Prediction of Phishing websites through Recurrent Neural Network with Autoencoders

V. Ragu[1], N. Pazhaniraja[2], B. Ezhil Kumaran[3], K. Sathish Kumar[4]

[1,3,4]UG Student, Department of Computer Science and Engineering, Madagadipet, Puducherry, India

[2]Associate Professor, Department of Computer Science and Engineering, Madagadipet, Puducherry, India

***Abstract*—Phishing websites are a growing threat to online security, exploiting users' trust by imitating legitimate platforms to steal sensitive data such as login credentials and financial information. The prevalence of these malicious sites presents a continuous challenge, as attackers frequently adapt their methods to bypass detection. Traditional defenses, like blacklisting, struggle to keep pace with the rapid evolution of phishing techniques, leaving users more vulnerable to these attacks. This issue is further intensified by the demand for real-time detection and the difficulty of differentiating between authentic and malicious URLs. As phishing tactics become increasingly advanced, it is essential to develop sophisticated, adaptive approaches that can reliably and efficiently identify and counter these threats to enhance online security.**

***Keywords*—Phishing website Prediction, Deep Learning, Recurrent Neural Networks, Autoencoders, URL pattern recognition, Cyber Security, Feature Extraction, Data analysis.**

I. INTRODUCTION

In today's digital era, Uniform Resource Locators (URLs) serve as essential pathways for accessing websites, applications, and various online services. These web addresses are structured to direct users to specific resources on the internet. However, cybercriminals increasingly exploit URLs by creating deceptive links that appear genuine but aim to mislead users and steal sensitive information. Such fraudulent URLs are a core component of phishing attacks, which mimic legitimate websites, including banking portals, social networks, or e-commerce platforms, to trick users into revealing personal credentials, financial data, or other confidential details.

Phishing websites often use URLs manipulated to closely resemble authentic ones, differing only slightly in characters or leveraging domains that appear credible. Unsuspecting users who click these links are directed to fake websites designed to collect their data for malicious purposes, such as identity theft or financial crimes. Traditional defenses, like URL blacklisting, aim to block known phishing links but struggle against newly created, dynamic URLs. This evolving threat landscape necessitates advanced detection systems that can identify suspicious URLs in real-time, reducing risks for users.

Machine learning (ML) techniques, such as Multilayer Perceptron (MLPs), have been used to detect phishing URLs, yet they face limitations in accurately predicting new and emerging threats. These methods often lack the capability for real-time analysis, leaving users vulnerable. To overcome these limitations, deep learning models, particularly Recurrent Neural Networks (RNNs) combined with autoencoders, offer a promising solution. RNNs excel at processing sequential data, such as URL structures, while autoencoders support dimensionality reduction and feature extraction, improving detection accuracy and efficiency.

The rapid growth of internet activity and the advanced tactics employed by cybercriminals, such as domain spoofing, obfuscation, and algorithmically generated URLs, further challenge traditional security systems. These evolving phishing strategies demand that detection models continuously adapt. By leveraging deep learning methods, these systems can provide a more robust defense, enhancing user safety and promoting a secure online environment globally.

II. LITERATURE SURVEY

A. AN EFFECTIVE DETECTION APPROACH FOR PHISHING URL USING RESMLP BY S. REMYA, MANU J. PILLAI, KAJAL K. NAIR, YONG YUN CHO[1]

Year: 2024

Phishing attacks, which use deceptive URLs to replicate legitimate websites, have become a major cybersecurity concern. These attacks aim to steal sensitive information by tricking users into engaging with fraudulent websites. This study presents a novel phishing URL detection method based on a residual pipeline architecture. The method utilizes convolutional and inverted residual blocks to extract key features from URLs, such as lexical attributes and domain age. These features are then passed through a Multi-Layer Perceptron (MLP) for classification. The residual pipelining enhances the model's ability to capture complex URL patterns, boosting detection accuracy. The proposed method was tested on a Kaggle dataset, achieving an impressive accuracy rate of 98.29%. Additionally, it outperformed traditional detection methods in other metrics, such as precision, recall, and F1 score. However, certain limitations were identified during the study that could affect the practical implementation of the ResMLP model. The approach's high computational demand could limit its performance in real-time applications, particularly on devices with limited resources. Moreover, its reliance on specific URL features, like domain age and lexical structure, may make it vulnerable to adversarial attacks or highly obfuscated phishing URLs. The model may also struggle to generalize to new datasets without retraining, which could hinder its adaptability to the evolving nature of phishing threats.

