iJRASET
International Journal For Research in Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 14 **Issue:** IV **Month of publication:** April 2026

DOI: https://doi.org/10.22214/ijraset.2026.80890

www.ijraset.com

Call: 08813907089 | E-mail ID: ijraset@gmail.com

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 14 Issue IV Apr 2026- Available at www.ijraset.com

AI-Based Network Intrusion Detection System

Anuradha Tiwari[1], Sandeep Bind[2], Ajeet Kanaujia[3], Neeraj Pal[4], Sameer Awasthi[5]
Department of Computer Science and Engineering (AIML) Bansal Institute of Engineering and Technology, Lucknow, India

Abstract: ***The fast development of digital infrastructure has contributed to the rising number of sophisticated cyberattacks. Classical Intrusion Detection Systems are predominantly based on signatures, which makes it easy to detect previously known cyberattacks but fail to find the novel ones. In this paper, we develop an AI-based Network Intrusion Detection System (NIDS) using the combination of the LSTM algorithm and Random Forest.***
Long Short-Term Memory helps in capturing temporal char-acteristics of network attacks, whereas Random Forest enhances the robustness of the classification. We use NSL-KDD dataset for testing the model and conduct an evaluation using accuracy, precision, recall, and F1 score. The hybrid model outperforms the models trained separately by detecting intrusions faster and minimizing the false positives.
The research shows the effectiveness of using combined ma-chine learning and deep learning approaches in creating adaptive and scalable intrusion detection systems.
Index Terms: Network Intrusion Detection System, Artificial Intelligence, Machine Learning, Deep Learning, LSTM, Random Forest, Cybersecurity.

I. INTRODUCTION

As a consequence of increasing the usage of internet ser-vices, cloud computing, and devices interconnected through networks, the problem of network security becomes crucial. As organizations and private users rely on digital communication infrastructure, they become increasingly vulnerable to attacks, including such types as DoS (Denial of Service) attack, intrusion into the system, malware attacks, and data theft. The purpose of NIDS technologies is to analyze traffic and detect suspicious and harmful actions on the network. Signature-based NIDS solutions have been proven highly efficient in detecting well-known types of cybercrime. Never-theless, they are useless in cases of sophisticated and complex attacks. This leads to the necessity of developing advanced intrusion detection solutions.

In this context, AI techniques prove to be an innovative solution to the described problem. Machine learning classifiers allow learning discriminative features from traffic, whereas deep learning models help to identify hidden features and dependencies. One of the efficient deep learning algorithms used in this research is LSTM. Also, Random Forest algorithm has been chosen as a machine learning classifier due to high performance and resistance to overfitting.Thus, the proposed approach involves building a NIDS architecture based on LSTM and Random Forest algorithms.

The contribution of this

1) A hybrid LSTM–Random Forest intrusion detection framework is presented
2) A preprocessing and feature preparation pipeline is de-signed for network traffic analysis
3) The model is evaluated using the NSL-KDD benchmark dataset.
4) The work highlights the benefit of combining deep learn-ing and machine learning for network security.

II. OBJECTIVES

The objectives of this research are:

1) To analyze the limitations of traditional intrusion detec-tion approaches.
2) To design an AI-based system for identifying malicious network traffic.
3) To preprocess and normalize traffic data for improved learning.
4) To apply LSTM for temporal dependency analysis.
5) To use Random Forest for robust classification.
6) To reduce false positive rates while improving overall accuracy.
7) To detect both known and previously unseen attacks.
8) To develop a scalable framework for future real-time deployment.

ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 14 Issue IV Apr 2026- Available at www.ijraset.com

III. PROBLEM STATEMENT

The increased complexity, volume, and variety of network traffic have made intrusion detection a challenging task. To-day's cyber attacks are highly adaptive and usually evade tradi-tional detection strategies based on rules. Traditional intrusion detection systems find it hard to detect zero-day attacks, while they also tend to generate false alarms, thereby complicating the job of network managers and security experts. Moreover, high-speed network traffic needs a smart moni-toring system. There is an urgent requirement for developing an intruder detection strategy that can not only understand the nature of the network but can also detect the malicious attacks on the basis of learning and adaptation.

