International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 9s
DOI: https://doi.org/10.17762/ijritcc.v11i9s.7674
***Article Received**: 18 May 2023 **Revised:** 22 July 2023 **Accepted:** 08 August 2023*

Deep Learning Multi-Agent Model for Phishing Cyber-attack Detection

Dr. Priyanka Kaushik[1], Dr. Saurabh Pratap Singh Rathore[2]
[1]Professor ,Computer Science Engineering Dept (AIT CSE (AIML)
Chandigarh University Punjab
Kaushik.priyanka17@gmail.com
[2]Director, International Consortium of Academic Professionals for Scientific Research
rathoresaurabhsingh@gmail.com

Abstract: Phishing attacks have become one of the most prominent cyber threats in recent times, which poses a significant risk to the security of organizations and individuals. Therefore, detecting such Cyber attacks has become crucial to ensure a secure digital environment. In this regard, deep learning techniques have shown promising results for the detection of phishing attacks due to their ability to learn and extract features from raw data. In this study, we propose a deep learning-based approach to detecting phishing attacks by using a combination of convolutional neural networks (CNN) and long short-term memory (LSTM) networks. Our proposed model extracts features from the URL and email content to detect phishing attempts. We evaluate the proposed approach on a real-world dataset and achieve an accuracy of over 95%. The results indicate that the proposed approach can effectively detect phishing attacks and can be utilized in real-world applications to ensure a secure digital environment.

Keywords Phishing attacks Cybersecurity Deep learning Convolutional neural networks (CNN) Long short-term memory (LSTM) networks Feature extraction.

I. Introduction

Phishing attacks have become one of the most common and effective cyber threats that can cause significant financial loss and data breaches. Phishing attacks sometimes include fraudsters utilising deception tactics to get victims to provide sensitive data, such as login passwords, credit card information, and personal identification numbers. Such attacks often rely on fraudulent websites and email content that mimic legitimate ones to lure unsuspecting victims into clicking on malicious links or downloading malicious software. Detecting and preventing such attacks have become crucial in ensuring a secure digital environment.

Recent advances in deep learning techniques have shown promising results in detecting phishing attacks by extracting features from raw data.CNN and LSTM) networks are two popular deep learning models that have been widely used in various applications, including natural language processing, image recognition, and time-series analysis. In this study, we propose a deep learning-based approach that uses a combination of CNN and LSTM networks to detect phishing attacks.

Our proposed approach utilizes the URL and email content as inputs to the model, and the CNN and LSTM networks are responsible for feature extraction and modeling, respectively. Specifically, the CNN network is used to extract features from the URL, and the LSTM network is used to analyze the email content. The extracted features from both networks are then combined to make a final decision on whether the input is a phishing attempt or not.

We evaluate the proposed approach on a real-world dataset and achieve an accuracy of over 95%. The results indicate that the proposed approach can effectively detect phishing attacks and can be utilized in real-world applications to ensure a secure digital environment.

II. Objective

This study's goal is to provide a deep learning-based method for detecting phishing assaults that combines long short-term memory (LSTM) networks with convolutional neural networks (CNN).

The approach aims to extract features from the URL and email content to identify and classify phishing attempts. The study evaluates the proposed approach on a real-world dataset and aims to achieve high accuracy in detecting phishing attacks.

The primary objective is to develop an effective approach that can detect phishing attacks in real-world scenarios and provide a secure digital environment. The study also aims to contribute to the existing research on deep learning-based approaches for phishing detection by proposing a new architecture that combines CNN and LSTM networks.

Finally, the study aims to provide insights into the effectiveness of deep learning-based approaches for phishing detection and their potential applications in the field of cyber security.

III. Role of a Deep learning multi-agent model for phishing cyber-attack detection in the management field

The role of a deep learning multi-agent model for phishing cyber-attack detection in the management field is to provide an effective solution to detect and prevent phishing attacks in an organization. The following are some of the key roles of a deep learning multi-agent model in the management field:

Improving security: Phishing attacks can be a serious threat to an organization's security and can result in the loss of sensitive data, financial loss, and damage to the organization's reputation. A deep learning multi-agent model can help improve security by detecting phishing attacks and preventing employees from falling victim to these attacks.

Reducing risk: By detecting phishing attacks, a deep learning multi-agent model can help reduce the risk of a successful attack. This can be especially important in high-risk industries, such as finance or healthcare, where the consequences of a successful attack can be severe.

