Current Journal of Engineering and Science Research
Volume 3, Issue 3, pp 28-33, March 2026
ISSN (online): 3049-1738
DOI: https://doi.org/10.47001/CJESR/2026.303005
Available Online at https://www.cjesr.com

Detecting Harmful Weblinks Using Intelligent Machine Learning Models

[1]S Iliyaz, [2]S Kaveri, [3]G Vishnu Medhas, [4]V Anjali, [5]G Pavan Sai Reddy, [6]A Navadeep, [7]D Sharfuddin

[1,2,3,4,5,6,7]Department of Computer Science Engineering (Cyber Security), GATES Institute of Technology, Gooty, Andhra Pradesh, India
E-mails: [1]shaikiliyaz242003@gmail.com, [2]kaverisakekaverisake@gmail.com, [3]vishnumedhas11@gmail.com, [4]anjalivadde@gmail.com, [5]pavanreddy6255@gmail.com, [6]navadeep8657@gmail.com, [7]dsharfu21@gmail.com

Abstract: In detecting malicious websites, a common approach is the use of blacklists which are not exhaustive in them-selves and are unable to generalize to new malicious sites. Detecting newly encountered malicious websites automatically will help reduce the vulnerability to this form of attack. In this study, we explored the use of ten machine learning models to classify malicious websites based on lexical features and understand how they generalize across datasets. Specifically, we trained, validated, and tested these models on different sets of datasets and then carried out a cross-datasets analysis. From our analysis, we found that K-Nearest Neighbour is the only model that performs consistently high across data. Other models such as Random Forest, Decision Trees, Logistic Regression, and Support Vector Machines also consistently outperform a baseline model of predicting every link as malicious across all metrics and datasets. Also, we found no evidence that any subset of lexical features generalizes across models or datasets. This research should be relevant to cybersecurity professionals and academic researchers as it could form the basis for real-life detection systems or further research work.

Keywords: Lexical features, machine learning, malicious URLs.

I. INTRODUCTION

The rapid growth of the internet has increased the use of online services for communication, business, and information sharing. However, this growth has also led to an increase in cyber threats such as malicious websites. These websites attempt to steal sensitive information, spread malware, or perform unauthorized activities.

Traditional detection methods like blacklist systems can identify known malicious URLs, but they often fail to detect newly created harmful links. To overcome this limitation, machine learning techniques are used to analyze patterns and features of URLs. In this project, ma-chine learning models are used to classify URLs as *benign or malicious* using lexical features of URLs. The system aims to improve cybersecurity by automatically detecting harmful web links and protecting users from potential online attacks.

Importance and Prior Work:

Detecting malicious URLs is an important task in cyber-security because harmful links can cause data theft, malware infections, and financial loss. Researchers have proposed several techniques such as data mining and ma-chine learning to detect malicious websites.

Earlier studies used features like network traffic information, webpage content, DNS data, and URL characteristics. However, extracting network and content-based features can be time-consuming and costly for real-time detection. Therefore, many researchers focused on URL lexical features such as URL length, special characters, and suspicious keywords. Machine learning algorithms like K-Nearest Neighbor (KNN), Support Vector Machine (SVM), Random Forest, Logistic Regression, and Naive Bayes have been used to classify malicious URLS.

II. RELATED WORK

Several researchers have studied different techniques to detect malicious or harmful web links. Earlier methods mainly relied on blacklist databases and rule-based detection systems to block known phishing or malware websites. Although these methods were useful for identifying previously reported malicious URLs, they were not effective in detecting new or unknown harmful links that were not yet included in the

Current Journal of Engineering and Science Research
Volume 3, Issue 3, pp 28-33, March 2026
ISSN (online): 3049-1738
DOI: https://doi.org/10.47001/CJESR/2026.303005
Available Online at https://www.cjesr.com

blacklist.

To overcome this limitation, researchers introduced machine learning approaches that automatically analyze the characteristics of URLs. These systems extract different URL features, such as lexical features (URL length, special characters, number of subdomains), domain-related information, and hosting details. By analyzing these features, machine learning models can identify suspicious patterns and classify links as safe or malicious.