IV. LITERATURE REVIEW

There has been extensive utilization of machine learning as well as deep learning in the area of intrusion detection studies. The technique known as Random Forest has gained popularity owing to its high degree of accuracy in classification as well as stability along with the ability to manage high dimensional feature spaces. Similarly, LSTM algorithms prove to be very useful in terms of handling sequential and temporal data analysis. Recently, it has been shown that hybrid methods may perform better than individual methodologies. Machine learn-ing offers effective and accurate classification, whereas deep learning helps in capturing temporal relations. Taking this concept into account, this study utilizes LSTM algorithms along with the Random Forest technique.

V. METHODOLOGY

A. System Overview

The proposed AI-based Network Intrusion Detection System classifies network traffic into two classes:

- Normal Traffic (0)
- Attack Traffic (1)

The full process involves data preprocessing, feature selec-tion, temporal processing using LSTM, classification using random forest, and lastly, decision-making using ensemble-based reasoning. The full process involves data preprocessing, feature selection, temporal processing using LSTM, classifi-cation using random forest, and lastly, decision-making using ensemble-based reasoning.

B. Data Preprocessing

Data preprocessing improves the quality and consistency of the dataset before model training. The following operations are performed:

- Removal of missing values
- Elimination of duplicate records
- Handling of infinite or inconsistent values
- Conversion of labels into binary classes
- Feature normalization using Min-Max scaling

These preprocessing steps improve learning stability and reduce noise in the input data.

C. Feature Engineering

Meaningful numerical attributes are extracted from network traffic to improve classification performance. These may include packet-level statistics, duration-based information, flow characteristics, and other indicators representing communication behavior. Proper feature engineering improves model reliability and reduces irrelevant noise.

D. LSTM Model

LSTM is used to capture temporal dependencies in network traffic. Because traffic behavior unfolds over time, LSTM is suitable for identifying attack-related patterns that develop sequentially. The model processes reshaped input sequences and generates probability scores representing the likelihood of malicious behavior.

E. Random Forest Model

Random Forest is employed as a supervised classifier for robust decision making. It constructs multiple decision trees and combines their outputs through majority voting. This improves stability, reduces overfitting, and provides strong classification performance for high-dimensional data.

F. Hybrid Ensemble Model

The outputs of LSTM and Random Forest are combined through a weighted ensemble strategy. The final probability is computed as:

$$P_{final} = 0.7 \times P_{LSTM} + 0.3 \times P_{RF} \quad (1)$$

The final class label is determined as:

$$\text{Prediction} = \begin{cases} 1, & \text{if } P_{final} \geq 0.5 \\ 0, & \text{otherwise} \end{cases} \quad (2)$$

This combination allows the system to capture both tempo-ral and feature-based characteristics of network traffic.

VI. IMPLEMENTATION

A. Development Environment

The model is implemented using Python and widely used AI libraries:

- Python
- TensorFlow / Keras
- Scikit-learn
- NumPy
- Pandas
- Matplotlib

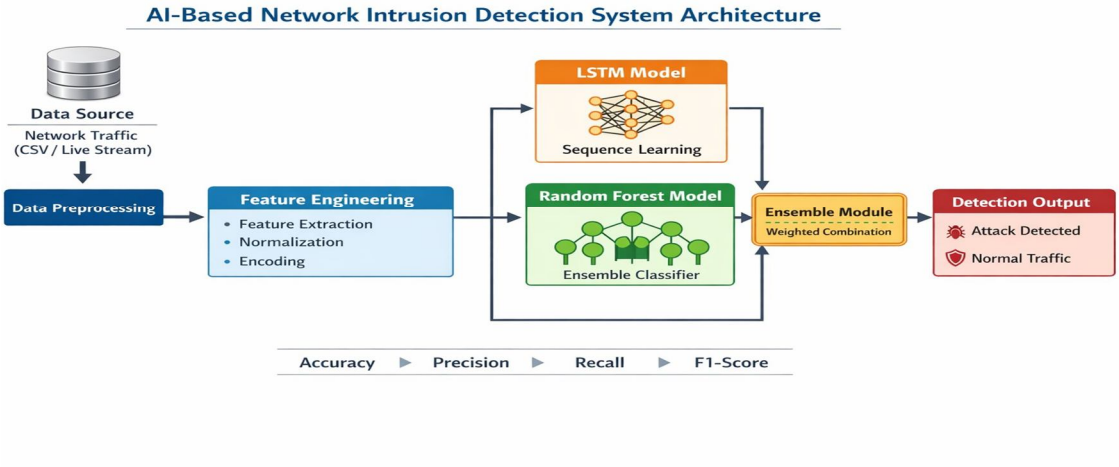


Fig. 1. Overall architecture of the proposed AI-based Network Intrusion Detection System.

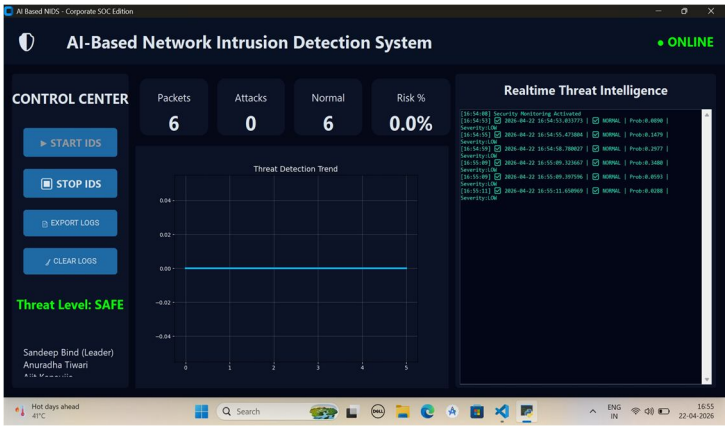


Fig. 2. Actual model working system.

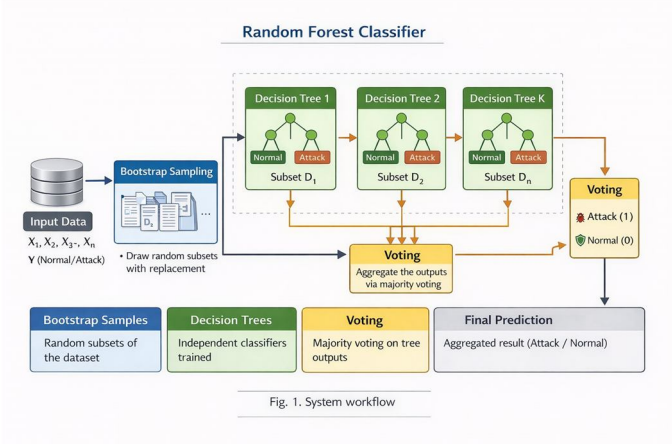


Fig. 3. Working principle of the Random Forest classifier for attack and normal traffic prediction.

B. Dataset Description

The proposed system is evaluated on the NSL-KDD dataset, which is an improved version of the KDD Cup 99 dataset. It is widely used in intrusion detection research because it reduces redundancy and offers a balanced benchmark for experimental evaluation.
The dataset includes the following categories:

- Denial of Service (DoS)
- Probe
- Remote-to-Local (R2L)
- User-to-Root (U2R)

Each record contains multiple features describing network behavior such as protocol type, connection attributes, and traffic statistics.

C. Model Configuration

LSTM model employs hidden layers and dropout for dealing with sequential data. On the other hand, the Random Forest classifier makes use of trees to classify data. Their results are aggregated by using the ensemble technique. Dataset is split into training and test datasets, while conventional measures are employed to evaluate performance.

VII. SYSTEM WORKFLOW

The overall workflow of the proposed system is:

1) Input network traffic data
2) Perform preprocessing and normalization
3) Extract and prepare relevant features
4) Apply LSTM for temporal analysis
5) Apply Random Forest for classification
6) Combine predictions using ensemble logic
7) Apply threshold to generate final output
8) Classify traffic as Attack or Normal

The framework can also support batch processing and maintain logs for analysis and monitoring.

