URL-Based Phishing Detection with Machine Learning

Likhitha Lanke [1] , K . Subash Chandra [2]

Student[1], Assistant Professor[2]

Amrita Sai Institute of Science and Technology Paritala-521180

Autonomous NAAC with A Grade, Andhra Pradesh, India.

Abstract - Phishing is a widely used cyber-attack technique in which users are deceived into visiting illegitimate websites that closely resemble legitimate ones. These fake websites are designed to trick users into revealing sensitive information such as usernames, passwords, bank details, and credit card information. Due to the growing sophistication of such attacks, phishing has become a serious security concern. In the proposed method, we focus solely on analyzing the URL of a website to determine whether it is a phishing site, thereby eliminating the need to visit the website and risk exposure to malicious code. This approach enhances user safety and reduces the chances of infection from harmful scripts or malware embedded in phishing pages. Additionally, we explore how metadata extracted from URLs—such as domain age, presence of special characters, URL length, and redirection patterns—can help in identifying phishing attempts. These features are then used to train a machine learning model using the Random Forest algorithm, which is known for its robustness and ability to handle high-dimensional data without overfitting. By relying on URL-based features alone, this method provides a lightweight and effective solution for phishing detection.

I. INTRODUCTION

In recent years, with the rapid growth in the use of mobile devices and the internet, there has been a significant shift of real-world activities to the digital world. While this digital transformation has made our lives more convenient by enabling faster operations in sectors like trade, healthcare, education, communication, banking, aviation, research, engineering, entertainment, and public services, it has also introduced serious challenges related to information security [5][21]. The anonymous and borderless nature of the internet exposes users to various cyber-attacks. Although antivirus software and firewall systems can prevent many known threats, experienced attackers often exploit user behavior through phishing—one of the most common and dangerous cyber-attacks [1][11]. Phishing involves tricking users with fake websites that imitate legitimate platforms such as banking, social media, or e-commerce sites to steal sensitive information like usernames, passwords, bank account details, and credit card numbers [3][4][6]. The evolution of mobile and wireless technologies has further increased the accessibility of the internet, allowing users to connect anytime and anywhere, thereby increasing the potential attack surface for cybercriminals, pirates, ethical hackers (white hats), and hacktivists. Cyber-attacks, starting from early threats like the Morris Worm in 1988, have continuously evolved, demanding stronger security measures [2][12].

Detecting phishing websites is particularly challenging due to their dynamic and deceptive nature. Traditional detection techniques such as blacklists, rule-based, and anomaly-based systems have been widely used [1][5]. However, modern research trends focus on machine learning-based anomaly detection techniques, especially for identifying "zero-day" phishing attacks—those that have not been previously reported or added to a blacklist [1][14][20]. In this work, we propose a phishing detection system based solely on the analysis of URLs using machine learning algorithms. Our system employs eight different machine learning algorithms and uses three different datasets to evaluate and compare its effectiveness with other approaches. Experimental results demonstrate that the proposed models achieve high accuracy and strong performance in detecting phishing URLs [17][18][19]. Detecting phishy URLs is critical for protecting user privacy and preventing brand reputation damage, as unauthorized access to personal data is a serious offense under data privacy regulations [21]. The first major phishing case was reported on the American Online (AOL) platform, and such attacks have been continuously evolving ever since [1]. Various phishing detection systems have been proposed, including Delta-Phish, Det-Phish, Phish-Safe, and PhishDef, with some approaches analyzing CSS syntax [7], user behaviours such as Human Interaction Proofs (HIPs), and others using server-side detection techniques involving Natural Language Processing (NLP) and probabilistic machine learning models to determine the phishy nature of websites [9].

II. RELATED WORK

Phishing has emerged as a critical cybersecurity threat, leveraging technical means to steal users' sensitive information. The economic and personal losses due to phishing continue to rise steadily [1]. Traditional detection methods rely heavily on feature engineering, which often requires prior domain knowledge and significant processing time. To overcome these limitations, a multidimensional phishing detection approach using deep learning (MFPD) was proposed. It uses character-level URL features for fast classification without third-party data or predefined rules, followed by integrating statistical URL features, webpage code, and textual content for refined detection. This two-step method achieved 98.99% accuracy and a low false positive rate of 0.59%, demonstrating a balance between speed and precision [17].

