Impact Factor (QJIF): 8.47 E-ISSN: 2583-1615, P-ISSN: 3049-3498

International Journal of Research in Academic World

Received: 25/January/2026 **IJRAW:** 2026; 5(3):225-232 **Accepted:** 07/March/2026

Cybersecurity and Phishing Detection using AI/ML

[1]Sakshi Gaikwad, [2]Bhumika Patil, [*3]Roshani Kinge, [4]Sakshi S Jawale, [5]Vaidai Jambhule and [6]Dr. Aparna Vaidyanathan

[1, *2, 3, 4, 5]Student, Department of MSc Computer Science (SY), Fergusson College, Savitribai Phule Pune University, Maharashtra, India.

[6]Professor, Department of Computer Science, Fergusson College, Savitribai Phule Pune University, Maharashtra, India.

Abstract

Phishing attacks are one of the most common and dangerous cyber threats that target internet users by stealing sensitive information such as usernames, passwords, and financial details. Traditional security mechanisms like blacklists and rule-based filters are often unable to detect newly created phishing websites. Therefore, intelligent detection techniques are required to improve cyber security.

This research proposes a machine learning based approach for detecting phishing websites by analysing various website and URL features. Different machine learning algorithms can be used to classify websites as legitimate or phishing based on patterns found in the data. The proposed system focuses on improving detection accuracy and reducing false positives by analysing multiple characteristics of websites such as URL structure, domain information, and webpage behaviour.

Experimental results demonstrate that machine learning techniques can effectively identify phishing websites and enhance user protection against cyber-attacks. The study highlights the importance of intelligent security systems in combating modern cyber threats and protecting users from online fraud.

Keywords: Phishing Detection, Cyber Security, Machine Learning, URL Analysis, Website Classification, Network Security, Online Fraud Detection, Web Security.

1. Introduction

The rapid growth of the internet and digital technologies has significantly transformed the way individuals and organizations communicate, conduct business, and access information. Online services such as banking, e-commerce, social networking, and cloud-based platforms have become an essential part of everyday life. However, this increasing dependence on internet-based services has also led to a rise in cyber security threats. Cybercriminals constantly develop new methods to exploit vulnerabilities in online systems and deceive users for malicious purposes.

One of the most common and dangerous cyber threats is phishing. Phishing is a fraudulent activity in which attackers attempt to trick users into revealing sensitive information such as usernames, passwords, credit card details, and other personal data. These attacks are typically carried out through deceptive emails, fake websites, or malicious links that appear to come from trusted organizations. Unsuspecting users may unknowingly provide their confidential information, which can then be misused by attackers for financial fraud, identity theft, or unauthorized access to online accounts.

Over the years, phishing attacks have become increasingly sophisticated. Modern phishing websites are designed to closely imitate legitimate websites, making them difficult for users to identify. Attackers often use techniques such as domain spoofing, URL manipulation, and website cloning to deceive victims. Due to the rapid creation of new phishing websites, traditional detection techniques such as blacklist-based methods and rule-based filtering systems are often insufficient. These methods mainly rely on previously reported malicious websites and may fail to detect newly generated phishing attacks.

To overcome these limitations, researchers have explored the use of advanced technologies such as machine learning for phishing detection. Machine learning techniques can analyse large volumes of data and identify hidden patterns that help distinguish between legitimate and phishing websites. Various machine learning algorithms such as decision trees, support vector machines, and neural networks have been studied to improve phishing detection accuracy.

This research focuses on analysing existing machine learning based phishing detection approaches and understanding their effectiveness in identifying malicious websites. The study reviews different techniques used for phishing detection and evaluates their strengths, limitations, and overall performance. By examining previously proposed methods, this research

***Corresponding Author:** Roshani Kinge

aims to provide a better understanding of how machine learning contributes to improving cyber security and protecting users from phishing attacks.

2. Problem Statement

Phishing attacks have become one of the most serious cyber security threats in recent years. Attackers create fraudulent websites and deceptive messages that closely resemble legitimate platforms in order to trick users into revealing sensitive information such as usernames, passwords, and financial details. As internet usage continues to grow, phishing attacks are also increasing in frequency and sophistication.

