Journal of Computer Science and Technology Studies
ISSN: 2978-8048
DOI: 10.32996/jcsts.2026.5.8.1
Journal Homepage: www.al-kindipublisher.com/index.php/jcsts

JCSTS
AL-KINDI CENTER FOR RESEARCH AND DEVELOPMENT

| RESEARCH ARTICLE

Intelligent Artificial Intelligence-Powered Cyber Threat Detection System for Cloud Security Enhancement

Naveen Reddy Thumma
Technical Account Manager(Devops/SRE)
Corresponding Author: Naveen Reddy Thumma, **E-mail**: reddynaveenthumma@gmail.com

| ABSTRACT

Network intrusion detection is a key part of keeping computer networks and the Internet of Things safe from cyber dangers. When it comes to decision-making, standard machine learning methods can only improve detection accuracy (ACC). This work introduces an explainable intrusion detection method that utilizes the CICIDS2017 dataset and depends on a Random Forest classifier (RF), LightGBM (Light Gradient Boosting Machine), and Voting Ensemble (VE). The data processing, feature selection, data balancing with SMOTE, and ensemble learning techniques of the proposed method are the foundation of intrusion detection ACC. At an ACC of 98.22%, the Voting Ensemble model outperforms the other models mentioned in the discussion (Multiple Baseline models), including the Autoencoder (AE), Decision Tree (DT), Deep Neural Network (DNN), and Logistic Regression (LR) classifiers. To further enhance model transparency and determine the most important network traffic features for attack prediction, SHAP-based explainable AI analysis is used. The proposed framework is able to identify influential traffic aspects that lead to attack prediction, subsequently leading to an accurate, scalable and trustworthy solution for current intrusion detection applications, thereby increasing the interpretability of cybersecurity.

| KEYWORDS

Cyber security, Anomaly Detection, Network Security, Threat Detection, Intrusion Detection System, Machine Learning, CIC-IDS2017.

| ARTICLE INFORMATION

ACCEPTED: 01 June 2026 **PUBLISHED:** 27 July 2026 **DOI:** 10.32996/jcsts.2026.8.8.15

I. Introduction

The recent proliferation of mobile devices, cloud computing, and smart technologies has greatly improved the functioning of networks and allowed for real-time data sharing and smarter decision-making. But the same speedy advancement of technology has also caused a significant increase in advanced cyber-attack techniques – cloud-based attacks [1]. Additionally, new security issues have emerged due to the proliferation of smart devices and the growing complexity of cloud network topologies [2]. Many smart devices are created with restricted functions and frequently do not have the fundamental security capabilities, which makes them extremely susceptible to cyberattacks [3]. There is a higher risk of exploitation for these devices since security is often not considered during design. Better cloud security is possible with an intrusion detection system. Malware, unauthorized access, and network attacks are among the dangerous acts it detects and blocks [4][5].

ML and DL, two branches of AI, have substantially enhanced IDS capabilities, enabling them to identify novel threats like zero-day vulnerabilities [6][7]. Data imbalance and high-dimensional feature spaces are two of the many issues that IDS still faces. Biased forecasts favouring the dominant class and disregarding crucial but uncommon assaults may be the outcome of imbalanced datasets [8][9]. Also, high-dimensional data makes the computer more complicated to use and can create the danger

of over-fitting, that can diminish the effectiveness of the IDS. It is necessary to solve these problems in order to enhance the effectiveness of IDS and the whole network security [10]. Enhancing cyber threat detection and providing decision-making transparency may be achieved by the construction of an intrusion detection system that is resilient, accurate, and explainable. Thus, by combining Ensemble Learning with XAI approach, this work seeks to enhance the detectability, explain ability, and usability of existing cyber security systems. This paper is summarized in the following way:

- Introduces a robust and interpretable cyber threat detection framework capable of handling complex network intrusion patterns with high classification reliability.
- Demonstrates the effectiveness of ensemble-based intelligent learning for improving cybersecurity detection performance over conventional machine learning approaches.
- Enhances model transparency and trustworthiness through Explainable AI analysis, enabling better understanding of feature-level attack prediction behavior.
- Achieves superior detection capability and balanced classification performance compared with several existing baseline intrusion detection models.
- Provides a scalable and practical cybersecurity solution suitable for real-world intelligent network monitoring and threat analysis environments.

The proposed work is innovative because it combines Explainable AI with ensemble learning to reliably detect cyber threats. The proposed method has the capability to achieve excellent detection performance and explainable decision-making due to the use of RF Classifier, LightGBM and interpretability through SHAP, which differ from traditional intrusion detection methods that rely primarily on prediction ACC. The study is further justified by the increasing need for accurate and explainable cybersecurity systems capable of identifying complex network attacks while supporting trustworthy real-world deployment and security analysis.

