International Journal of
Human Computations & Intelligence
Vol. 04, Issue. 04, 2025

RESEARCH ARTICLE OPEN ACCESS

Deep Defender: Smart detection of phishing websites

Lavanya N L . Mani Prasad K R . Manjunatha Prasad G R . Monish Gowda V . Sachin Krishna K U

Department of Computer Science and Engineering,
East West College of Engineering,
Visveswaraya Technological University, Yelahanka new Town,
Bengaluru, Karnataka-560064

DOI: **10.5281/zenodo.15429934**
Received: 28 April 2025 / Revised: 09 May 2025 / Accepted: 16 May 2025

Abstract -- Phishing is a consistent threat causing internet users to provide sensitive details in fictitious network environments. Current detection tools tend to sacrifice accuracy and timeliness of response, in doing which the threat exposure level is increased for the users. Here is presented a system based on machine learning intended to detect phishing URLs in the moment, with the aim of enhancing general online footprint safety. Based on the RNN-GRU algorithm, the system tries to maximize the effectiveness and promptness of phishing URL detection. The introduction of this approach brings an effective shield against phishing, a considerable increase in users protection in the digital era.

Index Terms – phishing attacks, cybersecurity, malicious nodes, cyber laws.

I. INTRODUCTION

Phishing is a cyberattack technique where malicious actors deceive users into revealing confidential information by mimicking legitimate websites. These attacks often target sensitive data such as login credentials, banking information, and personal identity details. The rise of internet services has made phishing one of the most prevalent and dangerous forms of online threats. Attackers continuously evolve their tactics, making it difficult for static, rule-based security systems to detect and prevent these attacks effectively. As illustrated in Figure 1, the number of unique phishing sites has seen a dramatic increase over the past decade. From just under 150,000 in 2013, the count escalated to over 1.6 million in Q1 2023, before slightly declining but remaining alarmingly high. This sharp growth highlights the pressing need for intelligent and adaptive phishing detection systems.

MILESTONE RESEARCH.IN OPEN ACCESS

ISSN (Online): 2583-5696
Int. Jr. of Hum Comp. & Int.

Phishing is a method of attack that involves tricking people in order to steal personal or monetary data by impersonating authentic sites. In most cases, attackers aim at exploiting sensitive data such as user passwords and banking records (as well as personally identifiable information). Phishing has become one of the most common and risky threats on the internet because of the proliferation of Internet services. As attacks get personalities and evolve with craftiness, static, rule based security systems find it difficult in the identification and blocking of attacks. Figure 1 indicates that there has been an astonishing increase in the number of unique phishing sites since 2013. The total grew from about 150,000 in 2013. The dramatic increase in phishing efforts necessitates the design of complex and adaptive phishing detection systems. Recent developments in machine learning and deep learning have greatly boosted attempts to detect phishing sites in the cybersecurity field. In contrast to conventional security systems, deep learning approaches help to remove the need for extensive manual feature selection and provide the opportunity to discover complex structures hidden within raw data of the information source. Among different methods of deep learning, the Recurrent Neural Networking (RNNs), and in particular, the Gated Recurrent Unit (GRU) models have been proven superior in their capabilities of detecting phishing attempts.