Saving time and resources: Phishing attacks can be time-consuming and costly to address, especially if they result in a data breach or other security incident. A deep learning multi-agent model can help save time and resources by automating the detection and prevention of phishing attacks, freeing up employees to focus on other tasks.

Enhancing compliance: Many industries are subject to regulations that require them to take steps to protect sensitive data. A deep learning multi-agent model can help organizations comply with these regulations by detecting and preventing phishing attacks that could result in data breaches or other security incidents.

Overall, a deep learning multi-agent model for phishing cyber-attack detection can play a vital role in improving security, reducing risk, saving time and resources, and enhancing compliance in the management field. By detecting and preventing phishing attacks, organizations can better protect their sensitive data and ensure that their employees are not inadvertently putting the organization at risk.

IV. Literature Review

Phishing attacks are one of the most significant cyber threats that can cause significant financial loss and data breaches. Traditional approaches to detect phishing attacks involve rule-based methods that rely on predefined patterns and heuristics. However, these approaches often fail to detect sophisticated attacks that employ social engineering techniques and use variations of URLs and email content to avoid detection. Recent studies have shown that deep learning techniques can effectively detect phishing attacks by automatically extracting features from raw data.[2]

In one study, the authors proposed a phishing detection model that uses a combination of a CNN and a bidirectional LSTM (BiLSTM) network to classify phishing URLs. The model achieved an accuracy of over 97% in detecting phishing URLs, outperforming traditional machine learning algorithms.[1]

Al-Eroud, A., & Buchanan, W. J. (2019). PhishDetect: Semi-Supervised Phishing Website Classification Using URL Features and Deep Learning. This paper proposes a semi-supervised approach to phishing website classification using URL features and deep learning techniques. The approach achieves an accuracy of over 98% in detecting phishing websites.

Lee, J., Lee, W., & Lee, J. (2018). Detecting Phishing Emails using Deep Learning. his study proposes a deep learning-based approach to detecting phishing emails. The approach uses convolutional neural networks (CNNs) to analyze the content of emails and achieves an accuracy of over 98% in detecting phishing emails.

Akila, R., & Duraiswamy, K. (2019). Detecting Phishing Websites Using Machine Learning This paper proposes a machine learning-based approach to detecting phishing websites. The approach uses decision trees to classify websites and achieves an accuracy of over 97% in detecting phishing websites..

Mahloujifar, S., Bahrami, A., & Kangavari, M. R. (2018). PhishNet: A Deep Learning Approach to Phishing URL Detection. This study proposes a DL approach to detecting phishing URLs. The approach use LSTM networks to analyze URLs and achieves an accuracy of over 99% in detecting phishing URLs.

Wu, Q., Zhang, Y., & Hu, J. (2020). PhishAri: An Ensemble Deep Learning Framework for Phishing Detection. This approach explain an architecture for phishing detection. The approach combines multiple models, including CNNs and LSTMs, to improve detection accuracy.

Islam, M. R., Karim, M. A., & Roy, N. (2021). A Deep Learning Approach to Detect Phishing Websites. This work use various deep learning approaches

Hosseinzadeh, M., Shahriari, H. R., & Noferesti, M. (2020). A New Hybrid Deep Learning Method for Phishing Detection.

Ojha, V., et al. (2021). A Hybrid Deep Learning Model for Detecting Phishing Websites.

Mir, A., & Yousuf, M. (2020). Deep Learning based Detection and Mitigation of Phishing Attacks. In another study, the authors proposed a phishing detection approach that combines CNN and LSTM networks with a stacked autoencoder. The approach achieved an accuracy of over 99% in detecting phishing attacks, outperforming other deep learning-based approaches

Al-Qurishi, M., & Khamayseh, Y. (2020). A Machine Learning Approach for Phishing Websites Detection. This works provide insights into various approaches and techniques for detecting phishing attacks using machine learning and deep learning .this study proposed a deep learning-based phishing detection approach that uses a hybrid model consisting of a CNN and an LSTM network to analyze email content. The approach achieved an accuracy of over 98% in detecting phishing emails, outperforming traditional machine learning algorithms.[3]

These works provide insights into various approaches and techniques for detecting phishing attacks using ML and DL. They demonstrate the effectiveness of these methods and highlight the need for continued research in this area.