A. Organization of Paper

The structure of the paper is as follows: Other relevant works are reviewed in Section II. The suggested approach is elaborated upon in Section III. The experimental data and their interpretation are presented in Section IV. The paper is wrapped up in Section V.

II. LITERATURE REVIEW

Discuss the prior work on intrusion detection systems in this section. By means of a number of recently suggested methods and technologies. S *et al.* (2026) provide near-real-time detection and analysis of cyber threats using an AI-based honeypot IDS in conjunction with ML and DL technologies. network traffic and attacker-honeypot interactions using CNNs, RNNs, and LSTM models. The suggested system was tested experimentally and found to have an intrusion detection ACC rate of 96.8%, a false positive rate of 2.1%, and an F1 score of 96.2% [11]. Ileana, Petrov and Milev, (2025) use edge-level, lightweight XAI models for interpretability and real-time detection, while transferring complicated threat analysis to a secure, cloud-backed infrastructure made possible by scalable distributed web technologies. The system outperforms conventional models in terms of detection rate (up to 96%), latency (sub 50 milliseconds), and resource consumption (less CPU and memory used) in experimental validation on the CIC-IDS2017 dataset [12].

Santhoshkumar *et al.* (2024) provide a long-term and extensible solution to protect critical infrastructures from evolving cyber risks. When intrusion detection systems (IDS) and digital twin technology (DTT) were used for cyber-physical security, the identification of threats was much improved. The results showed that early intrusion detection ACC increased by 40% as compared to the standard models of intrusion detection systems [13]. Tocci, Zhou and Zhang, (2023) The training process makes use of a Deep-Q Learning strategy, with training data generated from the CICIDS2017 dataset. The agent achieves a 90% success rate in distinguishing between port scan attacks, DDoS attacks, and benign traffic after training [14]. Mouti *et al.*, (2022) Physical network management has been examined using a multi-connect variational auto-encoder. In comparison to centralized DL-based detection systems, distributed attack detection systems perform better in tests. The suggested MCVAE_PBNN worked as expected on the UNBS-NB-15 dataset, with 96% ACC, 71% FPR, 92% sensitivity, 82% specificity, 63% false positive rate, and 75% area under the curve. On the KDD99 dataset, it was 95% accurate, 68% false positive rate, 92% sensitivity, 84% specificity, 61% false positive rate, and 78% AUC [15].

Research Gap: Although existing IDS achieve high ACC using AI, DL, XAI, and reinforcement learning techniques, many models still face challenges related to high computational complexity, false alarm rates, limited scalability, and poor real-time adaptability. Moreover, previous research typically uses a single-model approach and standard datasets, which fail to address the ever-changing, constantly evolving threats to cyber systems. Furthermore, most studies focus on single-model approaches with standard datasets, making them ineffective against growing and new cyber threats. In today's network environment, a lightweight, scalable and efficient ensemble-based IDS framework with better detection ACC and lower FPR is needed.

III. METHODOLOGY

The suggested intrusion detection system's workflow is illustrated in Figure 1, utilizing the CIC-IDS2017 dataset. Initially, the dataset is collected and analyzed through Exploratory Data Analysis (EDA), then data undergoes pre-processing, label transformation, feature selection, and feature normalization. After pre-processing, the dataset is split into training and testing sets, then balanced using SMOTE to address class imbalance. Subsequently, RF, LightGBM, and Voting Ensemble models are trained and evaluated using ACC, PRE, REC, F1, and ROC-AUC, followed by SHAP-based explainable AI analysis for feature interpretation.

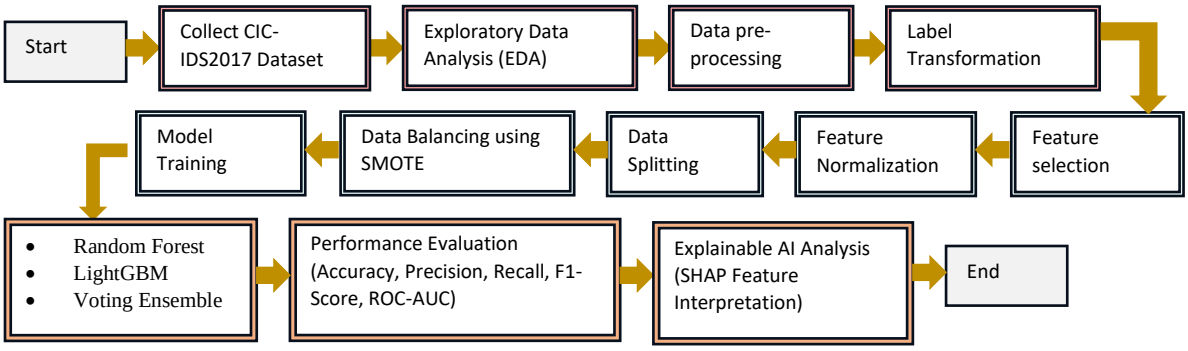

Fig. 1. Proposed Flowchart for Cyber Threat Detection

Cyber threat detection implementation steps are briefly explained below:

A. Data Acquisition

The CIC-IDS2017 Dataset, a standard dataset for cybersecurity studies, is utilized by the suggested intrusion detection framework. Various attack scenarios are represented in the dataset, which includes actual benign and malicious network traffic. The starting dataset contained 620,000 traffic instances with 78 network flow features, such as communication behavior, packet statistics, and protocol characteristics. These features present detailed knowledge to identify unusual or malicious behavior in network environments.

B. Exploratory Data Analysis (EDA)

The EDA is done to gain insight into the nature of data and network traffic. To build an effective intrusion detection model, EDA can help identify significant features, data imbalance and hidden relationships among these data.

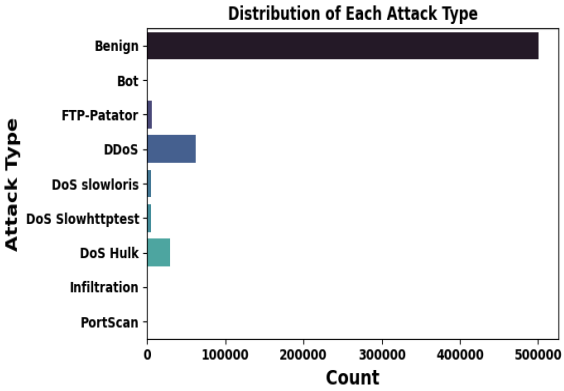

Fig. 2. Distribution of Network Attack Types

The network traffic categories are shown in the frequency distribution in Figure 2. There are many more benign traffic instances than malicious ones in the dataset, with almost 500,000 samples of benign traffic present. Comparing the attacks, PortScan and DDoS are quite common, whereas Bot, FTP-Patator, DoS variants (Slowloris, Slowhttptest, Hulk), and Infiltration are relatively uncommon, with fewer occurrences of each. The visualization features a difference between benign traffic and attack traffic which is a big issue in detecting rare, but valuable attacks within the network security dataset.

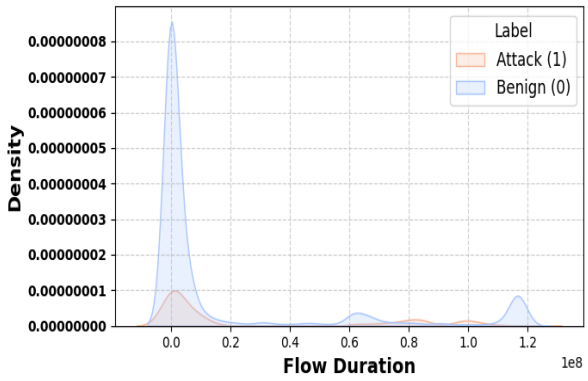

Fig. 3. Kernel Density Estimate of Flow Duration by Class

Figure 3 shows that the flow duration distribution of benign traffic is different from that of attack traffic. The blue curve (benign) has a large peak near 0 duration, and smaller peaks around $1.0x10^8$ and $1.2x10^8$, indicating that durations of the flow are short but with some longer durations. The overall density of the red curve (attack) is not as dense as the green curve (defense) but the shape is similar and fewer flows are of equal length. The visualization shows the differences in temporal behavior between normal and malicious traffic, which are useful for discriminating attack signatures in intrusion detection.

C. Pre-processing and Label Transformation

The obtained dataset underwent to a variety of pre-processing procedures intended to enhance data quality and get it ready for training machine learning models. The first step in reducing bias and eliminating redundancy is to identify and delete duplicate traffic data. The original dataset included 10,744 duplicate occurrences that were found and eliminated. Following the pre-processing, the final dataset size is 609,256 × 78. After that, Label Encoding is used for binary classification to turn the class labels into numerical values.

D. Feature Optimization

In order to enhance intrusion detection performance and decrease dimensional complexity, feature optimization is performed. The Mutual Information Score technique is employed to identify the most impactful attributes connected to the target labels. The top 20 features are chosen for further investigation and model training according to their ranking scores. Reducing the amount of duplicated data, increasing computation performance, and improving the model's generalizability are all achieved by this technique.

