

# PROJECT: Machine Learning

## Phishing Detector using KNN

### **Purpose of the Document**

The document has to specify the requirements for the project "Build a detector for Phishing websites (KNN)." Apart from specifying the functional and non-functional requirements for the project, it also serves as an input for project scoping.

### **Problem Statement**

The purpose of the project is to use one or more of the classification algorithms to train a model on the Phishing website dataset.

You are provided with the following resources that can be used as inputs for your model:

1. A collection of website URLs for 11000+ websites. Each sample has 30 website parameters and a class label identifying it as a phishing website or not (1 or -1).
2. Code template containing these code blocks:
  - a) Import modules (part 1)
  - b) Load data function + input/output field descriptions

You are expected to write the code for a binary classification model using Python Scikit-Learn that trains on the data and calculates the accuracy score on the test data.

## Project Guidelines

#	Exercises	Process
1	Initiation	Begin by extracting the ipynb file.
2	Exercise	<p>Build a phishing website classifier using KNN with number of neighbors as 5 and distance metric as "minkowski." Use 70% of data as training data and remaining 30% as test data.</p> <p>(Hint: Use Scikit-Learn library KNeighboursClassifier and refer to the KNN tutorial.)</p> <p>Print count of misclassified samples in the test data prediction as well as accuracy score of the model.</p>



