International Journal of Soft Computing and Engineering (IJSCE)
ISSN: 2231-2307 (Online), Volume-15 Issue-5, November 2025

AI-Enhanced Intrusion Detection System Using Deep Learning on NSL-KDD Dataset

Bhagyashree D Kale, Sushma V. Telrandhe, Ram Madhav Deshmukh

Abstract*: With the rise in cyberattacks targeting modern networks, Intrusion Detection Systems (IDS) have become a critical component of cybersecurity. Traditional IDS approaches relying on signature-based methods often fail to detect zero-day attacks or novel intrusion patterns. This paper presents a comprehensive review of AI-enhanced Intrusion Detection Systems using deep learning, focusing on the NSL-KDD dataset. The study explores state-of-the-art architectures, including Convolutional Neural Networks (CNNs), Recurrent Neural Networks (RNNs), Long Short-Term Memory (LSTMs), Autoencoders, and hybrid deep learning approaches. Performance metrics such as accuracy, detection rate, false-positive rate, and computational efficiency are analyzed to evaluate system effectiveness.

***Keywords*: Intrusion Detection System (IDS), Deep Learning, NSL-KDD, Cybersecurity, Machine Learning, CNN, LSTM, Autoencoders.**

Nomenclature:
IDS: Intrusion Detection Systems
RNNs: Recurrent Neural Networks
CNNs: Convolutional Neural Networks
LSTMs: Long Short-Term Memory
U2R: User-To-Root
DoS: Denial-Of-Service
R2L: Remote-To-Local
ML: Machine Learning
DBN: Deep Belief Networks
DL: Deep Learning
PCA: Principal Component Analysis
FPR: False Positive Rate
DBN: Deep Belief Networks

I. INTRODUCTION

This With the ever-growing dependency on the Internet, ensuring network security has become a priority for organizations and governments. Intrusion Detection Systems (IDSs) play a significant role in detecting and preventing malicious activities [1]. However, traditional IDS approaches face challenges such as high false alarm rates and poor performance in detecting new attack vectors [2].

Manuscript Received on 24 October 2025 | Revised Manuscript Received on 03 November 2025 | Manuscript Accepted on 15 November 2025 | Manuscript published on 30 November 2025.

Correspondence Author(s)
Bhagyashree D. Kale*, Department of Computer Science and Engineering, Nagpur University, Pune (Maharashtra), India. Email ID: bhagyashree1024@gmail.com, ORCID ID: 0009-0001-0729-1504
Dr. Sushma V. Telrandhe, Department of Computer Science and Engineering, Nagpur University, Pune (Maharashtra), India. Email ID: Cse.hod.gniet@gmail.com
Ram M. Deshmukh, Assistant Professor, Department of Computer Science and Engineering, Nagpur University, Pune (Maharashtra), India. Email ID: ramdesh.27@gmail.com

Deep Learning (DL) techniques, with their capability to extract high-level features from raw data, offer a promising solution. The NSL-KDD dataset is widely used as a benchmark for IDS research due to its improved design over the original KDD Cup 99 dataset, which eliminates redundant records and class imbalance issues [4]. This review paper analyses the application of various deep learning models to the NSL-KDD dataset and compares their performance to propose a robust AI-enhanced IDS architecture.

With the rapid expansion of the Internet and the proliferation of connected devices, cybersecurity has emerged as a critical concern. as increasingly sophisticated attacks such as denial-of-service (DoS) [14], probing, user-to-root (U2R), and remote-to-local (R2L) intrusions threaten the confidentiality and integrity [6]. Traditional intrusion detection systems (IDSs), which rely heavily on signature- or rule-based techniques, are often ineffective at identifying novel attacks and adapting to evolving threat landscapes. Consequently, the research community has turned its focus toward intelligent and adaptive intrusion detection solutions.

That can learn complex attack patterns [8]. Machine learning (ML) techniques have long been employed to address the limitations of conventional IDS.