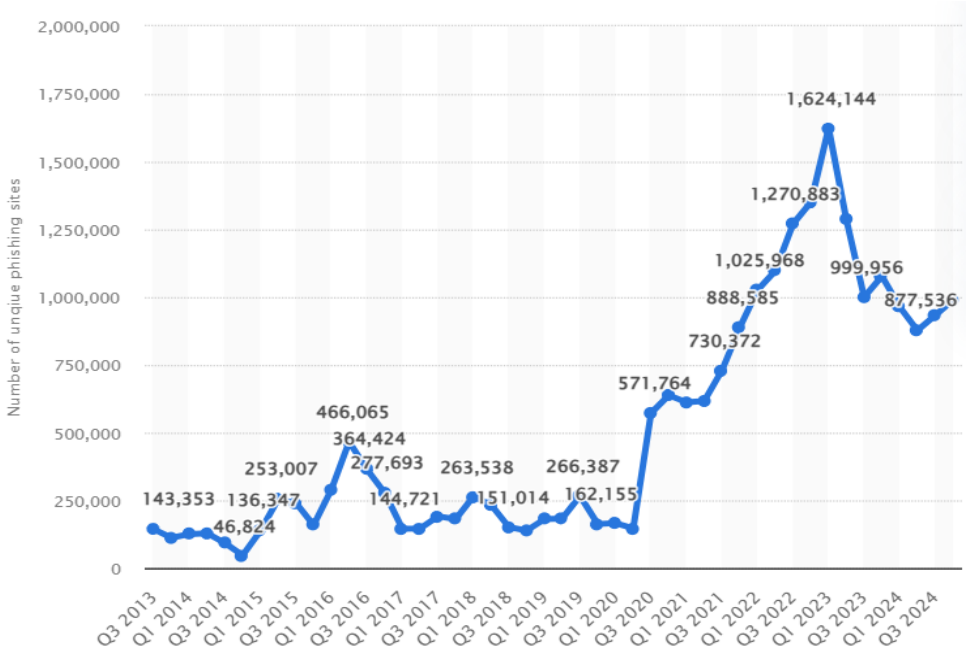

Fig. 1 : Number of global phishing attacks Q3 2013- Q4 2024*

*(*source :https://www.statista.com/graphic/1/266155/number-of-phishing-attacks-worldwide.jpg)*

The process starts with data preprocessing, followed by model training and evaluation. The trained model is then deployed to classify new URLs as either phished or legitimate. A sample GRU-based model architecture is provided to demonstrate the core implementation used in training. This methodology enables real-time classification and supports continuous learning by updating the model with new data. Such a system is capable of adapting to emerging phishing strategies and provides a scalable solution for securing web users. In summary, integrating deep learning techniques like GRU networks with a systematic detection pipeline provides an effective strategy for combating phishing attacks. The ability to learn and adapt in real-time makes this approach a robust defense in the evolving landscape of cybersecurity.

II. LITERATURE REVIEW

Phishing attacks remain a significant threat in cybersecurity, prompting extensive research into machine learning and deep learning-based detection methods. Various studies have explored different approaches to improving the accuracy, efficiency, and adaptability of phishing detection models as below;

- Abdul Basit et al. [1] in 2020 introduced ensemble approach using KNN, Decision Tree and Random Forest classifiers. This specific strategy improves detection results, in particular, when used in combination with KNN and RFC. However, the method is computationally demanding and difficult since it employs various classifiers.
- In 2021, a detailed study on phishing detection methods was carried out by the Mohammed Hazim Alkawaz and colleagues[2]. The researchers focused on integrated systems that combine Random Forest with machine learning. These techniques are good at reducing false positives but suffer in performance with a large dataset.
- Ngueut Quang Do et al. in 2022 carried out a systematic literature review of 2022 that aimed at deep learning innovations, such as CNNs and hybrid models. The methods allow hierarchical learning and simplify feature extraction. However, deep learning models are computationally expensive and have limited capacity for transparency-based interpretability.
- ✦Lizhen Tang&Qusay H. Mahmoud (2022)[4] developed a phishing detection model based on GRUs. GRUs are sequential data processing recurrent neural networks, called Gated Recurrent Units. The model demonstrated a greater level of precision than typical methods, although current concerns are the risk of overfitting and DEXA reliance. ✦
- Umezara et al. (2024)[5] combined RNN and LSTM models to improve their framework for deep learning. The research mainly involved analysis of URL properties and network dynamics to sort good from phishing domains. As expected, the high levels of detection accuracy attained by the approach were attributed to its high demands on processing capabilities and high demands in terms of volume of data.

