AI-Powered URL Security and Phishing Detection System

Akileshwari A.[1]; Akash A.[2]; Eswar M. S.[3]; Abishek G. P.[4]

[1] Assistant Professor, Dept. of Computer Science Engineering, K.L.N College of Engineering, Tamil Nadu, India
[2,3,4] Student, Dept. of Computer Science Engineering, K.L.N College of Engineering, Tamil Nadu, India

Publication Date: 2026/03/28

Abstract: Phishing attacks are among the most prevalent cybersecurity threats, exploiting users through deceptive URLs and malicious websites to steal sensitive information such as login credentials, financial data, and personal details. This paper proposes an AI-powered URL security and phishing detection system that analyzes URLs using machine learning techniques and feature-based evaluation. The system extracts critical attributes such as HTTPS usage, SSL certificate validity, domain characteristics, URL structure, and WHOIS information. A trained machine learning model processes these features to classify URLs as legitimate or phishing. The application is implemented as a web-based system using Flask, providing real-time analysis and user-friendly interaction. Experimental results demonstrate improved accuracy and efficiency in detecting phishing URLs, making the system a reliable tool for enhancing online security and safe browsing practices.

Keywords: *Phishing Detection; URL Security; Machine Learning; Cybersecurity; Flask; SSL Certificate; WHOIS.*

How to Cite: Akileshwari A.; Akash A.; Eswar M. S.; Abishek G. P. (2026) AI-Powered URL Security and Phishing Detection System. *International Journal of Innovative Science and Research Technology*, 11(3), 2440-2445. https://doi.org/10.38124/ijisrt/26mar1467

I. INTRODUCTION

The rapid expansion of internet services has significantly increased the dependency of users on online platforms for communication, banking, shopping, and information exchange. However, this growth has also led to a rise in cyber threats, particularly phishing attacks. Phishing is a fraudulent practice in which attackers create deceptive websites or URLs that mimic legitimate ones to trick users into revealing sensitive information such as passwords, credit card details, and personal data. These attacks are becoming increasingly sophisticated, making it difficult for traditional detection techniques to identify them effectively.

Conventional approaches such as blacklist-based detection and rule-based filtering are limited in their ability to detect newly generated phishing URLs. As attackers continuously evolve their strategies, there is a need for intelligent systems that can analyze and identify malicious patterns dynamically. In this context, the integration of artificial intelligence and machine learning provides a powerful solution for detecting phishing attempts in real time.

➢ *Objective and Scope of the Project*

The primary objective of this project is to develop an AI-powered system capable of detecting phishing URLs with high accuracy using machine learning techniques. The system aims to analyze various URL features such as security protocols, domain characteristics, and structural patterns to classify URLs as safe or malicious. Another important objective is to design a user-friendly web interface that allows users to easily input URLs and obtain real-time results. Additionally, the project seeks to improve cybersecurity awareness by providing detailed insights into the analyzed URL, enabling users to understand potential risks and make informed decisions while browsing.

The scope of this project is focused on the development of a web-based phishing detection system that operates using URL-based analysis. The system incorporates feature extraction methods such as HTTPS verification, SSL certificate validation, domain age analysis through WHOIS data, and URL structure examination. It is designed to provide real-time detection and enhance user safety during online activities. However, the system is limited to URL-level analysis and does not extend to email-based phishing detection, network traffic monitoring, or advanced behavioral analysis. Future enhancements can expand the scope to include these additional security layers.

II. SYSTEM MODULES

The proposed system is composed of multiple interconnected modules that collectively perform phishing detection. The first module focuses on user interaction, where the user provides a URL through a web interface. This module ensures ease of use and accessibility by allowing users to quickly submit URLs and view results.

- *User Interface Module*

The user interface module serves as the entry point of the system, allowing users to interact with the application through a web-based platform. It provides a simple and intuitive interface where users can input URLs for analysis. The module is designed to ensure ease of use, responsiveness, and clear visualization of results. It also displays the final output in an understandable format, including the classification of the URL and additional security details.

