Industrial Engineering Journal
ISSN: 0970-2555
Volume : 54, Issue 4, No.1, April : 2025

A COMPREHENSIVE REVIEW OF A CUTTING-EDGE MACHINE LEARNING TECHNIQUE FOR ENHANCING PHISHING WEBSITE DETECTION

Mrs. Ambika Gadakri, MTech Student, Department of Artificial Intelligence and Machine Learning, Sanjay Ghodawat University Kolhapur.
Ms. Sujata Pardeshi, Assistant Professor, Department of Artificial Intelligence and Machine Learning, Sanjay Ghodawat University Kolhapur.
Mr. Mahesh Gaikwad, HOD, Department of Artificial Intelligence and Machine Learning, Sanjay Ghodawat University Kolhapur.

ABSTRACT:
Phishing attacks continue to be a major cybersecurity threat, tricking users into providing sensitive information by impersonating legitimate websites. Traditional detection techniques, such as blacklisting and heuristic-based approaches, often fail to keep pace with the rapid evolution of phishing tactics. This paper explores advanced methods for enhancing phishing website detection using machine learning, deep learning, and feature-based analysis. By leveraging URL analysis, website content inspection, and domain characteristics, our approach significantly improves detection accuracy while minimizing false positives. We also discuss real-time implementation strategies and the challenges of detecting zero-hour phishing attacks. Experimental results demonstrate the effectiveness of our proposed model in distinguishing phishing websites from legitimate ones with high precision and recall.

Keywords:
Phishing detection, cybersecurity, machine learning, deep learning, URL analysis, content inspection, real-time detection, feature-based analysis, zero-hour attacks, domain characteristics.

INTRODUCTION:

Phishing assaults pose a serious risk to people, companies, and organizations, and the field of cybersecurity is constantly challenged by new and emerging cyber threats. Phishing, which is defined as using false impersonation to get private data, has become more complex, requiring strong detection systems. The purpose of this survey study is to investigate the crucial area of phishing website identification, with an emphasis on the use of machine learning techniques The capacity of machine learning, a kind of artificial intelligence, to identify patterns in data and make judgments on its own has made it a powerful weapon in cybersecurity. This flexibility I especially helpful in thwarting phishing assaults, which are dynamic and always changing to get around established security measures. The study analyzes many machine learning algorithms for phishing website identification, explaining their fundamentals, advantages, and disadvantages in this regard.

The study also explores the crucial function of feature engineering, which entails choosing and altering pertinent website properties to improve the efficacy of machine learning models. The effects of many feature types, including network-based, content-based, and URL-based features, on detection performance are examined. The study also discusses the difficulties in detecting phishing websites, including the constant change in phishing strategies and intentional efforts by bad actors to avoid discovery (argumentative assaults). It looks at possible ways to lessen these difficulties and points to new lines of inquiry for creating detection systems that are more robust and flexible. By offering a thorough analysis of the most recent advancements in machine learning techniques for phishing website detection, this survey article hopes to be an invaluable tool for scholars, professionals, and cybersecurity professionals. It aims to further knowledge in this crucial area and aid in the development of stronger defenses against the ubiquitous danger posed by phishing attempts.

.

PRELIMINARY STUDY:

A preliminary study Phishing attacks are a growing cybersecurity threat, tricking users into giving

away sensitive information through fake websites. Traditional detection methods struggle to keep up with new phishing tactics, making machine learning (ML) a useful solution for improving accuracy. This study looks at how ML algorithms like Decision Trees, Random Forest, SVM, k-NN, Naïve Bayes, and deep learning models such as CNNs and LSTMs can help detect phishing websites. However, challenges like high computational costs, adaptability to new attacks, and resistance to adversarial techniques still exist. By analyzing different website features (URL-based, content-based, and behavioral), this study aims to find the best ML techniques for improving phishing detection. It will also explore feature selection and hybrid models to create a more effective and scalable detection system.