B. PHISHING DETECTION SYSTEM THROUGH HYBRID MACHINE LEARNING BASED ON URL BY ABDUL KARIM, MOBEEN SHAHROZ, KHABIB MUSTOFA, SAMIR BRAHIM BELHAOUARI, S. RAMANA KUMAR JOGA[2]

Year: 2023

Phishing continues to be one of the most widespread and damaging forms of cybercrime, utilizing deceptive URLs to imitate legitimate websites with the goal of stealing sensitive personal information. This study aims to combat the increasing threat of phishing attacks by leveraging machine learning techniques for detecting malicious URLs. The research introduces a hybrid model that integrates several machine learning algorithms, including Decision Tree (DT), Linear Regression (LR), Random Forest (RF), Naive Bayes (NB), Gradient Boosting Machine (GBM), K-Neighbors Classifier (KNN), and Support Vector Classifier (SVC). The main contribution of this work is the development of a hybrid LSD (Logistic Regression, Support Vector Classifier, and Decision Tree) model, which utilizes both soft and hard voting strategies to improve accuracy and efficiency in detecting phishing URLs.

The dataset used in this study includes over 11,000 URLs, with features derived from both phishing and legitimate websites. Multiple preprocessing techniques were applied, followed by the use of the canopy feature selection method, cross-fold validation, and Grid Search Hyperparameter Optimization to enhance the model's performance. The model's effectiveness was evaluated using metrics such as precision, accuracy, recall, F1-score, and specificity. The results show that the hybrid LSD model significantly outperforms other machine learning techniques, providing a marked improvement in phishing detection accuracy.

However, a notable drawback of the hybrid LSD model, as well as other complex models like gradient boosting, is the risk of overfitting. Overfitting happens when a model becomes excessively tailored to the training data, including noise and outliers, impairing its ability to generalize to new, unseen data. While these models may show high performance on training datasets, their real-world application accuracy can decline as they struggle to differentiate between genuine patterns and noise. To mitigate this challenge, techniques like regularization, cross-validation, and careful feature selection are necessary, but overfitting remains a significant concern for ensuring the robustness and scalability of phishing detection.

C. PHISHCATCHER: CLIENT-SIDE DEFENSE AGAINST WEB SPOOFING ATTACKS USING MACHINE LEARNING BY MUZAMMIL AHMED, AHME ALTAMIMI, WILAYAT KHAN, MOHAMMAD ALSAFFAR[3]

Year: 2023

With the rising incidence of cybercrimes like phishing, ensuring the confidentiality and integrity of user data has become a crucial aspect of cybersecurity. Phishing attacks, where malicious entities create fake websites that resemble legitimate ones to steal sensitive information (e.g., passwords and PINs), present a serious threat. These attacks are commonly executed through methods such as phishing emails, fake ads, and click-jacking, exposing billions of users to risk daily. While various security measures have been proposed to counter phishing, many of these solutions face challenges related to latency and accuracy, particularly in real-time applications.

To tackle these issues, the authors of this study introduce PhishCatcher, a client-side defense tool designed to identify spoofed websites and protect users from phishing attacks using machine learning techniques. PhishCatcher is developed as a Google

Chrome extension that classifies URLs as either suspicious or safe. At its core, the tool employs a machine learning algorithm that examines four key web features to detect phishing attempts. A random forest classifier is used to determine if a login page is genuine or fake. The tool was evaluated on a dataset containing 800 URLs (400 phishing and 400 legitimate) to test its accuracy and precision. The results show that PhishCatcher achieved an impressive 98.5% accuracy and precision, highlighting its effectiveness in phishing detection. Moreover, the extension operates efficiently, with an average response time of only 62.5 milliseconds when tested on forty phishing URLs, demonstrating minimal latency and real-time detection.

Despite its strong performance in terms of accuracy and efficiency, PhishCatcher has some limitations. Its reliance on a fixed set of web features may not be sufficient to detect more sophisticated or evolving phishing techniques, particularly those employing advanced obfuscation. Furthermore, the random forest classifier may not always generalize effectively to new or unseen data without frequent retraining. The tool is also limited to Google Chrome, meaning users of other browsers are excluded. While it performs well in low-latency environments, its real-time detection capabilities could be hindered in high-traffic or complex scenarios. These factors could impact its ability to detect novel or intricate phishing attacks effectively.

D. INTRUSION DETECTION BASED ON PHISHING DETECTION WITH MACHINE LEARNING BY R. JAYARAJ, A. PUSHPALATHA, K. SANGEETHA, T. KAMALESHWAR, S. UDHAYA SHREE, DEEPA DAMODARAN[4]

Year: 2023

Phishing attacks, where cybercriminals impersonate legitimate websites to steal sensitive information like usernames, passwords, and credit card details, remain a significant cybersecurity threat. In response, machine learning techniques, particularly artificial neural networks, are being leveraged to detect and prevent phishing attempts. This study introduces a Hybrid Ensemble Feature Selection (HEFS) method to enhance phishing detection accuracy. The HEFS method starts with the novel Cumulative Distribution Function gradient (CDF-g) algorithm to generate initial feature subsets, which are then refined through a data perturbation ensemble to create secondary subsets. These refined features are used to detect phishing URLs. The approach is evaluated and compared with existing methods, demonstrating its effectiveness in identifying malicious phishing sites through a more precise feature selection process. The paper emphasizes the role of advanced machine learning techniques in improving phishing URL detection and strengthening intrusion detection systems.

