

Phishing Detection

Import numpy as np

```
from sklearn.ensemble import RandomForestClassifier as rfc from  
sklearn.model_selection import train_test_split
```

import feature extraction

```
def getResult(url):
```

#Importing dataset

```
data= np.loadtxt("dataset.csv", delimiter = "")
```

#Seperating features and labels

```
X = data[:, -1]
```

```
y = data[:, -1]
```

#Seperating training features, testing features, training labels & testing labels
x_train, x_test, y_train, y_test train_test_split(x, y, test_size = 0.2)

```
clf= rfc()
```

```
clf.fit(x_train, y_train) score clf.score(x_test, y_test)
```

```
print(score 100)
```

```
X_new = []
```

```
X_input url
```

```
X_new=feature extraction.generate_data_set(X_input)
```

```
X_new np.array(X_new).reshape(1,-1)
```

```
try:
```

```
prediction clf.predict(X_new)
```

```
if prediction -1:
```

```
    return "Phishing Url"
```

```
else:
```

```
    return "Legitimate Url"
```

```
except:
```

```
    return "Phishing Url"
```