iJRASET
International Journal For Research in Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 13 **Issue:** VI **Month of publication:** June 2025

DOI: https://doi.org/10.22214/ijraset.2025.72761

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 13 Issue VI June 2025- Available at www.ijraset.com

Cybersentinel AI: An Intelligent Cybersecurity Framework Using Artificial Intelligence

Ritik[1], Dr. Amandeep[2]
[1]*M.sc CS, (Department of Artificial Intelligence and Data Science), Guru Jambheshwar University of Science and Technology, Hisar, Haryana, India*
[2]*Assistant Professor, Department of Artificial Intelligence and Data Science, Guru Jambheshwar University of Science and Technology, Hisar, Haryana, India*

Abstract: ***"CyberSentinel AI: An Intelligent Cybersecurity Framework Using Artificial Intelligence," we proposed a scalable AI based solution for detecting and preventing cyberattacks which has been implemented and evaluated using machine learning methods and NSL-KDD dataset which is one of the most popular benchmark dataset in this domain.The introduced approach can be applied to detect malicious behavior in network traffic through preprocessing data, feature extraction, and probabilistic model training used for binary classification of normal and attack data. The pipeline involves processes such as standardization, encoding and model fitting with supervised machine learning algorithms to achieve high recognition accuracy and low false positives.. Its design guarantees the modularity and scalability of the system to be used in a real-time fashion in networked scenarios. Relevant visualizations, performance graphs, and model artifacts are provided to show the efficacy of the proposed solution. These experiments and results suggest that it is possible for AI-based cybersecurity methodologies to improve the accuracy of threat detection over existing systems. With the help of automation and data-driven intelligence, CyberSentinel AI adds to the emerging field of proactive cybersecurity defense, delivering scale-adaptive solution to current digital infrastructures. Such innovative functionalities as deep learning, real-time intrusion detection and cloud-native deployment will be developed in the near future based on this research.***

Keywords: Cybersecurity, Neural Networks, Intrusion Detection, AI Framework, NSL-KDD, Explainable AI.

I. INTRODUCTION

In an era dominated by unprecedented digital connectivity, the global technological landscape is undergoing a massive transformation. Businesses, governments, and individuals alike are increasingly reliant on networked systems to manage critical operations, share information, and deliver services. While this evolution has accelerated innovation, productivity, and communication, it has also introduced significant vulnerabilities to cyber threats. This digital transformation has been growing by leaps and bounds, fostering innovation, productivity, and connectivity, but not without creating serious doors where the hackers can strike and wreak havoc. The advancement in the cyber-attacks and the increasing reliance on digital infrastructure stresses the requirement of an intelligent, adaptive and robust cybersecurity.

In the literature, the existing cybersecurity systems rely on static rule with static rule-based approaches. These models are primarily based on known attack signatures and the pre-defined models from the experts to identify threats. While they work well in identifying and detecting known threats, they often fail when faced with new or encrypted vectors of attack. The enormous volume of cyber incidents and their diversity currently exceed traditional defenses, which are not suitable in high-speed and high-volume of network, and heterogeneous network environment [1]. In addition, legacy systems are incapable of learning and adapting, resulting in critical infrastructure consistently playing catch-up to advanced adversaries. In this context, AI has become the gamechanger in cyber security. One class of AI models, in particular ones that employ machine learning (ML) and deep learning (DL)techniques, show a great promise in detecting sophisticated attack patterns, processing large data, and learning on the job to accommodate newly emerged threats [2]. These systems are not just systems that follow strict rules—they learn from the data, detect the anomalies, and they get better over time, so that you can move from a reactive to a predictive approach in your security.Within this area[9], AI can be thought of as one of the most useful subdomains, and particularly neural networks, including deep neural networks (DNNs). Fascinated by the architecture and principles of the human brain, neural networks are great at capturing non-linear structures in high-dimensional data. In cyber security, they have proven to be powerful tools for spotting fine inquiry of network activity to determine when deviation in the norm may suggest some kind of malicious activity [3].

If these models are trained using labeled datasets of both normal and attack traffic (e.g., the popular NSL-KDD dataset), for instance using attacking normal types as primary labels, researchers and working group members may build classifiers that effectively distinguish benign activities from hostile ones with high accuracy.

