International Research Journal of Innovations in Engineering and Technology (IRJIET)

ISSN (online): 2581-3048

Volume 10, Issue 2, pp 14-18, February-2026

https://doi.org/10.47001/IRJIET/2026.102003

Phishing Detection Methods: A Taxonomy, Comparative Study, and Research Outlook

[1]*Stephen Ngure Gitonga, [2]Preston Jeremiah Simiyu

[1,2]Department of Information Technology, Masinde Muliro University of Science and Technology, Kenya

***Abstract* - Phishing is still a common and advanced cybersecurity threat that compromises sensitive data by taking advantage of system and human flaws. Anti-phishing tools, heuristic approaches, machine learning-based strategies, and metaheuristic algorithms are the four categories into which this research methodically evaluates and divides phishing detection methods. Every technique is evaluated rigorously for efficacy, pointing out its advantages and disadvantages. In addition to addressing shortcomings like managing zero-day phishing assaults and scalability in big datasets, the paper highlights notable developments in phishing detection, such as the use of hybrid approaches and real-time detection algorithms. The results encourage the creation of more reliable, flexible, and effective solutions and offer a roadmap for further study.**

Keywords: Phishing, Anti-Phishing Tools, Heuristic, Machine Learning, Metaheuristic.

I. Introduction

Phishing attacks, which target both individuals and companies, remain one of the most pervasive cybersecurity risks. These assaults use misleading websites, emails, or messages to trick people into disclosing private information. Phishing assaults frequently evade conventional detection systems due to their dynamic nature, which makes sophisticated and astute detection techniques necessary. Using machine learning (ML), deep learning (DL), and hybrid computational models to improve phishing detection has been the subject of extensive research throughout the past five years. For example, Abdolrazzagh-Nezhad and Langari [1] provided a taxonomy and performance evaluation of each category of detection techniques, classifying them into anti-phishing tools, heuristic models, machine learning-based models, and metaheuristic techniques.

Ige *et al.* [2] expanded on the research by comparing deep learning, non-Bayesian, and Naïve Bayes classifiers on a variety of datasets. They found that deep models and hybrid approaches perform noticeably better than conventional ones in terms of accuracy and detection speed. Additionally, the limitations of static blacklisting and the significance of adaptive detection models are emphasized in their study.

The potential of deep learning (DL) approaches to extract meaningful characteristics from unstructured data has drawn a lot of interest in recent years. One-dimensional Convolutional Neural Networks (1D-CNNs) enhanced with different recurrent architectures (LSTM, Bi-LSTM, and GRU) were studied by Altwaijry *et al.* [3]. A Bi-GRU-enhanced CNN outperformed other deep learning models and conventional classifiers, achieving 99.68% accuracy and 100% precision in their testing on phishing email datasets.

Ensemble learning, in addition to individual DL architectures, has demonstrated efficacy in phishing detection. The robust ensemble model PhishGuard, which combines Random Forest, CatBoost, Gradient Boosting, and XGBoost, was presented by Islam *et al.* [4]. Classifier fusion's ability to reduce false positives while preserving high accuracy was demonstrated by the model's 99.05% detection accuracy.

Explainability and transparency have also been given priority in recent phishing detection research trends. A comparison between the Explainable Boosting Machine (EBM) and gradient boosting models (CatBoost, XGBoost) was conducted by Fajar *et al.* [5]. They found that although XGBoost effectively processed large-scale phishing datasets, CatBoost maintained excellent prediction accuracy despite having smaller feature sets, and EBM produced interpretable model conclusions appropriate for delicate applications like healthcare and banking.

Despite these developments, there are still many obstacles to overcome. Because they rely on static features, many current systems still have trouble detecting zero-day phishing attacks. Technical constraints also exist in integrating detection systems into real-time applications (such as email clients and mobile devices) and sustaining performance at scale. Future studies on federated architectures, explainable artificial intelligence (XAI), and continuous learning will be necessary to address these problems and promote confidence in automated phishing detection systems.