Figure 4 displays the twenty most informative network characteristics ranked by their mutual information ratings. Other essential metrics are average bandwidth segment size (0.347) and average packet length (0.347), with average packet size (0.392) being the most important. Counts of bytes sent and received, as well as the standard deviation and variance in packet length, are other highly regarded characteristics. Top 20 smallest: Fwd Packet Length Std (0.270). With this visual aid, selecting characteristics for intrusion detection models becomes much easier, as it becomes immediately apparent which traffic features are best at differentiating between safe and dangerous data transfers.

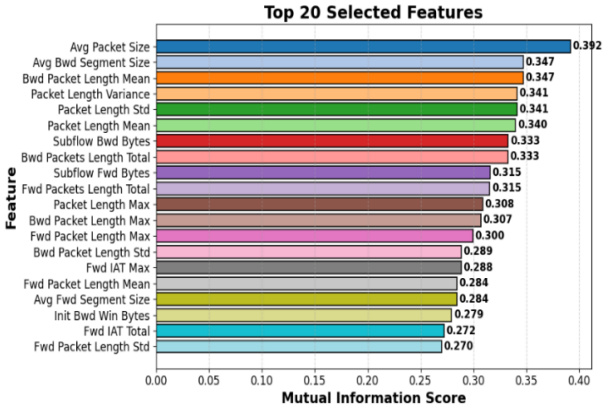

Fig. 4. Top 20 Selected Features by Mutual Information Score

E. Feature Scaling and Splitting

Feature scaling (FS) is the process of changing the values in a dataset to a certain scale (range) so that features with bigger scales don't take over the learning process. For the purpose of this analysis, the raw data have been transformed into a scale from 0 to 1 using the conventional scaling technique. Simply plugging in any dataset variable X into Equation (1) gives the standard value.

$$X_{stand} = \frac{X - mean(X)}{\text{standard deviation(X)}} \quad (1)$$

The efficiency of ML algorithms is enhanced by scaling the features, which helps to avoid issues associated with features having huge ranges. Training and testing sets are created from the normalized dataset by dividing it in half with a ratio of 80:20. Both the training set and the testing set contain 487,404 and 121,852 examples, respectively, that used to evaluate the model.

F. Class Imbalance Handling with SMOTE

Data balancing is a method for reducing the impact of class imbalance. SMOTE is employed to circumvent the problem of data imbalance. It uses the characteristics of the minority class to create simulated samples. After receiving the whole dataset as input, SMOTE increases the percentage of cases belonging to the minority exclusively. When it comes to balancing datasets, SMOTE is the finest option.

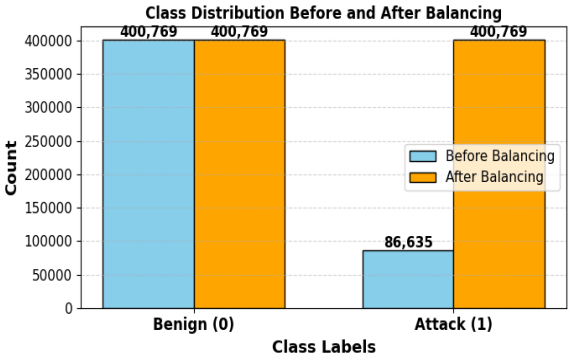

Fig. 5. Class Distribution Before and After Balancing

The class counts before and after balancing the dataset are shown in Figure 5. A large class imbalance occurred because there were 400,769 instances of benign traffic and 86,635 instances of attack traffic. The two classes are balanced and have 400,769 samples in each class, which is more representative. The visualization highlights the significance of balancing techniques in addressing skewed distributions, thereby promoting fairness and reliability in the development of IDS models.

G. Machine Learning Model Development

This section presents the creation of models for intelligent cyber threat detection using machine learning and ensemble learning. The study uses RF Classifier, LightGBM, and Voting Ensemble Classifier for the analysis of malicious traffic and benign traffic.

1) Random Forest (RF)

RF is a supervised ML method for fixing classification problems [16]. The model builds a number of DTs during training and then employs a majority vote to decide the outcome during prediction, both of which improve accuracy and reliability. The RandomForestClassifier model is optimized with 50 estimators, max_depth 3 and random_state 42 to provide a stable and reproducible result. The classifier learns complex traffic patterns from the balanced training dataset during training and makes reliable predictions in intrusion detection.

2) Light Gradient Boosting Machine (LightGBM)

LightGBM is used for high-speed IDS, as being trained quickly with low memory. The model is taught to better classify attacks in the attack category classification job by learning the intricate patterns seen in network data using gradient boosting decision trees. During training, the ideal environment is fine-tuned to minimize overfitting and maximize prediction performance, such as learning rate = 0.1, maximum depth = 5, and random state. The LightGBM classifier can accurately identify malicious network activities with high ACC and also effectively process big data in the cybersecurity information field.

