

Survey paper on Phishing Website Detection using Supervised Machine Learning Technique

¹Ms. Aditi Patel, ²Dr. Nikhil Chaurasia, ³Mr. Nitin Choudhary

M. Tech. Scholar, School of Computer Technology, Sanjeev Agrawal Global Educational
University Bhopal, M.P.¹

Associate Professor, School of Computer Technology, Sanjeev Agrawal Global Educational
University Bhopal, M.P.²

Assistant Professor, School of Computer Technology, Sanjeev Agrawal Global Educational
University Bhopal, M.P.³

Abstract: Phishing websites have become a major cybersecurity threat, targeting users by impersonating legitimate websites to steal sensitive information such as usernames, passwords, banking credentials, and personal data. Traditional blacklist- and rule-based detection methods often struggle to identify newly generated and rapidly changing phishing websites. In recent years, supervised machine learning (ML) techniques have emerged as effective approaches for detecting phishing websites by learning patterns from labeled datasets containing legitimate and malicious website characteristics. This survey paper presents a comprehensive review of supervised ML techniques used for phishing website detection, including Decision Tree, Random Forest, Support Vector Machine, K-Nearest Neighbors, Naïve Bayes, Logistic Regression, and ensemble learning methods. The survey examines commonly used phishing-related features, such as URL structure, domain information, hyperlink characteristics, webpage content, security indicators, and HTML-based attributes. Furthermore, the study compares existing approaches based on evaluation metrics including accuracy, precision, recall, F1-score, and ROC-AUC. The advantages and limitations of different supervised learning methods are discussed, with particular emphasis on feature selection, class imbalance, dataset quality, and generalization to previously unseen phishing websites. The survey also identifies current research challenges and potential future directions, including hybrid ML models, automated feature engineering, explainable artificial intelligence, and real-time phishing detection systems. This study provides a structured overview of supervised machine learning approaches and highlights promising directions for developing accurate, scalable, and robust phishing website detection systems.

Keywords: Phishing Website Detection, Supervised Machine Learning, Cybersecurity

I. INTRODUCTION

The rapid growth of the Internet and web-based services has significantly increased the use of online banking, e-commerce, social networking, cloud services, and digital communication. Along with these developments, cyberattacks have also become more sophisticated, creating serious security and privacy challenges for individuals, organizations, and businesses. Among different types of cyber threats, phishing is one of the most common and effective attacks.

Phishing attacks generally involve fraudulent websites that imitate legitimate websites to deceive users into providing confidential information such as usernames, passwords, credit card details, banking credentials, and other personal information. Attackers frequently modify domain names, URLs, webpage structures, and website content to make malicious websites appear trustworthy and convincing [1, 2].

Traditional phishing detection approaches mainly depend on manually maintained blacklists, whitelists, predefined rules, and signature-based techniques. Although these approaches can successfully identify previously reported phishing websites, their effectiveness is limited against newly created or rapidly changing phishing websites. A newly generated phishing website may remain undetected until its characteristics are identified and added to a detection database. Furthermore, attackers continuously develop new techniques to bypass conventional security mechanisms, creating a need for more adaptive and intelligent detection methods [3]. Machine Learning (ML) provides an effective solution by automatically identifying patterns associated with phishing and legitimate websites from previously labeled data. In supervised machine learning, a classification model is trained using datasets containing known phishing and legitimate website samples. Various features can be extracted from URLs and webpages, including URL length, number of special characters, domain age, presence of an IP address, HTTPS usage, hyperlink characteristics, redirections, HTML elements, and webpage content. These features can then be used by supervised classification algorithms to distinguish between legitimate and malicious websites [4, 5].

Several supervised machine learning algorithms have been investigated for phishing website detection, including Decision Tree (DT), Random Forest (RF), Support Vector Machine (SVM), K-Nearest Neighbors (KNN), Naïve Bayes (NB), Logistic Regression (LR), and ensemble learning techniques. Among these approaches, ensemble methods such as Random Forest can provide robust classification performance by combining the predictions of multiple decision trees. However, the performance of a machine learning-based phishing detection system depends strongly on the quality of the dataset, relevance of extracted features, feature-selection strategy, class distribution, and ability of the model to generalize to previously unseen phishing attacks [6, 7].

Therefore, a systematic survey of supervised machine learning techniques is important for understanding the current state of phishing website detection research. This survey reviews commonly used datasets, website and URL features, supervised classification algorithms, preprocessing and feature-selection techniques, and evaluation metrics used in existing studies. It also discusses the major challenges associated with phishing detection, including evolving attack patterns, zero-day phishing websites, imbalanced datasets, adversarial manipulation, and real-time detection requirements. Finally, the survey identifies potential research directions involving hybrid machine learning, automated feature engineering, explainable artificial intelligence, deep learning, and real-time browser-based phishing detection systems [8, 9].