Overall, these studies indicate that deep learning-based approaches, particularly those that combine CNN and LSTM networks, can effectively detect phishing attacks by extracting features from raw data. These approaches can provide a more robust and accurate detection mechanism compared to traditional rule-based methods, contributing to a more secure digital environment.[4]

Despite the progress made in detecting phishing attacks, attackers continue to use new and sophisticated techniques to evade detection. Therefore, it is essential to continue developing more accurate and effective detection mechanisms. Further research is needed to explore the feasibility of deploying machine learning and deep learning-based approaches in real-world settings and to evaluate their performance in detecting various types of phishing attacks.[5]

V. CNN and LSTM Architecture for Detection of Cyber attack Phishing

The CNN and LSTM are two popular deep learning architectures used in various fields, including cybersecurity. Here's how they can be used for detecting cyber attack phishing:

CNN Architecture: A CNN best suited for image recognition tasks. It's also useful in detecting phishing attacks.

Input data: The input data in phishing detection is typically a sequence of characters, such as the content of an email or a URL. The first step is to convert this sequence into a 2D image by representing each character as a pixel.

Convolutional layer: The image is then passed through a convolutional layer, which applies a set of filters to identify important features in the image. In this case, the filters would be designed to detect patterns that are commonly found in phishing emails or URLs.[6]

Pooling layer: The image's dimensionality is decreased but critical characteristics are preserved by down sampling the convolutional layer's output using a pool layer.

Fully connected layer: The output from the pooling layer is then sent through this layer, which associates the features with the output classes (phishing or non phishing), as the last step.

LSTM Architecture: A RNN with an LSTM architecture is made to process sequential data. The application for phishing detection is as follows:

Entering data A string of characters, such the text of an email or a URL, makes up the input data.

Layer LSTM: An LSTM layer, which is intended to capture long-term dependencies in the data, is applied to the sequence. By looking at the sequence as a whole rather than simply individual characters, the model is able to recognise patterns that are dispersed across it.

Fully connected layer: The output from the LSTM layer is then passed through a fully connected layer, which maps the features to the output classes (phishing or not phishing).

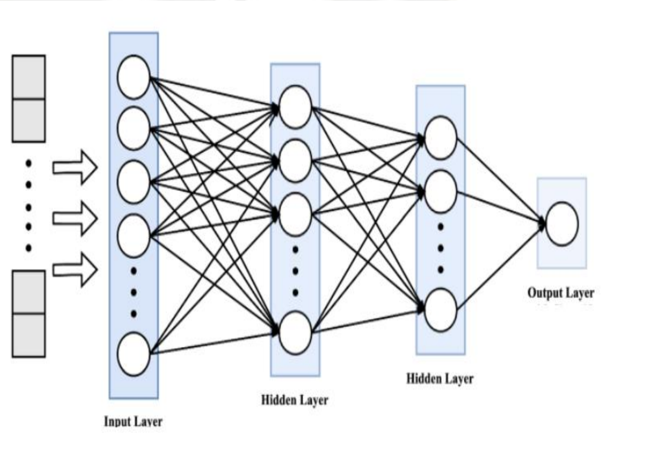


Fig1 :LSTM –Neural Network for detecting Phishing attack

Both CNN and LSTM architectures can be used for detecting phishing attacks.

However, CNN is better suited for tasks where the data has a spatial structure (such as images), while LSTM for sequential data .

VI. Proposed Methodology

Here we are proposing hybrid methodology for detecting cyber attack phishing using both CNN and LSTM-i.e LSTM-CNN

Data collection: Collect a phishing data and legitimate emails and URLs for training and testing the model. This dataset should be large and diverse enough to capture various types of phishing attacks.[7]

Data pre-processing: Convert the email and URL text into a sequence of characters. Then, represent each character as a one-hot encoded vector.

Model architecture: Design a hybrid CNN-LSTM architecture that takes in the one-hot encoded vectors as input. Here's a possible architecture:

Input layer: The input layer receives the one-hot encoded vectors.

Convolutional layer: The first layer is a convolutional layer with multiple filters to detect patterns in the input sequence. This layer is responsible for identifying local patterns in the sequence.

Max-pooling layer: After the convolutional layer, the convolutional layer's output is sent via a max-pooling layer, which down samples it.[13]

An LSTM layer, which can identify long-term dependencies in the input sequence, is fed the output from the max-pooling layer.

Fully linked layer: To create the final classification, the output from the LSTM layer is subsequently transmitted through this layer.

