PHISHING EMAIL DETECTION USING A CONTEXT-AWARE HYBRID DEEP LEARNING FRAMEWORK

[1]SUJEERA JASMIN M, [2]Ms.SREEJI K B

[1]MCA Scholar, [2]Assistant Professor
[1]Department of Computer Applications,
[1]Nehru College of Engineering and Research Centre
Thrissur, India

Abstract:

Phishing emails are one of the most common cybersecurity threats used by attackers to obtain sensitive information such as passwords, banking details, and personal data. Traditional detection systems rely mainly on rule-based filtering techniques and blacklist approaches, which are often ineffective against modern phishing attacks. Attackers continuously modify email content to bypass these security mechanisms. Therefore, intelligent detection techniques are required to identify phishing attempts accurately. This paper presents a deep learning-based phishing email detection system that uses Natural Language Processing (NLP) techniques to analyze the textual content of emails. The proposed model combines Convolutional Neural Networks (CNN) and Gated Recurrent Units (GRU) to identify hidden patterns within email messages. CNN helps extract important textual features while GRU captures sequential relationships between words in the email body. By learning contextual patterns from large datasets, the deep learning model is capable of distinguishing phishing emails from legitimate ones.
Experimental results demonstrate that the proposed approach provides improved detection accuracy and reduced false positives compared with traditional machine learning techniques. The system can be effectively used in email security platforms to enhance protection against phishing attacks.

Index Terms **- Phishing Detection, Deep Learning, CNN, GRU, Natural Language Processing, Email Security, Cybersecurity.**

I. INTRODUCTION

Email communication plays a critical role in modern digital ecosystems, serving individuals, enterprises, and governmental organizations. Despite improvements in cybersecurity infrastructure, phishing attacks continue to exploit email platforms due to their wide accessibility and user dependency. Phishing emails are carefully crafted messages intended to deceive recipients into revealing confidential information such as login credentials, banking details, or organizational data.
Conventional phishing detection systems rely on predefined rules, blacklist databases, and keyword-based filtering. Although effective against known attack signatures, these approaches struggle to identify sophisticated phishing emails that utilize contextual deception and semantic modification.
Recent advancements in deep learning and natural language processing have introduced transformer-based models capable of capturing contextual relationships within textual data. These architectures provide improved semantic understanding compared to traditional feature engineering techniques. Motivated by these developments, this study proposes a hybrid deep learning framework that integrates contextual embedding, convolutional feature extraction, sequential modeling, and attention mechanisms to strengthen phishing detection capabilities.

II.RELATED WORK

Deep learning methodologies have significantly enhanced text classification tasks in cybersecurity. Transformer-based models generate contextual embeddings that capture bidirectional semantic dependencies within textual sequences. This capability enables better differentiation between legitimate and malicious communication patterns.

Convolutional Neural Networks (CNN) have demonstrated effectiveness in extracting localized textual features, particularly short phrases associated with phishing attempts. Gated Recurrent Units (GRU) contribute by modeling sequential dependencies within structured text while maintaining computational efficiency.
Recent research trends favor hybrid architectures that combine transformer embeddings with convolutional and recurrent layers. Attention mechanisms further improve performance by assigning adaptive importance to linguistically significant segments of email content.

III.PROPOSED METHODOLOGY

The proposed phishing detection system follows several stages including data collection, preprocessing, feature extraction, and classification.
First, a dataset containing phishing emails and legitimate emails is collected from publicly available sources. The dataset is then pre processed to remove noise and irrelevant information. Preprocessing steps include converting text to lowercase, removing punctuation marks, eliminating stop words, and tokenizing sentences into individual words.
After preprocessing, the textual data is converted into numerical representations using word embedding techniques. Word embeddings transform textual information into vectors that can be processed by neural networks.
The deep learning model used in this research combines Convolutional Neural Networks (CNN) and Gated Recurrent Units (GRU). The CNN layer extracts important textual features from the email content. These features represent patterns that are useful for identifying phishing attempts.
The GRU layer processes the extracted features and captures sequential relationships between words in the email text. This helps the model understand contextual patterns within the message.
Finally, a dense output layer performs the classification task and predicts whether the email is phishing or legitimate. The model is trained using labeled email datasets to improve its detection accuracy.