Across the reviewed literature, a common trend is the growing reliance on deep learning for its ability to automatically extract and learn from complex data patterns. While machine learning models are still used, especially in hybrid formats, deep learning offers superior adaptability and accuracy in identifying zero-hour phishing attacks. The limitations of deep learning approaches include long training times, increased resource demands, and difficulty in explaining model decisions. Despite these challenges, deep learning remains a preferred solution due to its scalability and potential for real-time threat identification. The evolution from ensemble and hybrid models to advanced deep learning architectures reflects the increasing need for models that not only perform well but also adapt quickly to new phishing techniques.

This review highlights that future work should aim to balance detection accuracy with computational efficiency, while also improving model interpretability and reducing dependency on third-party services.

Problem Statement

Phishing attacks impersonate legitimate websites to steal user data, challenging traditional detection methods. Deep learning models like RNN and GRU can analyze URLs, HTML, and network behavior to improve detection. However, their computational cost hinders real-time application. An efficient and adaptive deep learning-based system is needed for accurate, real-time phishing detection.

III. RESEARCH METHODOLOGY

Data collection is the initial step in phishing detection, figure 2 illustrates, meaning the process involves the creation of a dataset containing both real and fake web addresses. After data collection, pre-processing is carried out on data using activities like cleaning to remove errors and anomalies, and ensuing meaningful features such as the URL length, the use of special characters and keyword identifiers are extracted. Text selection allows for stray bits to be retrieved from the URL, and tokenization converts selected sequences to model inputs. MinMax scaling is used in the transformation process to normalize the data. After relationship analysis to determine correlated features, part of the process includes retaining only the most relevant ones at feature extraction step. Once data preprocessing is completed, a deep learning model is then applied, and RNN or GRU layers are used since they can operate on sequential URLs. The model is taught by using the labeled data to identify indicative patterns in URLs that are related with the phishing attacks. When the training is over, accuracy metric is included for evaluation and verification of the model's reliability. After the training, the model is deployed into a web application environment such as Django, such that URL phishing can be assessed instantly. To maintain effectiveness, the system is designed so that it can ingest new URLs is regularly retrained in the face of changing fellitation methods.

