IJCRT.ORG ISSN : 2320-2882

INTERNATIONAL JOURNAL OF CREATIVE RESEARCH THOUGHTS (IJCRT)

An International Open Access, Peer-reviewed, Refereed Journal

Malicious URL Detection: Methods, Challenges, And Comparative Insights

Divya N [1*], Dr. Bhagyajyothi K L[1], Dr. Divya A K[1], Prof. Naseema C A[1]

[1]Department of Computer Science, K.V.G College of Engineering, Sullia, D. K., Karnataka, India.

Abstract— Phishing, a prevalent cyber threat, involves the deceptive attempt to obtain sensitive information such as usernames, passwords, and financial details by masquerading as a trustworthy entity in electronic communication. With the exponential growth of online activities, phishing has emerged as a significant concern, posing serious security risks to individuals, organizations, and society at large. The impacts of phishing are multifaceted, ranging from financial losses and identity theft to reputational damage and compromised cybersecurity infrastructures. Addressing problem of phishing necessitates robust detection and prevention mechanisms. This paper is focuses on the classification of phishing URLs, a crucial aspect of mitigating the phishing threat. By leveraging machine learning and data mining techniques, the aim is to develop efficient algorithms capable of accurately identifying and categorizing phishing URLs. This project seeks to contribute to the ongoing efforts in enhancing cybersecurity and safeguarding users against the perils of phishing attacks. The research underscores the DNN's potential as a robust solution for real-world phishing detection systems. Its capacity to effectively navigate the complexities of diverse URLs positions it as a promising ally in the ongoing battle against cyber threats. The findings contribute significantly to the advancement of cybersecurity measures, emphasizing the practical efficacy of the DNN model. By shedding light on its consistent performance, this study underscores the DNN's pivotal role in mitigating the ever-growing menace of phishing attacks, offering a beacon of hope in fortifying digital landscapes against evolving cybersecurity challenges.

Keywords— Phishing, cybersecurity, neural network models, Deep Neural Network (DNN), diverse dataset, malicious URLs, legitimate URL.

1. INTRODUCTION

The word “phishing” is believed to have first appeared around 1995, when it was included in the hacker toolkit known as AOHell. Some accounts suggest that the expression might have been circulating even earlier in underground hacker circles, such as those connected with the magazine 2600. The term itself is derived from the idea of “fishing,” where bait is cast to trick someone into giving up valuable information—in this case, sensitive data like passwords or financial details. Efforts to counter phishing attacks generally fall into four main areas: law enforcement, user training, public awareness campaigns, and stronger technical safeguards. Over the past few years, the significance of phishing awareness has grown sharply in both personal and workplace contexts. For example, reports show that the percentage of businesses facing phishing attempts increased from 72% in 2017 to 86% in 2020. By 2021, the average click rate for a phishing email stood at 17.8%, but when attackers combined email with follow-up phone calls, the rate jumped to 53.2%, making such campaigns nearly three times more successful. Large-scale scans of millions of web addresses have also revealed that about 12% of suspicious URLs carried malware.

In today’s digital world, phishing has become one of the most persistent and dangerous cybersecurity threats. It works by tricking people into believing they are interacting with a legitimate source, only to steal sensitive details such as login credentials or financial information. What makes phishing particularly challenging is its constantly evolving nature—attackers continuously change their strategies, making it hard for traditional defense methods to keep up. This reality highlights the urgent need for smarter, adaptable, and more resilient security measures to protect individuals, organizations, and even critical infrastructure from these increasingly sophisticated attacks.

Recent advancements in intelligent transportation systems have shown promise in enhancing road safety and traffic management. For instance, Smith et al. (Year) proposed a Convolutional Neural Network (CNN) integrated with the YOLO framework to detect, classify, and count vehicles from videos captured under various climatic conditions. Their approach achieved an impressive average accuracy of 94.4% in vehicle detection, classification, and counting, showcasing the potential of deep learning techniques in addressing complex transportation challenges. With the growing need for immediate medical interventions, especially in rural areas lacking adequate healthcare facilities, telemedicine and efficient patient transportation have gained significance. Jones et al. (Year) introduced an IoT-based mobile medical edge (IM2E) node for continuous monitoring of emergency patients during ambulance transport. Their work addresses key challenges such as real-time risk assessment and dynamic medical interventions, underscoring the importance of technology-enabled solutions in improving healthcare delivery in emergency situations. Ensuring driver safety and preventing vehicular accidents remain critical concerns in modern society. Brown et al. (Year) proposed an innovative approach, the Unsafe Driving Detection System Using IoT (UDDSUI), aimed at mitigating accidents caused by driver errors and fatigue. By monitoring factors such as proximity to other vehicles, alcohol intake, and driver exhaustion, the UDDSUI system issues timely warnings to drivers and stores relevant data for analysis, emphasizing the role of IoT in enhancing road safety measures [1] [2] [3]. Mobile sink-driven data acquisition has emerged as a promising approach to enhance efficiency and overcome hotspot issues in wireless sensor networks (WSNs). For instance, A bug algorithm based on an obstacle-aware trajectory (CSOBUG) for efficient data acquisition by mobile sinks in WSNs. Their approach integrates cat swarm optimization (CSO) for constructing obstacle-aware trajectories, achieving superior performance in terms of throughput, lifetime, energy efficiency, delay, and packet loss compared to existing techniques. Energy harvesting techniques offer a solution to extend the lifespan of batteries in wireless sensor networks (WSNs). Investigated the classification of energy-efficient sensor nodes for EHWSNs using machine learning algorithms such as KNN and SVM. Their study evaluates the effectiveness of different sensor nodes and energy harvesters, providing valuable insights for optimizing energy utilization and prolonging the operational lifespan of WSNs in remote or inaccessible areas. In the field of materials science, (Year) employed a simple solution combustion technique to synthesize $SrCeO_3$ nanopowders doped with europium ions (Eu^{3+}). Through detailed characterization using various analytical techniques, they observed promising properties such as orthorhombic structure, needle-like morphology, and efficient photoluminescence with red emission. These nanopowders exhibit potential applications in display devices and advanced forensic technologies, highlighting their significance in both scientific and practical domains [4][5][6].