However, a limitation of the HEFS method is its reliance on the CDF-g algorithm for generating initial feature subsets, which may not always capture the most relevant features in the case of highly dynamic phishing attacks. Although the hybrid ensemble approach improves feature selection, the model's complexity could lead to longer computation times and higher resource consumption, especially when handling large datasets. Moreover, like many machine learning-based systems, the HEFS method’s effectiveness depends on the quality and diversity of the training data, meaning it might face challenges in detecting novel phishing techniques or adapting to evolving attack strategies without regular retraining.

E. PHISHING WEBSITE PREDICTION USING BASE AND ENSEMBLE CLASSIFIER TECHNIQUES WITH CROSS-VALIDATION BY ANJANEYA AWASTHI AND NOOPUR GOEL[5]

Year: 2022

The modern internet environment is fraught with vulnerabilities, making it easy for inexperienced or careless users to fall victim to numerous threats. Malicious actors exploit a variety of tools and techniques to steal personal information, resulting in significant data breaches and financial losses. Despite ongoing efforts by users, software developers, and application creators to secure IT infrastructure with encryption, digital signatures, and certificates, phishing remains a pervasive and difficult challenge. This paper takes a targeted approach to tackle phishing, focusing on the detection and prediction of phishing website URLs. Using machine learning classifiers and advanced ensemble methods applied to two separate datasets, the study proceeds in three stages. First, it involves classification using base classifiers, followed by ensemble classifiers, and finally, testing the performance of these ensembles with and without cross-validation. The analysis of classifier performance offers important insights that lay the groundwork for further research in strengthening cybersecurity measures against phishing threats.

The research approach follows a structured exploration of phishing detection through machine learning, combining both traditional classifiers and novel ensemble methods. By conducting experiments with two distinct datasets and merging them for a more comprehensive analysis, the study ensures an in-depth examination of phishing URL prediction. Performance evaluation is carried out by comparing the effectiveness of individual classifiers with that of ensemble classifiers, both with and without cross-validation. The findings presented in the paper serve as a reference point for future research, providing useful information for enhancing understanding and developing strategies to combat phishing in the rapidly changing digital world.

F. A DEEP EARNING-BASED FRAMEWORK FOR PHISHING WEBSITE DETECTION BY LIZHEN TANG; QUSAY H. MAHMOUD[6]

Year: 2021

Phishing attacks are a significant threat as attackers use various methods, including email, text messages, and social media platforms, to distribute malicious links. Through social engineering tactics, they trick users into visiting fraudulent websites, where sensitive personal information is unknowingly disclosed. This stolen data is then exploited to undermine trust in legitimate websites or financial institutions, leading to financial losses. Recognizing the growing sophistication of these attacks, this paper presents an innovative approach that utilizes deep learning for detecting phishing websites. The proposed system is designed as a browser plug-in that provides real-time warnings to users when they visit potentially phishing sites.

To improve accuracy and reduce false positives, the real-time prediction service incorporates several strategies, such as whitelist filtering, blacklist interception, and machine learning (ML) prediction. The paper evaluates multiple machine learning models using different datasets, and the results reveal the Gated Recurrent Unit (GRU) model as the most effective, achieving an accuracy of 99.18%. This demonstrates the feasibility and reliability of the proposed solution, highlighting the potential of deep learning techniques to strengthen cybersecurity defenses against phishing threats. The integration of real-time detection into a browser plug-in not only raises user awareness but also serves as a practical implementation of advanced machine learning for proactive defense against evolving cyber threats in the digital world.

G. MALICIOUS URL DETECTION: A COMPARATIVE STUDY BY SHANTANU, JANET B[7]

Year: 2021

In the ever-changing digital world, cybersecurity has become a critical issue, requiring innovative solutions to combat emerging threats. This study aims to introduce a novel approach based on transfer learning within deep neural networks, focusing on phishing URLs and malicious links, which are key sources of cyber-attacks. To support this, a comprehensive dataset containing a variety of phishing and legitimate URLs is carefully assembled and subjected to thorough pre-processing techniques to ensure the quality of input for the models. Departing from traditional methods, the proposed approach leverages transfer learning to identify complex patterns in URLs and their content.

The result of this approach is the development of a hybrid model that combines deep learning techniques with the advantages of transfer learning. This model is further enhanced by incorporating both soft and hard voting strategies, which help improve the accuracy and efficiency of phishing threat detection. Advanced feature selection and hyperparameter optimization methods are used during the fine-tuning stage to strengthen the model’s performance, ensuring that the proposed cybersecurity solution is robust and effective. The model is also designed to adapt to the constantly evolving nature of phishing attacks, ensuring its continued effectiveness. Furthermore, the hybrid model enhances generalization, minimizing the risk of overfitting while maintaining high accuracy across various datasets. The study uses rigorous evaluation metrics to assess the performance of the developed models, ensuring the effectiveness of the hybrid model in strengthening cybersecurity defenses against the dynamic and evolving threats, particularly those related to phishing URLs and malicious links.