II. SURVEY METHODOLOGY

This survey follows a structured methodology to review research studies related to phishing website detection using supervised machine learning techniques. The methodology consists of literature identification, study selection, information extraction, classification of approaches, and comparative analysis. The main objective of the survey is to identify commonly used machine learning algorithms, phishing-related features, datasets, evaluation metrics, advantages, limitations, and emerging research directions.

A. Literature Search Strategy

Relevant research studies related to phishing website detection and supervised machine learning were identified from research literature published in the field of cybersecurity and machine learning. The search was focused on studies addressing phishing website detection, phishing URL classification, machine learning-based detection, and comparative analysis of supervised learning algorithms. The survey primarily considers recent studies published between 2022 and 2025, while selected earlier studies were also considered to understand the development of phishing detection techniques. The keywords considered for identifying relevant studies include “Phishing Website Detection,” “Phishing URL Detection,” “Machine Learning for Phishing Detection,” “Supervised Machine Learning,” “Phishing Detection using SVM,” “Phishing Detection using Random Forest,” and “URL-based Phishing Detection.”

B. Study Selection Criteria

The identified studies were screened according to their relevance to the research objective of this survey. Studies were included when they focused on phishing website or phishing URL detection using machine learning techniques and provided information about the applied methodology, features, datasets, or evaluation results. Studies that were not directly related to phishing detection or did not provide sufficient information about the detection approach were excluded from the detailed analysis. The final selected studies were then examined to identify the commonly used algorithms, feature types, performance measures, and research challenges.

C. Data Extraction and Classification

For each selected study, important information was extracted and organized according to several parameters. These parameters include publication year, machine learning algorithm, type of phishing-related features, dataset characteristics, detection approach, and evaluation metrics. The extracted studies were classified according to the machine learning techniques used, including Decision Tree (DT), Random Forest (RF), Support Vector Machine (SVM), K-Nearest Neighbors (KNN), Naïve Bayes (NB), Logistic Regression (LR), and ensemble learning methods. The feature categories considered in the survey include URL-based features, domain information, webpage and HTML characteristics, hyperlink properties, and security-related indicators.

D. Performance Analysis

The selected studies were analyzed using commonly reported performance metrics such as accuracy, precision, recall, F1-score, and ROC-AUC. Particular attention was given to recall because incorrectly classifying a phishing website as legitimate may expose users to security

and privacy risks. The survey also considers factors that can influence model performance, including dataset quality, feature selection, class imbalance, preprocessing techniques, and the ability of models to generalize to previously unseen phishing websites.

E. Research Gap Identification

Finally, the reviewed studies were analyzed to identify limitations and research gaps in existing phishing website detection approaches. The major issues identified include the detection of newly generated and zero-day phishing websites, dynamically changing URLs, imbalanced datasets, feature selection, adversarial manipulation, and the requirement for real-time detection. Based on these observations, potential future research directions such as hybrid and ensemble learning, automated feature engineering, explainable artificial intelligence, deep learning, continuously updated datasets, and real-time browser-based detection are identified.

III. LITERATURE REVIEW

Selvan.R et al. [1] proposed a phishing URL detection approach using combined machine learning methods to improve cybersecurity and protect users from malicious websites. The study focused primarily on URL-based characteristics that can be extracted without requiring extensive analysis of webpage content. By applying multiple machine learning techniques, the authors aimed to improve the classification of URLs into phishing and legitimate categories. The use of combined learning methods is particularly useful because different classifiers can capture different patterns present in phishing URLs, thereby improving the robustness of detection. The study highlights the importance of URL characteristics such as URL structure, length, special characters, domain-related information, and other suspicious patterns in identifying phishing attempts. The proposed approach demonstrates the potential of combining machine learning models for enhancing phishing detection performance and reducing the limitations associated with relying on a single classifier. However, URL-based detection may still face challenges when attackers generate sophisticated URLs that closely resemble legitimate website addresses. The work provides a useful foundation for developing more accurate and adaptive phishing URL detection systems using supervised machine learning techniques.

K. K et al. [2] presented a machine learning-based approach for detecting phishing websites. The study addresses the increasing threat posed by fraudulent websites that attempt to deceive users and obtain confidential information. The authors investigated machine learning models for distinguishing phishing websites from legitimate websites using relevant website characteristics. The general methodology involves collecting website-related features, preprocessing the dataset, training classification models, and evaluating their ability to correctly identify phishing instances. Machine learning provides an advantage over conventional blacklist-based approaches because models can learn patterns from previously identified phishing and legitimate websites. The study emphasizes the importance of selecting informative features that represent suspicious characteristics of phishing websites. Such characteristics may include URL-based properties, domain information, security indicators, and webpage-related attributes. The research demonstrates that machine learning can provide

an effective mechanism for automated phishing website detection. However, the effectiveness of the models depends on the quality and diversity of the training dataset and their ability to recognize newly generated phishing websites. This study therefore contributes to the development of automated and data-driven phishing detection systems.