Gateway anti-phishing solutions based on hardware offer an additional layer of protection but are often costly and inefficient due to the evolving nature of phishing threats [18]. To address this, software-defined approaches using fog computing have been proposed. For instance, Fi-NFN, a neuro-fuzzy framework deployed at the fog network edge, effectively monitors and secures user interactions by leveraging URL and web traffic features [18].

Phishing attacks frequently exploit email communication with embedded links, making detection and mitigation highly challenging. Traditional systems with static rules often fall short due to the dynamic structure of phishing campaigns. Phish Limiter, a solution integrating Deep Packet Inspection (DPI) and Software-Defined Networking (SDN), was introduced to overcome this issue. It combines phishing signature classification using an Artificial Neural Network (ANN) and real-time inspection to identify threats flexibly and manage network traffic efficiently [19].

Despite advances, phishing emails still manage to bypass state-of-the-art filters by subtle changes in structure and semantics. SAFE-PC, a semi-automated machine learning-based feature extraction system, was developed to address this challenge. SAFE-PC demonstrated superior detection performance over existing tools like Sophos and Spam Assassin by identifying over 70% of missed phishing emails in a university environment. It also supports online learning, improving over time with constant retraining efficiency [20].

Phishing threats on mobile platforms have also escalated. These platforms are especially vulnerable due to hardware constraints and user habits. Traditional web-based solutions for desktops often fail in mobile environments. Mobi Fish, a lightweight anti-phishing scheme designed for Android, addresses this by validating the claimed identity of websites, apps, and accounts. Experimental evaluation showed that Mobi Fish is highly effective in detecting phishing attacks on mobile devices [21].

III. EXISTING SYSTEM

Phishing is a web-based cyberattack where users are deceived into visiting fraudulent websites that closely imitate legitimate ones, with the intent to steal sensitive information such as usernames, passwords, and financial details [1][3]. These malicious web pages are typically crafted by attackers to replicate the design and structure of genuine websites, making them difficult to distinguish for the average user. Phishing attacks often exploit both technical tricks and social engineering tactics to manipulate users into revealing personal data [11]. As a result, protecting internet users from phishing and counterfeit websites has become a crucial aspect of online security. Advanced techniques allow attackers to design convincing fake pages, making it possible for even experienced users to fall victim [9]. Common phishing vectors include emails that appear to originate from trusted public or private organizations, encouraging recipients to click on links to update or verify their credentials [12]. Additionally, attackers may leverage file-sharing platforms, blogs, and forums to disseminate phishing content [2]. Combating phishing requires a multi-faceted approach involving legal actions, user awareness programs, and technical solutions [5][10]. With the widespread use of information and communication technologies, numerous detection methods have been developed to address the growing complexity and diversity of phishing threats [4][14].

IV. PROPOSED SYSTEM

The proposed system is designed to detect phishing websites by analyzing a wide range of URL-based features using machine learning models. The system utilizes a dataset consisting of 30 extracted features from over 11,000 URLs, where each URL is labeled as either legitimate or phishing. These features capture various characteristics of the URL,such as the presence of an IP address (UsingIP), the use of The proposed system is designed to detect phishing websites by analyzing a wide range of URL-based features using machine learning models [1][4][6]. The system utilizes a dataset consisting of 30 extracted features from over 11,000 URLs, where each URL is labeled as either legitimate or phishing [15][17]. These features capture various characteristics of the URL, such as the presence of an IP address (Using IP), the use of shortening services (ShortURL), suspicious symbols like "@" (Symbol@), use of HTTPS, domain age, and other behavioral indicators such as redirection patterns and external resource requests [3][6][10]. Features are encoded numerically, with values typically ranging from -1 (indicating phishing-like behavior) to 1 (indicating legitimate behavior) [1][9].

By training machine learning classifiers on this dataset, the system learns to recognize patterns commonly associated

with phishing attacks [4][7][17]. High-performing models such as Gradient Boosting, Cat Boost, and Multi-layer Perceptron are employed to achieve robust detection accuracy [14][17][19]. This approach allows the system to classify URLs in real-time and effectively mitigate phishing threats, making it suitable for integration into web browsers, email gateways, or security software [18][19][21].