Traditional phishing detection techniques such as blacklist-based systems and rule-based filters are commonly used to identify malicious websites. However, these approaches have several limitations. Blacklist systems rely on previously identified phishing websites, which means they cannot effectively detect newly created or unknown phishing attacks. Similarly, rule-based methods depend on predefined patterns and rules, which may not adapt quickly to evolving phishing techniques.

With the increasing complexity of phishing attacks, there is a need to analyse and understand more advanced detection techniques that can improve the identification of malicious websites. Machine learning has emerged as a promising approach for phishing detection because it can analyse large datasets and identify patterns that distinguish phishing websites from legitimate ones.

Therefore, the main problem addressed in this research is the need to analyse existing machine learning based phishing detection approaches and evaluate their effectiveness in identifying phishing websites while improving cyber security.

3. Objectives of the Study

The main objective of this research is to analyse existing machine learning based techniques used for phishing website detection and understand their role in improving cyber security. The study aims to examine different methods used by researchers and evaluate their effectiveness in identifying malicious websites.

The specific objectives of this study are:

- To study the concept of phishing attacks and their impact on cyber security.
- To analyse different machine learning techniques used for phishing website detection.
- To examine various features used for identifying phishing websites such as URL structure, domain information, and webpage behaviour.
- To evaluate the strengths and limitations of existing phishing detection methods.
- To understand how machine learning techniques contribute to improving the accuracy of phishing detection systems.
- To highlight the importance of advanced security mechanisms in protecting users from online fraud and cyber-attacks.

4. Literature Review

Mohamed (2024): Mohamed [1] provides a comprehensive overview of how Artificial Intelligence (AI) and Machine Learning (ML) are transforming cybersecurity by improving intrusion detection, malware classification, behavioral analysis, and threat intelligence. Traditional security mechanisms such as firewalls and signature-based systems are often ineffective against modern threats including zero-day attacks, ransomware, and advanced persistent threats. AI and ML provide adaptive and data-driven approaches that continuously learn from evolving attack patterns.

The study highlights the use of supervised, unsupervised, and reinforcement learning techniques, along with deep neural networks, to detect anomalies in real time and automate incident response. However, challenges such as adversarial machine learning, scalability issues, and the "black-box" nature of deep learning models remain significant concerns. The author also emphasizes the importance of explainable AI, federated learning, and advanced intelligent frameworks for developing more reliable cybersecurity systems.

Mandora, Mehta and Mehta (2024): Mandora, Mehta, and Mehta [2] focused specifically on phishing detection, which remains one of the most damaging forms of cybercrime. Their research shows that AI-based detection systems outperform traditional approaches such as blacklists, rule-based heuristics, and signature matching techniques.

The authors applied machine learning algorithms including Support Vector Machines (SVM), Random Forest, and deep learning models such as Convolutional Neural Networks (CNN) and Recurrent Neural Networks (RNN) to classify phishing attempts across multiple communication channels including email, SMS, social media, and websites. In addition, Natural Language Processing (NLP) techniques were used to analyse writing styles, sentiment, and deceptive linguistic patterns within phishing messages.

Although these approaches achieved high detection accuracy, the study identified challenges such as false positives, adversarial phishing techniques, and the high computational cost of deep learning models.

Singh and Maurya (2024): Singh and Maurya (2024) reviewed various machine learning algorithms used for phishing detection, including Random Forest, XGBoost, Na¨ıve Bayes, and Support Vector Machine. Their study emphasized the importance of proper data preprocessing techniques such as data cleaning, normalization, and feature extraction to improve model reliability and performance.

Using a dataset containing more than 11,000 samples, the researchers found that Gradient Boosting achieved the highest accuracy of 97.6% in detecting phishing websites. The study also highlighted that ensemble learning methods are generally more effective than traditional classifiers because they can capture complex relationships within large datasets.

However, the authors noted challenges related to dataset diversity and the ability of models to detect newly emerging phishing attacks in real time.

