

Project Development Phase

Model Performance Test

Date	10 November 2022
Team ID	PNT2022TMID06350
Project Name	Smart Lender - Applicant Credibility Prediction for Loan Approval
Maximum Marks	10 Marks

Model Performance Testing:

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

S.No.	Parameter	Values	Screenshot
1.	Metrics	Regression Model: MAE – 0.0476 MSE – 0.0046 RMSE – 0.0682 R2 score – 0.78	
2.	Tune the Model	Hyperparameter Tuning - Validation Method -	

1. Metrics :

```
In [55]: from math import sqrt
from sklearn.metrics import mean_absolute_error, mean_squared_error

print("Mean Absolute Error :", mean_absolute_error(y_test, y_pred))
print("Mean Squared Error :", mean_squared_error(y_test, y_pred))
print("Rooted Mean Squared Error :", sqrt(mean_squared_error(y_test, y_pred)))
print("R2 Score :", r2_score(y_test, y_pred))
```

Mean Absolute Error : 0.04766740707308981
Mean Squared Error : 0.004658892249358201
Rooted Mean Squared Error : 0.06825607847919628
R2 Score : 0.7777563100798979

Hyper-parameter Tuning

```
In [83]: # Hyper-parameter Tuning is for improvement of the accuracy.

In [84]: RFmodel = RandomForestClassifier(n_estimators=50, min_samples_split=25, max_depth=7, max_features=1)
RFmodel.fit(x_train, y_train)
RF_CV_score = cross_val_score(RFmodel, x, y, cv=5)

print("Accuracy using Random Forest classifier is ", RFmodel.score(x_test, y_test)*100,"%")
print("Cross validation for Random Forest is ", np.mean(RF_CV_score)*100, "%")
```

Accuracy using Random Forest classifier is 79.87012987012987 %
Cross validation for Random Forest is 78.50593096094896 %

Confusion Matrix

```
In [85]: # A confusion matrix is a summary of prediction results on a classification problem.

In [86]: # importing
from sklearn.metrics import confusion_matrix

In [87]: # using Logistic Regression Model
y_pred = LRmodel.predict(x_test)
cm = confusion_matrix(y_test, y_pred)
cm
```

Out[87]: array([[19, 24],
[2, 109]], dtype=int64)