VIII. EXPECTED OUTCOMES

The proposed model is expected to provide the following benefits:

1) Improved detection accuracy compared to standalone models
2) Reduced false positive rates
3) Better identification of sequential and complex attack patterns

4) Strong generalization across multiple attack categories
5) Efficient processing of network traffic data
6) Support for future near real-time deployment

By combining LSTM and Random Forest, the system aims to achieve balanced performance in terms of accuracy, preci-sion, recall, and F1-score.

IX. RESULTS AND DISCUSSION

The hybrid model improves intrusion detection by com-bining the strengths of deep learning and machine learning. LSTM captures temporal patterns in traffic streams, while Random Forest provides stable feature-based classification. Together, these components improve prediction quality and reduce the limitations observed in single-model approaches.

The integrated design is especially useful in dynamic net-work environments where attacks do not always follow fixed patterns. As a result, the proposed framework offers a more adaptive and intelligent strategy for intrusion detection.

Fig. 5 and Fig. 6 show the output screens of the proposed AI-based Network Intrusion Detection System.

X. LIMITATIONS AND FUTURE WORK

Although the proposed model provides a strong basis for intrusion detection, some limitations remain:

Increased computational complexity due to the hybrid framework

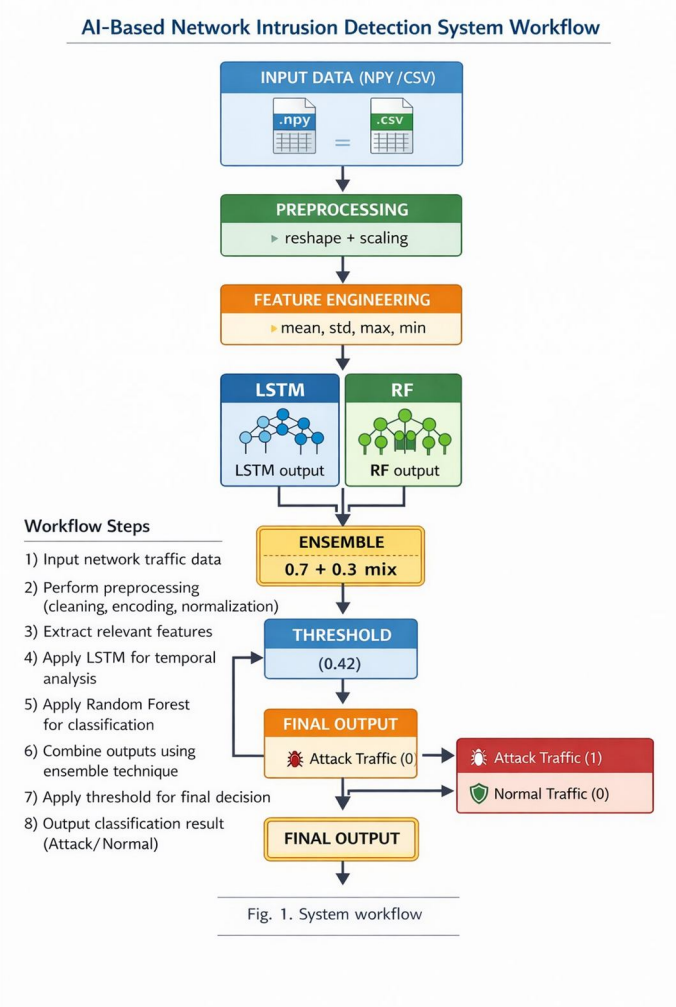


Fig. 4. Workflow of the proposed hybrid NIDS from input traffic to final attack or normal classification.

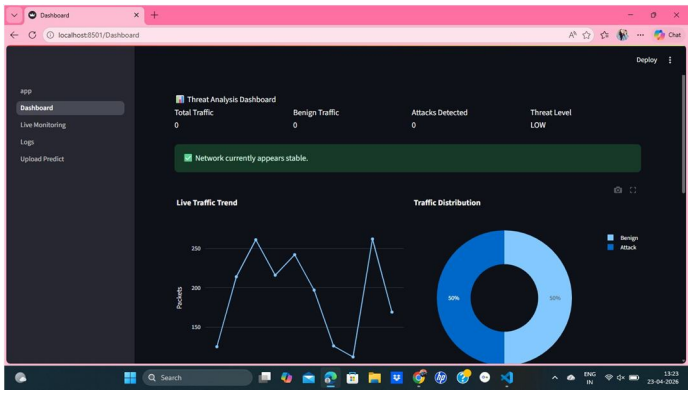


Fig. 5. Output screen 1 of the proposed AI-based intrusion detection system.

