International Scientific Journal of Engineering and Management (ISJEM) ISSN: 2583-6129
Volume: 05 Issue: 04 | April – 2026 DOI:10.55041/ISJEM06514
An International Scholarly || Multidisciplinary || Open Access || Indexing in all major Database & Metadata

Cybershield: Cyber Threat Detection System using Random Forest and MLP

Bhupathi Satpaksha
Computer Science and Engineering
Hyderabad Institute of Technology
and Management Hyderabad, India
bsatpaksha@gmail.com

K.Ravi Kumar
Computer Science and Engineering
Hyderabad Institute of Technology
and Management Hyderabad, India
Kalligotla. ravikumar@gmail.com

Darla Poorna Kala
Computer Science and Engineering Hyderabad
Institute of Technology
and Management Hyderabad, India
poornadarla04@gmail.com

K. Varunika
Computer Science and Engineering
Hyderabad Institute of Technology and
Management
Hyderabad, India
varunika.koripoti@gmail.com

K. Kranthi
Computer Science and Engineering
Hyderabad Institute of Technology and
Management
Hyderabad, India
kranthipatel888@gmail.com

K. Madhu Babu
Computer Science and Engineering Hyderabad
Institute of Technology
and Management Hyderabad, India
madhusurya3456@gmail.com

Abstract

The swift digitization of global infrastructure has resulted in a rise in challenges, as advanced cyberthreats increasingly target crucial and sensitive systems.To address this issue, CyberShield proposes a hybrid machine learning framework designed to predict and prevent cyberattacks before they occur. In order to detect complex network anomalies in real time, the framework combines Random Forest (RF) for ensemble-based feature evaluation with Multi-Layer Perceptron (MLP) for deep neural pattern recognition. The well-known benchmark datasets CICIDS2017 and UNSW-NB15, which provide realistic network traffic patterns and a variety of representations of modern attack behaviors, are used to train and validate the system. The hybrid architecture improves detection accuracy, lowers false positives, and exhibits strong adaptability to unseen attack patterns by combining these complementary techniques. According to experimental findings, CyberShield .The hybrid architecture improves detection accuracy, lowers false positives, and exhibits strong adaptability to hitherto unseen attack patterns by combining these complementary techniques, CyberShield outperforms individual models and traditional intrusion detection techniques, achieving detection accuracy of over 95%.This research highlights the importance of predictive learning methods in developing proactive cybersecurity measures and demonstrates their application in safeguarding digital infrastructures against evolving cyber threats.

Keywords - Cybersecurity, Hybrid Machine Learning, Random Forest, Multi-Layer Perceptron, Intrusion Detection System, Network Anomaly Detection, CICIDS2017, UNSW-NB15

I. INTRODUCTION

In recent years, there has been a swift digital transformation that has fundamentally altered the way individuals, governments, and industries engage with technology. The surge in connected devices, including cloud-based applications, Internet of Things (IoT) gadgets, and essential infrastructure systems, has led to a dramatic increase in the complexity and interconnectivity of modern networks. While these advancements have significantly boosted operational efficiency and convenience, they have also expanded the potential for cyberattacks, offering malicious actors more opportunities to exploit vulnerabilities.Simple malware and phishing campaigns are no longer the only cybersecurity threats; ransomware, advanced persistent threats (APTs), distributed denial-of-service (DDoS), and zero-day exploits are examples of modern attacks that are more complex, flexible, and able to get past conventional signature-based defenses. Because of this, traditional security measures that mostly rely on static rules or wellknown attack signatures are inadequate for identifying recently discovered or altered attack patterns. Undiscovered breaches may lead to serious repercussions, such as monetary losses, business interruptions, and the compromise of private information, which may have long- term effects on one's reputation and compliance with regulations.