Research shows that phishing campaigns are primarily driven by data theft (85%), followed by financial motives (26%). A well-documented case occurred in Q3 2023, when MGM Resorts International, a leading U.S. hospitality and entertainment company, fell victim to a sophisticated attack.

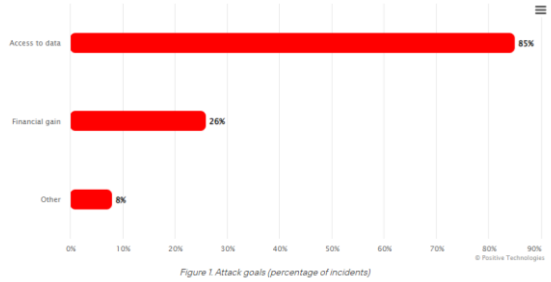

Fig. 1: Attacks Goals according to year 2023

The incident unfolded in several stages. It began with a fake hotel reservation, after which attackers replied to the confirmation email. They then sent follow-up messages crafted to create urgency and appeal to the recipient's emotions. Once trust was established, the attackers delivered a malicious URL disguised as a link to important documents. When the victim accessed the link, malware was downloaded, infiltrating the system and stealing sensitive data. Such stolen information was later exploited, with compromised hotel profiles being leveraged to target customers directly.

To deliver these malicious payloads, compressed file formats (zip, rar, 7z, etc.) are commonly used (37%). These files are difficult for many security tools to scan thoroughly, allowing attackers to conceal malware within what appears to be a harmless document or image. Because employees are generally familiar with archives, they are more likely to open them without suspicion.

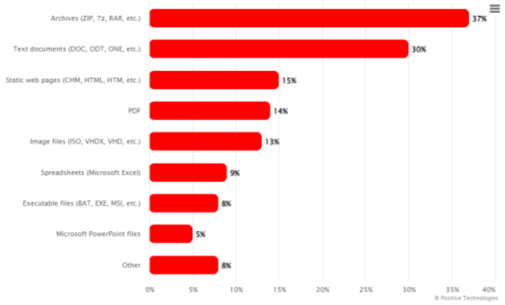

Fig. 2: Types of attachments

Another favored tactic is the use of static web pages instead of PDFs in phishing emails. This approach often relies on HTML smuggling, a technique that embeds encoded malicious code into an HTML attachment or page. When the file is opened, the victim's browser decodes the script and reconstructs the payload locally—bypassing the need to send an executable over the network and slipping past firewalls. A notable case in July 2023 involved a phishing campaign targeting 615 employees with personalized HTML attachments that appeared to be salary payment notifications. The encoded script avoided detection by email scanners, tricking recipients into opening it.

HTML smuggling has proven effective not only in stealing credentials but also in deploying malware, making it one of the more resilient phishing techniques. In response to the rising use of such methods, organizations have begun strengthening defenses. For example, Microsoft introduced a policy to block macros in Office files from the internet by default, forcing attackers to increasingly rely on malicious links instead of direct file delivery.

This paper embarks on comprehensive exploration of state-of-the-art neural network models, specifically Convolutional Neural Network (CNN), Recurrent Neural Network (RNN), Deep Belief Network (DBN), Deep Neural Network (DNN), and Long Short-Term Memory (LSTM), in the specific context of phishing URL detection. The overarching objective is to meticulously evaluate the effectiveness of these sophisticated models in discerning phishing URLs from legitimate ones, aiming to contribute to the development of advanced and proactive defense mechanisms against cyber threats. The introduction serves as the foundational chapter, contextualizing the gravity of the phishing threat within the broader cybersecurity landscape. It acknowledges the limitations of conventional security measures in addressing the nuanced and evolving tactics employed by cyber adversaries in orchestrating phishing campaigns. These campaigns often exploit human vulnerabilities, capitalizing on factors such as social engineering, deceptive links, and sophisticated spoofing techniques, thereby necessitating a more advanced and adaptive approach to detection. As technology advances, cybercriminals exploit increasingly sophisticated methods, necessitating a deeper exploration of neural network models capable of learning intricate patterns within URL data. Understanding the nuanced features is essential for improving the precision and effectiveness of detection systems linked to phishing URLs. The techniques utilised, the neural network model architectures that are being examined, the training and evaluation datasets, and—most importantly—the comparison of CNN, RNN, DBN, DNN, and LSTM models.

In this application, we prioritize the implementation of SSL vendor information to enhance the security and transparency of data transmissions. By retrieving details such as "Issued From," "Issued To," certificate issuance date, and expiration date, our system ensures a comprehensive overview of the SSL certificate's validity and origin. This approach not only contributes to a secure communication environment but also empowers users and administrators with essential insights into the credibility and timeframe of SSL certificates, reinforcing trust in the integrity of data exchanges within our application.

