Feature-Optimized CNN-LSTM-BiLSTM Model for SMS Threat Detection

[1]Mr. E. Narasimhulu Naidu, [2]Dr. K. Venkataramana
[1]PG Scholar, VEMU Institute of Technology
[2]Professor, Dept. of MCA, VEMU institute of Technology

Abstract: **Smishing, a combination of SMS and phishing, is a growing cybersecurity threat where malicious messages trick users into revealing sensitive information. Traditional machine learning models struggle with high false-positive rates due to similarities between legitimate and fraudulent messages. This study enhances smishing detection by integrating CNN, LSTM, and Bidirectional LSTM (BiLSTM) to improve accuracy. The CNN extracts features, while LSTM learns patterns, and BiLSTM enhances predictive performance by considering both past and future contexts. This hybrid approach achieves superior accuracy compared to existing models, reducing false positives. The extended model enhances security in mobile communications, providing a more reliable and efficient solution for smishing detection.**

Keywords: Smishing, Cybersecurity

INTRODUCTION

The increasing reliance on smartphones for banking, messaging, and daily activities has led to a rise in cyber threats, particularly smishing attacks. Smishing involves sending deceptive SMS messages containing malicious URLs that, when clicked, steal sensitive user data or install harmful software. As mobile users are more likely to trust SMS over emails, attackers exploit this trust to execute phishing schemes effectively.

Various rule-based and machine-learning approaches have been proposed to detect smishing messages. However, many existing models suffer from high false-positive rates due to similarities between legitimate (HAM) and fraudulent messages. This study addresses this limitation by leveraging deep learning techniques, specifically CNN and LSTM, to enhance detection accuracy. CNN optimizes feature extraction, while LSTM improves sequence learning for better classification. Additionally, integrating a Bidirectional LSTM (BiLSTM) further enhances performance. This research evaluates and compares the proposed model against traditional methods, demonstrating its effectiveness in minimizing false-positive rates.

LITERATURE SURVEY

- Goel and Jain (2017) - Smishing-Classifer Model
  - Focus: Detection of smishing attacks in mobile environments using a novel rule-based approach.
  - Method: Implemented a classification framework using Naïve Bayes and keyword-based feature extraction.
- Sonowal and Kuppusamy (2018) - SmiDCA Model
  - Focus: Machine learning-based smishing detection using a correlation algorithm.
  - Method: Feature extraction from 39 attributes in smishing messages combined with Random Forest classification.
- Joo et al. (2017) - S-Detector Model
  - Focus: Enhanced security model for detecting smishing attacks on mobile devices.
  - Method: Employed an SMS monitoring system with a Naïve Bayes classifier.
- Jain and Gupta (2018) - Rule-Based Smishing Detection
  - Focus: Framework for smishing message detection using nine predefined classification rules.
  - Method: Integrated heuristic-based detection with classification models like PRISM.
- Mishra and Soni (2019) - Content-Based Smishing Detection
  - Focus: Detection of smishing messages using keyword analysis and contextual assessment.
  - Method: Applied machine learning models with natural language processing (NLP) for message classification.

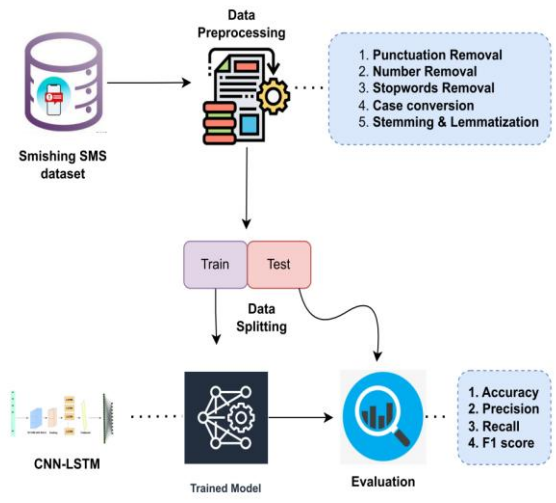
PROBLEM STATEMENT

Smishing attacks have become a significant cybersecurity challenge, exploiting SMS platforms to deceive users into revealing sensitive data. Traditional detection methods, including rule-based and machine learning algorithms, struggle with high false-positive rates, often misclassifying legitimate messages as threats. This results in decreased trust in detection systems and inefficiencies in cybersecurity defenses. The challenge lies in developing a model that accurately distinguishes between normal and fraudulent SMS while minimizing false positives.