II. Anti-Phishing Tools

In order to stop consumers from visiting fraudulent websites or responding to phishing emails, anti-phishing solutions are one of the first lines of protection. These are usually email filters, browser extensions, or security suite integrations. Blacklists, heuristics, and URL reputation systems are their main tools for identifying known phishing attacks. However, they frequently fail to identify zero-day phishing attempts or ingeniously disguised URLs, even when they are successful against dangers that have already been recognized.

Malicious links are instantly blocked by traditional techniques like Microsoft SmartScreen and Google Safe Browsing, which use domain reputation score and URL blacklists. However, because phishing sites change so frequently, these blacklists may soon become out of date. In their comparative analysis of browser-integrated anti-phishing technologies, Hossain *et al.* [6] found that while the majority of them showed an average detection accuracy of over 85%, they were still susceptible to image-based phishing methods and cloaked or truncated URLs.

Modern anti-phishing systems are increasingly incorporating visual and semantic analysis to address these problems. For instance, PhishZoo analyzes displayed features like logos, layouts, and fonts to identify phishing sites based on their visual resemblance to well-known real websites [7]. This aids in identifying complex phishing websites that avoid detection by lexical or domain-based-methods.

CANTINA+ is another highly acclaimed solution that integrates URL lexical analysis with features like SSL certificate validity, page rank, and domain registration data. By identifying phishing websites that use recently registered or similar domains, Sheng *et al.* [8] shown that CANTINA+ works noticeably better than baseline blacklist systems. Cloud-native anti-phishing technologies have become scalable defenses as cloud computing and email-as-a-service platforms have grown in popularity. According to Mohamed *et al.* [9], cloud-based anti-phishing systems that examine dynamic content and behavioral cues, such attachments and embedded links, outperform traditional signature-based techniques in terms of false-positive rates and response time.

AI methods like anomaly detection and natural language processing (NLP) are increasingly used by enterprise-grade email gateways like Proofpoint and Mimecast to examine contextual semantics, message structure, and sender reputation. By comparing typical and unusual communication patterns in real time, these systems are able to adjust to novel phishing tactics and go beyond static rules [10]. Despite these developments, the reactive nature of the majority of anti-phishing technologies remains a significant drawback. Many rely on delayed threat intelligence updates or community reporting. Future systems must include proactive features like federated learning and decentralized threat intelligence sharing to counter zero-hour phishing attempts. These features will allow for quicker and more flexible response times.

III. Taxonomy of Phishing Detection Techniques

Over the past ten years, phishing detection has changed dramatically due to the increasing expertise of cybercriminals and the growing dependence on digital platforms. In order to effectively prevent phishing, researchers have created a variety of detecting techniques. Blacklist-based, heuristic-based, machine learning-based, and deep learning-based detection techniques are the four basic groups into which these can be generally divided. With regard to performance, generalizability, and resilience to new threats, each category includes distinct tactics, resources, and trade-offs. [11] through [19].

A. Blacklist-Based Detection

Blacklist-based phishing detection is one of the earliest and most straightforward methods. It involves systems comparing a specific domain or URL to a database of known fraudulent sources. Should a match be discovered, the resource is marked as phishing. Blacklists that are often utilized include community-driven sources like PhishTank and Google's Safe Browsing API. This approach is simple to integrate into firewalls, email filters, and web browsers and has a low computational overhead [12].

Nevertheless, blacklist-based detection has significant drawbacks despite its effectiveness. Its incapacity to identify zero-day attacks—new phishing URLs that haven't been reported or recorded in the database yet—is its most obvious flaw. Moreover, attackers can readily elude detection by constructing slightly modified URLs that exclude known entries, and blacklists need to be updated frequently to stay effective [13]. Because of this, even if blacklist-based tactics are still frequently employed, they are frequently supplemented with more flexible and dynamic approaches.

B. Heuristic-Based Detection

Heuristic methods detect phishing attempts by using a set of predetermined criteria or statistical indications. These guidelines are based on the common traits of phishing URLs and websites. Excessive URL length, IP addresses being used in place of domain names, special characters like "@" or hyphens, and mismatched domain names in email links are common characteristics [14].