3) Voting Ensemble Classifier

Voting Ensemble Classifier is an ensemble of Random Forest Classifier and LightGBM that improves IDS performance and classification stability. The ensemble model combines prediction probabilities of individual classifiers using soft voting to obtain the final prediction. The ensemble framework comprises two models: RF (n_estimators = 50, max_depth = 3, random_state = 42) and LightGBM (n_estimators = 100, learning_rate = 0.1, max_depth = 5, random_state = 42), which are used during training for robust attack detection. This hybrid method helps in increasing the generalization ability as well as decreasing the variance in the prediction and thus improves the overall cybersecurity threat classification ACC.

H. Performance Evaluation and Explainable AI Analysis

This study evaluates the performance of the ML models with the help of different evaluation indicators. These are based on the results of the confusion metrics. The evaluation measures, such as REC, ACC, PRE, f1, and ROC, are shown in Equations (2-5)

$$Accuracy = \frac{\mathrm{TN} + \mathrm{TP}}{\mathrm{TP} + \mathrm{TN} + \mathrm{FP} + \mathrm{FN}} \quad (2)$$

$$\mathrm{Precision} = \frac{\mathrm{TP}}{\mathrm{TP+FP}} \quad (3)$$

$$\mathrm{Recall} = \frac{\mathrm{TP}}{\mathrm{TP+FN}} \quad (4)$$

$$\mathrm{F1} = \frac{2*(\mathrm{precision}*\mathrm{recall})}{\mathrm{precision}+\mathrm{recall}} \quad (5)$$

The ACC rate is the probability of correctly performing positive and negative class predictions on all samples. The ACC rate is the percentage of positive samples that were accurately detected out of all the positive samples that were anticipated. The probability of a sample being appropriately identified as a positive sample out of all the positive samples is called the REC rate. The result of a harmonic mean of PRE and REC is the F1, which can have a value between 0 and 1.

Additionally, the classifiers are evaluated using the ROC-AUC methodology on both malicious and benign traffic. Moreover, using XAI analysis with SHAP (SHapley Additive Explanations) identifies the most impacting attributes of network traffic and explains the model predictions. The XAI framework helps with improved cybersecurity decision-making by increasing openness and making models more interpretable.

IV. Results and Discussion

A Windows 10 machine with 16 GB of RAM and a 64-bit Intel(R) Core(TM) i7-7500U CPU is used for all the testing. To complete this project, I used Python 3.7.2 and the PyTorch module version 1.9.0. Table I shows the experimental performance comparison of the cyber threat detection models based on data using the RF, LightGBM and Voting ensemble models. The Voting classifier had the best classification performance and balanced detection performance with an ACC of 98.22%, PRE of 94.49%, REC of 95.57% and F1 of 95.03%, among the evaluated classifiers. The LightGBM model also performed well with 97.73% ACC and high REC rate compared with the RF model in most of the evaluation metrics. The Voting model is more time-consuming and requires more training, but its results when it came to detection ACC are very impressive, showing that it can be very useful for reliable cyber threat identification.

TABLE I. Experiment Results of ML Models for Cyber Threat Detection

Metric	RF	LightGBM	Voting
Accuracy	96.46	97.73	98.22
Precision	84.71	90.25	94.49
Recall	97.81	97.86	95.57
F1-Score	90.79	93.90	95.03
Training Time (s)	0.0000	0.00001	131.0698
Prediction Time (s)	0.1460	0.0400	7.6422

The confusion matrices for RF, LightGBM and Voting Classifier models used for intrusion detection classification are presented in Figure. 6. The RF model achieved 96,294 true negatives and 21,248 true positives, with 3,834 false positives and 476 false negatives. The LightGBM model successfully classified 97,831 benign samples and 21,259 attack samples producing better

classification results. Similarly, the Voting model had 98,918 true negatives and predicted a lower number of false negatives, indicating that the Voting model possessed a higher detection efficiency than the RF model. The overall performance of Voting model and LightGBM model was better than the RF model in terms of classification capability and ACC of detection.