Yadav *et al.* (2025): Yadav *et al.* (2025) proposed an advanced AI-based phishing detection framework that combines Natural Language Processing (NLP) with ensemble machine learning techniques for comprehensive threat analysis. The model utilizes DistilBERT embeddings to analyse textual content such as emails and webpages, while URL and HTML features are processed using Random Forest and XGBoost classifiers.

This hybrid model achieved an accuracy of 97.2%, demonstrating strong performance compared to traditional detection methods. The study also discussed previous research involving deep learning models such as Long Short-Term Memory (LSTM), Convolutional Neural Networks (CNN), and Transformer architectures.

While these models improve contextual understanding, they often require high computational resources. To address this issue, the authors developed a lightweight and scalable

architecture that can be deployed across platforms using technologies such as Flask and Flutter.

WJAETS (2025): The WJAETS (2025) study reviewed the growing importance of Artificial Intelligence and Machine Learning in enhancing cybersecurity systems. According to the study, traditional security solutions struggle to address the increasing complexity and diversity of cyber threats such as phishing attacks, malware, and ransomware.

AI and ML techniques enable systems to analyse large volumes of data and detect previously unknown threats by identifying unusual patterns and behaviours. The paper also examined deep learning and reinforcement learning approaches used for detecting phishing websites and malicious activities. Although these techniques significantly improve detection accuracy, the study highlighted challenges related to high computational requirements, the need for large datasets, and difficulties in interpreting AI-based decisions.

AI in Cyber Security (2024): This study provides a detailed overview of how Artificial Intelligence and Machine Learning enhance cyber threat detection and response mechanisms. Traditional security systems often fail to detect new attack patterns due to their static and rule-based nature.

In contrast, machine learning models use historical data to identify suspicious activities and detect phishing attacks more effectively. The study also presents several real-world case studies where AI-based tools successfully identified phishing attempts and other cyber threats.

However, the research emphasizes ongoing challenges such as limited availability of quality datasets, adversarial attacks targeting AI models, and the lack of transparency in complex machine learning systems.

Kumar *et al.* (2023–2024): Kumar *et al*. (2023–2024) discussed the role of Artificial Intelligence and Machine Learning in developing intelligent cybersecurity frameworks capable of detecting sophisticated cyber-attacks. Their research highlighted that traditional rulebased security systems are unable to cope with rapidly evolving threats.

AI-driven models can process large datasets, identify hidden attack patterns, and respond to threats in real time. Supervised learning models are particularly effective for detecting known attack signatures, while unsupervised and reinforcement learning techniques help identify unknown threats and automate defense strategies.

Despite these advantages, the study also pointed out challenges such as data bias, adversarial attacks, and privacy concerns associated with large-scale data analysis.

Ajayi *et al.* (2025): Ajayi *et al*. (2025) examined the role of Artificial Intelligence and Machine Learning in strengthening cybersecurity systems and protecting critical infrastructure. Their research highlighted how ML algorithms such as Random Forest, XGBoost, Support Vector Machines, and ensemble learning methods achieve high accuracy in detecting phishing attacks across different platforms including emails, websites, and social media.

The integration of Natural Language Processing further improves phishing detection by analysing linguistic patterns and user behaviour. However, the authors identified several challenges including adversarial attacks on AI models, high computational requirements, limited dataset diversity, and the lack of explainability in complex models.

The study suggests that future research should focus on explainable AI, hybrid detection frameworks, and scalable architectures to enhance the reliability and transparency of AI-based cybersecurity systems.

5. Research Methodology

The research methodology describes the systematic approach followed in this study to analyse machine learning based phishing website detection techniques. The purpose of this methodology is to understand how different website features and machine learning algorithms can be used to identify phishing websites and improve cyber security mechanisms. This research focuses on analysing existing approaches rather than developing a completely new system.

The methodology consists of several stages including dataset collection, data visualization, data preprocessing, feature analysis, machine learning algorithm analysis, model evaluation, and result analysis. Each stage contributes to understanding how phishing detection systems operate and how different machine learning models perform in identifying malicious websites.

a) Dataset Collection: The first step in the research methodology is the collection of datasets related to phishing websites. A dataset is a collection of records that contains different attributes describing the characteristics of websites and URLs. These attributes help in identifying whether a website is legitimate or phishing.