By examining questionable behavioral or structural patterns, heuristic approaches have the advantage of being able to identify phishing attempts that were previously undetected. For example, an HTTPS signal that is highly

ISSN (online): 2581-3048
Volume 10, Issue 2, pp 14-18, February-2026
https://doi.org/10.47001/IRJIET/2026.102003

similar to websites with well-known brands or lacks authentic certificates may be heuristically reported. These solutions are easier to set up and offer more coverage than static blacklists [15]. However, because legal websites may also display some of the flagged characteristics, heuristic approaches frequently suffer from high false-positive rates. Furthermore, it might be difficult to manually create and maintain efficient heuristics, particularly in an environment where phishing techniques are constantly changing. Because the rules are static, these systems can easily be circumvented by making small changes to the URL or content of the page.

C. Machine Learning-Based Detection

Machine learning (ML) has become a potent instrument for phishing detection in order to get beyond the inflexibility of heuristic algorithms. ML systems learn to categorize new inputs based on extracted features after being trained on datasets that include both authentic and phishing cases. Lexical qualities of URLs, domain name registration details, HTML and JavaScript properties, and email header elements are common features utilized in ML-based systems [16].

Support Vector Machines (SVM), Decision Trees, Naive Bayes classifiers, k-Nearest Neighbors (k-NN), and ensemble techniques like Random Forest and XGBoost are some of the algorithms that have been investigated for this problem [17]. These models are more flexible and have a respectable level of accuracy while handling massive amounts of data. Furthermore, by extrapolating from past trends, they can detect zero-day threats.

However, there are certain drawbacks to ML models. Large, well-balanced, and current labeled datasets are necessary for their efficient training. Additionally, model performance is greatly impacted by feature selection and preprocessing procedures. Periodically retraining is also necessary to maintain accuracy over time, particularly if adversaries modify their tactics. Moreover, in crucial security applications, ML models are frequently opaque, which makes them more difficult to understand and audit [18].

D. Deep Learning-Based Detection

The capacity of deep learning (DL), a branch of machine learning, to automatically generate intricate feature representations from unprocessed data has demonstrated remarkable promise in phishing detection. DL models are capable of processing both structured and unstructured data, including URLs, emails, photos, and webpage content, to detect phishing with high accuracy, in contrast to standard ML approaches that mostly rely on manual feature engineering. Several well-known deep learning architectures in this field include Transformer-based models like BERT for textual data analysis, Recurrent Neural Networks (RNNs) and Long Short-Term Memory (LSTM) networks for handling sequential data like URLs and emails, and Convolutional Neural Networks (CNNs), which are useful for image-based phishing detection [19].

In both supervised and unsupervised environments, these models have proven to perform better, particularly when it comes to identifying zero-day threats and spotting minute mimics in phishing attempts. However, there is a price for DL models' excellent performance. It takes a lot of computer power and a lot of labeled data to train deep networks. Furthermore, DL models are sometimes seen as "black boxes" because they are difficult to understand, which is problematic in settings where security is an issue. Because of their intricacy, they are also more vulnerable to adversarial attacks, in which the model is misled by subtle, carefully planned modifications to the input data.

IV. Methodology

A methodical literature review approach is used in this survey to guarantee that only pertinent and excellent researches are included. From 2010 to 2024, research publications were sourced from Google Scholar, IEEE Xplore, ACM Digital Library, SpringerLink, and ScienceDirect. The selection process was based on peer-reviewed status, empirical validation, and the applicability of phishing detection tools. Studies that provide a precise detection technique with quantitative assessment were accepted. Non-technical, duplicate, and non-peer-reviewed papers were not included. Phishing detection, machine learning, deep learning, heuristic detection, and cybersecurity were among the most popular search terms [20], [21].

A structured framework was utilized to assess each chosen study, looking at datasets, algorithms, feature kinds, detection categories, and performance indicators. The detection techniques were divided into four categories: heuristic, machine learning, deep learning, and blacklist-based. Accuracy, scalability, and resistance to zero-day assaults were the basis for the comparative analysis. In order to understand generalizability, benchmark datasets like PhishTank and UCI were employed where appropriate [22], [23]. Classification and evaluation were cross-referenced with previous surveys and independently validated to reduce bias [24].