Various algorithms such as K-Nearest Neighbour (KNN), Random Forest, Decision Trees, Logistic Regression, and Support Vector Machines (SVM) have been widely used for malicious URL detection. These models are trained using datasets containing both benign and malicious URLs and then tested to evaluate their detection accuracy.

Recent research focuses on improving detection performance by using large datasets, advanced learning models, and better feature extraction techniques. These studies show that machine learning-based methods can significantly improve the detection of harmful web links and provide better protection against phishing and other web-based threats

III. PROPOSED SYSTEM

The proposed system is an intelligent web link security analyzer that uses machine learning models to detect harmful URLs in real time. When a user enters a URL, the system predicts a harmfulness percentage score (0–100%) indicating how likely the link is malicious (phishing, malware, etc.).

It integrates Explainable AI (XAI) to provide actionable suggestions based on risk level—such as "Open in secure browser only," "Avoid entering personal information," or "This link is unsafe – do not proceed."

The system maintains a scan history for reviewing past checks and offers a visualization dashboard with inter-active charts (pie charts, bar graphs, trend lines) to help users understand browsing safety patterns over time.

This combination of ML-based scoring, real-time XAI suggestions, historical tracking, and visual analytics provides transparent, accurate, and user-friendly protection against web-based threats.

IV. SYSTEM ARCHITECTURE

The system architecture processes URL security analysis through several stages. First, the user enters a URL into the system. The system performs feature extraction by analyzing characteristics such as URL length, do-main age, and special characters. These features are then sent to a machine learning model, which predicts the harmfulness percentage (0–100%) and classifies the link as Safe, Suspicious, or Harmful.

Next, the Explainable AI (XAI) module provides explanations by identifying the key features that influenced the prediction. The scan details, including the URL, score, and classification, are stored in the history data-base. Finally, the visualization dashboard displays the results and charts to help users understand the security status of scanned links.

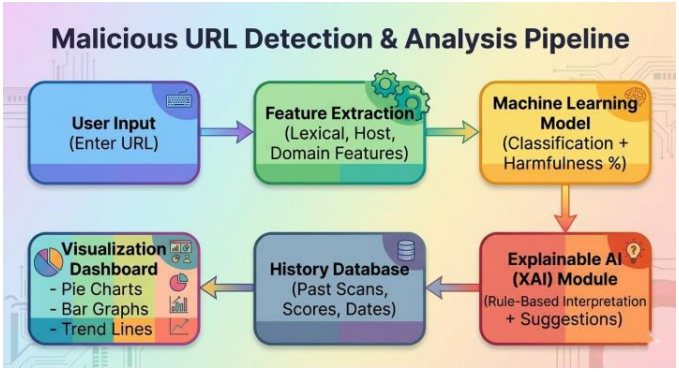

Figure 1: System Architecture

V. METHODOLOGY

The research problem can be subdivided into five main sub-problems: data collection, data pre-processing, lexical feature engineering, machine learning modeling, cross datasets analysis.

A. Data Collection:

This involved gathering the datasets that were used in this research. The sub-steps were:

1) Manually searching through online datasets repositories such as datasetsearch.google.com and Kaggle.com to find potential

Current Journal of Engineering and Science Research
Volume 3, Issue 3, pp 28-33, March 2026
ISSN (online): 3049-1738
DOI: https://doi.org/10.47001/CJESR/2026.303005
Available Online at https://www.cjesr.com

datasets that could be used for analysis.

2) Pruning datasets using metrics such as size, uniqueness, and the credibility of the source. Uniqueness is measured by a maximum of 20% overlap with the other datasets used in this analysis. The credibility of a source is measured by verifiable explanations of how the dataset was curated from. This would give a shortlist of datasets that would go through data integrity checking.

3) Performing data integrity checking by manually verifying the labels of a random subset of size 100 for each of the datasets. Any dataset that failed this manual verification of 80% accuracy as ascertained by a human verifier was dropped.

B. Data Pre-processing:

This entailed standardizing the datasets to ensure they all have the form – URL (a string) and label (benign or malicious). This stage involved writing python scripts that do this standardization based on the dataset that we are dealing with. Using these scripts, each of the datasets will then be standardized for analysis.

C. Lexical Feature Engineering:

This involved generating values for the feature space of this analysis using the URL lexical properties. Lexical feature engineering had the following steps:

1) Searching literature to obtain the lexical features used in this kind of analysis

2) Ranking these features in terms of popularity, literature importance, and novelty.