In a typical phishing attempt, the attacker carefully gathers details about the target. This may begin with basic information such as the person's name, job title, or workplace, and then extend to more private data, including hobbies, favorite places, financial details, email addresses, phone numbers, or even social media activity on platforms like Facebook and Twitter. Because phishing is widespread and constantly evolving, it cannot be eliminated overnight—and given its persistence, it is unlikely to vanish in the foreseeable future. Tackling this issue requires comprehensive and ongoing research.

Recognizing the seriousness of phishing threats has motivated researchers to explore the use of advanced machine learning models for detection. The solution proposed in this project is a web-based application where users input a website's URL. The system then analyzes and extracts relevant features from the URL and applies deep learning techniques to determine whether the site is legitimate or malicious.

This study is organized to present a thorough examination of how advanced neural network models can be applied to identify malicious URLs effectively. The paper opens with an introduction that emphasizes the growing impact of phishing threats on cybersecurity. Following this, the methodology section details the step-by-step approach used, including URL feature extraction, evaluation of SSL certificates, and the subsequent classification of potential threats. A comparison of the Convolutional Neural Network (CNN), Recurrent Neural Network (RNN), Deep Belief Network (DBN), Deep Neural Network (DNN), and Long Short-Term Memory (LSTM) models is then provided, backed up by comparison graphs and classification report tables. The findings are examined in detail in the results and discussion, which highlight the DNN model's superiority. The conclusion highlights the DNN's potential to strengthen digital defences against phishing attacks and provides a summary of important findings.

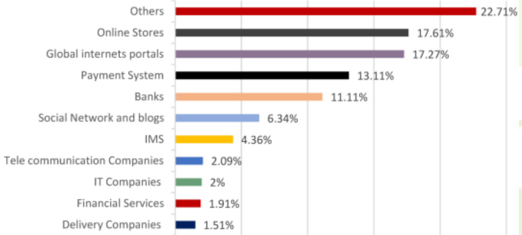

Fig. 3: Percentage of phishing attack organizations

Identifying phishing websites is a challenging task that requires intelligent techniques capable of analyzing and classifying their features to determine whether a site is genuine or fraudulent. Because phishing often relies on social engineering and involves subtle, complex patterns, traditional AI methods struggle to deliver accurate and efficient detection. Developing an intelligent system for this purpose is therefore critical. Such a system must achieve very high accuracy and minimize errors, since even a small mistake—such as misclassifying a fraudulent website as legitimate—can have serious consequences.

This study not only seeks to assess the performance of these models but also aims to shed light on why one model, the Deep Neural Network (DNN), emerges as the most accurate in distinguishing between legitimate and malicious URLs. Through this exploration, the research contributes significantly to the ongoing efforts to fortify digital landscapes against the ever-evolving tactics of cyber adversaries. Emphasizing the pivotal role of advanced neural network models in contemporary cybersecurity, this paper endeavors to provide insights that propel the development of robust, intelligent, and adaptive defenses in the face of the growing menace of phishing attacks.

2. RELATED WORK

Gold Wejinya et al. introduced the MuD model, which integrates three supervised machine learning classifiers—Support Vector Machine (SVM), Logistic Regression, and Naïve Bayes—to improve the detection of malicious URLs. Among these, Naïve Bayes produced the most accurate results [3]. Similarly, Do Xuan et al. emphasized the importance of URL behavior and attribute analysis, incorporating big data technologies to enhance detection capability against abnormal patterns. Their work demonstrated that combining feature engineering, machine learning models, and big data can significantly boost detection accuracy, making their approach practical and efficient [4].

B. Janet et al. framed URL classification as a binary classification problem, testing several machine learning classifiers using a Kaggle dataset of 450,000 URLs. Their results showed that the chosen best-performing classifier was effective in identifying malicious links from sources such as OpenPhish [5]. Aljabri et al. provided an extensive review of machine learning-based malicious URL detection, identifying the gaps in current studies, including challenges in detecting Arabic phishing websites, and suggesting future research directions [6].

JT Yuan et al. proposed a joint neural network model that combines the attention mechanism, bidirectional independent recurrent neural networks, and capsule networks, achieving strong detection performance [7]. In parallel, S. Abad et al. applied traditional algorithms such as SVM, Random Forest, Decision Trees, and kNN, enhanced with Bayesian optimization and instance selection techniques to improve both accuracy and computational efficiency [8]. Tie Li et al. combined linear and nonlinear transformation methods—using singular value decomposition and kernel approximation—to improve distance metrics for URL classification [9].

Further contributions include Yu-Chen Chen et al., who designed a model using 41 extracted URL features with ANOVA testing and XGBoost, achieving over 99% accuracy [10]. Mohammed Alsaedi et al. built a two-stage ensemble model combining Random Forest for preliminary classification and Multilayer Perceptron (MLP) for final decision-making [11]. Zhiqiang Wang et al. and Dipankar Kumar Mondal et al. explored the role of character-based word embeddings for URL representation, showing higher accuracy compared to traditional embedding techniques [12][13].

Swagat M. Karve et al. offered a structured survey that examined feature representation and algorithm design, providing researchers and practitioners with insights into effective malicious URL detection [14]. Malak Aljabri et al. compared ML and DL models, finding Naïve Bayes to be the most accurate in their dataset (96%) while emphasizing the importance of feature engineering [15]. Clayton Johnson et al. compared traditional ML methods like Random Forest and kNN against deep learning models (Fast.ai, TensorFlow/Keras) across different architectures (CPU, GPU, TPU) using the ISCX-URL-2016 dataset, showing how deep learning can outperform conventional approaches in certain contexts [16].