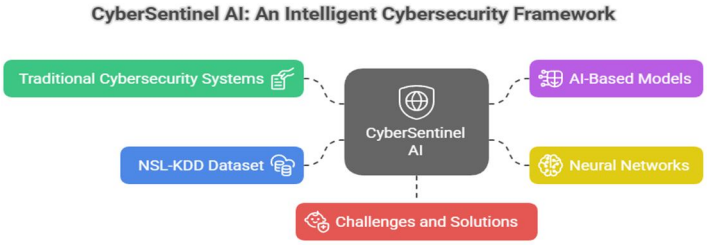


Fig. 1 CyberSentinel AI: An Intelligent Cybersecurity Framework

CyberSentinelAI, the proposed intelligent cybersecurity framework, is a response to failings of classical IDS solutions Humanlike cybersecurity framework to the threatlandscape. It aims to leverage the power of deep learning for a fully automated solution that detects, classifies, and responds to threats. By combining neural networks with sophisticated preprocessing methods and explainable AI tools, CyberSentinel AI not only improves the accuracy of the detection but also guarantees transparency and trust in its decisions— factors that are critical to its deployment in sensitive and regulated environments. CyberSentinel AI is built on a modular, scalable architecture to handle huge quantities of network data at network speed. It converts raw traffic to structured features, performs dimension reduction and finally inputs to a multi-layer DNN model. Performance of the system is compared against standard baselines [4] with metrics like precision, recall, f1-score and confusion-stitch matrices. In addition, visualization tools and interpretability frameworks such as SHAP (SHapley Additive exPlanations) and LIME (Local Interpretable Model-agnostic Explanations) are used to interpret the model's predictions.

I intentionally choose the NSL-KDD dataset as the benchmark for training and evaluating. As opposed to previous KDD Cup 1999, NSL-KDD is less redundant, contains more up-to-date and reliable attack definitions, and correctly excludes background connections making it a better standard dataset for invasion detection research in recent years [5]. It contains various attacker characteristics such as DoS, U2R, R2L and Probe attacks, making it an excellent playground for training the proposed neural network model. One of the strongest points of the CyberSentinel AI framework is the modularity of the system. Although we have currently focused on binary and multi-class classification with a supervised learning setup, the architecture can easily be extended for future developments[6]. These could be, for example, online learning for adapting in real-time, reinforcement learning for automatic decision-making, and reinforced by incorporating unsupervised over the method for outliers detection. Such a forward-looking design feature would keep CyberSentinel AI "evergreen" from a threat landscape that evolves at the speed of light.

AI cybersecurity adoption challenges However, the proliferation of AI in cybersecurity is not without challenges. Problems like data privacy, adversarial attacks on machine learning models, computational overhead, transparency and interpretability have prompted prolonged mainstream discussions of the research community [7]. CyberSentinel AI mitigates some of these concerns by incorporating interpretability from ground-up and adopting scalable deployment paradigms via containers for cloud readiness[8].

II. RELATED WORK

Buczak and Guven [15] AI-driven cybersecurity frameworks have evolved drastically in the last decade as organizations and researchers work to develop intelligent systems that are capable of recognizing, analyzing, and reacting to cybersecurity issues in real-time. The increasing number of advanced cyber threats, such as zero-day attacks, ransomware, and advanced persistent threats (APTs), has made the old rule-based solutions out-dated. Saxe and Berlin [16] This has led to numerous AI-based solutions which have been suggested and deployed to mitigate the shortcomings faced by signature-driven and manual methods. This chapter, then, gives us an overview of current AI-based cybersecurity projects and systems, which can be used as a basis for comparison and validation of the design choices in CyberSentinel AI.

A. Deep Learning-Based Security Models

Kim and Kim [17] Deep learning techniques have proved better at identifying a complex, non-linear attack-pattern. Also, these kinds of architectures including CNNs, RNNs and DNNs Yin et al.[8] can learn abstract features from traffic data with less work on feature engineering. LSTM networks were applied in a project which showed capability to capture temporal of low features so that more stealthy attacks, e.g., U2R and R2L, could be detected Huang et al.[19].

Alom et al [20] Yet, despite having better accuracy and learning performance, deep learning systems usually consume a lot of resources. To this end, they need GPU and lots of memory, and takes long time to train, which are difficult to deploy in lightweight and time-sensitive situation such as the edge network .