III. TABLE

SL.NO	TITLE	AUTHOR	YEAR	ADVANTAGE	DISADVANTAGE

1	An Effective Detection Approach for Phishing URL Using ResMLP	S. Remya, Manu J. Pillai, Kajal K. Nair, Yong Yun Cho	2024	Captures complex URL patterns, achieving 98.29% accuracy and outperforming traditional methods in precision, recall, and F1 score.	High computational complexity affects real-time efficiency, and on specific URL features reduces robustness against obfuscated phishing URLs,
2	Phishing Detection System Through Hybrid Machine Learning Based on URL	Abdul Karim, Mobeen Shahroz, Khabib Mustofa, Samir Brahim Belhaouari, S. Ramana Kumar Joga	2023	Integrates multiple machine learning algorithms and voting strategies, enhancing accuracy and efficiency, and outperforming other models in phishing detection.	Overfitting in complex models like gradient boosting can reduce real-world accuracy.
3	Phishcatcher: Client-Side Defense Against Web Spoofing Attacks Using Machine Learning	Muzammil Ahmed, Ahme Altamimi, Wilayat Khan, Mohammad AlSaffar	2023	Detects phishing with 98.5% accuracy and 62.5 ms latency, providing an efficient real-time client-side defence.	The tool's reliance on specific web features limits detection of advanced phishing, and its performance may decline without retraining
4	Intrusion Detection Based on Phishing Detection with Machine Learning	R. Jayaraj, PushpalathaSangeetha, Deepa Damodaran, Kamalesh.	2023	Enhances phishing URL detection accuracy by refining feature selection, improving intrusion detection system robustness.	Complexity increases computation time and resource consumption, struggling to detect new phishing techniques without retraining.
5	Phishing Website Prediction Using Base and Ensemble Classifier Techniques with Cross-Validation	Anjaneya Awasthi and Noopur Goel	2022	Cross-validation aids in evaluating the performance of each individual base classifier across different subsets of the data.	The computational cost can become a limitation, especially when dealing with large datasets or complex ensemble models.
6	A Deep Learning-Based Framework for Phishing	Lizhen Tang; Qusay H. Mahmoud	2021	Inherent Feature Extraction and Learning	Sufficiently large and diverse dataset of labeled phishing websites can be challenging.

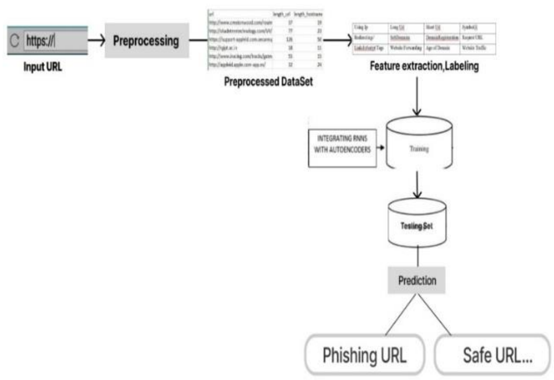
	Website Detection				
7	Malicious URL Detection: A Comparative Study	Shantanu, Janet B	2021	Ensuring scalability against evolving threats.	Despite reducing overfitting, the approach may struggle with novel or highly obfuscated phishing techniques.

IV. PROPOSED SYSTEM

The proposed system enhances phishing detection through a deep learning model that leverages Recurrent Neural Networks (RNNs) and autoencoders to address the limitations of traditional blacklist methods. RNNs are particularly well-suited for processing sequential data, making them ideal for recognizing the complex URL patterns that often differentiate legitimate links from phishing attempts. By examining the sequence of characters and structure within URLs, RNNs can more accurately identify suspicious patterns, which static blacklists may miss, enabling the model to stay effective even as phishing tactics evolve. Autoencoders add another layer of efficiency by performing dimensionality reduction and feature extraction. Through this process, the model isolates only the most essential characteristics of URLs, minimizing unnecessary data and simplifying the detection process. By focusing on key patterns within URLs, autoencoders help the model reduce noise and increase detection speed, resulting in a streamlined and optimized approach to identifying phishing threats. This capability is particularly advantageous for real-time applications, where quick and accurate detection is critical.

Integrating RNNs with autoencoders allows the system to offer real-time, precise phishing URL detection, which significantly outperforms traditional blacklist-based methods in both response time and reliability. As phishing techniques become more sophisticated, this model's adaptive nature provides a future-proof solution, equipping it to handle emerging threats effectively. Overall, this proposed deep learning approach strengthens the system's resilience and ensures robust protection against cyber threats, safeguarding users in an ever-evolving digital landscape.

