IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.102 ⁑ Peer-reviewed & Refereed journal ⁑ Vol. 14, Issue 4, April 2025

DOI: 10.17148/IJARCCE.2025.14480

AI-Driven Phishing Detection and Awareness

Mr. S. Dinkar Jose[1], Sri Arvind M[2], Shyam S[3]

Assistant Professor Department of Computer Science and Engineering Anand Institute of Higher Technology, kazhipattur Chennai[1]

Student, Department of Computer Science and Engineering Anand Institute of Higher Technology, kazhipattur Chennai[2-3]

Abstract: As cyber threats continue to rise, individuals and businesses face growing risks from phishing attacks. This project introduces an AI-driven Phishing Detection and Awareness Platform, designed to provide a comprehensive and interactive approach to phishing awareness, detection, and prevention. Developed as a web-based solution, the platform integrates phishing detection (AI), password breach checking, email analysis, URL scanning, and user awareness training. The phishing detection module analyzes emails and URLs using fine- tuned transformer models, while the breach checker leverages the HaveIBeenPwned API. quizzes and interactive simulations enhance user training, ensuring engaging and practical learning. Additionally, the platform includes real-time alerts, detailed reports, and secure data storage to strengthen cybersecurity measures. By offering an all-in-one, accessible, and AI-powered security solution, this project empowers users to recognize, prevent, and respond to phishing threats more effectively.

Keywords: AI-powered phishing detection, Phishing simulation, cybersecurity training, Real-time threat alerts, AI-driven email analysis, Automated phishing awareness, Secure password management, URL threat analysis, Cybersecurity education platform, AI-powered risk assessment, Interactive phishing scenarios, Data-driven phishing insights, AI in cybersecurity awareness, Smart anti-phishing solutions.

I. INTRODUCTION

The rise of artificial intelligence (AI) has revolutionized cybersecurity, making threat detection and prevention more efficient. This project introduces an AI-powered web platform designed to combat phishing attacks through detection, simulation,and awareness training.The system integrates AI- driven phishing detection to analyze emails and URLs, identifying potential threats with high accuracy. Integrated password breach checking and URL scanning enhance real-time threat identification using trusted cybersecurity APIs.

A phishing simulation module generates and tracks phishing awareness tests, helping users recognize deceptive tactics. training features, including interactive quizzes and AI-generated phishing scenarios, provide an engaging learning experience to strengthen user vigilance. To ensure security and scalability, the platform incorporates secure data storage for logs, user activity, and AI model improvements.

A real-time alert and reporting system notifies users of detected threats and generates detailed security insights. The web-based dashboard offers an intuitive interface, allowing individuals and businesses to access phishing detection tools, training modules, and threat reports seamlessly. A review of existing cybersecurity approaches highlights the need for an integrated and accessible anti-phishing solution. By combining AI technology and user-focused training, this project aims to enhance phishing awareness, improve threat detection accuracy, and empower users to mitigate cyber risks effectively.