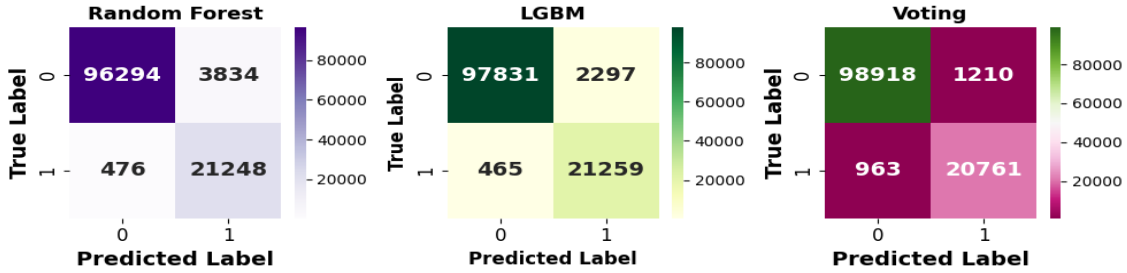

Fig. 6. Confusion Matrix Comparison of Proposed Models

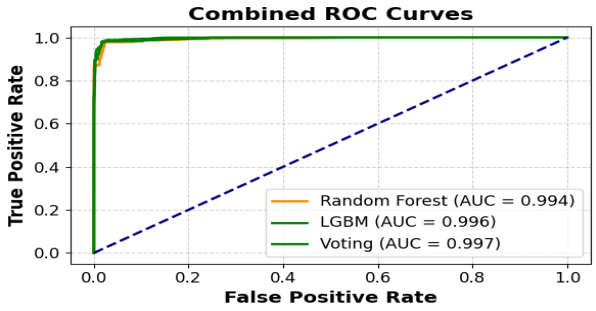

Fig. 7. Combined ROC Curves for Classifier Comparison

The ROC curves of three different models are displayed in Figure 7. When all curves are converging towards the upper left corner, the classification performance is high. While LightGBM and RF both obtain AUC values of 0.996 and 0.994, respectively, the Voting model comes out on top with an AUC of 0.997. The line of random chance is the diagonal dashed line. What is emphasized in this visualization is that all models achieve near-optimal accuracy and that ensemble learning only has a marginal improvement in terms of benign and attack traffic accuracy.

The SHAP values, shown in Figure 8, provide light on the relative importance of each component in the model's forecasts. Dots stand for sample locations, and the colors of the features match the corresponding values; blue denotes low feature values and red high feature values. The average packet size, average packet length, forward packet length standard, and forward packet length maximum are the attributes with the highest SHAP values. The points along the x-axis are used to show that high and low values of features can affect the predictions of the model and give information about how the model gets its interpretability as to the traffic characteristics that are the most important to the model's decision.

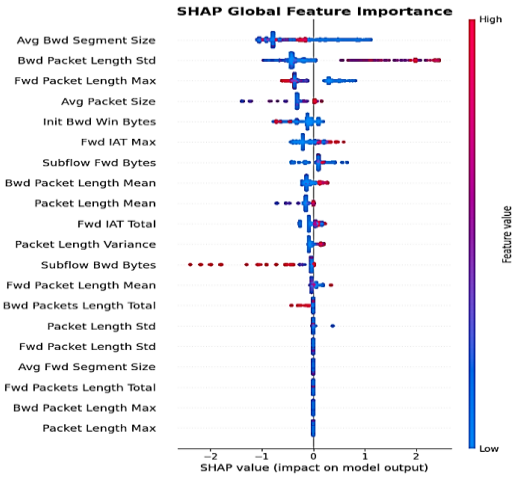

Fig. 8. Plot SHAP Global Feature Importance

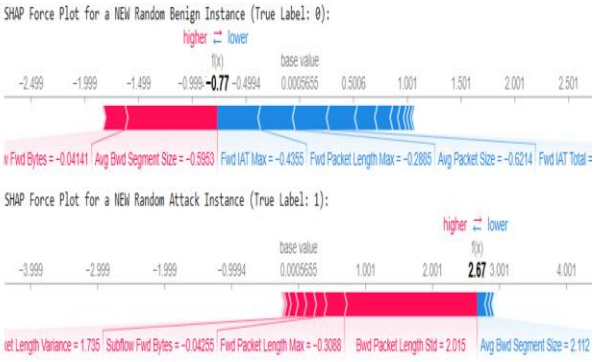

Fig. 9. SHAP Force Plot Analysis for Benign and Attack Traffic Instances

Figure 9 displays the analysis report for feature contributions using the RF model for both benign and attack network traffic examples. With the help of Avg Bwd Segment Size and Fwd Packet Length Max, the attack prediction probability was reduced for the benign situation. On the other hand, the attack case shows that features like Init Win bytes ahead and Bwd packet length standard led the model to classify attacks. This allows one to observe how the SHAP values characterize the model visually and determine which qualities are most important for the model's predictions.

A. Comparison and Discussion

Table II displays the results of the performance comparison between the suggested cyber threat detection model and the current baseline classifiers. This suggested model outperformed all of the others in terms of ACC, coming in at 92.3%, DT at 87.52%, DNN at 87%, and LR at 89.20%. Table II displays the results of the comparison between the suggested cyber threat detection model and the current baseline classifiers. The proposed model achieved the highest ACC, outperforming AE (92.3%), DT (87.52%), DNN (87%), and LR (89.20%) models. The proposed framework outperforms existing ones with PRE of 94.49% and REC of 95.57%, significantly improving the identification ACC of cyber threats with minimal misclassification. The proposed model proved to be more robust and reliable in cyber threat detection with a highly balanced overall performance of the F1 score of 95.03% compared to the DT + LR hybrid approach with a high PRE of 96.37%.

