



International Journal of Innovative Research in Science Engineering and Technology (IJIRSET)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)



Impact Factor: 8.699

Volume 14, Issue 12, December 2025

Predictive Analytics for Fuel Consumption Optimization using Data Science Techniques

P Shirisha¹, Maheshwari M Desai²

PG Student, Dept. of MCA, City Engineering College, Bengaluru, India¹

Assistant Professor, Dept. of MCA, City Engineering College, Bengaluru, India²

ABSTRACT: Fuel consumption optimization has become essential due to rising fuel costs, increasing energy demand, and environmental concerns. Inefficient fuel usage in transportation and industrial sectors leads to higher operational costs and increased carbon emissions. Traditional fuel management approaches are static and ineffective for large and dynamic datasets. This project presents a predictive analytics-based framework using data science This as well as initiatives like machine learning techniques. Historical vehicle data, driving behavior, and environmental conditions are analysed after data cleaning, normalization, and feature selection. Regression models are applied to accurately predict fuel consumption. Model performance is evaluated using MAE, MSE, and R-squared metrics. The predictive insights support efficient fuel planning and sustainable transportation.

KEYWORDS: Fuel Consumption Optimization, Analytical prediction, Data and Machine Learning, Regression Analysis, Energy Management, Sustainable Transportation.

I. INTRODUCTION

The amount of fuel used is a growing challenge due to transportation growth, rising fuel costs, and environmental concerns. Excessive fuel usage increases operational expenses and carbon emissions, making efficient management essential. Traditional methods are rule-based and cannot adapt to dynamic real-world conditions. Fuel consumption depends on factors like vehicle type, engine capacity, load, driving behavior, and environment. This project proposes a predictive analytics-based system using data science techniques. Historical fuel data is pre-processed with cleaning, feature selection, and scaling, and analysed through Exploratory Data Analysis. Machine learning regression models predict fuel consumption and identify key factors affecting efficiency. A Streamlit interface enables easy input and visualization, making the system a practical tool for sustainable fuel management.

II. LITERATURE SURVEY

1) Title: Fuel Usage Making Forecasts Using Machine Learning Techniques

Authors: Mohammad A. Rahman, S. Ahmed, R. Hasan

Abstract: This paper presents a machine learning-based approach to forecast fuel usage by Using Vehicle and engine-related parameters. Different models of regression are evaluated to Ascertain the most precise prediction method. The study demonstrates that data-driven Models significantly outperform. Traditional estimation techniques, helping reduce fuel Consumption and operational costs.

2) Title: A Data-Driven Approach for Vehicle Fuel Efficiency Analysis

Authors: Li Wang, Chen Yu, Zhang Wei

Abstract: The authors propose a data-driven framework to analyze fuel efficiency based on Vehicle Specifications and driving conditions. Statistical analysis and predictive modeling are Used to identify Key factors affecting fuel usage. The results highlight the significance of the Engine size, vehicle Weight, and driving behaviour in fuel consumption optimization.

3) Title: Predictive Analytics for Fuel Usage within Transportation Systems

Authors: John Miller, Kevin Brown, Emily Johnson

Abstract: This research focuses on applying predictive analytics techniques to transportation

Datasets to forecast fuel consumption. Regression and group learning models are employed To improve Prediction Accuracy. The study emphasizes the role of predictive systems in Supporting energy Efficient Transportation planning.

4) Title: Machine Learning-Based Estimating Vehicle Fuel Consumption

Authors: Ahmed Al-Masri, Noor Khalid

Abstract: This document investigates algorithms for supervised machine learning Estimating vehicle Fuel Consumption. Feature engineering and model evaluation are Performed to enhance Prediction accuracy. The proposed approach assists users in Understanding fuel efficiency under different operating Conditions.

III. METHODOLOGY

The methodology focuses on predicting and optimizing fuel consumption using Data science as well as machine as a learning techniques. Initially, fuel consumption datasets are collected containing parameters such as vehicle type, engine specifications, distance travelled, speed, load conditions, and driving behavior. The collected data is pre-processed through cleaning data and dealing with missing values, outlier removal, feature selection, and scaling to improve data quality. Regression models for machine learning like as Decision-making and linear regression Tree, as well as Random the forest are then trained on historical data to learn fuel consumption patterns. Model performance is assessed using standard metrics for accuracy to select the best-performing model. The selected model predicts fuel consumption for new operational conditions. Prediction results are analysed to identify factors contributing to excessive fuel usage. Optimization strategies such as efficient driving behavior, speed control, and timely maintenance are suggested. Finally, results are visualized through a Streamlit-based user interface for easy interpretation and decision-making.