Architecture Diagram of Proposed System:



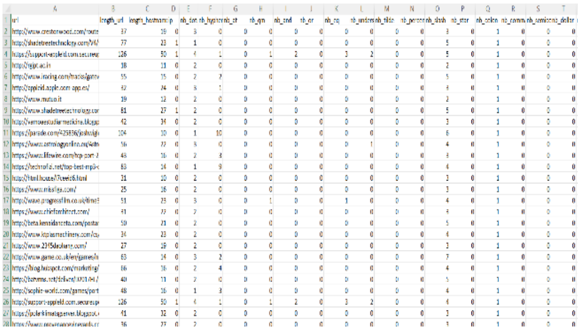
The proposed system for malicious URL detection operates through a structured, two-stage approach that utilizes deep learning techniques, specifically Recurrent Neural Networks (RNNs) and autoencoders. In the training stage, URLs are initially processed to extract features and label the data as either "Malicious" or "Safe." This stage involves training a model that integrates RNNs, which are adept at handling sequential patterns inherent in URL data, with autoencoders, which perform dimensionality reduction and improve feature representation. By reconstructing URL data, autoencoders help highlight relevant patterns and anomalies, enhancing the system's ability to discern between safe and malicious URLs. This training phase equips the model with a comprehensive understanding of the characteristics of different URLs. During the detection stage, the system applies the trained model to new, unseen URLs to determine their safety. As new URLs enter the system, features are extracted using the learned patterns from the training phase. The model then classifies each URL as "Malicious" or "Safe," based on the previously acquired knowledge. This classification process leverages the deep learning model's capacity to recognize suspicious patterns quickly and accurately, making it highly effective in real-time scenarios. By integrating RNNs with autoencoders, the system demonstrates enhanced detection accuracy and resilience, providing robust protection against evolving malicious URL threats.

a. Data collection:

The initial stage involves gathering a dataset of URLs from Kaggle, a popular platform offering diverse open-source data. This dataset contains a wide range of labeled URLs, including both safe and malicious ones, which forms the foundation for training and evaluating the detection model. Kaggle datasets are valuable because they are large, diverse, and frequently updated, capturing various phishing patterns and types. Collecting data from Kaggle ensures that the model is exposed to a broad spectrum of URLs, enhancing its ability to generalize and recognize new phishing tactics effectively.

b. Pre-processing:

In this stage, raw URL data is prepared for analysis. Pre-processing involves cleaning the data by removing duplicates, irrelevant entries, or incomplete URLs that may affect the accuracy of the model. It may also include normalizing the data to ensure consistent formats across different URLs, such as converting all URLs to lowercase.



Additionally, pre-processing could involve splitting URLs into their component parts (such as domain, subdomain, and path) to enable more detailed feature extraction. Proper pre-processing is essential to minimize noise and focus on the most relevant URL elements, facilitating the extraction of meaningful patterns.

c. Feature Extraction:

After pre-processing, the next step is to extract critical features from each URL that can indicate malicious activity. Features might include the length of the URL, the presence of certain keywords, the number of dots, special characters, and information on the domain's age. Extracting these features allows the model to identify key patterns that often distinguish malicious URLs from safe ones. This phase is particularly crucial as it creates a refined dataset with essential attributes, helping the model to focus on significant factors and enabling the RNN and autoencoder to learn effectively.

d. Model creation using RNN and Autoencoder:

The core of the detection system is created by combining Recurrent Neural Networks (RNNs) with autoencoders. RNNs are designed to handle sequential data, making them suitable for analyzing URLs, which are character sequences with specific patterns. The autoencoder, on the other hand, reduces dimensionality, focusing on relevant data patterns while filtering out noise. This combination allows the model to capture complex relationships within URL sequences, providing a more nuanced analysis than simpler models. The RNN identifies and learns sequential patterns, while the autoencoder helps in reconstructing data to improve feature representation, resulting in a robust model capable of distinguishing between safe and malicious URLs.

e. Test data:

Once the model is trained, it is tested on a separate subset of data that the model has not seen before. This test dataset, also sourced from Kaggle, includes both safe and malicious URLs and serves as a benchmark to evaluate the model's performance. Testing involves feeding URLs into the model to check its classification accuracy. By analyzing results like accuracy, precision, and recall on this test data, researchers can assess how well the model generalizes and whether any adjustments are needed for improvement. Effective testing ensures the model's reliability in real-world applications.

f. Prediction:

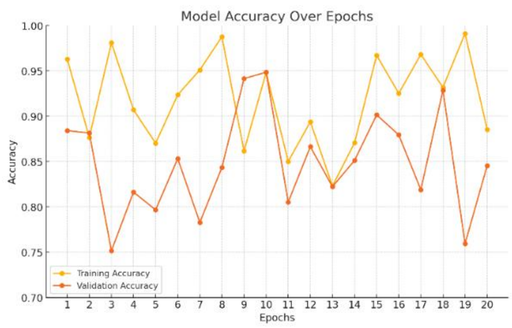
In the final stage, the trained model uses its learned patterns to predict whether new URLs are "Safe" or "Malicious." Each URL input goes through the same feature extraction process and is then classified by the RNN-autoencoder model. This prediction phase is critical, as the model needs to apply its training to unseen data in real-time. The ultimate goal is for the system to accurately flag malicious URLs quickly and reliably, providing proactive security measures against phishing attempts and minimizing risks for users.