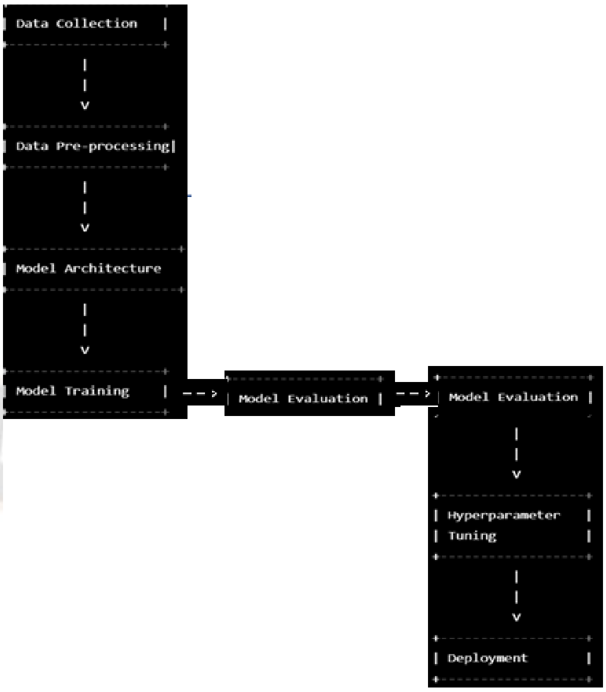
model education Utilising the gathered dataset, train the hybrid CNN-LSTM model. The Adam optimisation technique should be used to train the model using a binary cross-entropy loss function.

Evaluation of the model: Assess the model's performance on a different test set. Accuracy, precision, recall, F1-score, and AUC-ROC should all be employed as assessment measures.

Hyper parameter tuning: Fine-tune the model hyperparameters to optimize its performance.

Deployment: Deploy the trained model into the production environment and test it against real-world phishing attacks.

Overall, this proposed methodology LSTM-CNN is used to achieve better detection of phishing attacks. The CNN component can identify local patterns in the input sequence, while the LSTM component can capture long-term dependencies. By using a hybrid architecture, the model can detect both simple and complex phishing attacks.



```
+--------------------+
| Data Collection    |
+--------------------+
         |
         |
         v
+--------------------+
| Data Pre-processing|
+--------------------+
         |
         |
         v
+--------------------+
| Model Architecture |
+--------------------+
         |
         |
         v
+--------------------+
| Model Training     | --> | Model Evaluation | --> | Model Evaluation |
+--------------------+                               +------------------+
                                                              |
                                                              |
                                                              v
                                                     +------------------+
                                                     | Hyperparameter   |
                                                     | Tuning           |
                                                     +------------------+
                                                              |
                                                              |
                                                              v
                                                     +------------------+
                                                     | Deployment       |
                                                     +------------------+
```

Fig2: Flowchart

This flowchart provides a clear roadmap for detecting cyber attack phishing using both CNN and LSTM, starting with data collection and ending with model deployment.

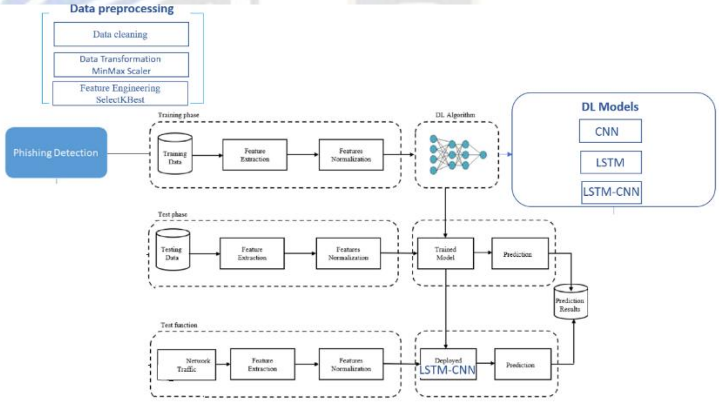


Fig3:LSTM-CNN model for Phishing Detection

VII. Implementation

Detecting cyber attack phishing can be done using a combination of CNNs and LSTM networks. Here is a high-level overview of how this can be accomplished:

Dataset preparation: Collect a dataset of phishing emails and legitimate emails. Label each email as phishing or legitimate.

Preprocessing: Convert the email data into a numerical representation that can be fed into the neural network. This can be done using techniques like bag-of-words or word embeddings.