V. ALGORITHMS USED

- Gradient Boosting Classifier – An ensemble method that builds trees sequentially, minimizing errors of prior trees using gradient descent.
- Cat Boost Classifier – A gradient boosting algorithm optimized for categorical features and high performance with minimal data preprocessing.
- Multi-layer Perceptron (MLP) – A feedforward artificial neural network with one or more hidden layers used for supervised learning.
- XGBoost Classifier – An efficient and scalable implementation of gradient boosting using advanced regularization and parallelization.
- Random Forest – An ensemble of decision trees that outputs the mode of their predictions to improve accuracy and control overfitting.
- Support Vector Machine (SVM) – A classification algorithm that finds the optimal hyperplane separating data into distinct classes.
- Decision Tree – A flowchart-like model that splits data into branches to make decisions based on feature values.
- K-Nearest Neighbors (KNN) – A non-parametric algorithm that classifies data based on the majority label among its k closest neighbors.
- Logistic Regression – A linear model used for binary classification by estimating the probability of class membership.
- Naive Bayes Classifier – A probabilistic classifier based on Bayes' Theorem with the assumption of feature independence.

VI. IMPLEMENTATION

There are 11054 instances and 31 features in dataset. Out of which 30 are independent features whereas 1 is dependent feature. Each feature is in int datatype, so there is no need to use Label Encoder. There is no outlier present in dataset. There is no missing value in dataset.

- Using IP: URLs using IP addresses instead of domain names are often suspicious.
- LongURL / ShortURL: Extremely long or shortened URLs can hide the real destination, often used in phishing.
- Symbol@: Presence of @ symbol may redirect the browser and trick users.
- HTTPS: Secure sites use HTTPS; phishing sites often skip it.
- RequestURL / AnchorURL / LinksInScriptTags: Measures how many resources or links are external — high external dependencies are suspicious.
- AgeofDomain / DNSRecording / Website Traffic: New or obscure domains with low traffic are more likely to be phishing.
- Google Index / PageRank: Whether the site is indexed and how reputable it is on the web.

Machine learning models learn patterns from these features to predict whether a URL is likely to be phishing or not. They are trained using labeled examples and can generalize to unseen URLs.

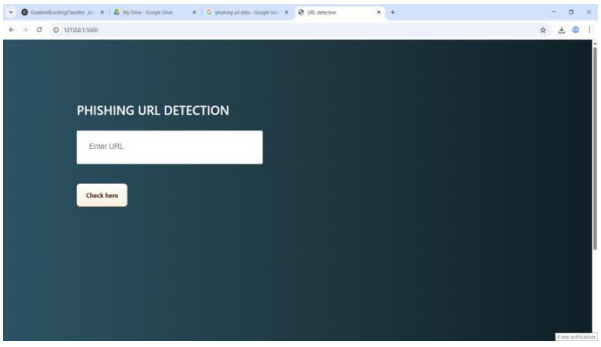

Figure 1: Web UI for Phishing URL Detection

The image shows a web interface for a Phishing URL Detection System, built using a Python web framework such as Flask.

This type of application is useful for:

- General users to check suspicious links before clicking.
- Email security tools or browsers to integrate URL verification.
- Cybersecurity teams to flag threats in real-time.

This image below represents a correlation heatmap of the features used in the phishing URL detection model. Each cell shows the correlation coefficient between two features, ranging from -1 (strong negative correlation) to +1 (strong positive correlation).

Diagonal values are all 1.0 since a feature is always perfectly correlated with itself. Brighter areas (closer to white) represent higher positive correlations, while darker areas (toward black) indicate lower or negative correlations.