V. RESULT AND DISCUSSION

The results of the proposed phishing detection system reveal significant improvements over traditional methods, particularly those based solely on blacklisting. The evaluation metrics indicate a high accuracy rate in detecting phishing URLs, with the hybrid model integrating RNNs and autoencoders demonstrating superior performance in identifying both known and novel threats. Notably, the system's capability to detect phishing attempts at the zero-hour mark—before they gain widespread recognition—underscores its proactive nature and effectiveness in real-time applications. Additionally, the integration of advanced feature extraction techniques allowed for the isolation of critical characteristics that distinguish legitimate links from malicious ones, leading to fewer false positives. This balance between detection accuracy and minimizing false alarms is crucial for maintaining user trust and ensuring a seamless browsing experience. The discussions surrounding these results highlight the model's adaptability in an evolving cybersecurity landscape, illustrating its resilience against increasingly sophisticated phishing tactics. Furthermore, the findings suggest that employing deep learning methodologies, specifically transfer learning, not only enhances detection rates but also equips the system to continually learn from emerging threats, thereby reinforcing the importance of innovative approaches in the field of cybersecurity. Overall, the results affirm the proposed system's potential as a robust and reliable tool in the ongoing battle against phishing and other cyber threats.

a. Accuracy:

In the context of using Recurrent Neural Networks (RNNs) and autoencoders for phishing detection, accuracy serves as a crucial performance metric to assess how well the model identifies phishing URLs versus legitimate ones. The accuracy of a model built with RNNs hinges on its ability to effectively process sequential data, such as the characters and structure within URLs. By leveraging RNNs, the model can recognize complex patterns that distinguish phishing attempts from legitimate links.



Autoencoders complement this by performing dimensionality reduction and feature extraction, enabling the model to focus on the most critical features of URLs while minimizing noise. The combined capabilities of RNNs and autoencoders enhance the model's predictive accuracy by ensuring that it not only identifies phishing URLs but does so with a high degree of precision. The formula for accuracy remains consistent across models:

$$\text{Accuracy} = \frac{\text{True Positives} + \text{True Negatives}}{\text{Total Predictions}}$$

This formula highlights the model's effectiveness in correctly classifying URLs, emphasizing its reliability in detecting phishing threats in real-time applications. A high accuracy percentage indicates that the RNN and autoencoder model can generalize well to new data, making it a robust solution for combating phishing attacks in the ever-evolving digital landscape.

b. F1-score:

The F1 score is a crucial evaluation metric used in assessing the performance of classification models, especially in contexts where the distribution of classes is imbalanced, such as phishing detection. In such scenarios, traditional metrics like accuracy can be misleading, as they may fail to adequately reflect a model's ability to correctly identify minority classes, which in this case are phishing attempts. The F1 score combines both precision and recall into a single metric, providing a balanced view of a model's performance. Precision measures the proportion of true positive predictions among all positive predictions, indicating how many of the flagged URLs are genuinely phishing.

$$F1 = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}}$$

Interpretation:

- The F1 score ranges from 0 to 1, where 1 indicates perfect precision and recall (i.e., no false positives or false negatives).
- A higher F1 score suggests a better between precision and recall, making it a useful metric in the context of phishing detection, where both false positives and false negatives can have significant consequences.

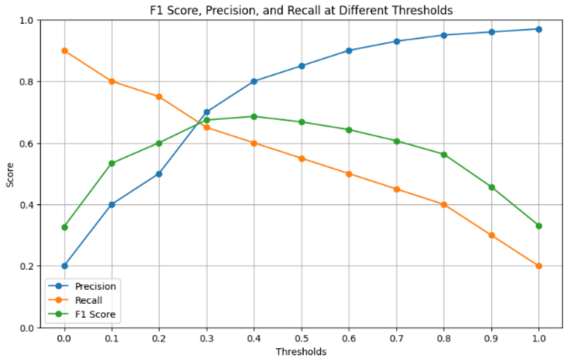
Recall, on the other hand, assesses the model's ability to identify actual phishing URLs among all legitimate and phishing instances. By harmonizing these two metrics, the F1 score offers a more nuanced understanding of a model's effectiveness in detecting phishing threats while minimizing false positives and negatives. This is particularly important in cybersecurity, where a high F1 score not only reflects accuracy in identifying threats but also helps maintain user trust and ensures a seamless online experience.

c. Precision:

The ratio of true positive predictions to the total predicted positives, which includes both true positives and false positives, is known as precision. This metric is critical in evaluating the effectiveness of a phishing detection model, as it quantifies the accuracy of the model's positive predictions. In the context of phishing detection, precision indicates the proportion of URLs flagged as phishing that are, in fact, legitimate threats.

$$\text{Precision} = \frac{TP}{TP + FP}$$

A high precision score signifies that the model is effective in distinguishing between phishing and legitimate URLs, resulting in fewer false positives—instances where legitimate URLs are incorrectly identified as phishing. This is particularly important for maintaining user trust; if users frequently encounter false positives, they may become skeptical of the system's reliability. Therefore, precision is a vital measure of a model’s performance, reflecting its ability to provide accurate phishing predictions while minimizing unnecessary disruptions for users.



d. Recall:

Recall is an essential metric for assessing the performance of classification models, especially in situations involving imbalanced datasets where the positive samples, such as phishing URLs, are significantly outnumbered by negative samples, like legitimate URLs. It quantifies the model's ability to accurately identify all relevant instances by focusing on true positives—those cases where the model correctly classifies a phishing attempt. In this context, recall provides insight into how effectively the model captures the positive class, highlighting its capacity to detect phishing threats while minimizing the risk of overlooking potentially harmful URLs.

$$\text{Recall} = \frac{\text{True Positives (TP)}}{\text{True Positives (TP)} + \text{False Negatives (FN)}}$$

Components of Recall:

- True Positives (TP): The number of instances that are correctly identified as positive (e.g., the number of phishing URLs correctly flagged as phishing).
- False Negatives (FN): The number of instances that are incorrectly identified as negative (e.g., legitimate URLs that were incorrectly flagged as safe but are actually phishing).

A high recall is particularly critical in security applications, where the consequences of missing a phishing attempt can lead to severe repercussions for users and organizations. By prioritizing the identification of true positives, recall ensures that the model is effective in safeguarding against cyber threats, even as phishing tactics evolve and become more sophisticated.

VI. CONCLUSION

In conclusion, the proposed deep learning-based approach that integrates Recurrent Neural Networks (RNNs) and autoencoders significantly enhances the detection of phishing URLs by addressing the limitations of traditional blacklisting methods. By leveraging RNNs' ability to analyze sequential data, the system effectively identifies intricate patterns in URL structures, thereby improving predictive accuracy and real-time responsiveness to emerging threats. The incorporation of autoencoders further refines the process by performing efficient feature extraction, allowing the model to focus on the most relevant characteristics of the data. Together, these technologies provide a proactive and robust solution capable of adapting to the ever-evolving tactics employed by phishing attackers, ultimately offering enhanced protection for users against cyber threats. This innovative approach not only improves detection rates but also lays a solid foundation for future developments in cybersecurity, highlighting the critical need for continuous advancements to combat sophisticated phishing techniques.

VII. FUTURE WORK

The scope for future work includes enhancing the model's robustness against adversarial attacks and exploring the integration of real-time monitoring systems for proactive threat detection. Additionally, incorporating more diverse datasets and leveraging transfer learning could improve model performance across various phishing tactics. Further research could also focus on developing user-friendly interfaces for better interaction and visualization of detection results, along with the implementation of collaborative detection frameworks to share insights across different platforms.

VIII. REFERNCES

[1] S. Remya, M. J. Pillai, K. K. Nair, S. Rama Subbareddy and Y. Y. Cho, "An Effective Detection Approach for Phishing URL Using ResMLP," in IEEE Access, vol. 12, pp. 79367-79382, 2024, doi: 10.1109/ACCESS.2024.3409049.

[2] A. Karim, M. Shahroz, K. Mustofa, S. B. Belhaouari and S. R. K. Joga, "Phishing Detection System Through Hybrid Machine Learning Based on URL," in IEEE Access, vol. 11, pp. 36805-36822, 2023, doi: 10.1109/ACCESS.2023.3252366.

[3] M. Ahmed et al., "PhishCatcher: Client-Side Defense Against Web Spoofing Attacks Using Machine Learning," in IEEE Access, vol. 11, pp. 61249-61263, 2023, doi: 10.1109/ACCESS.2023.3287226.

[4] R. Jayaraj, A. Pushpalatha, K. Sangeetha, T. Kamaleshwar, S. Udhaya Shree, Deepa Damodaran,Intrusion detection based on phishing detection with machine learning,Measurement: Sensors,Volume 31,2024,101003,ISSN2665-9174,https://doi.org/10.1016/j.measen.2023.101003.(https://www.sciencedirect.com/science/article/pii/S2665917423003392)

[5] Awasthi, A., Goel, N. Phishing website prediction using base and ensemble classifier techniques with cross-validation. Cybersecurity **5**, 22 (2022). https://doi.org/10.1186/s42400-022-00126-9

[6] L. Tang and Q. H. Mahmoud, "A Deep Learning-Based Framework for Phishing Website Detection," in IEEE Access, vol. 10, pp. 1509-1521, 2022, doi: 10.1109/ACCESS.2021.3137636.

