

---

## **Dual-Task Learning Framework for Next-Gen Network Security and Performance Optimization Using Machine Learning**

Latta Akash<sup>1</sup>, Kiran Gadapaka<sup>2\*</sup>, Jampala Chandrika<sup>1</sup>, Peyyala Anand<sup>1</sup>, Malyala Sritej<sup>1</sup>, Jannu Sathvik<sup>1</sup>

<sup>1</sup>UG Student, <sup>2</sup>Assistant Professor, <sup>1,2</sup>Department of Computer Science and Engineering (AI&ML)

<sup>1,2</sup>Vaagdevi Engineering College, Bollikunta, Warangal, 506005, Telangana, India

\*Correspondence: Kiran Gadapaka(gadapakakiran@gmail.com)

### **Abstract**

The evolution of 6G networks introduces new challenges for ensuring secure, adaptive, and high-performance communication systems. Traditional network monitoring and intrusion detection approaches, which rely on static rules and manual inspection, are inadequate for handling the complexity and scale of modern network environments. These methods often fail to detect sophisticated cyber threats and are limited in optimizing network performance under dynamic conditions. To address these challenges, this study proposes a machine learning–driven framework based on Classification and Regression Tree (CART) principles for dual-purpose analysis, including attack classification and throughput prediction. The framework integrates multiple machine learning models such as Support Vector Machine (SVM), k-Nearest Neighbors (KNN), and a novel Tree-based Adaptive Optimization (TAO) ensemble model inspired by Random Forest (RF). These models are designed to identify malicious traffic patterns while simultaneously predicting network throughput. To enhance performance and generalization, the system incorporates preprocessing techniques such as Label Encoding, feature standardization, and class balancing using the Synthetic Minority Over-sampling Technique (SMOTE). This ensures robustness across diverse and complex network scenarios. Experimental results demonstrate that the TAO ensemble model outperforms individual models in both classification accuracy and regression reliability. Additionally, the framework is deployed through a web-based interface using Flask, enabling real-time monitoring and user interaction. The proposed system offers a scalable, intelligent, and efficient solution for strengthening cybersecurity and optimizing performance in next-generation communication networks.

**Keywords:** 6G networks, classification, cybersecurity, machine learning, regression, throughput prediction

### **1. Introduction**

Network management and monitoring rely on data collection protocols such as SNMP (Simple Network Management Protocol), NetFlow, IPFIX, and NETCONF (Network Configuration Protocol) [1]. Organizations develop customized tools based on these protocols to gain insights into network conditions and respond effectively to faults and performance issues. With the rapid growth in demand for network services, the requirements for network components and monitoring systems have significantly increased as shown in Fig 1. Consequently, network devices now generate vast volumes of data, including control information, statistical metrics, and user traffic, at an accelerating rate [2]. Analysing such large-scale and complex data using manual or human-assisted approaches requires substantial expertise, time, and operational effort. As network environments continue to expand in scale and complexity, these methods become less efficient and harder to sustain. This has led to a growing shift toward automated and intelligent

solutions capable of processing high-dimensional data and enabling faster, more accurate decision-making in modern network infrastructures.

A key approach enabling advanced management and monitoring of telecommunication networks is the concept of Software-Defined Networking (SDN) [3], which introduces a centralized and programmable architecture for network control. This paradigm simplifies network configuration and enhances flexibility; however, it still requires continuous monitoring and efficient anomaly detection mechanisms. Centralized data collection offers significant advantages by enabling better coordination, aggregation, and correlation of network information. Despite these benefits, processing and analyzing such large-scale datasets presents substantial computational challenges, often referred to as “big data analytics” [4]



Fig 1: 6G Smart Network Monitoring with Multi-Layer Data Capture.

To address these challenges, various tools and platforms have been developed to support the implementation, deployment, and utilization of large datasets. One such example is Platform for Network Data Analytics (PNDA), an open-source solution that enables efficient data collection, storage, and real-time analysis. Given the vast volume of network data, integrating Artificial Intelligence and Machine Learning techniques becomes essential for extracting meaningful insights and improving decision-making processes. Continuous data collection plays a crucial role in effective network monitoring; however, it can increase device workload and impact network throughput, particularly within transmission and management networks. To mitigate these issues, adaptive data collection strategies can be employed, where polling intervals are dynamically adjusted based on network conditions. For instance, increasing the frequency of data collection upon detecting anomalies can enhance responsiveness while maintaining overall system efficiency [5].

## 2. Literature Survey

Saeed, et al. [6], developed a novel anomaly detection system for 6G networks (AD6GNs) based on ensemble learning (EL) for communication networks. The first stage in the EL-ADCN process was pre-

processing. The second stage was the feature selection approach. Rekkas, et al. [7] presented a summary of ML methods, as well as an up-to-date review of ML approaches in 6G wireless communication systems. These methods included supervised, unsupervised and reinforcement techniques. Additionally, they discussed open issues in the field of ML for 6G networks and wireless communications in general, as well as some potential future trends to motivate further research into this area.

Ismail, et al. [8] explored the application of an adaptive optimized machine learning-based framework to improve intrusion detection system (IDS) performance in wireless network access scenarios. The framework used involved developing a lightweight model based on a convolution neural network with 11 layers, referred to as CSO-2D-CNN, which demonstrated fast learning rates and excellent generalization capabilities. Zahid, et al. [9] introduced unprecedented challenges in security, particularly concerning the threat of unconventional drones and swarm attacks. To deal with threats, drones needed to be classified by intercepting their Radio Frequency (RF) signals. With the arrival of Sixth Generation (6G) networks, it was required to develop sophisticated methods to properly categorize drone signals.

Puspitasari, et al. [10] examined the potential of ML algorithms and their derivatives in optimizing emerging technologies to align with the visions and requirements of the 6G network. It was crucial in ushering in a new era of communication marked by substantial advancements and required grand improvement. Rzym, et al. [11] introduced a meticulously crafted solution designed explicitly for 6G software-defined networks (SDNs). The approach employed deep neural networks for anomaly detection within network traffic, contributing to a more robust security framework. Kaur, et al. [12] the progress introduced significant security risks, as technologies like O-RAN, terahertz communication, and native AI posed threats such as eavesdropping, supply chain vulnerabilities, model poisoning, and adversarial attacks. The increased exposure of sensitive data in 6G applications further intensified these challenges.

Okere, et al. [13] presented 6G-enabling technologies including digital twin (DT), intelligent reflecting surface (IRS), visible light communication (VLC), quantum computing (QC), blockchain, unmanned aerial vehicles (UAVs), and non-orthogonal multiple access (NOMA), among others. Optimal network performance required that machine learning (ML) techniques be integrated over the 6G wireless network to provide solutions to highly complex networking problems, massive users, high overhead, and computational complexity. Damaševičius, et al. [14] proposed an ensemble classification-based methodology for malware detection. The first-stage classification was performed by a stacked ensemble of dense (fully connected) and convolutional neural networks (CNN), while the final stage classification was performed by a meta-learner. For a meta-learner, they explored and compared 14 classifiers.

Mahmoud, et al. [15] presented an advanced framework for securing 6G communication by integrating deep learning and physical layer security (PLS). The proposed model incorporated multi-stage detection mechanisms to enhance security against various attacks on the 6G air interface. Deep neural networks and a hybrid model were employed for sequential learning to improve classification accuracy and handle complex data patterns. Nassreddine, et al. [16] recent advancements across various sectors had resulted in a significant increase in the utilization of smart gadgets. This augmentation had resulted in an expansion of the network and the devices linked to it.

Rekha Gangula et al. [17] proposed an intrusion detection approach using Firefly Optimization Algorithm with an ensemble classification model. The optimization algorithm selected optimal feature subsets. The

ensemble classifier enhanced detection performance in complex network environments. Rekha Gangula et al. [18] proposed an intelligent intrusion prevention framework for network applications. The system integrated detection and prevention mechanisms within a unified architecture. The framework reduced attack propagation and improved response efficiency. Rekha Gangula et al. [19] proposed a deep learning and optimization-based intrusion detection method for IoT systems. The model combined deep neural architectures with optimization techniques. The approach improved detection accuracy and minimized computational overhead.