- *URL Normalization Module*

The template is used to format your paper and style the text. All margins, column widths, line spaces, and text fonts are prescribed; please do not alter them. You may note peculiarities. For example, the head margin in this template measures proportionately more than is customary. This measurement and others are deliberate, using specifications that anticipate your paper as one part of the entire proceedings, and not as an independent document. Please do not revise any of the current designations.

- *Feature Extraction Module*

The feature extraction module is one of the most critical components of the system. It analyzes the normalized URL and extracts various attributes that help in identifying phishing behavior. These features include the presence of HTTPS, SSL certificate validity, domain age obtained through WHOIS data, URL length, number of subdomains, and the presence of suspicious symbols or patterns. The extracted features are then prepared as input for the machine learning model.

- *External and Internal Threat Analysis Module*

This module enhances the system by analyzing both external and internal threat factors associated with the URL. External threat analysis involves checking the URL against online threat intelligence sources, examining domain reputation, and verifying SSL and WHOIS data. Internal threat analysis focuses on identifying suspicious patterns within the URL itself, such as abnormal structure, excessive subdomains, and hidden malicious indicators. By combining both analyses, the system improves its ability to detect sophisticated phishing attacks.

- *Machine Learning Prediction Module*

This module is responsible for classifying the URL as either legitimate or phishing. It uses a trained machine learning model that has learned patterns from a dataset of known phishing and safe URLs. Based on the extracted features, the model predicts the nature of the input URL with high accuracy. The use of machine learning enables the system to adapt to new and evolving phishing techniques.

- *Data Integration Module*

The data integration module connects the system with external data sources to gather additional information about the URL. It retrieves WHOIS details such as domain registration date and ownership information, as well as SSL certificate data to verify the security of the website. This real-time data enhances the reliability and effectiveness of the phishing detection process.

- *Result Visualization Module*

The result visualization module presents the final output to the user in a clear and informative manner. It displays whether the URL is safe or phishing along with supporting details such as SSL status, domain information, and risk level. This module helps users understand the reasoning behind the classification and enables them to make informed decisions while browsing.

III. LITERATURE REVIEW

- *Ume Zara, Kashif Ayyub, Hikmat Ullah Khan, Ali Daud, Tariq Alsahfi, Saima Gulzar Ahmad, "Phishing Website Detection Using Deep Learning Models," IEEE Access, Vol. 12(2), October 2024.*

This study highlights the use of deep learning models for detecting phishing websites by learning complex patterns from large datasets. It improves detection accuracy compared to traditional methods but requires high computational resources and large training data.

- *A. K. Jain and B. B. Gupta, "PHISH-SAFE: URL Features-Based Phishing Detection System Using Machine Learning," Springer, Vol. 729, April 2018.*

This research focuses on a machine learning approach using URL-based features to classify websites. It shows good detection accuracy but requires frequent updates to handle newly emerging phishing attacks.

- *J. K. Lee, Y. Chang, H. Y. Kwon, and B. Kim, "Reconciliation of Privacy with Preventive Cybersecurity: The Bright Internet Approach," Information Systems Frontiers, Vol. 22, No. 1, February 2020.*

This paper discusses secure and privacy-preserving cybersecurity systems. It emphasizes detecting threats like phishing while protecting user data, though implementation is complex.

- *S. Asiri, Y. Xiao, S. Alzahrani, S. Li, and T. Li, "A Survey of Intelligent Detection Designs of HTML URL Phishing Attacks," IEEE Access, Vol. 11, January 2023.*

This survey reviews various phishing detection methods including machine learning and heuristic approaches. It concludes that hybrid techniques provide better accuracy and reliability.