TABLE II. COMPARATIVE ANALYSIS BETWEEN BASE AND PROPOSED MODELS

Classifiers	**Accuracy**	**Precision**	**Recall**	**F1-Score**
AE [17]	92.3	91.7	91.5	91.6
DT [18]	87.52	87.23	-	87.90
DNN [19]	87	92	92	92
LR[20]	89.20	87.50	86.80	87.10
DT + LR[21]	94.91	96.37	94.19	95.12
Propose Model	**98.22**	**94.49**	**95.57**	**95.03**

The suggested study works well and can be trusted to find cyber threats in today's network environments. When compared to more traditional models, the ensemble-based classification approach improves harmful traffic detection while decreasing misclassification issues. The framework also provides better flexibility for managing complex and large amounts of network traffic. Furthermore, Explainable AI improves trust and transparency by making it easy to see how various traffic characteristics affect intrusion predictions. The proposed system is more suitable for practical use in cybersecurity applications, which require reliable detection and interpretability.

V. CONCLUSION AND FUTURE WORK

Traditional signature-based security mechanisms are difficult to deal with the increasing sophistication and volume of cyber-attacks. In this paper, an AI-based cybersecurity method is introduced which is based on machine learning techniques for intelligent threat detection to overcome these drawbacks. By combining machine learning with Explainable AI (XAI) methods, the proposed ensemble approach aims to achieve higher interpretability, classification ACC, and intrusion detection effectiveness. The results of the experimental analysis show that Voting Ensemble model outperforms individual classifiers with 98.22% ACC, 94.49%

PRE, 95.57% REC and 95.03% F1. The LightGBM model and the RF model also deliver good results with respective ACC values of 97.73% and 96.46%. Further comparative analysis reveals that the proposed model is superior to the other models, including AE, DT, DNN, LR and DT+LR. Moreover, SHAP analysis provides greater transparency of the model, allowing the identification of the most relevant characteristics of the network traffic for predicting attacks. But it has only binary classification and takes longer time to be trained in the ensemble. Research on future applications of XAI techniques in large-scale cybersecurity solutions, lightweight deep learning architectures, multi-class attack analysis, and real-time intrusion detection can be pursued.

Funding: This research received no external funding.
Conflicts of Interest: The authors declare no conflict of interest.
Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers.