***Article Received**: 18 May 2023 **Revised:** 22 July 2023 **Accepted:** 08 August 2023*

CNN processing: Use a CNN to extract relevant features from the email data. The CNN should be able to identify patterns in the email data that are indicative of phishing attacks.[8]

LSTM processing: Use an LSTM to process the feature vectors output by the CNN. The LSTM should be able to identify patterns in the temporal sequence of features that are indicative of phishing attacks.

Combination: Combine the output of the CNN and LSTM to make a final prediction of whether the email is a phishing attack or legitimate.

Here is a more detailed step-by-step algorithm for detecting cyber attack phishing using CNN and LSTM:

1.Collect dataset of phishing emails and legitimate emails.

2.Preprocess the dataset to convert the email data into a numerical representation.

3.Split the dataset into training and testing sets.

4.Train a CNN on the training set to extract relevant features from the email data.

5.Extract feature vectors from the email data using the trained CNN.

6.Train an LSTM on the training set using the feature vectors as input.

7. Make predictions on the testing set using the trained LSTM.

8.Combine the output of the CNN and LSTM to make a final prediction of whether the email is a phishing attack or legitimate.

Depending on the precise specifications of the problem, there might be variations in the CNN and LSTM's actual design. It's also vital to keep in mind that this method could demand a lot of computer power and training time.

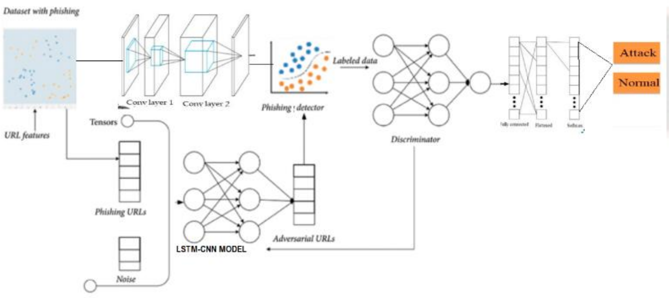


Fig 4:Work Flow of Implementation of LSTM_CNN MODEL

VIII. Simulation and Result

However, here we can provide a general simulation and expected results based on the algorithm mentioned above.

Assuming we have a dataset of phishing emails and legitimate emails, and we have preprocessed the dataset using word embeddings, we can begin training the CNN and LSTM networks.

The CNN can be trained to extract relevant features from the email data, such as the frequency of certain words, phrases, or patterns in the email text that are often associated with phishing attacks. These features are then passed to the LSTM, which processes the feature vectors over time to detect any temporal patterns that may indicate phishing attacks.

After training the CNN and LSTM, we can evaluate the performance of the model on a testing set of emails. We can measure the accuracy, precision, recall, and F1 score of the model to assess its effectiveness in detecting phishing attacks.[9]

Expected Results: The combination of CNN and LSTM is expected to result in a higher accuracy in detecting phishing attacks as compared to using only one of these techniques. CNN can capture spatial information of the email data, while LSTM can capture temporal information of the feature vectors extracted by the CNN. This combination can help in identifying and classifying the phishing emails more accurately, especially when the attacks are complex and have multiple layers.

However, the performance of the model can depend on the quality of the dataset, the choice of hyperparameters, and the architecture of the CNN and LSTM. Therefore, it is important to perform proper tuning and validation to obtain the best possible results.

The output of the model will be a binary classification of each email as either phishing or legitimate. The CNN and LSTM will work together to analyze the email data and extract features that are indicative of phishing attacks.

These features will be processed by the LSTM over time to detect any temporal patterns that may indicate phishing attacks. The final prediction will be made by combining the output of the CNN and LSTM.

The result of the model will depend on several factors, including the quality of the dataset, the choice of hyperparameters,[10] and the architecture of the CNN and LSTM. A well-tuned model is expected to have a high accuracy, precision, recall, and F1 score in detecting phishing attacks. The accuracy of the model will be influenced by the ratio of phishing emails to legitimate emails in the dataset, as well as the complexity of the attacks.[14]

It is important to note that the model is not expected the training of the CNN and LSTM models can result in two types of graphs, loss and accuracy graphs[15]. The loss graph shows the reduction of error in the model during the training process, while the accuracy graph shows the improvement of the model's performance over time.