- *N. A. Azeez, S. Misra, I. A. Margaret, L. Fernandez-Sanz, and S. M. Abdulhamid, "Adopting Automated Whitelist Approach for Detecting Phishing Attacks," Computers & Security, Vol. 108, September 2021.*

This study explains the whitelist-based detection method, which is fast and reduces false positives but fails to detect new phishing websites effectively.

IV. EXISTING SYSTEM

Existing phishing detection systems primarily rely on blacklist-based methods and traditional rule-based techniques. In blacklist systems, URLs are compared against a database of known malicious websites. While this approach is simple and efficient, it is ineffective against newly created phishing URLs that are not yet listed in the database.

Rule-based systems analyze predefined characteristics such as URL length, presence of special characters, and domain structure. Although these systems can detect certain types of phishing attacks, they lack adaptability and may produce false positives or false negatives.

Furthermore, many existing systems do not provide real-time analysis or detailed insights into the detected threats. This limits their effectiveness in dynamic environments where phishing techniques are constantly evolving. As a result, there is a need for more advanced and intelligent systems that can overcome these limitations.

V. PROPOSED SYSTEM

The proposed system introduces an AI-powered approach to phishing detection by combining machine learning with feature-based analysis. Unlike traditional systems, this approach does not rely solely on static databases or predefined rules. Instead, it dynamically analyzes URLs and identifies patterns associated with phishing attacks.

The system extracts multiple features from the input URL and processes them using a trained machine learning model. By incorporating real-time data such as SSL certificate status and WHOIS information, the system enhances its ability to detect malicious URLs accurately.

The proposed solution also emphasizes user experience by providing a web-based interface that delivers instant results along with detailed explanations. This not only improves detection accuracy but also increases user awareness of cybersecurity threats.

VI. SYSTEM ARCHITECTURE

The system architecture consists of a structured workflow that connects the frontend interface with backend processing components. The process begins with the user entering a URL into the web interface, which sends the input to the backend server developed using Flask.

The backend first performs URL normalization to ensure consistency, followed by feature extraction where relevant attributes are identified. These features are then passed to the machine learning model, which analyzes the data and generates a prediction.

Simultaneously, the system retrieves additional information from external services such as WHOIS databases and SSL verification tools. This information is combined with the model's output to produce a comprehensive analysis.

Finally, the results are sent back to the frontend, where they are displayed to the user in an organized and visually appealing format. This architecture ensures efficient processing, scalability, and real-time response.

