UNIVERSITY OF GHANA

COLLEGE OF BASIC AND APPLIED SCIENCES

FACULTY OF PHYSICAL AND MATHEMATICAL SCIENCES



ADAPTIVE PHISHING DETECTION USING MACHINE LEARNING

BY

NAME

(ID. NO. XXXXXXX)

A DISSERTATION SUBMITTED TO THE SCHOOL OF GRADUATE STUDIES IN PARTIAL FULFILMENT OF THE AWARD OF DEGREE OF MASTER OF SCIENCE IN COMPUTER SCIENCE

DEPARTMENT OF COMPUTER SCIENCE

MONTH 2025

UNIVERSITY OF GHANA

COLLEGE OF BASIC AND APPLIED SCIENCES

FACULTY OF PHYSICAL AND MATHEMATICAL SCIENCES

ADAPTIVE PHISHING DETECTION USING MACHINE LEARNING

BY

NAME

(ID. NO. XXXXXXX)

DEPARTMENT OF COMPUTER SCIENCE

MONTH 2025

# DECLARATION

I hereby declare that except where speciﬁc references are made to the work of others, the contents of this dissertation are original and have not been submitted in whole or in part for consideration for any other degree or qualiﬁcation in this, or any other University.

NAME

Month Year

………………………………………………… Date:

NAME (Student)

This dissertation was under the supervision of:

………………………………………………… Date:

DR. JUSTICE KWAME APPATI

# DEDICATION

# ACKNOWLEDGMENTS

fff

# ABSTRACT

# TABLE OF CONTENT

[DECLARATION iii](#_Toc199388677)

[DEDICATION iv](#_Toc199388678)

[ACKNOWLEDGMENTS v](#_Toc199388679)

[ABSTRACT vi](#_Toc199388680)

[TABLE OF CONTENT vii](#_Toc199388681)

[LIST OF FIGURES ix](#_Toc199388682)

[LIST OF TABLES x](#_Toc199388683)

[LIST OF ABBREVIATIONS xi](#_Toc199388684)

[CHAPTER 1 - INTRODUCTION 1](#_Toc199388685)

[1.1 Background of Study 1](#_Toc199388686)

[1.2 Problem Statement 2](#_Toc199388687)

[1.3 Objectives 3](#_Toc199388688)

[1.4 Outline of Methodology 3](#_Toc199388689)

[1.5 Justification 4](#_Toc199388690)

[1.6 Outline of Dissertation 5](#_Toc199388691)

[CHAPTER 2 - LITERATURE REVIEW 6](#_Toc199388692)

[2.1 Overview 6](#_Toc199388693)

[2.2 Machine Learning Models for Phishing Detection 6](#_Toc199388694)

[2.3 Feature Engineering in Phishing Detection 7](#_Toc199388695)

[2.4 State-Of-The-Art Deployment Tools in Phishing Detection 7](#_Toc199388696)

[CHAPTER 3 - METHODOLOGY 9](#_Toc199388697)

[3.1 Overview 9](#_Toc199388698)

[3.1 Data Collection 9](#_Toc199388699)

[3.1 Feature Engineering 9](#_Toc199388700)

[3.1 Proposed Adaptive Phishing Detection Model 9](#_Toc199388701)

[CHAPTER 4 – RESULTS AND DISCUSSION 11](#_Toc199388702)

[4.1 Overview 11](#_Toc199388703)

[CHAPTER 5 – CONCLUSION AND FUTURE WORKS 12](#_Toc199388704)

[5.1 Overview 12](#_Toc199388705)

[REFERENCES 14](#_Toc199388706)

# LIST OF FIGURES

# LIST OF TABLES

# LIST OF ABBREVIATIONS

# CHAPTER 1 - INTRODUCTION

## 1.1 Background of Study

Phishing attacks are among the most persistent and damaging forms of cybercrime. They exploit both technical vulnerabilities and human psychology to deceive individuals into divulging sensitive information such as login credentials and financial data (Mosa et al., 2023; Opara et al., 2024). These attacks leverage sophisticated emails, websites, and instant messages that impersonate trusted organizations, using tactics such as spoofing, domain squatting, and URL obfuscation to evade detection (Atawneh & Aljehani, 2023; Karim et al., 2023).