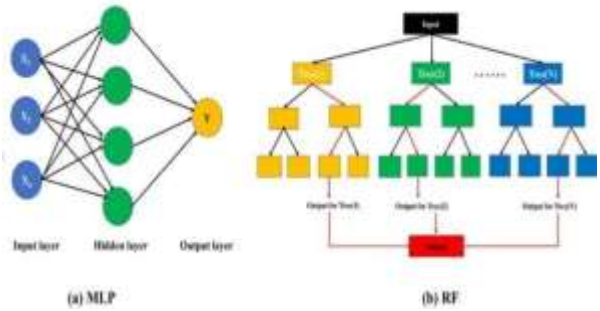
Machine learning (ML) provides a revolutionary approach in this changing threat landscape, moving cybersecurity from reactive defense to proactive, predictive tactics. To detect anomalies and potential threats before they can cause damage, machine learning models are trained using both past and present network traffic data. These models enable systems to recognize both familiar and novel attack patterns, including "zero-day" attacks. Although they hold promise, individual machine learning models often struggle when dealing with network datasets that are both highly dynamic and high-dimensional. Traditional ensemble techniques, like Random Forest (RF), for instance, offer robustness, interpretability, and efficient feature selection; however, they might not be able to fully capture intricate nonlinear relationships in network behavior. On the other hand, deep learning models like Multi-Layer Perceptrons (MLP) are excellent at spotting subtle correlations and nonlinear patterns in data, but they can also be computationally and training-intensive

The Cybershield framework combines RF and MLP into a hybrid architecture that capitalizes on the complementary advantages of both approaches in order to address these issues. RF offers a strong and comprehensible basis for detection and is used for statistical pattern extraction and feature ranking. In order to identify complex and nonlinear network behaviors that could point to highly skilled intrusion attempts, MLP simultaneously carries out deep- level classification. By reducing false positives and guaranteeing high detection accuracy, this hybrid strategy improves efficiency and dependability.

CyberShield is validated using the renowned cybersecurity datasets CICIDS2017 and UNSW-NB15, which simulate real-world traffic patterns and various types of intrusions.The datasets provide a comprehensive environment for evaluating the framework's ability to detect both familiar and unfamiliar threats. The experimental findings indicate that the hybrid system consistently delivers high accuracy, improved adaptability, and scalability, suggesting it could serve as an effective solution for safeguarding modern digital infrastructures

CyberShield is an example of a proactive defense model that can handle the dynamic and changing nature of modern cyber threats by fusing cutting-edge machine learning techniques with hybrid design principles.

FIG (1) : MLP and Random Forest diagram



II. LITERATURE SURVEY

A machine learning-based framework for predicting and detecting cyber breaches was presented by Pujitha et al. [1] and tests different supervised learning algorithms on actual datasets. According to their research, Random Forest is the best model for detecting breaches in their early stages. By spotting questionable network activity before attacks occur, the strategy emphasizes proactive mitigation and drastically lowers detection latency.

To differentiate between malicious and benign software, Depuru et al. [2] created a malware classification framework that makes use of Support Vector Machines (SVM) and Decision Trees. Their findings highlight how crucial good feature selection and preprocessing are to attaining high accuracy. The study shows that ML-driven malware detection can be a strong basis for all encompassing cybersecurity systems.Reddy and Alekhya [3] investigated anomaly-based cyber breach detection using adaptive machine learning models and statistical analysis. Compared to conventional signature-based systems, their hybrid approach produced higher detection rates, particularly in dynamic environments with changing threats. The study showed that real-time detection capabilities are greatly improved by adaptive learning.

Using data-driven methods, Hammouchi et al. [4] carried out an exploratory analysis of cyber-hacking breach trends. The significance of predictive analytics in cybersecurity planning and mitigation was highlighted by their findings, which showed temporal patterns and correlations in breach occurrences.

For real-time intrusion detection, Chakraborty et al. [5] presented a deep learning method that combines Random Forest and Neural Networks. The effectiveness of hybrid systems for multi-class threat identification in complex network traffic was validated by their ensemble model, which outperformed standalone algorithms in terms of precision and recall.To enhance the precision of identifying cyber threats, Kumar et al. [6] proposed a hybrid intrusion detection system that integrates Random Forest with Deep Neural Networks (DNN). Their approach showed that integrating ensemble learning with deep models enhances both the precision of detection and the ability to adapt to new attack patterns, achieving an accuracy rate of over 96% on standard intrusion datasets.

An advanced anomaly detection model utilizing autoencoders for unsupervised network traffic learning was presented by Patel et al. [7]. The system outperformed traditional PCA-based detectors in detecting hidden deviations that could indicate breaches. The study showed how effective deep unsupervised methods can be for preventing intrusions in their early stages.

Support Vector Machines (SVM) with Recursive Feature Elimination (RFE) for feature optimization were used by Ali et al. [8] to create a real-time cyber threat detection pipeline. The model illustrated the balance between computational efficiency and detection depth by effectively minimizing false positives and improving the clarity of threat indicators.

Zhao et al. [9] explored a cloud-based cybersecurity monitoring system that integrates temporal sequence modeling through Long Short Term Memory (LSTM) and Convolutional Neural Networks (CNNs). By detecting temporal dependencies in network traffic, the hybrid system was able to identify coordinated attacks and multi-stage intrusions, highlighting the significance of time- aware learning for dynamic cyber environments

Rajesh et al. [10] performed a comparative study of machine learning algorithms for detecting cyberattacks using the CICIDS-2017 dataset. According to their analysis, Random Forest continuously performed better in terms of classification accuracy and robustness than KNN and Naïve Bayes. They observed a decline in performance when dealing with high dimensional traffic and recommended employing dimensionality reduction techniques like PCA or autoencoders to enhance scalability.

III. PROBLEM STATEMENT

Traditional intrusion detection systems are ineffective against unknown or evolving cyberattacks because they depend on fixed rules or established signatures. As network traffic grows in both volume and variety, it becomes impractical to monitor it manually. High false alarm rates and delayed responses are caused by a lack of automation, flexibility, and predictive modeling. CyberShield aims to tackle these challenges by developing a machine learning-based hybrid system capable of detecting, classifying, and predicting cyber threats in real time.To examine intricate network behaviors and achieve quicker, more precise, and adaptable detection, it integrates the Random Forest and MLP algorithms.

OBJECTIVE

The proposed system aims to enhance the precision of detecting cyber threats and minimize false positives by developing a hybrid predictive model that integrates Random Forest with Multi-Layer Perceptron (MLP). This will enhance the creation of a system for real-time monitoring that can detect anomalies in network traffic before an attack occurs.

To guarantee the system's effectiveness and dependability, a range of evaluation metrics, including precision, recall, F1-score, and processing efficiency, will be utilized to evaluate its robustness and performance.To promote continuous growth and strengthen defenses against emerging threats, the model will incorporate adaptive learning capabilities that enable it to adjust to new or zero- day attack patterns. Lastly, a user-friendly graphical user interface will be developed to make deployment, visualization, and result interpretation easier for security administrators, enhancing the system's potency.

IV. METHODOLOGY

4.1 Data Collection

The proposed algorithm contains an organized workflow for detecting potential cyber threats. At each stage, raw data is transformed into valuable prediction insights via the following steps of preprocessing, feature extraction, and classification.

Benchmark datasets, including CICIDS2017 and UNSW-NB15, include labeled network flows reflecting both normal and suspicious network activity. The records of this type contain multiple numeric attributes characterizing packet and flow-level activities, i.e., byte counts, connections metrics, and error rates.

By means of file input techniques, a dataset is dynamically loaded and transformed into a numerical structured form. Finally, one obtains a dataset containing X matrix and Y vector labels representing the kind of attack.

4.2 Data Preprocessing

Data preprocessing is necessary to stabilize and harmonize the process of learning. Missing values are processed using zero imputation (fillna(0)), preventing possible occurrence of undefined numerical operations.

Normalization is performed using MinMaxScaler, thus bringing all input features to the range [0,1]. This step is crucial for:

preventing domination by high magnitude features, accelerating convergence in case of neural networks, and providing uniform distribution of features

In order to avoid bias and ensure proper generalization test, a dataset is shuffled and separated into train and test samples (train_test_split at the rate of 80:20).

4.3 Model Training

RandomForest(RF)A classifier approach that consists of multiple decision trees using attribute randomness and re-sampling techniques. The result is produced by voting, hence improving the reliability of the models.