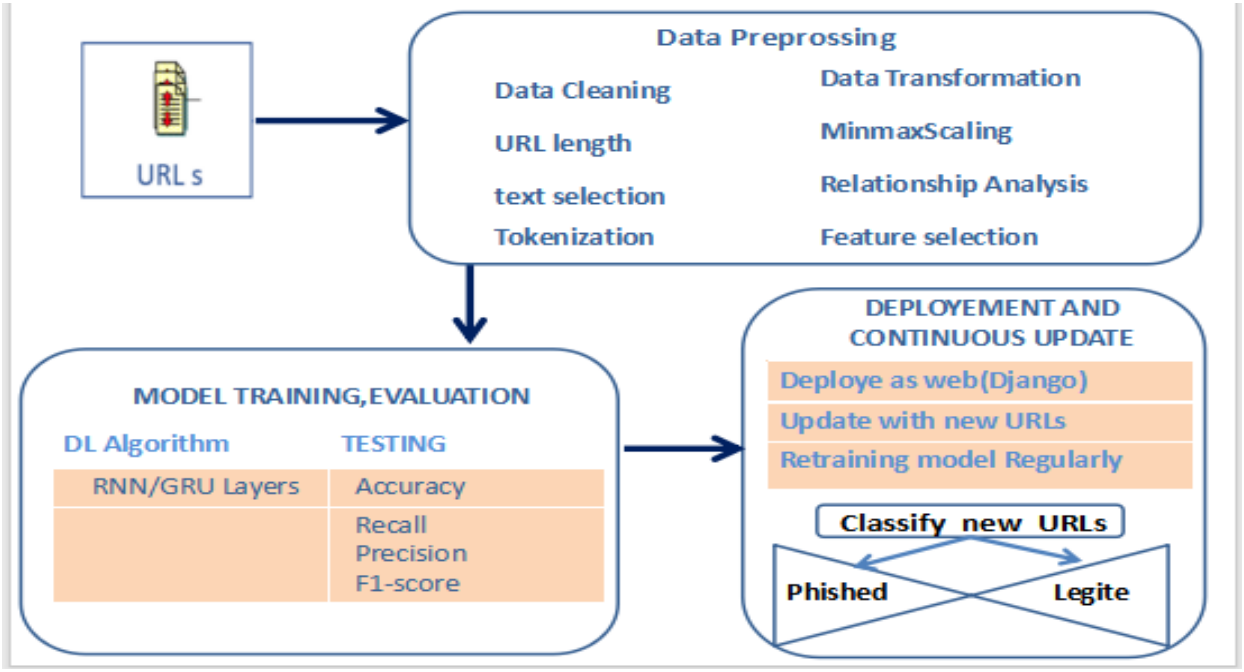

Fig. 2 : Phishing detection methodology

A.APPLIED ALGORITHM

A Gated Recurrent Unit (GRU) is a type of recurrent neural network that is well-suited for processing sequential data such as URLs. In the context of phishing website detection, the GRU can be used to analyze the sequence of characters or tokens in a URL to determine if it is malicious or legitimate. First, the input URL is preprocessed ref. figure[3] and tokenized—either character by character or using n-gram tokens—and then transformed into vector representations through an embedding layer. These embedded vectors are fed sequentially into the GRU, which processes one token at a time while maintaining a hidden state that captures contextual information from the sequence.

The GRU architecture as in figure [3] consists of two main gates: the update gate and the reset gate. The update gate determines how much of the previous information should be retained, helping the model remember important parts of the URL, such as suspicious domain patterns. The reset gate decides how much of the previous state to forget, allowing the model to ignore irrelevant past information, such as common or benign URL structures. A candidate hidden state is calculated using the current input and the reset-modified hidden state, which is then blended with the previous hidden state based on the update gate's output. This mechanism enables the GRU to dynamically control the flow of information through the sequence.

As the model processes the URL, the hidden state evolves to capture meaningful patterns. At the end of the sequence, the final hidden state represents a condensed summary of the entire URL's characteristics. This state is passed through a dense layer and an activation function (typically sigmoid or softmax) to produce a binary classification: phishing or legitimate. GRUs are particularly effective for this task because they handle sequential dependencies well and are computationally more efficient than LSTMs. This makes GRUs a powerful tool in automated phishing detection systems based on URL analysis.

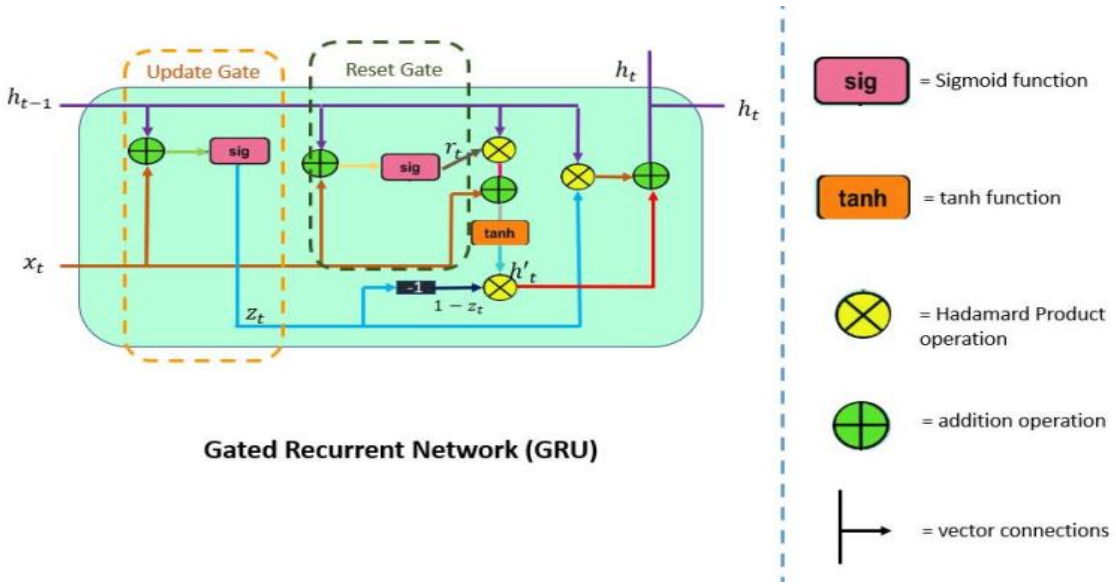

Fig. 3 : Gated recurrent Unit.