Overall, the literature reveals that while attackers constantly evolve their techniques, machine learning and deep learning provide effective countermeasures by learning from URL patterns and behaviors. Since phishing URLs are often constructed randomly, URL-based detection methods offer a safer and more scalable approach than visiting or analyzing the malicious websites directly.

In summary, reviewing prior work shows that researchers have developed diverse solutions, from lightweight machine learning classifiers to advanced deep neural models, all aimed at improving URL detection accuracy. Many online tools now exist that can quickly analyze a suspicious link entered by a user. Building on this foundation, our system applies machine learning to scan and classify URLs in real time, offering a fast, one-click solution to identify whether a link—be it from email, social media, or another source—is safe or malicious.

4. Research Gap

The field of cybersecurity is always changing as adversaries use more advanced tactics. There is still a significant research gap in the area of malicious URL detection, despite significant advancements in the field. Although somewhat successful, current approaches frequently fall behind the ever-evolving strategies used by cybercriminals. The difficulty is in identifying subtle, quickly changing patterns that point to malicious URLs.

Conventional methods might not be flexible enough to adjust to new threats, which could cause a delay in detection. Furthermore, detecting malicious URLs gets harder as attackers continue to hone their tactics. By investigating how neural network models, in particular the Deep Neural Network (DNN), can improve the precision and speed of malicious URL detection, this study seeks to close this gap. Closing this gap is essential to creating proactive defences that can successfully combat the constantly changing cyberthreat landscape.

3. Proposed Methods

The architecture of the proposed system is designed to detect malicious websites through a structured and multi-layered process that begins when a user submits a URL. Once entered, the system immediately performs feature extraction and analysis. This involves breaking down the URL into measurable components, including domain length, presence of special characters, and structural layout. In parallel, the system evaluates the SSL certificate linked to the site, checking details such as the issuer, expiration period, and certificate validity, which are critical indicators of trustworthiness.

To strengthen detection, the system combines traditional attributes (hostname, IP address, SSL details, and URL depth) with advanced machine learning methods. A supervised learning model is trained using a labeled dataset that contains both legitimate and phishing URLs. To improve reliability and prevent overfitting, feature selection techniques such as recursive feature elimination (RFE) or principal component analysis (PCA) can be applied, ensuring only the most significant attributes are retained. Additionally, ensemble learning strategies like bagging and boosting are incorporated, leveraging multiple classifiers to achieve higher accuracy and robustness.

Continuous refinement plays a vital role in the framework. The system undergoes cross-validation and hyperparameter tuning to maximize key performance indicators, including accuracy, precision, recall, and F1-score. Beyond training, the architecture also emphasizes real-time monitoring and adaptive feedback loops, allowing the model to evolve alongside new and sophisticated phishing tactics.

In essence, this design integrates comprehensive feature extraction, advanced machine learning, and rigorous evaluation to build a resilient phishing detection framework. By doing so, it not only classifies URLs with high accuracy but also provides an adaptive shield against the ever-changing landscape of phishing threats.

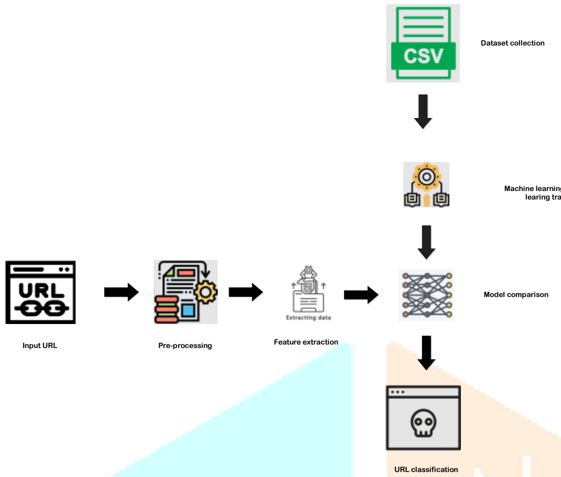

Fig. 4: System Architecture

The system also considers the depth of the URL within the website structure, evaluating its hierarchical position. This depth analysis aids in understanding the potential complexity and strategic placement of the threat within the website. Finally, employing machine learning models, particularly focusing on a Deep Neural Network (DNN), the system classifies the input URL as either malicious or legitimate based on the amalgamation of extracted features and their corresponding threat levels. The user is promptly notified of the outcome, ensuring transparency and empowering users with information about the security status of the websites they engage with. This holistic architecture underscores the system's adaptability and efficiency in countering the diverse and evolving landscape of malicious online activities.

SSL (Secure Socket Layer):

In our project, we have implemented a series of steps akin to the SSL certificate acquisition process to fortify our defenses against malicious URLs. Just as in the SSL workflow, our system begins by generating a unique identifier for each URL, akin to a Certificate Signing Request (CSR), to initiate the validation process. This identifier facilitates rigorous verification steps, similar to those performed by trusted Certificate Authorities Through this process, we confirm the authenticity of URLs and ensure they aren't malicious or deceptive. By validating control over domains and cross-referencing with authoritative records, we emulate the verification steps crucial for SSL certificates. Upon successful validation, URLs deemed safe receive a unique identifier, akin to a public key, which is then integrated into our system for future reference. This methodology mirrors the SSL certificate issuance process, thereby bolstering our ability to identify and mitigate potential threats posed by malicious URLs within our project.

