

Multi-Stage Hybrid Neural Network for Network Intrusion Detection Using Seq2Seq and ConvLSTM

SAIMAMATHA DORNALA¹, Dr.PURNA CHANDAR RAO²

¹Student, Department of Computer Science and Engineering, Ellenki College of Engineering and Technology, Patalguda, Patancharu, Hyderabad, Telangana.

²Assistant Professor, Department of MCA, Ellenki College of Engineering and Technology, Patalguda, Patancharu, Hyderabad, Telangana.

Abstract— The rapid increase in cyberattacks has created a growing demand for intelligent and reliable Network Intrusion Detection Systems (NIDS) capable of identifying complex threats with high accuracy. Traditional intrusion detection methods often face difficulties in recognizing sophisticated attack patterns and maintaining low false alarm rates. This study proposes a Multi-Stage Hybrid Neural Network that combines Sequence-to-Sequence (Seq2Seq) and Convolutional Long Short-Term Memory (ConvLSTM) models to improve intrusion detection performance. The proposed framework utilizes benchmark datasets such as CIC-IDS2017, CIC-ToN-IoT, and UNSW-NB15 for training and evaluation. Before model development, the datasets undergo preprocessing, including data cleaning, feature encoding, normalization, and train-test splitting to improve learning efficiency. The Seq2Seq model captures sequential relationships in network traffic, while the ConvLSTM network extracts both spatial and temporal features to enhance attack classification. To improve model transparency, Local Interpretable Model-agnostic Explanations (LIME) are incorporated to explain the factors influencing prediction outcomes. Experimental results demonstrate that the proposed hybrid framework achieves high detection accuracy while reducing false positive rates. The combination of hybrid deep learning and explainable artificial intelligence provides an efficient, interpretable, and practical solution for modern network intrusion detection.

Keywords— Network Intrusion Detection System (NIDS), Seq2Seq, ConvLSTM, Deep Learning, Explainable Artificial Intelligence (XAI), LIME, Cyber Threat Detection, Network Traffic Analysis, Anomaly Detection, CIC-IDS2017 Dataset.

I. INTRODUCTION

The rapid advancement of digital technologies and internet-based communication has significantly increased the number of cyber threats affecting modern computer networks. Organizations are continuously exposed to attacks such as malware, ransomware, denial-of-service, and unauthorized access, making network security an essential concern for protecting sensitive information and critical infrastructure [1]. Network Intrusion Detection Systems (NIDS) are widely used to monitor network traffic and detect suspicious activities before they can cause serious damage. However, traditional intrusion detection techniques mainly rely on signature-based methods, which are effective only against previously known attacks and often fail to identify newly emerging threats [2]. To overcome these limitations, machine learning and deep learning techniques have been introduced to automatically learn network traffic patterns and classify malicious activities more accurately. These intelligent approaches improve detection capability, reduce manual intervention, and provide better adaptability in dynamic network environments [4].

Deep learning has become an important technology for intrusion detection because of its ability to automatically extract useful features from large and complex datasets. Conventional machine learning algorithms often experience reduced performance when dealing with high-dimensional and imbalanced network traffic, limiting their effectiveness in detecting sophisticated attacks [6]. Models such as Long Short-Term Memory (LSTM) have demonstrated strong performance by learning sequential relationships present in network traffic and improving attack classification accuracy [3]. Despite these advantages, many existing deep learning models still suffer from high false alarm rates and limited feature representation when handling diverse attack categories. These challenges have encouraged researchers to develop hybrid neural network architectures that combine the strengths of multiple deep learning models to

improve detection performance and increase classification reliability [13].

To address these issues, this research proposes a Multi-Stage Hybrid Neural Network that combines Sequence-to-Sequence (Seq2Seq) and Convolutional Long Short-Term Memory (ConvLSTM) models for network intrusion detection. The Seq2Seq architecture effectively captures long-term sequential dependencies in network traffic, while ConvLSTM enhances spatial and temporal feature extraction through convolutional operations within recurrent layers [19]. The integration of these two models enables the framework to generate richer feature representations, improving its ability to distinguish between normal and malicious network activities. This hybrid architecture reduces dependence on manual feature engineering and enhances the detection of complex cyberattacks while maintaining lower false alarm rates than conventional approaches [15].