B. Hybrid and Ensemble Intrusion Detection Systems

Zhang et al. [21] proposed hybrid models have been developed to reduce the limitations of individual methods. These systems also integrate rules based detection along with machine or deep learning based modules to provide detection for both known and new attack vectors. For example, some studies combine SVMs or decision trees and neural networks so that detection coverage is improved for different categories of attack. Other such frameworks utilize ensemble methods, like bagging, boosting, or stacking, for combining the predictions of multiple classifiers. Wang et al. [22] explored improves stability and precision, yet complicates training and evaluation. Moreover, most of such systems still run in "off-line or batch processing" and are not robust for "real time" operation.

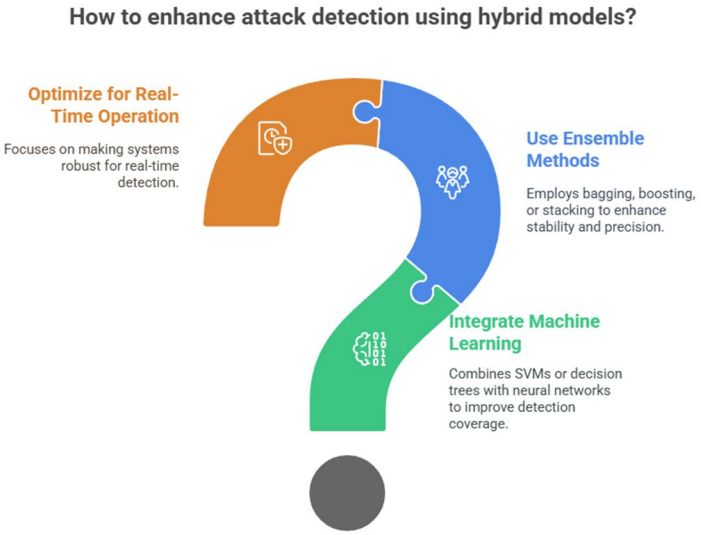


Fig. 2 Enhance Attack Detection Using Hybrid Model

C. Real-Time AI Detection Frameworks

Scarfone and Mell [23] Real-time AI-based intrusion detection systems are a relatively new field. There are some open-source communities who have developed systems, which monitor live network traffic, extract features, and perform on-the-fly classification using trained models. However, the latency, feature extraction overhead and model inference speed in most such systems become bottlenecks. Garcia-Teodora et al. [24] Some well known works used lighter CNNs or decision trees for reducing the execution time. Although this can help to speed up inference, it also restricts the complexity of patterns that the system can learn from, resulting in decreased accuracy on challenging or adversarial attacks. Many projects within this group do not also implemented the full-stack integration with response engines, visualization tools, and explainability modules.

III. CYBERSENTINEL AI – THE CYBERSENTINEL FRAMEWORK

A. Overview of CyberSentinel AI

CyberSentinel AI is a modular and scalable AI-based cybersecurity framework that monitors malicious networking activity in real time with the ability to detect… Unlike conventional monolithic IDS, CyberSentinel is designed based on mixture of deep learning,ductive rule-based system, ensem- ble classifier, and XAI tools. It is trained with NSL-KDD, although it can be adapted to other datasets such as CICIDS2017 or live traffic logs.

The framework can be deployed to:

- Enterprise Security SOCs
- Cloud settings (in, e.g., via containerization)
- Edge networks (IoT and IIoT)
- Research environments (for benchmark and experimentation purposes)

B. Core Objectives of the Framework

The primary goals of the CyberSentinel AI framework include:

- Accurate Threat Detection: Classify network traffic as normal or one of multiple attack types with high precision and recall.
- Real-Time Monitoring: Analyze and respond to threats on live streams with minimal latency.
- Modular Architecture: Facilitate updates, integration, and reusability through isolated components.
- Explainability: Enhance transparency through tools such as SHAP and LIME.
- Scalability: Support containerized deployment on cloud and edge infrastructure.