Gujar [3] investigated different machine learning algorithms for detecting phishing websites. The study focuses on the application of classification algorithms to distinguish malicious phishing websites from legitimate websites based on their observable characteristics. The research demonstrates the importance of machine learning in cybersecurity because phishing websites continuously change their appearance, URLs, and technical characteristics to avoid conventional detection mechanisms. Various machine learning algorithms can be trained using labeled phishing datasets to learn relationships between website features and their corresponding classes. The study provides an overview of the effectiveness of machine learning algorithms in phishing website classification and highlights the role of appropriate feature selection in achieving reliable detection. The comparison of different algorithms is valuable because individual classifiers may exhibit different levels of accuracy, computational complexity, and generalization capability. Nevertheless, machine learning models may experience reduced performance when tested on previously unseen phishing websites or datasets with different characteristics from the training data. The work establishes the importance of evaluating multiple machine learning algorithms for developing reliable phishing website detection systems.

Sharma et al. [4] presented a review of real-time phishing website detection using machine learning techniques. The study emphasizes the need for real-time detection mechanisms because phishing websites can be created and distributed rapidly, potentially causing significant financial and privacy-related losses. Traditional approaches based on static databases may not be sufficiently effective against newly generated phishing websites. Machine learning provides an opportunity to analyze website characteristics automatically and classify websites according to their likelihood of being malicious. The authors reviewed machine learning-based approaches for phishing detection and discussed the importance of real-time identification. A real-time detection system must not only provide high classification accuracy but should also maintain low detection latency so that users can be protected before entering sensitive information. The study also highlights challenges related to feature extraction, continuously changing phishing strategies, and the requirement for updated training data. This review provides useful insights into the development of intelligent and real-time phishing detection frameworks and indicates that future systems should combine accurate classification with rapid detection capabilities.

Jain and Gupta [5] proposed a Support Vector Machine (SVM)-based machine learning technique for detecting phishing websites. SVM is a supervised classification algorithm that attempts to identify an optimal decision boundary between different classes of data. In phishing website detection, website characteristics can be represented as feature vectors and provided to the SVM classifier to differentiate legitimate websites from phishing websites. The approach is particularly relevant for cybersecurity applications because SVM can perform effectively in

high-dimensional feature spaces. The study demonstrates the applicability of SVM for automated phishing website classification and emphasizes the importance of extracting appropriate features from websites. The performance of an SVM-based detector depends significantly on the selected features and model parameters. Although SVM can provide effective classification, its computational requirements can increase for very large datasets, and its performance may be affected by noisy or highly overlapping features. Nevertheless, the study demonstrates that SVM is a promising supervised learning technique for phishing website detection and provides a foundation for comparing SVM with other classification algorithms.

Bhavani et al. [6] investigated the detection of legitimate and phishing websites using machine learning techniques. The study focuses on the binary classification problem of determining whether a website is trustworthy or malicious. The proposed machine learning-based approach uses website characteristics to identify patterns associated with phishing websites. This approach can reduce dependence on manually maintained phishing databases and provide automated classification of websites. The research highlights the significance of website feature extraction and preprocessing before applying machine learning algorithms. Features representing URL properties, webpage characteristics, and security-related information can provide useful indicators for distinguishing between legitimate and phishing websites. The study demonstrates the potential of machine learning to improve automated website security and assist users in avoiding malicious websites. However, because phishing strategies evolve continuously, models trained on older datasets may require periodic retraining and updated feature sets. The work contributes to the broader research area by demonstrating the applicability of supervised machine learning for legitimate and phishing website classification.

Anakal et al. [7] studied phishing website detection using machine learning methods. The research addresses the increasing sophistication of phishing attacks and the limitations of traditional security mechanisms in detecting newly generated malicious websites. The authors explored machine learning-based classification to identify suspicious website characteristics and categorize websites as phishing or legitimate. Machine learning models can learn patterns from historical examples and subsequently apply these patterns to unseen website samples. The study demonstrates the importance of preprocessing and feature extraction in constructing an effective phishing detection model. URL-related and webpage-related characteristics can be used as input features for classification algorithms. The application of multiple machine learning methods also enables researchers to compare their classification performance and determine suitable models for phishing detection. Despite the advantages of machine learning, the study area continues to face challenges such as dynamic phishing URLs, imbalanced datasets, and adversarial manipulation of website features. Overall, this work supports the use of machine learning as an effective approach for automated phishing website detection.

