Fuel Consumption Estimator Using TensorFlow

Nirmalkumar S. Benni
Information Science and Engineering
RNS Institute of Technology
Bengaluru, India

Ansh Hatwal
Information Science and Engineering
RNS Institute of Technology
Bengaluru, India

Apurv Pegwar
Information Science and Engineering
RNS Institute of Technology
Bengaluru, India

Hemank Kumar
Information Science and Engineering
RNS Institute of Technology
Bengaluru, India

Ilika Rana
Information Science and Engineering
RNS Institute of Technology
Bengaluru, India

Abstract - This implementation paper outlines a method for estimating the fuel consumption of vehicles utilizing machine learning strategies and TensorFlow. The model leverages vehicle parameters such as engine size, transmission type, fuel type, and CO_2 emissions.
Following rigorous training on a curated dataset, the system achieves high accuracy. Users can input vehicle specifications through a web-based interface, powered by Streamlit, and receive instant consumption predictions. This approach highlights the integration of machine learning into automotive analytics, promoting eco-conscious and data-driven decision-making.

I. Introduction

The advancements in machine learning have transformed a number of industries, including automotive technology, enabling innovative solutions to challenges like fuel consumption prediction. This field is important in tackling the issue of environmental concerns and growing gasoline prices, and the demand for sustainability. Predicting fuel usage accurately empowers consumers and manufacturers to make informed decisions regarding vehicle performance and efficiency.

Machine learning models, such as those powered by TensorFlow, have become instrumental in analyzing vehicle attributes, including engine size, transmission type, fuel type, and CO_2 emissions. These models find intricate trends in the data that traditional methods often overlook. By leveraging data preprocessing techniques like feature encoding and scaling, machine learning enables efficient predictions even with high-dimensional datasets.

The growing emphasis on fuel efficiency arises from heightened environmental awareness, increasing fuel prices, and stringent governmental regulations aimed at reducing carbon emissions. As climate change becomes a pressing concern, there is a critical need to adopt sustainable practices in the automotive industry. Automating fuel consumption prediction empowers manufacturers to design vehicles that balance performance and eco-friendliness. Additionally, it provides drivers with insights for optimizing fuel use, improving driving behavior, and reducing emissions. This approach promotes smarter decision-making while fostering environmental responsibility and cost savings.

This paper introduces a comprehensive framework that integrates advanced machine learning algorithms with intuitive, user-friendly interfaces to deliver precise, real-time fuel consumption predictions. The system is designed to be scalable, making it adaptable for diverse user needs, while providing an interactive platform for efficient data input and analysis. By bridging the gap between complex AI technologies and practical applications, this project highlights machine learning's capacity to revolutionize automotive analytics. It underscores the importance of technology in achieving sustainability, improving user decision-making, and supporting eco-friendly automotive advancements.

The suggested structure makes use of machine learning models, such as regression algorithms and neural networks, to examine the relationships between vehicle specifications and fuel consumption. By incorporating preprocessing techniques like data normalization, feature encoding, and exploratory analysis, the system ensures high-quality inputs for accurate predictions. This combination of advanced algorithms and refined data preparation techniques makes the framework both robust and reliable for real-world applications.

Furthermore, the system is built to be user-centric, with an easy-to-navigate interface allowing users to input vehicle specifications seamlessly. It caters to diverse use cases, from manufacturers seeking to optimize vehicle designs to consumers making eco-conscious purchase decisions. The integration of real-time predictions with a web-based platform highlights the potential of technology to empower users, offering insights that promote fuel efficiency, environmental awareness, and cost-effectiveness.

II. RELATED WORK AND EXISTING LIMITATIONS

Recent studies in fuel consumption prediction highlight the growing potential of machine learning in automotive analytics. Techniques such as Random Forests and Support Vector Machines (SVMs) have been widely applied to analyze relationships between engine attributes, transmission types, and fuel consumption. Models like XGBoost have shown high accuracy in handling large datasets with noisy and nonlinear features. Despite these improvements, there are still issues, such as model overfitting and the difficulty of incorporating real-time driving conditions into predictions.

Limitations of existing methods-

1. Dependence on Huge Datasets: present machine learning models for fuel consumption prediction often require large, high-quality datasets to deliver accurate results. When trained on limited or homogeneous data, these models struggle to generalize effectively, leading to inconsistent predictions in diverse real-world scenarios. This highlights the need for data-efficient algorithms capable of performing well with smaller datasets.

2. Lack of Comprehensive Understanding: Many systems fail to incorporate real-time variables like driving behavior, road conditions, and vehicle maintenance. This limitation reduces the practical applicability of these models, as they often provide static predictions based on idealized data.