The threat landscape has evolved with the emergence of advanced phishing techniques, including spear phishing (targeting specific individuals), whaling (targeting high-profile executives), vishing (voice phishing), and smishing (SMS phishing), all of which expand the attack surface and challenge traditional security solutions (Gholampour & Verma, 2023; Hassan, 2024). Attackers exploit weaknesses in email protocols, browser vulnerabilities, and user behavior, often automating campaigns using readily available phishing kits (Al-Subaiey et al., 2024; Doshi et al., 2023).

Recent years have seen a surge in the use of machine learning and AI by both attackers and defenders, increasing the complexity and sophistication of phishing schemes (Gholampour & Verma, 2023). The proliferation of digital communications and online transactions has amplified the impact of phishing, resulting in significant financial losses, data breaches, and reputational harm (Doshi et al., 2023; Mosa et al., 2023). Phishing remains a leading vector for cyber incidents globally, highlighting the urgent need for adaptive, robust, and scalable detection mechanisms that leverage advanced machine learning techniques (Karim et al., 2023; Linh et al., 2024).

To address these evolving challenges, the development of adaptive phishing detection models is essential. Such models require comprehensive datasets, advanced feature engineering, and the integration of state-of-the-art machine learning algorithms to enhance detection accuracy and adaptability to emerging threats (Goud & Mathur, n.d.; Liu et al., 2021).

Despite the proliferation of machine learning and AI-based solutions, a persistent challenge in phishing detection research is the reliance on static or outdated datasets. Many existing studies use datasets that do not reflect the rapidly evolving tactics of attackers, resulting in models that may perform well in controlled experiments but fail to generalize to real-world scenarios. While this research acknowledges the importance of comprehensive and up-to-date datasets, it specifically addresses the need for advanced feature engineering and real-time adaptability to improve detection accuracy and resilience against emerging phishing threats.

## 1.2 Problem Statement

Despite advancements in phishing detection, attackers continuously adapt their tactics, making many existing systems ineffective against new and sophisticated threats. Current limitations include insufficient diversity in feature engineering and a lack of real-time adaptability. As a result, detection systems often suffer from high false positive rates, limited scalability, and poor generalization to novel phishing strategies. There is a critical need for a robust and adaptive phishing detection approach that leverages advanced feature engineering (including URL, HTML, and behavioral features), and rigorous validation to achieve high accuracy, low false positive rates, and resilience against evolving phishing techniques.

This research seeks to bridge these gaps by systematically integrating state-of-the-art machine learning and feature engineering methods to create a scalable and adaptive phishing detection system.

## Objectives

The objectives of this research are:

1. To design and implement advanced feature engineering methods, including URL, HTML, and behavioural features, to enhance model performance over established baselines.
2. To develop a machine learning-based phishing detection model using the engineered features and diverse datasets.
3. To deploy and evaluate the adaptive phishing detection system in a simulated real-time environment, assessing its scalability, adaptability, and response to emerging phishing tactics.

## Outline of Methodology

The methodology used directly builds on and extends the state-of-the-art in phishing detection. It is consisting of three subsections - a baseline replication and data preparation, advanced feature engineering and model enhancement, deployment and real-time evaluation. A baseline for phishing detection is replicated and data prepared. The dataset used was acquired from PhishTank, Kaggle repository. The data is cleaned, normalized, split into train and test data. The feature extraction techniques and baseline model (XGBoost) were used to establish a performance benchmark as described by Aljofey et al. (2022). The next is the feature engineering and model enhancement. Over here, we design and implemented additional feature extraction methods, including behavioral and language-independent features. Then we integrated these with URL and HTML features. We then Trained and validated machine learning models, comparing results to the baseline and targeting measurable improvements in F1-score and accuracy. Finally, we deployed monitored in real time. At this stage, only the best-performing model in a simulated real-time environment was deployed. We also measured system scalability, latency, and adaptability to new simulated phishing tactics.

