INTERNATIONAL JOURNAL OF RESEARCH, EXPLORATION & KNOWLEDGE (IJREK)
Official Scholarly Journal Portal: **https://www.ijrek.org** | DOI: **https://doi.org/10.5281/zenodo.23857410**

AI BASED PHISHING WEBSITE DETECTION USING MACHINE LEARNING

Mrs.Dhage T.S[1], Ashwini Jagdale[2]

[1] Faculty Guide & Supervisor | [2] Dattakala Group Of Institutions, India

Corresponding author: **ashwinijagdale2006@gmail.com**

ARTICLE ID	ARTICLE TYPE	RESEARCH AREA	PUBLICATION
IJREK-2026-00184	Research Article	Machine Learning & Deep Learning	Volume 1, Issue Issue 1 (2026)
RECEIVED	**REVISED**	**ACCEPTED**	**PUBLISHED**
31 October 2026	02 November 2026	31 October 2026	02 November 2026

Abstract

Phishing websites are a common cybersecurity threat that trick users by pretending to be legitimate websites and attempting to obtain sensitive information. Traditional methods, such as blocklists and rule-based techniques, may not always detect newly created or modified phishing websites. Machine learning provides an effective approach to phishing detection by learning patterns from examples of legitimate and phishing websites. This paper presents a conceptual framework for AI-based phishing website detection using machine learning. The framework includes collecting phishing and legitimate website data, extracting important URL and website features, preprocessing the data, selecting relevant features, applying machine learning classification, and evaluating the detection performance. Recent research on feature engineering, feature selection, machine learning, deep learning, XGBoost-based detection, and phishing campaign detection is also reviewed. The proposed framework aims to provide a simple and structured approach for distinguishing phishing websites from legitimate websites while considering the changing nature of phishing attacks. Since this study is research and visualization oriented, the proposed system is presented as a conceptual framework rather than a developed application or trained software model.

Keywords: *Phishing Website Detection; Machine Learning; Artificial Intelligence; Cybersecurity; Feature Engineering; Feature Selection; URL Analysis; Classification*

INTRODUCTION

The rapid growth of internet-based services has made websites an essential part of everyday activities such as online banking, shopping, education, communication, and digital transactions. Along with these benefits, the increasing dependence on online services has also resulted in the growth of cybersecurity threats. One of the common threats is phishing, in which attackers create fraudulent websites or links that imitate legitimate services to deceive users. Phishing websites are designed to appear similar to trusted websites. Attackers may use misleading domain names, suspicious URLs, copied webpage designs, redirects, and other techniques to make fraudulent websites appear legitimate. When users interact with such websites, they may unintentionally provide sensitive information. Therefore, detecting phishing websites before users interact with them is an important cybersecurity requirement. Traditional phishing detection methods commonly use blacklists, manually defined rules, and previously identified malicious URLs. These approaches can be useful for detecting known threats, but they may face difficulties when attackers create new phishing websites or frequently modify existing URLs. This creates a need for detection techniques that can identify previously unseen patterns. Artificial Intelligence and Machine Learning provide an alternative approach to this problem. Machine learning models can learn patterns from labelled examples of legitimate and phishing

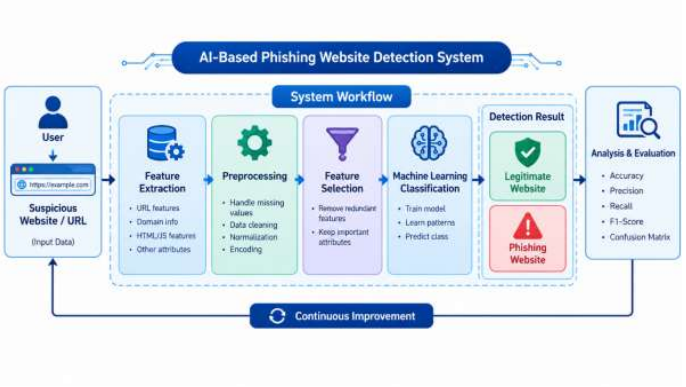


Fig. 1: Conceptual Overview of AI-Based Phishing Website Detection

websites. Various characteristics related to URLs, domains, and webpage behaviour can be converted into features and provided to a classification model. The model can then learn to distinguish between legitimate and phishing websites. Recent research has explored different approaches for phishing detection, including feature engineering, feature selection, traditional machine learning, deep learning, and tree-based models such as XGBoost. Research has also investigated the problem of generalization, where a model trained on one dataset may perform differently when tested on another dataset. These studies demonstrate the importance of selecting informative features and evaluating detection approaches using reliable and representative datasets. This paper focuses on AI-based phishing website detection using machine learning. The study reviews recent research in this area and presents a conceptual framework involving dataset preparation, feature extraction, preprocessing, feature selection, machine learning classification, and performance analysis. The objective is to provide a clear understanding of how machine learning techniques can be applied to phishing website detection while considering the changing nature of phishing attacks. The proposed work is presented through research analysis and system visualization without developing a separate website, application, or deployed machine learning model.

