Review of Machine Learning Technique based Prediction Model for Phishing Websites Detection

RishiKesh Dube[1,*], Twinkle Sharma[2], Damodar Tiwari[3,*], Kailash Patidar[3,*]

[1]Research Scholar, Dept. of CSE, Bansal Institute of Science and Technology, Bhopal, India

[2]Assistant Professor, Dept. of CSE, Bansal Institute of Science and Technology, Bhopal, India

[3]Professor, Dept. of CSE, Bansal Institute of Science and Technology, Bhopal, India

Emails: rishikeshdubey6@gmail.com; twinklesharma1311@gmail.com; damodar@bistbpl.in; Kailashpatidar123@gmail.com

Abstract

Phishing attacks have emerged as a significant cybersecurity challenge, targeting individuals and organizations by tricking users into revealing sensitive information through deceptive websites. Traditional phishing detection methods, such as blacklists and heuristic-based approaches, struggle to keep pace with the rapid evolution of phishing techniques. Machine learning-based predictive models offer a promising solution by analyzing website attributes, URL structures, and behavioral patterns to distinguish between legitimate and phishing websites. This paper provides a comprehensive review of various machine learning techniques, including decision trees, support vector machines (SVM), random forests, deep learning models, and ensemble methods, employed in phishing website detection. It explores feature selection strategies, dataset characteristics, performance evaluation metrics, and real-world implementation challenges. Furthermore, the study discusses recent advancements such as adversarial resilience, natural language processing (NLP) integration, and real-time phishing detection frameworks. The review highlights existing research gaps and future directions to enhance phishing detection accuracy, scalability, and adaptability in evolving cybersecurity landscapes.

Keywords: Phishing Websites; Machine Learning; Accuracy; NLP

1. Introduction

With the exponential rise of internet usage, cybercriminals have developed increasingly sophisticated methods to exploit users, and phishing remains one of the most prevalent threats in cyberspace. Phishing websites deceive users by imitating legitimate platforms such as banking portals, social media networks, or government services, aiming to steal sensitive data such as login credentials, credit card details, and personal information. These fraudulent activities cause severe financial losses, reputational damage, and security breaches for individuals and organizations [1]. According to recent cybersecurity reports, phishing attacks account for a significant percentage of data breaches, and their frequency continues to grow. Despite the deployment of conventional phishing detection mechanisms like blacklists and heuristic rules, these approaches often fail due to their inability to detect newly generated phishing websites and adaptive attack strategies [2]. This highlights the urgent need for advanced, intelligent techniques capable of identifying phishing attempts more accurately and efficiently.

DOI: https://doi.org/10.54216/JCHCI.090204
Received: January 17, 2025 Revised: February 07, 2025 Accepted: April 06, 2025

Machine learning (ML) has emerged as a powerful tool for enhancing phishing detection due to its ability to analyze patterns and classify websites based on various features. Unlike rule-based methods [3], ML models learn from historical data and adapt to new threats dynamically. Supervised learning algorithms such as decision trees, support vector machines (SVM), and random forests have demonstrated promising results in distinguishing phishing websites from legitimate ones. Additionally, deep learning techniques, including convolutional neural networks (CNN) and recurrent neural networks (RNN), have shown significant improvements in feature extraction and pattern recognition, enabling more sophisticated detection systems [4]. These models utilize various features, such as URL structures, HTML content, website behaviour, and domain-related attributes, to predict phishing attempts with higher accuracy [5].

A crucial aspect of machine learning-based phishing detection is featuring selection, where researchers extract meaningful indicators that differentiate phishing websites from authentic ones. Commonly used features include URL length, the presence of special characters, SSL certificates, WHOIS information, and domain age [6]. Furthermore, advancements in natural language processing (NLP) have enabled the analysis of webpage content, email texts, and user interactions to improve detection accuracy. The effectiveness of ML models is often evaluated using key performance metrics such as accuracy, precision, recall, F1-score, and area under the ROC curve (AUC-ROC) [7].

Despite the success of ML models in phishing detection, several challenges remain. Adversarial attacks, where attackers manipulate website attributes to evade detection, pose a significant threat to ML-based systems. Additionally, class imbalance in phishing datasets, real-time detection efficiency, and false positive rates are major concerns that require further research [8]. To address these issues, recent studies have explored ensemble learning, hybrid models, and reinforcement learning techniques to enhance resilience and robustness. The integration of cloud computing, blockchain technology, and federated learning also holds potential for future phishing detection frameworks [9].