Phishing detection datasets generally contain features such as URL length, presence of special characters in URLs, use of HTTPS protocol, domain age, website traffic ranking, and redirection behaviour. Each entry in the dataset is labelled as either phishing or legitimate.

These datasets are usually obtained from publicly available cyber security repositories and research platforms. Using such datasets enables researchers to analyse the behaviour of phishing websites and identify patterns that distinguish them from legitimate websites.

b) Data Visualization: Data visualization is an important step that helps researchers understand the structure and distribution of the dataset. Visualization techniques are used to represent data graphically so that patterns, trends, and relationships between features can be easily observed.

Graphical representations such as bar charts, histograms, and feature distribution graphs are commonly used to analyse phishing datasets. These visualizations help in identifying the characteristics that frequently appear in phishing websites.

For example, phishing websites often contain unusually long URLs, multiple subdomains, or suspicious domain structures. By visualizing the dataset, researchers can determine which features play a significant role in phishing detection.

c) Data Preprocessing: Data preprocessing is performed to improve the quality of the dataset before applying machine learning algorithms. Raw datasets may contain missing values, duplicate records, or inconsistent data formats that can affect the performance of machine learning models.

During the preprocessing stage, several techniques are applied to clean and prepare the data. Duplicate entries are removed, missing values are handled appropriately, and categorical attributes are converted into numerical formats suitable for machine learning models.

Another important step in preprocessing is data normalization, which ensures that all feature values fall within a similar range. This prevents features with large numerical values from dominating other features during model training.

d) Feature Analysis: Feature analysis involves examining the different attributes present in the dataset to determine their importance in identifying phishing websites.

Features are generally divided into three categories:

i). URL Based Features: URL based features are derived directly from the website address.

- Length of the URL
- Presence of special characters such as @, -, or _
- Use of IP address instead of domain name
- Number of subdomains in the URL

ii). Domain Based Features: Domain based features provide information about domain registration and ownership.

- Domain age
- Domain registration length
- DNS record availability
- Website traffic ranking

iii). Webpage Based Features: Webpage based features analyse the structure and behaviour of webpage content.

- Presence of iframe elements
- Number of external links
- Redirection behaviour
- Loading of external scripts

e) Machine Learning Algorithm Analysis

Machine learning algorithms play an important role in phishing website detection because they can analyse large datasets and identify patterns that distinguish phishing websites from legitimate ones.

Several classification algorithms are analysed in this study, including Logistic Regression, Support Vector Machine (SVM), Random Forest, and K-Nearest Neighbour (KNN).

i). Logistic Regression: Logistic Regression is used for binary classification problems.

$$P(y = 1|x) = \frac{1}{1 + e^{-(\beta 0 + \beta 1\, x1 + \beta 2\, x2 + \cdots + \beta nxn)}} \quad (1)$$

Where:

- $P(y = 1\ x)$ represents the probability that the website is phishing
- $x_1, x_2, \ldots, x_n$ are input features
- β_0 is the intercept
- $\beta_1, \beta_2, \ldots, \beta_n$ are feature coefficients
- e is Euler's constant

ii). Support Vector Machine (SVM): SVM separates data using an optimal hyperplane.

$$w \cdot x + b = 0 \quad (2)$$

The classification function is:

$$f(x) = \text{sign}(w \cdot x + b) \quad (3)$$

Where w is the weight vector and b is the bias.

iii). Random Forest: Random Forest is an ensemble learning algorithm that combines multiple decision trees.

$$\hat{y} = \frac{1}{N}\sum_{i=1}^{N} T_i(x)$$

Where:

$T_i(x)$ is the prediction of the i^{th} tree

- N is the total number of trees
- $\hat{y}$ is the final predicted output

iv). K-Nearest Neighbour (KNN): KNN classifies data based on the distance between feature vectors.

$$d(x, y) = \sqrt{\sum_{i=1}^{n}(x_i - y_i)^2}$$

Where x_i and y_i represent feature values and n is the number of features.

f) Model Evaluation

The performance of machine learning models is evaluated using the following metrics.

