

## Project Initialization and Planning Phase

Date	10 june, 2024
Team ID	739879
Project Title	Detection of phishing websites from URLs
Maximum Marks	3 Marks

### Project Proposal (Proposed Solution) report

The proposal report aim to develop an intelligent system that can accurately detect phishing websites from URLs in real-time, leveraging machine learning techniques.

Project Overview	
Objective	The primary goal is to create a system that can correctly identify phishing websites and distinguish them from legitimate ones with a high degree of precision. This objective is the most critical aspect of the project, as it directly impacts the effectiveness of the phishing detection system.
Scope	The project focuses solely on developing a system to detect phishing websites from URLs using machine learning.
Problem Statement	
Description	The goal is to achieve high accuracy and minimal false positives, enabling effective protection of users from phishing threats
Impact	Improve online security, enhance user trust, reduce financial losses, increased awareness, contribution to cybersecurity research, protection of sensitive information by achieving these impacts, the project will make a significant contribution to the fight against phishing and cyber threats, enhancing online security and protecting users from these malicious attacks.
Proposed Solution	
Approach	Develop a machine learning model that analyzes various features of URLs, including domain name, subdomain, path, parameters, and HTML content, to detect phishing websites with high accuracy. Integrate the trained model into a web application or API that can take a URL as input and output a phishing detection result, continuous updating the model to adapt to evolving phishing tactics

Key Features	It is the combination of machine learning algorithms with various URL analysis techniques, including, domain name, URL structure, HTML,CSS, pythonscript to detect phishing websites with high accuracy
--------------	---

## Resource Requirements

Resource Type	Description	Specification/Allocation
<b>Hardware</b>		
Computing Resources	CPU/GPU specifications, number of cores	T4 GPU
Memory	RAM specifications	8 GB
Storage	Disk space for data, models, and logs	1 TB SSD
<b>Software</b>		
Frameworks	Python frameworks	Flask
Libraries	Additional libraries	scikit-learn, pandas, numpy, matplotlib, seaborn
Development Environment	IDE	Jupyter Notebook
<b>Data</b>		
Data	Source, size, format	Kaggle dataset, 614, csv UCI dataset, 690, csv