Zin et al. [8] investigated machine learning techniques for phishing website detection, focusing on the identification of characteristics that can differentiate malicious websites from legitimate ones. The study demonstrates how supervised machine learning can be applied to cybersecurity problems by training classification models using previously labeled website samples. The

detection process generally involves dataset preparation, feature extraction, preprocessing, model training, and performance evaluation. Selecting meaningful features is particularly important because phishing websites may exhibit different combinations of suspicious characteristics. Machine learning models can analyze these characteristics and identify complex relationships that may be difficult to capture using manually defined rules. The study contributes to the development of automated phishing detection by demonstrating the usefulness of machine learning techniques in website classification. However, model performance can be influenced by the quality of the dataset, feature representation, and distribution of phishing and legitimate samples. The research indicates that future phishing detection systems should focus on robust feature engineering and models capable of adapting to continuously evolving phishing attacks.

Uddin et al. [9] presented a comparative analysis of machine learning-based website phishing detection using URL information. Unlike approaches that require complete webpage content analysis, the study focuses on URL-based information for identifying phishing websites. URL features are attractive for phishing detection because they can potentially be extracted quickly and analyzed before a user fully interacts with a webpage. The research compares different machine learning approaches to determine their effectiveness in classifying phishing and legitimate URLs. Important URL characteristics can include URL length, number of special characters, domain-related properties, suspicious tokens, and structural patterns. The comparative analysis provides useful information about the relative performance of different machine learning algorithms and demonstrates that URL information can serve as an important source of phishing indicators. However, relying exclusively on URL characteristics may be insufficient when attackers construct highly convincing URLs that closely resemble legitimate addresses. Therefore, combining URL-based features with domain, HTML, webpage, and security-related features may provide more comprehensive detection. The study is particularly relevant to lightweight and rapid phishing detection systems.

Soumya et al. [10] proposed a Logistic Regression-based machine learning technique for phishing website detection. Logistic Regression is a supervised classification algorithm that estimates the probability of an instance belonging to a particular class. In phishing detection, website features can be provided as input to the model, which then predicts whether the corresponding website is legitimate or phishing. The study demonstrates the applicability of a relatively simple and computationally efficient machine learning algorithm to cybersecurity-based classification. Logistic Regression can be useful for phishing detection because it provides a straightforward probabilistic interpretation of the classification outcome and can serve as an effective baseline model for comparison with more complex algorithms. However, its performance may be limited when phishing and legitimate websites exhibit highly nonlinear relationships between their features. The study highlights the potential of supervised learning for automated phishing detection and provides a basis for comparing Logistic Regression with advanced classification and ensemble techniques. The research also indicates the importance of suitable feature selection and preprocessing for improving classification performance.

Table 1. Comparative Analysis of Existing Phishing Website Detection Studies

Ref.	Author(s) / Year	Detection Approach	Major Features / Information Used	ML Technique	Major Observation	Limitation / Research Gap
[1]	Selvan.R et al. (2025)	Phishing URL Detection	URL structure, URL length, special characters, domain-related information	Combined ML Methods	Combined learning can improve phishing URL classification by capturing different patterns	Sophisticated URLs may closely resemble legitimate URLs
[2]	K. K et al. (2025)	Phishing Website Detection	Website characteristics, URL, domain and security- related attributes	Machine Learning	ML provides automated classification of phishing and legitimate websites	Performance depends on dataset quality and diversity
[3]	Gujar (2024)	Phishing Website Classification	Observable website characteristics and selected features	Multiple ML Algorithms	Comparison of different classifiers helps analyze classification capability	Generalization may decrease for unseen phishing websites
[4]	Sharma et al. (2023)	Real-Time Phishing Detection	Website characteristics and extracted phishing indicators	Machine Learning	Highlights the importance of rapid and real- time phishing identification	Feature extraction, changing phishing strategies and updated training data remain challenging
[5]	Jain & Gupta (2023)	Phishing Website Detection	Website characteristics represented as feature vectors	SVM	SVM is applicable to phishing classification and high-dimensional feature spaces	Computational requirements may increase for large datasets
[6]	Bhavani et al. (2023)	Legitimate vs. Phishing Website Detection	URL, webpage and security- related features	Machine Learning	ML can automate classification and reduce dependence on manually maintained databases	Models require updated datasets and feature sets
[7]	Anakal et al. (2023)	Phishing Website Detection	URL-related and webpage- related characteristics	Multiple ML Methods	Multiple ML models enable comparative evaluation of	Dynamic URLs, imbalanced datasets and adversarial

					phishing detection	manipulation remain challenges
[8]	Zin et al. (2023)	Phishing Website Detection	Website characteristics and extracted features	Supervised ML	Meaningful feature selection is important for phishing classification	Dataset quality, feature representation and changing attacks affect performance
[9]	Uddin et al. (2022)	URL-Based Phishing Detection	URL length, special characters, domain properties, suspicious tokens and URL patterns	Comparative ML Approaches	URL information can provide useful indicators for lightweight and rapid detection	URL-only detection may fail against highly convincing URLs
[10]	Soumya et al. (2022)	Phishing Website Detection	Website-related features	Logistic Regression	LR provides a simple and computationally efficient baseline	Performance may be limited for highly nonlinear feature relationships