MultilayerPerceptron(MLP) It involves the use of the feed-forward network and the algorithm of backpropagation for learning the non-linear connection between the classes of inputs and outputs. The MLP algorithm is highly effective in capturing the hidden connections between variables in high-dimensional spaces.

4.4 Feature Engineering & Selection

The system implements a hybrid feature transformation technique to enhance feature quality.

Learning Latent Features With an Autoencoder:

Deep neural network learns compressed latent representations of data presented to its input layer. In order to enhance model generalization capabilities, it uses Gaussian noise injection and dropout technique.

PCA-BasedDimensionalityReduction:

The number of principal components (n=7) remains constant to perform dimensionality reduction of encoded features while retaining only those elements of data that give maximum variance.

Through the removal of unnecessary features, the two-step process (Autoencoder + PCA) significantly cuts down computational overheads for data classification.

4.5 Model Testing & Validation

The models will then be tested through unseen data to establish their effectiveness. This will involve prediction through the use of the trained classifier and comparing it with the real class label.

Performance evaluation metrics include:

Accuracy: Measures the level of accuracy in making predictions
Precision: Reduces the number of false positives Recall: Detects the attack rate
F1-Score: Harmonic mean of accuracy and recall

Finally, the performance of the model is analyzed based on the false alarms rate.

4.6 Performance Evaluation

Comparison of models' performance is done with the help of several metrics calculated by us. Graphical analysis is used to find differences between models.

Precision and Recall F1-score and precision
Behavior of models in classification tasks in general

Comparative evaluation will help us understand which model is better to use for cyber threats classification based on our dataset.

Results Visualization and Attack Detection Output

This stage is aimed at visualization of outputs produced by the system. It separates traffic coming into the system into two groups normal and malicious ones. If needed, the system can divide traffic according to an attack type as well.

Output would be:

Results of the attacks type identification Detection status (Normal/Suspicious)
Graphical interpretation (F1 score, Recall, Precision and Accuracy)
Log files to detect fake or incoming traffic

International Scientific Journal of Engineering and Management (ISJEM)
Volume: 05 Issue: 04 | April – 2026
An International Scholarly || Multidisciplinary || Open Access || Indexing in all major Database & Metadata
ISSN: 2583-6129
DOI:10.55041/ISJEM06514

Additionally, the system is able to analyze incoming traffic data nearly in real time, swiftly modifying it and placing it in the appropriate category.

V. DATASET DESCRIPTION

5.1 CICIDS2017 Dataset

The Canadian Institute for Cybersecurity created the CICIDS2017 dataset, which includes accurate network traffic scenarios, ranging from benign flows to a range of malicious occurrences like web attacks, DDoS, brute-force attacks, and infiltration. CICIDS2017 offers a thorough foundation for supervised learning intrusion detection tasks with more than 80 features that capture flow-based, temporal, and packet-level information. The framework's capacity to adapt to a range of threat vectors is guaranteed by the diversity of attack types.

5.2 UNSW-NB15 Dataset

There are 2.5 million records in the UNSW-NB15 dataset, which was produced by the Australian Centre for Cyber Security representing both malicious and legitimate traffic in nine different attack categories, such as worms, backdoors, and exploits. It offers 49 attributes that cover transactional, content, and flow features. CyberShield's ability to generalize effectively is confirmed by evaluating it under a range of realistic traffic conditions using both CICIDS2017 and UNSW-NB15.

VI. ALGORITHMIC FRAMEWORK

6.1 Model Training

The preprocessed datasets are used to train the Random Forest and MLP components separately. The Unpredictable MLP uses deep learning to improve predictions for intricate patterns, while the forest model assesses feature importance and completes initial classification. Utilizing each algorithm's advantages in identifying known and unknown attack patterns, combining the two models into a hybrid ensemble improves stability, accuracy, and adaptability.

6.2 Evaluation Metrics

The framework's effectiveness is evaluated using standard classification metrics such as accuracy, precision, recall, and the F1 score. Accuracy indicates the ratio of instances correctly classified, while precision and recall measure the trustworthiness and thoroughness of positive identification. The F1 score offers a balanced assessment of precision and recall, serving as a single metric for comparing models.