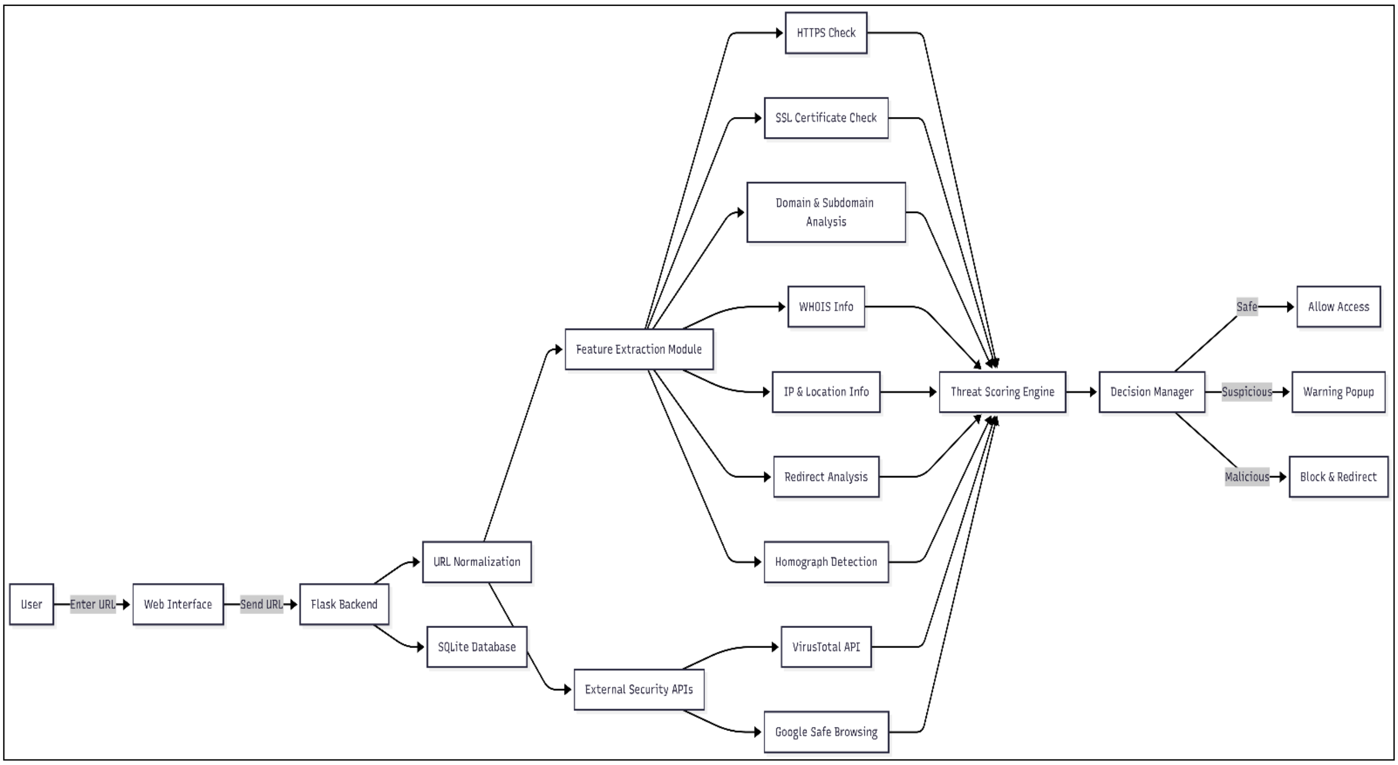
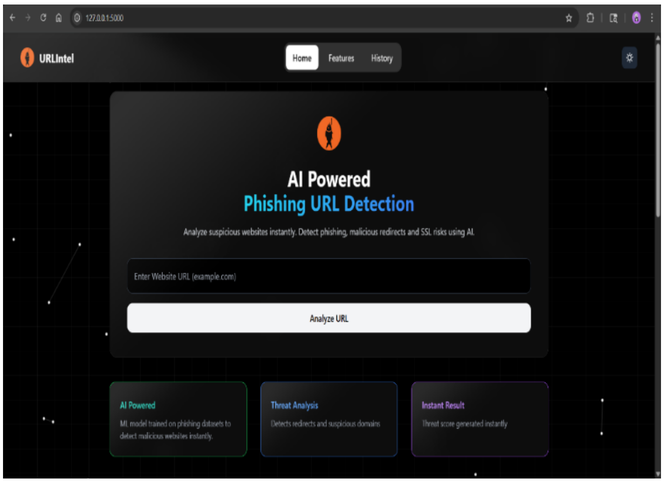


Fig - 1: Architecture Diagram

VII. RESULTS AND DISCUSSION

The developed system, AI-Powered URL Security and Phishing Detection System (URLIntel), was successfully tested with multiple types of URLs including safe, suspicious, and phishing websites. The system analyzes the given URL in real-time and generates a threat score (0–100) along with a classification: Safe, Suspicious, or Phishing.



AI Powered
Phishing URL Detection

Analyze suspicious websites instantly. Detect phishing, malicious redirects and SSL risks using AI.