Figure 1: Architecture of the Proposed Phishing Email Detection System Using Deep Learning

RESULTS AND DISCUSSION

The performance of the proposed deep learning model was evaluated using standard classification metrics such as accuracy, precision, recall, and F1-score.
Experimental results show that the hybrid CNN-GRU model provides high detection accuracy when compared with traditional machine learning approaches. The model successfully identifies phishing patterns within email text and reduces false positive rates.
Deep learning models are particularly effective because they automatically learn features from the dataset. This eliminates the need for manual feature engineering and improves the system's ability to detect new phishing strategies.
The results demonstrate that intelligent deep learning models can significantly enhance email security by detecting phishing emails with high reliability

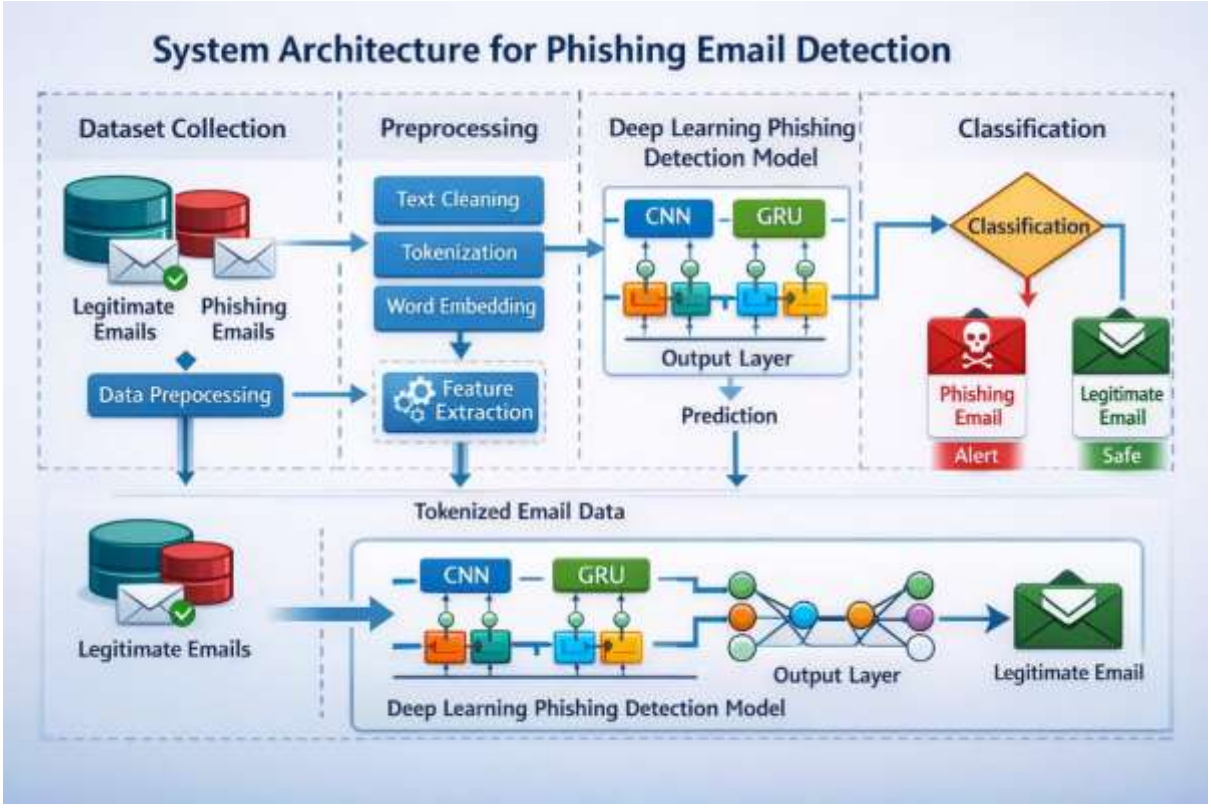

Figure 2: System Architecture of the Proposed Deep Learning Based Phishing Email Detection Model

The above figure illustrates the overall working process of the proposed phishing email detection system. It shows the stages involved in detecting phishing emails, including dataset collection, preprocessing, feature extraction, deep learning model processing using CNN and GRU, and final classification into phishing or legitimate emails.

IV.ANALYTICAL DISCUSSION

The integration of contextual semantic modeling with hierarchical neural processing offers a theoretically robust framework for phishing detection. Transformer embeddings enhance semantic depth, while convolutional layers identify localized malicious cues. Sequential modeling strengthens contextual continuity analysis, and attention mechanisms dynamically highlight high-risk linguistic patterns.

Although empirical implementation is beyond the scope of this conceptual study, prior research indicates that hybrid deep learning architecture consistently outperforms traditional statistical classifiers in phishing detection tasks. Future work will involve experimental validation using standard evaluation metrics such as accuracy, precision, recall, and F1-score.

V. ADVANTAGES OF THE PROPOSED SYSTEM

The proposed phishing detection system provides several advantages:

- Automated feature extraction from email content
- Improved phishing detection accuracy
- Reduced false positive rates
- Ability to detect new phishing patterns
- Scalable for large email datasets
- Reduced dependency on manual feature engineering

VI. FUTURE DIRECTIONS

Future research can focus on integrating advanced transformer-based models such as BERT and GPT for improved contextual understanding of email content. These models can further enhance phishing detection accuracy.