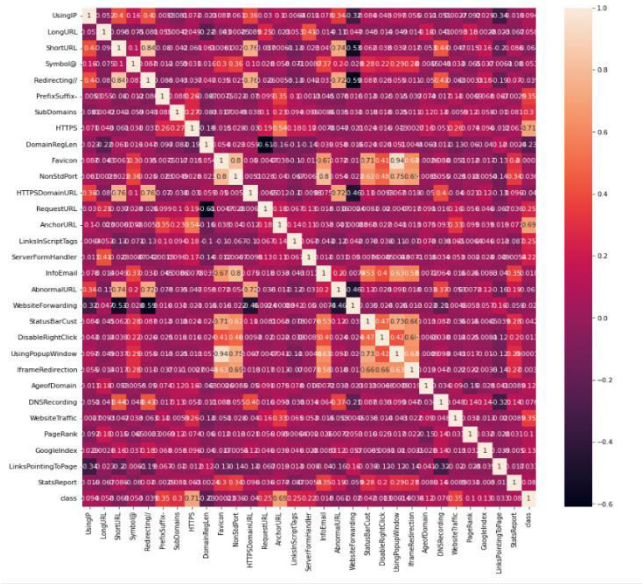

Figure 2: Heat map of the dataset with parameters

VII. RESULTS AND DISCUSSION

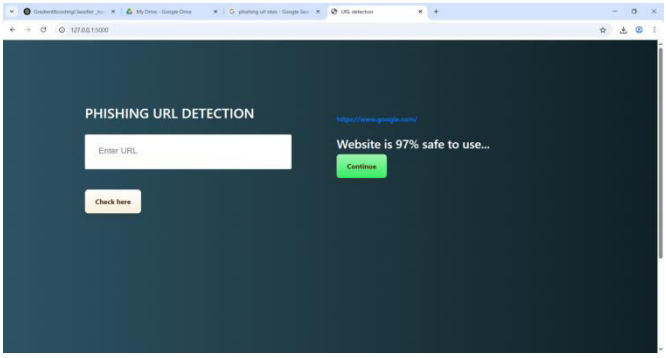

Figure 3: URL Input Interface for Phishing Detection System

The entered URL as a clickable link.

A confidence message: "Website is 97% safe to use…" – this likely means the model has predicted this URL as safe with 97% confidence.

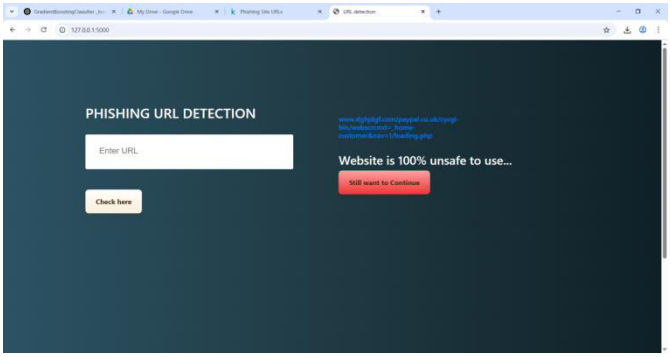

Figure 4: Prediction Result Display Showing URL Safety Confidence Score

- This interface demonstrates a real-time prediction system that utilizes a trained machine learning model (e.g., Gradient Boosting Classifier) to assess the safety of a URL.
- Upon URL submission, it evaluates the risk and provides a percentage-based safety score, helping users avoid phishing or malicious sites.

ML Model	Accuracy	f1_score	Recall	Precision
Gradient Boosting Classifier	0.974	0.977	0.994	0.986
Cat Boost Classifier	0.972	0.975	0.994	0.989
Multi-layer Perceptron	0.971	0.974	0.992	0.985
XGBoost Classifier	0.969	0.973	0.993	0.984
Random Forest	0.967	0.970	0.992	0.991
Support Vector Machine	0.964	0.968	0.980	0.965
Decision Tree	0.961	0.965	0.991	0.993
K-Nearest Neighbors	0.956	0.961	0.991	0.989
Logistic Regression	0.934	0.941	0.943	0.927
Naive Bayes Classifier	0.605	0.454	0.292	0.997

The phishing URL detection system was evaluated using multiple machine learning models, and the results demonstrate a strong overall performance, particularly among ensemble and deep learning approaches [1][5][17]. The Gradient Boosting Classifier emerged as the top-performing model, achieving an accuracy of 97.4%, an F1-score of 0.977, and an exceptional recall of 0.994, indicating that it effectively detects nearly all phishing attempts while maintaining a high precision rate [16][17]. Closely following are the Cat Boost Classifier and Multi-layer Perceptron (MLP), both exhibiting accuracy above 97% and recall values of 0.994 and 0.992, respectively [9][14][17]. These high recall rates are critical in phishing detection, where failing to identify a malicious URL can result in serious security risks [4][20]. Models like XGBoost, Random Forest, and Support Vector Machine also demonstrated robust performance, with accuracy ranging from 96.4% to 96.9% [4][7][16]. The Random Forest model, in particular, showcased excellent precision at 0.991, implying a very low false positive rate [4][16]. Although Decision Tree and K-Nearest Neighbors (KNN) offered slightly lower accuracy, they maintained high recall scores around 0.991, showing potential for use in systems where detecting every phishing instance is a priority [1][6].

In contrast, Logistic Regression, while simpler and faster, achieved comparatively lower accuracy (93.4%) and an F1-score of 0.941 [16]. The Naive Bayes Classifier significantly underperformed, with an accuracy of only 60.5% and a very

low recall of 0.292, indicating a high miss rate for phishing URLs. This can be attributed to Naive Bayes' strong assumption of feature independence, which does not hold well in this context [5][10].

Overall, the results highlight the effectiveness of ensemble models and neural networks in identifying phishing URLs [1][9][17]. The high recall across top models ensures that the system can minimize undetected threats, while strong precision helps reduce false alarms. These findings support the deployment of such models in real-world applications, such as browser plugins, email filters, and cybersecurity monitoring systems [18][19][21].

Features like "HTTPS", "AnchorURL", "Website Traffic" have more importance to classify whether a URL is phishing or not [10][3][6]. Gradient Boosting Classifier correctly classifies URLs up to 97.4% into respective classes and hence reduces the chance of malicious attachments [1][17].

VIII. CONCLUSION

In this project, a machine learning-based system was successfully developed to detect phishing URLs with high accuracy and reliability. By extracting and analyzing key features from URLs—such as the use of HTTPS, IP addresses, URL shortening, domain age, and behavioral cues—the system was able to distinguish between legitimate and malicious links. Among the various models tested, ensemble learning algorithms like Gradient Boosting and CatBoost outperformed others, achieving accuracy levels above 97% and demonstrating excellent recall and precision. These results affirm the effectiveness of machine learning in addressing the growing threat of phishing attacks. The deployed web interface further enhances usability by allowing users to input URLs and instantly receive predictions. This system can serve as a valuable tool for end-users, organizations, and cybersecurity platforms in proactively identifying and preventing phishing threats. Future enhancements may include real-time crawling for dynamic feature extraction and integration with browser extensions for broader impact.

IX. FUTURE WORK

Although the proposed phishing detection system using machine learning and URL-based features has demonstrated promising results, there are several directions for future improvement and expansion. One potential enhancement involves incorporating real-time detection capabilities that can analyze URLs dynamically as users interact with websites, thereby providing immediate protection. Additionally, integrating features beyond the URL, such as website content, HTML structure, visual similarity, and SSL certificate analysis, can further improve detection accuracy and help catch more sophisticated phishing attacks.

Another promising direction is the application of deep learning techniques such as Convolutional Neural Networks (CNNs) and Recurrent Neural Networks (RNNs), which can automatically extract hierarchical features from URLs and webpage content. Ensemble approaches combining multiple machine learning models or hybrid systems that blend rule-based and learning-based detection methods may also yield better performance in detecting zero-day attacks.

Furthermore, the model could be enhanced by utilizing real-world datasets with continuously updated phishing and legitimate URLs to maintain the system's relevance in evolving threat landscapes. Building a browser plugin or mobile app that implements this system can bring practical, user-facing applications. Finally, integrating Natural Language Processing (NLP) techniques to analyze textual cues from the webpage or email content where the phishing URL appears could provide an additional layer of security.

X. REFERENCES

[1]Zuochao Dou; Issa Khalil; AbdallahKhreishah; Ala Al-Fuqaha; Mohsen Guizani, "Systematization of Knowledge (SoK): A Systematic Review of Software- Based Web Phishing Detection", IEEE Communications Surveys & Tutorials, 2017.