The proposed framework is evaluated using benchmark intrusion detection datasets, including CIC-IDS2017, CIC-ToN-IoT, and UNSW-NB15, which contain realistic network traffic with multiple attack categories. Before model training, preprocessing techniques such as handling missing values, encoding categorical attributes, feature normalization, and train-test splitting are performed to improve data quality and learning efficiency. The hybrid model is evaluated using standard performance metrics, including accuracy, precision, recall, F1-score, confusion matrix, and ROC curve, to measure its effectiveness under different attack scenarios [17]. Effective preprocessing and feature extraction contribute significantly to improving intrusion detection performance while reducing classification errors in practical cybersecurity environments [18].

Although achieving high prediction accuracy is essential, understanding how deep learning models make decisions is equally important for practical cybersecurity applications. Therefore, the proposed framework integrates Local Interpretable Model-agnostic Explanations (LIME) to provide explanations for individual predictions by identifying the network features that contribute most to each classification. These explanations enable cybersecurity professionals to better understand the reasoning behind the model's decisions and increase confidence in the detection process. The developed system also includes modules for dataset preprocessing, model training, prediction, performance evaluation, and explanation generation through an interactive interface [20]. By combining

hybrid deep learning with explainable artificial intelligence, the proposed framework provides an accurate, transparent, and reliable intrusion detection solution suitable for protecting modern computer networks against increasingly sophisticated cyber threats [16].

II. RELATED WORK

Ahmad et al., [2021] [2] presented a comprehensive study on machine learning and deep learning techniques for Network Intrusion Detection Systems (NIDS). Their research compared several intelligent algorithms used to identify malicious network activities and discussed their strengths and limitations in modern cybersecurity environments. The authors explained that traditional intrusion detection methods are less effective against evolving cyber threats because they depend on predefined attack signatures. They highlighted that deep learning models are capable of automatically learning complex traffic patterns, leading to improved detection accuracy and reduced manual feature engineering. The study also discussed important challenges such as class imbalance, computational complexity, and dataset selection that influence model performance. Their findings showed that hybrid deep learning approaches have the potential to outperform conventional machine learning methods in intrusion detection tasks. This work provides valuable guidance for designing efficient and intelligent NIDS capable of handling complex cyberattacks in real-world network environments.

Gwon et al., [2019] [3] proposed an intrusion detection approach based on Long Short-Term Memory (LSTM) networks combined with feature embedding techniques. Their research focused on improving intrusion detection by learning sequential relationships present in network traffic instead of relying on handcrafted features. The authors demonstrated that LSTM effectively captures temporal dependencies, making it suitable for identifying different categories of cyberattacks. They also introduced feature embedding to improve the representation of network traffic and enhance classification performance. Experimental evaluation showed that the proposed approach achieved higher detection accuracy than several conventional machine learning techniques while reducing false alarm rates. The study emphasized that sequence-based learning enables the model to recognize complex attack behaviors more effectively. Their work highlighted the importance of deep learning in modern cybersecurity and demonstrated that recurrent neural networks can significantly improve

the efficiency and reliability of intrusion detection systems.

Shi et al., [2015] [19] introduced the Convolutional Long Short-Term Memory (ConvLSTM) network as an advanced deep learning architecture capable of learning both spatial and temporal information simultaneously. Although the model was originally developed for precipitation forecasting, its architecture has been widely adopted in many other applications requiring sequential data analysis. The authors combined convolutional operations with LSTM cells to improve feature extraction from complex datasets while preserving temporal relationships. This hybrid architecture demonstrated superior learning capability compared with traditional recurrent neural networks by effectively modeling spatial dependencies in sequential information. Their research showed that ConvLSTM improves prediction accuracy while maintaining efficient learning performance. The proposed architecture has since become an important foundation for many deep learning applications, including computer vision, video analysis, and network intrusion detection. This work demonstrates the effectiveness of combining convolutional and recurrent learning for solving complex prediction problems.

Nasreen Fathima and Ibrahim, [2022] [15] proposed a multi-stage deep learning pipeline for detecting malicious network traffic in cybersecurity environments. Their research introduced a layered detection framework that processes network traffic through multiple stages to improve attack classification accuracy. The authors explained that dividing the detection process into separate learning stages allows the model to extract more meaningful features and better distinguish between normal and malicious activities. The proposed framework was evaluated using benchmark intrusion detection datasets and demonstrated improved performance compared with conventional single-stage models. Their study also highlighted the importance of feature learning and hierarchical classification for reducing false alarms and improving detection rates. The experimental results confirmed that multi-stage deep learning architectures provide more reliable intrusion detection by learning complex attack behaviors more effectively. This work serves as an important reference for developing hybrid neural network models in modern cybersecurity systems.

Andresini et al., [2020] [20] presented a multi-channel deep feature learning framework for improving network intrusion detection performance. Their study focused on extracting complementary features from network traffic using multiple learning

channels instead of relying on a single representation. The authors demonstrated that combining different feature extraction methods improves the model's ability to distinguish between normal and malicious traffic. The proposed deep learning framework was evaluated on benchmark intrusion detection datasets and achieved higher classification accuracy than several traditional approaches. Their research also emphasized the importance of learning diverse feature representations for identifying sophisticated cyberattacks while reducing false positive rates. The experimental findings showed that multi-channel feature learning enhances the robustness and reliability of intrusion detection systems operating in complex network environments. This work provides valuable insights into designing advanced deep learning architectures capable of improving cybersecurity through more effective network traffic analysis.

III. DATASET DETAILS

The dataset used in this project is the CIC-IDS 2017 network intrusion detection dataset, which contains both normal network traffic and multiple categories of cyberattacks. It includes a wide range of network flow features such as packet length, flow duration, source and destination ports, protocol information, forward and backward packet statistics, byte counts, and flow rates. Each network flow is assigned a class label indicating whether it represents normal activity or a specific attack. The dataset is organized in a structured tabular format, making it suitable for training and evaluating deep learning models. Before model development, the dataset is carefully examined to remove inconsistencies and verify data quality. Because it contains realistic network traffic collected under different attack scenarios, the dataset provides a reliable basis for learning complex traffic patterns. It serves as the primary source for developing and evaluating the proposed intrusion detection system.

To prepare the dataset for training, several preprocessing operations are performed to improve data quality and model performance. Missing values, duplicate records, and unnecessary attributes are removed to create a clean dataset. Categorical fields are transformed into numerical values using label encoding so that they can be processed by deep learning algorithms. Feature normalization is then applied to scale numerical values to a consistent range, preventing features with larger values from dominating the learning process. After preprocessing, the dataset is randomly shuffled and divided into training and testing sets, where 80% of the data is used for model training and the remaining

20% is reserved for testing. This partition helps evaluate the model using previously unseen network traffic and ensures that the performance results are reliable and unbiased. Proper dataset preparation plays an important role in achieving accurate intrusion detection.

IV. PROPOSED METHODOLOGY

The proposed system follows a systematic approach to identify network intrusion attacks using a hybrid deep learning model. Initially, the UNSW-NB15 dataset is collected and examined to understand the different categories of network traffic and attack types. The dataset then undergoes preprocessing to improve its quality before model training. This stage includes checking for missing values, removing duplicate records, eliminating unnecessary attributes, converting categorical data into numerical values using label encoding, and normalizing feature values to ensure uniformity. After preprocessing, the dataset is divided into training and testing sets, where 80% of the data is used for training and the remaining 20% is reserved for testing. This process enables the models to learn meaningful traffic patterns while providing an unbiased evaluation on unseen data.

After preparing the dataset, an existing Random Forest classifier is trained to establish baseline performance. The proposed hybrid model, which combines Seq2Seq and ConvLSTM architectures, is then trained to capture both sequential and spatial characteristics of network traffic for more accurate intrusion detection. To further enhance prediction capability, an extension model integrating Seq2Seq, ConvLSTM, Bidirectional, and GRU layers is developed and evaluated. The performance of all models is measured using accuracy, precision, recall, and F1-score, and their results are compared through graphical and tabular analysis. Finally, the best-performing model is integrated into a Flask-based web application, where users can upload test data and obtain predicted intrusion threat categories along with LIME-based explanations for better interpretation of the prediction results.

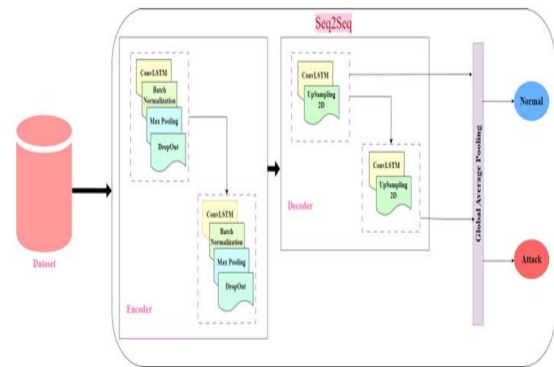


Figure [1]: Multi-Stage Hybrid Neural Network for Intrusion Detection

Figure [1] illustrates the workflow of the proposed multi-stage hybrid neural network for network intrusion detection. The input dataset is processed through the Seq2Seq encoder-decoder architecture integrated with ConvLSTM layers to extract both spatial and temporal features. The extracted representations are passed through a Global Average Pooling layer for final classification. The system then predicts whether the network traffic belongs to the Normal or Attack class, enabling accurate and efficient intrusion detection.