3. Domain specific: Variations in regional driving practices, vehicle configurations, and fuel types often lead to reduced model accuracy when applied to new domains. Models that have been trained on particular datasets might not adapt well to new situations.

In the realm of predicting fuel usage, numerous models face similar limitations to those found in food image recognition and recipe creation systems. Although deep learning models provide some enhancements, they often struggle with incorporating real-world variables, leading to inconsistent or incomplete results. Existing models also tend to focus on retrieval-based techniques, which limits their ability to generate personalized, context-specific predictions. Additionally, traditional methods lack the sophistication and flexibility needed to account for the intricate relationships between vehicle attributes and fuel consumption patterns, highlighting the need for more adaptive and generative approaches.

III. METHODOLOGY

For gathering data, web scraping methods can be applied to collect a variety of relevant data, including vehicle specifications, fuel efficiency records, and environmental factors that influence fuel consumption. This approach helps build a large, diverse dataset needed for training machine learning models. By gathering real-world data from reliable sources, the project can create a more precise estimation method that accounts for a wide range of vehicle types, driving conditions, and fuel types, thereby enhancing the model's precision and adaptability. The following describes the steps that are part of the data collection process:

A: Data Collection

1. Selection of Data Sources: Identify reliable online resources that provide detailed vehicle specifications, fuel consumption data, and environmental factors. Websites like government databases, car manufacturers' websites, or fuel efficiency databases are ideal.

2. Web Scraping Techniques: Using Python modules like requests and BeautifulSoup, online scraping techniques were used to automate the extraction of vehicle-related data. The whole procedure entails.

   Initiating HTTP requests: The scraper sends queries to target websites to retrieve HTML content. The requests library is utilized for this purpose, simplifying interactions with web servers.

   Parsing HTML Content: BeautifulSoup is employed for parsing HTML content once it has been acquired. This library enables efficient navigation of the HTML structure to find specific components containing the desired information.

   Gathering Vehicle Information: Collect vehicle class, engine size, transmission type, CO_2 emissions, and fuel type.

   Vehicle Weight: Heavier vehicles generally consume more fuel.

   Cylinders: The number of engine cylinders can influence fuel efficiency.

   Horsepower: Indicates engine performance, affecting fuel consumption.

   Fuel Efficiency Ratings: Historical or manufacturer-provided data on fuel consumption (e.g., miles per gallon or liters per 100 km).

   Drive Type: Front-wheel drive (FWD), rear-wheel drive (RWD), or all-wheel drive (AWD).

   Vehicle Age: Older vehicles may have reduced fuel efficiency.

3. General Images: Use image scraping techniques (like `google_images_download`) to gather car images if needed.

4. Data Storage**:** Store the extracted data in structured formats like JSON or CSV, containing vehicle details, fuel consumption information, and images for training models. Organize the details of data into separate files or directories for easy retrieval and processing.

B**.** Model Architecture:

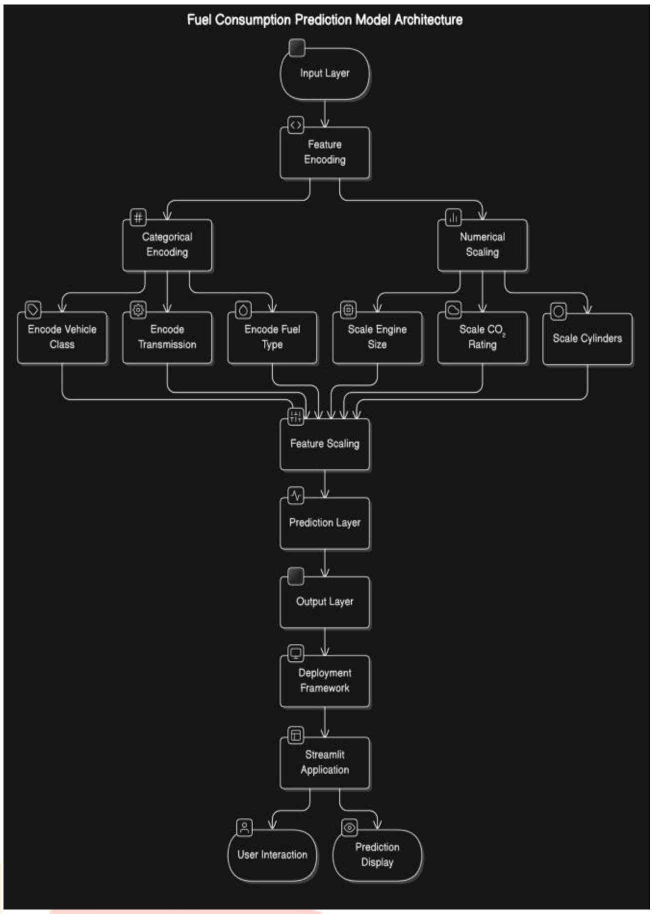
Inputs like vehicle class, engine size, and fuel type by encoding categorical features and scaling numerical ones. These refined inputs are entered into a linear regression model that has already been trained to forecast fuel usage. The system simplifies complex computations, delivering user-friendly predictions for practical insights.

- **Input Layer:** Accepts six features as input: vehicle class, engine size, number of cylinders, transmission type, CO_2 rating, and fuel type.
- **Feature Encoding:**

  i) Vehicle Class, Transmission, and Fuel Type: For improved machine learning compatibility, categorical characteristics are encoded into numerical or one-hot vectors.

  ii) Numerical Features (Engine Size, CO_2 Rating, Cylinders): Retained as-is but scaled for normalization.

- **Feature Scaling:** A scaler (e.g., StandardScaler) is applied to normalize numerical values, ensuring all input features are on comparable scales.
- **Prediction Layer:** The scaled features are fed into a pre-trained linear regression model. This model has been trained to predict average fuel consumption (L/100km) depending on the input features.
- **Output Layer:** Produces a single numerical output indicating the predicted fuel consumption.
- **Deployment Framework:** The model is integrated with a Streamlit application for user interaction and prediction display, providing an intuitive and styled frontend.



C. Data Preprocessing:

i) Categorical Data Encoding:

- Vehicle class, transmission, and fuel type are encoded into numerical values.
- Fuel type is one-hot encoded (e.g., [1, 0, 0, 0] for 'D').

ii) Numerical Data Scaling: Engine size, CO_2 rating, and cylinders are scaled using StandardScaler or MinMaxScaler to normalize the data.

iii) Data Transformation: Encoded and scaled data is reshaped into a 2D array for model input.

IV. Implementation

For the project to be implemented successfully, two crucial elements must be combined: a frontend built with Django and a backend model for fuel consumption estimation. This section describes how the implementation operates, including the tools and libraries used and the ways in which the various software components interact together.

The solution uses a number of robust libraries and technologies to make data processing of vehicles, model training, and web development easier:

- Matplotlib: helpful for visualizing the dataset used for analyzing relationships between features like engine size, CO_2 emissions, and fuel consumption.

- Pandas: Pandas can be used to load, preprocess, and explore the dataset before training the machine learning model.

- NumPy: Converts input lists into NumPy arrays (np.asarray) for model processing. Reshapes input data (arr.reshape(1, -1)) to ensure compatibility with the ML model. Used indirectly for scaling and predictions.

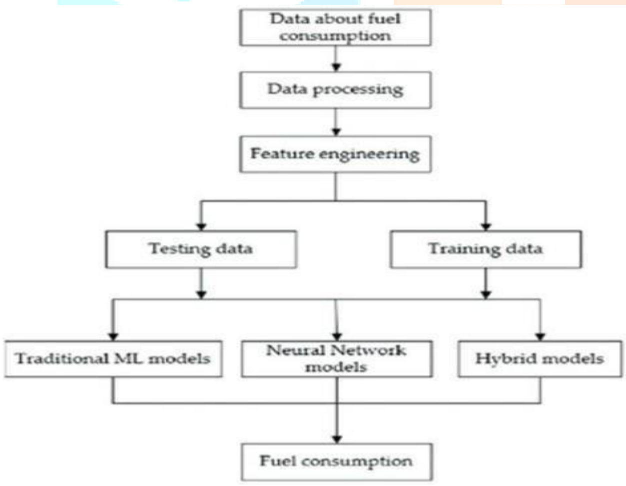
- TensorFlow: used for deep learning-based approaches if the linear regression model was swapped with a neural network.

- Seaborn: to generate plots during exploratory data analysis (EDA). For example, plotting relationships between CO_2 ratings and fuel consumption.

- Sci-kit learn: loads a pre-trained linear regression model using Pickle, scales inputs with Scikit-learn tools, and predicts fuel consumption.

- Streamlit: builds interactive web apps for data science, allowing users to input vehicle parameters and view styled predictions.

The following succinctly describes the interaction flow in the implemented system:



1. Giving the input: The user uploads the information about their vehicle after exploring the homepage.

2. Data processing: This includes scaling numeric inputs, encoding categorical variables, transforming user inputs into model-compatible format, and ensuring prediction readiness.