Accuracy

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \quad (6)$$

Precision

$$Precision = \frac{TP}{TP + FP} \quad (7)$$

Recall

$$Recall = \frac{TP}{TP + FN} \quad (8)$$

F1 Score

$$F\,1 = \frac{2 \times Precision \times Recall}{Precision + Recall} \quad (9)$$

Where:

- TP = True Positives
- TN = True Negatives
- FP = False Positives
- FN = False Negatives

g) Methodology Workflow

The overall workflow of the research methodology is as follows:

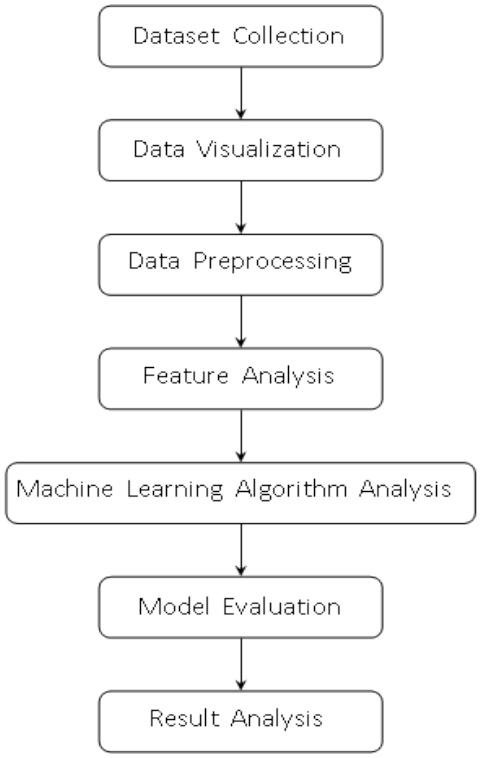

Fig 1: Research Methodology Workflow for Phishing Detection

6. System Architecture

The system architecture describes the overall structure of the phishing detection analysis process and the interaction between different components involved in identifying phishing websites. Although this research focuses on analysing existing machine learning techniques rather than developing a completely new system, a conceptual architecture helps in understanding how phishing detection systems generally operate.

The architecture consists of several modules including data collection, feature extraction, data preprocessing, machine learning analysis, and result evaluation. These modules work together to analyse website data and classify websites as phishing or legitimate.

The architecture begins with collecting datasets containing information about website characteristics such as URL features, domain information, and webpage behaviour. These features are then processed and analysed using machine learning algorithms to determine whether a website is malicious or safe. The architecture helps illustrate the workflow followed in phishing detection systems and highlights how machine learning models contribute to improving cyber security by identifying fraudulent websites.

a) Architecture Components

The proposed architecture consists of the following main components:

i). **Data Collection Module:** This module gathers datasets containing information about phishing and legitimate websites. The dataset typically includes features such as URL length, domain age, use of HTTPS protocol, number of external links, and presence of suspicious characters in URLs.
The collected dataset serves as the input for further processing and analysis.

ii). **Feature Extraction Module:** In this stage, relevant features are extracted from the dataset. Feature extraction focuses on identifying important attributes that can help distinguish phishing websites from legitimate ones.
These features are generally categorized into:

- URL based features
- Domain based features
- Webpage based features

Extracting meaningful features improves the effectiveness of machine learning models.

iii). **Data Preprocessing Module:** The preprocessing module prepares the dataset for machine learning analysis. It removes duplicate records, handles missing values, and converts data into a structured format suitable for machine learning algorithms.
Data normalization and transformation techniques may also be applied to ensure consistency among feature values.

iv). ***Machine Learning Analysis Module:*** In this module, different machine learning algorithms are applied to analyse the dataset and classify websites.
Algorithms commonly used in phishing detection include:

- Logistic Regression
- Support Vector Machine (SVM)
- Random Forest
- K-Nearest Neighbour (KNN)

Each algorithm analyses the patterns present in the dataset and predicts whether a website is phishing or legitimate.

