A Deep Learning-Based Phishing Detection System Using CNN,

LSTM, and LSTM-CNN Sandhya Mishra, Devpriya Soni

Zainab Alshingiti 1, Rabeah Alaqel 1, Jalal Al-Muhtadi 1,2, Qazi Emad Ul Haq 3,\*, Kashif Saleem 2

and Muhammad Hamza Faheem 3

Published: 3 January 2023

Abstract:

The abstract discusses a deep learning-based phishing detection system that utilizes Convolutional Neural Networks (CNN), Long Short-Term Memory (LSTM), and a combined LSTM-CNN approach to identify phishing websites. The system demonstrates high accuracy rates for each method: 99.2% for CNN, 97.6% for LSTM-CNN, and 96.8% for LSTM. The study highlights the effectiveness of these deep learning techniques in detecting phishing attacks, with the CNN-based method showing superior performance in comparison.

Introduction:

NEWS : The New Indian Express. Increasing cybercrime: UN reports 350 per cent rise in phishing websites during a pandemic. 2020. https://www.newindianexpress.com/business/2020/aug/08/increasing -cybercrime-un-reports-350-per-cent-rise-in-phish

URL phishing is a cyber-attack that uses URLs and e-mails as a technique to trick

users into believing that the URL or e-mail is a trustworthy mechanism in electronic

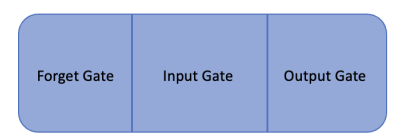
communication, such as a note from their company or a request from their bank, for

instance, to download the attachment or to click a link. At that moment, attackers are able

to access the user’s data. Furthermore, phishing websites or e-mails are designed to mimic

the look of a real company webpage/email.

**LSTM (long short-term memory)** is a form of recurrent neural network (RNN) that gains superior results when dealing with time-series data, removing vanishing gradients and long-term dependencies. The architecture of LSTM is made up of a cell and three gates (input, output, and forget) [17,18] as shown in Figure 1.



**A convolutional neural network** is a kind of neural network that requires large, labeled data for training. CNNs play a significant role in many problems such as image classification, object recognition, phishing detection, and diagnosis of medical diseases. Input, convolution, pooling, and fully connected layers are the main layers needed to construct a CNN as shown in Figure 2. Accelerating the learning process has led CNN to accomplish great and high results for man.

A diagram of a structure

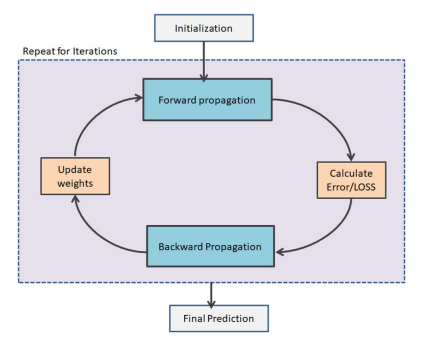
Description automatically generated

**LSTM–CNN architecture involves both CNN and LSTM** methods as shown in Figure 3 in order to make use of the benefits of both methods and accomplish excellent performance. Since CNN and LSTM show high performance in overcoming classification, detection, and recognition tasks [17], to using these three methods for the phishing detection task is promising.

A diagram of a diagram of a structure

Description automatically generated with medium confidence

An examination of the methods currently used to identify phishing websites.• Analysis and use of three state-of-the-art deep learning methods, LSTM, CNN, andLSTM–CNN, to predict phishing URLs.• Presentation of an efficient deep learning architecture based on CNN due to its capacityto identify patterns, extract features, and automatic and accurate classification of URLs.• Comparison and evaluation of suggested LSTM, CNN, and LSTM–CNN models.• Consideration of a dataset with 30 features after a feature selection process.• Highlighting several restrictions based on the conclusions of earlier investigations andsuggestion of potential fixes for these issues.



The introduction section mentions that Artificial Neural Networks (ANNs) are employed as a method to improve the detection of smishing attacks. ANNs are highlighted for their ability to process and classify large volumes of data, making them suitable for distinguishing between benign and malicious text messages. The section underscores the potential of ANNs in enhancing the accuracy and efficiency of smishing detection mechanisms, contributing to the development of more effective cybersecurity solutions.

