**Cybersecurity Project**

**Phishing Detection Using Machine Learning**

Goals and Objectives:

The primary objective of this project is to create a machine learning-based classifier that is both effective and robust, capable of accurately distinguishing between legitimate (safe) emails and fraudulent emails. The classifier endeavors to safeguard users from phishing assaults and improve email security by employing sophisticated machine learning algorithms.

Research Question:

How effective can machine learning algorithms be to classify the incoming mails into phishing and normal category to help protect user against frauds?

Which machine learning algorithm is efficient (SVM, Random Forest) in classifying the received emails based on accuracy, precision, recall and f1-score?

Data source:

Kaggle has a collection of emails that will be used to train our design. The link to the data source is: https://www.kaggle.com/datasets/subhajournal/phishingemails

The datasource has 11322 safe emails and 7312 phishing emails which will be adequate to conduct this research.