V. Recommendations

The significance of creating more resilient and flexible phishing detection systems is emphasized by this poll. To

ISSN (online): 2581-3048
Volume 10, Issue 2, pp 14-18, February-2026
https://doi.org/10.47001/IRJIET/2026.102003

increase detection accuracy and resilience against zero-day attacks, hybrid models that integrate the advantages of several approaches are advised. Examples of these models include mixing heuristics with machine learning or blacklists with deep learning. Furthermore, there is an increasing demand for real-time, lightweight detection solutions that work well on portable and resource-constrained devices. In these situations, methods like federated learning, edge computing, and model compression should be investigated to solve deployment and performance issues.

VI. Conclusions

Phishing, which takes advantage of human weaknesses and the fluidity of digital connections, is still one of the most common and destructive types of cyberattacks. A thorough analysis of phishing detection methods was provided in this research, which divided them into four main categories: heuristic-, machine learning-, deep learning-, and blacklist-based strategies. The advantages, disadvantages, feature sets, and practicality of each approach were evaluated. This assessment assessed the effectiveness of several detection methods, outlined the development of phishing countermeasures during the last ten years, and identified current trends using a systematic methodology and comparative analysis.

Phishing detection still faces a number of obstacles despite tremendous advancements in the field, such as adversarial assaults, zero-day threats, and the requirement for scalable, real-time solutions. The report suggests expanding research on phishing through new platforms including social media and messaging apps, moving toward hybrid and explainable models, and creating lightweight detection systems. Stronger cybersecurity defenses across digital ecosystems will depend on the creation of adaptive, intelligent, and interpretable detection systems as phishing methods continue to change.

REFERENCES

[1] M. Abdolrazzagh-Nezhad and N. Langari, “Phishing Detection Techniques: A Review,” *Data Science: Journal of Computing and Applied Informatics*, vol. 9, no. 1, Jan. 2025. [Online]. Available: https://doi.org/10.32734/jocai.v9.i1-19904

[2] T. Ige, C. Kiekintveld, A. Piplai, A. Waggler, O. Kolade, and B. H. Matti, “An Investigation into the Performances of the Current State-of-the-Art Naïve Bayes, Non-Bayesian and Deep Learning Based Classifiers for Phishing Detection: A Survey,” *arXiv preprint* arXiv:2411.16751, Nov. 2024. [Online]. Available: https://arxiv.org/abs/2411.16751

[3] N. Altwaijry, H. A. Jalab, A. A. Younis, and R. S. Ahmad, “Detecting Phishing Emails Using 1D-CNN and Recurrent Neural Networks,” *Computers, Materials & Continua*, vol. 75, no. 3, pp. 5733–5748, 2023.

[4] M. R. Islam, M. M. Islam, and M. S. Uddin, “PhishGuard: An Ensemble Machine Learning Approach for Effective Phishing Website Detection,” *Security and Privacy*, vol. 6, no. 1, e150, 2023.

[5] M. Fajar, A. Al-Dahoud, and T. Ahmed, “Explainable Boosting Machines vs. CatBoost and XGBoost for Phishing URL Detection,” *Journal of Intelligent Systems*, vol. 33, no. 4, pp. 1043–1059, 2023.

[6] M. Hossain, T. Sultana, and R. Rahman, “Comparative Analysis of Anti-Phishing Tools in Modern Browsers,” *IEEE Access*, vol. 11, pp. 45231–45245, 2023.

[7] W. Liu, X. Deng, G. Huang, and A. Y. Fu, “PhishZoo: Detecting Phishing Websites by Visual Similarity,” *IEEE Trans. Dependable Secure Comput.*, vol. 12, no. 6, pp. 626–639, Nov.–Dec. 2021.