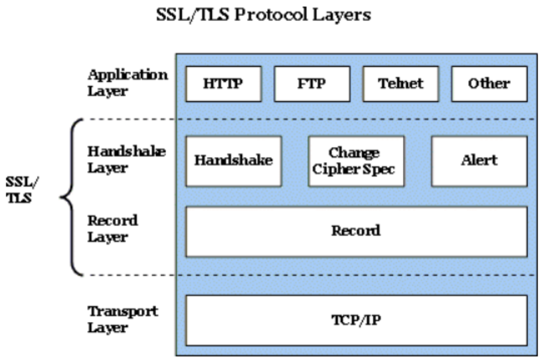

Fig. 5: SSL working architecture

Record, Handshake, Change Cipher Spec & Alert are four different protocols that Secure Socket Layer defines to accomplish its operations. Let us have a detailed look on how SSL work to accomplish its task:

The Handshake Protocol: This protocol provides security parameters for the Record protocol. It uses a message to negotiate the cipher suite and provides keys and security parameters. It also authenticates the server to client and client to server whenever needed. It exchanges information for building the cryptographic secrets. The Handshaking is done in 4 phases: Establishing Security Capabilities, Server Authentication & Key Exchange, Client Authentication & Key Exchange, and Finalizing the Handshake Protocol.

The Change Cipher Spec Protocol: Cipher Suite & Cryptographic Secrets are not used by the client or the server until Change Cipher Spec message is not delivered by them. Before exchanging this message only pending columns will have values.

The Alert Protocol: SSL uses this protocol for reporting any abnormal condition or error. It contains only message type and the alert message that describe the level of problem (warning or fatal).

The Record Protocol: This protocol carries the message from the upper layer. The message is fragmented and then compressed (if required). MAC address is then added using negotiated hash algorithm. Fragmented message and MAC are then encrypted by using a lossless negotiable encrypted algorithm and message blocks are framed by adding headers.

SSL/TLS Encryption and Keys:

Asymmetric keys – The public and private key pair are used to identify the server and initiate the encrypted session. The private key is known only to the server, while the public key is shared via a certificate.

Symmetric session keys – Disposable keys are generated for each connection and used to encrypt/decrypt transmitted data. The symmetric keys are securely exchanged using asymmetric encryption.

SSL/TLS supports multiple symmetric ciphers and asymmetric public key algorithms. For example, AES with 128-bit keys is a common symmetric cipher, while RSA and ECC commonly use asymmetric algorithms.

Dataset Collection:

As part of our extensive endeavour to strengthen cybersecurity, we have meticulously selected datasets of total 970 instances each of three distinct malicious URL types: Social Engineering, Spoofing, and Phishing websites and Legitimate.

Table 1: Dataset information

Types	Counts
Social Engineering websites	300
Spoofing websites	250
Phishing websites	220
Legitimate	200
Total	970

Our system can learn and adjust to different attack vectors thanks to this careful collection, which guarantees a varied and representative range of malicious activities. Our strategy improves the effectiveness of our cybersecurity measures by concentrating on these particular categories, which enables targeted analysis and detection. The basis for training and validating our models to strengthen against the constantly changing landscape of cyber threats is this dataset, which has been painstakingly curated with attention to the subtleties of each malicious type.

Convolutional Neural Network (CNN):

CNNs excel in capturing spatial patterns within data. In the context of URL detection, they can effectively recognize and learn hierarchical features present in the structure of URLs, allowing them to identify distinctive patterns associated with phishing or malicious links.

Recurrent Neural Network (RNN):

RNNs are designed to capture sequential dependencies within data. In the case of URL detection, RNNs can analyze the sequential nature of characters in URLs, capturing patterns that may be indicative of malicious intent or phishing attempts.

Deep Belief Network (DBN):

DBNs are generative models capable of capturing complex hierarchical relationships in data. In URL detection, DBNs can learn and represent intricate features within URLs, aiding in the identification of malicious patterns.

Deep Neural Network (DNN):

DNNs are deep-layered networks designed for feature abstraction. In the URL context, DNNs can autonomously learn and extract abstract features from URLs, enabling them to discern between benign and malicious links with a high degree of accuracy.

Long Short-Term Memory (LSTM): LSTMs are a type of recurrent neural network with specialized memory cells. In URL detection, LSTMs can effectively capture long-term dependencies within sequences of characters, enabling the model to recognize patterns that may be indicative of phishing or malicious URLs.

5. RESULTS AND DISCUSSION

Five sophisticated neural network algorithms—Convolutional Neural Network (CNN), Recurrent Neural Network (RNN), Deep Belief Network (DBN), Deep Neural Network (DNN), and Long Short-Term Memory (LSTM)—are shown in the comparison graph in terms of how well they classify malicious URLs. Each algorithm's accuracy in distinguishing between legitimate and malicious URLs is presented, offering a visual representation of their comparative strengths. The Convolutional Neural Network (CNN) demonstrates notable efficiency in capturing spatial patterns within URL data, showcasing competitive accuracy. Meanwhile, Recurrent Neural Network (RNN) leverages sequential dependencies for nuanced understanding, yielding commendable results. Deep Belief Network (DBN) showcases a strong performance in learning hierarchical relationships, contributing to effective classification.

