International Journal of Contemporary Research in Computer Science and Technology (IJCRCST) ISSN: 2395-5325
Volume 9, Special Issue 1 (April '2026)
Special Issue on International Conference on Innovation, Engineering & Management (ICIEM-2026)
Conference held at Garden City University, Bangalore, India.

MACHINE LEARNING-BASED PHISHING DETECTION SYSTEM FOR ENHANCING CYBERSECURITY

Greeshma BL
Department of Computer Science
Garden City University, Bengaluru, India

Prof. Pallabee Padhi
Department of Computer Science
Garden City University, Bengaluru, India

Abstract: *Machine learning is rapidly transforming the cybersecurity landscape by enabling intelligent systems to detect and prevent online threats. One of the most significant threats today is phishing, where attackers create fraudulent websites to steal sensitive information such as passwords and financial data. Traditional rule-based detection systems often fail to identify new and sophisticated phishing attacks. This paper aims to provide a comprehensive overview of a Machine Learning-Based Phishing Detection System for enhancing cybersecurity. It highlights recent advancements in machine learning techniques for phishing detection from 2022–2025, including improved accuracy, real-time detection, and adaptive learning models. Furthermore, the challenges such as dataset imbalance, feature selection, evolving attack patterns, and model generalization are discussed, along with possible solutions to overcome these limitations.*

Keywords: *Machine Learning, Phishing Detection, Cybersecurity, Feature Extraction, URL Analysis, Classification Algorithms, Random Forest, Logistic Regression, Deep Learning, Real-time Detection*

I. INTRODUCTION

In the early development of cybersecurity systems, traditional rule-based and signature-based approaches were widely used to detect malicious activities. However, with the rapid growth of the internet and digital services, cyber threats such as phishing attacks have become increasingly sophisticated and difficult to detect using conventional methods [1]. Over the past decade, Machine Learning (ML) has emerged as a powerful tool for enhancing cybersecurity by enabling systems to learn patterns from data and adapt to new threats dynamically [2].

Unlike traditional detection systems that rely on predefined rules, machine learning-based approaches analyze large volumes of data to identify hidden patterns and anomalies. In phishing detection, features such as URL structure, domain information, website content, and user behavior are used to classify websites as legitimate or malicious. Algorithms such as Logistic Regression, Decision Trees, Random Forest, and Neural Networks have demonstrated significant improvements in detection accuracy and efficiency [3]. These models can generalize from training data and identify previously unseen phishing attacks, making them highly effective in real-world scenarios [4].

Between 2022 and 2025, significant advancements have been observed in the application of machine learning for phishing detection. Researchers have developed hybrid models combining traditional machine learning and deep learning techniques, improved feature extraction methods, and implemented real-time detection systems integrated with web browsers and security tools [5]. Additionally, large-scale datasets and cloud-based computing have enhanced model performance and scalability [6]. Despite these advancements, challenges such as dataset imbalance, evolving phishing strategies, feature selection complexity, and model interpretability continue to affect the reliability and robustness of these systems [7].

II. LITERATURE REVIEW

Despite a growing body of publications, five key gaps motivate the present review [8].

A. Limited Unified Model Evaluation

Most existing studies focus on individual machine learning algorithms such as Logistic Regression, Decision Trees, or Random Forest, without providing a comprehensive comparison across multiple models using consistent datasets and evaluation metrics. This lack of standardized evaluation makes it difficult to identify the most efficient and reliable model for phishing detection in real-world scenarios [9].

B. Incomplete Coverage of Real-Time Detection Systems

Many research works report high accuracy in controlled or offline environments; however, limited attention has been given to real-time phishing detection. Integration of machine learning models into live systems such as web browsers, email filtering systems, and network security tools remains insufficiently explored. This creates a gap between theoretical performance and practical usability [10].

C. Disconnect Between Model Performance and Real-World Implementation

Although machine learning models show promising results in experimental settings, their performance often degrades in real-world environments due to noisy data, evolving phishing techniques, and system constraints. Factors such as latency, computational cost, and scalability are rarely considered, resulting in a gap between algorithmic capability and practical deployment [11].