Another possible improvement is the development of real-time phishing detection systems that can be integrated into enterprise email servers and cloud-based email platforms. Additionally, explainable artificial intelligence techniques can be used to improve transparency in model predictions.

Mobile-friendly phishing detection systems can also be developed to protect users on smartphones and portable devices.

VII. CONCLUSION

Phishing attacks continue to be a major cybersecurity challenge in modern digital communication. Traditional detection methods are often insufficient to detect sophisticated phishing attempts. This research proposed a deep learning-based phishing email detection system that uses Natural Language Processing techniques and neural network architectures to identify malicious emails. The hybrid CNN-GRU model effectively analyses textual patterns and contextual relationships within email messages. Experimental results demonstrate that the proposed approach provides improved detection accuracy compared with traditional machine learning techniques.
The study highlights the importance of intelligent automated systems in strengthening email security and protecting users from phishing attacks.

VIII. ACKNOWLEDGEMENT

The author would like to express sincere gratitude to Ms. Sreeji K.B for her valuable guidance, encouragement, and continuous support throughout the completion of this research work. The author also thanks the Department of MCA, Nehru College of Engineering and Research Centre for providing the necessary academic support and resources to carry out this study successfully.

IX. REFERENCES

[1]. Egozi, G. & Verma, R. Phishing email detection using robust nlp techniques. In 2018 IEEE International Conference on Data Mining Workshops (ICDMW). IEEE. (2018).
[2]. Fang, Y. et al. Phishing email detection using improved RCNN model with multilevel vectors and attention mechanism. Ieee Access.7, 56329–56340 (2019).
[3]. Valecha, R., Mandaokar, P. & Rao, H. R. Phishing email detection using persuasion cues. IEEE Trans. Dependable Secur. Comput.19 (2), 747–756 (2021).
[4]. Bountakas, P. & Xenakis, C. Helphed: hybrid ensemble learning phishing email detection. J. Netw. Comput. Appl. 210, 103545 (2023).
[5]. Che, H. et al. A content-based phishing email detection method. In 2017 IEEE International Conference on Software Quality, Reliability and Security Companion (QRS-C). IEEE. (2017).
[6]. Devlin, J. Bert: Pre-training of deep bidirectional transformers for language understanding. arXiv preprint arXiv:1810.04805 (2018).
[7]. Yamashita, R. et al. Convolutional neural networks: an overview and application in radiology. Insights into Imaging. 9, 611–629 (2018).
[8]. Chung, J. et al. Empirical evaluation of gated recurrent neural networks on sequence modeling. arXiv preprint arXiv:1412.3555(2014).
[9]. Li, J. et al. On the diversity of multi-head attention. Neurocomputing 454, 14–24 (2021).
[10]. Mirjalili, S., Mirjalili, S. M. & Lewis, A. Grey Wolf optimizer. Adv. Eng. Softw. 69, 46–61 (2014).
[11] Mirjalili, S. & Lewis, A. The Whale optimization algorithm. Adv. Eng. Softw. 95, 51–67 (2016).
[12]. Mirjalili, S. et al. Salp swarm algorithm: A bio-inspired optimizer for engineering design problems. Adv. Eng. Softw. 114, 163–191 (2017).
[13]. Salloum, S. et al. A systematic literature review on phishing email detection using natural Language processing techniques. IEEE Access. 10, 65703–65727 (2022).
[14]. Wang, J., Li, Y. & Rao, H. R. Overconfidence in phishing email detection. J. Association Inform. Syst. 17 (11), 1 (2016).
[15]. Al-Subaiey, A. et al. Novel interpretable and robust web-based AI platform for phishing email detection. Comput. Electr. Eng. 120, 109625 (2024).
[16]. Qi, Q. et al. Enhancing phishing email detection through ensemble learning and undersampling. Appl. Sci. 13 (15), 8756 (2023).
[17]. Champa, A. I., Rabbi, M. F. & Zibran, M. F. Curated datasets and feature analysis for phishing email detection with machine learning. In 2024IEEE 3rd International Conference on Computing and Machine Intelligence (ICMI). IEEE. (2024).
[18]. Opara, C., Modesti, P. & Golightly, L. Evaluating spam filters and stylometric detection of AIgenerated phishing emails. Expert Syst. Appl. 276, 127044 (2025). NCERC 14
[19]. Yang, Y. et al. Servenet: A deep neural network for web services classification. in 2020 IEEE international conference on web services (ICWS). IEEE. (2020).
[20]. Tang, B. et al. Co-attentive representation learning for web services classification. Expert Syst. Appl. 180, 115070 (2021).