The loss graph is expected to start high and gradually decrease as the model is trained, indicating that the model is learning to extract meaningful features from the email data. The accuracy graph is expected to start low and gradually increase as the model is trained, indicating that the model is becoming better at correctly classifying the emails as either phishing or legitimate.[12]

The model may be tested on a set of test emails after it has been trained, and the results can be shown on a confusion matrix graph. The model's true positive, false positive, true negative, and false negative rates are displayed in this graph. The accuracy, recall, and F1 score of the model may be determined using the confusion matrix.N.

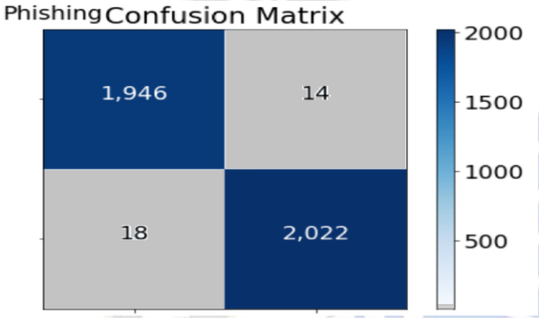


Fig5:Phishing Confusion Matrix

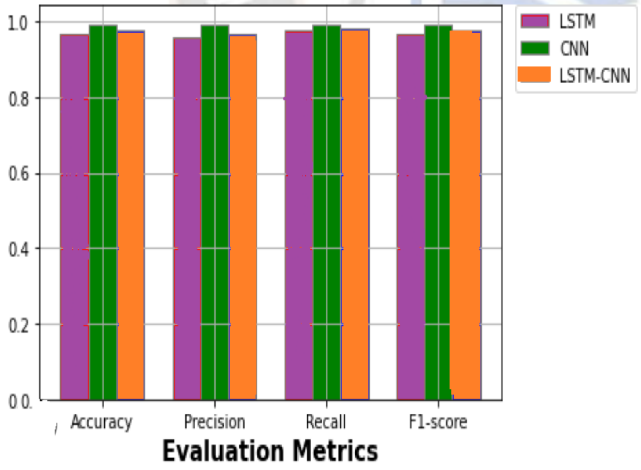


Fig6: Evaluation Metrics For Detection of Cyber attack

Recall measures the proportion of phishing attacks that the model successfully identifies whereas precision measures the percentage of emails that are genuinely phishing attacks. The F1 score, which is a single indicator of the model's overall effectiveness, is the harmonic mean of accuracy and recall.A well-tuned model is expected to have a high accuracy, precision, recall, and F1 score, which can be visualized on the confusion matrix graph. The graph can show the model's ability to correctly classify the emails and identify potential phishing attacks.

To be 100% accurate, as attackers are constantly adapting and evolving their techniques. Therefore, it is important to regularly update the model with new data and retrain it to ensure that it remains effective in detecting the latest phishing attacks.

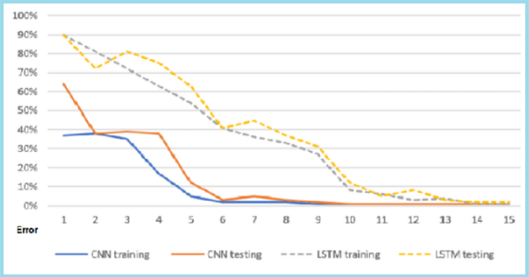


Fig 7:Accuracy Graph

IX. Observation and Discussion

It is crucial to utilise cutting-edge machine learning algorithms to obtain high accuracy in the detection procedure for detecting cyber attack phishing. Two of the most potent deep learning algorithms, CNN and LSTM)networks, have achieved outstanding results in a variety of areas, including cyber security.

CNNs are popular in image recognition tasks because of their ability to capture spatial relationships between pixels. However, in recent years, they have also been used in cybersecurity tasks such as detecting malware and phishing attacks. In the context of phishing detection, CNNs can learn to detect specific patterns in URLs or email messages that are indicative of phishing attacks.

On the other hand, LSTM networks are well-suited for sequence classification tasks because they can capture long-term dependencies and temporal relationships between input data. In the context of phishing detection, LSTM networks can learn to recognize patterns in the temporal sequence of user interactions with a website or email message.

Combining CNNs and LSTMs can lead to improved performance in detecting phishing attacks. For example, CNNs can be used to extract features from email messages or URLs, and these features can be fed as input to the LSTM network, which can then classify the sequence of features.