## Justification

Phishing remains one of the most prevalent and damaging cyber threats faced by businesses today. Existing detection systems often struggle to keep up with the rapidly evolving tactics of attackers, leading to costly breaches and loss of customer trust. This research directly addresses these challenges by developing an adaptive phishing detection system that leverages advanced feature engineering—including URL, HTML, and behavioral features—and real-time adaptability. By providing a scalable and robust solution that can detect novel and sophisticated phishing attacks, this work offers industry practitioners a practical tool to enhance their cybersecurity posture, reduce false positives, and respond more effectively to emerging threats. The integration of user feedback mechanisms further ensures that the system can continuously learn and adapt, making it highly relevant for deployment in dynamic, real-world environments.

This research advances the academic field of cybersecurity and machine learning by systematically addressing gaps in feature diversity and model adaptability for phishing detection. The study introduces and rigorously evaluates new feature engineering strategies, including the integration of behavioral and language-independent features, which are underexplored in current literature. By benchmarking against established baselines and deploying the system in a simulated real-time environment, the research provides valuable empirical evidence and methodological innovations. These contributions not only extend the theoretical understanding of adaptive security systems but also offer a foundation for future research on robust, data-driven approaches to cyber threat detection.

## Outline of Dissertation

This dissertation is organized into seven chapters. Chapter 1 is introduction. It provides the background, problem statement, objectives, methodology overview, justification, and the overall structure of the dissertation. Chapter 2 is literature review. It reviews existing research on phishing detection, machine learning techniques, feature engineering, and highlights the gaps addressed by this work. Chapter 3 is methodology. It details the research design, data collection, feature engineering, model development, evaluation strategies, and user feedback integration. Chapter 4 is proposed model. It describes the architecture and workflow of the adaptive phishing detection system, including the integration of advanced features and machine learning algorithms. Chapter 5 is results and discussion. It presents the experimental setup, evaluation metrics, results, and a discussion comparing the findings with existing systems and literature. Chapter 6 is the conclusion and future works. It summarizes the main findings, contributions, and limitations of the research, and outlines directions for future work. Chapter 7 is research timeline. It provides a detailed timeline for the completion of the research project, including key milestones and deliverables.

# CHAPTER 2 - LITERATURE REVIEW

## 2.1 Overview

Phishing detection research has rapidly evolved, with a diverse array of machine learning and deep learning approaches proposed in recent years. This review synthesizes findings from a broad set of studies and how this research aims to extend the state of the art.

## 2.2 Machine Learning Models for Phishing Detection

Numerous studies demonstrate the effectiveness of machine learning and deep learning models for phishing detection. Aljofey et al. (2022) achieved 96.76% accuracy and a 1.39% false positive rate using XGBoost on a large dataset with URL and HTML features. Other works, such as Mosa et al. (2023), explored neural networks and ensemble methods, reporting similar or higher performance. Deep learning models (CNN, LSTM, BERT) are increasingly used for both website and email phishing detection (Atawneh & Aljehani, 2023; Hassan, 2024), often exceeding 97% accuracy. However, several authors (Gholampour & Verma, 2023; Misra & Rayz, 2022) stress the importance of model robustness and generalizability, as overfitting to specific datasets can lead to poor real-world performance.

Data diversity and validation strategies are critical. While many studies use public datasets (PhishTank, Kaggle, Enron, etc.), Aljofey et al. (2022) combined benign and phishing samples from multiple sources for a more representative dataset. Recent literature also emphasizes the use of adversarial and out-of-domain samples to test model resilience (Gholampour & Verma, 2023; Opara et al., 2024).

## 2.3 Feature Engineering Techniques for Phishing Detection

Numerous studies demonstrate the effectiveness of machine learning and deep learning models for phishing detection. Aljofey et al. (2022) achieved 96.76% accuracy and a 1.39% false positive rate using XGBoost on a large dataset with URL and HTML features. Other works, such as Mosa et al. (2023), explored neural networks and ensemble methods, reporting similar or higher performance. Deep learning models (CNN, LSTM, BERT) are increasingly used for both website and email phishing detection (Atawneh & Aljehani, 2023; Hassan, 2024), often exceeding 97% accuracy. However, several authors (Misra & Rayz, 2022; Gholampour & Verma, 2023) stress the importance of model robustness and generalizability, as overfitting to specific datasets can lead to poor real-world performance.

