AI Based Network Intrusion Detection System – Multi Layer Perceptron Approach

By Seran Gemechu
College of Science – Department of Computer Science (Artificial Intelligence)
San Jose State University
San Jose, CA
Seran.gemechu@sjsu.edu
March 13, 2025

Abstract - This paper presents an AI-based network intrusion detection system (IDS) using a deep learning approach—specifically, a multilayer perceptron (MLP) using PyTorch. Leveraging the NSL-KDD dataset, an improved variant of the original KDD'99 dataset, my proposed method addresses data redundancy and imbalance issues inherent in the older dataset. I provide a comprehensive literature review and detail the advantages of the NSL-KDD dataset, including its elimination of duplicate records and more balanced representation of attack types. Experimental results indicate an overall accuracy of 98.86%, with high precision, recall, and F1-scores for the detection of intrusions. I further discuss the limitations of the current approach, outline potential future enhancements, and explore its applications in real-world network security environments.

1. Introduction

Network security is an ever-growing concern as cyber-attacks become increasingly sophisticated. Intrusion detection systems (IDS) are pivotal in identifying and mitigating such threats. Traditional IDS methods have relied on signature-based techniques, which often fail to detect novel or evolving attack patterns. To overcome these limitations, anomaly-based IDS approaches have emerged, wherein deviations from normal (0) network behavior are flagged as potential intrusions (1).

This paper focuses on an anomaly-based IDS utilizing a deep learning framework with a multilayer perceptron (MLP). By applying the NSL-KDD dataset, I exploit its improved characteristics over the traditional KDD'99 dataset, thereby enabling more accurate training and evaluation of my detection model. The remainder of the paper is organized as follows: Section 2 reviews related work; Section 3 details the methodology; Section 4 presents experimental results; Section 5 discusses my findings, limitations, and real-world applications; and Section 6 concludes the paper with future research directions.

2. Related Work

In recent years, several studies have investigated the use of machine learning and deep learning for network intrusion detection. Early work centered on classical algorithms such as support vector machines and decision trees [1]. However, these methods were often limited by their reliance on pre-defined features and the inability to adapt to complex attack patterns.

The evolution towards deep learning has brought forth models such as convolutional neural networks (CNNs) and recurrent neural networks (RNNs) for IDS applications [2]. Among these, the multilayer perceptron (MLP) has been widely adopted due to its relative simplicity and effectiveness in capturing non-linear relationships in data. Several studies have demonstrated that MLPs, when paired with proper preprocessing and normalization techniques, can significantly enhance detection rates for both known and unknown attacks [3].

Furthermore, the NSL-KDD dataset has been embraced as a benchmark for evaluating IDS models. Unlike the KDD'99 dataset, which suffers from high redundancy and imbalance, NSL-KDD offers a more rigorous testbed by

removing duplicate records and ensuring a more equitable distribution of attack types [4]. This dataset, along with a variety of preprocessing techniques and neural network architectures, has paved the way for robust IDS solutions in both academic and practical settings.

3. Methodology

3.1 Data Collection and Preprocessing

The NSL-KDD dataset, derived from the KDD'99 dataset, serves as the cornerstone for my experiments. This dataset contains network traffic records that include both normal and attack behaviors. Key improvements of NSL-KDD over KDD'99 include:

- Elimination of Redundant Records: The training set is free from duplicate records, reducing bias towards more frequent instances.
- Balanced Test Sets: Test sets are curated without duplicate records, ensuring that evaluation metrics are not skewed by frequency biases.
- Proportional Sampling: Records are selected inversely proportional to their frequency in the original dataset, leading to a wider range of classification outcomes [4].

Data preprocessing involves merging the training and test sets, splitting the combined dataset into new training 67%, validation 13%, and test 20% sets, encoding categorical features using one-hot encoding, and normalizing numerical features to ensure uniformity across input dimensions.

3.2 Neural Network Model: MLP Architecture

The core of my IDS is an MLP model designed to classify network traffic as normal (0) or anomalous (1). The model architecture includes:

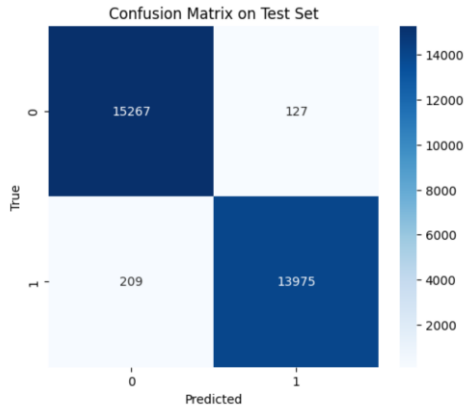
- Input Layer: Receives preprocessed features from the NSL-KDD dataset.
- Hidden Layers: Multiple hidden layers with non-linear activation functions (e.g., ReLU) to capture complex patterns.
- Output Layer: A Sigmoid layer that outputs the probability of each class (normal or intrusion).

The model is trained using standard forward propagation, with the loss computed by a categorical cross-entropy function. Backpropagation is employed to update weights iteratively, using an optimization algorithm such as Adam.