PROPOSED METHOD

This study proposes an advanced smishing detection model by integrating CNN, LSTM, and BiLSTM for superior accuracy and reduced false positives. CNN extracts key features from SMS text, optimizing data representation. LSTM then processes sequential information, identifying hidden patterns in text-based phishing attacks. To further enhance accuracy, BiLSTM incorporates bidirectional processing, capturing dependencies from both past and future inputs. The model is trained on a combined dataset from Kaggle and Mendeley, processed using TF-IDF for feature extraction. Comparative analysis with traditional machine learning models, such as SVM and Random Forest, highlights the proposed model's superior performance in reducing false-positive classifications.

ARCHITECTURE



METHODOLOGY

Dataset Collection and Integration

The first step in the methodology involves gathering datasets to train and evaluate the smishing detection model. The datasets used in this study were collected from Kaggle and Mendeley, which contain SMS messages labeled as HAM (legitimate), SPAM, and SMISHING (fraudulent). Since smishing messages are relatively fewer in most datasets, combining multiple sources helps in creating a more balanced and diverse dataset. The merged dataset ensures that the model has sufficient examples to learn patterns associated with smishing attacks while reducing bias in classification.

Data Preprocessing and Cleaning

Before training the model, the raw SMS data undergoes preprocessing to remove inconsistencies and improve model performance. The preprocessing steps include removing special characters, numbers, and stopwords from the SMS text. Additionally, stemming and lemmatization techniques are applied to standardize the text, ensuring that words with similar meanings are treated uniformly. Tokenization is then performed to convert the text into structured input for the deep learning model. These steps help in improving the accuracy of the feature extraction process.

Feature Extraction Using TF-IDF

After preprocessing, the SMS messages are transformed into numerical representations using the Term Frequency-Inverse Document Frequency (TF-IDF) method. TF-IDF assigns weights to words based on their importance within a message, helping the model differentiate between commonly used words and crucial terms related to smishing. This transformation ensures that the text data is structured in a format suitable for deep learning models, improving classification performance.

Splitting Data into Training and Testing Sets

To evaluate the model effectively, the dataset is divided into training and testing sets using an 80-20 split. The training set is used to teach the model patterns associated with smishing, while the testing set is reserved for validation. A randomized split is applied to ensure that both datasets contain a representative distribution of HAM, SPAM, and SMISHING messages. This step helps in preventing overfitting,

where the model memorizes training data but fails to generalize to new SMS messages.

Implementation of CNN for Feature Extraction

The first deep learning component of the proposed model is a Convolutional Neural Network (CNN), which is used for feature extraction. CNN processes the transformed SMS data and identifies crucial patterns that distinguish smishing messages from legitimate ones. The extracted features are then passed to the next layers of the model for sequence learning and classification. CNN's role in the model is to optimize feature selection, reducing the chances of misclassification due to word similarities.

Integration of LSTM for Sequence Learning

Long Short-Term Memory (LSTM) networks are used to process the sequential dependencies within SMS messages. Since smishing messages often follow specific patterns, LSTM helps the model capture long-range dependencies between words. By analyzing the relationships between different terms in a message, LSTM improves the classification accuracy of the model, making it more effective in detecting fraudulent messages.

Enhancement with BiLSTM for Improved Accuracy

To further enhance performance, the study extends the CNN-LSTM model by integrating a Bidirectional LSTM (BiLSTM) layer. Unlike standard LSTM, which processes data in one direction, BiLSTM analyzes text sequences in both forward and backward directions. This bidirectional approach allows the model to capture more contextual information, improving its ability to distinguish between smishing and legitimate messages. The BiLSTM component significantly reduces the false-positive rate, ensuring that real SMS messages are not mistakenly classified as smishing.