**Research Gap:** Research presents a hierarchical exploration of 6G networks, poised at the forefront of the next revolution in wireless technology. This study delves into the technological advancements that underpin the need for 6G, examining its key features, benefits, and key enabling technologies. We dissect the intricacies of cutting-edge innovations like terahertz communication, ultra-massive MIMO, artificial intelligence (AI), ML, quantum communication, and reconfigurable intelligent surfaces. Through a meticulous analysis, we evaluate the strengths, weaknesses, and state-of-the-art research in these areas, offering a wider view of the current progress and potential applications of 6G networks.

### 3. Proposed Methodology

This research focuses on developing an integrated analytical framework capable of performing both security threat classification and network performance prediction within a unified system. Instead of treating intrusion detection and performance evaluation as separate tasks, the proposed approach combines classification and regression techniques to establish a relationship between network attacks and their impact on throughput. This enables simultaneous analysis of security conditions and performance behavior in modern communication networks. The framework uses ML models such as SVM, KNN, and a custom TAO Tree ensemble model inspired by RF. These models are designed to classify different types of network attacks while predicting throughput based on network conditions. The system is implemented as a web-based platform using Flask, providing an interactive and user-friendly interface. It offers a complete pipeline that includes dataset upload, preprocessing, Exploratory Data Analysis (EDA), model training, performance evaluation, and real-time prediction as demonstrate in Fig. 2. This integrated design ensures efficient handling of large-scale network data and supports informed decision-making for improving both network security and performance in next-generation environments.

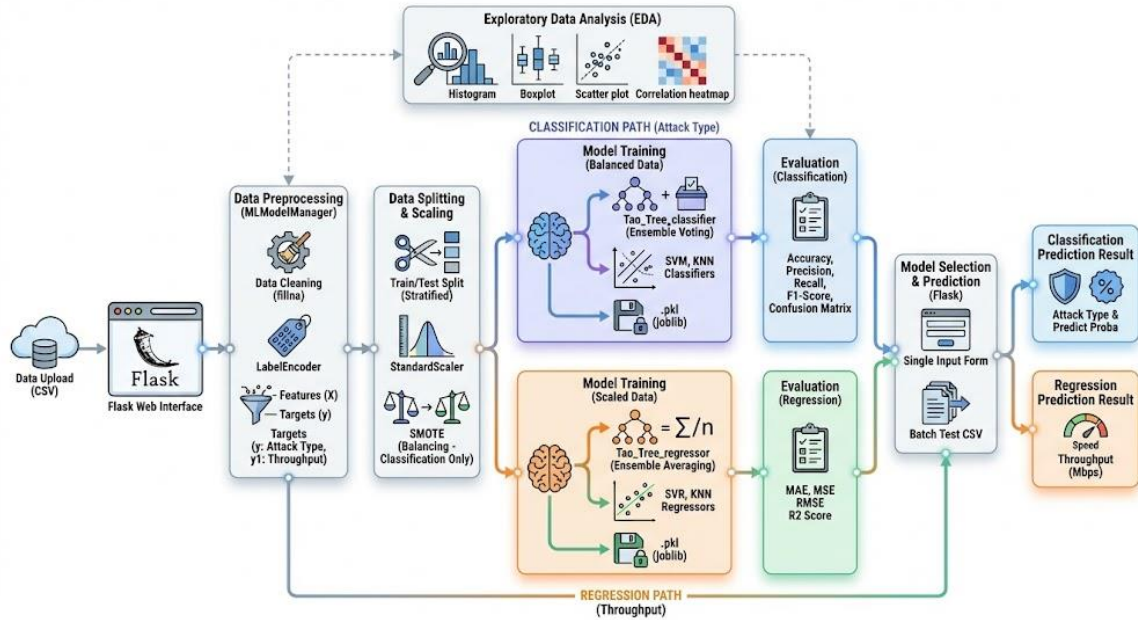


Figure 2: Proposed system architecture

**Data Acquisition and Upload:** The process begins with the user uploading a network traffic dataset, typically in a CSV format, through the web application's main interface. The system is configured to validate the file type to ensure it is a .csv file, and it uses secure filename handling to prevent malicious uploads. This step acts as the primary entry point for all subsequent analysis and model training activities. Upon successful upload, the system saves the dataset to a dedicated uploads directory and stores its path and basic information, like its shape, in the user's session for continuous access throughout the workflow.