3.1 Dataset-Wise Analysis

Datasets play an important role in the development and evaluation of machine learning-based phishing website detection systems. The quality, size, diversity, feature representation, and class distribution of a dataset directly influence the performance and generalization capability of a detection model. The reviewed studies primarily utilize information extracted from phishing and legitimate URLs, domains, webpages, HTML structures, hyperlinks, and security-related characteristics. Based on the reviewed literature, the datasets can broadly be analyzed according to the type of information used for phishing detection.

URL-Based Datasets: URL-based datasets contain characteristics extracted directly from website addresses. Common information includes URL length, special characters, suspicious tokens, domain-related information, IP address usage, and URL structure. Such datasets are useful for lightweight and rapid phishing detection because features can potentially be extracted before the complete webpage is accessed. The study by Uddin et al. [9], for example, focuses on URL information for phishing detection. However, URL-only datasets may have limitations when phishing URLs are designed to closely resemble legitimate URLs.

Website and Webpage Feature Datasets: Some studies use broader website-related characteristics instead of relying only on URLs. These datasets may represent webpage properties, HTML characteristics, hyperlinks, and other observable website attributes. Such datasets provide more information for classification and can capture phishing indicators that may not be visible from the URL alone. However, extracting webpage-level features may require additional processing and can increase the computational requirements of the detection

system. Your paper identifies URL, domain, webpage structure, HTML content, hyperlinks, and security information as important feature categories.

Domain and Security Information: Domain-related and security-related information can also be used to distinguish phishing websites from legitimate websites. Examples discussed in your review include domain information, domain age, HTTPS usage, SSL/TLS-related characteristics, and suspicious domain properties. These features can provide additional information about the trustworthiness of a website. However, individual security indicators should not be considered sufficient by themselves because, as your paper notes, HTTPS alone does not prove that a website is legitimate.

Combined Feature Datasets: A more comprehensive dataset can combine URL, domain, webpage, HTML, hyperlink, and security-related features. Combining multiple feature categories can provide a broader representation of phishing behavior and potentially improve robustness against different types of phishing attacks. However, increasing the number of features can also introduce redundancy, computational complexity, and the need for effective feature-selection techniques. Your paper specifically identifies feature selection and dataset quality as important factors affecting machine-learning performance.

Dataset Challenges: The reviewed literature indicates several dataset-related challenges. Phishing websites continuously change their URLs, domains, webpage designs, and technical characteristics. Consequently, a model trained using older datasets may not generalize effectively to newly generated or zero-day phishing websites. Class imbalance, insufficient diversity, outdated samples, and differences between training and real-world data can further affect model performance. Therefore, future phishing detection research should consider continuously updated, diverse, and representative datasets to improve the generalization capability of machine learning models.

3.2 Taxonomy of Phishing Detection Approaches and Features

Phishing website detection approaches can be broadly categorized according to the type of detection mechanism and the characteristics extracted from websites. Based on the reviewed literature, phishing detection can be organized into traditional detection approaches and machine learning-based approaches. Machine learning-based detection can further be classified according to the type of features used for identifying phishing websites.

Table 2. Taxonomy of Phishing Detection Approaches and Features

Category	Sub-category	Representative Features / Techniques	Main Purpose
Traditional Detection	Blacklist	Known malicious URLs/domains	Detect previously reported phishing websites
	Whitelist	Known legitimate websites	Allow trusted websites
	Rule/Signature-based	Predefined patterns and rules	Identify known suspicious characteristics
Machine Learning	Supervised ML	DT, RF, SVM, KNN, NB, LR	Classify phishing and legitimate websites
	Ensemble ML	Combined classifiers	Improve robustness through multiple models

URL Features	URL structure	URL length, special characters, suspicious tokens	Identify suspicious URL patterns
	URL obfuscation	Encoded/manipulated URLs, redirects	Detect hidden or misleading destinations
Domain Features	Domain characteristics	Domain age, domain information	Identify suspicious domains
Webpage/HTML Features	Page structure	HTML elements, forms, webpage content	Detect fraudulent webpage characteristics
Hyperlink Features	Link characteristics	Hyperlinks and link structures	Identify suspicious linking behavior
Security Features	Security indicators	HTTPS, SSL/TLS information	Provide additional security information
Combined Features	Multi-feature approach	URL + domain + webpage + HTML + security	Comprehensive phishing detection

IV. PHISHING WEBSITE DETECTION

A phishing website is a fraudulent or malicious website created by attackers to imitate a legitimate website and deceive users into providing sensitive information. These websites commonly copy the appearance, logos, colors, login forms, and content of trusted platforms such as online banking, e-commerce, email, and social-media websites. The primary objective is to steal confidential information such as usernames, passwords, banking credentials, credit-card details, and personal information.