Enter Website URL (example.com)

Analyze URL

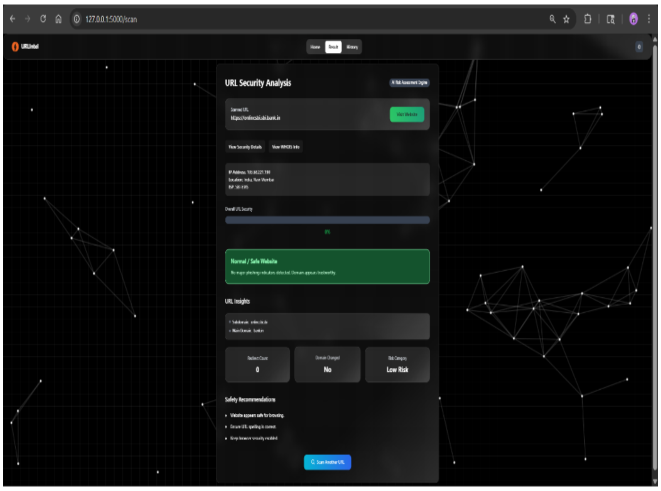
AI Powered
ML model trained on phishing datasets to detect malicious websites instantly.

Threat Analysis
Detects redirects and suspicious domains

Instant Result
Threat score generated instantly

Fig - 2: User Web Interface

The index page of the URLIntel system serves as the main entry point where users can input a website URL for analysis. It features a clean and modern dark-themed interface with a central input field labeled "Enter Website URL" and an "Analyze URL" button for initiating the scan. The page also highlights key features such as AI-powered detection, threat analysis, and instant results, helping users understand the system's capabilities. The design is user-friendly and intuitive, allowing even non-technical users to easily submit URLs and begin the security analysis process efficiently.



URL Security Analysis

Scanned URL
https://onlinesbi.sbi.bank.in

IP Address: 103.68.221.190
Location: India, Navi Mumbai
ISP: SBI EMS

Overall URL Security
0%

Normal / Safe Website
No major phishing indicators detected. Domain appears trustworthy.

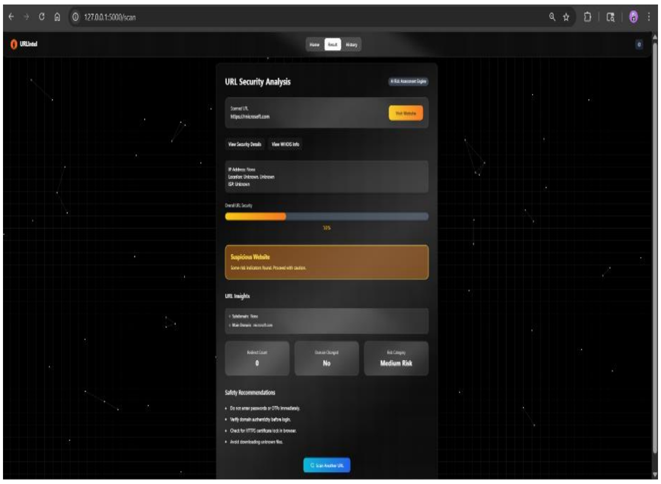
URL Insights
- Subdomain: onlinesbi.sbi
- Main Domain: bank.in

Redirect Count	Domain Changed	Risk Category
0	No	Low Risk

Safety Recommendations
- Website appears safe for browsing.
- Ensure URL spelling is correct.
- Keep browser security enabled.

Fig - 3: Safe Website Result

The result page for a safe website displays a well-structured and visually clear interface indicating a low threat score (0%) with a green progress bar, representing a secure and trustworthy domain. The system identifies no suspicious behavior such as redirects or domain mismatches, and the classification "Normal / Safe Website" is prominently shown. Additional details like IP address, location, ISP, and domain structure are provided, reinforcing transparency.



URL Security Analysis

Scanned URL
https://mlcrosoft.com

IP Address: None
Location: Unknown, Unknown
ISP: Unknown

Overall URL Security
30%

Suspicious Website
Some risk indicators found. Proceed with caution.