- Dependence on dataset quality and representativeness
- Limited validation in live real-time environments
- Difficulty in detecting rare classes without further bal-ancing techniques

Future work may focus on:

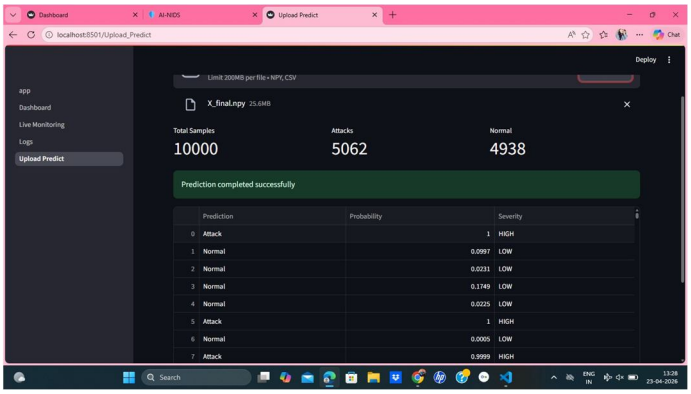


Fig. 6. Output screen 2 showing intrusion detection results and system behavior.

- Evaluation on modern datasets such as CICIDS2017 and UNSW-NB15
- Hyperparameter optimization
- Integration with real-time traffic capture systems
- Use of Explainable AI (XAI)
- Exploration of attention-based and transformer-based models

XI. CONCLUSION

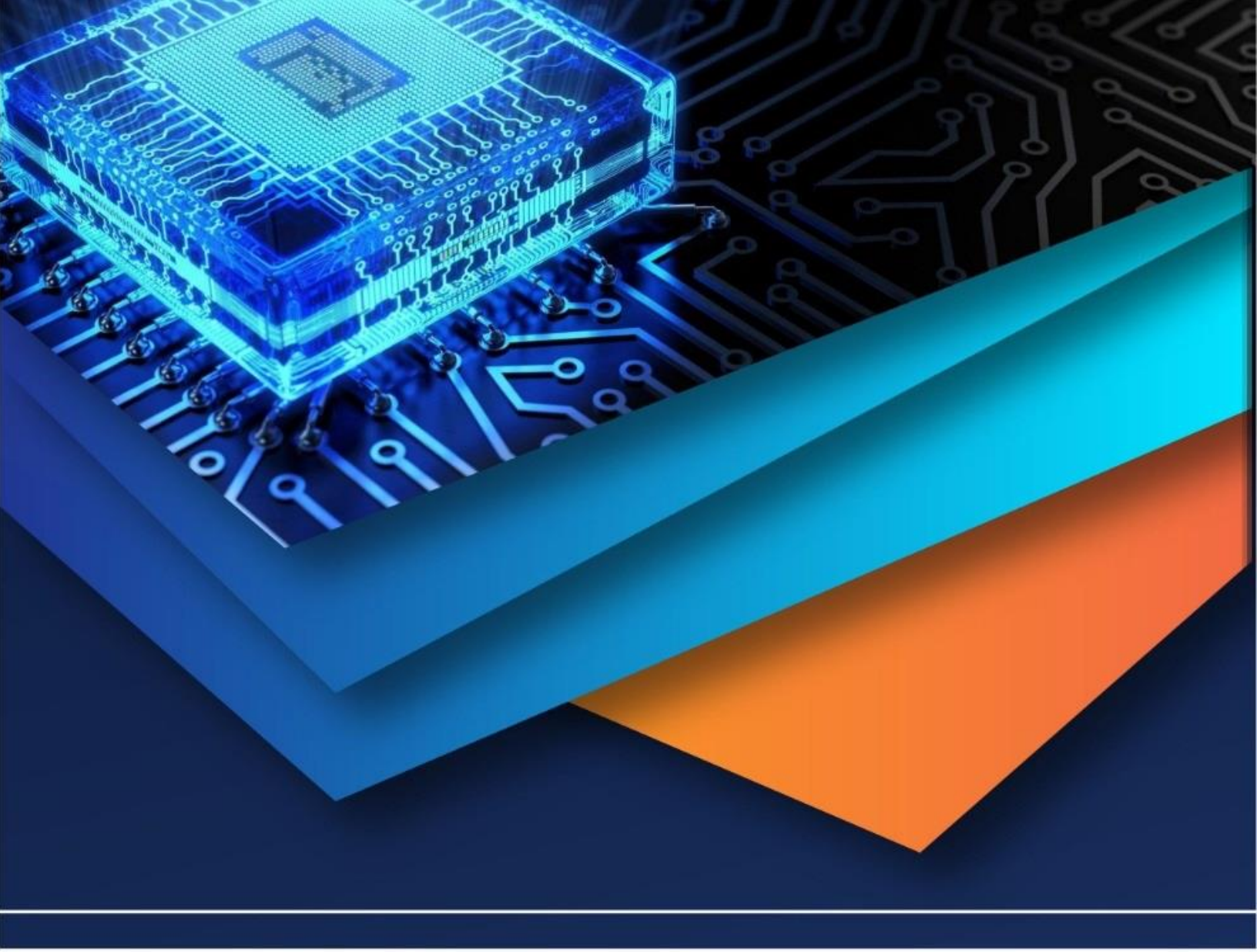
The above-discussed paper has proposed a novel hybrid AI-based Network Intrusion Detection System that utilizes the Long Short-Term Memory technique and the Random Forest for the network traffic classification problem. The discussed technique combines two different approaches for the improvement of intrusion detection.

The offered system solves most drawbacks that are faced when using other conventional techniques and allows us to consider network traffic as a time series data. It gives a deeper understanding of the importance of using both deep and machine learning techniques for cybersecurity.

REFERENCES

[1] M. Tavallaee, E. Bagheri, W. Lu, and A. A. Ghorbani, “An intrusion detection dataset and intrusion traffic characterization,” in Proc. Int. Conf. Information Systems Security and Privacy, 2009, pp. 108–116.
[2] KDD Cup 1999 Data, UCI KDD Repository. [Online]. Available: http://kdd.ics.uci.edu/databases/kddcup99/kddcup99.html
[3] L. Breiman, “Random forests,” Machine Learning, vol. 45, no. 1, pp. 5–32, 2001.

[4] S. Mukkamala, G. Janoski, and A. Sung, "Intrusion detection using neural networks and support vector machines," in Proc. IEEE Int. Joint Conf. Neural Networks, 2002, pp. 1702–1707.
[5] S. Hochreiter and J. Schmidhuber, "Long short-term memory," Neural Computation, vol. 9, no. 8, pp. 1735–1780, 1997.
[6] C. Yin, Y. Zhu, J. Fei, and X. He, "A deep learning approach for intrusion detection using recurrent neural networks," IEEE Access, vol. 5, pp. 21954–21961, 2017.
[7] G. Kim, S. Lee, and S. Kim, "A novel hybrid intrusion detection method integrating anomaly detection with misuse detection," Expert Systems with Applications, vol. 41, no. 4, pp. 1690–1700, 2014.
[8] A. Javaid, Q. Niyaz, W. Sun, and M. Alam, "A deep learning approach for network intrusion detection system," in Proc. ACM Int. Conf. Bioinformatics and Computational Biology, 2016, pp. 21–26.
[9] I. Sharafaldin, A. H. Lashkari, and A. A. Ghorbani, "Toward generating a new intrusion detection dataset and intrusion traffic characterization," in Proc. Int. Conf. Information Systems Security and Privacy, 2018, pp. 108–116.
[10] M. Ring, D. Wunderlich, D. Grudl, D. Landes, and A. Hotho, "A survey of network-based intrusion detection data sets," Computers & Security, vol. 86, pp. 147–167, 2019.
[11] H. Hindy et al., "A taxonomy of network threats and the effect of current datasets on intrusion detection systems," *IEEE Access*, vol. 8, pp. 104650–104675, 2020.



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY