

Fuel Efficiency Prediction using Machine Learning

Project Overview

This project focuses on predicting vehicle fuel efficiency using a Random Forest Regressor. The key components of the project include data preprocessing, outlier detection and removal, feature encoding, and comprehensive model evaluation. The model was trained on a dataset sourced from Kaggle and achieved a strong R^2 score of 0.90, demonstrating its predictive accuracy.

Deployment

The trained Random Forest model was deployed using Streamlit. Streamlit enables users to input vehicle features and receive instant fuel efficiency predictions. The deployment features a clean and interactive user interface suitable for real-time use.

Dataset

The dataset used for this project is publicly available on Kaggle. It contains information about various vehicle attributes such as engine size, cylinders, fuel type, and more. Link to the dataset: <https://www.kaggle.com/datasets/sahirmaharajj/fuel-economy>

Required Python Packages

The following Python libraries are required to run the project:

1. NumPy - for numerical computations
2. Pandas - for data manipulation and analysis
3. Matplotlib - for data visualization
4. Seaborn - for statistical data visualization
5. Scikit-Learn - for building and evaluating machine learning models
6. Pickle - for saving the trained model
7. Streamlit - for deploying the model with a user interface