Gated Recurrent Units (GRUs) are particularly effective for phishing website detection using URLs due to their ability to model sequential data and learn complex patterns. Unlike traditional approaches that rely on handcrafted features, GRUs can learn directly from raw URL sequences, automatically identifying suspicious patterns such as obfuscated domain names, use of misleading keywords, or the presence of IP addresses instead of domain names. One of the key strengths of GRUs is their ability to handle URLs of varying lengths, which is important since URLs differ significantly in structure. Their gating mechanisms—specifically the update and reset gates—help the model focus on relevant information while filtering out noise, such as benign query parameters or common URL prefixes.

Table 1 : Variuos features of url considered

Feature Name	Feature Description
SFH (Server Form Handler)	Indicates how form data is handled. If “about:blank” is used or if there is no handler, it is suspicious. Value: -1, 0, 1.
URL_of_Anchor	Examines the anchor tags (<a href>) in the webpage. If many refer to different domains or are empty (#), it is suspicious. Value: -1, 0, 1.
Web_traffic	Based on the volume of traffic a website receives. Higher traffic typically implies legitimacy. Value: -1, 0, 1.
URL_Length	Long URLs (over 54 characters) are more likely to be phishing. Short URLs are typically legitimate. Value: -1, 0, 1.
age_of_domain	Older domains are more trustworthy. Newly registered domains are considered suspicious. Value: -1, 1.
having_IP	Checks whether the URL uses an IP address instead of a domain name, which is a phishing indicator. Value: -1, 1.
Favicon	If the favicon is missing or not displayed in the browser tab, it may indicate phishing. Value: -1, 1.
IFrame	Presence of invisible iframes can hide malicious content. If found, it’s suspicious. Value: -1, 1.
Google_Index	Checks if the page is indexed by Google. Non-indexed pages may be suspicious. Value: -1, 1.
DNSRecord	Valid DNS records indicate legitimacy. Anomalies or absence may suggest phishing. Value: -1, 1.
Prefix_Suffix	Use of “-” in domain names is often a phishing trick to mimic legitimate domains. Value: -1, 1.
having_At_Symbol	Use of “@” in URL is suspicious, as it can redirect to different sites. Value: -1, 0, 1.

This table lists various URL-based features used to detect phishing websites. Each feature represents a specific characteristic, such as use of IP address, presence of special symbols (like "@"), URL length, domain age, and web traffic. These features are numerically encoded (e.g., -1, 0, 1) to indicate suspicious, neutral, or safe behavior. For example, “having_IP” checks for IP addresses in

the URL (suspicious), while "Google_Index" and "DNSRecord" confirm if a site is indexed or has valid DNS data (legitimate indicators). These features help the model identify phishing attempts based on abnormal URL structure and metadata.

IV. RESULT AND ANALYSIS

The results demonstrate that the system has adequately classified a potential phishing URL as hazardous. This result corresponds with the procedure shown in figure [4]. To begin with, raw URLs are transformed through necessary steps such as cleansing, tokenization, length analysis and feature extraction. This process allows the extraction that provides relevant signals of phishing attempts, for instance, the use of symbols, structure of the domain name, or a pattern of specific keywords. After preprocessing is done, the preprocessed features are the input of training a neural network that is based on GRU layers. After being trained on an extensive set of labeled URLs, this model develops an ability to distinguish between authentic and phishing URLs from the understanding of time-based dependencies and feature interplay.

After performance evaluation using metrics like accuracy, precision, recall, and F1-score, the model is put on a web framework (e.g., Django) so as to offer a real-time phishing detections. After examining the input URL, the model clearly identifies it as a phishing site, and the system warns the user right away by showing a warning message. With the help of data preprocessing, feature extraction, deep learning algorithms, and user interaction, this task demonstrates the high accuracy and merciless response time of the system for detecting phishing attempts.

