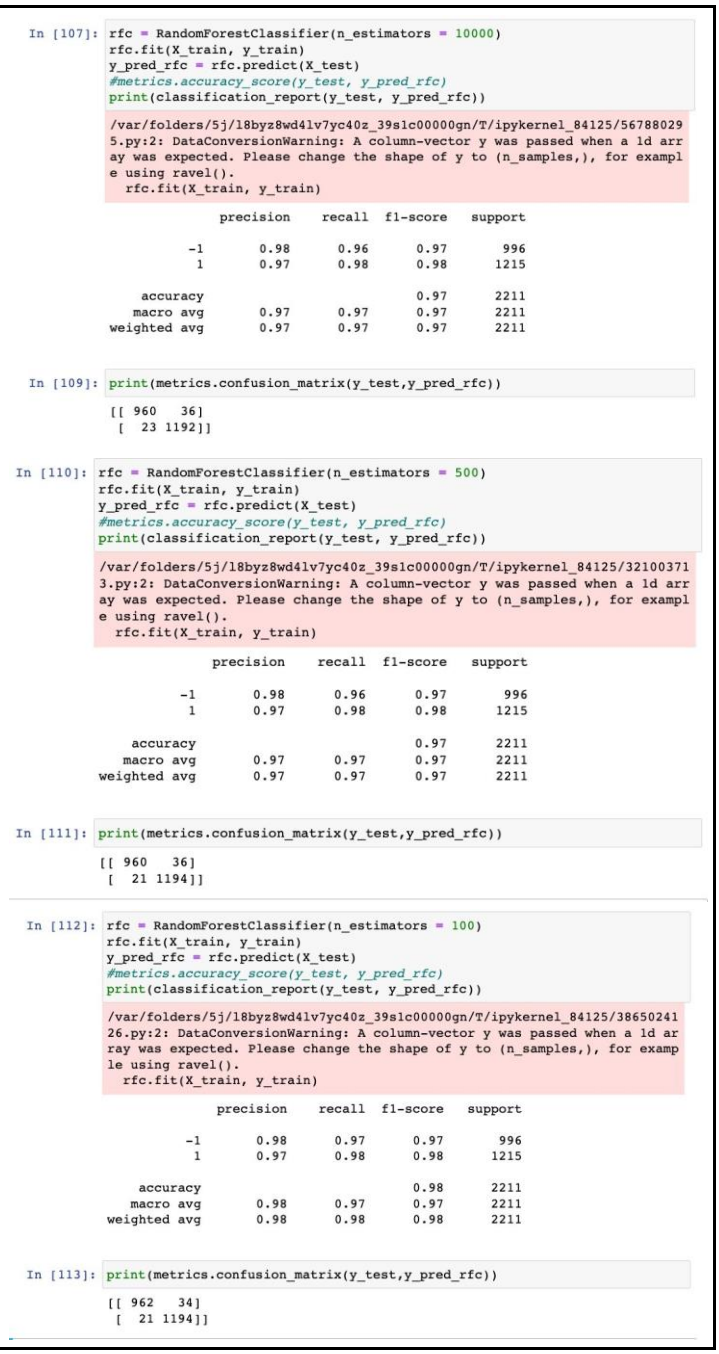
Project Development Phase Model
Performance Test

Date	19 November 2022
Team ID	PNT2022TMID24070
Project Name	Project - Web Phishing Detection
Maximum Marks	10 Marks

Model Performance Testing:

S.No.	Parameter	Values	Screenshot

1.	Metrics	**Classification Model:**
		Confusion Matrix -
		[[1401 50]
		[31 1835]],
		Accuracy Score-
		0.9755803436840519
		& Classification Report
		(included in screenshot)
			
			`X_train, X_test, y_train, y_test = train_test_split(X, y, train_size = 0.7, shuffle = True)`
			`X_train.shape, X_test.shape, y_train.shape, y_test.shape`
			✓ 0.9s
			((7738, 30), (3317, 30), (7738,), (3317,))
			
			`from sklearn.ensemble import RandomForestClassifier`
			`from sklearn.metrics import classification_report`
			`from sklearn.metrics import confusion_matrix`
			`from sklearn.metrics import accuracy_score`
			✓ 0.9s
			