Data diversity and validation strategies are critical. While many studies use public datasets (PhishTank, Kaggle, Enron, etc.), Aljofey et al. (2022) combined benign and phishing samples from multiple sources for a more representative dataset. Recent literature also emphasizes the use of adversarial and out-of-domain samples to test model resilience (Gholampour & Verma, 2023; Opara et al., 2024).

## 2.4 State-Of-The-Art Deployment Tools in Phishing Detection

Real-world deployment is a growing focus. Linh et al. (2024) built a browser extension for real-time phishing detection, while Doshi et al. (2023) and Al-Subaiey et al. (2024) evaluated web-based and dual-layer systems for scalability and adaptability. Aljofey et al. (2022) highlighted the need for client-side, real-time detection. Robustness to adversarial attacks and adaptability to new phishing tactics are ongoing challenges (Gholampour & Verma, 2023).

The literature shows strong progress in phishing detection, but persistent challenges include dataset diversity, feature robustness, adaptability, and real-world deployment. Building on Aljofey et al. (2022), this research will replicate and extend their approach by incorporating advanced, language-independent, and behavioral features, and by evaluating the system in a simulated real-time environment.

# CHAPTER 3 - METHODOLOGY

## 3.1 Overview

This chapter details the stepwise methodology for developing, validating, and deploying an adaptive phishing detection system, directly building on the base paper and addressing identified gaps.

## 3.1 Data Collection

Public phishing and legitimate datasets (PhishTank, Kaggle, Enron, SpamAssassin, IWSPA, etc.), and the dataset from Aljofey et al. (2022) is used for this study. The data types include URLs, emails, HTML content, behavioral logs, and metadata. We pre-processed the data by cleaning (deduplication, normalization), handling missing values, and class balancing (SMOTE/ADASYN).

## 3.1 Feature Engineering

The baseline features were URL and HTML features (TF-IDF, hyperlink statistics) were replicated according to Aljofey et al. (2022). Advanced features were added. These include behavioral (e.g., user interaction, click patterns) and language-independent features, plus third-party reputation. We used feature selection and dimensionality reduction (recursive elimination, genetic algorithms, PCA) tehniques.

## 3.1 Proposed Adaptive Phishing Detection Model

The baseline modele, XGBoost is reproduced as in Aljofey et al. (2022) for benchmarking. We trained and compared Random Forest, SVM, CNN, LSTM, BERT, and hybrid/ensemble models. After training we performed cross-validation, hyperparameter tuning, and direct comparison with baseline.The performance is assessed using metrics such as accuracy, F1-score, and ROC-AUC. We Integrated into a simulated environment and incoporated user feedbackloop. The workflow of the proposed model is summariesd is figure 3.1.

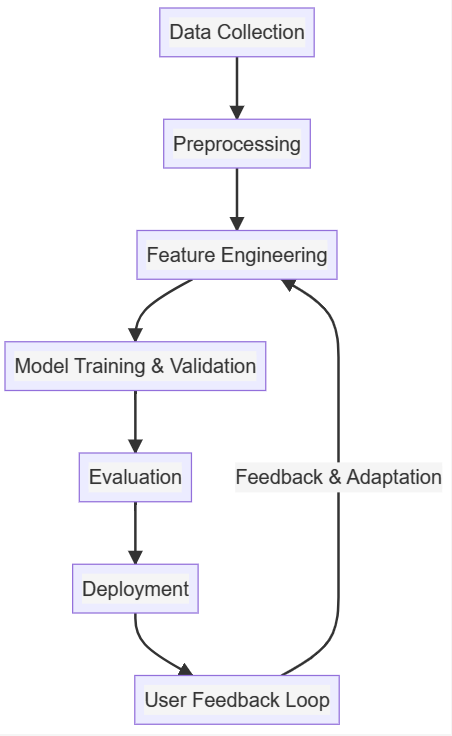


Figure : Figure 3.1 is a flowchart of the proposed adaptive phishing detection model pipeline. The system starts with data collection, proceeds through preprocessing and feature engineering, then model training and evaluation, and finally deployment. A user feedback loop enables continuous adaptation to new phishing threats.