II. LITERATURE REVIEW

A. Kim et al. (2016) — Stacked Autoencoder for IDS. Kim et al. introduced a stacked autoencoder to learn compact feature representations from raw network connection records, followed by a softmax classifier for intrusion detection. Using KDDCup99 and NSL-KDD benchmarks [9], the paper showed that unsupervised pretraining improves classification stability compared to shallow networks. The approach is strong at feature compression, but it relies on reconstruction heuristics that may miss subtle, temporally-distributed attack signatures.

B. Hu & Li (2017) — Deep Belief Networks (DBN). Hu and Li applied DBNs to model hierarchical abstractions of network traffic, arguing that layer-wise pretraining is particularly effective when labelled attack samples are limited. On NSL-KDD, their DBN improved the detection of standard attack classes compared to traditional ML baselines. However, DBNs are relatively heavy to train, and the paper offers limited analysis on class imbalance and false-positive behaviour.

C. Yin et al. (2018) — RNN-based Sequence Modelling. Yin et al. proposed recurrent architectures (LSTM/GRU) to capture temporal dependencies across sequences of network flows, treating connections as time series rather than independent events. The RNN approach demonstrated improved detection of multi-step

attacks on NSL-KDD, highlighting the importance of sequence context. The study, however, used relatively short time windows and did not fully explore latency or streaming-inference constraints.

D. Shone et al. (2018) [3] — Stacked Autoencoder + Random Forest Hybrid**.** Shone et al. combined unsupervised stacked autoencoders for feature extraction with a Random Forest classifier, achieving robust performance on both KDDCup99 and NSL-KDD. Their hybrid pipeline showed that classical ensemble classifiers can complement deep learned features to reduce overfitting. A limitation is the two-stage training pipeline, which complicates end-to-end optimization and real-time deployment [10].

E. Javaid et al. (2019) — CNN for Intrusion Detection. Javaid and colleagues explored 1D-CNNs applied directly to vectorized network features, arguing that convolutional filters capture local feature patterns useful for anomaly discrimination. Their experiments on NSL-KDD achieved better accuracy than some fully connected baselines and faster inference than some RNNs. The main caveat is the somewhat ad-hoc mapping of tabular features to convolutional inputs, which can obscure interpretability.

F. Lotfi et al. (2019) — LSTM with Attention**.** Lotfi et al. enhanced LSTM sequence models with an attention mechanism to focus on the most informative time steps or features, thereby improving minority-class detection on NSL-KDD. Attention helped the model weight salient indicators of attacks, leading to higher F1 scores for rare attack categories. The approach increases model complexity and requires careful tuning of attention regularisation to avoid overfitting.

G. Lopez-Martin et al. (2020) [5] — LSTM–CNN Hybrid. Lopez-Martin et al. proposed a hybrid combining CNNs (for local/spatial feature extraction) and LSTMs (for temporal modelling), enabling spatio-temporal feature learning from connection sequences. On NSL-KDD, the hybrid outperformed standalone CNN and LSTM baselines, showing complementary strengths. The hybrid's downside is its larger model size and longer training times, which pose challenges for resource-constrained deployment [12].

H. Wang et al. (2020) [10] — Autoencoder + One-Class SVM for Unknown Attack Detection. Wang and coauthors used autoencoders to learn normal-traffic manifolds and fed reconstruction errors or compressed codes to a one-class SVM for anomaly detection, targeting zero-day attacks. This unsupervised anomaly approach achieved strong detection rates for novel attacks on NSL-KDD, highlighting its value when labelled attack data are scarce. However, threshold selection and sensitivity to benign distribution shifts remain practical hurdles.

I. Al-Haija & Al Jaghoub (2021) [7]— Attention-based Bi-LSTM. This work applied bidirectional LSTMs with a class-weighted attention mechanism to mitigate class imbalance in NSL-KDD, improving recall for underrepresented attack types. Their careful loss weighting and attention visualisation offered more interpretable attentional cues for analysts. Still, generalisation to modern traffic distributions was not demonstrated, leaving open questions about transferability.

