Project Development Phase
Model Performance Test

Date	18-Nov-22
Team ID	PNT2022TMID52926
Project Name	Project - Web Phishing Detection
Maximum Marks	4 marks

A Classification report is used to measure the quality of predictions from a classification algorithm. How many predictions are True and how many are False. More specifically, True Positives, False Positives, True negatives and False Negatives are used to predict the metrics of a classification report as shown below.

- Support Vector Machine Classifier

	precision	recall	f1-score	support
-1	0.97	0.94	0.96	976
1	0.96	0.98	0.97	1235
accuracy			0.96	2211
macro avg	0.97	0.96	0.96	2211
weighted avg	0.96	0.96	0.96	2211

- Decision Trees

	precision	recall	f1-score	support
-1	0.95	0.95	0.95	976
1	0.96	0.96	0.96	1235
accuracy			0.96	2211
macro avg	0.96	0.96	0.96	2211
weighted avg	0.96	0.96	0.96	2211

- Logistic Regression

	precision	recall	f1-score	support
-1	0.94	0.91	0.92	976
1	0.93	0.95	0.94	1235
accuracy			0.93	2211
macro avg	0.93	0.93	0.93	2211
weighted avg	0.93	0.93	0.93	2211

- K Nearest Neighbors Method

	precision	recall	f1-score	support
-1	0.95	0.95	0.95	976
1	0.96	0.96	0.96	1235
accuracy			0.96	2211
macro avg	0.96	0.96	0.96	2211
weighted avg	0.96	0.96	0.96	2211

- RNN-LSTMs

```
f1_score on training Data: 0.985
f1_score on test Data: 0.985

Recall on training Data: 0.978
Recall on test Data: 0.544

precision on training Data: 0.993
precision on test Data: 0.544
```

Hyper Parameter Tuning:

- Using Grid Search to tune the **best performing model - SVC**:

```
# Support Vector Classifier model
from sklearn.svm import SVC
from sklearn.model_selection import GridSearchCV

# defining parameter range
param_grid = {'gamma': [0.1],'kernel': ['rbf','linear']}

svc = GridSearchCV(SVC(), param_grid)

# fitting the model for grid search
svc.fit(X_train, y_train)
```

```
GridSearchCV(estimator=SVC(),
             param_grid={'gamma': [0.1], 'kernel': ['rbf', 'linear']})
```