# CHAPTER 4 – RESULTS AND DISCUSSION

## 4.1 Overview

# CHAPTER 5 – CONCLUSION AND FUTURE WORKS

In this section, we provide details of the various experiments we performed. We also present out findings and discuss our observations as well.

## 5.1 Overview

# REFERENCES

Aljofey, A., Jiang, Q., Rasool, A., Chen, H., Liu, W., Qu, Q., & Wang, Y. (2022). An effective detection approach for phishing websites using URL and HTML features. *Scientific Reports*, *12*(1). https://doi.org/10.1038/s41598-022-10841-5

Al-Subaiey, A., Al-Thani, M., Abdullah Alam, N., Antora, K. F., Khandakar, A., & Uz Zaman, S. A. (2024). Novel interpretable and robust web-based AI platform for phishing email detection. *Computers and Electrical Engineering*, *120*. https://doi.org/10.1016/j.compeleceng.2024.109625

Atawneh, S., & Aljehani, H. (2023). Phishing Email Detection Model Using Deep Learning. *Electronics (Switzerland)*, *12*(20). https://doi.org/10.3390/electronics12204261

Doshi, J., Parmar, K., Sanghavi, R., & Shekokar, N. (2023). A comprehensive dual-layer architecture for phishing and spam email detection. *Computers and Security*, *133*. https://doi.org/10.1016/j.cose.2023.103378

Gholampour, P. M., & Verma, R. M. (2023). Adversarial Robustness of Phishing Email Detection Models. *IWSPA 2023 - Proceedings of the 9th ACM International Workshop on Security and Privacy Analytics*, 67–76. https://doi.org/10.1145/3579987.3586567

Goud, N. S., & Mathur, A. (n.d.). Feature Engineering Framework to detect Phishing Websites using URL Analysis. In *IJACSA) International Journal of Advanced Computer Science and Applications* (Vol. 12, Issue 7). www.ijacsa.thesai.org

Hassan, M. (2024). Evaluation of Deep Learning Algorithms in Comparison for Phishing Email Identification. In *Applied Mathematics on Science and Engineering* (Vol. 1, Issue 1).

Karim, A., Shahroz, M., Mustofa, K., Belhaouari, S. B., & Joga, S. R. K. (2023). Phishing Detection System Through Hybrid Machine Learning Based on URL. *IEEE Access*, *11*, 36805–36822. https://doi.org/10.1109/ACCESS.2023.3252366

Linh, D. M., Hung, H. D., Chau, H. M., Vu, Q. S., & Tran, T. N. (2024). Real-time phishing detection using deep learning methods by extensions. *International Journal of Electrical and Computer Engineering*, *14*(3), 3021–3035. https://doi.org/10.11591/ijece.v14i3.pp3021-3035

Liu, D. J., Geng, G. G., Jin, X. B., & Wang, W. (2021). An efficient multistage phishing website detection model based on the CASE feature framework: Aiming at the real web environment. *Computers and Security*, *110*. https://doi.org/10.1016/j.cose.2021.102421

Misra, K., & Rayz, J. T. (2022). LMs go Phishing: Adapting Pre-trained Language Models to Detect Phishing Emails. *Proceedings - 2022 IEEE/WIC/ACM International Joint Conference on Web Intelligence and Intelligent Agent Technology, WI-IAT 2022*, 135–142. https://doi.org/10.1109/WI-IAT55865.2022.00028

Mosa, D. T., Shams, M. Y., Abohany, A. A., El-Kenawy, E. S. M., & Thabet, M. (2023). Machine Learning Techniques for Detecting Phishing URL Attacks. *Computers, Materials and Continua*, *75*(1), 1271–1290. https://doi.org/10.32604/cmc.2023.036422

Opara, C., Chen, Y., & Wei, B. (2024). Look before you leap: Detecting phishing web pages by exploiting raw URL and HTML characteristics. *Expert Systems with Applications*, *236*. https://doi.org/10.1016/j.eswa.2023.121183