Phishing websites are generally distributed through fraudulent emails, SMS messages, social-media posts, instant messaging, advertisements, and malicious URLs. Attackers often use deceptive domain names, URL redirection, shortened links, fake security warnings, and similar webpage designs to make the malicious website appear genuine. When a user visits such a website and enters confidential information, the information can be captured by the attacker.



Figure 1: Phishing Website

Important characteristics that can indicate a phishing website include:

1. **Suspicious URL:** Unusual domain names, excessive URL length, or misleading subdomains.
2. **IP Address in URL:** Some phishing websites directly use an IP address instead of a conventional domain name.
3. **URL Obfuscation:** Attackers may use special characters, redirections, or encoded URLs to hide the actual destination.
4. **Fake Login Forms:** Fraudulent login pages are designed to collect usernames and passwords.
5. **Poor or Copied Content:** Spelling errors, unusual language, or copied content may indicate a malicious website.
6. **Missing or Suspicious HTTPS:** The absence of appropriate security mechanisms can be an indication of risk, although HTTPS alone does not prove that a website is legitimate.
7. **Unusual Domain Information:** Newly registered or suspicious domains may be associated with phishing activities.
8. **Excessive Redirects:** Multiple redirects can be used to hide the final malicious destination.

A phishing website is a fraudulent website designed to imitate a legitimate website with the intention of deceiving users into revealing sensitive and confidential information. Such information may include usernames, passwords, banking credentials, credit card details, personal identification information, and other authentication data. Phishing websites often use misleading domain names, shortened or manipulated URLs, copied logos, similar webpage layouts, fake login forms, and deceptive content to appear trustworthy. Attackers commonly distribute phishing links through emails, social media platforms, instant messaging applications, advertisements, and other online communication channels.

Phishing website detection refers to the process of identifying whether a website is legitimate or malicious before the user interacts with it or provides sensitive information. Traditional detection techniques mainly rely on blacklists, whitelists, signatures, and manually defined security rules. Although these methods can effectively identify known phishing websites, they may have difficulty detecting newly generated phishing websites because attackers can rapidly create new domains and modify website characteristics.

Machine learning-based detection provides a more adaptive approach by learning patterns from previously identified phishing and legitimate websites. A typical detection framework consists of data collection, preprocessing, feature extraction, feature selection, model training, classification, and performance evaluation. Features may be extracted from the URL, domain, webpage structure, HTML content, hyperlinks, JavaScript behavior, SSL/TLS information, and other security-related characteristics. Supervised machine learning algorithms such as Decision Tree, Random Forest, Support Vector Machine, Logistic Regression, K-Nearest Neighbors, and ensemble methods can then classify websites into phishing and legitimate categories.

The effectiveness of phishing website detection is generally evaluated using accuracy, precision, recall, F1-score, and ROC-AUC. Among these metrics, precision and recall are

particularly important because incorrectly classifying a phishing website as legitimate can expose users to serious security risks. Therefore, an effective phishing detection system should provide high detection accuracy while maintaining low false-negative and false-positive rates. Combining multiple feature types and advanced machine learning techniques can further improve the robustness of phishing website detection against evolving cyber threats.

V. MACHINE LEARNING

Machine Learning (ML) is a branch of Artificial Intelligence (AI) that enables computer systems to learn patterns from existing data and make predictions or decisions without being explicitly programmed for every task. In phishing website detection, machine learning is used to learn the characteristics of both legitimate and phishing websites and then classify a new website as Legitimate or Phishing.

A typical ML-based phishing detection system consists of several stages: data collection, data preprocessing, feature extraction, feature selection, model training, classification, and performance evaluation. Features can be extracted from URLs, domains, HTML content, hyperlinks, security certificates, and webpage behavior. The extracted features are provided to a machine learning classifier that learns the relationship between website characteristics and their corresponding classes.