DOI: 10.17148/IJARCCE.2025.14480

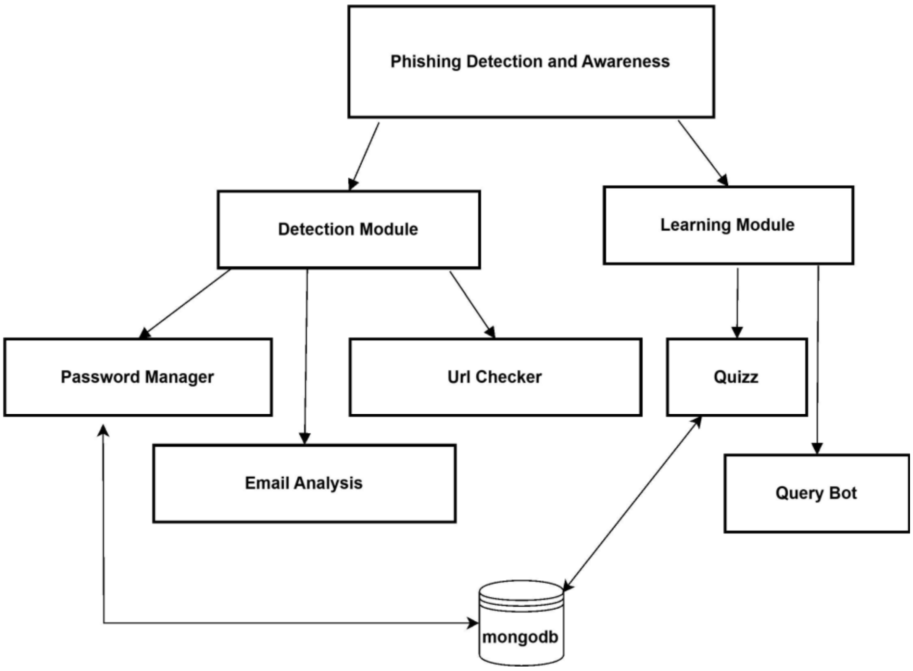

Figure 1:Architecture Diagram

II. RELATEDWORK

Artificial Intelligence (AI) has significantly transformed cybersecurity by enhancing phishing detection, awareness training, and user protection. Several studies have explored AI-driven solutions for phishing detection, awareness programs, and phishing simulations. Rahim et al. (2020) investigate phishing susceptibility among undergraduate students through phishing simulations, revealing that younger internet users are highly vulnerable. Their findings highlight the importance of phishing simulations in improving awareness, which aligns with this project's goal of providing AI-driven phishing simulations and real-time alerts to enhance learning and protection. [1]

Similarly, the CyberPhishing platform (2015) introduces a game-based phishing awareness system, demonstrating that gamification improves engagement and learning outcomes. This supports the phishing training module in this project, which integrates interactive quizzes, real-world phishing scenarios, and adaptive difficulty levels. Azman & Kob (2020) further emphasize that phishing simulations improve awareness over time, validating the project's approach of continuous phishing testing and adaptive learning. [2]

In terms of AI-based phishing detection, an AI-enhanced phishing detection system (2024) explores the use of machine learning models such as SVM, Random Forest, LSTM, and CNN to analyze phishing emails, achieving 96-99% accuracy. The study confirms the effectiveness of AI-powered email and URL analysis, aligning with this project's AI-driven phishing detection module, which automates real-time email and URL analysis. [3]

Furthermore, comparative studies on phishing awareness among employees (2021) highlight that non-technical staff are more vulnerable to phishing compared to IT professionals. The findings support this project's adaptive, role-based phishing awareness training, which tailors simulations and lessons based on user knowledge levels. Similarly, research on phishing education (2022) shows that interactive training improves awareness and retention better than traditional simulations. This supports the need for an AI-driven platform integrating training, AI phishing detection, and continuous simulations for enhanced cybersecurity. [4][5]

Overall, existing research underscores the transformative role of AI in phishing detection, awareness training, and real-time cybersecurity solutions. By integrating AI-powered phishing detection, continuous simulations, learning, and enterprise- level insights, this project builds on prior studies to offer a comprehensive, proactive, and adaptive cybersecurity solution.

Author(s)	Tool	Description	Keyfeatures
Rahim & Feninferi na (2021)	Phishing Simulation	Analyzed phishing susceptibility among undergraduat e students through simulated phishing attacks.	Behavioral phishing analysis User susceptibility tracking Email-based phishing simulation
M.L. Hale, F.R Gamble	Game-Based Phishing Awareness Platform	Developed an interactive game- based phishing awareness system to test users' ability to recognize phishing attempts.	phishing training Real-time feedback on phishing attempts Adaptive difficulty levels
P. Chinna samy,P. Krishna moorthy	AI Models (SVM, Random Forest, LSTM, CNN)	Integrated machine learning models to detect phishing emails, URLs, and attachments with high accuracy..	AI-driven phishing detection Email and URL scanning Threat score assignment
T. Daengsi, P. Wuttiditt achotti	Phishing Awareness Study	Investigated phishing awareness differences among IT and non-IT employees across various departments	Phishing susceptibility analysis Department- based security training recommendations

Table 2.1 Related works of authors

III. TECHNOLOGIES USED IN EXISTING PLATFORM

AI-Powered Phishing Detection

Existing phishing detection systems, including Google Safe Browsing, Microsoft Defender SmartScreen, and Cisco Umbrella, leverage AI models like Random Forest, SVM, and Neural Networks (LSTM, CNN) to detect phishing emails and URLs. These platforms analyze email headers, domain reputation, text patterns, and behavioral anomalies to classify phishing threats in real time, supporting robust and adaptive detection strategies.