3) Designing at least one entirely new feature based on our domain knowledge.

4) Writing python scripts that take in the standardized dataset and return all the engineered features that will then be used for further analysis.

5) Partition of the datasets into training partition (34%), validation partition (33%), and test partition (33%).

D. Machine Learning Modelling:

This was the heart of our analysis. It involved converting the standardized dataset into relevant models.

This sub-problem involved:

1) Shortlisting machine learning models based on literature use for malicious website detection or similar tasks.

2) Pruning the shortlisted machine models based on empirical performance both in the literature related to malicious website detection and other well-known tasks to 10 models.

3) Determining what metrics would be used for training and validating the models.

4) Building the machine learning models, training them on the training partition, and then carrying out validation with hyperparameter optimization on the validation set, to ensure that the models are well fine-tuned to the task at hand.

5) Saving these fine-tuned models for cross-dataset analysis.

E. Cross-Datasets Analysis:

1) The cross-datasets analysis involved testing the models on the test partition and performing comparative analysis on all the models' performances across every dataset in the test partition. To guarantee that our modeling choices were appropriate, we used the same models to train, validate and test on one of the datasets in a single dataset analysis.

2) If the results of the single dataset analysis were okay, it mean that our modeling choices were solid, and the results obtained from the cross-datasets analysis were justifiable. Ranking tables were obtained to show the models that generalized best across the datasets based on the comparative analysis using the mean rank score.

3) In this step, the trained machine learning models are tested on different datasets to evaluate their performance. This analysis helps to check whether the model can accurately detect harmful and safe URLs on new data. It ensures that the system works effectively across multiple datasets and improves the reliability of malicious web link detection.

Current Journal of Engineering and Science Research
Volume 3, Issue 3, pp 28-33, March 2026
ISSN (online): 3049-1738
DOI: https://doi.org/10.47001/CJESR/2026.303005
Available Online at https://www.cjesr.com

VI. RESULTS

Figure 2: Malicious URL Detection system interface

This image shows a web application interface designed to detect malicious website URLs using intelligent ma-chine learning models. The screen displays a local development environment with navigation buttons for "About," "Contact Details," and the core URL detection tool.

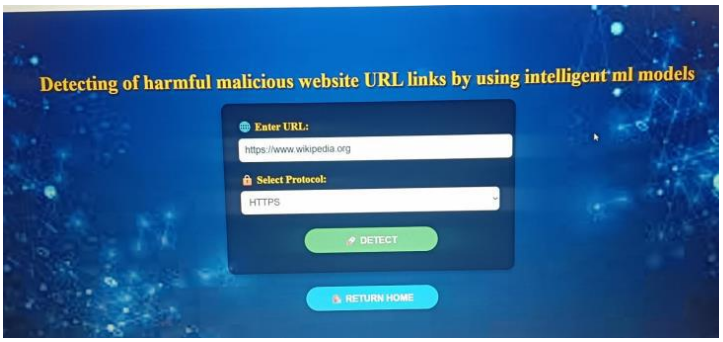

Figure 3: URL input page for checking harmful links

In this step, the image shows the active input form where a user specifies the URL they wish to analyze for potential security threats. The interface features an "Enter URL" field (currently containing a Wikipedia link) and a "Select Protocol" dropdown, allowing the machine learning model to process the specific components of the web address.

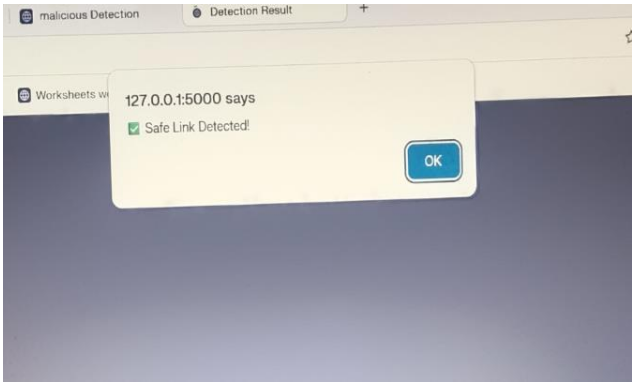

Figure 4: Result page showing safe or malicious link detection

In this step, the image shows the final output alert generated by the machine learning model after analyzing the provided URL. The application displays a popup message from the local server (127.0.0.1:5000) confirming "Safe Link Detected!", which indicates the link success-fully passed the security classification.