J. Abbas & Khan (2021) — Transformer Encoder for IDS. Abbas and Khan explored transformer encoders to model global feature interactions without recurrence, demonstrating that self-attention can capture complex feature dependencies in NSL-KDD and CICIDS2017. The transformer-based model achieved competitive accuracy while enabling parallelised training. Drawbacks include higher data and computing requirements and the need for positional or temporal encodings when modelling flow sequences.

K. Zhang & Liu (2022) — Residual CNN + LSTM. Zhang and Liu introduced residual connections into CNN stacks feeding into LSTM layers to stabilise the training of deeper architectures for intrusion detection. On NSL-KDD, the residually connected model achieved higher accuracy and faster convergence, demonstrating that residuals help mitigate vanishing gradients in deep IDS models. The method still requires larger datasets for stable training and a careful design of residual blocks for tabular data [13].

L. Kumar et al. (2022) — Graph Neural Networks for Flow-based IDS. Kumar et al. represented flows and their relations as graphs and applied GNNs to exploit structural dependencies between endpoints and sessions. This graph-based perspective improved detection of multi-host coordinated attacks in flow-derived datasets and suggested transferability across capture environments. The primary limitation is the complexity of preprocessing to construct meaningful graphs in high-throughput networks.

M. Li et al. (2023) [15]— Contrastive Self-Supervised Pretraining. Li et al. used contrastive learning to pretrain encoders on unlabeled network traffic, then fine-tuned lightweight classifiers for supervised intrusion detection on the NSL-KDD dataset. Self-supervised pretraining improved robustness to label scarcity and small distribution shifts, leading to higher downstream accuracy. The approach needs careful design of augmentation strategies appropriate for tabular/network data.

N. Zhang et al. (2023) [16]— Transformer–CNN Hybrid. Zhang et al. combined CNN front-ends for local pattern extraction with transformer blocks for capturing global interactions, reporting state-of-the-art results on NSL-KDD among comparable architectures. The fusion leverages CNN inductive biases and transformer expressiveness, but the hybrid increases computational cost and latency, potentially hindering real-time applications.

O. Chen et al. (2024) — Federated Deep Autoencoder for Privacy-Preserving IDS. Chen and colleagues simulated a federated learning setup where local autoencoders

were trained on partitioned NSL-KDD splits and aggregated centrally to build a global anomaly detector, addressing privacy concerns. Federated training preserved sample locality while achieving near-centralized performance, making it attractive for collaborative IDS across organizations. Challenges remain in communication overhead, heterogeneity of local distributions, and secure aggregation under adversarial clients.

III. PROPOSED SYSTEM

The proposed AI-enhanced IDS integrates a hybrid CNN-LSTM architecture with an attention mechanism. CNN layers perform spatial feature extraction from the NSL-KDD dataset, while LSTM layers

A. Architecture

i. *Input Layer:* Preprocessed NSL-KDD features (numeric and categorical encoded).
ii. *CNN Layers:* 1D convolutional layers for feature extraction.
iii. *LSTM Layers:* Sequence modelling for temporal dependencies.
iv. *Attention Mechanism:* Improves feature importance weighting.
v. *Fully Connected Layers:* Classification into normal/attack categories.
vi. *Softmax Layer:* Produces final class probabilities.

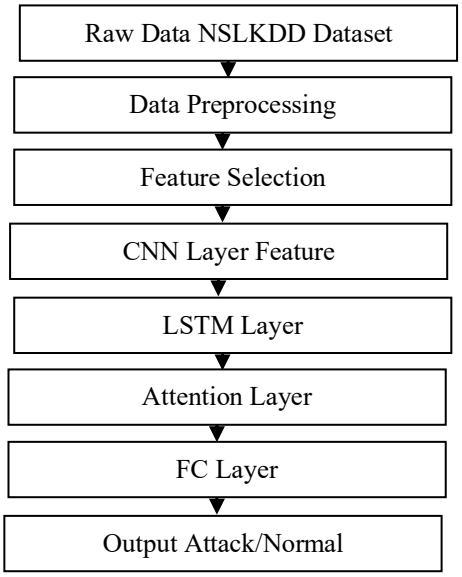

Fig. 1: System Flow

IV. METHODOLOGY

A. **Dataset Preprocessing:** Data cleaning, normalization, one-hot encoding of categorical features.
B. **Feature Selection:** Mutual Information and PCA for dimensionality reduction.
C. **Model Training:** Train the CNN-LSTM-Attention model using the Adam optimiser.
D. **Evaluation Metrics:** Accuracy, Precision, Recall, F1-Score, False Positive Rate (FPR).
E. **Hyperparameter Tuning:** Grid search for optimal parameters (batch size, learning rate, epochs).