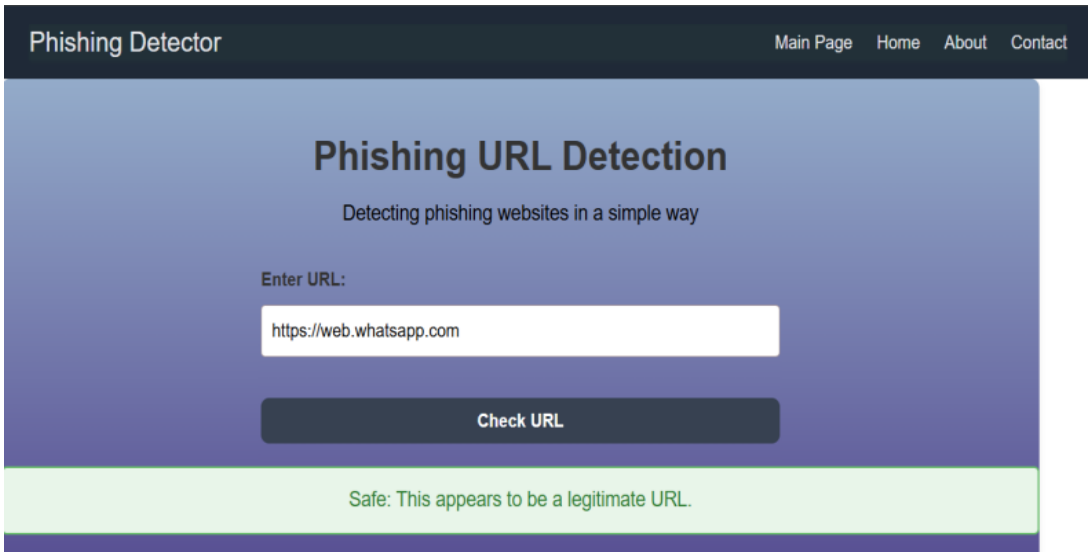

Fig. 4 : Phishing detection web page

Sample code :

```
model = Sequential()
model.add(GRU(64, input_shape=(1, X_train.shape[2])))
model.add(Dropout(0.2)) # Updated input shape
model.add(Dense(1, activation='sigmoid'))
```

```
model.compile(loss='binary_crossentropy', optimizer='adam', metrics=['accuracy'])
```

This code snippet defines a **GRU-based deep learning model** using Keras for binary classification — in this case, for **phishing URL detection**. Here's a brief explanation of each part:

- Initializes a **sequential model**, meaning layers are added one after another in a linear stack.
- Adds a **GRU (Gated Recurrent Unit)** layer with 64 units.
  - input_shape=(1, X_train.shape[2]): The model expects sequences of length 1 (since each URL is treated as a single timestep) with multiple features (e.g., extracted URL features).
- GRUs are excellent at learning **temporal dependencies** and **feature interactions** with fewer parameters than LSTMs.
- Adds a **Dropout layer** to prevent overfitting by randomly setting 20% of input units to 0 during training.
- Final **Dense (fully connected)** layer with 1 neuron and a **sigmoid activation function**.Suitable for **binary classification** (phishing or legitimate).
- Compiles the model with:
  - **Binary cross-entropy loss**: Appropriate for binary classification.
  - **Adam optimizer**: Efficient and adaptive gradient-based optimizer.
  - **Accuracy**: Used as a metric to evaluate model performance during training.

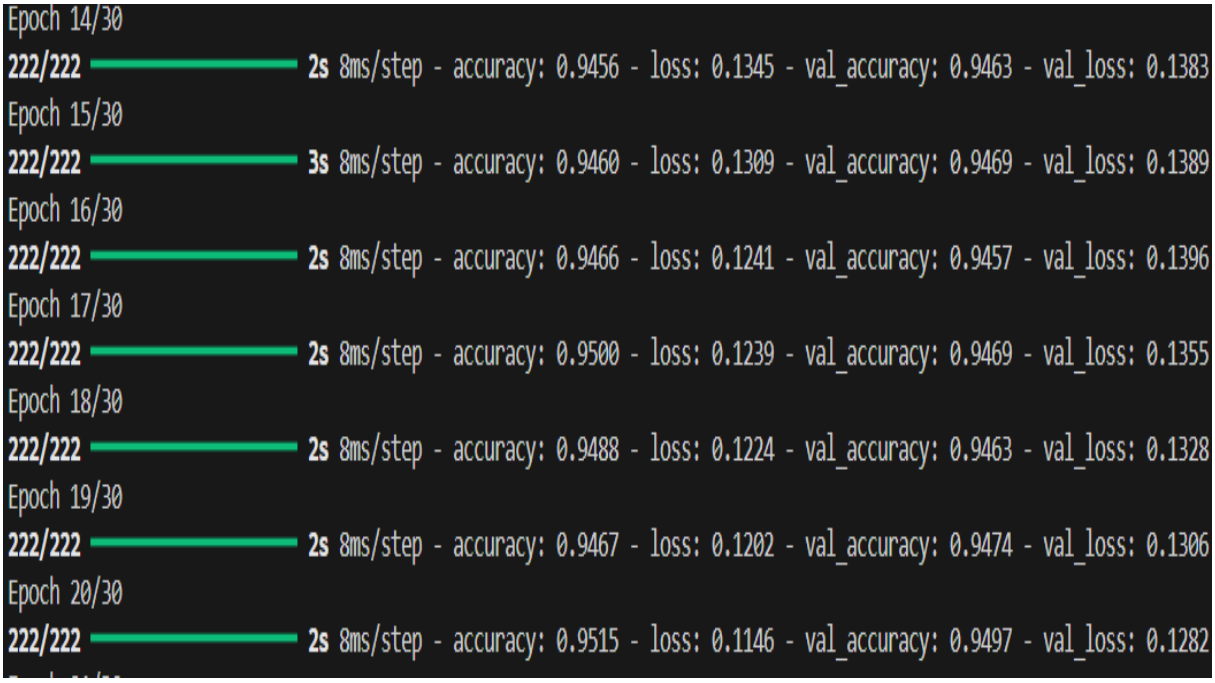

Fig. 5 : model view at various epochs

Figure [5] shows how a phishing URL detection model based on a GRU-based deep learning strategy was developed during training. With training, it is observed that, as epochs increase (14-20), the model keeps improving to read a training accuracy of 95.15% and validation accuracy that ranges from 84.6% to 94.9%. The fact that the number of loss for training and validation data decreased indicates the efficiency of learning by the model, with the model staying reasonably calibrated. respond Eventually, the observed results indicate that the model works at a high degree of consistency and is appropriate for phishing URLs identification. The combined results reveal that the model shows a high performance and is capable of reliably discriminating between phishing URLs and non-phishing URLs.

V. CONCLUSION AND FUTURE ENHANCEMENT

Employing best of the breed machine learning approaches like Recurrent Neural Networks (RNN) and Gated Recurrent Units (GRU), the implemented phishing URL detection system proves to be quite effective in suppressing growing threat of phishing attacks. Powered by a robust deep learning model, the system allows for reliable differentiation of legitimate URLs from potentially deceptive ones – thus, providing the users with prompt immunity to harmful web content. The simple web-based system simplifies URL checking so that non-technical individuals can evaluate sites for validity online. The high accuracy performance and low error rates confirmed by the performance test also reinforces the suitability of the model for real-time, high-traffic environments which prepares the way for scalability. Paying homage to the faceted nature of the problem, the solution becomes a significant addition to the cybersecurity sphere, with improving user safety, diminishing the risks of phishing, and digital security promotion among its main functions.