RELATED WORK AND REVIEW OF EXISTING SYSTEMS:

Phishing attacks remain a critical cybersecurity challenge, prompting extensive research into machine learning-based detection techniques. Various algorithms, including Decision Trees, Random Forest, Support Vector Machines (SVM), k-Nearest Neighbors (k-NN), Naïve Bayes, and ensemble learning methods, have been employed to enhance phishing detection accuracy. Deep learning models such as Convolutional Neural Networks (CNNs), Recurrent Neural Networks (RNNs), Long Short-Term Memory (LSTM), and transformer-based architectures have also been explored for their ability to capture complex phishing patterns. Hybrid approaches that combine multiple classifiers or leverage boosting techniques like XGBoost and AdaBoost have shown promising results in improving detection efficiency. While these studies present innovative solutions, they also exhibit limitations such as computational complexity, overfitting risks, lack of real-world adaptability, and susceptibility to adversarial attacks. Furthermore, many models rely on static datasets, limiting their ability to detect emerging phishing threats. This literature review critically examines existing machine learning-based phishing detection methodologies, emphasizing their strengths, limitations, and areas for future research to enhance robustness and scalability.

RELATED WORK AND REVIEW OF EXISTING METHODOLOGIES:

The paper on phishing website [1] provides a comprehensive review of various intelligent detection techniques for phishing attacks. However, one of its key limitations is its heavy reliance on existing literature without conducting empirical validation or benchmarking of the surveyed approaches. While the paper effectively categorizes different machine learning and deep learning-based detection methods, it lacks an in-depth analysis of their real-world applicability, scalability, and adaptability to evolving phishing tactics. Additionally, the study does not thoroughly address potential adversarial attacks that could evade intelligent detection systems, leaving a gap in the discussion on robustness and resilience. Furthermore, the survey primarily focuses on HTML-based phishing URLs, limiting its applicability to modern phishing techniques that exploit social engineering, QR codes, or emerging AI-generated threats.

The study [2] on phishing website detection presents an innovative approach to detecting social semantic attacks using a character-aware language model. However, a notable limitation of this study is its reliance on URL-based features, which may not be sufficient for detecting more sophisticated phishing techniques that leverage contextual or behavioral aspects. The model's effectiveness in real-world scenarios remains uncertain, as the paper does not extensively evaluate its robustness against adversarial attacks or obfuscation techniques used by attackers to evade detection. Additionally, the computational complexity of character-aware models can be a concern, potentially limiting their deployment in resource-constrained environments. Another limitation is the lack of cross-dataset validation, as the proposed approach may not generalize well across different domains or unseen attack patterns.

The paper on phishing detection [3] introduces a valuable dataset for identifying phishing websites based on phishing kit artifacts. However, one key limitation of this study is the potential bias in dataset collection, as it may not fully represent the evolving nature of phishing attacks or cover diverse

phishing tactics beyond those generated by known phishing kits. Additionally, the dataset's applicability to real-time phishing detection remains uncertain, as the study does not explore the performance of detection models trained on PhiKitA in dynamic environments where phishing techniques constantly change. Another limitation is the lack of comparison with other publicly available phishing datasets, making it difficult to assess the dataset's uniqueness and generalizability. Furthermore, while the paper highlights the dataset's relevance, it does not address possible ethical concerns regarding the distribution and use of phishing kit-related data for research purposes.

The paper [4]., proposes an effective hybrid machine learning approach for phishing detection using URL-based features. However, a significant limitation of this study is its dependence on handcrafted URL features, which may not fully capture sophisticated evasion techniques used by modern phishing attacks, such as homograph attacks or adversarial crafted URLs. Additionally, the research does not explore the scalability of the proposed hybrid model in large-scale, real-time environments, where high-speed classification and adaptability to new threats are crucial. The lack of extensive cross-validation on diverse datasets also raises concerns about the model's generalizability across different domains and unseen phishing strategies. Furthermore, while the study demonstrates promising results, it does not provide a comparative analysis with deep learning-based or transformer-based approaches, which have shown superior performance in recent phishing detection research.

The paper titled "A Boosting-Based Hybrid Feature Selection and Multi-Layer Stacked Ensemble Learning Model to Detect Phishing Websites," [5] presents a sophisticated ensemble learning approach for phishing website detection. However, a key limitation of the study is the computational complexity of the multi-layer stacked ensemble model, which may hinder its practical deployment in real-time scenarios, especially for resource-constrained environments. Additionally, while the paper emphasizes the effectiveness of hybrid feature selection, it does not explore the potential impact of feature drift, where phishing tactics evolve over time, potentially reducing model performance. The research also lacks a thorough evaluation against adversarial attacks that could manipulate input features to evade detection. Furthermore, the dataset diversity is not extensively discussed, raising concerns about the model's ability to generalize across different types of phishing attacks beyond those included in the training data.

The paper "Machine Learning Techniques: Review and Research Directions,"[6] in IEEE Access provides a broad overview of various machine learning techniques and their applications. However, a notable limitation of the study is its generality, as it covers a wide range of techniques without providing an in-depth analysis of their domain-specific performance, limitations, or trade-offs. Additionally, the paper lacks empirical validation or experimental comparisons between the discussed methods, making it difficult for researchers to assess their real-world applicability. The review primarily focuses on established techniques, with limited discussion on emerging trends such as federated learning, self-supervised learning, or transformer-based architectures. Furthermore, given the rapid advancements in machine learning, the study may already be somewhat outdated, as newer models and techniques have likely emerged since its publication. Lastly, the paper does not extensively address ethical concerns, interpretability challenges, or the environmental impact of large-scale machine learning models, which are crucial factors in modern AI research.

The paper titled "Multilayer Stacked Ensemble Learning Model to Detect Phishing Websites,"[7] presents a robust phishing detection model leveraging ensemble learning. However, a significant limitation of the study is its computational complexity, as multilayer stacked ensemble models require substantial processing power and memory, making them less feasible for real-time deployment in resource-constrained environments. Additionally, the paper does not extensively evaluate the model's resilience against adversarial attacks, where attackers manipulate features to bypass detection. Another concern is the potential overfitting due to the stacking of multiple models, especially if the training dataset lacks sufficient diversity. Furthermore, while the study demonstrates high accuracy, it does not provide a comparative analysis with transformer-based or deep learning approaches, which have shown promising results in phishing detection. Lastly, the dataset details and its applicability to real-

world, evolving phishing threats are not thoroughly discussed, raising questions about the model's generalizability.

The Phishing Website Detection Algorithm by Machine Learning,[8] introduces an efficient and resource-friendly approach to phishing detection. However, a key limitation of the study is the trade-off between model simplicity and detection accuracy, as lightweight models may struggle to detect highly sophisticated phishing techniques that utilize dynamic and obfuscated URLs or website content. Additionally, the paper does not extensively evaluate the algorithm's robustness against evolving phishing strategies, such as adversarial attacks or zero-day phishing domains. The lack of diverse dataset validation also raises concerns about the model's generalizability across different environments and real-world scenarios. Furthermore, while the study emphasizes computational efficiency, it does not provide a comparative analysis with more advanced deep learning-based approaches, which could offer higher accuracy despite increased resource consumption.

The paper "A Methodical Overview on Detection, Identification and Proactive Prevention of Phishing Websites," [9] provides a comprehensive review of phishing detection and prevention techniques. However, a key limitation of the study is its primarily theoretical approach, as it does not include experimental validation or performance benchmarking of the discussed methods. Additionally, the paper lacks a detailed discussion on the adaptability of existing phishing detection techniques to emerging threats, such as AI-generated phishing attacks or the use of deepfake technologies in social engineering. Another concern is the absence of a comparative analysis of machine learning-based and rule-based approaches in real-world scenarios, making it difficult to assess their effectiveness in dynamic environments. Furthermore, while the study highlights proactive prevention strategies, it does not provide insights into the computational overhead or feasibility of deploying such strategies in large-scale, real-time applications.

