Project Development Phase
Model Performance Test

Date	14 November 2022
Team ID	PNT2022TMID34845
Project Name	Project – web phishing detection
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	**Classification Model:**
		XGBooster Classification
		Accuracy score-97.4%
			Performance Evaluation:
			`#computing the accuracy of the model performance`
			`acc_train_xgb = accuracy_score(y_train,y_train_xgb)`
			`acc_test_xgb = accuracy_score(y_test,y_test_xgb)`
			`print("XGBoost: Accuracy on training Data: {:.3f}".format(acc_train_xgb))`
			`print("XGBoost : Accuracy on test Data: {:.3f}".format(acc_test_xgb))`
			XGBoost: Accuracy on training Data: 0.965
			XGBoost : Accuracy on test Data: 0.923

2.	Tune the Model	**Hyperparameter Tuning –** 97%
		Validation Method – KFOLD & Cross validation method
			`#computing the classification report of the model`
			`print(metrics.classification_report(y_test, y_test_gbc))`
			precision recall f1-score support
			-1 0.99 0.96 0.97 976
			1 0.97 0.99 0.98 1235
			accuracy 0.97 2211
			macro avg 0.98 0.97 0.97 2211
			weighted avg 0.97 0.97 0.97 2211

METRICS:

TESTING THE MODEL:

Testing the saved model:

```
# load model from file
loaded_model = pickle.load(open("model.pkl", "rb"))
loaded_model
```

```
XGBClassifier(learning_rate=0.4, max_depth=7, missing=nan)
```

COMPARING THE MODEL PERFORMANCE:

```
#Sorting the datafram on accuracy
results.sort_values(by=['Test Accuracy', 'Train Accuracy'], ascending=False)
```

	ML Model	Train Accuracy	Test Accuracy
2	XGBoost	0.965	0.923
0	Decision Tree	0.929	0.923
1	Random Forest	0.916	0.908
3	SVM	0.899	0.892

For the above comparision, it is clear that the **XGBoost Classifier** works well with this dataset.

So, saving the model for future use.