IV. SYSTEM DESIGN

The system is designed with a modular and scalable architecture for efficient fuel consumption prediction and optimization. It consists of six core modules: data collection, data pre-processing, model training, prediction engine, optimization analysis, and visualization interface. The data collection module gathers historical fuel consumption data from datasets and vehicle logs. The pre-processing module cleans the data, handles missing values, removes outliers, and applies feature selection and normalization. The model training module makes use of a machine learning algorithms such as decision-making, linear regression & Unpredictable, with performance evaluation on instruction and assessment datasets. The prediction engine consumption for estimations for new fuel user inputs in real time. The optimization analysis module identifies factors causing high fuel usage and provides actionable recommendations. A Streamlit-based visualization interface allows users to input data and view predictions through interactive graphs and charts. Overall, the system ensures accuracy, scalability, and effective decision support for fuel optimization.

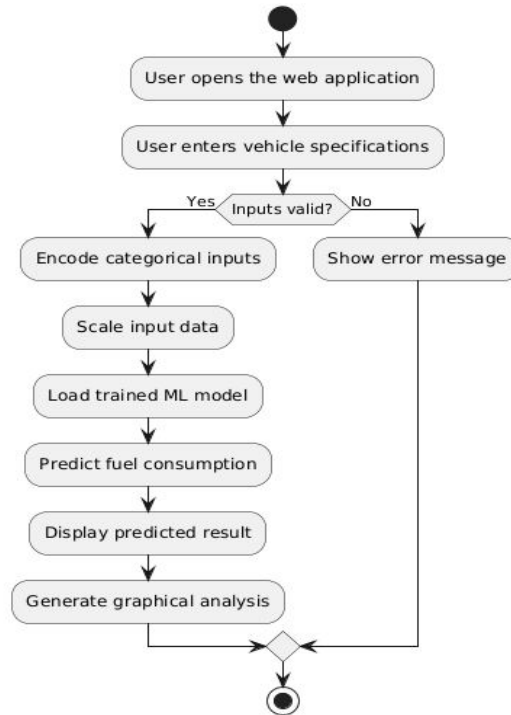


Fig 1: Activity Diagram

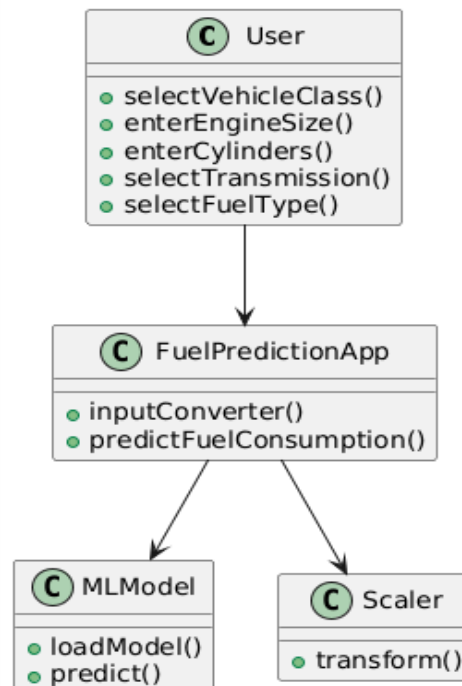


Fig 2: Class Diagram

V. SYSTEM ARCHITECTURE & DESIGN

The proposed system follows a modular and layered architecture for predicting and optimizing fuel consumption using data science techniques. The data layer collects historical fuel consumption data from vehicle logs and datasets, including vehicle type, engine capacity, distance travelled, speed, and load conditions. The pre-processing layer cleans and transforms raw data by managing and eliminating outliers, missing values, and applying feature scaling and the machine learning layer implements regression algorithms such as Decision-making and regression, as well as Random Forests to learn fuel consumption patterns. The prediction layer generates fuel consumption estimates for new input conditions using trained models. The optimization layer analyses prediction results to identify factors causing high fuel usage and suggests efficiency improvements. The presentation layer is developed using Streamlit to provide user interaction and visualization. Users can view predictions through charts and graphs for better understanding. Overall, the architecture ensures scalability, accuracy, and effective decision support for fuel consumption optimization.