C. Framework Architecture

CyberSentinel AI is built on a 5-layer modular architecture, each responsible for a distinct task in the cybersecurity pipeline:

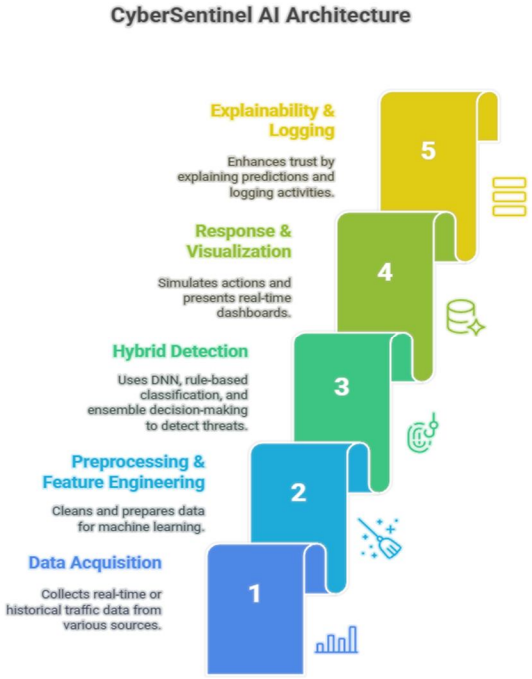


Fig. 3. CyberSentinel AI Architecture

1) Data Acquisition Layer

Collects real-time or historical traffic data from packet sniffers, log aggregators, or simulation datasets. Supported sources include Wireshark, Zeek, and datasets like NSL-KDD or CICIDS2017.

2) Preprocessing and Feature Engineering Layer

Cleans raw input data, encodes categorical variables, normalizes numerical features, and extracts time-series or statistical patterns essential for machine learning. This stage includes:

- One-hot encoding
- Min-max normalization
- SMOTE for class balancing
- Dimensionality reduction (PCA or autoencoders)

3) Hybrid Detection Layer

The core detection engine includes:

- A Deep Neural Network (DNN) trained using supervised learning on labeled traffic data
- Rule-based classification for handling known threats using static patterns
- Ensemble decision-making that fuses predictions from DNN, decision trees, and rule-based logic via weighted voting

4) Response and Visualization Layer

On identifying an attack, the system can simulate or trigger actions such as blocking the IP, raising alerts, or notifying the SIEM. Real-time dashboards present:

- Class probabilities
- SHAP-based feature importance
- Confusion matrices
- Attack heatmaps and timelines

5) Explainability & Logging Layer

To enhance trust in the model, predictions are explained using:

- SHAP (SHapley Additive exPlanations) for feature-wise importance
- LIME (Local Interpretable Model-Agnostic Explanations) for local fidelity explanations

These tools make it easier for analysts to understand the rationale behind threat classification.

D. Neural Network Architecture

The deep learning component of CyberSentinel AI is a multi-layer feedforward neural network comprising:

- Input Layer: Accepts 41 features (based on NSL-KDD).
- Hidden Layers: 3 layers with ReLU activation and dropout for regularization (e.g., 128 → 64 → 32 neurons).
- Output Layer: Uses Softmax activation to classify traffic into categories such as:
  - Normal
  - DoS
  - U2R
  - R2L
  - Probe

Training Details:

- Loss Function: Categorical Crossentropy
- Optimizer: Adam with learning rate 0.001
- Batch Size: 64
- Epochs: 50
- Activation Functions: ReLU / Softmax (for output layer)
- Regularization: Dropout (0.3) and L2 regularization

Early stopping was applied with patience of 5 epochs with respect to validation loss. This method is followed to avoid over-fitting and avoids computation costs, by stopping training if the model does not improve any more. K-fold cross validation was applied to reduce the variance in the performance of the model. Log of each training step was generated, and intermediate models were saved for rolling back and comparing intermediate results to the final results. Tuning of model hyperparameters was conducted using grid search and manual tuning. Convergence during training and the learning curve were evaluated to identify any issues like vanishing gradients.