v). **Result Evaluation Module:** The final module evaluates the performance of machine learning algorithms using evaluation metrics such as accuracy, precision, recall, and F1 score.
These metrics help determine how effectively the algorithms identify phishing websites and reduce false predictions.

b) System Architecture Diagram

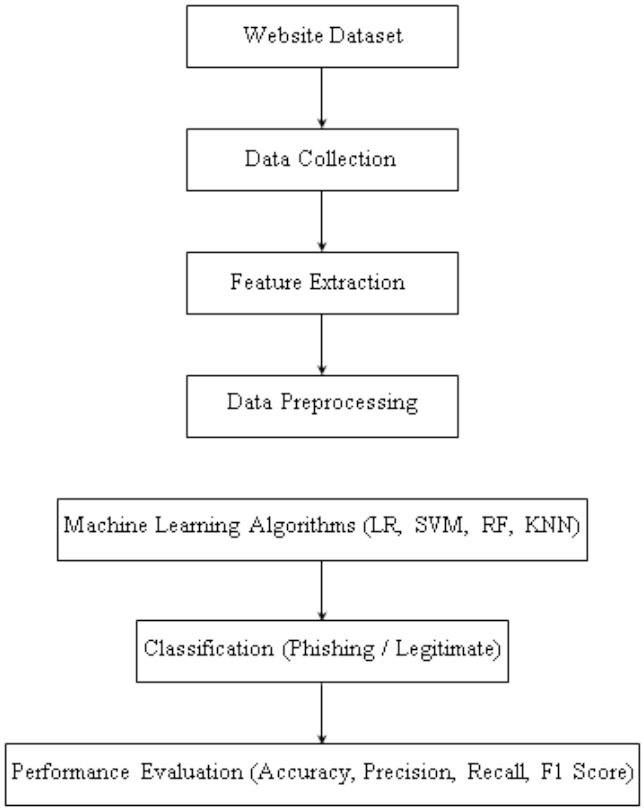

Fig 2: System Architecture for Phishing Website Detection

Results and Discussion

The results and discussion section presents the analysis of different machine learning algorithms used for phishing website detection. The purpose of this analysis is to understand how effectively these algorithms can classify websites as phishing or legitimate based on the extracted features.

Different machine learning algorithms have been widely used by researchers for phishing detection. In this study, the performance of algorithms such as Logistic Regression, Support Vector Machine (SVM), Random Forest, and K-Nearest Neighbour (KNN) is analysed based on commonly used evaluation metrics including accuracy, precision, recall, and F1 score.

The results obtained from various studies indicate that machine learning techniques significantly improve the detection of phishing websites compared to traditional rule-based approaches. By analysing patterns in URL structure, domain information, and webpage behaviour, machine learning models can effectively identify malicious websites.

a) Performance Comparison of Algorithms

The performance of different machine learning algorithms can be compared using evaluation metrics. Table I presents a comparison of commonly used algorithms for phishing detection.

Table 1: Performance comparison of machine learning algorithms

Algorithm	Accuracy (%)	Precision	Recall	F1 Score
Logistic Regression	90	0.89	0.88	0.88
Support Vector Machine	93	0.92	0.91	0.91
Random Forest	96	0.95	0.94	0.94
K-Nearest Neighbour	88	0.87	0.86	0.86

From Table I, it can be observed that Random Forest provides the highest accuracy among the analysed algorithms. This is because ensemble methods combine multiple decision trees, which improves classification performance and reduces overfitting.

Support Vector Machine also performs well due to its ability to handle high-dimensional data and identify optimal decision boundaries. Logistic Regression provides reasonable performance but may struggle with complex feature relationships. K-Nearest Neighbour is simple and effective but can become computationally expensive for large datasets.

b) Graphical Representation of Results

To better understand the performance of different machine learning algorithms, the results can also be represented graphically.