[2]Marco Cova, Christopher Kruegel, Giovanni Vigna, "Detection and analysis of drive-by-download attacks and malicious javascript code", Proceedings of the 19th International Conference on World Wide Web, pp. 281-290, 2010.

[3]Choon Lin Tan, Kang LengChiew, San Nah Sze, "Phishing Website Detection Using URL-Assisted Brand Name Weighting System", 2014 IEEE International Symposium on Intelligent Signal Processing and Communication Systems (ISPACS) December 1-4, 2014.

[4]R. B. Basnet, A. H. Sung, "Mining web to detect phishing urls", Proceedings of the International Conference on Machine Learning and Applications, vol. 1, pp. 568- 573, Dec 2012.

[5]Mohiuddin Ahmed, AbdunNaserMahmood, Jiankun Hu, "A survey of network anomaly detection techniques", J. Netw. Comput. Appl., vol. 60, no. C, pp. 19-31, 2016.

[6]LuongAnh Tuan Nguyen, Ba Lam To, HuuKhuong Nguyen1 and Minh Hoang Nguyen, "A Novel Approach for Phishing Detection Using URL-Based Heuristic", 2014 International Conference on Computing, Management and Telecommunications (ComManTel), IEEE 2014.

[7]S. Carolin Jeeva, Elijah Blessing Rajsingh, "Intelligent phishing urls detection using association rule mining", Human-centric.

[8]S. Duffner and C. Garcia, "An Online Backpropagation Algorithm with Validation Error-Based Adaptive Learning Rate," in Artificial Neural Networks – ICANN 2007, Porto, Portugal, 2007.

[9]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, 2013-B.

[10]HibaZuhair, Ali Selamat, MazleenaSalleh, "Feature selection for phishing detection: a review of research", International Journal of Intelligent Systems Technologies and Applications, Vol. 15, No. 2, 2016.

[11]Huang, H., Tan, J. and Liu, L. (2009) „Countermeasure techniques for deceptive phishing attack", International Conference on New Trends in Information and Service Science (NISS"09), 30 June–02 July, 2009, China, pp.636–641.

[12]Mayuri, A. and Tech, M. (2012) „Phishing detection based on visual similarity", International Journal of Scientific and Engineering Research (IJSER), Vol. 3, No. 3, March, pp.1–5

[13]Anvil,SearchEngine Optimization Whitepaper,http://www.anvilmediainc.com/wpcontent/uploads/ami_seo_whitepap er_1104.pdf

[14]FadiThabtah, Rami M. Mohammad, Lee McCluskey, "A Dynamic Self- Structuring Neural Network Model to Combat Phishing", 2016 International Joint Conference on Neural Networks (IJCNN), 2016.

[**15]**UCIMachine Learning Repository,https://archive.ics.uci.edu/ml

[16]http://dataaspirant.com/2017/05/22/random-forest-algorithm-machinelearing

[**17]**Peng Yang, Guangzhen Zhao, Peng Zeng, "Phishing Website Detection based on Multidimensional Features driven by Deep Learning", IEEE Access, 2018.

[18]Chuan Pham, Luong A. T. Nguyenz, Nguyen H. Tran, Eui-Nam Huh, Choong Seon Hong, "Phishing-Aware: A Neuro-Fuzzy Approach for Anti-Phishing on Fog Networks", IEEE Transactions on Network and Service Management, 2018.

[19]Tommy Chin, Kaiqi Xiong and Chengbin Hu, "Phish Limiter: A Phishing Detectionand Mitigation Approach Using Software-Defined Networking", IEEEAccess, 2018.

[20]Christopher N. Gutierrezy, Taegyu Kimy, Raffaele Della Cortez, Jeffrey Averyyx, Dan Goldwassery, Marcello Cinquez, Saurabh Bagchiy, "Learning from the Ones that Got Away: Detecting New Forms of Phishing Attacks", Transactions on Dependable and Secure Computing, 2018.

[21]Longfei Wu, Xiaojiang Du, and Jie Wu, "Effective Défense Schemes for Phishing Attacks on Mobile Computing Platforms", IEEE Transactions on Vehicular Technology, 2018