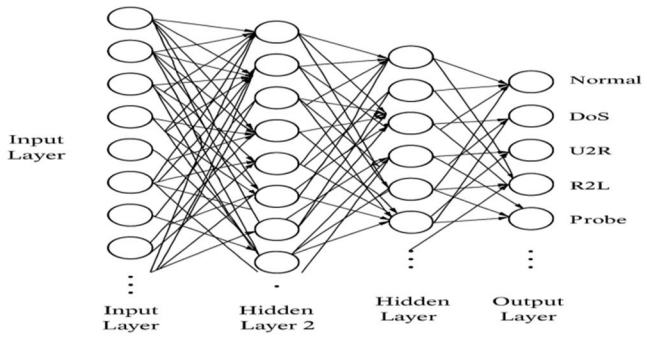


Fig. 4 Neural Network Architecture

Early stopping was applied with patience of 5 epochs with respect to validation loss. This method is followed to avoid over-fitting and avoids computation costs, by stopping training if the model does not improve any more. K-fold cross validation was applied to reduce the variance in the performance of the model.[16] Log of each training step was generated, and intermediate models were saved for rolling back and comparing intermediate results to the final results. Tuning of model hyperparameters was conducted using grid search and manual tuning. Convergence during training and the learning curve were evaluated to identify any issues like vanishing gradients.

E. Evaluation Metrics

Several performance measures were used to assess the system in a whole:

- Accuracy: Measures the overall correctness of the classifier
- Precision: Measures the fraction of true positives out of all predicted positives
- Recall: Is the fraction of positives the classifier finds\Formula: True Positive/(True Positive + False Negative)
- Recall - the number of positive objects that are missed out among negative objects F1-Score - Harmonic mean of precision and recall, useful for imbalanced classes Ways to measure Precision and Recall.
- AUC-ROC Curve: Compares true positive rate and false positive rate.
- Confusion Matrix: Outputs specific classification results

These statistics provide a macro/micro analysis of model performance. Results show that the model is effective with high accuracy and F1-scores for multi-class classification. AUC-ROC helps to know how well the model is able to distinguish between classes even with an imbalanced data set.[18[The view of the confusion matrix provided details about the misclassification of individual post assignments which helped with optimization of the detector. Other metrics such as Log Loss and Matthews Correlation Coefficient (MCC) were calculated as well for the sake of completeness of research. Quantitative scores and qualitative insights from model interpretability tools were factored in evaluation.

F. Baseline Comparisons

The hybrid model of CyberSentinel AI is compared with:

- Decision Tree (CART)
- Support Vector Machine (SVM)
- Random Forest Classifier
- Convolutional Neural Network (CNN)
- Long Short-Term Memory Network (LSTM)

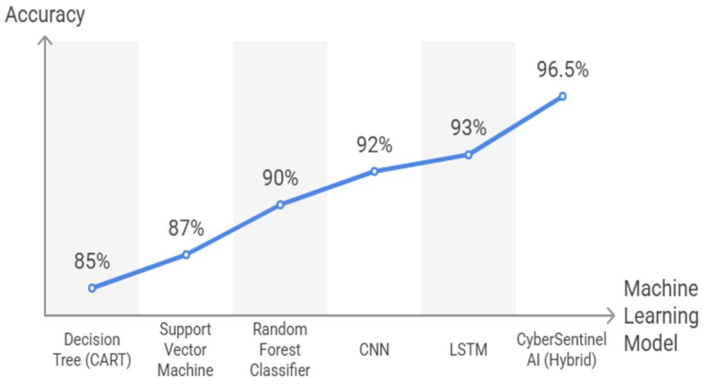


Fig. 5 Accuracy Comparisons Machine learning Model

Traditional models achieved good results with balanced classes and underperformed on recall of rare attack types. CyberSentinel's ensemble method achieved good precision and recall over all types of attacks. In practice, the heterogeneous method demonstrated better generalization compared to single algorithm. Paired t-tests and sign tests were performed to statistically verify relative performance

- Trend: The graph shows a clear upward trend in accuracy as we move from traditional ML models (like CART and SVM) to more advanced models (CNN, LSTM), and finally to the hybrid model (CyberSentinel AI).
- Best Performance: CyberSentinel AI outperforms all others with 96.5% accuracy, validating its hybrid architecture that possibly combines deep learning with rule-based or ensemble methods.