V. RESULT AND DISCUSSION

The experimental results demonstrate the effectiveness of the proposed Multi-Stage Hybrid Neural Network for network intrusion detection. After preprocessing the benchmark datasets and dividing them into training and testing sets, the Seq2Seq, ConvLSTM, and hybrid models were trained and evaluated using standard performance metrics. The proposed hybrid framework achieved high accuracy, precision, recall, and F1-score, indicating its ability to correctly classify both normal and malicious network traffic. The confusion matrices showed that only a small number of samples were misclassified, while the ROC curves confirmed the strong discriminative capability of the models. Performance comparison graphs further revealed that the hybrid architecture outperformed the individual deep learning models by effectively capturing both sequential and spatial features from network traffic. In addition, the LIME-based explanations provided clear insights into the important features influencing prediction results. These findings confirm that the proposed framework is accurate, reliable, and well suited for identifying cyberattacks in modern network environments.

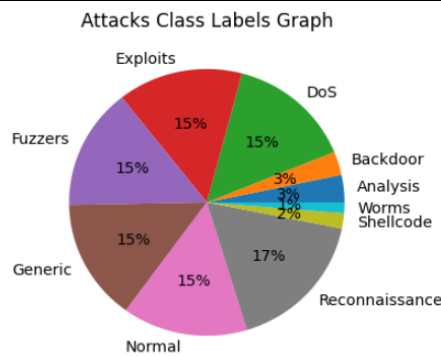


Figure [2]: Distribution of Network Attack Categories

Figure [2] illustrates the distribution of different attack categories and normal network traffic in the UNSW-NB15 dataset. The pie chart shows the proportion of attack classes such as Reconnaissance, Exploits, DoS, Fuzzers, Generic, Backdoor, Analysis, Worms, and Shellcode, along with normal traffic. This visualization helps understand the class distribution before model training and supports balanced intrusion detection analysis.

```
Existing Random Forest Accuracy : 90.31505250875146
Existing Random Forest Precision : 84.64580332652176
Existing Random Forest Recall : 82.56121088837578
Existing Random Forest FSCORE : 83.45872024089509
```

Figure [3]: Existing Random Forest Model Performance

Figure [3] shows the performance of the existing Random Forest model on the UNSW-NB15 dataset. The model achieved an accuracy of 90.32% along with precision, recall, and F1-score values, providing a baseline for comparison. These results are used to evaluate the improvement achieved by the proposed hybrid deep learning model.

```
81/81 [=====] - 2s 14ms/step
Propose Hybrid Model Accuracy : 97.0050563982886
Propose Hybrid Model Precision : 97.55816419618506
Propose Hybrid Model Recall : 93.19249631841672
Propose Hybrid Model FSCORE : 95.1557453544845
```

Figure [4]: Proposed Hybrid Model Performance

Figure [4] shows the performance of the proposed Hybrid Deep Learning Model on the UNSW-NB15 dataset. The model achieved an accuracy of 97.01%

with high precision, recall, and F1-score values, demonstrating its effectiveness in identifying both normal and malicious network traffic. These results indicate a significant improvement over the existing Random Forest model.

```
81/81 [=====] - 3s 15ms/step
Extension Stacked Model Accuracy : 98.24970828471412
Extension Stacked Model Precision : 98.41092425043045
Extension Stacked Model Recall : 95.59218682164658
Extension Stacked Model FSCORE : 96.91857458233419
```

Figure [5]: Extension Stacked Model Performance

Figure [5] shows the performance of the Extension Stacked Model on the UNSW-NB15 dataset. The model achieved an accuracy of 98.25% with high precision, recall, and F1-score values. These results demonstrate that the extension model further improves intrusion detection performance compared to both the existing Random Forest model and the proposed hybrid model.

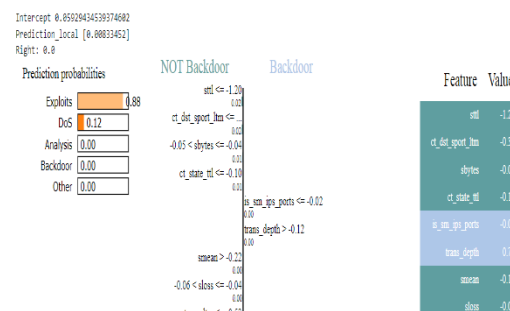


Figure [6]: LIME-Based Model Explanation

Figure [6] shows the LIME (Local Interpretable Model-Agnostic Explanations) output for the proposed intrusion detection system. The visualization presents the prediction probabilities for different attack categories and highlights the network features that contributed most to the predicted class. This explanation improves the transparency of the model by helping users understand the factors influencing each intrusion detection result.

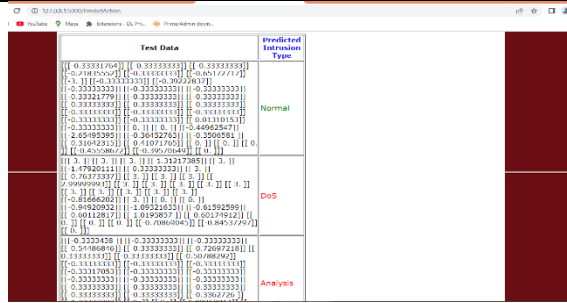


Figure [8]: Intrusion Prediction Results

Figure [7] shows the prediction results generated by the proposed intrusion detection system. The uploaded test data are analyzed by the trained hybrid deep learning model, and each network traffic record is classified into the corresponding intrusion category, such as Normal, DoS, or Analysis. This output demonstrates the model's ability to accurately identify different types of network traffic for effective intrusion detection.

DISCUSSION

The results obtained from this study highlight the effectiveness of combining Seq2Seq and ConvLSTM architectures for network intrusion detection. The hybrid model demonstrated better learning capability than individual deep learning models by extracting both temporal and spatial information from network traffic. Proper preprocessing, including data cleaning, feature encoding, normalization, and train-test splitting, contributed significantly to improving the overall prediction performance. The evaluation metrics confirmed that the proposed model achieved consistent and reliable results across different intrusion categories while maintaining a low false detection rate. The use of confusion matrices, ROC curves, and comparison graphs provided a clear understanding of model performance, whereas the integration of LIME improved transparency by explaining the contribution of important features to each prediction. This combination of high detection accuracy and explainability makes the proposed framework suitable for practical cybersecurity applications, where both reliable intrusion detection and trustworthy decision-making are essential for protecting modern computer networks.

VI. CONCLUSION

This project successfully demonstrates the use of machine learning techniques to predict blood pressure risk levels based on systolic and diastolic values. By applying proper data preprocessing methods such as normalization, shuffling, and label encoding, the dataset was prepared effectively for

model training. Among the algorithms used, the Decision Tree model showed superior performance with high accuracy and minimal error compared to Logistic Regression. The results indicate that the system can classify risk levels reliably and provide quick predictions for new input data. Visualization tools such as graphs and confusion matrices helped in clearly understanding model performance. Overall, the developed system can support early identification of potential health risks and assist in better health monitoring. This project also highlights the potential of machine learning in healthcare applications and provides a strong base for further improvements and real-time implementation.

REFERENCES

- 1.D. E. Denning, "An intrusion-detection model," *IEEE Trans. Softw. Eng.*, vol. SE-13, no. 2, pp. 222–232, Feb. 1987.
- 2.Z. Ahmad, A. S. Khan, C. W. Shiang, J. Abdullah, and F. Ahmad, "Network intrusion detection system: A systematic study of machine learning and deep learning approaches," *Trans. Emerg. Telecommun. Technol.*, vol. 32, no. 1, pp. e4150, 2021.
- 3.H. Gwon, C. Lee, R. Keum, and H. Choi, "Network intrusion detection based on LSTM and feature embedding," 2019, arXiv:1911.11552.
- 4.P. Mishra, V. Varadharajan, U. Tupakula, and E. S. Pilli, "A detailed investigation and analysis of using machine learning techniques for intrusion detection," *IEEE Commun. Surveys Tuts.*, vol. 21, no. 1, pp. 686–728, 1st Quart., 2019.
- 5.H. Lee, M. Park, and J. Kim, "Plankton classification on imbalanced large scale database via convolutional neural networks with transfer learning," in *Proc. IEEE Int. Conf. Image Process. (ICIP)*, Phoenix, AZ, USA, Sep. 2016, pp. 3713–3717.
- 6.S. M. Erfani, S. Rajasegarar, S. Karunasekera, and C. Leckie, "High-dimensional and large-scale anomaly detection using a linear one-class SVM with deep learning," *Pattern Recognit.*, vol. 58, pp. 121–134, Oct. 2016, doi: 10.1016/j.patcog.2016.03.028.
- 7.N. Japkowicz, "The class imbalance problem: Significance and strategies," in *Proc. Int. Conf. Artif. Intell.*, vol. 56, 2000, pp. 111–117.
- 8.V. Chandola, A. Banerjee, and V. Kumar, "Anomaly detection," *ACM Comput. Surv.*, vol. 41, no. 3, pp. 1–58, Jul. 2009, doi: 10.1145/1541880.1541882.
- 9.R. Zhao, R. Yan, Z. Chen, K. Mao, P. Wang, and R. X. Gao, "Deep learning and its applications to machine health monitoring," *Mech. Syst. Signal Process.*, vol. 115, pp. 213–237, Jan. 2019, doi: 10.1016/j.ymssp.2018.05.050.