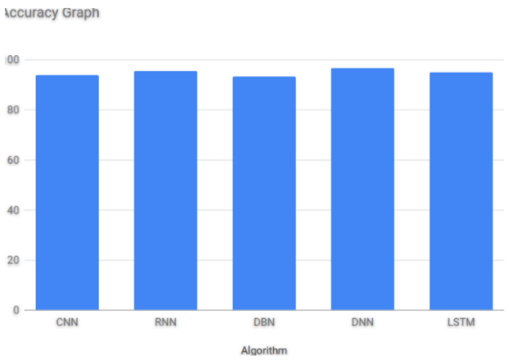

Fig. 5: Comparison Graph

The graph emphasizes the remarkable accuracy of the Deep Neural Network (DNN), positioning it as a standout performer among the algorithms. Its ability to autonomously extract abstract features from URLs proves to be a significant factor in achieving high accuracy. Long Short-Term Memory (LSTM), with its specialized memory cells, also exhibits commendable performance, particularly in capturing long-term dependencies within URL sequences. In essence, the comparison graph provides valuable insights into the relative strengths of each algorithm, guiding the selection of the most effective model for classifying malicious URLs in real-world cybersecurity applications. The high accuracy achieved by the DNN underscores its potential as a powerful tool in the ongoing efforts to fortify digital landscapes against the ever-evolving threat of phishing attacks.

The bar graph shows how well various models perform in comparison across a range of research papers in the field of malicious URL detection. Each bar shows the accuracy attained by a particular model, giving an illustration of how effective they are.

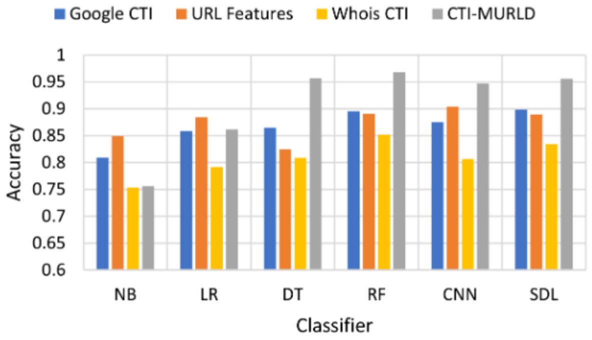

Fig. 6: Comparison in terms of the detection-accuracy performance.

By analyzing multiple papers, the graph offers insights into the consistency and reliability of these models in different experimental contexts. This comparison aids in identifying trends and determining the most effective approaches for robust malicious URL detection in cybersecurity applications.

The classification report table presents a detailed overview of the performance metrics for each proposed model in context of malicious URL detection. The table includes precision, recall, F1-score, and accuracy metrics, offering a comprehensive evaluation of the models' capabilities. The precision metric signifies the ratio of correctly predicted positive instances to the total instances predicted as positive, providing insights into the models' accuracy in identifying malicious URLs. Recall, on the other hand, represents the ratio of correctly predicted positive instances to the total actual positive instances, indicating the models' sensitivity to identifying malicious URLs.

The following table presents the set of features extracted from URLs to support the classification of phishing activities. Each feature serves as an indicator that helps in evaluating whether a URL is legitimate or malicious. Together, these attributes form the foundation for identifying suspicious patterns and strengthening the accuracy of phishing detection. Below is a concise description of each feature:

Table 2: Features extracted from URL

URL Features	Information
Hostname	https://www.randilion.com
IP address	192.168.29.18
Severity	8.11839771919782
Potential threat	7.6159073890
Level	1
Depth	5
Issuer (SSL)	Lets Encrypt
Issued to (SSL)	randilion
License (SSL)	0337D636D278B250B447
Valid from (SSL)	Feb 2024
Valid to (SSL)	May 2024

Hostname: Refers to the domain name extracted from the URL, serving as the primary identifier of the website.
IP Address: The unique numerical label assigned to the server hosting the website, which

helps in identifying its geographical or network location.

Severity: Represents the threat level posed by the URL, categorized on a predefined scale ranging from low to high.

Potential Threat: An evaluation of the probability and possible consequences of the URL being linked to phishing activities, aiding in overall risk assessment.

Level: The degree of risk involved in visiting the URL, commonly classified into categories such as low, medium, or high.

Depth: Indicates the number of subdirectory layers within the URL’s structure, reflecting the complexity of the webpage hierarchy.

Issuer (SSL): The certificate authority or organization that issued the SSL certificate, confirming the authenticity of secure communication.

Issued to (SSL): The individual or organization to whom the SSL certificate was granted, verifying the ownership and legitimacy of the website.

License (SSL): Specifies the type of SSL validation (e.g., domain, organization, or extended validation) and the corresponding level of encryption.

Valid From (SSL): The starting date when the SSL certificate becomes active and trusted for secure communications.

Valid To (SSL): The expiration date of the SSL certificate, after which secure communication can no longer be guaranteed.

This comprehensive table assists in the systematic analysis and classification of URLs, aiding in the identification and mitigation of phishing threats.

Table 3: Classifications reports of models

Model	Precision	Recall	F1 Score
CNN	0.87	0.86	0.92
DBNS	0.91	0.90	0.95
DNN	0.89	0.88	0.93
LSTM	0.85	0.91	0.91

Accuracy: Accuracy is a metric that measures how often a machine learning model correctly predicts the outcome. You can calculate accuracy by dividing the number of correct predictions by the total number of predictions.

$$\text{accuracy} = (TP + TN) / (TP + TN + FP + FN)\ldots(1)$$

Precision: In the confusion matrix in the preceding illustration, these metrics are calculated in the following way:

$$\text{Precision} = TP \div (TP + FP)\ \ldots(2)$$

Recall: Represents the number of positive instances correctly identified by the model.

$$\text{Recall} = TP \div (TP + FN)\ldots(3)$$

For the purpose of summarising model performance, the counts of true positive (TP), false positive (FP), false negative (FN), and true negative (TN) ground truths and inferences are crucial. These metrics are the building blocks of many other metrics, including accuracy, precision, and recall Metric.

The F1-score represents the harmonic mean of precision and recall, offering a balanced indicator of a model’s performance by accounting for both false positives and false negatives. In contrast, accuracy reflects the percentage of correctly classified samples out of the total dataset, providing an overall measure of effectiveness. Together, these metrics form the basis of the classification report table, which serves as a critical resource for evaluating and comparing different models. By analyzing these results, researchers and practitioners can make well-informed decisions when selecting the most effective algorithm for reliable malicious URL detection in real-world cybersecurity scenarios.

Confusion Matrix:

In this project, multiple machines learning algorithms, including Convolutional Neural Networks (CNN), Deep Belief Networks (DBN), Deep Neural Networks (DNN), Long Short-Term Memory (LSTM) networks, and Recurrent Neural Networks (RNN), were employed for phishing URL classification. The confusion matrix serves as a vital evaluation tool for each algorithm. It presents a tabular summary of the model's performance, showing the number of correct and incorrect predictions for each class. Analyzing the confusion matrix provides insights into the strengths and weaknesses of each algorithm in accurately classifying phishing URLs, aiding in model selection, optimization, and overall improvement of the phishing detection system. The intensity of color in each cell corresponds to the frequency or proportion of instances, offering insights into the model's performance across different classes.

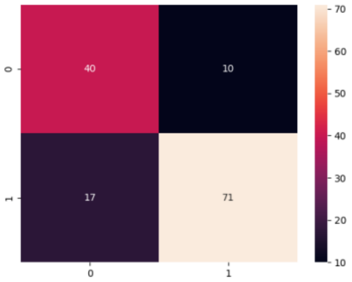

Fig. 7: RNN confusion matrix

The heatmap in Figure 7 illustrates the confusion matrix generated from the recurrent neural network (RNN) model's classification of phishing URLs. Each cell in the heatmap represents the frequency of instances where the actual class labels differ from the predicted ones. The diagonal cells depict correct predictions, while off-diagonal cells indicate misclassifications.

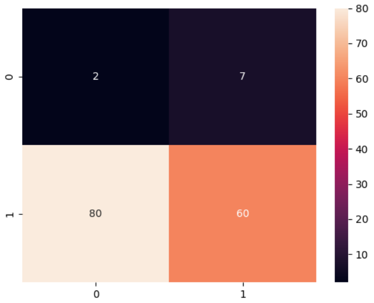

Fig. 8: CNN confusion matrix

The heatmap in Figure 8 displays the confusion matrix generated by the convolutional neural network (CNN) model's classification of phishing URLs. Each cell in the heatmap represents the frequency of instances where the predicted class labels deviate from the actual ones.

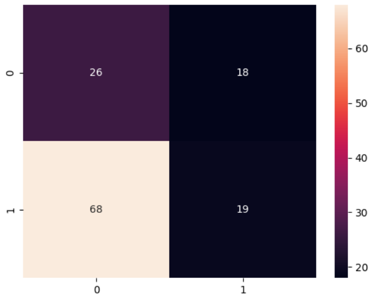

Fig. 9: DBN confusion matrix

The heatmap in Figure 9 illustrates the confusion matrix resulting from the deep belief network (DBN) model's classification of phishing URLs. Each cell in the heatmap represents the frequency of instances where the predicted class labels differ from the actual ones.

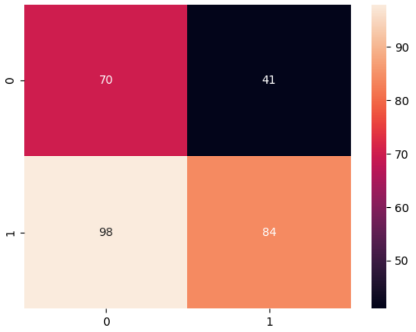

Fig. 10: DNN confusion matrix

The heatmap in Figure 10 illustrates the confusion matrix generated by the deep neural network (DNN) model's classification of phishing URLs. Each cell in the heatmap represents the frequency of instances where the predicted class labels differ from the actual ones.

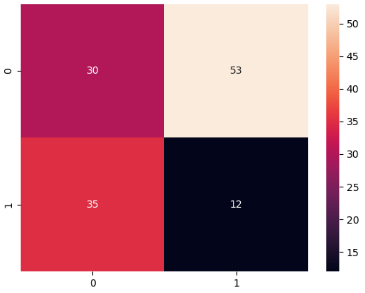

Fig. 11: LSTM confusion matrix

The heatmap in Figure 11 depicts the confusion matrix derived from the long short-term memory (LSTM) model's classification of phishing URLs. Each cell represents the frequency of instances where the predicted class labels deviate from the actual ones.

6. CONCLUSION

In conclusion, our study underscores the pivotal role of advanced neural network models in the realm of malicious URL detection. The comprehensive evaluation of Long Short-Term Memory (LSTM), Convolutional Neural Network (CNN), Recurrent Neural Network (RNN), Deep Belief Network (DBN), and Deep Neural Network (DNN) all demonstrate how much more accurate the Deep Neural Network. This finding positions the DNN as a robust solution for real-world phishing detection systems. The study contributes valuable insights into the nuanced features associated with malicious URLs, emphasizing the efficacy of advanced models in countering the evolving tactics of cyber adversaries. As the cyber threat landscape continues to advance, the DNN emerges as a promising ally in fortifying digital defenses, showcasing its potential as a cornerstone in the ongoing efforts to mitigate the ever-growing menace of phishing attacks.