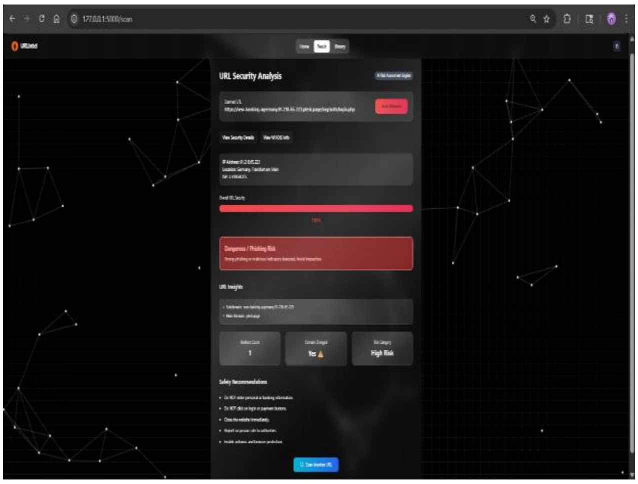
URL Insights
- Subdomain: None
- Main Domain: mlcrosoft.com

Redirect Count	Domain Changed	Risk Category
0	No	Medium Risk

Safety Recommendations
- Do not enter passwords or OTPs immediately.
- Verify domain authenticity before login.
- Check for HTTPS certificate lock in browser.
- Avoid downloading unknown files.

Fig - 4: Suspicious Website Result

In the case of a suspicious website, the result page highlights a moderate threat score (around 30%) using a yellow/orange progress bar, indicating potential risk. The system labels the site as "Suspicious Website" and identifies minor irregularities such as incomplete IP data or subtle domain issues. While no major phishing indicators are detected, the system provides cautionary recommendations, encouraging users to verify authenticity before interacting. This demonstrates the system's ability to detect early warning signs without misclassifying the site as fully malicious.



URL Security Analysis

Scanned URL
https://www.banking-agermany91-218-65-221.plesk.page/log/auth/login.php

IP Address: 91.218.65.221
Location: Germany, Frankfurt am Main
ISP: [illegible]

Overall URL Security
100%

Dangerous / Phishing Risk
Strong phishing or malicious indicators detected. Avoid interaction.

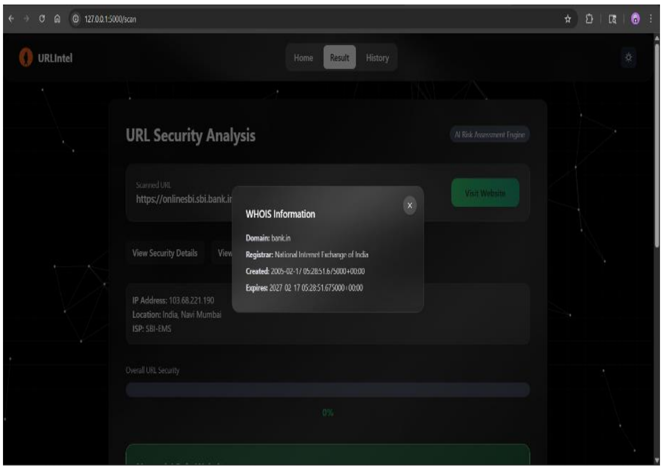
URL Insights
- Subdomain: www.banking-agermany91-218-65-221
- Main Domain: plesk.page

Redirect Count	Domain Changed	Risk Category
1	Yes ⚠	High Risk

Safety Recommendations
- Do NOT enter personal or banking information.
- Do NOT click on login or payment buttons.
- Close the website immediately.
- Report suspicious site to authorities.
- Enable antivirus and browser protection.

Fig - 5: Phishing Website Result

For a phishing or dangerous website, the result page clearly emphasizes a high threat score (100%) with a red progress bar, signaling severe risk. The classification "Dangerous / Phishing Risk" is prominently displayed, along with critical indicators such as suspicious subdomains, redirect activity, and domain mismatches. The system also provides detailed IP information and flags high-risk behavior, helping users understand the threat. Strong safety recommendations are included, advising users to avoid interaction detecting and warning against malicious websites.