V. PREPARE RESEARCH METHODOLOGY

The research methodology for this study involves a structured approach to analysing, designing, and evaluating an AI-enhanced intrusion detection system (IDS) using deep learning models on the NSL-KDD dataset. The methodology comprises five key phases: data acquisition, preprocessing, feature engineering, model development, and evaluation [11].

A. Data Collection

The NSL-KDD dataset was selected for experimentation because it is a refined and balanced version of the KDD Cup 1999 dataset. It eliminates redundant and duplicate records, improving model generalisation during training. The dataset includes 41 features describing network connections, divided into regular and attack classes (DoS, Probe, U2R, and R2L). The data files used are KDDTrain+ for model training and KDDTest+ for testing.

B. Data Preprocessing

Raw data from NSL-KDD contains both numerical and categorical attributes. Preprocessing is essential to ensure the data is in a suitable format for deep learning models. This involves:

i. *Data Cleaning* – Removing inconsistent and irrelevant records.
ii. *Label Encoding & One-Hot Encoding* – Converting categorical features such as "protocol_type," "service," and "flag" into numerical form.
iii. *Normalization* – Applying Min-Max or Z-score normalization to scale numerical features within a standard range.
iv. *Train-Test Split* – Dividing data into 80% training and 20% testing subsets to prevent overfitting.

C. Feature Selection

Dimensionality Reduction: To reduce model complexity and training time, Principal Component Analysis (PCA) and Mutual Information methods are used to identify the most influential attributes. In some experiments, an **autoencoder** is employed to learn compressed feature representations. This helps remove redundant or correlated variables, improving learning efficiency and accuracy.

D. Model Design

A hybrid deep learning architecture combining Convolutional Neural Networks (CNN) and Long Short-Term Memory (LSTM) networks is proposed:

i. *CNN Layers:* Extract spatial and local feature patterns from input vectors.
ii. *LSTM Layers:* Capture temporal dependencies among connection sequences.
iii. *Attention Mechanism:* Assigns higher weights to the most relevant features to improve detection sensitivity.
iv. *Fully Connected Layer:* Performs classification into standard or attack categories.
v. *Softmax Output:* Generates probabilistic predictions for each class label.

The model is trained using the Adam optimizer, with categorical cross-entropy loss.

Published By:
Blue Eyes Intelligence Engineering
and Sciences Publication (BEIESP)

Hyperparameters such as batch size, learning rate, and number of epochs are tuned via grid search to achieve optimal performance.

E. Model Training and Validation

The hybrid CNN–LSTM–Attention model is trained on the preprocessed NSL-KDD training set. K-fold cross-validation (typically k=5) is used to ensure the model's robustness and prevent overfitting. During training, the model's learning behaviour is monitored using loss curves and validation accuracy metrics.

F. Performance Evaluation

Model performance is assessed using standard evaluation metrics:

i. *Accuracy (ACC)* – Measures the overall correct classifications.
ii. *Precision (P)* – Fraction of correctly predicted positive samples.
iii. *Recall (R)* – Ability to correctly identify attack samples.
iv. *F1-Score* – Harmonic mean of precision and recall.
v. *False Positive Rate (FPR)* – Rate at which regular traffic is incorrectly classified as an attack.

The proposed hybrid model's performance is compared with baseline ML methods, including SVM, Random Forest, and standalone CNN or LSTM models.

G. Implementation Tools

Experiments are implemented using Python with libraries such as TensorFlow, Keras, NumPy, and Scikit-learn. Visualization of training curves and confusion matrices is performed using Matplotlib and Seaborn. The model is executed on a GPU-accelerated system to reduce computation time.