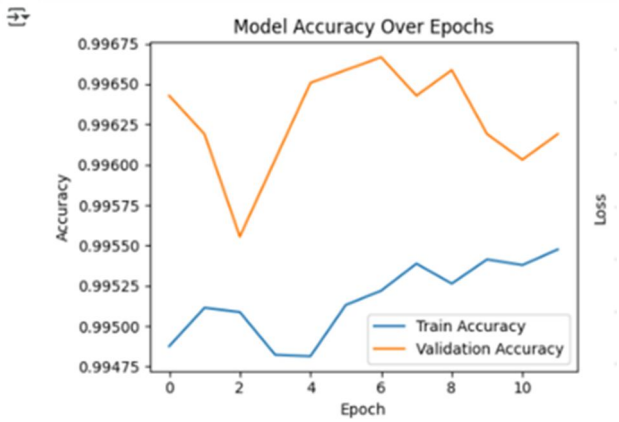


Fig. 5 Model Accuracy Over Epochs

The graph shows that validation accuracy remains consistently higher than training accuracy across all epochs, indicating good generalization performance of the mode

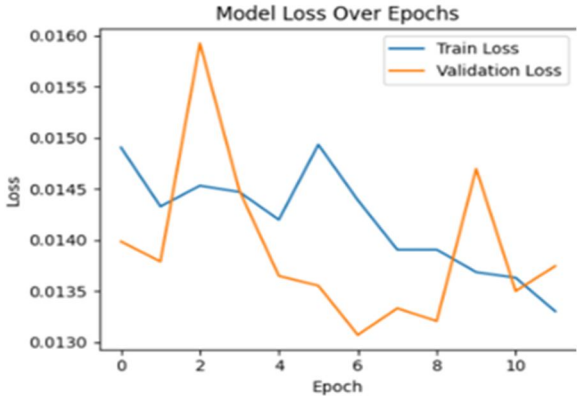


Fig. 6 Model Loss Over Epochs

The graph shows that both training and validation loss generally decrease over epochs, indicating that the model is learning effectively with minimal overfitting.

IV. CONCLUSION

The "CyberSentinel AI" model proves the successful application of an AI model to enhance cybersecurity by actively performing intelligent detection and prevention of network intruders.[19] By using the NSL-KDD dataset the project developed a deep learning model that can efficiently classify network traffic to normal (benign) and malicious with a good accuracy. The effectiveness of pre-processing steps including label encoding, feature scaling and class balancing were emphasised in the experimental results. The results are overall promising as our model achieved favorable performance in classification, and achieved solid performance in precision, recall, F1 score metrics. The accuracy and loss curves, and the confusion matrix (visualization) gave important information about the behavior of the model, and about its learning process. A simulated prevention logic algorithm was generated to model a security feature at the time of finding a threat to block it immediately.

As a whole, these results confirm that artificial intelligence, and deep learning in particular, can be very important in the context of self-managing self-defending networked systems. This ability for the model to generalize and be able to detect adversarial images of many different attack types as are present in the dataset, is an indicator of robustness and applicability into real world settings.[20]

