**MPG Fuel Project**

**Overview:**

The MPG-Fuel-Project focuses on predicting miles per gallon (mpg) for vehicles using various regression models. The project explores data preprocessing, model building, and performance evaluation to derive insights into fuel efficiency.

**Data Exploration and Preprocessing:**

The dataset comprises features related to automobiles, and the target variable is mpg. Initial steps involve examining numerical and non-numerical values, checking for missing data, and ensuring a consistent numerical format. Data preprocessing includes handling missing values, scaling features, and conducting exploratory data analysis (EDA).

**Model Building:**

**Linear Regression Model:**

A Linear Regression model is employed to establish a baseline for prediction. The dataset is split into training and testing sets, features are scaled, and the model is evaluated on both sets. The results provide an initial understanding of predictive capabilities.

Linear Regression Results:

Train score: 0.755

Test score: 0.727

Mean Squared Error: 17.43

**Ridge Regression Model:**

To address potential overfitting, a Ridge Regression model is introduced. Hyperparameter tuning using GridSearchCV further refines the model. The evaluation includes train and test scores, highlighting the impact of regularization.

Ridge Regression Results:

Best alpha: 1

Train score: 0.755

Test score: 0.727

Mean Squared Error: 17.43

**Decision Tree Regressor:**

A Decision Tree Regressor is implemented to capture non-linear relationships in the data. Hyperparameter tuning is performed using GridSearchCV to optimize model performance. The decision tree's interpretability is considered along with predictive accuracy.

Decision Tree Results:

Best parameters: {'max\_depth': 6, 'max\_features': 'log2', 'min\_samples\_leaf': 3, 'min\_samples\_split': 7}

Train score: 0.860

Test score: 0.817

Mean Squared Error: 14.02

**Random Forest Regressor:**

The Random Forest Regressor, an ensemble method, is chosen for its ability to handle complex relationships. Hyperparameter tuning via RandomizedSearchCV enhances the model's accuracy. The final model demonstrates superior predictive capabilities.

Random Forest Results:

Best parameters: {'n\_estimators': 500, 'min\_samples\_split': 3, 'min\_samples\_leaf': 1, 'max\_features': 'sqrt', 'max\_depth': 6}

Train score: 0.928

Test score: 0.868

Mean Squared Error: 12.59

**Model Comparison and Evaluation:**

The performance of each model is compared based on training and testing scores, emphasizing accuracy and generalization to unseen data. The Random Forest Regressor emerges as the most effective model for predicting mpg, showcasing its versatility in handling complex relationships within the dataset.

**Conclusion:**

The MPG-Fuel-Project concludes with a comprehensive analysis of regression models for predicting miles per gallon. The Random Forest Regressor stands out as the preferred choice, providing a robust framework for future predictions and further refinement. The project highlights the importance of model selection, hyperparameter tuning, and evaluation metrics in the context of regression tasks.