DOI: 10.17148/IJARCCE.2025.14480

Phishing Awareness Training
Phishing simulation and awareness platforms like KnowBe4, PhishMe (now Cofense), and Wombat Security use interactive phishing simulations, training modules, and behavioral analytics to educate users. These systems employ adaptive learning algorithms and real-world phishing scenarios to improve phishing recognition skills among employees and individuals.

Phishing Simulation and Awareness Testing :
Platforms such as Gophish, Lucy Security, and Infosec IQ enable organizations to conduct phishing attack simulations. These tools provide phishing email templates, customizable simulations, and behavior tracking dashboards to measure phishing susceptibility and build user resilience against cyber threats.

Natural Language Processing (NLP) for Email Analysis Services like Google Cloud Natural Language API, IBM Watson NLP, and OpenAI's GPT-based models process email content, sentiment, and contextual patterns to detect phishing attempts. NLP-driven solutions help identify social engineering tactics, impersonation strategies, and fraudulent communications embedded within emails.

Secure Password Management and Authentication Existing platforms like LastPass, Bitwarden, and 1Password provide AES-256 encryption, biometric authentication, and zero-knowledge architecture for secure password storage. Additionally, Google’s Password Manager and Microsoft Edge Password Monitor integrate phishing prevention features by alerting users about compromised credentials.

Cloud-Based Data Security and Threat Intelligence Phishing detection platforms rely on cloud security services such as AWS Shield, Microsoft Defender for Cloud, and Google Chronicle to store, analyze, and secure threat intelligence data. These platforms use threat intelligence feeds, behavior-based anomaly detection, and automated response mechanisms to mitigate phishing risks.

Web-Based Dashboards for Phishing Analytics and Reporting
Enterprise phishing protection solutions, including Microsoft Security Center, Google Workspace Security Dashboard, and Cofense Vision, provide AI-driven analytics, phishing campaign tracking, and user susceptibility reporting. These platforms use Power BI, Tableau, and Google Data Studio to help security teams monitor phishing threats, assess training effectiveness, and optimize cybersecurity strategies.

IV. IMPLEMENTATION

The implementation of the AI-driven Phishing Detection and Awareness Platform follows a well-structured approach utilizing modern technologies to ensure accuracy, scalability, and security in phishing detection, awareness training, and phishing simulations. The system is divided into two main components: real-time phishing detection and interactive phishing awareness training.

The detection module is developed using Streamlit, offering a lightweight and interactive interface for users to perform both manual and automated email analysis. This module integrates a fine-tuned DistilBERT model accuracy of ~ 80% figure 3 pointing out the training and loss .The model is used to analyze the email body text, identifying phishing attempts through semantic understanding and contextual cues. In parallel, VirusTotal's API is used to analyze URLs extracted from emails, allowing real-time detection of malicious links using a multi-engine threat database.

Additionally, the platform includes a password management feature that promotes strong credential hygiene. It combines a password strength checker with integration of the Have I Been Pwned (HIBP) API, which alerts users if their credentials have appeared in known data breaches.

The phishing awareness and training module is designed using vanilla HTML and CSS, ensuring fast loading and broad compatibility. This module includes quizzes to test phishing knowledge and a query-based educational chatbot, powered by NLP, which helps users understand phishing concepts in conversational form.

The backend is implemented in Python, supporting seamless integration of machine learning inference, email parsing, URL scanning, and password validation. MongoDB is used as the primary database for storing email logs, phishing detection results, quiz responses, and user-related activity, offering flexibility and horizontal scalability for growing datasets.