II. LITERATURE REVIEW

Recent research has explored several machine learning and artificial intelligence techniques for detecting phishing websites and malicious URLs. The existing studies focus on different aspects such as feature engineering, feature selection, classification algorithms, deep learning, crossdataset validation, and campaign-level phishing analysis. These studies provide the foundation for developing a conceptual framework for AI-based phishing website detection. A systematic review on feature engineering for phishing website detection examined the importance of selecting and designing appropriate features for machine learning-based phishing detection [1]. The study highlights that the quality and relevance of features have a significant influence on the performance of phishing detection systems. URLrelated characteristics and other website properties can provide useful information for distinguishing between legitimate and phishing websites. The review also indicates that feature engineering is an important stage in developing effective phishing detection approaches. Duarte et al. investigated the early detection of phishing URLs in parked domains using machine learning techniques [2]. The study considered a large dataset containing 211,659 URLs and evaluated several machine learning algorithms, including Logistic Regression, Random Forest, XGBoost, LightGBM, and Gradient Boosting. The results demonstrated the potential of machine learning for identifying phishing URLs at an early stage. However, the study was based on data associated with a particular financial institution, which may limit the generalization of the approach to other Another study presented a machine learning approach combined with feature selection and deep learning models for phishing detection [3]. The research used a dataset containing 58,645 URLs and 111 features. Feature selection was applied to reduce the number of features and identify the most informative characteristics for detection. The study demonstrated that feature selection can simplify the input space while maintaining useful detection capability. This work is particularly relevant to the proposed framework because it emphasizes the importance of preprocessing and selecting meaningful features before classification. Misiek and Hyla proposed an XGBoost-based URL phishing detection method with cross-dataset validation [4]. The study investigated the performance of phishing detection across multiple datasets and selected a smaller set of stable features that could provide better cross-dataset generalization. The research demonstrated that a model may perform well on its training or source dataset but experience reduced performance when evaluated on a different dataset. This highlights the importance of generalization and robust evaluation in phishing detection systems. PhishLumos introduced a broader approach for phishing mitigation by moving from individual URL detection toward campaign-level analysis [5]. The approach considers relationships among URLs, domains, IP addresses, certificates, and historical information to identify related phishing activities. The study demonstrates that phishing attacks can occur as part of larger campaigns rather than as isolated websites. Although this approach provides a broader cybersecurity perspective, it involves a more complex infrastructure and is beyond the scope of the conceptual machine learning framework presented in this paper. Overall, the reviewed studies show that phishing detection can benefit from appropriate feature engineering, feature selection, machine learning classification, and robust evaluation. They also indicate that dataset variation and evolving phishing techniques remain important challenges. Based on these observations, this paper presents a conceptual AI-based phishing website detection framework that combines data preprocessing, feature extraction, and visualization-based analysis.

III. PROBLEM STATEMENT

Phishing websites continue to pose a significant cybersecurity risk because they are designed to imitate legitimate websites and deceive users. Attackers can frequently modify URLs, domain names, webpage structures, and other characteristics, making it difficult for traditional detection methods to identify newly created phishing websites. Conventional approaches such as blacklists and manually defined rules mainly depend on previously known phishing information. As phishing attacks continuously evolve, these approaches may not be sufficient for identifying previously unseen or modified phishing websites. Furthermore, the large number of website-related characteristics makes it important to identify the most relevant features for effective detection. Machine learning provides a promising approach by learning patterns from legitimate and phishing website data. However, an effective phishing detection framework requires suitable feature extraction, data preprocessing, feature selection, classification, and proper evaluation. Another important challenge is generalization, as detection performance may vary when the data distribution changes. Therefore, there is a need for a structured AI-based phishing website detection framework that represents the complete detection process from data collection to classification and analysis. The proposed work focuses on developing such a conceptual framework using machine learning techniques and visualizing its working process without developing a separate website or deployed software model.