It is expected that the system will go through further development so as to be fitted with improvements that will respond to the continued change in online threat pattern. Priority is given to enhancing detection abilities using the inclusion of different training samples and the use of robust computational resources. Moreover, the number of language supported will strengthen the system's worldwide usability and attractiveness. Connecting the system to social media sites will add to the capability of the system and thus immediate detection/elimination of any phishing attempt on user feeds is possible. Enacting automated retraining and update processes will advance the observed online phishing threats to the detection algorithms in the system. Through the use of user feedback and community-contributed flag systems, decision-making processes in the model can be improved consistently. The improvements aim to drive the system towards ultimate detection and highlight its part of the cybersecurity program that identifies and develops with phishing attacks with a finalized broader protection of the internet for everyone.

REFERENCES

1. Basit, A., Zafar, M., Javed, A. R., & Jalil, Z. (2020, November). A novel ensemble machine learning method to detect phishing attack. In *2020 IEEE 23rd International Multitopic Conference (INMIC)* (pp. 1–5). IEEE. https://doi.org/10.1109/INMIC50486.2020.9318105
2. Alkawaz, M. H., Steven, S. J., Mohammad, O. F., & Johar, M. G. M. (2022, May). Identification and analysis of phishing website based on machine learning methods. In *2022 IEEE 12th Symposium on Computer Applications & Industrial Electronics (ISCAIE)* (pp. 246–251). IEEE. https://doi.org/10.1109/ISCAIE53661.2022.9803903

3. Do, N. Q., Selamat, A., Krejcar, O., Herrera-Viedma, E., & Fujita, H. (2022). Deep learning for phishing detection: Taxonomy, current challenges and future directions. *IEEE Access, 10*, 36429–36463. https://doi.org/10.1109/ACCESS.2022.3162951
4. Tang, L., & Mahmoud, Q. H. (2021). A deep learning-based framework for phishing website detection. *IEEE Access, 10*, 1509–1521. https://doi.org/10.1109/ACCESS.2021.3138631
5. Zara, U., Ayub, K., Khan, H. U., Daud, A., Alsahfi, T., & Gulzar, S. (2024). Phishing website detection using deep learning models. *IEEE Access*. https://doi.org/10.1109/ACCESS.2024.3392073
6. Jain, A. K., & Gupta, B. B. (2018). PHISH-SAFE: URL features-based phishing detection system using machine learning. In *Cyber Security: Proceedings of CSI 2015* (pp. 467–474). Springer Singapore. https://doi.org/10.1007/978-981-10-8536-9_37
7. Rajasekar, V., Premalatha, J., Sathya, K., Raakul, S. D., & Saracevic, M. (2021, February). An enhanced anti-phishing scheme to detect phishing website. In *IOP Conference Series: Materials Science and Engineering* (Vol. 1055, No. 1, p. 012077). IOP Publishing. https://doi.org/10.1088/1757-899X/1055/1/012077
8. Wang, W., Zhang, F., Luo, X., & Zhang, S. (2019). PDRCNN: Precise phishing detection with recurrent convolutional neural networks. *Security and Communication Networks, 2019*(1), Article 2595794. https://doi.org/10.1155/2019/2595794
9. Lakshminarayana, L. N., & Gangadharaiah, S. R. (2023). Trust based multi objective-pelican optimization algorithm for mobile ad hoc networks. *International Journal of Intelligent Engineering and Systems, 16*(6). https://doi.org/10.22266/ijies2023.1231.17
10. Lavanya, N. L., & Nagarathna, C. (2024). Transformer model to evaluate subjective script. *International Journal of Human Computations & Intelligence, 3*(4), 350–357.
11. Pasha, A., Ahmed, S. T., Painam, R. K., Mathivanan, S. K., Mallik, S., & Qin, H. (2024). Leveraging ANFIS with Adam and PSO optimizers for Parkinson's disease. *Heliyon, 10*(9), e23976. https://doi.org/10.1016/j.heliyon.2024.e23976
12. Kumar, S. S., Ahmed, S. T., Sandeep, S., Madheswaran, M., & Basha, S. M. (2022). Unstructured oncological image cluster identification using improved unsupervised clustering techniques. *Computers, Materials & Continua, 72*(1), 1061–1077. https://doi.org/10.32604/cmc.2022.020884