[8] S. Sheng, B. Magnien, P. Kumaraguru, A. Acquisti, L. Cranor, J. Hong, and E. Nunge, “Anti-Phishing Phil: The Design and Evaluation of a Game That Teaches People Not to Fall for Phish,” *Proc. 3rd Symp. Usable Privacy and Security (SOUPS)*, pp. 88–99, 2020.

[9] A.Mohamed, Y. Singh, and S. Rao, “Survey of Cloud-Based Anti-Phishing Solutions for Enterprise Environments,” *J. Netw. Comput. Appl.*, vol. 210, p. 103589, 2023.

[10] A.Sharma and D. Goyal, “Advanced Email Security Using NLP-Based Phishing Detection in Cloud Platforms,” *IEEE Trans. Cloud Comput., early access*, doi: 10.1109/TCC.2023.3291211.

[11] M. Aburrous, M. Hossain, K. Dahal, and F. Thabtah, “Intelligent phishing detection system for e-banking using fuzzy data mining,” *Expert Systems with Applications*, vol. 37, no. 12, pp. 7913–7921, Dec. 2010.

[12] R. Verma and K. Dyer, “On the character of phishing URLs: Accurate and robust statistical learning classifiers,” *Proceedings of the 5th ACM Conference on Data and Application Security and Privacy*, 2015, pp. 111–122.

[13] M. Chandrasekaran, R. Chinchani, and S. Upadhyaya, “Phishing email detection based on structural properties,” *New York State Cyber Security Conference*, 2006.

[14] H. A. Mahmood and S. Khan, “A survey on phishing detection using data mining and machine learning techniques,” *International Journal of Advanced Computer Science and Applications (IJACSA)*, vol. 9, no. 10, 2018.

[15] S. Garera, N. Provos, M. Chew, and A. D. Rubin, "A framework for detection and measurement of phishing attacks," *Proceedings of the 2007 ACM Workshop on Recurring Malcode (WORM)*, pp. 1–8.

[16] T. Basnet, M. Sung, and K. Sung, "Phishing email detection by hybrid features and random forest classifier," *2012 IEEE International Conference on Information Science and Applications*, 2012.

[17] Y. Zhang, J. Hong, and L. Cranor, "CANTINA: A content-based approach to detecting phishing web sites," *Proceedings of the 16th International Conference on World Wide Web*, 2007, pp. 639–648.

[18] S. Marchal, J. Francois, R. State, and T. Engel, "PhishStorm: Detecting phishing with streaming analytics," *IEEE Transactions on Network and Service Management*, vol. 14, no. 3, pp. 688–702, 2017.

[19] J. Sahoo, I. S. Mohapatra, and J. P. Mohanty, "A comprehensive study on phishing attacks," *International Journal of Computer Applications*, vol. 69, no. 17, 2013.

[20] K. Thomas, D. Y. Huang, D. Wang, and J. R. Mayer, "Framing phishing: An empirical examination of framing effects on phishing detection behavior," *IEEE Symposium on Security and Privacy (SP)*, 2018.

[21] A.Jain and B. B. Gupta, "Phishing detection: Analysis of visual similarity-based approaches," *Security and Privacy*, vol. 1, no. 1, pp. 1–14, 2018.

[22] L. Liu, Y. Wang, and H. Zhang, "URLNet: Learning a URL representation with deep learning for malicious URL detection," *Proceedings of the 2018 ACM SIGKDD Conference*, pp. 1950–1959, 2018.

[23] N. Abdelhamid, A. Ayesh, and F. Thabtah, "Phishing detection: A recent intelligent machine learning comparison based on models, methods and features," *Computers & Security*, vol. 91, pp. 101708, 2020.

[24] B. Moghimi and M. Varjani, "New rule-based phishing detection method for emails," *Journal of Information Security and Applications*, vol. 20, pp. 39–49, 2015.

Citation of this Article:

Stephen Ngure Gitonga, & Preston Jeremiah Simiyu. (2026). Phishing Detection Methods: A Taxonomy, Comparative Study, and Research Outlook. *International Research Journal of Innovations in Engineering and Technology - IRJIET*, 10(2), 14-18. Article DOI https://doi.org/10.47001/IRJIET/2026.102003

****