[7] Shantanu, B. Janet and R. Joshua Arul Kumar, "Malicious URL Detection: A Comparative Study," 2021 International Conference on Artificial Intelligence and Smart Systems (ICAIS), Coimbatore, India, 2021, pp. 1147-1151, doi: 10.1109/ICAIS50930.2021.9396014.

[8] Alshingiti, Zainab, Rabeah Alaqel, Jalal Al-Muhtadi, Qazi Emad Ul Haq, Kashif Saleem, and Muhammad Hamza Faheem. 2023. "A Deep Learning-Based Phishing Detection System Using CNN, LSTM, and LSTM-CNN" Electronics 12, no. 1: 232. https://doi.org/10.3390/electronics12010232

[9] M. Sánchez-Paniagua, E. F. Fernández, E. Alegre, W. Al-Nabki and V. González-Castro, "Phishing URL Detection: A Real-Case Scenario Through Login URLs," in IEEE Access, vol. 10, pp. 42949-42960, 2022, doi: 10.1109/ACCESS.2022.3168681.

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

[11] P. Yang, G. Zhao and P. Zeng, "Phishing Website Detection Based on Multidimensional Features Driven by Deep Learning," in IEEE Access, vol. 7, pp. 15196-15209, 2019, doi: 10.1109/ACCESS.2019.2892066.

[12] M. Sameen, K. Han and S. O. Hwang, "PhishHaven—An Efficient Real-Time AI Phishing URLs Detection System," in IEEE Access, vol. 8, pp. 83425-83443, 2020, doi: 10.1109/ACCESS.2020.2991403.

[13] S. He, B. Li, H. Peng, J. Xin and E. Zhang, "An Effective Cost-Sensitive XGBoost Method for Malicious URLs Detection in Imbalanced Dataset," in IEEE Access, vol. 9, pp. 93089-93096, 2021, doi: 10.1109/ACCESS.2021.3093094.

[14] S. Marchal, J. François, R. State and T. Engel, "PhishStorm: Detecting Phishing With Streaming Analytics," in IEEE Transactions on Network and Service Management, vol. 11, no. 4, pp. 458-471, Dec. 2014, doi: 10.1109/TNSM.2014.2377295.

[15] M. Aljabri et al., "Detecting Malicious URLs Using Machine Learning Techniques: Review and Research Directions," in IEEE Access, vol. 10, pp. 121395-121417, 2022, doi: 10.1109/ACCESS.2022.3222307.

[16] A. C. Bahnsen, E. C. Bohorquez, S. Villegas, J. Vargas and F. A. González, "Classifying

phishing URLs using recurrent neural networks," 2017 APWG Symposium on Electronic Crime Research (eCrime), Scottsdale, AZ, USA, 2017, pp. 1-8, doi: 10.1109/ECRIME.2017.7945048.

[17] Weiping Wang, Feng Zhang, Xi Luo, Shigeng Zhang, and Angel M. Del Rey. 2019. PDRCNN: Precise Phishing Detection with Recurrent Convolutional Neural Networks. Sec. and Commun. Netw. 2019 (2019). https://doi.org/10.1155/2019/2595794

[18] Hassan Y.A. Abutair, Abdelfettah Belghith,Using Case-Based Reasoning for Phishing Detection,Procedia Computer Science,Volume 109,2017,Pages 281-288,ISSN 1877-0509,https://doi.org/10.1016/j.procs.2017.05.352(https://www.sciencedirect.com/science/article/pii/S1877050917310219)

[19] Xi Xiao, Dianyan Zhang, Guangwu Hu, Yong Jiang, Shutao Xia,CNN–MHSA: A Convolutional Neural Network and multi-head self-attention combined approach for detecting phishing websites,Neural Networks,Volume 125,2020,Pages 303-312,ISSN 0893-6080,https://doi.org/10.1016/j.neunet.2020.02.013.(https://www.sciencedirect.com/science/article/pii/S0893608020300587)

[20] Rami M. Mohammad, Fadi Thabtah, and Lee Mccluskey. 2014. Predicting phishing websites based on self-structuring neural network. Neural Comput. Appl. 25, 2 (August 2014), 443–458. https://doi.org/10.1007/s00521-013-1490-z

[21] A. Subasi, E. Molah, F. Almkallawi and T. J. Chaudhery, "Intelligent phishing website detection using random forest classifier," 2017 International Conference on Electrical and Computing Technologies and Applications (ICECTA), Ras Al Khaimah, United Arab Emirates, 2017, pp. 1-5, doi: 10.1109/ICECTA.2017.8252051.