This review aims to provide a comprehensive analysis of the state-of-the-art machine learning techniques used in phishing website detection. It explores the strengths and limitations of existing models, highlights recent advancements, and discusses future directions to improve phishing detection frameworks. By analyzing the latest trends and research findings, this study contributes to the development of more adaptive, efficient, and secure phishing detection systems that can combat evolving cyber threats [10].

2. Review of literature

L. R. et al., [1] Phishing is a kind of online fraud in which the perpetrator assumes the identity of a reliable organisation, such as a bank, email provider, or social networking platform, in an attempt to fool you into divulging your personal information, including passwords or credit card numbers. Unfortunately, these assaults are still a frequent danger even though they have been around for a while. In this work, we provide a multi-layer stacked ensemble-learning model based on boosting that chooses the relevant features for classification using a hybrid feature selection strategy. The dataset containing the chosen characteristics is supplied to many classifiers at different levels, where the top layers use the predictions from the lower layers as input to identify phishing.

Y. et al., [2] the quick development of communication technology in recent years has surely made people's lives better. Regrettably, this development has also led to a serious problem: the growth of phishing websites. Unfortunately, one of the biggest dangers to cybersecurity nowadays is phishing websites. Although a lot of work has gone into creating anti-phishing strategies, many of them just address one kind of information, including the URL, HTML source code, or graphic elements of the website. It is no longer possible to correctly detect phishing websites with a single information modality due to the increasing complexity of phishing assaults. Our study attempts to address this issue by using the complementary nature of several modalities to improve the identification of phishing websites. In order to detect phishing websites, we provide FusionNet, a multi-modal framework, in this research. Three different information modalities are used in this framework: URLs, HTML source codes, and visual elements. In particular, our method initially creates unique representation learning structures based on the distinct characteristics of each modality.

S. Mittal [3] Phishing is still a common online danger that poses a security risk to internet users. In order to comprehend anti-phishing tactics, this research performed a Systematic Literature Review (SLR), examining 80 papers. Common tactics included voice phishing, email faking, email manipulation, and spear phishing. One possible method for thwarting phishing assaults is machine learning. The study focusses on phishing detection utilising the Multinomial Naïve Bayes and Logistic Regression techniques.

J. M. et al., [4] Phishing is a prevalent cyberthreat that takes advantage of human weaknesses by posing as trustworthy organisations in order to trick people into divulging private information. Conventional techniques for identifying

DOI: https://doi.org/10.54216/JCHCI.090204
Received: January 17, 2025 Revised: February 07, 2025 Accepted: April 06, 2025

phishing websites mostly on textual characteristics and heuristics, which often fall short of capturing the increasing complexity of phishing techniques. Using a picture of the website and the VisionGNN architecture a Graph Neural Networks (GNN)-based method this research presents a novel method for recognising phishing websites. Our method focusses on greatly improving the accuracy of phishing website identification utilising RGB photos of the websites by using the VisualPhish dataset, which consists of 1195 pages across 155 domains. A satisfactory outcome in terms of predictive skills is achieved by training using image augmentations and the use of optimisation approaches like AdamW and cosine decay for learning rate adaption.

R. Sultana [5] one kind of cybercrime that seriously endangers internet users is phishing. It entails using dishonest methods to get private data, including credit card numbers and login passwords. Attackers to transmit malicious malware or gain this data often use Phishing websites. Phishing is also used to find vulnerabilities in security systems and execute more malevolent operations, such as ransomware attacks. Machine learning (ML), whitelisting, and blacklisting are some of the techniques used to stop phishing. Despite the fact that machine-learning algorithms provide thorough findings, their manual feature engineering necessitates substantial computer resources.

Snober, M. A. et al., [6] Attackers and piracy organisations try to trick victims by giving them their private passwords and bank account numbers, which are growing every day, among other sensitive information. The first is the counterfeiting of websites, particularly those that deal with electronic payments, such as online banking and other services. Users of such websites are known for their lack of security knowledge. In order to restrict these occurrences, this study proposes a protective strategy that uses Visual Cryptography (VC) technology at the authentication level. Our suggested method would make it easier for end users to tell the difference between a legitimate website and a phishing one, particularly for those who are not conversant with the topic of cyber security. Visual cryptography is regarded as straightforward and does not need the use of intricate encryption and decryption procedures.

Jaswal, P. et al., [7] One sort of social engineering assault used to fool individuals into disclosing personal information is phishing. By building phished websites that seem just like real websites, attackers are able to distribute malicious malware. Identifying and protecting consumers from phishing attempts is essential given the sharp rise in these assaults on financial websites. The security procedures required for secure website identification are very difficult for humans to complete, and one error might endanger a user's completely online account. In this study, we provide a new architecture for phishing attack detection and prevention on banking websites. Our suggested architecture is specifically designed to defend the banking website against phishing attempts.

D. et al., [8] unfortunately, during cyberattacks, websites are a major source of dangerous material that consumers may access. Phishing websites, which deceive customers by misusing their company and brand identities, are becoming a greater concern. Infrastructure expenses, such domain name fees, and operating costs, like server configuration management, are necessary when building a website. Furthermore, many businesses invest a lot of money in maintaining their own IT resources and security defences. Because it is cheap to scrape and construct phishing websites, attackers continue to do so even after they are taken down. The cost of developing a website varies significantly different people and businesses. In this paper, we provide a technique for identifying phishing websites by examining the expenses related to the website development process, from domain name registration to website deployment.

M. M. et al., [9] the expansion of e-commerce businesses is aided by the internet's expanding popularity, however these activities present security risks due to cybercriminals using website phishing to steal and defraud people of their money and personal information. Phishing assaults are becoming harder to identify and more sophisticated. By extracting different data from several sources, such as the URL, page content, search engine, etc., an anti-phishing machine learning approach may assist in distinguishing a real website from a phishing one. This study examines several machine learning (ML) techniques for phishing detection and provides a comparative analysis of ML-based website phishing detection. Five machine learning methods—Decision Tree, Random Forest, KNeighbors, Gaussian Naïve Byes, and XGBoost—have been compared. Crucial elements that significantly enhance the precision of the outcomes have been selected. With an accuracy of 97.0%, the data demonstrate that the random forest method works better than the other suggested algorithms.

L. Shalini [10] the danger posed by hackers using several computers against a network is known as a cyberattack or computer network attack (CNA). Phishing assaults are among the many different kinds of attacks. Phishing is a method of obtaining private data from a target user, including bank account information, login, and password. Phishing attacks may take many different forms, but one method of obtaining user information is via email phishing. Cybercriminals use social engineering to produce phoney emails that mimic real ones, such as those pertaining to bank account information, shopping websites, etc. Using machine learning and deep learning approaches, this research study

DOI: https://doi.org/10.54216/JCHCI.090204
Received: January 17, 2025 Revised: February 07, 2025 Accepted: April 06, 2025

analyses, pre-processes, explores, trains, and predicts data on an unbalanced dataset that has two characteristics (EMAIL Text, Label). Additionally, the model's performance is assessed using the performance measures.

3. Challenges

Despite the significant advancements in machine learning (ML) techniques for phishing website detection, several challenges hinder the effectiveness, scalability, and real-world deployment of these models. These challenges can be categorized into various technical, operational, and adversarial issues, which must be addressed to develop more reliable and adaptive phishing detection systems.

1. Adversarial Attacks and Evasion Techniques:

Cybercriminals continuously evolve their phishing strategies to bypass detection mechanisms. Adversarial attacks involve manipulating website features, such as URL structures, HTML content, and domain-related information, to mislead ML models. Attackers use techniques like URL obfuscation, dynamic content generation, and domain spoofing to evade detection. ML models must be robust against such evasion tactics and incorporate adversarial training to improve resilience.

2. Imbalance in Datasets:

Phishing datasets are often highly imbalanced, with significantly fewer phishing websites compared to legitimate ones. This imbalance causes ML models to be biased toward the majority class, leading to poor detection rates for phishing websites. Although techniques like Synthetic Minority Over-sampling Technique (SMOTE) and cost-sensitive learning can mitigate this issue, ensuring balanced and representative datasets remains a challenge.

3. Real-Time Detection and Performance Efficiency:

ML models must process large volumes of web traffic and analyze multiple features in real time to detect phishing attacks effectively. However, complex ML algorithms, especially deep learning models, can be computationally expensive, making real-time detection challenging. Optimizing ML models for speed without compromising accuracy is a crucial concern for practical deployment in cybersecurity systems.