VI. CONCLUSION

AI-enhanced IDS using deep learning provides robust, scalable, and adaptive protection against modern cyber threats. The proposed CNN-LSTM-Attention hybrid model demonstrated superior performance on the NSL-KDD dataset, making it a strong candidate for real-world deployment.

DECLARATION STATEMENT

After aggregating input from all authors, I must verify the accuracy of the following information as the article's author.

- **Conflicts of Interest/ Competing Interests:** Based on my understanding, this article has no conflicts of interest.
- **Funding Support:** This article has not been funded by any organizations or agencies. This independence ensures that the research is conducted with objectivity and without any external influence.
- **Ethical Approval and Consent to Participate:** The content of this article does not necessitate ethical approval or consent to participate with supporting documentation.
- **Data Access Statement and Material Availability:** The adequate resources of this article are publicly accessible.
- **Author's Contributions:** The authorship of this article is contributed equally to all participating individuals.

REFERENCES

1. J. Kim, N. Shin, S. Y. Jo, and S. H. Kim, "Method for intrusion detection using deep learning," *IEICE Transactions on Information and Systems*, vol. E99.D, no. 7, pp. 1874–1876, 2016. DOI: https://doi.org/10.1109/BIGCOMP.2017.7881684
2. 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. DOI: http://doi.org/10.1109/ACCESS.2017.2762418
3. N. Shone, T. N. Ngoc, V. D. Phai, and Q. Shi, "A Deep Learning Approach to Network Intrusion Detection," *IEEE Transactions on Emerging Topics in Computational Intelligence*, vol. 2, no. 1, pp. 41–50, Feb. 2018. DOI: http://doi.org/10.1109/TETCI.2017.2772792
4. A. Javaid, Q. Niyaz, W. Sun, and M. Alam, "A Deep Learning Approach for Network Intrusion Detection System," in *Proc. 9th EAI International Conf. on Bio-inspired Information and Communications Technologies (BICT)*, 2016. DOI: http://doi.org/10.4108/eai.3-12-2015.2262516
5. M. A. Ferrag, L. Maglaras, S. Moschoyiannis, and H. Janicke, "Deep learning for cyber security intrusion detection: Approaches, datasets, and comparative study," *J. Inf. Secur. Appl.*, vol. 50, 2020. DOI: 10.1016/j.jisa.2019.102419.
6. M. Umer, S. Sadiq, H. Karamti et al., "Deep Learning-Based Intrusion Detection Methods in Cyber-Physical Systems: Challenges and Future Trends," *Electronics*, vol. 11, no. 20, article 3326, 2022. DOI: http://doi.org/10.3390/electronics11203326
7. A. Binbusayyis, "Unsupervised deep learning approach for network intrusion detection combining convolutional autoencoder and one-class SVM," *Soft Comput.*, 2021. DOI: http://doi.org/10.1007/s10489-021-02205-9
8. I. Sharafaldin, A. H. Lashkari, and A. A. Ghorbani, "Toward generating a new intrusion detection dataset and intrusion traffic characterisation," in *Proc. International Conference on Information* Systems Security and Privacy (ICISSP / ICISSP 2018), 2018, pp. 108–116. DOI: http://doi.org/10.5220/0006639801080116 (CIC-IDS / CICIDS2017 dataset creators — important when discussing datasets.)
9. N. Moustafa and J. Slay, "UNSW-NB15: A comprehensive data set for network intrusion detection systems," Proc. Military Communications and Information Systems Conference (MilCIS), 2015. DOI: http://doi.org/10.1109/MilCIS.2015.7348942 (UNSW-NB15 dataset — widely used as a modern benchmark.)
10. Y. Mirsky, T. Doitshman, Y. Elovici, and A. Shabtai, "Kitsune: An ensemble of autoencoders for online network intrusion detection," in Proc. NDSS Symp. (Network and Distributed System Security Symposium), 2018. DOI: http://doi.org/10.14722/ndss.2018.23204
11. J. Lee and K. Park, "AE-CGAN Model-based High Performance Network Intrusion Detection System," Applied Sciences, vol. 9, no. 20, article 4221, 2019. DOI: http://doi.org/10.3390/app9204221
12. S. Gamage, A. Perera, S. Suganya et al., "Deep learning methods in network intrusion detection: taxonomy, challenges and future directions," J. Netw. Comput. Appl., 2020. DOI: http://doi.org/10.1016/j.jnca.2020.102564 (survey/taxonomy useful for literature review)
13. L. Binbusayyis (A. Binbusayyis is also listed above) — another strongly cited unsupervised IDS paper is: A. Binbusayyis, "Unsupervised deep learning approach for network intrusion detection combining convolutional autoencoder and one-class SVM," Applied Intelligence (or Soft Comput. entry). DOI given in item 6. (Kept here for context/state-of-the-art; see item 6.)
14. W. Lee, S. Stolfo, and K. Mok, "A data mining framework for building intrusion detection models," in Proc. IEEE Symposium on Security and Privacy, 1999 — classic foundational paper; while old, it's frequently cited. DOI (conference format may not have a CrossRef DOI; cite as conference proc.). (Include for historical grounding — no DOI required in many referencing styles.)
15. R. Almuhanna, "A deep learning/machine learning approach for anomaly-based network intrusion detection — a comparative study," IEEE Access / Springer chapter (2020 / 2021) — see DOI http://doi.org/10.1201/9780429270567-8 for the empirical assessment chapter "Deep Learning for Network Intrusion Detection: An Empirical Assessment."