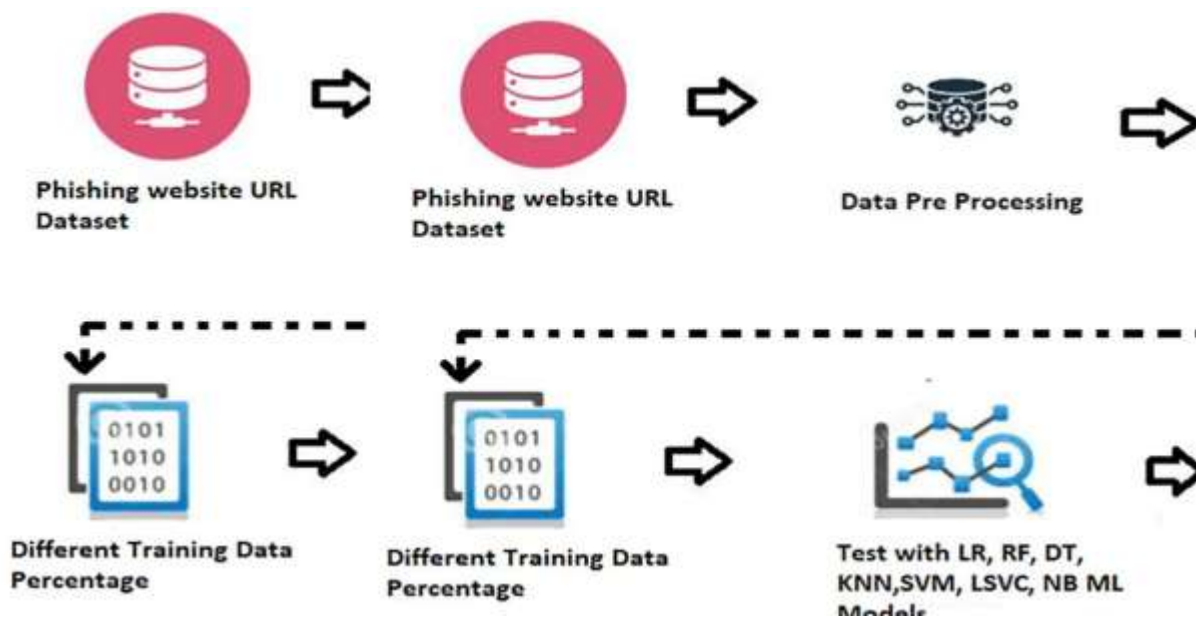


Figure 2: ML Approach

For supervised machine learning, the training dataset contains two or more labeled classes, such as Phishing and Legitimate. Common algorithms used for phishing website detection include:

- Decision Tree (DT)
- Random Forest (RF)
- Support Vector Machine (SVM)

- Logistic Regression (LR)
- K-Nearest Neighbors (KNN)
- Naïve Bayes (NB)
- Ensemble Learning

The trained model is evaluated using metrics such as Accuracy, Precision, Recall, F1-Score, and ROC-AUC. Among these, recall is particularly important because a false classification of a phishing website as legitimate may expose users to credential theft, financial loss, and privacy violations. Thus, machine learning provides an adaptive and automated approach for detecting both known and potentially previously unseen phishing websites.

VI. CONCLUSION

Phishing websites continue to pose a significant cybersecurity threat because attackers continuously develop sophisticated techniques to deceive users and steal sensitive information. This survey reviewed recent research on phishing website detection using supervised machine learning techniques, covering studies published between 2022 and 2025. The reviewed literature demonstrates that machine learning algorithms such as Support Vector Machine, Logistic Regression, Decision Tree, Random Forest, and other combined or ensemble approaches can effectively classify websites and URLs as legitimate or phishing. URL-based features, domain characteristics, webpage attributes, and security-related features are commonly used for developing detection models.

The literature analysis also indicates that no single machine learning algorithm is universally optimal for all phishing detection scenarios. Model performance depends on dataset quality, feature selection, preprocessing, class distribution, and the ability to detect previously unseen phishing patterns. Combined machine learning methods can provide improved robustness by utilizing the strengths of multiple classifiers, while lightweight models such as Logistic Regression and SVM can be useful for rapid and computationally efficient detection. However, challenges remain in detecting zero-day phishing websites, handling dynamically changing URLs, maintaining updated datasets, addressing imbalanced data, and resisting adversarial attacks.

Future research should therefore focus on developing hybrid and ensemble machine learning frameworks that integrate URL, domain, HTML, and webpage-content features. Feature-selection and automated feature-engineering techniques can further improve model efficiency and reduce unnecessary computational complexity. The integration of explainable artificial intelligence can also provide users and security administrators with understandable reasons for phishing predictions. Furthermore, real-time browser-based detection, continuously updated datasets, deep learning, and adaptive learning techniques represent promising directions for improving the accuracy, scalability, and robustness of future phishing website detection systems.