DOI: 10.17148/IJARCCE.2025.14480

Security is strengthened through OTP-based authentication using Twilio or Firebase Auth, ensuring that only verified users can access sensitive modules. All stored phishing data is protected using AES-256 encryption, supporting compliance with data security standards and policies.

A key component is the automated report generation feature, developed using Python NLP libraries and ReportLab, which extracts insights from phishing analyses and outputs them into well-structured PDF reports. These are securely stored in AWS S3 or Google Cloud Storage, ensuring accessibility and durability.

With MongoDB as the core database, the system efficiently handles logs, training data, and detection outcomes while supporting fast query execution. Its NoSQL structure also allows easy scalability as phishing analysis volumes grow.
Overall, the platform unifies AI-driven phishing detection, phishing simulations, interactive training, secure password practices, and real-time analytics into a single, cohesive system. By leveraging DistilBERT, VirusTotal, HIBP, and modern web technologies, the solution helps users and organizations stay resilient against phishing threats while improving cybersecurity awareness through automation, data- driven insights, and proactive training mechanisms.

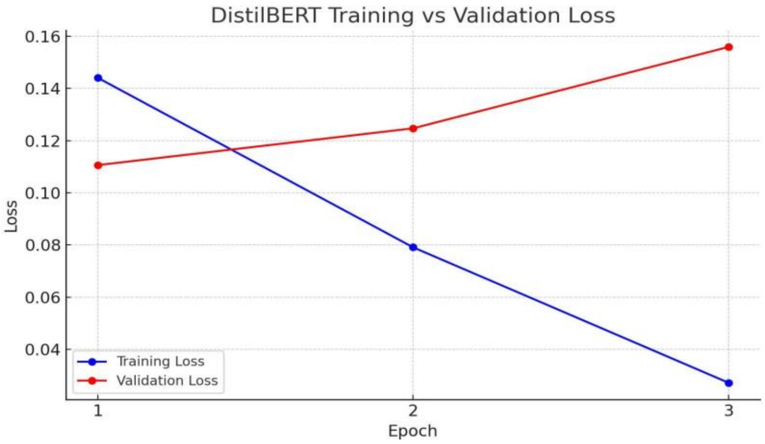

Figure 3 Training data of finetunning distilbert

Table 3.1 Training data table

Epoch	Training Loss	Validation
1	0.1441	0.1106
2	0.0791	0.1247
3	0.0272	0.1559

Table 3.2Tools used in implementation

Tools	Reason	Usage
Streamlit	Lightweight and interactive UI framework	Build the phishing detection interface for manual and automated email analysis
Node.js / Django	Scalable and efficient back- end framework	API development, request handling, and business logic implementation
MongoDB	Flexible NoSQL database with high scalability	Store inspection data, user details, and AI generated reports

DOI: 10.17148/IJARCCE.2025.14480

DistilBERT	Advanced AI model for phishing detection	Analyze phishing emails, URLs, and file attachments using AI-based detection models
Natural Language Processing (NLP)	Text and content analysis for phishing detection	Detect social engineering tactics in phishing emails and messages
Twilio / Firebase Auth	Secure user authentication	Implement OTP-based verification for login and approvals
Python (FastAPI / Flask)	Fast and scalable backend framework	Handle API requests, machine learning inference, VirusTotal queries, and password checks
Virus Total API	External URL analysis service	Scan email-extracted URLs for malicious threats in real time
HIBP API	Breached credential detection	Alert users if their passwords or emails appear in known data breaches

V. CONCLUSION

The AI-driven Phishing Detection and Awareness Platform enhances cybersecurity through AI-powered threat detection, automated simulations, and interactive training. Fine-tuned on the DistilBERT model using a hybrid dataset of ERAON and custom phishing data, the system ensures accurate detection of phishing emails, URLs, and suspicious content. Python-based backend services manage model inference, breach checks via Have I Been Pwned (HIBP), and VirusTotal integration for URL analysis. MongoDB provides scalable data management, while real-time alerts and visual analytics enable timely threat response. The React-based admin dashboard delivers actionable AI-driven insights, allowing users and organizations to monitor threats and improve cybersecurity strategies. By combining advanced NLP, machine learning, and secure authentication, this platform modernizes phishing awareness, reduces cyber risks, and empowers users to defend against evolving phishing attacks.

