1. Problem Definition and Scope

Phishing attacks are a very common cybersecurity threat. The attacker often impersonates an organization that is trusted. Such attacks often lead to fraud, identity theft, and data breaches. Many people struggle to identify phishing attempts due to the increase in sophistication.

Target Application Domain and Significance

PhishGuard is designed for email security in an effort to detect phishing in Gmail and Outlook. It leverages machine learning and NLP to analyze email content, metadata, and embedded links to mitigate phishing threats.

Scope of Research

- The system will analyze email headers, content, and embedded links in an attempt to detect phishing
- Use machine learning models for classification
- Perform URL scanning
- **Limitation:** The accuracy will be dependent on the training dataset, rather than the evolving phishing attacks.

2. Project Division into Subtasks

1) **Data Collection**
    a) Collect and process phishing and legitimate emails
2) **Data Preprocessing**
    a) Avoid feature redundancy: avoid using variables that have strong correlations with each other to 1) keep the model simpler and improve interpretability (avoid overfitting), 2) because email datasets are large, using fewer features can speed up computation time.
    b) Splitting the data: split our data into an array containing features and another containing labels.
3) **Model Training**
    a) Normalizing the data: normalize the train and test features and then reduce the dimensionality of our data for further analysis.
    b) Analysis and visualization of data: Use a variety of plots and graphs to determine the number of components to use in analysis
    c) Balance the data: Analyze the data and weight the value of a correct classification in each class inversely to the occurrence of data points for that class in order to reduce the skewing of distinguishing between phishing and non-phishing.
4) **Email and URL Analysis**
    a) Scan embedded URLs for malicious domains and flag potential threats
    b) Analyze SPF (Sender Policy Framework), DKIM (DomainKeys Identified Mail), and DMARC (Domain based Message Authentication, Reporting, and Conformance) records to detect spoofing (impersonating)
5) **Deployment and Real Time Detection**
    a) Integrate PhishGuard with Gmail and Outlook as a browser extension for live detection

b) Implement real time alerts and indicators

3. Research Existing Technologies and Methodologies

a) Data Collection and Model Training

i) SpamAssassin

- Data Collection:
  SpamAssassin gathers spam and non-spam emails from a variety of sources. Its community of volunteers contributes samples from public spam archives, user submissions, and network monitoring.
- Model Training:
  Rather than using complex machine learning frameworks, SpamAssassin employs heuristic rules and a Bayesian spam filter. The Bayesian component learns from provided email examples, calculating the probability that certain words or patterns indicate spam. This learning is continuously refined as more data is added, enabling the filter to adapt to evolving spam tactics.

ii) PhishTank

- Data Collection:
  PhishTank is built on a community-driven model where users submit suspected phishing URLs. Once a URL is reported, a community of researchers and volunteers verifies its legitimacy. This crowdsourced verification process results in a robust, frequently updated database of phishing links.
- Model Training:
  While PhishTank itself does not train complex models, its curated data serves as a vital resource. Organizations can integrate PhishTank's data into their training pipelines, using the verified phishing URLs to train or fine-tune machine learning models that detect malicious links and identify phishing patterns in emails.

b) Email and URL Security Analysis

i) Email Header Analysis: Check SPF, DKIM, and DMARC records to verify sender authenticity. This helps in identifying spoofed emails and unauthorized senders.

ii) URL Inspection: Develop methods to parse and analyze embedded URLs within emails. Techniques include comparing URLs against known phishing blacklists and evaluating link structure or redirection patterns.

iii) Real-Time Scanning: Integrate real-time analysis tools that scan email content and URLs at the moment of receipt, providing immediate alerts if anomalies or potential threats are detected.

c) Deployment and Real Time Detection

i) Google Workspace Security - Provides strong phishing protection but is a paid feature, and is not customizable

ii) Microsoft Defender - Enterprise phishing protection, but lacks open source implementation; only works on microsoft ecosystem

iii) OpenPhish - Publicly available phishing detection, it is free but is limited to URL based detection

iv) References:

https://workspace.google.com/blog/identity-and-security/protecting-you-against-phishing
https://learn.microsoft.com/en-us/defender-office-365/anti-phishing-protection-about
https://openphish.com/