D. Feature Selection Challenges and Evolving Attack Patterns

Phishing attacks continuously evolve, making static feature-based detection less effective over time. Many existing systems rely on predefined features such as URL length, domain age, and keyword presence, which may not capture new and sophisticated phishing strategies. Adaptive and dynamic feature selection methods are still underexplored [12].

E. Lack of Updated and Comprehensive Studies

Recent years (2022–2025) have seen rapid advancements in machine learning techniques, including deep learning and hybrid models for phishing detection. However, many existing surveys and studies do not fully incorporate these recent developments, creating a gap in up-to-date and comprehensive analysis of current techniques and their effectiveness [13].

III. METHODOLOGY

This paper adopts a systematic literature review methodology, with scope restricted to publications from January 2022 through March 2026, focusing on advancements in machine learning techniques for phishing detection and cybersecurity [17].

A. Data Collection and Preprocessing

A structured dataset consisting of phishing and legitimate URLs was collected from reliable sources such as publicly available repositories and security databases. The data was cleaned by removing duplicate and irrelevant entries, handling missing values, and converting categorical data into numerical form. Data normalization was also performed to ensure consistency and improve model performance. In addition, a structured search was conducted across major academic databases including IEEE Xplore, ACM Digital Library, ScienceDirect, SpringerLink, and Google Scholar to identify relevant datasets and research contributions. Primary search terms included machine learning, phishing detection, cybersecurity, URL classification, feature extraction, deep learning, and real-time detection systems. Boolean operators were used to refine results (e.g., "phishing detection AND machine learning AND 2023") [18]. Inclusion required peer-reviewed publications or reputable conference papers within the 2022–2026 window, addressing phishing detection techniques, machine learning models, feature engineering, or real-world cybersecurity applications. Studies lacking experimental validation or not directly related to phishing detection were excluded [19].

B. Feature Extraction and Selection

Relevant features were extracted from the dataset to effectively identify phishing patterns. These features include URL-based attributes such as length, presence of special characters, and structure; domain-based attributes such as domain age, DNS records, and HTTPS usage; and content-based attributes such as HTML elements, scripts, and embedded links. Feature selection techniques were applied to eliminate redundant and less significant features, thereby improving model efficiency and accuracy. In addition, existing literature was reviewed to identify commonly used and effective feature sets for phishing detection, ensuring that the selected features align with recent advancements and real-world attack patterns [20].

C. Model Development and Training

Multiple machine learning algorithms were implemented, including Logistic Regression, Decision Tree, and Random Forest, to classify websites as phishing or legitimate. The dataset was divided into training (80%) and testing (20%) sets to ensure proper evaluation. The models were trained using the training dataset, and hyperparameter tuning was performed to optimize their performance. Furthermore, insights from recent studies were incorporated to select appropriate algorithms and training strategies, ensuring that the models are robust and capable of handling diverse phishing patterns [21].

D. Model Evaluation

The performance of the trained models was evaluated using standard evaluation metrics such as Accuracy, Precision, Recall, and F1-Score. These metrics provide a comprehensive assessment of the model's ability to correctly classify phishing and legitimate websites while minimizing false positives and false negatives. Comparative analysis was also conducted to identify the best-performing model. In addition, evaluation approaches discussed in recent literature were considered to ensure that the assessment methods are consistent with current research standards [22].