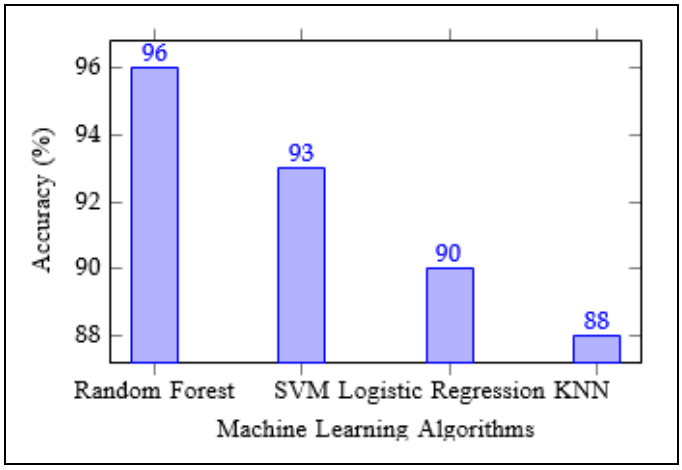

Fig 3: Accuracy Comparison of Machine Learning Algorithms for Phishing Detection

The graph clearly shows that Random Forest achieves the highest classification accuracy, followed by Support Vector Machine.

c) Discussion

The analysis of machine learning algorithms shows that intelligent models significantly improve phishing website detection compared to traditional security mechanisms. Traditional approaches such as blacklist-based detection rely on previously identified phishing websites, making them ineffective against newly created phishing attacks.

Machine learning models overcome this limitation by analysing patterns in website features and identifying suspicious behaviour automatically. Features such as abnormal URL structures, suspicious domain registration patterns, and unusual webpage behaviour play an important role in detecting phishing websites.

Among the analysed algorithms, Random Forest demonstrates the best performance because it combines multiple decision trees and reduces overfitting. Support Vector Machine also provides strong classification performance due to its ability to create optimal decision boundaries in high-dimensional feature spaces.

Overall, the results highlight the effectiveness of machine learning techniques in improving phishing detection and enhancing cyber security systems.

7. Advantages and Limitations

The use of machine learning techniques for phishing website detection provides several advantages over traditional security

approaches. At the same time, certain limitations still exist that need to be addressed for improving the effectiveness of phishing detection systems.

A). Advantages

- One of the major advantages of machine learning based phishing detection is its ability to identify patterns in large datasets. Unlike traditional blacklist-based systems that rely on previously reported phishing websites, machine learning models can analyse various website features and detect suspicious behaviour even in newly created phishing websites.
- Another advantage is the ability of machine learning algorithms to improve detection accuracy. By analysing features such as URL structure, domain information, and webpage behaviour, these models can effectively distinguish between legitimate and phishing websites.
- Machine learning techniques also provide flexibility because they can be trained using different datasets and adapted to changing phishing techniques. As cyber attackers continuously develop new phishing strategies, machine learning models can be retrained to improve detection capabilities.
- In addition, machine learning models such as Random Forest and Support Vector Machine can handle high-dimensional datasets and complex relationships between features, which enhances the effectiveness of phishing detection systems.

B). Limitations

- One of the main challenges is the dependency on large and high-quality datasets. If the dataset used for training is incomplete or contains biased samples, the performance of the model may decrease.
- Another limitation is the possibility of false positives and false negatives. False positives occur when legitimate websites are incorrectly classified as phishing, while false negatives occur when phishing websites are mistakenly classified as legitimate.
- Machine learning models also require computational resources for training and analysis, especially when dealing with large datasets. Some algorithms may take longer processing time, which can affect real-time phishing detection systems.
- Additionally, cyber attackers continuously evolve their phishing techniques, which means detection models must be regularly updated to remain effective.

8. Conclusion and Future Work

Phishing attacks continue to be one of the most significant cyber security threats affecting internet users worldwide. These attacks attempt to deceive users by creating fraudulent websites that closely resemble legitimate platforms in order to steal sensitive information such as login credentials, financial details, and personal data.

This research analysed the use of machine learning techniques for phishing website detection. Various machine learning algorithms including Logistic Regression, Support Vector Machine, Random Forest, and K-Nearest Neighbour were studied to understand their effectiveness in identifying phishing websites based on different website features.

