Machine Learning Model Report

Objective:

To build a machine learning model to predict the class of cars based on a set of features in the "cars_class.csv" dataset.

Code Flow:

1.Import Required Libraries:

Required libraries like pandas, sklearn, seaborn, matplot were imported.

2.Data Exploration and Preprocessing:

Data Loading: The dataset was loaded into a Pandas DataFrame named "cars_class".

Feature and Target Variable Identification: Features (X) and the target variable (y) were identified. The target variable is 'Class,' and the features are all other columns

Data Splitting: The dataset was split into training and testing sets with a test size of 20% for evaluation.

3.Data Standardization:

Feature Scaling: StandardScaler from scikit-learn was applied to standardize the feature values. Standardization ensures that all features have a mean of 0 and a standard deviation of 1.

4.Model Selection and Training:

Classification: A Random Forest Classifier was chosen as the model for its ability to handle complex relationships in the data and provide feature importance.

Model Training: The model was trained on the scaled training data using the fit method.

5.Model Evaluation:

Model Prediction: The trained model was used to predict the target variable 'class' on the test set.

Evaluation Metrics: Model performance was assessed using accuracy, F1 score, and a confusion matrix to provide insights into the model's classification performance.

OUTPUT:

```
Accuracy:   0.7222222222222222
F1 Score:   0.7090956005757002
Confusion Matrix: [[35  0  0  0]
 [ 1 19 18  2]
 [ 3 13 15  3]
 [ 0  0  0 35]]
```

6.Feature Importance Analysis:

Feature Importance Calculation: The importance of each feature in the trained Random Forest model was determined using the feature_importances_ attribute.

Visualization: Feature importance scores were visualized using a bar plot. This plot helps identify the most influential features in predicting car classes.

OUTPUT:
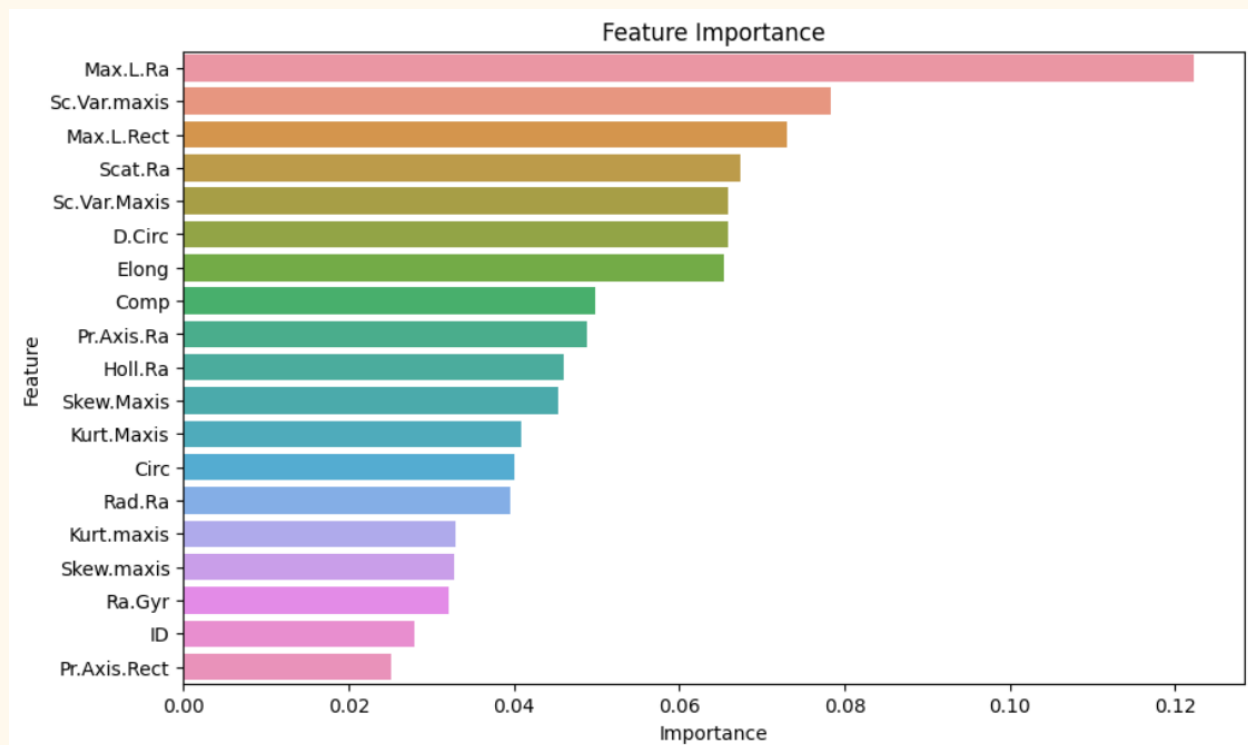
Conclusion:

In conclusion, the model, after standardizing features, achieves reasonable accuracy and F1 score on the test set. Feature importance analysis reveals key variables influencing predictions and is visualized through a bar plot. These findings offer valuable insights into the model's performance and highlight important features in the dataset.