REFERENCES

1. Sunny, Alfina, and N. Manohar. "Detection, Classification and Counting of Moving Vehicles from Videos." In International Conference on DATA ANALYTICS & LEARNING, pp. 231-242. Singapore: Springer Nature Singapore, 2022.

2. Mukhopadhyay, Adwitiya, Aryadevi Remanidevi Devidas, Venkat P. Rangan, and Maneesha Vinodini Ramesh. "A QoS-Aware IoT Edge Network for Mobile Telemedicine Enabling In-Transit Monitoring of Emergency Patients." Future Internet 16, no. 2 (2024): 52.

3. Tushaar, T. H., B. Shashank, Tanvi K. Jois, and Adwitiya Mukhopadhyay. "UDDSUI: An Unsafe Driving Detection System Using IoT." In 2022 IEEE 19th India Council International Conference (INDICON), pp. 1-6. IEEE, 2022.

4. Gowthami, Dasari, Ebenezer Jangam, and Pallavi Joshi. "Intelligent Trajectory for Mobile Element in WSNs with Obstacle Avoidance." Contemporary Mathematics (2024): 157-174.

5. Srinivas, Haripriya R., Priyanka Mohan, and Pallavi Joshi. "Detection of Energy Efficient Sensor Node in EHWSN." In 2023 World Conference on Communication & Computing (WCONF), pp. 1-6. IEEE, 2023.

6. Majani, Sanjay S., R. B. Basavaraj, K. N. Venkatachalaiah, Thalari Chandrasekhar, and Shiva Prasad Kollur. "Versatile deep red-emitting SrCeO3: Eu3+ nanopowders for display devices and advanced forensic applications." Journal of Solid State Chemistry 329 (2024): 124360.

7. Wejinya, G., & Bhatia, S. (2021). Machine learning for malicious URL detection. In ICT Systems and Sustainability: Proceedings of ICT4SD 2020, Volume 1 (pp. 463-472). Springer Singapore.

8. Do Xuan, C., Nguyen, H. D., & Tisenko, V. N. (2020). Malicious URL detection based on machine learning. International Journal of Advanced Computer Science and Applications, 11(1).

9. Janet, B., & Kumar, R. J. A. (2021, March). Malicious URL detection: a comparative study. In 2021 International Conference on Artificial Intelligence and Smart Systems (ICAIS) (pp. 1147-1151). IEEE.

10. Aljabri, M., Altamimi, H. S., Albelali, S. A., Al-Harbi, M., Alhuraib, H. T., Alotaibi, N. K., ... & Salah, K. (2022). Detecting malicious URLs using machine learning techniques: review and research directions. IEEE Access, 120, 121395-121417.

11. Yuan, J., Liu, Y., & Yu, L. (2021). A novel approach for malicious URL detection based on the joint model. Security and Communication Networks, 2021, 1-12.

12. Yuan, J., Liu, Y., & Yu, L. (2021). A novel approach for malicious URL detection based on the joint model. Security and Communication Networks, 2021, 1-12.

13. Li, T., Kou, G., & Peng, Y. (2020). Improving malicious URLs detection via feature engineering: Linear and nonlinear space transformation methods. Information Systems, 91, 101494.

14. Chen, Y. C., Ma, Y. W., & Chen, J. L. (2020, July). Intelligent malicious URL detection with feature analysis. In 2020 IEEE Symposium on Computers and Communications (ISCC) (pp. 1-5). IEEE.

15. Alsaedi, M., Ghaleb, F. A., Saeed, F., Ahmad, J., & Alasli, M. (2022). Cyber threat intelligence-based malicious URL detection model using ensemble learning. Sensors, 22(9), 3373.

16. Wang, Z., Ren, X., Li, S., Wang, B., Zhang, J., & Yang, T. (2021). A malicious URL detection model based on convolutional neural network. Security and Communication Networks, 2021, 1-12.

17. Alsaedi, M., Ghaleb, F. A., Saeed, F., Ahmad, J., & Alasli, M. (2022). Cyber threat intelligence-based malicious URL detection model using ensemble learning. Sensors, 22(9), 3373.

18. Karve, S. M., Kakad, S., Amol, S., Gavali, A. B., Gavali, S. B., & Shirkande, S. T. (2024). An Identification and Analysis of Harmful URLs through the Application of Machine Learning Techniques. International Journal of Intelligent Systems and Applications in Engineering, 12(17s), 456-468.

19. Aljabri, M., Alhaidari, F., Mohammad, R. M. A., Mirza, S., Alhamed, D. H., Altamimi, H. S., & Chrouf, S. M. B. (2022). An assessment of lexical, network, and content-based features for detecting malicious urls using machine learning and deep learning models. Computational Intelligence and Neuroscience, 2022.

20. Johnson, C., Khadka, B., Basnet, R. B., & Doleck, T. (2020). Towards Detecting and Classifying Malicious URLs Using Deep Learning. J. Wirel. Mob. Networks Ubiquitous Comput. Dependable Appl., 11(4), 31-48.

21. Aung, E. S., & Yamana, H. (2020). Malicious URL detection: a survey. In DEIM Forum F6-3 (Vol. 290).