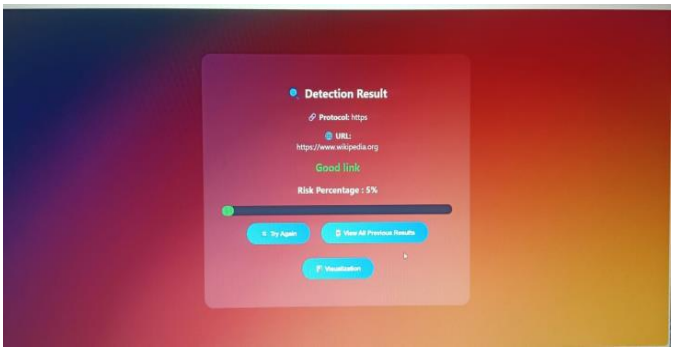

Figure 5: Analysis records and previous detection results

Here in this step, the image shows the detailed classification dashboard where the machine learning model provides a specific "Good link" verdict for the analyzed URL. The interface displays a Risk Percentage of 5% on a visual progress bar, indicating a very high confidence level that the link is safe and non-malicious.

Current Journal of Engineering and Science Research
Volume 3, Issue 3, pp 28-33, March 2026
ISSN (online): 3049-1738
DOI: https://doi.org/10.47001/CJESR/2026.303005
Available Online at https://www.cjesr.com

Figure 6: Previous URL detection results table

Here in this step, the image shows a comprehensive logs table titled "Previous Detection Results," which acts as a database record of all analyzed URLs. The table organizes historical data into columns such as Prediction, Risk Level, Risk %, and Suggestions, showing that while Wikipedia is flagged as a "Good link" (5% risk), other entries like an IP address are flagged as "Harmful" with a higher 30% risk.

Figure 7: Risk level distribution of analyzed URLS

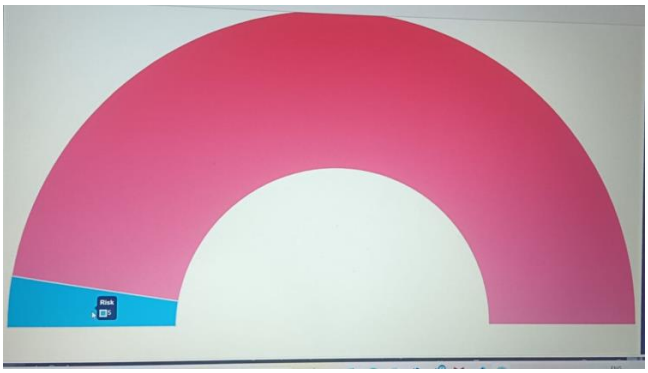

Figure 8: Risk level gauge visualization of analyzed URLS

Here in this step, the image shows the Visualization Dashboard, which uses a semi-circular doughnut chart to represent the risk analysis of the URL.

The small blue segment on the left, labelled with a "Risk" value of 5, visually reinforces that the probability of the link being harmful is extremely low compared to the safe (red) portion of the scale.

VII. CONCLUSION

This project presents a system for detecting harmful or malicious web links using intelligent machine learning models. The system analyzes different features of URLs and classifies them as safe or harmful. By using machine learning techniques, the model can identify suspicious links more effectively than traditional blacklist methods. The proposed system helps improve user security by warning users before they access dangerous websites. Overall, this project demonstrates how machine learning can be used to enhance web safety and protect users from online threats.

REFERENCES

[1] B. Eshete, A. Villafiorita and K. Weldemariam, "Malicious Website Detection: Effectiveness and Efficiency Issues", 2011 First SysSec Workshop, 2011. Available: 10.1109/syssec.2011.9.

[2] A.Ali Ahmed, "Malicious Website Detection: A Review", Journal of Forensic Sciences & Criminal Investigation, vol. 7, no. 3, 2018. Available:10.19080/jfsci.2018.07.555712.