URL Security Analysis

Scanned URL
https://onlinesbi.sbi.bank.ir

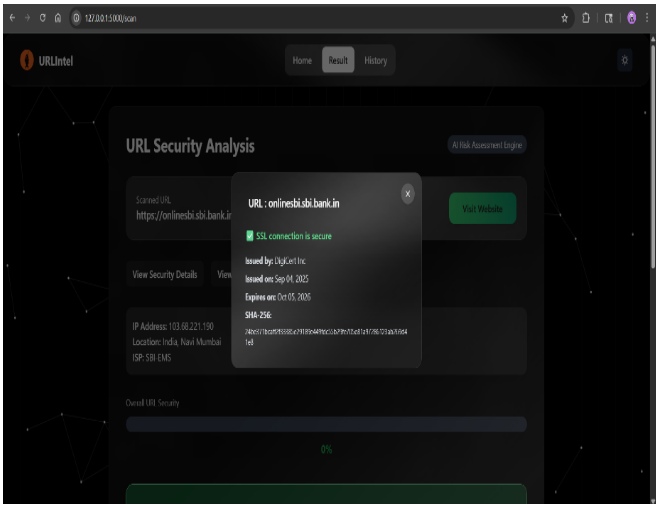
WHOIS Information

Domain: bank.in
Registrar: National Internet Exchange of India
Created: 2005-02-17 05:28:51.675000+00:00
Expires: 2027 02 17 05:28:51.675000+00:00

IP Address: 103.68.221.190
Location: India, Navi Mumbai
ISP: SBI-EMS

Fig - 6: WHOIS Information Result

The WHOIS information pop-up provides detailed domain registration data, helping users verify the authenticity and credibility of the website. It displays key details such as the domain name, registrar (National Internet Exchange of India), creation date, and expiration date, which are important indicators of domain legitimacy. In this case, the domain shows a long registration history and valid future expiry, suggesting it is a well-established and trusted domain. By presenting this information in a clear and structured format, the system enables users to assess whether a domain is newly created or suspicious, thereby improving decision-making and enhancing overall security awareness.



URL Security Analysis

Scanned URL
https://onlinesbi.sbi.bank.ir

URL : onlinesbi.sbi.bank.in

☑ SSL connection is secure

Issued by: DigiCert Inc
Issued on: Sep 04, 2025
Expires on: Oct 05, 2026
SHA-256:
24be371bcaff2f83388c29189e449fdc55b29fc705e81a97286123ab2f69d4
1e8

IP Address: 103.68.221.190
Location: India, Navi Mumbai
ISP: SBI-EMS

Fig -7: SSL Certification Result

The SSL details pop-up provides a focused and informative view of the website’s security certificate, enhancing transparency in the analysis process. It clearly indicates that the SSL connection is secure using a green status indicator, building user confidence. The panel displays key certificate information such as the issuing authority (DigiCert Inc), issue date, expiry date, and SHA-256 fingerprint, which are essential for verifying the authenticity and validity of the website. This feature allows users to inspect encryption details without cluttering the main result page, making the system both technically informative and user-friendly.

VIII. CONCLUSION

This paper presents an AI-powered URL security and phishing detection system designed to enhance online safety. By integrating machine learning techniques with feature-based analysis, the system effectively identifies phishing URLs and provides real-time feedback to users.

The proposed system overcomes the limitations of traditional detection methods by adapting to new and evolving threats. It offers a user-friendly interface and detailed insights, making it a practical solution for improving cybersecurity awareness. The successful implementation of this system demonstrates the potential of artificial intelligence in addressing modern security challenges.