VI. FUTURE ENHANCEMENTS MOBILE APPLICATION

A dedicated mobile app for Android and iOS platforms will be developed to allow users to analyze emails, check suspicious URLs, and access training modules directly from their smartphones. This will improve user reach and convenience, enabling phishing protection on the go.

Browser Extention

A lightweight browser extension (e.g., for Chrome and Firefox) will be added to provide real-time phishing detection while users browse the web. The extension will automatically scan URLs, detect fake login pages, and issue immediate warnings, reinforcing user protection during daily web activities.

Multilingual Nlp Capabilities

To cater to a global audience, support for multiple languages will be integrated. Multilingual versions of the phishing detection model and training content will help non-English speakers better understand and defend against phishing attacks.

DOI: 10.17148/IJARCCE.2025.14480

Expanded Threat Intelligence Integration

The platform will incorporate additional threat intelligence APIs and services beyond VirusTotal (e.g., Cisco Talos, Google Safe Browsing) to enhance the accuracy and depth of URL and domain reputation checks.

Behavioral Analytics for Anomaly Detection

User behavior tracking (e.g., click patterns, response times) will be employed to detect anomalies that may indicate potential phishing exposure or susceptibility. This data will also be used to tailor training content to individual user profiles.

Gamification of Awareness Training

To make phishing education more engaging, gamified elements such as leaderboards, achievement badges, and timed challenges will be added to the quiz and simulation modules. This approach will encourage repeated use and better knowledge retention.

Organizational Admin Dashboard

A centralized dashboard for administrators in enterprise environments will be introduced. It will provide visibility into phishing incidents, employee training progress, and risk scores across departments, aiding in informed decision-making and targeted awareness campaigns.

AI Chatbot Enhancement

The current educational chatbot will be upgraded with a more advanced transformer-based conversational model (e.g., fine- tuned GPT or T5), enabling more natural, informative, and context-aware dialogues with users seeking phishing guidance.

REFERENCES

[1]. T. N. Hanis binti Tuan Kob, F. A. Rahim and F. Azman, "Phishing Attack Simulation: Measuring Susceptibility among Undergraduate Students,"2020 8th International Conference on Information Technology and Multimedia (ICIMU), Selangor, Malaysia, 2020, pp. 132-137, doi: 10.1109/ICIMU49871.2020.9243426.

[2]. M. L. Hale, R. F. Gamble and P. Gamble, "CyberPhishing: A Game-Based Platform for Phishing Awareness Testing,"2015 48th Hawaii International Conference on System Sciences, Kauai, HI, USA, 2015, pp. 5260-5269, doi: 10.1109/HICSS.2015.670.

[3]. P. Chinnasamy, P. Krishnamoorthy, K. Alankruthi, T. Mohanraj, B. S. Kumar and L. Chandran, "AI Enhanced Phishing Detection System,"2024 Third International Conference on Intelligent Techniques in Control, Optimization and Signal Processing (INCOS), Krishnankoil, Virudhunagar district, Tamil Nadu, India, 2024,pp.1-5, doi:10.1109/INCOS59338.2024.10527485.

[4]. T. Daengsi, P. Wuttidittachotti, P. Pornpongtechavanich and N. Utakrit, "A Comparative Study of Cybersecurity Awareness on Phishing Among Employees from Different Departments in an Organization,"2021 2nd International Conference on Smart Computing and Electronic Enterprise (ICSCEE), Cameron Highlands, Malaysia, 2021, pp. 102-106, doi: 10.1109/ICS CEE50312.2021. 9498208.

[5]. N. Davis and E. S. Grant, "Simulated Phishing Training Exercises versus Phishing Education Games,"2022 Fourth International Conference on Emerging Research in Electronics, Computer Science and Technology (ICERECT), Mandya, India, 2022, pp. 1-8, doi: 10.1109/ICERECT56837.2022.10060595