* Artificial Neural Network is implemented for ‘Smishing Detector’ model.
* The best 7 features of the Smishing SMS are extracted using Neural Network.
* The accuracy of each feature is reported for efcient smishing detection.
* The algorithm of the ‘Smishing Detector’ [11]model is presented.
* The ‘Smishing Detector’[11] model is implemented using Neural Network and its evaluation using real-time

Proposed approach:

A diagram of data processing

Description automatically generatedEvaluation Results:

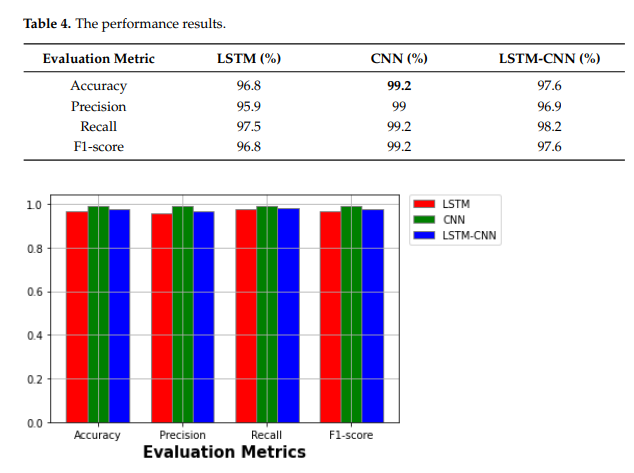
* **Evaluation Metric :**

**Precision** = TP / TP + FP = True positive Total / predicted positive

**Recall: =** Recall = TP / TP + FN = True positive Total / predicted positive

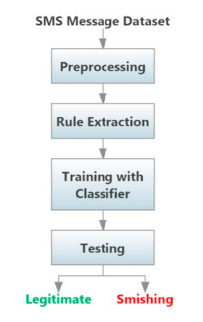
**Accuracy =** Accuracy = TP + TN / TP + TN + FN + FP

**F1-Score:** The process of taking the harmonic mean of a classifier’s precision and recall. It can be combined into a single metric. F1 − score = 2 × (Precision × Recall) (Precision + Recall)

****

**Text Normalization:** This step involves processing the SMS text to normalize its format. Given the informal nature of SMS language, this step ensures that the text is in a suitable form for analysis.

* **Application of Classification Rules:** The core of the framework lies here. Nine specific rules are applied to the normalized text. These rules are designed based on common characteristics observed in smishing attacks.
* **Rule Examples:**
  + Presence of URLs or shortened links.
  + Inclusion of phone numbers.
  + Urgent or alarming language prompting immediate action.
  + Requests for sensitive personal information.
  + Use of promotional or enticing language.
  + Classification Process: Using algorithms like Decision Tree, RIPPER, and PRISM, the SMS is analyzed under these rules to classify it.
* **Categorization into 'Smishing' or 'Legitimate':** Based on the analysis, each SMS is categorized as either a smishing message or a legitimate one.
* **Effective Detection:** The framework is designed to efficiently identify smishing messages, thereby protecting users from potential phishing attacks via SMS.
* **High True Negative Rate:** The proposed system is shown to have a high true negative rate, meaning it effectively identifies non-smishing messages correctly, reducing false positives.
* **Zero-hour Attack Detection:** The framework is also capable of detecting new, previously unseen smishing attacks (zero hour attacks).



Experimental Evaluation:

**Dataset:** The authors utilized a dataset comprising both smishing and legitimate SMS messages. This dataset was essential for training and testing the classification models.**Preprocessing and Normalization:** Before applying the classification rules, the SMS messages underwent preprocessing and normalization to standardize the text, making it suitable for analysis.**Application of Rules:** The nine specific rules identified for smishing detection were applied to the dataset. These rules were based on common characteristics of smishing messages, such as the presence of URLs, phone numbers, and certain keywords.**Classification Algorithms:** The paper details the use of several rule-based classification algorithms, including Decision Tree, RIPPER, and PRISM, to evaluate the effectiveness of the proposed framework.**Performance Metrics:** The evaluation focused on various performance metrics, such as accuracy, precision, recall, and the true negative rate. These metrics provided a quantitative measure of the framework's effectiveness in distinguishing between smishing and legitimate messages.