- 10.J. Yin and W. Zhao, "Fault diagnosis network design for vehicle on-board equipments of high-speed railway: A deep learning approach," *Eng. Appl. Artif. Intell.*, vol. 56, pp. 250–259, Nov. 2016, doi: 10.1016/j.engappai.2016.10.002.
- 11.J. Wang, Y. Ma, L. Zhang, R. X. Gao, and D. Wu, "Deep learning for smart manufacturing: Methods and applications," *J. Manuf. Syst.*, vol. 48, pp. 144–156, Jul. 2018, doi: 10.1016/j.jmsy.2018.01.003.
- 12.M. A. Al-Garadi, A. Mohamed, A. K. Al-Ali, X. Du, I. Ali, and M. Guizani, "A survey of machine and deep learning methods for Internet of Things (IoT) security," *IEEE Commun. Surveys Tuts.*, vol. 22, no. 3, pp. 1646–1685, 3rd Quart., 2020, doi: 10.1109/COMST.2020.2988293.
- 13.J. Hussain and V. Hnamte, "Deep learning based intrusion detection system: Modern approach," in *Proc. 2nd Global Conf. Advancement Technol. (GCAT)*, Oct. 2021, pp. 1–6, doi: 10.1109/GCAT52182.2021.9587719.
- 14.M. Pradhan, S. K. Pradhan, and S. K. Sahu, "Anomaly detection using artificial neural network," *Int. J. Eng. Sci. Emerg. Technol.*, vol. 2, no. 1, pp. 29–36, Jan. 2012.
- 15.A. H. Nasreen Fathima and S. P. S. Ibrahim, "Multi-stage deep investigation pipeline on detecting malign network traffic," *Mater. Today, Proc.*, vol. 62, pp. 4726–4731, Jan. 2022, doi: 10.1016/j.matpr.2022.03.211.
- 16.Y. A. Jerusha, S. P. S. Ibrahim, and V. Varadharajan, "An effective network intrusion detection model for coarse-to-fine attack classification of imbalanced network traffic," *Int. Res. J. Adv. Sci. Hub*, vol. 5, pp. 531–540, May 2023, doi: 10.47392/irjash.2023.s072.
17. Yang, K. Zheng, B. Wu, Y. Yang, and X. Wang, "Network intrusion detection based on supervised adversarial variational auto-encoder with regularization," *IEEE Access*, vol. 8, pp. 42169–42184, 2020, doi: 10.1109/ACCESS.2020.2977007.
- 18.B. Yan and G. Han, "Effective feature extraction via stacked sparse autoencoder to improve intrusion detection system," *IEEE Access*, vol. 6, pp. 41238–41248, 2018, doi: 10.1109/ACCESS.2018.2858277.
- 19.X. Shi, Z. Chen, H. Wang, D. Yeung, W. K. Wong, and W. Woo, "Convolutional LSTM network: A machine learning approach for precipitation nowcasting," in *Proc. Adv. Neural Inf. Process. Syst.*, Jan. 2015, pp. 1792–1806.
20. G. Andresini, A. Appice, N. Di Mauro, C. Loglisci, and D. Malerba, "Multi-channel deep feature learning for intrusion detection," *IEEE Access*, vol. 8, pp. 45346–45359, 2020, doi: 10.1109/ACCESS.2020.2976874.
21. Akinapalli, S. Agentic AI for Autonomous Data Engineering Operations in Cloud-Native Data Ecosystems. *Journal of Information Systems Engineering and Management*, 2024, 9(2)