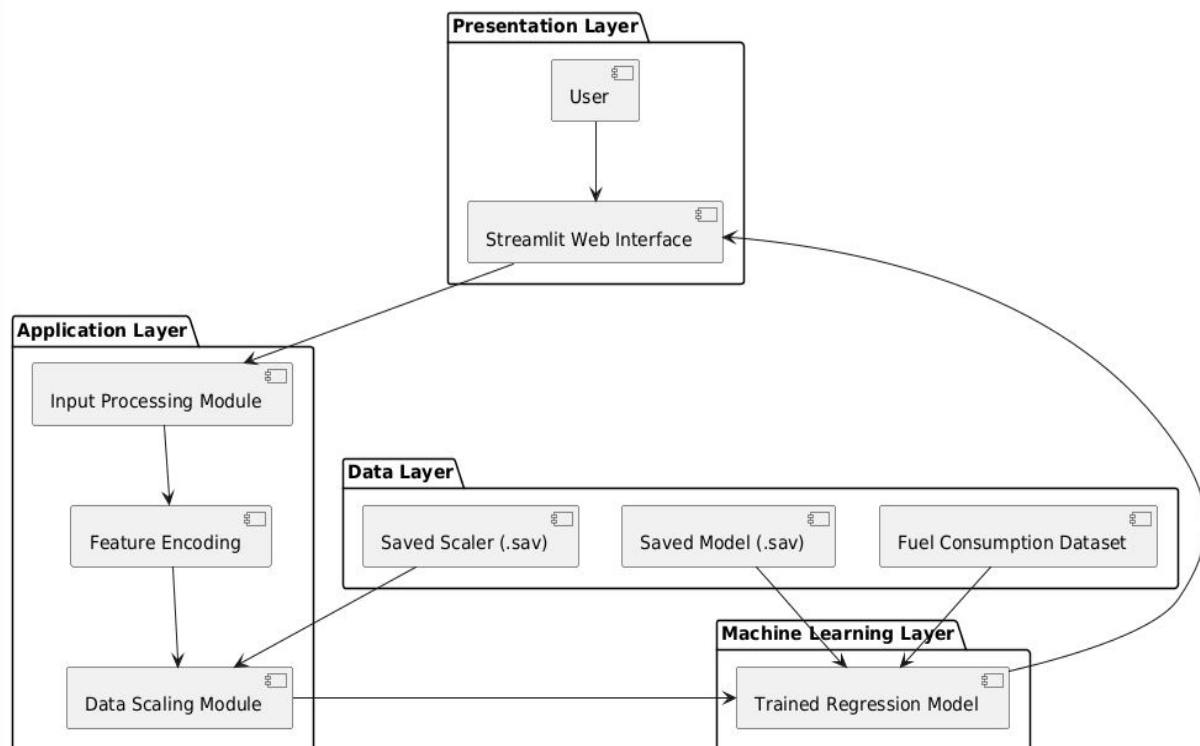


Fig 3: System Architecture & Design

VI. IMPLEMENTATION

The fuel consumption optimization system is implemented using Python and data science libraries in a modular manner. Fuel consumption datasets are collected in CSV format and processed using Pandas and NumPy. Data pre-processing comprises managing missing values, eliminating outliers, feature selection, and normalization. Models for machine learning such as regression with Linear, Decision Tree, in addition to Random Forest are put into practice utilising Scikit-learn. The dataset is split into instruction and testing sets and evaluated using MAE, MSE, and R-squared metrics. The best-performing model is selected and saved for future use. During prediction, the trained model estimates fuel consumption based on user inputs. A Streamlit-based interface enables interactive input, prediction, and visualization of results.