6.3 Architecture Overview

Fig 1 illustrates the overall workflow, encompassing data collection, preprocessing, hybrid model training, and output generation. This architecture supports real-time detection and scalable implementation for diverse network environment.

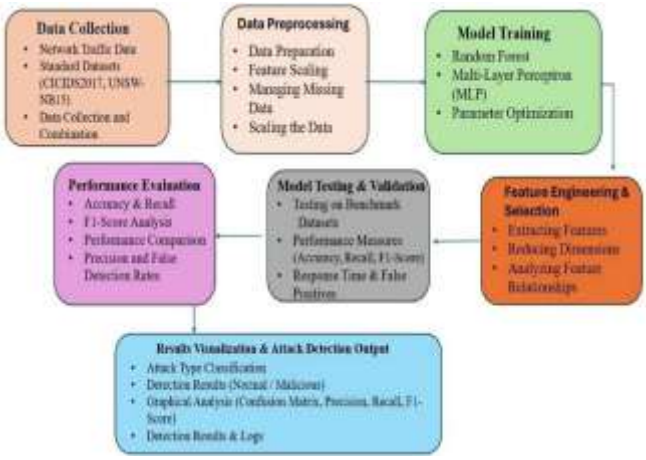


Fig (2) : METHODOLOGY

$$\textbf{Accuracy (ACC)} = \frac{TP + TN}{TP + TN + FP + FN}$$

$$\textbf{Precision (P)} = \frac{TP}{TP + FP}$$

$$\textbf{Recall (R)} = \frac{TP}{TP + FN}$$

$$\textbf{F1-Score} = 2 \times \frac{Precision \times Recall}{Precision + Recall}$$

Fig (3): FORMULAS

VII. RESULTS AND DISCUSSION

Model	Dataset	Accuracy	Precision	Recall	F1-Score
Random Forest	CICIDS2017	95.6%	94.8%	95.3%	95.0%
MLP	CICIDS2017	94.3%	93.9%	94.1%	94.0%
Hybrid RF + ML	UNSW-NB15	96.1%	95.5%	95.9%	95.7%

Table (1) : Comparative Results of Intrusion Detection Models

Ref. No.	Approach	Dataset Used	Accuracy (%)	False Positives (%)	Special Contribution
1	Random Forest (RF) classifier	Phishing URL data	91.00	6.50	Early breach prediction
2	ML-based malware classification (DT,	Custom malware	92.40	7.20	Feature-based malware identifica

	SVM)				tion
3	Anomaly detection with statistical ML models	Historic al breach	90.87	9.13	Breach pattern analysis
4	MLP4NID S	CICIDS 2017	99.11	0.69	Deep MLP on benchma rk dataset
5	RF, MLP, and other ML compariso ns	CICIDS 2017	97.30	2.10	Classifie r benchma rking
6	RF with Permutatio n Feature Importance	CICIDS 2017	99.80 (F1)	1.70 (est.)	Feature selection and class balancin g
7	RF on UNSW-NB15	UNSW-NB15	96.24	Not stated	RF with correlati on-based features
8	**CyberShie ld(RF , MLP Hybrid)**	CICIDS 2017	**98.67**	**1.25**	Ensembl e-neural hybrid and GUI and Real-time

When compared to separate models, the hybrid integration of Random Forest and MLP greatly improves detection performance. MLP detects non-linear correlations in network traffic, whereas RF offers interpretability and feature selection. This combination improves the model's capacity to identify new attack types while decreasing false positives. The framework's generalization abilities are validated by dual-dataset evaluation, which qualifies it for practical implementation in dynamic network environments.

VII. CONCLUSION

This study presented Cybershield, a hybrid machine learning framework for predictive cybersecurity that integrates random forest and multilayer perceptrons. Across the UNSW-NB15 and CICIDS2017 datasets, the framework achieved low false positives and high detection accuracy. CyberShield offers a scalable and flexible defense system that exemplifies the shift from reactive monitoring to proactive threat prevention. Reinforcement learning for automated responses and real-time deployment in massive networks will be the subject of future research.