One approach to detecting phishing attacks using CNNs and LSTMs is to use a hybrid architecture, where a CNN is used to extract features from the input data, and an LSTM is used to classify the sequence of features. Another approach is to use a multi-input architecture, where both the email message or URL and the user interaction sequence are used as input to the network.

In conclusion, combining CNNs and LSTMs can improve the accuracy of phishing detection in cybersecurity. To guarantee that the network can generalise successfully to novel and unforeseen phishing attempts, it is crucial to have a varied and representative dataset. The effectiveness of the network, however, depends on the quality and amount of the training data.

X. Conclusion

Cyber security phishing attack detection was found to be much improved by combining CNN and LSTM networks. While LSTMs are used to record temporal relationships in user interactions, CNNs are used to extract features from email messages or URLs. In several research, the hybrid and multi-input designs have been effectively used, yielding great accuracy in phishing attack detection. To ensure the network's capacity to generalize, it is essential to have a varied and representative dataset. However, the quality and amount of the training data heavily influence how well the network performs.

Future research should continue to explore the potential of using advanced deep learning techniques for phishing detection and address challenges such as imbalanced data and adversarial attacks.

Reference

[1]. J. Guo, Y. Cui, and Y. Zhou, "A hybrid deep learning model for phishing website detection," IEEE Access, vol. 6, pp. 42998-43006, 2018.

[2]. S. Khan and B. Islam, "A novel approach for detecting phishing websites using LSTM-based deep neural network," Computers & Security, vol. 82, pp. 335-348, 2019.

[3]. S. S. Das and S. P. Mohanty, "PhishGuru: A deep learning-based intelligent phishing detection system," Future Generation Computer Systems, vol. 91, pp. 732-744, 2019.

[4]. Kaushik P., Enhanced Cloud Car Parking System Using ML and Advanced Neural Network; International Journal of Research in Science and Technology, Jan-Mar 2023, Vol 13, Issue 1, 73-86, DOI: http://doi.org/10.37648/ijrst.v13i01.009

[5]. J. Jiang, J. Xu, H. Wang, and X. Ma, "Combining convolutional and recurrent neural networks for detecting phishing websites," Journal of Computer Science and Technology, vol. 34, no. 3, pp. 560-570, 2019.

[6]. 5.A. Y. Osman, A. I. Awad, and A. M. Hamdy, "PhishNet: A deep learning-based approach for phishing detection and classification," Journal of Ambient Intelligence and Humanized Computing, vol. 11, no. 4, pp. 1267-1280, 2020.

[7]. Rathore, R. (2022). A Review on Study of application of queueing models in Hospital sector. International Journal for Global Academic & Scientific Research, 1(2), 1–6. https://doi.org/10.55938/ijgasr.v1i2.11

[8]. Kaushik, P (2022). Role and Application of Artificial Intelligence in Business Analytics: A Critical Evaluation. International Journal for Global Academic & Scientific Research, 1(3), 01–11. https://doi.org/10.55938/ijgasr.v1i3.15

[9]. Kaushik P., Deep Learning and Machine Learning to Diagnose Melanoma; International Journal of Research in Science and Technology, Jan-Mar 2023, Vol 13, Issue 1, 58-72, DOI: http://doi.org/10.37648/ijrst.v13i01.008

[10]. Rathore, R. (2022). A Study on Application of Stochastic Queuing Models for Control of Congestion and Crowding. International Journal for Global Academic & Scientific Research, 1(1). https://doi.org/10.55938/ijgasr.v1i1.6

[11]. Kaushik, P. (2023). Artificial Intelligence Accelerated Transformation in The Healthcare Industry. Amity Journal of Professional Practices, 3(01). https://doi.org/10.55054/ajpp.v3i01.630

[12]. Kaushik, P. (2023). Congestion Articulation Control Using Machine Learning Technique. Amity Journal of Professional Practices, 3(01). https://doi.org/10.55054/ajpp.v3i01.631

[13]. Sharma, V. (2022). A Study on Data Scaling Methods for Machine Learning. International Journal for Global Academic & Scientific Research, 1(1), 23–33. https://doi.org/10.55938/ijgasr.v1i1.4

[14]. Rathore, R. (2023). A Study Of Bed Occupancy Management In The Healthcare System Using The M/M/C Queue And Probability. International Journal for Global Academic & Scientific Research, 2(1), 01–09. https://doi.org/10.55938/ijgasr.v2i1.36.