

SPRINT 3

Date	12 November 2022
Team ID	PNT2022TMID15890
Project Name	Web Phishing Detection

Sprint Objective:

In Sprint 3 the team members develop the model for the application, that is used to predict the Phishing website, is to be created. Then they should integrate it with the UI i.e. the webpage created in Sprint 2.

Screenshots:

```
In [8]: x = df.iloc[:, 1:31].values
        y = df.iloc[:, -1].values

In [9]: from sklearn.model_selection import train_test_split
        x_train, x_test, y_train, y_test = train_test_split(x, y, test_size = 0.2, random_state = 0)

In [10]: from sklearn.ensemble import RandomForestClassifier
         rf = RandomForestClassifier(n_estimators=10, random_state=0, n_jobs=-1)
         rf.fit(x_train, y_train)

Out[10]: RandomForestClassifier(n_estimators=10, n_jobs=-1, random_state=0)

In [11]: y_pred = rf.predict(x_test)
         from sklearn.metrics import accuracy_score
         random_forest_accuracy = accuracy_score(y_test, y_pred)
         random_forest_accuracy

Out[11]: 0.9669832654907262
```

Sprint Result:

As planned for Sprint 3 the team members developed the model to be used by the application. This model is created using Random Forest Classifier algorithm. The model uses 80 percentage of the dataset downloaded and pre-processed in Sprint 1 as training set and the remaining 20 percentage as testing set. Then the accuracy for the testing model is found to be 96.69%. Finally, the model is integrated with the web page using python flask.