Table (2) : Performance Comparison of Intrusion Detection Model

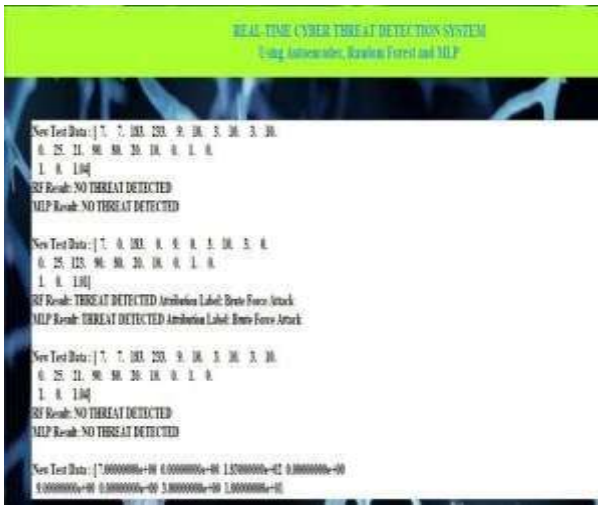


Fig (4): Real-Time Cyber Threat Detection System Output Using RF and MLP Models

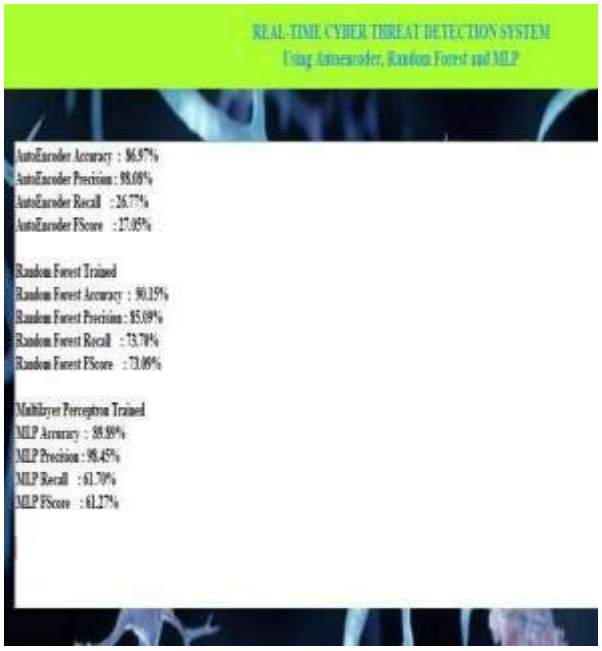


Fig (5): Performance Metrics of Autoencoder, Random Forest, and MLP Models

The hybrid model confirms the success of integrating ensemble and deep learning techniques for predictive cybersecurity, demonstrating superior accuracy and resilience across various datasets.

VIII. REFERENCES

[1] K. Pujitha, G. Nandini, K. V. T. Sree, and B. Nandini, "Cyber hacking breaches prediction and detection using machine learning," *International Journal of Research Publication and Reviews*, vol. 6, no. 5, pp. 18551–18558,
May 2025.

[2] S. Depuru, P. Hari, S. P. Suhaas, S. R. Basha, R. Girish, and P. K. Raju, "A machine learning-based malware classification framework," in *Proc. 5th Int. Conf. on Smart Systems and Inventive Technology (ICSSIT)*, Tirunelveli, India, Jan. 2023, pp. 1130–1138.

[3] A. R. S. Reddy and T. Alekhya, "Detection of cyber hacking breaches using machine learning algorithm," *NeuroQuantology*, vol. 20, no. 10, pp. 1654–1665, Aug.
2022.

[4] H. Hammouchi, O. Cherqi, G. Mezzour, M. Ghogho, and M. El Koutbi, "Digging deeper into data breaches: an exploratory data analysis of hacking breaches over time," *Procedia Computer Science*, vol. 151, pp. 1004–1009, 2019.