3.3 Training and Evaluation

The training process involves iterative learning over multiple epochs. The dataset is partitioned to include a validation set that aids in monitoring overfitting and fine-tuning hyperparameters. The evaluation metrics include:

- Accuracy
- Precision
- Recall
- F1-Score



My final experimental results reveal:
- True Negatives (TN) = 15,267
- False Positives (FP) = 127
- False Negatives (FN) = 209
- True Positives (TP) = 13,975

```
------------------Classification Report on Test Set------------------
              precision    recall  f1-score   support

         0.0       0.99      0.99      0.99     15394
         1.0       0.99      0.99      0.99     14184

    accuracy                           0.99     29578
   macro avg       0.99      0.99      0.99     29578
weighted avg       0.99      0.99      0.99     29578
```

Overall accuracy reached 98.86%, with precision, recall, and F1-scores approximating 99% for the intrusion detection class.

4. Experiments & Results

4.1 Experimental Setup

The experiments were conducted on Azure Machine Learning Studio, leveraging its scalable and efficient computational resources. This environment provided the necessary infrastructure to preprocess the NSL-KDD

dataset, optimize hyperparameters, and train the multilayer perceptron (MLP) model. Azure Machine Learning Studio facilitated rapid experimentation and iterative improvements, ensuring that the deep learning model was rigorously evaluated under real-world cloud-based conditions. The training process was executed using the preprocessed NSL-KDD dataset. Hyperparameters such as learning rate, number of epochs, and batch size were optimized using the validation set.

4.2 Performance Metrics

The performance metrics of the model are summarized as follows:

- Accuracy: 98.86%
- Precision (Intrusion Class): 99.09%
- Recall (Intrusion Class): 99%
- F1-Score (Intrusion Class): 99%

These metrics indicate a robust performance of the MLP in distinguishing between normal and anomalous network traffic.

4.3 Comparative Analysis

Compared to earlier approaches using traditional machine learning methods, the deep learning-based MLP model demonstrates superior accuracy and a more balanced detection rate. The NSL-KDD dataset's inherent improvements contribute significantly to these results by mitigating issues related to data redundancy and class imbalance.

5. Discussion

5.1 Significance of the NSL-KDD Dataset

The NSL-KDD dataset provides a more realistic and challenging environment for evaluating IDS models. Its elimination of redundant records and balanced test sets ensures that the detection performance is not artificially inflated by overrepresented classes. This dataset enables a more thorough evaluation of deep learning techniques, which require diverse and non-biased data to generalize effectively.

5.2 Model Performance and Limitations

While the MLP model demonstrates high accuracy, it is essential to acknowledge certain limitations:

- Data Dependency: The model's performance is closely tied to the quality and representativeness of the NSL-KDD dataset. Real-world networks may exhibit behaviors not captured in the dataset.
- Scalability: Deep learning models, including MLPs, can be computationally intensive, posing challenges for real-time intrusion detection in high-throughput environments.
- Feature Sensitivity: Although the preprocessing steps mitigate many issues, the selection and normalization of features play a critical role in the model's efficacy.

5.3 Future Work

Future research may explore:

- Hybrid Models: Combining MLPs with other deep learning architectures (e.g., CNNs, RNNs) to capture both spatial and temporal features of network traffic.
- Real-Time Implementation: Enhancing the model to operate in real-time, incorporating streaming data analysis techniques.
- Broader Datasets: Evaluating the approach on additional datasets and real-world network traffic to further validate the model's generalizability.
- Explainability: Integrating explainable AI methods to provide insights into the decision-making process of the IDS.

5.4 Real-World Applications

The proposed IDS can be integrated into various network security infrastructures, including:

- Enterprise Networks: To monitor and mitigate internal and external threats.
- Critical Infrastructures: Protecting vital systems such as power grids, transportation, and healthcare.
- Cloud Environments: Providing enhanced security for cloud-based applications and services by detecting anomalous behaviors promptly.

6. Conclusion

This paper presents an effective anomaly-based IDS using an MLP deep learning model trained on the NSL-KDD dataset. The enhancements offered by NSL-KDD over the original KDD'99 dataset enable a more accurate and reliable evaluation of the IDS. With an overall accuracy of 98.86% and high precision, recall, and F1-scores for intrusion detection, my approach

demonstrates significant promise. Nevertheless, limitations regarding data representativeness, scalability, and feature sensitivity suggest avenues for future research. Further work on hybrid models, real-time implementation, and explainability will be essential to advance IDS technology for practical, real-world applications.

References

[1] Lippmann, R. et al., "Evaluating intrusion detection systems: The 1998 DARPA off-line intrusion detection evaluation," Proceedings of the DARPA Information Survivability Conference and Exposition, 2000.
[2] Kim, G., Lee, S., & Kim, S., "A novel hybrid intrusion detection method integrating anomaly detection with misuse detection," Expert Systems with Applications, vol. 41, no. 4, pp. 1690-1700, 2014.
[3] Javaid, A. et al., "A deep learning approach for network intrusion detection system," Computers & Security, vol. 74, pp. 118-128, 2018.
[4] Tavallaee, M. et al., "A detailed analysis of the KDD CUP 99 data set," in Proceedings of the IEEE Symposium on Computational Intelligence for Security and Defense Applications, 2009.
[5] Mukkamala, S., Janoski, G., & Sung, A. H., "Intrusion detection using an ensemble of intelligent paradigms," Journal of Network and Computer Applications, vol. 28, no. 2, pp. 167-182, 2005.
[6] Bagheri, E. et al., "A survey on network anomaly detection: methods, techniques, and challenges," IEEE Communications Surveys & Tutorials, vol. 21, no. 4, pp. 3094-3118, 2019.