E. System Implementation and Limitations

The best-performing model was integrated into a detection system where users can input a URL, and the system analyzes its features to predict whether it is phishing or legitimate. The system is designed to support real-time detection and can be

extended to applications such as browser extensions and email filtering systems. However, the proposed system has certain limitations, including dependency on dataset quality, evolving phishing techniques, and potential performance degradation in real-world environments. These limitations highlight the need for continuous model updates and integration of adaptive learning techniques, as also emphasized in recent research studies [23]

IV. RESULTS AND DISCUSSION

A. Model Performance Analysis

The performance of different machine learning models, including Logistic Regression, Decision Tree, and Random Forest, was evaluated using metrics such as Accuracy, Precision, Recall, and F1-Score. Among these models, Random Forest achieved the highest accuracy due to its ability to handle complex feature interactions and reduce overfitting. Logistic Regression provided stable and interpretable results, while the Decision Tree model showed moderate performance and was more prone to overfitting. These findings highlight the effectiveness of ensemble methods in phishing detection [24**].**

B. Comparative Evaluation

A comparative analysis of the implemented models revealed differences in performance based on their complexity and feature-handling capabilities. Random Forest outperformed the other models in terms of accuracy and robustness, while Logistic Regression demonstrated better computational efficiency. The Decision Tree model, although simple and easy to interpret, showed limitations when applied to larger datasets. This comparison emphasizes the importance of selecting suitable algorithms for practical applications [25].

C. Impact of Feature Selection

Feature extraction and selection significantly influenced the performance of the machine learning models used for phishing detection. URL-based and domain-based features contributed most to accurate classification, while content-based features further enhanced detection capability. The removal of irrelevant and redundant features helped in reducing noise and improving model efficiency. This highlights the importance of selecting meaningful attributes for better prediction accuracy. Effective feature engineering plays a critical role in improving the reliability and performance of phishing detection systems [26***].***

D. System Effectiveness in Real-Time Detection

The developed system was capable of analyzing user-input URLs and generating predictions in real time. The integration of the trained machine learning model into a simple interface demonstrated its practical usability and ease of access for users. The system efficiently classified websites as phishing or legitimate with minimal delay, ensuring a smooth user experience. It provides quick and reliable responses, making it suitable for real-world cybersecurity applications such as browser extensions and email filtering systems. The implementation demonstrates that machine learning models can effectively support real-time threat detection and prevention. However, system performance may vary depending on computational resources, dataset size, and network conditions. High traffic and limited system resources may impact response time and prediction accuracy. This highlights the need for optimization techniques to improve speed, scalability, and reliability in deployment [27].

E. Limitations and Discussion

Despite achieving strong performance, the system has certain limitations. The accuracy of the model depends heavily on the quality and diversity of the dataset used for training. Evolving phishing techniques may reduce the effectiveness of the system over time. Additionally, the model may produce false positives and false negatives in some cases. The system's performance can also be affected by feature selection and data imbalance issues. Real-time implementation may introduce challenges related to computational cost and response time. These limitations highlight the need for continuous improvement, regular dataset updates, and adaptive learning approaches. Further enhancements are required to improve robustness and long-term reliability of the system [28].

V. CONCLUSION

This study demonstrates that machine learning techniques provide an effective and practical approach for detecting phishing websites and enhancing cybersecurity. By implementing classification algorithms such as Logistic Regression, Decision Tree, and Random Forest, the system was able to accurately distinguish between phishing and legitimate websites. The use of feature extraction, particularly URL-based, domain-based, and content-based features, significantly improved the model's ability to identify malicious patterns. The evaluation results using metrics such as accuracy, precision, recall, and F1-score confirmed the reliability and efficiency of the proposed system. Furthermore, the integration of the trained model into a real-time detection framework highlights its applicability in real-world environments such as web browsers and email filtering systems. The study also emphasizes the importance of proper data preprocessing and feature selection in achieving high model performance. However, challenges

such as evolving phishing techniques, dataset imbalance, and dependency on static features may affect long-term effectiveness. The presence of false positives and false negatives indicates the need for continuous model improvement and validation. Additionally, real-time deployment introduces constraints related to computational resources and system scalability. To address these issues, future enhancements should focus on adaptive learning, use of updated and diverse datasets, and integration of advanced techniques such as deep learning and hybrid models. Overall, the proposed system offers a reliable, scalable, and efficient solution for phishing detection and contributes significantly to strengthening modern cybersecurity systems.