IV. PROPOSED SYSTEM

The proposed system presents a conceptual machine learning framework for detecting phishing websites based on relevant website and URL characteristics. The framework is designed to analyse available website data, identify informative features, and classify websites as either legitimate or phishing. Core Components: Dataset Collection – A suitable dataset containing examples of legitimate and phishing websites is considered as the input for the proposed framework. Feature Extraction – Relevant characteristics are extracted from URLs and website information. These may include URL length, number of special characters, presence of suspicious patterns, domainrelated characteristics, HTTPS usage, and other relevant attributes. Data Preprocessing – The collected data is processed to handle missing values, inconsistent information, and unsuitable data formats so that it can be used for machine learning-based analysis. Feature Selection – Important features are selected from the available attributes to reduce unnecessary information and improve the efficiency of the detection process. Machine Learning Classification – Machine learning classification techniques are considered to learn patterns from the selected features and distinguish phishing websites from legitimate websites. Detection Output – Based on the learned patterns, the framework provides a classification indicating whether the analysed website is legitimate or potentially phishing. Process Flow: 1. Input: Website and URL-related data is collected from a suitable dataset. 2. Feature Extraction: Relevant characteristics are identified from the available data. 3. Preprocessing: The data is cleaned and prepared for analysis. 4. Feature Selection: Important features are selected for classification. 5. Classification: Machine learning-based classification is applied conceptually. 6. Output: The website is classified as legitimate or phishing. 7. Analysis: The detection framework is analysed using suitable evaluation measures discussed in existing research.



Fig. 2: Proposed Workflow for Phishing Website Detection

V. ARCHITECTURAL DIAGRAM

The architecture of the proposed AI-based phishing website detection framework consists of a sequence of stages that transform website and URL-related data into a phishing detection decision. The framework is designed to represent the complete detection process from input data to final classification. The architecture consists of five primary stages: Data Input Layer: This layer represents the collection of legitimate and phishing website data from a suitable dataset. The collected data contains URL and website-related characteristics required for further analysis. Feature Extraction Layer: In this stage, relevant characteristics are identified from the available website information. Examples include URL length, special characters, domain-related characteristics, HTTPS usage, and other URL or website properties. Preprocessing and Feature Selection Layer: The extracted data is cleaned and prepared for analysis. Relevant features are then selected to remove unnecessary information and provide a suitable feature set for classification. Machine Learning Classification Layer: The selected features are provided to a machine learning classification approach. The classifier learns patterns associated with legitimate and phishing websites and performs conceptual classification of the input data. Output and Analysis Layer: The final output represents the classification of a website as

legitimate or phishing. The framework also supports analysis using evaluation measures such as matrix as described in existing research. The overall flow of the proposed framework can therefore be represented as: Data Input → Feature Extraction → Data Preprocessing → Feature Selection → Machine Learning Classification → Detection Output → Analysis

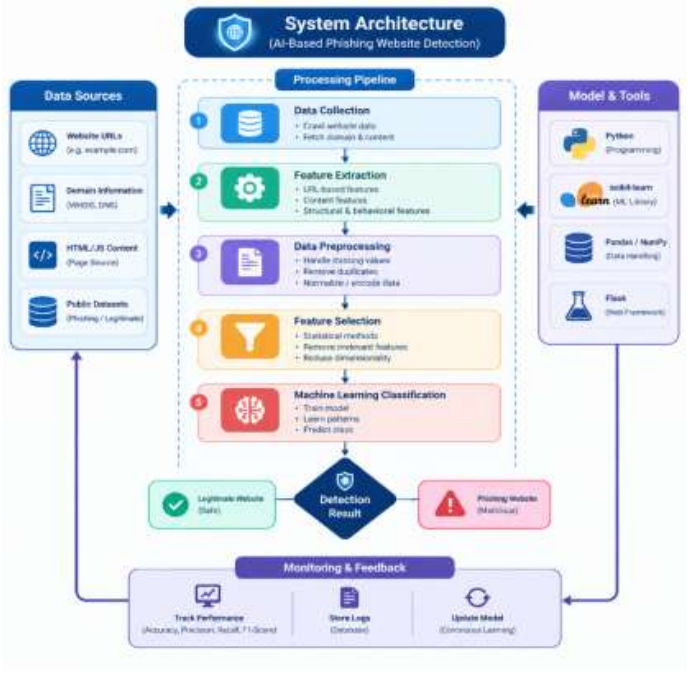


Fig. 3: System Architecture of the Proposed AI-Based Phishing Website Detection Framework

