Model Building

Start building Machine Learning Model

There are several Machine Learning algorithms to be used depending on the data you are going to process such as images, sound, text, and numerical values. The algorithms can be chosen according to the objective. As the dataset which we are using is a Classification dataset we can use the following algorithms,

- Logistic Regression
- Random Forest Regression / Classification
- Decision Tree Regression / Classification
- K-Nearest Neighbours
- Support Vector Machine

In order to get appropriate predictions, the dataset can be trained with any of the above algorithms.

Choose The Appropriate Model

Working with Logistic Regression model

Step - 1

Here, we will be initially considering Logistic Regression model and fit the data.

```
from sklearn.linear_model import LogisticRegression
lr=LogisticRegression()
lr.fit(x_train,y_train)
```

Step - 2

Check the metrics of the model

Here we will be evaluating the model built. We use the test set for evaluation. The test set is given to the model for prediction and prediction values are stored in another variable called y_pred1.

The actual and predicted values are compared to know the accuracy of the model using the accuracy score function from Sklearn metrics package.

Follow the below steps to find the accuracy of the model.

```
y_pred1=lr.predict(x_test)
from sklearn.metrics import accuracy_score
log_reg=accuracy_score(y_test,y_pred1)
log_reg
```

```
0.9167797376752601
```

The accuracy for logistic regression model for this dataset is 91.6%.

Note: You can use different classification models to know the performance and choose whichever works better.

Step - 3

Saving the model

The finalized model is now to be saved. We will be saving the model as a pickle or pkl file.

Use the command below to save the model.

```
import pickle
pickle.dump(lr,open('Phishing_Website.pkl','wb'))
```