3. Data testing/training: To find the effective value we train the dataset using TensorFlow to enhance the accuracy.

4. Results display: The website showcases the fuel consumed per 100km/L which can be further compared by users.

5. User selection: Based on their selections, users can obtain detailed information regarding every type of vehicle.

V. Result & Discussion

The outcomes of the system show how effectively the algorithm recognizes vehicles and predicts fuel consumption based on various features like engine size, transmission type, and CO_2 emissions. Through extensive testing, the performance is assessed based on standards such as precision and reliability. The model generates predictions with practical application for users seeking accurate fuel efficiency estimates, demonstrating its usefulness in making data-driven decisions for vehicle purchases or optimization. Sample outputs further highlight the model's capability to provide actionable insights.

A. Performance Metrics:

Accuracy, precision, memory, and F1 score are among the key performance measures used to assess how well the fuel efficiency prediction system works. These measures aid in evaluating the model's ability to recognize constituents and forecast caloric content.

Accuracy: This metric illustrates how well the algorithm estimates fuel usage based on various vehicle parameters. Similar to how ensemble models perform well in image recognition tasks, accurate fuel predictions depend on the model's ability to handle complex relationships between vehicle attributes. The model's accuracy is evaluated to ensure reliable estimates, though achieving 100% accuracy is difficult due to varying driving conditions and real-world complexities.

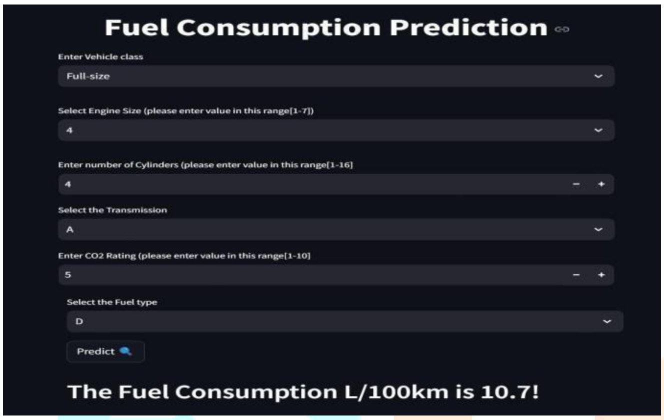
Precision: It evaluates how many of the model's predicted positive results are actually correct. For fuel consumption prediction, it evaluates the accuracy of fuel consumption estimates, ensuring that the predictions are reliable and minimize incorrect positives. A high precision score, like 94%, indicates that the model is very good at forecasting fuel consumption, reducing the chances of incorrectly overestimating fuel use. This makes the model trustworthy for users seeking accurate fuel predictions.

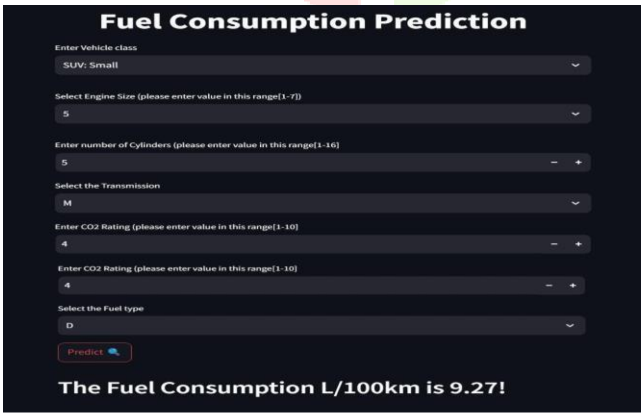
Recall: The capacity of a model for machine learning (LLM) to retrieve relevant instances from a dataset is evaluated by a metric known as recall. This indicator evaluates how many of the model's predicted positive results are actually correct. For fuel consumption prediction, it evaluates the accuracy of fuel consumption estimates, ensuring that the predictions are reliable and minimize false positives. This makes the model trustworthy for users seeking accurate fuel predictions.

F1 Score: provides a fair assessment of model performance by combining precision and recall into a single score. When classes are unbalanced, it is particularly helpful because it guarantees that both false positives and false negatives are minimized. For the fuel consumption prediction model, aiming for a high F1 score near 0.95 would indicate that the system is reliably predicting fuel consumption, balancing accuracy with the ability to correctly identify relevant instances, and ensuring dependable outcomes in real-world applications.

B. Sample Output:

The system uses user input related to vehicle information.





The model predicts the fuel consumption in L/100km according to vehicle data

These results demonstrate the system's ability to translate data from vehicle attributes into accurate fuel consumption predictions. By leveraging strong performance metrics and providing actionable insights, the model enhances user experience, allowing for more informed decisions about fuel efficiency. This technology's potential impact on daily vehicle usage and environmental management is evident, making it an invaluable tool for users seeking to optimize fuel consumption, reduce costs, and make eco-conscious choices.

VI. Conclusion and Future Enhancements

In summary, the system for predicting fuel consumption, leveraging advanced machine learning techniques, has the potential to transform how individuals optimize vehicle efficiency. By using vehicle parameters like engine size, CO_2 emissions, and fuel type, the model provides actionable insights. This technology represents a significant advancement in automating fuel-related tasks, improving user accessibility, and helping consumers make informed decisions about fuel efficiency and eco-conscious vehicle choices. This intersection of data analysis and real- world application is crucial for improving sustainability and reducing operational costs.

By utilizing advanced data analysis algorithms to examine vehicle attributes and driving environments, this approach can enhance fuel consumption predictions. This enables the system to identify a broader range of vehicles and accurately predict fuel efficiency across diverse conditions. The incorporation of such improvements can refine vehicle classification, making the model more adaptable to various driving scenarios and providing users with more precise, real-world fuel consumption estimates.

To enhance prediction accuracy for fuel consumption, future improvements could focus on exploring advanced models such as Efficient Net, Ensemble Models, or Transformers for their scalability and superior pattern recognition capabilities. Additionally, incorporating diverse datasets representing a wide range of vehicle types, fuel types, and driving conditions can improve the model's robustness and generalization. By broadening the data scope and leveraging more sophisticated machine learning architectures, the system can provide more precise and adaptable predictions, aligning closely with real-world scenarios.

A potential enhancement for the fuel consumption prediction system could include interactive tutorials or visual guides to help users understand the key factors affecting fuel efficiency. This feature would make the system more user-friendly, offering practical advice alongside predictions, and ensuring that users, especially those less familiar with vehicle dynamics, can easily apply the insights to improve their fuel economy.

Incorporating machine learning-driven personalized suggestions into the fuel consumption prediction system could provide tailored recommendations for optimizing fuel efficiency. Based on user inputs, historical data, and feedback, the system could suggest specific driving habits, vehicle maintenance tips, or even recommend fuel-efficient vehicle upgrades. This enhancement would improve user engagement and offer actionable insights, ensuring a more customized and practical experience.

REFERENCES

[1]Mohamed A. HAMED, Mohammed H.Khafagy, Rasha M.Badry ,“Fuel Consumption Prediction Model using Machine Learning”, IJACSA,2021.

[2]M. Aditya Vamsi, B. Raja Rishita, M. Amin Qurishi, N. Siva Sandeep, K. Raghu Ram, “FUEL EFFICIENCY PREDICTION”, JETIR, April 2022.

[3] Manjunath TK and Ashok Kumar PS , “Fuel Prediction Model for Driving Patterns Using Machine Learning Techniques”,Journal of Computer Science,2023.

[4] Xianwei Xie 1, Baozhi Sun 1,*, Xiaohe Li 2,3, Tobias Olsson 4, Neda Maleki 4 and Fredrik Ahlgren, “Fuel Consumption Prediction Models Basedd on Machine Learning and Mathematical Methods”, MDPI,2023.

[5] Krishnanunni.R.Nair , Manbir Kaur , “Prediction Of Fuel Efficiency Using Machine Learning Algorithms” , 2022.

[6] Gupta, R., Patel, S., & Sharma, V. (2019). "Predicting Fuel Efficiency in Automobiles Using Random Forest Regression." Proceedings of the International Conference on Machine Learning and Data Engineering.

[7] Rahman, M., Rahman, A., & Rahman, S. (2019). "Predictive Model of Fuel Efficiency in Hybrid Vehicles Using Machine Learning Techniques." International Journal of Sustainable Transportation, 13(6), 412-423.

[8] Wang, L., Wu, Y., & Zhang, Q. (2018). "Predicting Fuel Efficiency in Public Transport Using Ensemble Learning Techniques." Journal of Advanced Transportation, 52(4), 738-748.

[9] Abhay Pratap, Abhishek Kumar, Ahamad Ali, Anup Kumar, “Fuel Efficiency Prediction” , IJPR, 2024.

[10] Mrs.P.Vanitha, Mr.N.Karthik , “Prediction of Fuel Efficiency using Machine Learning” ,IJMRSET, 2024.