References

[1] M. Alotaibi *et al.*, "Hybrid GWQBBA model for optimized classification of attacks in Intrusion Detection System," *Alexandria Eng. J.*, 2025, doi: 10.1016/j.aej.2024.12.057.
[2] A. Thakkar and R. Lohiya, "Attack Classification of Imbalanced Intrusion Data for IoT Network Using Ensemble-Learning-Based Deep Neural Network," *IEEE Internet Things J.*, 2023, doi: 10.1109/JIOT.2023.3244810.
[3] V. Sharma, "Security and Threat Mitigation in 5G Core and RAN Networks," *Int. J. Multidiscip. Res.*, vol. 3, no. 5, pp. 1–10, Sep. 2021, doi: 10.36948/ijfmr.2021.v03i05.54992.
[4] V. Sanikal, "AI-Enhanced Network Security Framework for Cloud-Edge Integrated Environments," in *2026 Innovations in Machine, Engineering, and Digital Conference (IMED)*, 2026, pp. 1–6. doi: 10.1109/IMED68921.2026.11484417.
[5] M. A. M. Farzaan, M. C. Ghanem, A. El-Hajjar, and D. N. Ratnayake, "AI-Enabled System for Efficient and Effective Cyber Incident Detection and Response in Cloud Environments," *IEEE Trans. Mach. Learn. Commun. Netw.*, Jan. 2025, doi: 10.1109/TMLCN.2025.3564912.
[6] D. Jain and S. Jain, "Artificial Intelligence (AI)-Driven Network Traffic Anomaly Detection for IT Infrastructure Security," in *2026 IEEE 5th International Conference on AI in Cybersecurity (ICAIC)*, IEEE, Feb. 2026, pp. 1–6. doi: 10.1109/ICAIC67076.2026.11395685.
[7] A. A. Eshmawi, A. Aldrees, and R. Alharthi, "Smart framework for industrial IoT and cloud computing network intrusion detection using a ConvLSTM-based deep learning model," *Front. Comput. Sci.*, vol. 7, Aug. 2025, doi: 10.3389/fcomp. 2025.1622382.
[8] V. K. Bollu, "Threat Landscape in Artificial Intelligence Systems: Taxonomy, Attack Vectors and Security Implications," *World J. Adv. Res. Rev.*, vol. 29, no. 1, pp. 285–294, 2026, doi: 10.30574/wjarr. 2026.29.1.0007.
[9] A. Fausto, G. Gaggero, F. Patrone, and M. Marchese, "Reduction of the Delays Within an Intrusion Detection System (IDS) Based on Software Defined Networking (SDN)," *IEEE Access*, 2022, doi: 10.1109/ACCESS.2022.3214974.
[10] M. A. Talukder, M. Khalid, and N. Sultana, "A hybrid machine learning model for intrusion detection in wireless sensor networks leveraging data balancing and dimensionality reduction," *Sci. Rep.*, 2025, doi: 10.1038/s41598-025-87028-1.
[11] S. P. S, V. Lanjewar, R. Chospal, S. Dorjey, M.Malleswari, and Bharathi B., "AI-Driven Honeypot-based Intrusion Detection System for Intelligent Cyber Threat Analysis," in *2026 9th International Conference on Inventive Computation Technologies (ICICT)*, 2026, pp. 58–63. doi: 10.1109/ICICT68280.2026.11511083.
[12] M. Ileana, P. Petrov, and V. Milev, "Intelligent Intrusion Detection in IoT with Explainable AI and Distributed Web Systems," in *2025 International Conference on Cybersecurity and AI-Based Systems (Cyber-AI)*, IEEE, Sep. 2025, pp. 124–127. doi: 10.1109/Cyber-AI66431.2025.11233693.
[13] S. P. Santhoshkumar, P. Sakthivel, A. Seetha, M. R, and M. Soni, "Enhancing Intrusion Detection Systems with Digital Twin Technology for Cyber-Physical Security," in *2024 International Conference on Distributed Systems, Computer Networks and Cybersecurity (ICDSCNC)*, 2024, pp. 1–8. doi: 10.1109/ICDSCNC62492.2024.10939345.
[14] D. Tocci, R. Zhou, and K. Zhang, "FPGA Accelerated Decentralized Reinforcement Learning for Anomaly Detection in UAV Networks," in *2023 IEEE 16th International Symposium on Embedded Multicore/Many-core Systems-on-Chip (MCSoC)*, 2023, pp. 248–253. doi: 10.1109/MCSoC60832.2023.00044.
[15] S. Mouti, S. K. Shukla, S. A. Althubiti, M. A. Ahmed, F. Alenezi, and M. Arumugam, "Cyber Security Risk management with attack detection frameworks using multi-connect variational auto-encoder with probabilistic Bayesian networks," *Comput. Electr. Eng.*, vol. 103, p. 108308, Oct. 2022, doi: 10.1016/j.compeleceng.2022.108308.
[16] S. Singh, "Advancing Network Security in 5G: Leveraging the 5G-NIDD Dataset for Intrusion Detection and Mitigation," in *2025 IEEE 12th International Conference on Cyber Security and Cloud Computing (CSCloud)*, IEEE, Nov. 2025, pp. 1–6. doi: 10.1109/CSCloud66326.2025.00055.
[17] A. Mehrin, "AI-Based Cyber Security Using Machine Learning For Intelligent Threat Detection," *Int. J. Creat. Res. Thoughts*, vol. 14, no. 4, pp. 447–454, 2026.
[18] T. Saranya and S. Indra Priyadharshini, "A Dual-Strategy Framework for Cyber Threat Detection in Imbalanced, High-Dimensional Data Across Heterogeneous Networks," *IEEE Access*, vol. 13, pp. 125313–125331, 2025, doi: 10.1109/ACCESS.2025.3582788.
[19] P. Waghmode, M. Kanumuri, H. El-Ocla, and T. Boyle, "Intrusion detection system based on machine learning using least square support vector machine," *Sci. Rep.*, vol. 15, no. 1, p. 12066, Apr. 2025, doi: 10.1038/s41598-025-95621-7.
[20] P. V. Chavan and N. V. Alone, "Optimizing Intrusion Detection with Random Forest: A High-Accuracy Approach Using CIC-IDS 2017," *Int. J. Comput. Appl.*, 2025, doi: 10.5120/ijca2025924816.
[21] V. K. Jewani, P. V. Thawani, P. E. Ajmire, M. Atique, and I. Ben Dhaou, "Hybrid Machine Learning Models for DDoS Detection: A Performance Evaluation on the CICIDS2017 Dataset," 2026, pp. 206–214. doi: 10.1007/978-981-95-7607-4_23..