			`rfc = RandomForestClassifier(n_estimators = 100)`
			`rfc.fit(X_train, y_train)`
			`y_pred_rfc = rfc.predict(X_test)`
			`print(classification_report(y_test, y_pred_rfc))`
			✓ 0.7s
			
			precision recall f1-score support
			
			-1 0.98 0.95 0.97 1471
			1 0.96 0.99 0.97 1846
			
			accuracy 0.97 3317
			macro avg 0.97 0.97 0.97 3317
			weighted avg 0.97 0.97 0.97 3317
			
			`print(confusion_matrix(y_test, y_pred_rfc))`
			✓ 0.1s
			[[1403 68]
			[27 1819]]
			
			`print(accuracy_score(y_test, y_pred_rfc))`
			✓ 0.4s
			0.9755803436840519

2.	Tune the Model	Hyperparameter Tuning - Number of estimators tweaked and best model was identified: 100	

In [107]:
```
rfc = RandomForestClassifier(n_estimators = 10000)
rfc.fit(X_train, y_train)
y_pred_rfc = rfc.predict(X_test)
#metrics.accuracy_score(y_test, y_pred_rfc)
print(classification_report(y_test, y_pred_rfc))
```

```
/var/folders/5j/l8byz8wd41v7yc40z_39s1c00000gn/T/ipykernel_84125/56788029
5.py:2: DataConversionWarning: A column-vector y was passed when a 1d arr
ay was expected. Please change the shape of y to (n_samples,), for exampl
e using ravel().
  rfc.fit(X_train, y_train)
```

	precision	recall	f1-score	support
-1	0.98	0.96	0.97	996
1	0.97	0.98	0.98	1215
accuracy			0.97	2211
macro avg	0.97	0.97	0.97	2211
weighted avg	0.97	0.97	0.97	2211

In [109]:
```
print(metrics.confusion_matrix(y_test,y_pred_rfc))
```

```
[[ 960   36]
 [  23 1192]]
```

In [110]:
```
rfc = RandomForestClassifier(n_estimators = 500)
rfc.fit(X_train, y_train)
y_pred_rfc = rfc.predict(X_test)
#metrics.accuracy_score(y_test, y_pred_rfc)
print(classification_report(y_test, y_pred_rfc))
```

```
/var/folders/5j/l8byz8wd41v7yc40z_39s1c00000gn/T/ipykernel_84125/32100371
3.py:2: DataConversionWarning: A column-vector y was passed when a 1d arr
ay was expected. Please change the shape of y to (n_samples,), for exampl
e using ravel().
  rfc.fit(X_train, y_train)
```

	precision	recall	f1-score	support
-1	0.98	0.96	0.97	996
1	0.97	0.98	0.98	1215
accuracy			0.97	2211
macro avg	0.97	0.97	0.97	2211
weighted avg	0.97	0.97	0.97	2211

In [111]:
```
print(metrics.confusion_matrix(y_test,y_pred_rfc))
```

```
[[ 960   36]
 [  21 1194]]
```

In [112]:
```
rfc = RandomForestClassifier(n_estimators = 100)
rfc.fit(X_train, y_train)
y_pred_rfc = rfc.predict(X_test)
#metrics.accuracy_score(y_test, y_pred_rfc)
print(classification_report(y_test, y_pred_rfc))
```

```
/var/folders/5j/l8byz8wd41v7yc40z_39s1c00000gn/T/ipykernel_84125/38650241
26.py:2: DataConversionWarning: A column-vector y was passed when a 1d ar
ray was expected. Please change the shape of y to (n_samples,), for examp
le using ravel().
  rfc.fit(X_train, y_train)
```

	precision	recall	f1-score	support
-1	0.98	0.97	0.97	996
1	0.97	0.98	0.98	1215
accuracy			0.98	2211
macro avg	0.98	0.97	0.97	2211
weighted avg	0.98	0.98	0.98	2211

In [113]:
```
print(metrics.confusion_matrix(y_test,y_pred_rfc))
```

```
[[ 962   34]
 [  21 1194]]
```