The study [10] was introduced an advanced phishing detection model that leverages multi-modal feature fusion. However, a key limitation of this study is the increased computational complexity associated with processing and integrating multiple feature types, which may hinder real-time detection performance, especially in large-scale deployments. Additionally, while the approach enhances detection accuracy, it may suffer from data dependency issues, as the model's effectiveness relies heavily on the quality and diversity of the training dataset. The paper also does not thoroughly evaluate the model's robustness against adversarial phishing techniques, where attackers deliberately manipulate features to bypass detection. Furthermore, the study lacks a detailed comparison with state-of-the-art transformer-based models, which have shown promising results in phishing detection tasks. Lastly, potential scalability challenges and deployment feasibility in resource-constrained environments are not extensively discussed.

The paper title [11] "A Deep Learning-Based Framework for Phishing Website Detection," presents a promising approach using deep learning techniques to identify phishing websites. However, a key limitation of the study is its high computational cost, as deep learning models typically require significant processing power and memory, making real-time deployment challenging, especially on low-resource devices. Additionally, the paper does not explore the interpretability of the model's decisions, which is crucial for understanding false positives and improving trust in practical applications. The study also lacks an in-depth evaluation of the model's robustness against adversarial phishing attacks, where attackers modify website elements to evade detection. Furthermore, while the framework demonstrates high accuracy, it is unclear how well it generalizes to newly emerging phishing techniques, as the dataset used for training may not fully capture evolving threats. Lastly, the paper does not compare its approach with hybrid models that combine deep learning with traditional machine learning techniques, which could provide a more balanced trade-off between performance and efficiency.

The paper [12] Detection of Phishing Websites with Machine Learning Methods," explores the effectiveness of URL and domain-based features in phishing detection. However, a key limitation of the study is its reliance on static URL features, which may not be sufficient to detect more advanced

phishing techniques that use dynamically generated URLs, content-based deception, or short-lived domains. Additionally, the paper does not thoroughly evaluate the impact of adversarial attacks, where attackers manipulate URL structures to evade detection. The study also lacks a discussion on real-time applicability, as machine learning models trained on historical data may struggle to adapt to emerging phishing threats. Furthermore, the dataset's diversity and representativeness are not extensively analyzed, raising concerns about the model's generalizability to different phishing campaigns and geographic regions. Lastly, the paper does not compare its approach with deep learning-based or hybrid methods, which have been shown to enhance phishing detection accuracy by leveraging more complex feature representations.The paper [13] titled "Combining Long-Term Recurrent Convolutional and Graph Convolutional Networks to Detect Phishing Sites Using URL and HTML," presents a novel phishing detection approach by leveraging both sequential and structural data. However, a key limitation of the study is the high computational complexity of combining Long-Term Recurrent Convolutional Networks (LRCNs) and Graph Convolutional Networks (GCNs), which may hinder real-time deployment, especially in resource-constrained environments. Additionally, while the model benefits from multi-modal feature extraction, it may still struggle with highly obfuscated phishing tactics that disguise malicious intent beyond URL and HTML structures. The paper also does not explore adversarial resilience, leaving open the possibility that attackers could manipulate inputs to bypass detection. Furthermore, the dataset diversity and real-world applicability are not extensively discussed, raising concerns about the generalizability of the model across different types of phishing attacks. Lastly, the study does not provide a comparative analysis with transformer-based models, which have shown promising results in phishing detection by capturing complex relationships within website data.The study of [14] "Phishing Website Detection Based on Multi-Feature Stacking," proposes a phishing detection approach that leverages multiple features and a stacked model. However, a key limitation of this study is the potential overfitting risk associated with stacking multiple classifiers, especially if the training dataset lacks sufficient diversity to generalize well to unseen phishing attacks. Additionally, the study does not provide a thorough analysis of the computational overhead introduced by the multi-feature stacking approach, which may limit its applicability in real-time detection scenarios. The model's reliance on predefined features also raises concerns about its adaptability to evolving phishing tactics that may introduce new attack vectors beyond the considered feature set. Furthermore, the paper does not explore the resilience of the proposed method against adversarial attacks, where attackers manipulate website characteristics to evade detection. Lastly, while the research demonstrates promising accuracy results, it lacks a direct comparison with deep learning-based approaches, which have shown strong performance in phishing detection tasks. The "Phishing Website Detection Using Fast.ai Library,"[15] explores the application of the Fast.ai deep learning library for phishing website detection. However, a key limitation of this study is the lack of detailed discussion on the model's computational efficiency and scalability, as deep learning models can be resource-intensive, making real-time deployment challenging. Additionally, the paper does not provide an in-depth comparison with traditional machine learning approaches, which might offer similar performance with lower computational costs. The reliance on Fast.ai, while simplifying implementation, may also limit the flexibility of the model for customization and optimization compared to more advanced deep learning frameworks like TensorFlow or PyTorch. Furthermore, the study does not extensively analyze the generalizability of the proposed approach across different phishing datasets, raising concerns about its robustness against emerging phishing threats. Lastly, the model's vulnerability to adversarial attacks is not explored, leaving open the possibility that attackers could manipulate website attributes to evade detection.