VII. RESULT AND DISCUSSION

The fuel consumption optimization system was successfully implemented and evaluated using real-world vehicle datasets. Data pre-processing, feature encoding, normalization, and Regression models for machine learning were used for accurate fuel consumption prediction. Various vehicle parameters such as engine size, fuel, gearbox type, and cylinders kind and vehicle class were considered. The trained model produced consistent and reliable predictions with very little divergence from real values. Graphical visualizations highlighted key trends and correlations influencing fuel efficiency. The results showed that machine learning-based prediction outperforms traditional estimation methods. A Streamlit-based interface enabled real-time prediction and user-friendly interaction. Overall, the results confirm the effectiveness of predictive analytics in optimizing fuel consumption and supporting sustainable fuel management.

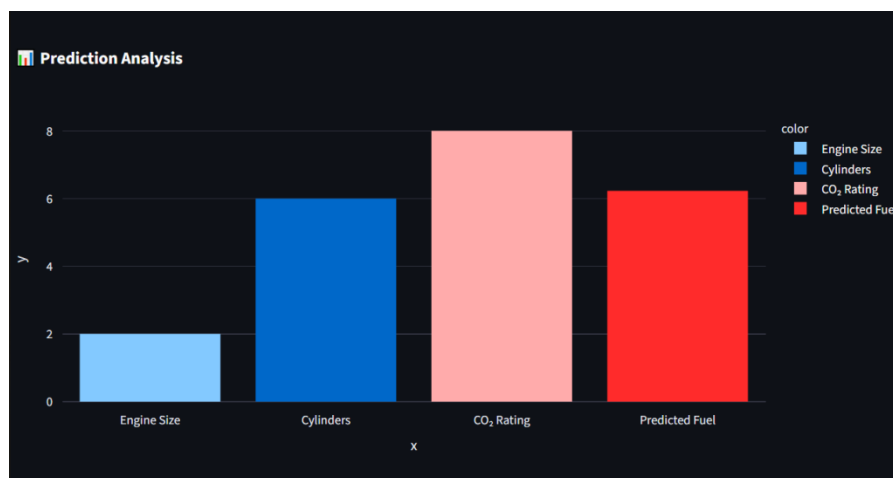


Fig 4: Prediction analysis graph

VIII. CONCLUSION

The project “Predictive Analytics for Fuel Consumption Optimization Using Data Science Techniques” applies machine learning to address real-world fuel management challenges. It predicts fuel consumption using factors such as vehicle characteristics, driving patterns, and operating conditions. Linear Algorithms Regression, Tree-like Decision, and Random Forest effectively learn patterns from historical data. Data pre-processing, feature selection, and normalization improve prediction accuracy and reliability. The system provides insights to reduce excessive fuel usage through better driving behavior and vehicle maintenance. A Streamlit-based interface ensures easy user interaction and visual interpretation of results. Overall, the project demonstrates the practical application of predictive analytics in improving fuel efficiency and sustainability.

REFERENCES

- [1] T. Hastie, R. Tibshirani, and J. Friedman, *the Elements of Statistical Learning: Data Mining, Inference, and Prediction*, Springer, 2nd ed., 2009.
- [2] I. H. Witten, E. Frank, and M. A. Hall, *Data Mining: Practical Machine Learning Tools and Techniques*, Morgan Kaufmann, 3rd ed., 2011.
- [3] L. Breiman, “Random forests,” *Machine Learning Journal*, vol. 45, no. 1, pp. 5–32, Oct. 2001.
- [4] S. R. Gunn, “Support vector machines for classification and regression,” *Technical Report*, University of Southampton, 1998.
- [5] A. Géron, *Hands-On Machine Learning with Scikit-Learn, Keras, and TensorFlow*, O’Reilly Media, 2nd ed., 2019.
- [6] Y. Freund and R. E. Schapire, “A decision-theoretic generalization of on-line learning and an application to boosting,” *Journal of Computer and System Sciences*, vol. 55, no. 1, pp. 119–139, Aug. 1997.
- [7] J. Zhang, F. Wang, and K. Li, “Fuel forecast of consumption Using machine learning techniques,” *International Journal of Advanced Computer Science and Applications*, vol. 10, no. 5, pp. 234–241, May 2019.



INTERNATIONAL
STANDARD
SERIAL
NUMBER
INDIA



INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN SCIENCE | ENGINEERING | TECHNOLOGY

 9940 572 462  6381 907 438  ijirset@gmail.com



www.ijirset.com

Scan to save the contact details