[3] NortonLifeLock, “Norton,” Norton.com, 2019. https://us.norton.com/internetsecurity-malware-what-are-maliciouswebsites.html.

[4] F. Vanhoenshoven, G. Nápoles, R. Falcon, K. Vanhoof and M. Köppen, "Detecting Malicious URLs using Machine Learning Techniques," 2016 IEEE Symposium Series on Computational Intelligence (SSCI), Athens, Greece, 2016, pp. 1-8, DOI:10.1109/SSCI.2016.7850079.

[5] M. Jordan and T. Mitchell, "Machine learning: Trends, Perspectives, and Prospects", Science, vol. 349, no. 6245, pp. 255-260, 2015. Available:10.1126/science.aaa8415

[6] A.S. Manjeri, K. R., A. M.N.V., and P. C. Nair, "A Machine Learning Approach for Detecting Malicious Websites using URL Features," 2019 3rd International Conference on Electronics, Communication, and Aerospace Technology (ICECA), Coimbatore, India, 2019, pp. 555-561, DOI: 10.1109/ICECA.2019.8821879.

Current Journal of Engineering and Science Research
Volume 3, Issue 3, pp 28-33, March 2026
ISSN (online): 3049-1738
DOI: https://doi.org/10.47001/CJESR/2026.303005
Available Online at https://www.cjesr.com

[7] Q. T. Hai and S. O. Hwang, “Detection of Malicious URLs Based on Word Vector Representation and Ngram,” Journal of Intelligent & Fuzzy Systems, vol. 35, no. 6, pp. 5889–5900, Dec. 2018, doi: 10.3233/jifs-169831.

[8] D. SAHOO, C. LIU, and S. HOI, "Malicious URL Detection using Machine Learning: A Survey", Arxiv.org, 2021. [Online]. Available: https://arxiv.org/pdf/1701.07179.pdf.

[9] A.Y. Daeef, R. B. Ahmad, Y. Yacob, and N. Y. Phing, "Wide Scope and Fast Websites Phishing Detection using URLs Lexical Features," 2016 3rd International Conference on Electronic Design (ICED), Phuket, Thailand, 2016, pp. 410-415, doi: 10.1109/ICED.2016.7804679.

[10] Y. Zhauniarovich, I. Khalil, T. Yu, and M. Dacier, "A Survey on Malicious Domains Detection through DNS Data Analysis", ACM Computing Surveys, vol. 51, no. 4, pp. 1-36, 2018. Available: 10.1145/3191329.

[11] K. Rieck, T. Krueger, and A. Dewald, “Cujo: Efficient detection and Prevention of Drive-by-download Attacks,” in Annual Computer Security Applications Conference (ACSAC), 2010, pp. 31–39.

[12] J. Ma, L. K. Saul, S. Savage, and G. M. Voelker, “Beyond Blacklists: Learning to detect Malicious Web-sites from Suspicious URLs.” Proceedings of the 15th ACM SIGKDD international conference on Knowledge discovery and data mining - KDD ’09, 2009, DOI: 10.1145/1557019.1557153.

[13] A.Joshi, L. Lloyd, P. Paul Westin, and S. Seethapathy, "Using Lexical Features for Malicious URL Detection -A Machine Learning Approach", 2019.

[14] H. Kazemian and S. Ahmed, "Comparisons of Ma-chine Learning Techniques for Detecting Malicious Web Pages", Expert Systems with Applications, vol. 42, no. 3, pp. 1166-1177, 2015. Available: 10.1016/j.eswa.2014.08.046.

[15] A.S. Manjeri, K. R., A. M.N.V., and P. C. Nair, "A Machine Learning Approach for Detecting Malicious Websites using URL Features," 2019 3rd International Conference on Electronics, Communication, and Aerospace Technology (ICECA), Coimbatore, India, 2019, pp. 555-561, DOI: 10.1109/ICECA.2019.8821879.

Citation of this Article:

S Iliyaz, S Kaveri, G Vishnu Medhas, V Anjali, G Pavan Sai Reddy, A Navadeep, & D Sharfuddin. (2026). Detecting Harmful Weblinks Using Intelligent Machine Learning Models. *Current Journal of Engineering and Science Research*. 3(3), 28-33. Article DOI: https://doi.org/10.47001/CJESR/2026.303005

***** End of the Article *****