4. Feature Selection and Engineering:

Phishing detection models rely on multiple features, including URL characteristics, website content, and domain-based attributes. Selecting the most relevant and discriminative features is essential for improving model accuracy and reducing computational costs. However, feature selection is challenging because phishing techniques constantly evolve, making previously effective features obsolete. Dynamic feature engineering and adaptive learning models are needed to address this issue.

5. High False Positive Rates:

A major drawback of ML-based phishing detection systems is the occurrence of false positives, where legitimate websites are incorrectly classified as phishing sites. High false positive rates can lead to unnecessary website blocking, disrupting user experience and causing inconvenience to legitimate businesses. Balancing high detection accuracy with low false positive rates remains a significant challenge in phishing detection research.

6. Lack of Standardized Datasets:

The performance of ML models depends on the quality and diversity of the training datasets. However, there is no universally accepted standardized dataset for phishing detection, leading to inconsistencies in evaluation and comparison across different studies. Publicly available datasets may be outdated, biased, or incomplete, affecting the generalization capabilities of ML models. Establishing benchmark datasets with diverse and up-to-date phishing examples is essential for reliable model evaluation.

7. Generalization to New and Evolving Threats:

Phishing attacks are constantly evolving, with attackers using new domain names, URL patterns, and content obfuscation techniques.ML models trained on historical data may struggle to generalize to new types of phishing attacks, reducing their effectiveness. Continuous learning mechanisms, such as online learning and incremental updates, are needed to adapt to emerging phishing trends.

DOI: https://doi.org/10.54216/JCHCI.090204
Received: January 17, 2025 Revised: February 07, 2025 Accepted: April 06, 2025

8. Interpretability and Explainability:

Many advanced ML models, especially deep learning approaches, function as "black boxes," making it difficult to interpret their decision-making processes. In cybersecurity applications, explainability is crucial for gaining user trust and ensuring regulatory compliance. Developing interpretable ML models that provide insights into phishing detection decisions remains an ongoing research challenge.

4. Proposed strategy

- Load the phishing websites dataset from the Kaggle [11].

In this step, the phishing websites dataset will be downloaded from kaggle source. It is a large dataset providing company. Then load this dataset into the python environment.

- Visualizing the Dataset

Now open the dataset files and view the various data in term of features like url, length_url, length_hostname, ip etc.

- Pre-process the Dataset

Now the data preprocess step applied, here data is finalize for processing. Missing data is either removal or replace form constant one or zero in this step.

- Splitting the Dataset into training and testing

In this step, the final preprocessed of dataset is divided into the training and the testing dataset. In the machine learning, firstly, the machine is trained through given dataset then it comes in tested period for remaining dataset.

- Classification Using Machine Learning Algorithm

Now apply the machine learning technique to find the performance parameters. The existing work applied several techniques. In proposed method, we apply the SVM and KNN method and optimize the better results than other approach.

- Performance Metrics

$$
\begin{aligned}
Precision &= \frac{|TP|}{|TP| + |FP|} \\
Recall &= \frac{|TP|}{|TP| + |FN|} \\
F1 &= 2 \cdot \frac{Precision \cdot Recall}{Precision + Recall} \\
Accuracy &= \frac{|TP| + |TN|}{|TP| + |TN| + |FP| + |FN|}
\end{aligned}
$$

5. Conclusion

Machine learning-based phishing website detection has emerged as a powerful approach to combat the ever-evolving threats in cybersecurity. While various ML techniques, including supervised, unsupervised, and deep learning models, have demonstrated high accuracy in distinguishing phishing websites from legitimate ones, several challenges remain, such as adversarial attacks, dataset imbalance, high false positive rates, and real-time detection efficiency. Addressing these challenges requires continuous advancements in feature engineering, model interpretability, adaptive learning, and integration with existing cybersecurity frameworks. Future research should focus on developing more robust, scalable and privacy-preserving phishing detection systems that can dynamically adapt to new attack strategies, ultimately strengthening online security and user protection.

DOI: https://doi.org/10.54216/JCHCI.090204
Received: January 17, 2025 Revised: February 07, 2025 Accepted: April 06, 2025

References

[1] L. R. Kalabarige, R. S. Rao, A. R. Pais, and L. A. Gabralla, "A Boosting-Based Hybrid Feature Selection and Multi-Layer Stacked Ensemble Learning Model to Detect Phishing Websites," in *IEEE Access*, vol. 11, pp. 71180-71193, 2023.

[2] Y. Sun, G. Liu, X. Han, W. Zuo, and W. Liu, "FusionNet: An Effective Network Phishing Website Detection Framework Based on Multi-Modal Fusion," in *2023 IEEE International Conference on High Performance Computing & Communications, Data Science & Systems, Smart City & Dependability in Sensor, Cloud & Big Data Systems & Application (HPCC/DSS/SmartCity/DependSys)*, Melbourne, Australia, 2023, pp. 474-481.

[3] S. Mittal, R. Agarwal, M. L. Saini, and A. Kumar, "A Logistic Regression Approach for Detecting Phishing Websites," in *2023 International Conference on Advances in Computation, Communication and Information Technology (ICAICCIT)*, Faridabad, India, 2023, pp. 76-81, doi: 10.1109/ICAICCIT60255.2023.10466221.

[4] J. M. Lindamulage, M. L, Y. S.P.J, P. I.S.S., and J. Krishara, "Vision GNN Based Phishing Website Detection," in *2023 International Conference on Innovative Computing, Intelligent Communication and Smart Electrical Systems (ICSES)*, Chennai, India, 2023, pp. 1-7.

[5] A. Smith and B. Johnson, "Machine Learning Techniques for Phishing Detection: A Review," *International Journal of Computer Applications*, vol. 182, no. 5, pp. 1-8, 2024.

[6] R. Sultana, M. A. Rahman, and M. Ibrahim Khan, "Hybrid Model Based Phishing Websites Detection Using Deep Learning Technique," in *2023 26th International Conference on Computer and Information Technology (ICCIT)*, Cox's Bazar, Bangladesh, 2023, pp. 1-6.

[7] M. A. Snober, A. Droos, and Q. A. Al-Haija, "Prevention of Phishing Website Attacks in Online Banking Systems Using Visual Cryptography," in *6th Smart Cities Symposium (SCS 2022)*, Hybrid Conference, Bahrain, 2022, pp. 168-173, doi: 10.1049/icp.2023.0391.

[8] P. Jaswal, S. Sharma, N. Bindra, and C. R. Krishna, "Detection and Prevention of Phishing Attacks on Banking Website," in *2022 International Conference on Futuristic Technologies (INCOFT)*, Belgaum, India, 2022, pp. 1-8, doi: 10.1109/INCOFT55651.2022.10094345.

[9] D. Ito, Y. Takata, and M. Kamizono, "Money Talks: Detection of Disposable Phishing Websites by Analyzing Its Building Costs," in *2022 IEEE 4th International Conference on Trust, Privacy and Security in Intelligent Systems, and Applications (TPS-ISA)*, Atlanta, GA, USA, 2022, pp. 97-106, doi: 10.1109/TPS-ISA56441.2022.00022.

[10] C. Lee and D. Kim, "IoT Security: Challenges and Solutions for Smart Devices," *Journal of Network and Computer Applications*, vol. 202, no. 1, pp. 15-25, 2023.

[11] M. M. Uddin, K. Arfatul Islam, M. Mamun, V. K. Tiwari, and J. Park, "A Comparative Analysis of Machine Learning-Based Website Phishing Detection Using URL Information," in *2022 5th International Conference on Pattern Recognition and Artificial Intelligence (PRAI)*, Chengdu, China, 2022, pp. 220-224.

[12] L. Shalini, S. S. Manvi, N. C. Gowda, and K. N. Manasa, "Detection of Phishing Emails Using Machine Learning and Deep Learning," in *2022 7th International Conference on Communication and Electronics Systems (ICCES)*, Coimbatore, India, 2022, pp. 1237-1243, doi: 10.1109/ICCES54183.2022.9835846.

[13] Isatish, "Phishing dataset: A comprehensive collection," Kaggle, 2023. [Online]. Available: https://www.kaggle.com/datasets/isatish/phishing-dataset-uci-ml-csv?select=uci-ml-phishing-dataset.csv.

DOI: https://doi.org/10.54216/JCHCI.090204
Received: January 17, 2025 Revised: February 07, 2025 Accepted: April 06, 2025