

Project Development Phase Model Performance Test

Date	9 November 2023
Team ID	Team-593390
Project Name	Online Payments Fraud Detection Using ML
Maximum Marks	10 Marks
Team Members	Rani Kushwaha Sachin Kajal

Model Performance Testing:

Since the outcome of our project needed to be predicted as either "is a fraud" or "is not a fraud," a classification-based model was necessary.

Random Forest classifier, Decision Tree classifier, Extra Tree classifier, SVM classifier were the models utilized in the projects.

The metrics reports for each model are as follows:

Random Forest classifier:

1. Test accuracy

```
y_pred=rfc.predict(X_test)
print("Training Score",accuracy_score(y_train_smote,rfc.predict(x_train_smote)))
print("Testing Accuracy",accuracy_score(y_test,y_pred))
print(X_test.shape)
```

Testing Accuracy 0.9837896584810781

2. Train accuracy

```
y_pred=rfc.predict(X_test)
print("Training Score",accuracy_score(y_train_smote,rfc.predict(x_train_smote)))
print("Testing Accuracy",accuracy_score(y_test,y_pred))
print(X_test.shape)
```

Training Score 0.9386852256321613

3. Confusion Matrix

```
pd.crosstab(y_test,y_pred)
```

		col_0	is Fraud	is not Fraud
		isFraud		
	is Fraud		1439	148
	is not Fraud		20480	1250454

4. Classification Report

```
print(classification_report(y_test,y_pred))
```

	precision	recall	f1-score	support
is Fraud	0.07	0.91	0.12	1587
is not Fraud	1.00	0.98	0.99	1270934
accuracy			0.98	1272521
macro avg	0.53	0.95	0.56	1272521
weighted avg	1.00	0.98	0.99	1272521

Decision tree Classifier:

1. Test accuracy

```
from sklearn.tree import DecisionTreeClassifier
dtc = DecisionTreeClassifier()
dtc.fit(X_train, y_train)
```

```
y_test_pred_2 = dtc.predict(X_test)
```

```
accuracy_test_2 = accuracy_score(y_test, y_test_pred_2)
accuracy_test_2
```

0.9997053093819277

2. Train accuracy

```
y_train_predict2=dtc.predict(x_train)
train_accuracy=accuracy_score(y_train,y_train_predict2)
train_accuracy
```

1.0

3. Confusion Matrix

```
pd.crosstab(y_test, y_test_pred_2)
```

	col_0	is Fraud	is not Fraud
isFraud			
is Fraud		1403	184
is not Fraud		191	1270743

4. Classification Report

ExtraTrees Classifier

1. Test accuracy

```
from sklearn.ensemble import ExtraTreesClassifier
etc=ExtraTreesClassifier()
etc.fit(x_train_smote,y_train_smote)
y_pred=etc.predict(X_test)
print("Training Score",accuracy_score(y_train_smote,etc.predict(x_train_smote)))
print("Testing Accuracy",accuracy_score(y_test,y_pred))
```

Testing Accuracy 0.9994451957963758

2. Train accuracy

Training Score 1.0

3. Confusion Matrix

```
pd.crosstab(y_test,y_pred)
```

	col_0	is Fraud	is not Fraud
isFraud			
is Fraud		1426	161
is not Fraud		545	1270389

4. Classification Report

```
print(classification_report(y_test,y_pred))
```

	precision	recall	f1-score	support
is Fraud	0.72	0.90	0.80	1587
is not Fraud	1.00	1.00	1.00	1270934
accuracy			1.00	1272521
macro avg	0.86	0.95	0.90	1272521
weighted avg	1.00	1.00	1.00	1272521

Final Prediction:

```
etc.predict([[0.00000000e+00, 0.00000000e+00, 0.00000000e+00, 1.00000000e+00,0.00000000e+00, 0.00000000e+00, 1.00000000e+00])
```

```
array(['is not Fraud'], dtype=object)
```