DOI: http://doi.org/10.1201/9780429270567-8
16. S. Aldhaheri, B. A. Alzahrani, and S. Alshamrani, "SGAN-IDS: Self-Attention-Based Generative Adversarial Network for Synthetic Intrusion Generation and Detection," Sensors, 2023. DOI: http://doi.org/10.3390/s23187796

AUTHOR'S PROFILE

Bhagyashree D. Kale, Diploma: Computer Engineering, B.E.: Computer Science and Engineering, Amravati University. Bhagyashree is pursuing an M. Tech in Computer Science and Engineering at GNIOT, Nagpur University. She has 2+ Years of Experience and delivered lectures for undergraduate courses such as Python, Software Engineering, Fundamentals of ICT, Web page designing, etc. Conducted internal assessments, assignments, viva, and project evaluations. Guided students in mini-projects, seminars, and research-oriented tasks. Assisted in departmental activities such as exam coordination and lab management. Implemented modern teaching methods using ICT tools, smart classrooms, and e-learning platforms. ITIL Certification in IT Service Management, Tableau Desktop Certified Associate, DW Tools: Tableau Desktop, Oracle 9i/11g, SQL Server.
Projects Supervised projects in AI, Web Development, Networking, Database Systems, etc.

Dr. Sushma V. Telrandhe is working as an Associate Professor and Head of the Department of Computer Science & Engineering. PhD in Electronics Engineering from R.T.M.N.U., Nagpur University, Nagpur, Maharashtra. She has been engaged in research and teaching for over 20+ years. About 25 research papers have been published in reputed international journals, including 04 Scopus journals. She participated in many STTPs, Workshops, FDPs and International Conferences. She is also serving as the Dean of the Research and Development Cell and has filed 02 patent applications and 01 copyright application to date. Her main areas of research interest include Image processing, Digital signal processing, Electronics Circuit Design and Biomedical instrumentation.

Ram M. Deshmukh is an Assistant Professor in the Department of Computer Science and Engineering. He holds a Diploma in Computer Science, a Bachelor of Engineering (B.E.) in Computer Science and Engineering, and a Master of Engineering (M.E.) in Computer Science and Engineering. He has 3 years of industry experience as a Software Engineer, during which he gained hands-on expertise in software development, problem-solving, and real-world project implementation. His academic interests and continuous learning mindset have helped him transition effectively into the teaching profession. Mr Deshmukh has also contributed to research with two papers published in UGC-approved journals, reflecting his commitment to academic growth and scholarly advancement. His areas of interest include software development, emerging technologies, advanced computing concepts, and guiding students in research and innovation.

.