**Project Development Phase**
**Model Performance Test**

| | |
|---|---|
| Date | 23 November 2023 |
| Team ID | Team-592817 |
| Project Name | Car purchase prediction using ML |
| Maximum Marks | 10 Marks |

**Model Performance Testing:**

Project team shall fill the following information in the model performance testing template.

| S.No. | Parameter | Values | Screenshot |
|---|---|---|---|
| 1. | Metrics | **Regression Model:**<br>MAE - , MSE - , RMSE - , R2 score -<br><br>**Classification Model:**<br>Confusion Matrix - , Accuray Score- & Classification Report - | Attached Below |
| 2. | Tune the Model | Hyperparameter Tuning -<br>Validation Method - | Attached Below |

```
Metrics for Logistic Regression:
Accuracy: 0.4958, Precision: 0.4958, Recall: 1.0000
F1-score: 0.6630, ROC-AUC: 0.5000

Metrics for Decision Tree:
Accuracy: 0.8917, Precision: 0.9115, Recall: 0.8655
F1-score: 0.8879, ROC-AUC: 0.8915

Metrics for Random Forest:
Accuracy: 0.8958, Precision: 0.8917, Recall: 0.8992
F1-score: 0.8954, ROC-AUC: 0.8959

Metrics for Gradient Boosting:
Accuracy: 0.9250, Precision: 0.9174, Recall: 0.9328
F1-score: 0.9250, ROC-AUC: 0.9251

Metrics for SVM:
Accuracy: 0.7125, Precision: 0.7551, Recall: 0.6218
F1-score: 0.6820, ROC-AUC: 0.7118
```

1. Metrics
    a. Logistic Regression

```
Logistic Regression Metrics:
Confusion Matrix:
[[  0 121]
 [  0 119]]

Accuracy Score: 0.4958

Classification Report:
              precision    recall  f1-score   support

           0       0.00      0.00      0.00       121
           1       0.50      1.00      0.66       119

    accuracy                           0.50       240
   macro avg       0.25      0.50      0.33       240
weighted avg       0.25      0.50      0.33       240
```

    b. Decision Tree

```
Decision Trees Metrics:
Confusion Matrix:
[[111  10]
 [ 16 103]]

Accuracy Score: 0.8917

Classification Report:
              precision    recall  f1-score   support

           0       0.87      0.92      0.90       121
           1       0.91      0.87      0.89       119

    accuracy                           0.89       240
   macro avg       0.89      0.89      0.89       240
weighted avg       0.89      0.89      0.89       240
```

    c. Random Forest

```
Random Forest Metrics:
Confusion Matrix:
[[108  13]
 [ 14 105]]

Accuracy Score: 0.8875

Classification Report:
              precision    recall  f1-score   support

           0       0.89      0.89      0.89       121
           1       0.89      0.88      0.89       119

    accuracy                           0.89       240
   macro avg       0.89      0.89      0.89       240
weighted avg       0.89      0.89      0.89       240
```

d. Gradient Boosting

```
Gradient Boosting Metrics:
Confusion Matrix:
[[111  10]
 [  8 111]]

Accuracy Score: 0.9250

Classification Report:
              precision    recall  f1-score   support

           0       0.93      0.92      0.93       121
           1       0.92      0.93      0.93       119

    accuracy                           0.93       240
   macro avg       0.93      0.93      0.93       240
weighted avg       0.93      0.93      0.93       240
```

e. Support Vector Machine

```
SVM Metrics:
Confusion Matrix:
[[97 24]
 [45 74]]

Accuracy Score: 0.7125

Classification Report:
              precision    recall  f1-score   support

           0       0.68      0.80      0.74       121
           1       0.76      0.62      0.68       119

    accuracy                           0.71       240
   macro avg       0.72      0.71      0.71       240
weighted avg       0.72      0.71      0.71       240
```

2. Cross Fold Validation

```
Mean Logistic Regression CV Accuracy: 0.5010471204188482
Mean Decision Tree CV Accuracy: 0.8713514397905758
Mean Random Forest CV Accuracy: 0.8912249127399651
Mean Gradient Boosting CV Accuracy: 0.902732329842932
Mean SVM CV Accuracy: 0.7019033595113439
```

3. Hyper Parameter Tuning
    a. Logistic Regression

```
Logistic Regression - Best Hyperparameters: {'C': 1.0, 'penalty': 'l1', 'solver': 'liblinear'}
Logistic Regression - Accuracy on test set: 0.8791666666666667
```

    b. Decision Tree

```
Decision Tree - Best Hyperparameters: {'max_depth': 5, 'min_samples_leaf': 4, 'min_samples_split': 10}
Decision Tree - Accuracy on test set: 0.9041666666666667
```

    c. Random Forest

```
Random Forest - Best Hyperparameters: {'max_depth': 7, 'min_samples_leaf': 1, 'min_samples_split': 10, 'n_estimators': 300}
Random Forest - Accuracy on test set: 0.9291666666666667
```

    d. Gradient Boosting

```
Gradient Boosting - Best Hyperparameters: {'learning_rate': 0.01, 'max_depth': 3, 'n_estimators': 300}
Gradient Boosting - Accuracy on test set: 0.9375
```