REFERENCES

- [1] M. Selvan.R, Bhavana.Jasti, Rushitha.Kuchi, ChakriSriRam. D and Venkataramanaiah. G, "Phishing URL's Detection for Enhancing Security with Combined Machine Learning Methods," *2025 IEEE International Conference on Innovate for Humanitarian: Tech Solutions for Global Challenge (ICIH)*, Indore, India, 2025, pp. 1-5.
- [2] K. K, A. Khan, R. M R, P. N K and P. S, "Phishing Websites Detection using Machine Learning Model," *2025 4th International Conference on Innovative Mechanisms for Industry Applications (ICIMIA)*, Tirupur, India, 2025, pp. 1868-1873.
- [3] S. S. Gujar, "Machine Learning Algorithms for Detecting Phishing Websites," *2024 International Conference on Innovative Computing, Intelligent Communication and Smart Electrical Systems (ICSES)*, Chennai, India, 2024, pp. 1-6.
- [4] K. Sharma, P. Rai and J. Chandel, "Review Paper Real-Time Phishing Website With Machine Learning," *2023 11th International Conference on Intelligent Systems and Embedded Design (ISED)*, Dehradun, India, 2023, pp. 1-5.
- [5] S. Jain and C. Gupta, "A Support Vector Machine Learning Technique for Detection of Phishing Websites," *2023 6th International Conference on Information Systems and Computer Networks (ISCON)*, Mathura, India, 2023, pp. 1-6.
- [6] A. Bhavani, R. S. Lakshmi, P. Harshavardhini, P. V. Prakash, N. V. Behara and V. A. Kumar, "Detection of Legitimate and Phishing Websites using Machine Learning," *2023 International Conference on Sustainable Computing and Smart Systems (ICSCSS)*, Coimbatore, India, 2023, pp. 366-371.
- [7] S. Anakal, K. Maka, A. Tadkal, S. Humanabad, S. Anakal and E. Laxmikant, "Phishing Website Detection Using Machine Learning Methods," *2023 International Conference on Integrated Intelligence and Communication Systems (ICIICS)*, Kalaburagi, India, 2023, pp. 1-5.
- [8] N. A. B. M. Zin, M. F. Ab Razak, A. Firdaus, F. Ernawan and N. S. A. Zulkifli, "Machine Learning Technique for Phishing Website Detection," *2023 IEEE 8th International Conference On Software Engineering and Computer Systems (ICSECS)*, Penang, Malaysia, 2023, pp. 235-239.
- [9] 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," *2022 5th International Conference on Pattern Recognition and Artificial Intelligence (PRAI)*, Chengdu, China, 2022, pp. 220-224.
- [10] T. R. Soumya, R. P, N. A. Rosy, N. Pughazendi, S. Padmapriya and R. Khilar, "Logistic Regression based Machine Learning Technique for Phishing Website Detection," *2022 4th International Conference on Inventive Research in Computing Applications (ICIRCA)*, Coimbatore, India, 2022, pp. 683-686.
- [11] A. N. S. Charan, Y. -H. Chen and J. -L. Chen, "Phishing Websites Detection using Machine Learning with URL Analysis," *2022 IEEE World Conference on Applied Intelligence and Computing (AIC)*, Sonbhadra, India, 2022, pp. 808-812.
- [12] N. Binti Md Noh and M. N. Bin M. Basri, "Phishing Website Detection Using Random Forest and Support Vector Machine: A Comparison," *2021 2nd International Conference on Artificial Intelligence and Data Sciences (AiDAS)*, IPOH, Malaysia, 2021, pp. 1-5.
- [13] S. Alharbi, H. Alshehri, and A. Alzahrani, "Machine Learning Techniques for Phishing Detection: A Review," *IEEE Access*, vol. 9, pp. 110063–110076, 2021.
- [14] K. S. Swarnalatha, K. C. Ramchandra, K. Ansari, L. Ojha and S. S. Sharma, "Steady Peril Information Block Phising Attacks," *2021 IEEE Overall Social occasion on Computation System and Information Advancement for Acceptable Courses of action (CSITSS)*, 2021, pp. 1–6.
- [15] B. J. S. Singh and M. S. Singh, "Phishing Detection Using Machine Learning Algorithms: A Review," *IEEE International Conference on Emerging Trends in Information Technology and Engineering*, pp. 1–5, 2020.
- [16] M. N. Alam, D. Sarma, F. F. Lima, I. Saha, R. E. Ulfath, and S. Hossain, "Phishing attacks detection using machine learning approach," in *Proceedings of the 3rd International Conference on Smart Systems and Inventive Technology, ICSSIT 2020, Institute of Electrical and Electronics Engineers Inc.*, Aug. 2020, pp. 1173–1179.
- [17] Choon Lin Tan, "Phishing Dataset for Machine Learning: Feature Evaluation," *Mendeley Data*, Mar. 24, 2018.
- [18] F. Vanhoenshoven, G. Napoles, R. Falcon, K. Vanhoof, and M. Koppen, "Detecting malicious URLs using machine learning techniques," in *2016 IEEE Symposium Series on Computational Intelligence, SSCI 2016, Institute of Electrical and Electronics Engineers Inc.*, Feb. 2017