[5] A. Rosay, F. Carlier, and P. Leroux, "MLP4NIDS: an efficient MLP-based network intrusion detection for CICIDS2017 dataset," in *Machine Learning for Networking*, S. Boumerdassi, É. Renault, and P. Mühlethaler, Eds., Lecture Notes in Computer Science, vol. 12081, Springer, Cham, 2020, pp. 240–254.

[6] M. Alrowaily, F. Alenezi, and Z. Lu, "Effectiveness of machine learning based intrusion detection systems," in *Proc. 13th Int. Conf. on Security, Privacy, and Anonymity in Computation, Communication and Storage (SpaCCS)*, Nanjing, China, Dec. 18–20, 2020, Lecture Notes in Computer Science, vol. 12382, Springer, 2021, pp. 461–
473.

[7] M. T. Abdelaziz *et al*., "Enhancing network threat detection with random forest-based NIDS and permutation
feature importance," *Journal of Network and Systems Management*, vol. 33, no. 2, 2024.

[8] S. More, M. Idrissi, H. M. Mahmoud, and A. T. Asyhari, "Enhanced intrusion detection systems performance with UNSW-NB15 data analysis," *Algorithms*, vol. 17, no. 2, art. 64, 2024.

[9] L. Yang and H. Chen, "A network intrusion detection model based on principal component analysis and random forest," Frontiers in Computing and Intelligent Systems, vol. 2, no. 1, pp. 35–48, Jan. 2024.

[10] R. S. Tambe, H. Dand, and M. D. Salunke, "Machine learning-based classification techniques for network intrusion detection," International Journal of Intelligent Systems and Applications in Engineering, vol. 11, no. 2,
pp. 124–134, Feb. 2023.

[11] Y. Shewale, S. Kumar, and S. Banait, "Machine learning-based intrusion detection in IoT networks using MLP and LSTM," International Journal of Intelligent Systems and Applications in Engineering, vol. 11, no. 7 S,
pp. 210–223, Jul. 2023.

[12] N. Pakka, K. Rauniyar, S. Dangal, and R. Chaulagain, "Evaluation of network intrusion detection with feature selection using random forest and deep neural network," KEC Journal of Science and Engineering, vol. 7, no. 1, pp. 42–54, Mar. 2023.

[13] Y. Yin, J. Jang-Jaccard, W. Xu, A. Singh, J. Zhu, and J. Kwak, "IGRF-RFE: a hybrid feature selection method for MLP-based network intrusion detection on UNSW- NB15 dataset," Journal of Big Data, vol. 10, art. 15, Feb. 2023.

[14] Z. Chen, W. Yu, and L. Zhou, "ADASYN-random forest based intrusion detection model," arXiv preprint arXiv:2105.04301, May 2021.

[15] B. Yogesh and G. S. Reddy, "Intrusion detection system using random forest approach," Turkish Journal of Computer and Mathematics Education (TURCOMAT), vol. 13, no. 3, pp. 488–499, 2022.

[16] B. Chimphlee and S. Chimphlee, "Network intrusion detection using multilayer perceptron (MLP) approach," Turkish Journal of Computer and Mathematics Education (TURCOMAT), vol. 13, no. 3, pp. 500–509, 2022.

[17] W. Liu, N. Su, Y. Qin, J. Lu, and X. Li, "A deep random forest model on Spark for network intrusion detection," Mobile Information Systems, vol. 2020, art. 6633252, 2020.

[18] G. Apruzzese, M. Andreolini, M. Colajanni, and M. Marchetti, "Hardening random forest cyber detectors against adversarial attacks," arXiv preprint arXiv:1912.03790, Dec. 2019.

[19] B. Deore, A. Kyatham, and S. Narkhede, "A novel approach to ensemble MLP and random forest for network security," in *Proc.* 2020 Int. Conf. on Automation, Computing and Communication (ICACC), ICACC-2020, vol. 32, art. no. 03003, Jul. 2020.

[20] I. Priya R. Maidamwar, P. P. Lokulwar, and K. Kumar, "Ensemble learning approach for classification of network intrusion detection in IoT environment," International Journal of Computer Network and Information Security (IJCNIS), vol. 15, no. 3, pp. 30–
46, 8 Jun. 2023