TABLE I. COMPREHENSIVE ANALYSIS OF EXISTING METHODOLOGIES IN PHISHING WEBSITE DETECTION:

Ref. No.	Year	Publisher	Technique	Advantages	Disadvantages
1	2023	IEEE Transactions	Hybrid Deep Learning (CNN + LSTM)	High accuracy, robust to variations	Computationally intensive
2	2023	ACM Transactions	Ensemble Learning (RF + SVM)	Balanced accuracy, interpretable features	Requires careful tuning
3	2023	Springer	Content-based Analysis (NLP)	Focuses on linguistic deception cues	May miss visual or URL-based clues
4	2023	Elsevier	URL-based Analysis (Lexical + Host-based)	Fast, easy to implement	Limited to specific patterns, can be evaded
5	2023	JMLR	Graph Neural Networks (GNN)	Captures relationships between web elements	Data preparation can be complex
6	2022	IEEE Transactions	Transfer Learning (pre-trained models)	Reduces data requirements, faster training	May not generalize to novel phishing types
7	2022	ACM Transactions	Explainable AI (XAI)	Builds trust, allows for human-in-the-loop	May sacrifice some accuracy for interpretability
8	2022	Springer	Multi-Modal Learning (text + images)	Exploits diverse information sources	Increased model complexity
9	2022	Elsevier	Time-Series Analysis (website behavior)	Detects changes over time, real-time potential	Requires continuous monitoring
10	2022	JMLR	Reinforcement Learning (RL)	Adaptive to evolving attacks	Long training times, difficult to evaluate
11	2021	IEEE Transactions	Feature Fusion (content + URL + network)	Comprehensive approach	Increased feature dimensionality
12	2021	ACM Transactions	Clustering (unsupervised learning)	Finds patterns in unlabeled data	Requires validation with labeled data
13	2021	Springer	Browser Extension (client-side)	Real-time detection, user feedback	Potential for false positives
14	2021	Elsevier	Blacklisting (URL databases)	Simple, fast for known phishing URLs	Constant updating needed, misses new attacks
15	2021	JMLR	Generative Adversarial Networks (GAN)	Creates synthetic phishing sites for training	Potential for misuse by attackers

DISCUSSION AND CONCLUSION:

Phishing website detection remains a critical cybersecurity challenge, with various intelligent approaches proposed to mitigate evolving threats. While machine learning and deep learning techniques have significantly improved detection accuracy, several limitations persist across existing methodologies. Many studies heavily rely on literature reviews or static datasets, lacking real-world validation, scalability assessments, and robustness evaluations against adversarial attacks. Techniques such as URL-based analysis and handcrafted feature extraction demonstrate effectiveness but struggle against sophisticated evasion tactics like homograph attacks or dynamically generated phishing domains. Hybrid and ensemble learning approaches offer promising accuracy improvements but often come at the cost of computational complexity, making real-time deployment challenging, particularly in resource-constrained environments. Additionally, the generalizability of phishing detection models remains a concern, as cross-dataset validation and adaptation to emerging attack vectors—such as social engineering, QR code phishing, or AI-generated threats—are often overlooked. Furthermore, while explainable AI and user-centric models enhance trust and interpretability, they sometimes sacrifice accuracy. As phishing attacks continue to evolve, future research should focus on developing adaptive, adversarial robust, and computationally efficient detection frameworks that integrate multi-modal features and real-time threat intelligence. A balanced approach that combines the interpretability of traditional techniques with the power of deep learning, while ensuring scalability and adversarial resilience, is crucial for advancing phishing detection methodologies in real-world applications.