Model Training and Optimization

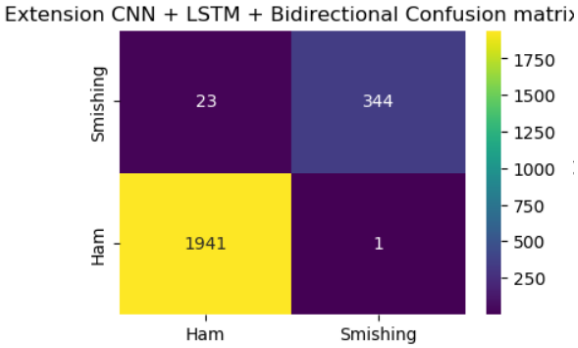
The CNN-LSTM-BiLSTM model is trained using the processed dataset, with optimization techniques applied to enhance its accuracy. The Adam optimizer is used for adjusting model weights, ensuring faster convergence and better performance. During training, hyperparameters such as learning rate, batch size, and the number of epochs are fine-tuned to achieve optimal results. Dropout layers are also included to prevent overfitting, making the model more robust when classifying new SMS messages.

Performance Evaluation and Comparison

After training, the model is evaluated using key performance metrics, including accuracy, precision, recall, and F1-score. A confusion matrix is generated to analyze misclassifications, particularly false positives and false negatives. The results of the CNN-LSTM-BiLSTM model are compared with traditional machine learning models such as Support Vector Machine (SVM) and Random Forest, as well as existing deep learning models like GRU, LSTM, and BiLSTM. The extended model demonstrates superior accuracy, with reduced false-positive rates, making it a reliable solution for smishing detection.

RESULTS

CNN + LSTM + Bidirectional



#implementing extension algorithm by combining 3 different models such as CNN + LSTM + Bidirectional where cnn will be used, CNN + LSTM + Bidirectional Accuracy : 98.96058899956691

Predict SMS as HAM or Smishing

Test Data = Think ur smart ? Win å£200 this week in our weekly quiz, text PLAY to 85222 now!T&Cs WinnersClub PO BOX 84, M26 3UZ. 16+. GBP1.50/week Predicted As ==➔ Smishing

Test Data = He says he'll give me a call when his friend's got the money but that he's definitely buying before the end of the week Predicted As ==➔ Ham

CONCLUSION

Smishing detection remains a critical challenge due to the increasing sophistication of cyber threats.

Traditional models often misclassify legitimate messages, leading to high false-positive rates. This study enhances smishing detection by integrating CNN, LSTM, and BiLSTM, achieving higher accuracy and reducing false alarms. By optimizing feature extraction and sequence learning, the proposed model outperforms existing machine-learning techniques. The results demonstrate improved precision and recall, making the model a reliable tool for mobile security. Future research could explore real-time deployment and integration with mobile security applications to further strengthen defenses against evolving smishing tactics.

REFERENCE

[1] S. Bajpai and D. Radha, ‘‘Smart phone as a controlling device for smart home using speech recognition,’’ in Proc. Int. Conf. Commun. Signal Process. (ICCSP), Apr. 2019, pp. 701–705.
[2] R. E. Madrid, F. A. Ramallo, D. E. Barraza, and R. E. Chaile, ‘‘Smartphone-based biosensor devices for healthcare: Technologies, trends, and adoption by end-users,’’ Bioengineering, vol. 9, no. 3, p. 101, Mar. 2022.
[3] K. K. Ibrahim and A. J. Obaid, ‘‘Fraud usage detection in internet users based on log data,’’ Int. J. Nonlinear Anal. Appl., vol. 12, no. 2, pp. 2179–2188, 2021.
[4] G. Sonowal and G. Sonowal, ‘‘Introduction to phishing,’’ in Phishing and Communication Channels: A Guide to Identifying and Mitigating Phishing Attacks. Springer, 2022, pp. 1–24.
[5] G. Desolda, L. S. Ferro, A. Marrella, T. Catarci, and M. F. Costabile, ‘‘Human factors in phishing attacks: A systematic literature review,’’ ACM Comput. Surv., vol. 54, no. 8, pp. 1–35, Nov. 2022.
[6] G. Brown and P. M. Greenfield, ‘‘Staying connected during stay-at-home: Communication with family and friends and its association with wellbeing,’’ Hum.