The study highlighted that machine learning techniques provide improved detection capabilities compared to traditional rule-based or blacklist-based security mechanisms. By analysing features such as URL characteristics, domain information, and webpage behaviour, machine learning models can successfully identify suspicious patterns associated with phishing websites.

Among the analysed algorithms, Random Forest and Support Vector Machine demonstrated higher performance in phishing detection due to their ability to handle complex datasets and identify hidden patterns.

Despite these advancements, challenges such as dataset quality, evolving phishing techniques, and computational requirements still exist. Addressing these challenges is essential for developing more reliable phishing detection systems.

A). Future Work

Future research in phishing detection can focus on improving detection accuracy by combining multiple machine learning algorithms and advanced techniques such as deep learning. Hybrid models that integrate feature-based detection with behavioural analysis may provide better protection against sophisticated phishing attacks.

Researchers can also explore real-time phishing detection systems that analyse website behaviour dynamically while users interact with webpages. In addition, integrating phishing detection mechanisms with browser security systems and email filtering tools can further enhance user protection against cyber threats.

Overall, machine learning based phishing detection systems play a crucial role in strengthening cyber security and protecting users from online fraud and malicious activities.

References

1. Mohamed. AI/ML Revolutionising Cybersecurity. 2024.
2. Mandora S, Mehta R, Mehta A. AI in Phishing Detection. 2024.
3. Singh R, Maurya S. ML Techniques for Phishing Detection. 2024.
4. Yadav R, *et al*. AI-driven Phishing Detection and Prevention Model. 2025.
5. WJAETS. Advancing Cybersecurity with AI/ML. 2025.
6. AI and Cyber Security: Enhancing Threat Detection. 2024.
7. Kumar, *et al*. AI and Machine Learning in Cyber Threat Detection. 2023–2024.
8. Ajayi, *et al*. AI/ML for Cybersecurity in Critical Infrastructure. 2025.
9. Krishna PG, Rajesh DS, Kumar KP. Phishing Detection: Machine Learning Implementation. *International Journal of Research in Engineering, Science and Management*. 2021;4(7):175–177.
10. Al-qasmi, Al-anazi A, Al-shehri L, Alshaman S, Al-atawi W, Abbass O. Machine Learning-Based Phishing Detection System. *International Journal of Intelligent Systems and Applications in Engineering*. 2024;12(4):4367–4372.
11. Almujahid NF, Haq MA, Alshehri M. Comparative Evaluation of Machine Learning Algorithms for Phishing Site Detection. *PeerJ Computer Science*. 2024.
12. Lokesh GH, BoreGowda G. Phishing Website Detection Based on Effective Machine Learning Approach. *Journal of Cyber Security Technology*. 2021;5(1):1–14.
13. Saxena, Degadwala S, Joshi M. Phishing URL Detection Using Machine Learning. *International Journal of Scientific Research in Science and Technology*. 2026;13(1):19–25.
14. Sharma. Machine Learning Approach for Phishing

Website Detection: A Literature Survey. *Journal of Discrete Mathematical Sciences and Cryptography*. 2022;25(3):817–827.
15. Sanjay Kumar G, Krishna G, R. G, Martin MA. Detection of Phishing Using Machine Learning Algorithms. *International Journal of Computational Learning & Intelligence*. 2022;1(1):31–36.
16. Pal R, Pandey MK, Pal S, Yadav DC. Phishing Detection: A Hybrid Model with Feature Selection and Machine Learning Techniques. *International Journal of Experimental Research and Review*. 2023;36:99–108.
17. Jain K, Gupta BB. Machine Learning Based Phishing Detection from URLs. *Expert Systems with Applications*. 2019;117:345–357.
18. Yerima SY, Alzaylaee MK. High Accuracy Phishing Detection Based on Convolutional Neural Networks. 2020.
19. Aslam S, Aslam H, Manzoor A, Hui C, Rasool A. AntiPhishStack: LSTM-Based Stacked Generalization Model for Optimized Phishing URL Detection. 2024.
20. Dubey R, Tripathi AM, Srivastava A, Singh S. Phishing Detection System: An Ensemble Approach Using Character-Level CNN and Feature Engineering. 2025.