Phishing URL Detection

The internet has integrated itself into our daily lives, but it has also given people the chance to engage in harmful acts like phishing in secret. Phishers attempt to trick their victims by employing social engineering techniques or building fake websites in order to obtain information such as account IDs, usernames, and passwords from people and businesses. Although several strategies have been put out to identify phishing websites, phishers have developed ways to circumvent these strategies. Machine learning is one of the best techniques for spotting these malicious behaviours. This is so that machine learning techniques can identify the common traits shared by the majority of phishing attacks.

The procedures shown in this notebook are:

1. Loading the data
2. Familiarizing with data & EDA
3. Visualizing the data
4. Splitting the data

Loading the data

Link: https://drive.google.com/file/d/18PuytIZWvNQCnMypKdaQjAqqQ0MRZ0uj/view?usp=sharing

```
#Loading data into data frame
```

Familiarizing with data & EDA

In this step, few data frame methods are used to look into the data and its features:

```
#Shape of data frame

#Listing the features of the dataset

#Information about the dataset

#Unique value in columns

#Droping index column

#Description of dataset
```

visualizing the data

Few plots and graphs are displayed to find how the data is distributed and the how features are related to each other:

```
#Correlation heatmap
#Pairplot for particular features
#Phishing Count in pie chart
```

Splitting the data

The data is split into train & test sets:

```
#Splitting the dataset into dependant and independent feature
#Splitting the dataset into train and test sets
```