CNN-LSTM: Hybrid Deep Neural Network for Network Intrusion Detection System

(i) Abstract

This research investigates creating a hybrid deep learning approach by merging Convolutional Neural Networks (CNN) and Long Short-Term Memory (LSTM) networks to enhance an Intrusion Detection System (IDS), tested on the UNSW-NB15 dataset. With the evolution of cyber threats, traditional IDS solutions are not effective in identifying new, sophisticated attacks. The model's goal is to improve detection accuracy and lower false alarm rates by combining CNN's spatial feature extraction with LSTM's temporal sequence learning abilities. The model underwent thorough testing and showed better performance than conventional machine learning methods, achieving high accuracy and detection rates.

(ii) Introduction

Project Objectives:

The main objective of this project is to create a successful IDS by implementing a combined CNN-LSTM model that can accurately detect various network intrusions while maintaining low false alarm rates. The precise goals are:

1. Develop a mixed deep learning model that combines CNN for extracting spatial features and LSTM for recognizing temporal patterns.
2. In order to assess the model with the UNSW-NB15 dataset, which consists of a wide range of network traffic, including different types of malicious behavior.
3. To evaluate the hybrid model's performance in relation to traditional machine learning methods and individual deep learning models such as CNN or LSTM.

Problem Formulation:

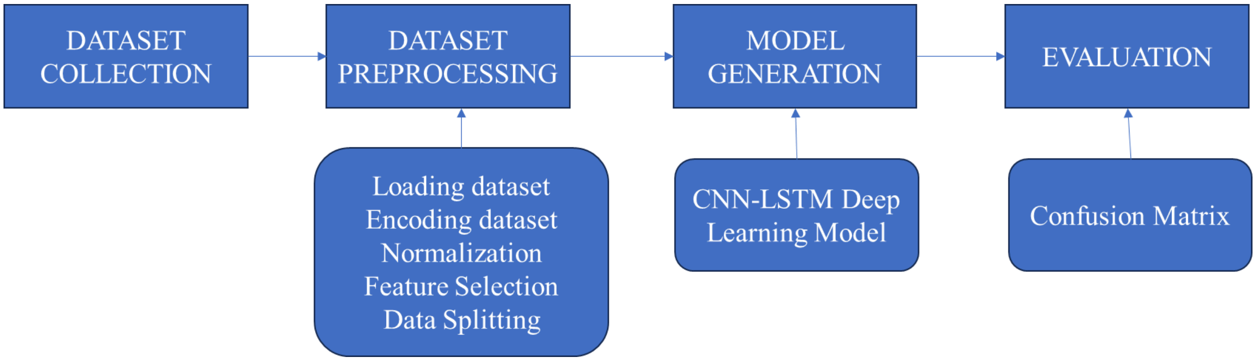
Intrusion Detection Systems (IDS) are vital in safeguarding network security, preventing unauthorized access, and upholding data integrity and confidentiality. Conventional intrusion detection systems, like those based on signatures, are restricted by their dependence on familiar attack patterns and frequently do not identify new or developing threats. Anomaly detection models can pinpoint unfamiliar attack methods, but often have a problem with too many false alarms. Deep learning techniques, specifically CNN and LSTM networks, have demonstrated potential in tackling these difficulties. CNNs are great at extracting spatial features, whereas LSTMs are specifically built to manage sequential relationships in time-series information. The goal of the hybrid CNN-LSTM model is to enhance IDS performance by incorporating spatial and temporal features of network traffic.

(iii) Methodology

Dataset:

The UNSW-NB15 dataset is well known for its detailed coverage of contemporary network traffic, encompassing both legitimate and harmful actions. It includes nine different forms of attacks including DoS, Exploits, Fuzzers, Backdoor, and Analysis. The data was gathered from actual traffic sources, making it very suitable for assessing the IDS effectiveness.

Model Architecture:



The suggested IDS model combines CNN and LSTM networks, aiming to utilize the advantages of both architectures.

- **CNN Layers:** They are in charge of extracting spatial features from the input data. Convolutional layers use filters to detect important patterns in the data, such as irregularities or uncommon traffic activities. Pooling layers are employed to decrease the size of feature maps, thereby diminishing dimensionality and safeguarding against overfitting.
- **LSTM Layers:** Following the CNN processing, the derived features are transmitted to LSTM layers, which have the ability to understand extended connections and time-related trends in the ordered data. LSTM networks utilize memory cells and gating mechanisms to preserve important information for extended periods, improving the model's capability to identify attacks occurring across multiple time steps.
- **Dropout and Batch Normalization:** Dropout layers are utilized during training to randomly disable neurons, aiding in the prevention of overfitting. Batch normalization layers help to stabilize the learning process by normalizing the inputs of every layer, which enables quicker and more dependable training.
- The last layer of the model is a fully connected layer that employs the Softmax activation function to categorize network traffic into benign or different malicious types.

Data Preprocessing:

Several preprocessing steps were carried out in order to ready the UNSW-NB15 dataset for training.

- **Data Cleaning:** The process included eliminating any missing values and duplicate records from the dataset, guaranteeing the data's quality and uniformity.
- **Encoding Data:** Categorical variables, like traffic labels, were transformed into numerical values via One-Hot Encoding. This change enables the deep learning model to efficiently handle the data.
- **Normalization:** Normalization of data involved using Standard Scaler to ensure that input features have a mean of zero and a standard deviation of one. This stage is crucial for enhancing the efficacy of training and the speed at which the neural network converges.
- **Feature Selection:** By utilizing SelectKBest from the sklearn library, significant features were chosen based on their statistical relevance. This process of selecting features decreased the data's dimensionality, enabling the model to train with increased efficiency while still achieving high performance.

The data was split into training, validation, and testing sets with an 80-20 split, and further partitioned for validation using Stratified K-Fold Cross-Validation. This method guarantees that the model is trained and validated on data that accurately represents the dataset, minimizing the likelihood of overfitting.

Training and Evaluation:

The Adam optimizer, famous for its adaptive learning rate capabilities, was utilized to train the CNN-LSTM model. Key metrics were used to assess the performance.

- **Precision:** Evaluates how well the model correctly identifies network traffic. Precision evaluates the model's capability to prevent false positives, specifically benign traffic being incorrectly labeled as attacks.
- **Detection Rate (DR):** This shows how many actual attacks are accurately detected by the model.
- **False Alarm Rate (FAR):** This is a crucial metric for practical IDS deployment as it measures the rate of benign traffic that is wrongly identified as malicious. The assessment was carried out using a confusion matrix, offering a precise analysis of how the model performed with various kinds of traffic.

(iv) Results

The hybrid CNN-LSTM model showed significant enhancements in intrusion detection compared to conventional and individual models. The outcomes on the UNSW-NB15 dataset were the following:

- **Precision:** The model obtained a general precision rate of 93.95%, surpassing alternative techniques like SVM (82.42%), Decision Trees (89.08%), and individual CNN or LSTM models.

- **Detection Rate:** The model showed a strong detection rate of 94.53%, successfully recognizing the majority of attack types present in the dataset.
- **Precision and FAR:** The model achieved a precision rate of 94.69%, effectively reducing false positives, and it also reached a critical false alarm rate of 6%, crucial for upholding the IDS's reliability.

Comparison with Other Models:

When compared to other models, the CNN-LSTM hybrid consistently demonstrated superior performance in accuracy and false alarm rates. For example, models like the Deep Belief Network (DBN) and Autoencoder-based Deep Neural Networks (ICVAE-DNN) exhibited lower accuracy and higher FAR in contrast to the hybrid method.