VI. IMPLEMENTATION

The proposed phishing website detection framework is presented as a conceptual implementation workflow rather than as a deployed website or software application. The implementation process is divided into a sequence of stages that describe how machine learning can be applied to phishing website detection. Phase 1 – Dataset Preparation: A suitable dataset containing legitimate and phishing website samples is considered for analysis. Each sample contains URL or website-related characteristics that can be used to distinguish between the two classes. Phase 2 – Feature Extraction: Relevant features are identified from the available website and URL information. These features may include URL length, number of special characters, suspicious URL patterns, domain-related characteristics, HTTPS usage, and other attributes reported in existing phishing detection research. Phase 3 – Data Preprocessing: The collected data is prepared for machine learning analysis. This stage includes handling unsuitable or missing values, organizing the feature representation, and preparing the data for subsequent feature selection and classification. Phase 4 – Feature Selection: The available features are analysed to identify the most informative characteristics for phishing detection. Feature selection can reduce unnecessary information and simplify the classification process. Phase 5 – Machine Learning Classification: The selected features are considered as input to a machine learning classification approach. The classification stage learns patterns associated with legitimate and phishing website samples and produces a classification decision. Phase 6 – Detection and Analysis: The output of the classification stage represents either a legitimate or phishing website. The proposed framework can be analysed using commonly used evaluation measures such as matrix. Since this research is visualization-oriented, the above phases describe the conceptual workflow of the proposed system. No separate website, deployed application, or independently trained machine learning model is developed as part of this work.



Fig. 4: Conceptual Implementation Process for Phishing Website Detection

VII. CONCLUSION

Phishing websites are an important cybersecurity threat because they can imitate legitimate websites and deceive users into providing sensitive information. Traditional detection techniques such as blacklists and rule-based methods may face limitations when dealing with newly created or modified phishing websites. This paper presented a conceptual framework for AI-based phishing website detection using machine learning. The proposed framework includes dataset preparation, feature extraction, data preprocessing, and detection analysis. Recent research shows that appropriate feature engineering and feature selection can support effective phishing detection, while robust evaluation is important for understanding model performance and generalization. The proposed work focuses on understanding and visualizing the complete phishing detection workflow rather than developing a deployed application or independently trained machine learning model. The framework provides a structured foundation for applying machine learning techniques to phishing website detection and highlights the importance of handling evolving phishing attacks and changing data distributions.

VIII. FUTURE SCOPE

The proposed conceptual framework provides a foundation for further development of AI-based phishing website detection systems. Future work can focus on developing and training machine learning and deep learning models using larger and more diverse datasets containing continuously updated phishing and legitimate website samples. Future research can also explore advanced URL and webpage features, real-time phishing detection, and improved feature selection techniques. Crossdataset evaluation can be used to study the generalization of detection models across different Since phishing techniques continuously evolve, adaptive learning approaches can also be investigated to identify newly emerging phishing patterns. In addition, future systems may integrate browserbased detection, real-time warning mechanisms, and automated threat intelligence sources. Advanced approaches such as ensemble learning, deep learning, and campaign-level phishing analysis can further improve the ability to detect complex and evolving phishing attacks. These developments can help in building more reliable and practical phishing website detection systems.

REFERENCES

[1] K. Kustiawan and A. Ghauth, “Feature Engineering for Phishing Website Detection Using Machine Learning: A Systematic Review,” IEEE Access, vol. 13, pp. 192080–192104, 2025, doi: 10.1109/ACCESS.2025.3630334.

[2] J. D. Duarte, P. Chagas Junior, J. P. J. da Costa, E. J. da Costa, et al., “Machine Learning for Early Detection of Phishing URLs in Parked Domains: An Approach Applied to a Financial Institution,” IEEE Access, vol. 13, pp. 145736–145753, 2025, doi: 10.1109/ACCESS.2025.3599454.

[3] G. S. Nayak, B. Muniyal, and M. C. Belavagi, “Enhancing Phishing Detection: A Machine Learning Approach With Feature Selection and Deep Learning Models,” IEEE Access, vol. 13, pp. 33308–33320, 2025, doi: 10.1109/ACCESS.2025.3543738.

[4] M. Misiek and T. Hyla, “XGBoost-Based URL Phishing Detection Method With Cross-Dataset Validation,” IEEE Access, vol. 14, pp. 43200– 43212, 2026, doi: 10.1109/ACCESS.2026.3672690.

[5] D. Chiba, H. Nakano, and T. Koide, “PhishLumos: From a Single URL to CampaignLevel Phishing Mitigation,” IEEE Access, vol. 14, pp. 79716–79739, 2026, doi: 10.1109/ACCESS.2026.3696597.

Publication archival note: Official scholarly contribution published under CC BY 4.0 International License. Permanent record on **https://www.ijrek.org** | Article ID: IJREK-2026-00184