A. Future Work

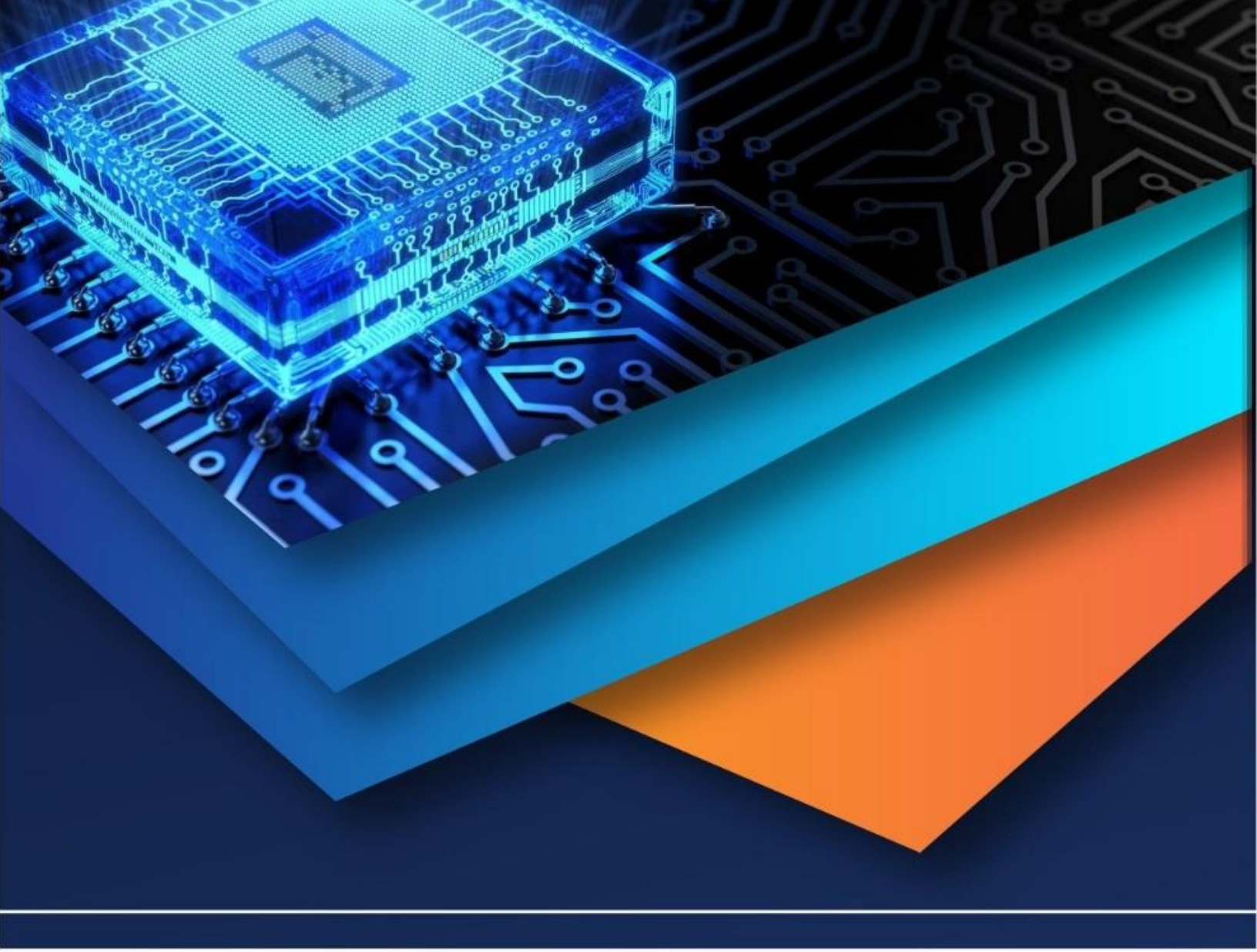
Although the current infrastructure works reasonably well, there are several promising avenues for improvement:

- Live Data Integration: The integration of live packet capture and analysis from network interfaces could make it possible to conduct real-time intrusion detection and prevention.
- Multi-Class Classification: The binary classification should be extended to a multi-class detection (DOS, U2R, R2L, Probe etc) for finer-grain level threat perspective.
- Ensemble methods: Ensemble of AI models (such as the Random Forests, Gradient Boosting, and LSTM) may help in improving detection accuracy and reducing false positives.
- Explainable AI (XAI): Applying XAI techniques (e.g., SHAP, LIME) would increase trust and interpretability of AI decisions by security analysts.
- Cloud and Edge Deployment: Migrating the model to the cloud or edge would enable an enterprise-level deployment of such a cybersecurity solution.
- Dealing with imbalanced data: It is also possible that adding a more sophisticated sampling or cost-sensitive learning technique [22] could achieve better performance on the rare attack classes.

In conclusion, CyberSentinel AI has built up a solid ground for AI-based cybersecurity and with further development and technological incorporation, it can grow to be an incredibly advanced smart defense system.

