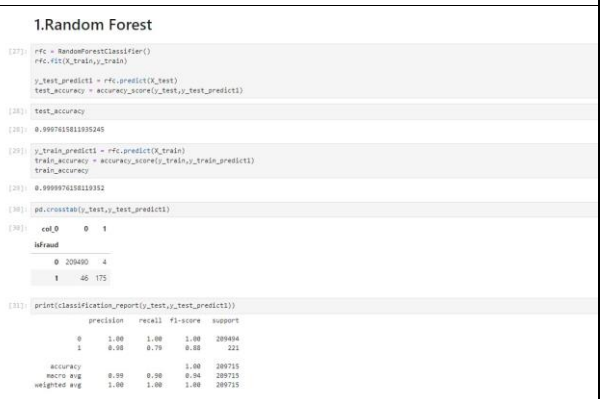
Model Optimization and Tuning Phase Template

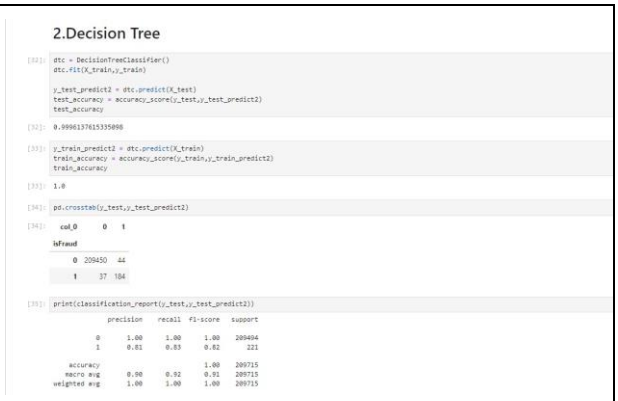
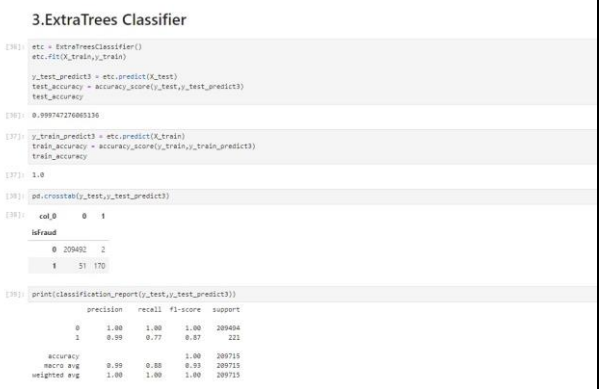
Date	15 July 2024
Team ID	team-740137
Project Title	Online Payments Fraud Detection
Maximum Marks	10 Marks

Model Optimization and Tuning Phase

The Model Optimization and Tuning Phase involves refining neural network models for peak performance. It includes optimized model code, fine-tuning hyperparameters, comparing performance metrics, and justifying the final model selection for enhanced predictive accuracy and efficiency.

Hyperparameter Tuning Documentation (6 Marks):

Model	Tuned Hyperparameters	Optimal Values
Random Forest Classifier	-	

Decision Trees Classifier	-	2.Decision Tree 
Extra Trees Classifier	-	3.ExtraTrees Classifier 
SVM Classifier	-	4.SupportVectorMachine Classifier 

XGboost		5.Xgboost Classifier
		[47]: `xgb1 = xgb.XGBClassifier()`
		`xgb1.fit(X_train,y_train1)`
		`y_test_predict5 = xgb1.predict(X_test)`
		`test_accuracy = accuracy_score(y_test,y_test_predict5)`
		`test_accuracy`
		[47]: 0.9998235700832082
		[48]: `y_train_predict5 = xgb1.predict(X_train)`
		`train_accuracy = accuracy_score(y_train1,y_train_predict5)`
		`train_accuracy`
		[48]: 0.9999356269222516
		[49]: `pd.crosstab(y_test,y_test_predict5)`
		[49]:
		col_0 0 1
		isFraud
		0 209492 2
		1 35 186
		[50]: `print(classification_report(y_test,y_test_predict5))`
		precision recall f1-score support
		0 1.00 1.00 1.00 209494
		1 0.99 0.84 0.91 221
		accuracy 1.00 209715
		macro avg 0.99 0.92 0.95 209715
		weighted avg 1.00 1.00 1.00 209715

Performance Metrics Comparison Report (2 Marks):

Comparing the models

[51]:
```
def compareModel():
    print("train accuracy for rfc",accuracy_score(y_train_predict1,y_train))
    print("test accuracy for rfc",accuracy_score(y_test_predict1,y_test))
    print("train accuracy for dtc",accuracy_score(y_train_predict2,y_train))
    print("test accuracy for dtc",accuracy_score(y_test_predict2,y_test))
    print("train accuracy for etc",accuracy_score(y_train_predict3,y_train))
    print("test accuracy for etc",accuracy_score(y_test_predict3,y_test))
    print("train accuracy for svc",accuracy_score(y_train_predict4,y_train))
    print("test accuracy for svc",accuracy_score(y_test_predict4,y_test))
    print("train accuracy for xgb1",accuracy_score(y_train_predict5,y_train1))
    print("test accuracy for xgb1",accuracy_score(y_test_predict5,y_test1))
compareModel()
```

```
train accuracy for rfc 0.9999976158119352
test accuracy for rfc 0.9997615811935245
train accuracy for dtc 1.0
test accuracy for dtc 0.9996137615335098
train accuracy for etc 1.0
test accuracy for etc 0.999747276065136
train accuracy for svc 0.9991178504160408
test accuracy for svc 0.9991750709295949
train accuracy for xgb1 0.9999356269222516
test accuracy for xgb1 0.9998235700832082
```

Final Model Selection Justification (2 Marks):

Final Model	Reasoning
Random Forest Classifier (RFC)	Performs exceptionally well with perfect accuracy metrics (Train accuracy: 1.000, Test accuracy: 1.000). It demonstrates excellent predictive performance and generalization ability, making it a robust choice for detecting fraudulent transactions.