VI. REFERENCES

[1]. A. K. Jain and B. B. Gupta, "Phishing detection: Analysis of visual similarity based approaches," *Security and Communication Networks*, vol. 2017, pp. 1–20.

[2]. S. Marchal, J. François, R. State, and T. Engel, "Phishing detection: A literature survey," *IEEE Communications Surveys & Tutorials*, vol. 18, no. 4, pp. 2360–2381, 2016.

[3]. M. Sahingoz, E. Buber, O. Demir, and B. Diri, "Machine learning based phishing detection from URLs," *Expert Systems with Applications*, vol. 117, pp. 345–357, 2019.

[4]. R. Verma and A. Das, "What's in a URL: Fast feature extraction and malicious URL detection," *Proceedings of ACM CCS Workshop*, 2017.

[5]. J. Ma, L. K. Saul, S. Savage, and G. M. Voelker, "Learning to detect malicious URLs," *ACM Transactions on Intelligent Systems and Technology*, vol. 2, no. 3, pp. 1–24, 2011.

[6]. N. Abdelhamid, A. Ayesh, and F. Thabtah, "Phishing detection using association classification," *Expert Systems with Applications*, vol. 41, no. 13, pp. 5948–5959, 2014.

[7]. T. Fawcett, "An introduction to ROC analysis," *Pattern Recognition Letters*, vol. 27, no. 8, pp. 861–874, 2006.

[8]. I. Goodfellow, Y. Bengio, and A. Courville, *Deep Learning*, MIT Press, 2016.

[9]. S. Garera, N. Provos, M. Chew, and A. D. Rubin, "A framework for detection and measurement of phishing attacks," *ACM Workshop on Recurring Malcode*, 2007.

[10]. A. Le, A. Markopoulou, and M. Faloutsos, "PhishDef: URL names say it all," *IEEE INFOCOM Workshops*, 2011.

[11]. M. Aburrous, M. A. Hossain, K. Dahal, and F. Thabtah, "Intelligent phishing detection system for e-banking using fuzzy data mining," *Expert Systems with Applications*, vol. 40, no. 12, pp. 5084–5092, 2013.

[12]. Y. Xiang, W. Zhou, M. Guo, and N. Xiong, "A hybrid approach for detection of phishing websites," *Journal of Network and Computer Applications*, vol. 36, no. 1, pp. 381–387, 2013.

[13]. UCI Machine Learning Repository, "Phishing Websites Dataset," [Online]. Available: https://archive.ics.uci.edu

[14]. PhishTank, "Phishing URL Dataset," [Online]. Available: https://www.phishtank.com

[15]. Kaggle, "Phishing Detection Dataset," [Online]. Available: https://www.kaggle.com

[16]. Google, "Google Safe Browsing," [Online]. Available: https://safebrowsing.google.com

[17]. Microsoft, "Microsoft Security Intelligence Report," 2023.

[18]. IBM Security, "Cost of a Data Breach Report," 2023.

[19]. B. Kitchenham and S. Charters, "Guidelines for systematic literature reviews in software engineering," EBSE Technical Report, 2007.

[20]. A. Aldweesh, A. Derhab, and A. Z. Emam, "Deep learning approaches for anomaly-based intrusion detection systems," *Journal of Ambient Intelligence and Humanized Computing*, 2020.

[21]. M. Alazab et al., "Phishing website detection using machine learning and deep learning," *Future Generation Computer Systems*, vol. 108, pp. 1097–1109, 2020.

[22]. S. Sahoo, B. B. Gupta, and K. R. Choo, "Real-time phishing detection using machine learning," *Computers & Security*, vol. 120, 2022.

[23]. A. Basnet, A. H. Sung, and Q. Liu, "Learning to detect phishing URLs," *International Journal of Research in Engineering and Technology*, 2015.

[24]. S. Adebowale et al., "Intelligent web phishing detection and classification using supervised learning," *Computers & Security*, 2022.