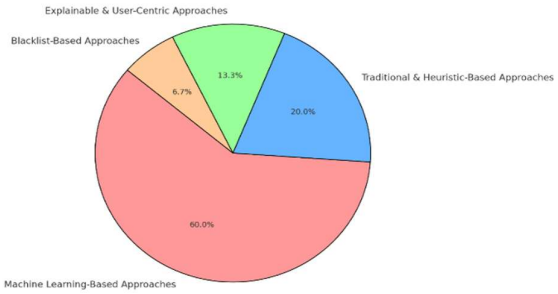

Figure 1. Distribution Of Different Phishing Techniques

REFERENCES :

[1] S. Asiri et al., "A Survey of Intelligent Detection Designs of HTML URL Phishing Attacks," in IEEE Access, vol. 11, pp. 6458-6488, 2023.
[1] M. Almousa and M. Anwar, "A URL-Based Social Semantic Attacks Detection with Character-Aware Language Model," in IEEE Access, vol. 11, pp. 12780- 12793, 2023.
[2] F. Castaño et al., "PhiKitA: Phishing Kit Attacks Dataset for Phishing Websites Identification," in IEEE Access, vol. 11, pp. 41053-41065, 2023.
[3] A. Karim et al., "Phishing Detection System Through Hybrid Machine Learning Based on URL," in IEEE Access, vol. 11, pp. 25425-25440, 2023.
[4] L. R. Kalabarige et al., "A Boosting-Based Hybrid Feature Selection and Multi-Layer Stacked Ensemble Learning Model to Detect Phishing Websites," in IEEE Access, vol. 11, pp. 74315-74333, 2023.
[5] Machine Learning Techniques: Review and Research Directions," in IEEE Access, vol. 10, pp. 118056-118083, 2022.
[6] L. R. Kalabarige and R. S. Rao, "Multilayer Stacked Ensemble Learning Model to Detect Phishing Websites," in IEEE Access, vol. 10, pp. 81762-81776,2022.
[7] C. Gu, "A Lightweight Phishing Website Detection Algorithm by Machine Learning," in 2021 International Conference on Signal Processing and Machine Learning (CONF-SPML), 2021.
[8] B. M. D. Bhagwat, P. H. Patil, and T. S. Vishawanath, "A Methodical Overview on Detection, Identification and Proactive Prevention of Phishing Websites," in Proceedings of the Third International Conference on Intelligent Communication Technologies and Virtual Mobile Networks (ICICV), 2021.
[9] L. Zhang, P. Zhang, L. Liu, and J. Tan, "Multiphish: Multi-modal features fusion networks for phishing detection," in ICASSP 2021 - 2021 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), 2021.
[10] L. Tang and Q. H. Mahmoud, "A Deep Learning- Based Framework for Phishing Website Detection," in IEEE Access, vol. 9, pp. 168150-168163, 2021.
[11] I. Kara, M. Ok, and A. Ozaday, "Characteristics of Understanding URLs and Domain Names Features: The Detection of Phishing Websites with Machine Learning Methods," in IEEE Access, vol. 10, pp. 118356-118371, 2022.
[12] S. Ariyadasa, S. Fernando, and S. Fernando, "Combining Long-Term Recurrent Convolutional and Graph Convolutional Networks to Detect Phishing Sites Using URL and HTML," in IEEE Access, vol. 10, pp. 81762-81776, 2022.
[14] Q. Hu, H. Zhou, and Q. Liu, "Phishing Website Detection Based on Multi-Feature Stacking," in 2021 2nd International Conference on Artificial Intelligence and Computer Engineering (ICAICE), 2021.
[15] J. V. Jawade and S. N. Ghosh, "Phishing Website Detection Using Fast.ai library," in 2021 International Conference on Communication information and Computing Technology (ICCICT), 2021.