# AI-Powered Phishing & Threat Intelligence Detection System

Step 1: Problem Statement & Scope Document

## 1. Problem Statement

Phishing has emerged as one of the most persistent and damaging forms of cyberattacks in recent years. It involves deceptive tactics—such as spoofed emails, fraudulent URLs, and malicious SMS messages (smishing)—to trick users into divulging sensitive information including passwords, banking credentials, and personal data. Despite awareness campaigns, phishing continues to be effective due to its evolving nature and the sophistication of social engineering techniques.  
  
Traditional detection systems rely on static rule-based filters or blacklists, which are often reactive and fail to adapt to zero-day phishing threats. There is a pressing need for an intelligent, real-time system that can proactively identify and neutralize phishing attempts before users fall victim.  
  
Our project aims to bridge this gap by designing an AI-powered phishing detection system that integrates machine learning, real-time threat intelligence, and cybersecurity principles to detect and prevent phishing threats across various mediums.

## 2. Project Objectives

The core objectives of this project are as follows:

* Develop an AI/ML-based model capable of detecting phishing in emails, messages, and URLs with high precision.
* Integrate external threat intelligence APIs (e.g., VirusTotal, AbuseIPDB) to enhance detection accuracy and credibility.
* Design a RESTful API and lightweight backend for real-time inference and detection.
* Provide a user-friendly dashboard for analyzing flagged threats, monitoring detection logs, and managing alerts.
* Build a browser extension that scans and flags malicious links on web pages in real time. (Stretch Goal)
* Integrate a feedback loop where users can report false positives/negatives to improve model accuracy over time. (Stretch Goal)

## 3. Scope of the Project

In-Scope:

* Detection of phishing via:
* - Email content (subject, body text, sender information)
* - URLs (domain reputation, redirection behavior)
* - Short messages (SMS/text-based phishing)
* Machine learning model training using real-world phishing datasets
* Integration with public threat intelligence APIs
* Backend API for real-time inference
* Visualization dashboard for threat tracking
* Secure data handling and user privacy compliance

Out of Scope (for the current version):

* Detection of file-based malware (e.g., malicious attachments like PDFs, .exe, etc.)
* End-to-end secure email gateway integration
* Corporate-level SOC (Security Operations Center) automation
* Deep web or dark web threat correlation

## 4. Deliverables

* A trained and validated AI/ML model for phishing classification
* API for real-time threat detection
* Threat intelligence enrichment module
* Frontend dashboard (web interface)
* Optional: Browser extension for live detection
* Technical documentation and deployment manual

## 5. Target Audience / Use Cases

* Individual users browsing or accessing emails
* Small businesses with limited cybersecurity infrastructure
* Educational institutions for awareness and early detection
* Researchers and developers studying phishing trends

## 6. Success Metrics

* Model Accuracy, Precision, Recall ≥ 90%
* False Positive Rate < 5%
* Real-time API latency < 2 seconds
* Usability score from feedback survey ≥ 4/5