REFERENCES

[1]. Ume Zara, Kashif Ayyub, Hikmat Ullah Khan, Ali Daud, Tariq Alsahfi, Saima Gulzar Ahmad, "Phishing Website Detection Using Deep Learning Models," IEEE Access, vol. 12, no. 2, Oct. 2024, DOI: 10.1109/ACCESS.2024.3486462.

[2]. K. Jain and B. B. Gupta, "PHISH-SAFE: URL Features-Based Phishing Detection System Using Machine Learning," in Cyber Security (Advances in Intelligent Systems and Computing), vol. 729, Singapore: Springer, 2018, pp. 467–474, DOI: 10.1007/978-981-10-8536-9_44.

[3]. J. K. Lee, Y. Chang, H. Y. Kwon and B. Kim, "Reconciliation of Privacy with Preventive Cybersecurity: The Bright Internet Approach," Information Systems Frontiers, vol. 22, no. 1, pp. 45–57, Feb. 2020, DOI: 10.1007/s10796-020-09984-5.

[4]. S. Asiri, Y. Xiao, S. Alzahrani, S. Li and T. Li, "A Survey of Intelligent Detection Designs of HTML URL Phishing Attacks," IEEE Access, vol. 11, pp. 6421–6443, 2023, DOI: 10.1109/ACCESS.2023.3237798.

[5]. N. A. Azeez, S. Misra, I. A. Margaret, L. Fernandez-Sanz and S. M. Abdulhamid, "Adopting Automated Whitelist Approach for Detecting Phishing Attacks," Computers & Security, vol. 108, Sep. 2021, Art. no. 102328, DOI: 10.1016/j.cose.2021.102328.

[6]. J. Ma, L. K. Saul, S. Savage and G. M. Voelker, "Beyond Blacklists: Learning to Detect Malicious Web Sites from Suspicious URLs," in Proc. ACM SIGKDD, 2009, pp. 1245–1254.

[7]. N. Abdelhamid, A. Ayesh and F. Thabtah, "Phishing Detection Based Associative Classification Data Mining," Expert Systems with Applications, vol. 41, no. 13, pp. 5948–5959, 2014.

[8]. R. Basnet, A. H. Sung and Q. Liu, "Learning to Detect Phishing URLs," International Journal of Research in Engineering and Technology, vol. 3, no. 6, pp. 11–24, 2014.

[9]. O. K. Sahingoz, E. Buber, O. Demir and B. Diri, "Machine Learning Based Phishing Detection from URLs," Expert Systems with Applications, vol. 117, pp. 345–357, 2019.

[10]. R. S. Rao and A. R. Pais, "Detection of Phishing Websites Using an Efficient Feature-Based Machine Learning Framework," Neural Computing and Applications, vol. 31, no. 8, pp. 3851–3873, 2019.

[11]. R. Verma and A. Das, "What's in a URL: Fast Feature Extraction and Malicious URL Detection," in Proc. IEEE Int. Conf., 2017, pp. 1–6.

[12]. A. C. Bahnsen, E. C. Bohorquez, S. Villegas, J. Vargas and F. A. Gonzalez, "Classifying Phishing URLs Using Recurrent Neural Networks," in IEEE Conf. Intelligence and Security Informatics, 2017, pp. 1–6.

[13]. S. Marchal, K. Saari, N. Singh and N. Asokan, "Know Your Phish: Novel Techniques for Detecting Phishing Sites and Their Targets," in IEEE Int. Conf. Distributed Computing Systems, 2016, pp. 323–333.

[14]. R. M. Mohammad, F. Thabtah and L. McCluskey, "Predicting Phishing Websites Based on Self-Structuring Neural Network," Neural Computing and Applications, vol. 25, no. 2, pp. 443–458, 2014.

[15]. Y. Zhang, J. I. Hong and L. F. Cranor, "Cantina: A Content-Based Approach to Detecting Phishing Web Sites," in Proc. WWW Conf., 2007, pp. 639–648.