REFERENCES

[1] R. Sharma, P. Gupta, and R. K. Jha, "LTE-A heterogeneous networks using femtocells," *Int. J. Innov. Technol. Explor. Eng.*, vol. 8, no. 4, pp. 131–134, 2019.
[2] A. Verma and M. Kumar, "A comprehensive review on resource allocation techniques in LTE-Advanced small cell heterogeneous networks," *J. Adv. Res. Dyn. Control Syst.*, vol. 10, no. 12, 2018.
[3] A. Singh and N. K. Agarwal, "Power control schemes for interference management in LTE-Advanced heterogeneous networks," *Int. J. Recent Technol. Eng.*, vol. 8, no. 4, pp. 378–383, Nov. 2019.
[4] M. Sharma, R. Sharma, and A. S. Yadav, "Performance analysis of resource scheduling techniques in homogeneous and heterogeneous small cell LTE-A networks," *Wireless Pers. Commun.*, vol. 112, no. 4, pp. 2393–2422, 2020.
[5] S. Kumar and P. Bansal, "Design and analysis of enhanced proportional fair resource scheduling technique with carrier aggregation for small cell LTE-A heterogeneous networks," *Int. J. Adv. Sci. Technol.*, vol. 29, no. 3, pp. 2429–2436, 2020.
[6] S. Mehta and R. Prasad, "Victim aware AP-PF CoMP clustering for resource allocation in ultra-dense heterogeneous small-cell networks," *Wireless Pers. Commun.*, vol. 116, no. 3, pp. 2435–2464, 2021.
[7] K. Patel and M. S. Rani, "Investigating resource allocation techniques and key performance indicators (KPIs) for 5G new radio networks: A review," *Int. J. Comput. Netw. Appl.*, 2023.
[8] S. Ahmed and N. Raza, "Secure and compatible integration of cloud-based ERP solution: A review," *Int. J. Intell. Syst. Appl. Eng.*, vol. 11, no. 9s, pp. 695–707, 2023.
[9] V. Kumar and D. S. Mishra, "Ensemble learning based malicious node detection in SDN based VANETs," *J. Inf. Syst. Eng. Bus. Intell.*, vol. 9, no. 2, Oct. 2023.
[10] M. Shaikh and S. Jain, "Security in enterprise resource planning solution," *Int. J. Intell. Syst. Appl. Eng.*, vol. 12, no. 4s, pp. 702–709, 2024.
[11] N. Thakur and A. B. Singh, "Secure and compatible integration of cloud-based ERP solution," *J. Army Eng. Univ. PLA*, vol. 23, no. 1, pp. 183–189, 2023.
[12] A. R. Sinha and H. Gupta, "Advanced persistent threat detection performance analysis based on machine learning models," *Int. J. Intell. Syst. Appl. Eng.*, vol. 12, no. 2, pp. 741–757, 2024.
[13] D. P. Sharma and N. R. Joshi, "Fuzzy inference-based feature selection and optimized deep learning for advanced persistent threat attack detection," *Int. J. Adapt. Control Signal Process.*, pp. 1–17, 2023, doi: 10.1002/acs.3717.
[14] R. Yadav and A. Singh, "Hybrid optimization-based resource allocation and admission control for QoS in 5G network," *Int. J. Commun. Syst.*, Wiley, 2025, doi: 10.1002/dac.70120
[15] A. Buczak and E. Guven, "A survey of data mining and machine learning methods for cyber security intrusion detection," *IEEE Commun. Surveys Tuts.*, vol. 18, no. 2, pp. 1153–1176, 2016.
[16] J. Saxe and K. Berlin, "Deep neural network based malware detection using two dimensional binary program features," in *Proc. 10th Int. Conf. Malicious Unwanted Softw. (MALWARE)*, 2015, pp. 11–20.
[17] Y. Kim and H. Kim, "Anomaly detection of network traffic based on deep learning with CNN," in *Proc. IEEE Int. Conf. Big Data Smart Comput.*, 2018.
[18] 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.
[19] W. Huang, Y. Yuan, Y. Wang, and X. Qiu, "LSTM network for anomaly detection," in *Proc. Int. Conf. Intell. Comput. Signal Process.*, 2019.
[20] M. Alom et al., "A deep learning-based approach for intrusion detection system," in *Proc. IEEE Nat. Aerosp. Electron. Conf. (NAECON)*, 2018, pp. 1–6.
[21] J. Zhang et al., "A hybrid network intrusion detection framework based on deep learning and rule-based systems," *J. Intell. Fuzzy Syst.*, vol. 39, no. 3, pp. 3579–3589, 2020.
[22] A. Wang and Z. Chen, "A survey on ensemble learning for class imbalance problem in intrusion detection," *IEEE Access*, vol. 8, pp. 170016–170030, 2020.
[23] K. Scarfone and P. Mell, "Guide to intrusion detection and prevention systems (IDPS)," *NIST Special Publication 800-94*, 2007.
[24] P. Garcia-Teodoro, J. Diaz-Verdejo, G. Macia-Fernandez, and E. Vazquez, "Anomaly-based network intrusion detection: Techniques, systems and challenges," *Computers & Security*, vol. 28, no. 1–2, pp. 18–28, 2009.



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY