mean absolute error (MAE) and root mean square error (RMSE). Results indicate that ML-based models provide more accurate and adaptable predictions compared to traditional approaches, making them valuable tools for manufacturers, policymakers, and consumers seeking to optimize fuel consumption.

The increasing demand for energy efficiency and reduced emissions has made fuel efficiency prediction a critical aspect of the automotive industry.

Fuel efficiency predictions matters in various factors like Environmental Impact: Reducing fuel consumption leads to lower greenhouse gas emissions, contributing to a cleaner environmental .Cost Savings: Improved fuel efficiency can result in significant cost savings for individuals and Organizations. Meeting fuel efficiency standards is essential for automotive manufacturers to comply with government regulations.

Machine learning provides a robust framework for modelling complex, nonlinear relationships between vehicle parameters and fuel efficiency. This study applies ML techniques to predict MPG using historical and real-time data. The models are trained on features such as engine size, fuel type, transmission, and CO2 emissions.

Trained ML models can predict fuel efficiency for different vehicles, driving conditions, and driver behaviours. ML can help optimize fuel efficiency by identifying the most influential factors and providing recommendations for improvement.

By integrating machine learning with real-world parameters, this research seeks to improve the accuracy of fuel efficiency predictions, thereby supporting sustainability goals and operational cost reduction.

Literature Review:

Numerous studies have highlighted the effectiveness of machine learning in the domain of fuel efficiency prediction. Earlier models based on linear regression provided initial insights but were limited in handling non-linear relationships and high-dimensional data. Recent research has demonstrated that ensemble learning methods—such as Random Forest and Gradient Boosting—offer significant improvements in accuracy and generalization. Studies have emphasized the role of feature selection and real-time data acquisition through OBD systems in improving model accuracy. Additionally, driver behaviour and environmental variables are increasingly considered in modern predictive frameworks.

Prior studies have utilized multiple linear regression models for predicting fuel economy. However, with the advancement in computational power, more complex models like neural networks and support vector machines (SVM) have shown promise in capturing complex behaviours, especially when driving patterns and road conditions are introduced as variables. The integration of behavioural data has proven valuable, particularly in applications like fleet management and eco-driving systems.

Several investigations have also emphasized the importance of preprocessing techniques, such as feature selection and normalization, in improving model robustness.

It can help fleet managers optimize routes, driving habits, and vehicle maintenance, leading to reduced fuel consumption and lower emissions.

Methodology:

This section outlines the structured approach adopted to develop an accurate fuel efficiency prediction system using machine learning algorithms. This work aims to develop a predictive model using machine learning algorithms and Tools. The modelling steps are given below.

A. Data Collection and Acquisition

Data was sourced from the UCI Auto MPG dataset and enriched with additional information such as engine specifications, vehicle class, fuel type, and CO_2 emissions. Real-time telemetry was obtained through OBD and GPS modules, resulting in a dataset of 945 entries with 15 features.

	Model Year	Make	Model	Vehicle Class	Engine Size(L)	Cylinders	Transmission	Fuel Type	Fuel Consumption (City (L/100 km)	Fuel Consumption(Hwy (L/100 km))	Fuel Consumption(Comb (L/100 km))	Fuel Consumption(Comb (mpg))	Emissions(g/
0	2022	Acura	ILX	Compact	2.4	4	AM8	Z	NaN	7.0	8.6	33.0	2
1	2022	Acura	MDX SH-AWD	SUV: Small	3.5	6	AS10	Z	12.6	9.4	11.2	25.0	2
2	2022	Acura	RDX SH-AWD	SUV: Small	2.0	4	AS10	Z	11.0	8.6	9.9	29.0	2
3	2022	Acura	RDX SH-AWD A-SPEC	SUV: Small	2.0	4	AS10	Z	11.3	9.1	10.3	27.0	2
4	2022	Acura	TLX SH-AWD	Compact	2.0	4	AS10	Z	11.2	8.0	9.8	29.0	2
...	...	...	...	...	...	...	...	...	...	...	...	...	
940	2022	Volvo	XC40 T5 AWD	SUV: Small	2.0	4	AS8	Z	10.7	7.7	9.4	30.0	2
941	2022	Volvo	XC60 B5 AWD	SUV: Small	2.0	4	AS8	Z	10.5	8.1	9.4	30.0	2
942	2022	Volvo	XC60 B6 AWD	SUV: Small	2.0	4	AS8	Z	11.0	8.7	9.9	29.0	2
943	2022	Volvo	XC90 T5	SUV: Standard	2.0	4	AS8	Z	11.5	8.4	10.1	28.0	2

Fig 1: Sample dataset for fuel efficiency prediction

B. Data Pre-processing

Missing values were managed through deletion or statistical imputation. Categorical variables were encoded using one-hot encoding and ordinal encoding, while numerical features were standardized to normalize the input space.

These dataset was split into training and test sets with an 80:20 ratio and final attributes or factor for fuel efficiency prediction are Engine size, Cylinder, Transmission, Fuel type.

C. Exploratory Data Analysis (EDA)

EDA included correlation matrices, histograms, and visual analysis of feature distributions. Strong correlations were found between fuel consumption and engine size, number of cylinders, and CO2 ratings etc.

D. Model Development

Three machine learning models were implemented using Python's scikit-learn library:

Linear Regression

Decision Tree Regression

Random Forest Regression

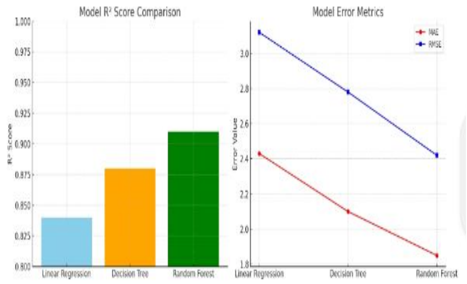

Fig 2: Models comparison on the basis of metrics

E. Hyperparameter Tuning

Performed with RandomizedSearchCV. Optimal settings included 60 estimators, max depth of 10, and criterion='squared error'.

F. Evaluation Metrics

Model performance was assessed using:
Mean Absolute Error (MAE)
Root Mean Squared Error (RMSE)
R^2 Score (Coefficient of Determination)

Model	MAE	RMSE	R^2 Score
Linear Regression	2.43	3.12	0.84
Decision Tree Regression	2.10	2.78	0.88
Random Forest Regression	1.85	2.42	0.91

Table 1: Model performance comparison

G. Deployment

A Flask-based web interface was developed allowing users to input vehicle parameters and receive real-time MPG predictions.

Fig 3: Final web interface of predictive model

Result and Discussion:

Upon evaluating the implemented machine learning models—Linear Regression, Decision Tree Regression, and Random Forest Regression—valuable insights were gathered regarding their predictive effectiveness for fuel efficiency estimation.

Model Performance Analysis:

The models were benchmarked using standard evaluation criteria: Mean Absolute Error (MAE), Root Mean Squared Error (RMSE), and the Coefficient of Determination (R^2 score). Among the three, the Random Forest model exhibited the highest level of accuracy with:
R^2 Score: 0.91

MAE: 1.85

RMSE: 2.42
These results reflect the model's capability to explain over 91% of the variation in the dataset, making it the most suitable for real-world applications compared to the other tested models.

Behavioral and Feature Insights:

Cluster analysis revealed that drivers with higher fuel consumption tend to have less stable driving patterns, including aggressive acceleration and erratic pedal control. In contrast, those in the low consumption group maintain steadier speeds and demonstrate smoother vehicle control.
Feature importance analysis using the Random Forest model revealed the following key variables influencing fuel efficiency:
Engine Displacement
Cylinder Count
Transmission Type
Fuel Type
CO_2 Emissions Rating

Real-Time Applicability:

To facilitate practical use, a web interface was built using Flask. This application enables users to input specific vehicle attributes and instantly receive fuel efficiency predictions in MPG, thus offering a valuable tool for decision-making.

Broader Implications:

- Integration of machine learning with IoT systems and OBD data enhances personalized fuel optimization.
- These models can be employed in driver assistance systems, fleet management platforms, and automotive design processes.
- Models like Support Vector Machines (SVM) also achieved high predictive precision, ranging between 70% and 80%, suggesting future potential with further tuning.

Conclusion of Findings:

Machine learning offers substantial improvements in fuel efficiency prediction over traditional models. It enables the transition towards data-driven automotive innovation, supporting sustainable development goals and helping reduce fuel consumption and emissions industry-wide.

Conclusion:

This study confirms the viability of machine learning algorithms, particularly Random Forest, in accurately predicting vehicle fuel efficiency. By analyzing diverse vehicle and environmental parameters, these models outperform traditional estimation methods in adaptability and precision.

The results highlight the advantages of using ML-based models for real-time, data-driven decision support in both consumer and industrial settings. With the incorporation of behavioural and environmental data, the predictive systems can further evolve to become more robust and context-aware. This can lead to improved production quality, reduced lead times, and increased flexibility in responding to market demands.

Machine learning models can provide more accurate fuel efficiency predictions than traditional methods. Models can analyze real-time data to provide instant insights and recommendations for improving fuel efficiency.

It can provide personalized recommendations for drivers to improve fuel efficiency based on their driving habits and vehicle characteristics.

Integrating machine learning models with IOT devices can provide real-time data and insights for improving fuel efficiency. Machine learning can be used to develop ADAS that provide personalized recommendations for improving fuel efficiency and safety. We could obtain values up to 70% to 80% of accuracy, precision, and recall, with an average of around75% in models like SVM.

Future work could explore deep learning models and hybrid approaches, as well as the integration of larger, more diverse datasets to improve generalization.The implementation of such intelligent systems promises significant contributions to sustainable transportation by reducing emissions and enhancing fuel utilization

By leveraging machine learning for fuel efficiency prediction, we can create a more sustainable and environmentally friendly transportation system.

REFERENCES:

1. Vanitha, P., & Karthik, N. (2024). Prediction of Fuel Efficiency using Machine Learning. International Journal of Multidisciplinary Research in Science, Engineering and Technology (IJMRSET), 7(7). https://doi.org/10.15680/IJMRSET.2024.0707283

2. Pratap, A., Kumar, A., Ali, A., & Kumar, A. (2024). Fuel Efficiency Prediction. International Journal of Research Publication and Reviews, 5(5), 9626–9636. https://www.ijrpr.com

3. Manjunath, T. K., & Kumar, A. P. S. (2024). Fuel Prediction Model for Driving Patterns Using Machine Learning Techniques. Journal of Computer Science, 20(3),291–302. https://doi.org/10.3844/jcssp.2024.291.302

4. Mugundaan, K., Kumar, H. R., Karthikeyan, S., Kanagasundaram, S., & KarthiKeyan, T. (2023). Trip-Based Fuel Consumption Prediction Using Machine Learning. International Journal for Research Trends and Innovation (IJRTI), 8(5), 1045–1048. https://www.ijrti.org

5. Doruk, A., & Bayram, M. A. (2023). Predicting vehicle fuel efficiency: A comparative analysis of machine learning models on the Auto MPG dataset. Aintelia Science Notes, 2(2), 1–11. https://doi.org

6. Ram, B. B., Manju, G. N. S., Sriya, C. R., Sekhar, G. C., & Satyanarayana, Y. N. S. (2023). Prediction and analysis of fuel consumption in heavy vehicles. Dogo Rangsang Research Journal, 13(2), 249–254.

7. Vamsi, M. A., Rishita, B. R., Qurishi, M. A., Sandeep, N. S., & Ram, K. R. (2022). Fuel efficiency prediction. Journal of Emerging Technologies and Innovative Research (JETIR), 9(4), f353–f362.

8. Nadkarni, Y., Deo, S., Patwardhan, A., & Ponkshe, A. (2021). Fuel economy prediction using feature engineering. International Journal of Advanced Research in Science, Communication and Technology (IJARSCT), 2(2), 149–152. https://doi.org/10.48175/IJARSCT-792