**Data Preprocessing and Splitting:** In the background, the ML Model Manager performs a series of essential preprocessing tasks to prepare the raw data for machine learning. First, it identifies and drops any irrelevant columns, such as "Unnamed" columns that often appear in datasets. Next, it handles categorical features, like different network services or protocols, by converting them into numerical representations using a Label Encoder. Missing values are filled using the mean of their respective columns. The prepared data is then split into two distinct parts: a training set (80%) and a testing set (20%). The data is also split into two target variables: a categorical one for Intrusion Detection (Attack Type) and a numerical one for Performance Prediction (Throughput). Finally, the system addresses class imbalance in the intrusion detection dataset by applying SMOTE on the training data. This ensures that the machine learning models do not become biased toward the majority class (normal traffic) and can effectively detect rare attack instances.

**EDA:** Once the data is loaded, the system automatically redirects the user to the EDA page. This crucial step provides a quick and informative overview of the dataset's characteristics. The system's ML Model Manager generates several key visualizations, including histograms to show the distribution of attack types and network throughput, box plots to analyze the relationship between categorical features (like signal strength) and numerical ones (like throughput), and scatter plots to visualize relationships between continuous variables. A correlation heatmap is also generated to identify which features are most strongly

related to the target variables (attack type and throughput). These plots are then displayed on the web page, allowing the user to gain an immediate understanding of the data's structure, potential issues like class imbalance, and key relationships before proceeding to model training.

**Model Training and Selection:** The user can choose from a variety of machine learning models to train, including SVM, KNN, and a custom ensemble model called TAO Tree. Upon selection, the ML Model Manager trains a classifier for intrusion detection and a regressor for throughput prediction for each chosen model. To enhance efficiency and reusability, the trained models are saved as .pkl files using joblib. This allows the system to load pre-trained models in subsequent sessions, eliminating the need for retraining and significantly reducing processing time.

**Performance Evaluation:** After training, the system evaluates the performance of each model on the held-out test data. For the classification models, it calculates key metrics such as accuracy, precision, recall, and F1-score. It also generates a confusion matrix, which visually represents the number of correct and incorrect predictions for each attack type. For the regression models, it calculates metrics like Mean Absolute Error (MAE), Mean Squared Error (MSE), and R-squared ( $R^2$ ) to assess the accuracy of throughput predictions. These detailed performance results, including the visual plots, are then displayed on dedicated pages, allowing the user to compare the effectiveness of different models and choose the best one for their specific needs.

**Prediction Interface:** The final step provides two methods for making predictions on new data. The single prediction interface allows users to manually input values for each network feature via a form. The system preprocesses this single data point and feeds it to the selected trained models to predict the attack type and network throughput. The batch prediction interface enables users to upload a new CSV file containing multiple instances of network data. The system automatically preprocesses the entire file and returns a table with the predicted attack type and throughput for each instance, providing a powerful tool for large-scale network security and performance analysis.

### 3.2 TAO Tree CART Model

The TAO Tree model is a proposed ML-based ensemble approach designed to perform both classification and regression tasks within a unified framework. It is inspired by RF and built upon CART principles, enabling simultaneous prediction of network attack types and throughput values using the same input features. Unlike individual models, the TAO model combines multiple decision trees trained on different subsets of data and features. This ensemble strategy improves prediction accuracy, reduces overfitting, and enhances generalization across complex network conditions as shown in Fig. 3. The model is specifically designed to handle non-linear relationships present in network traffic data, making it suitable for next-generation network environments.

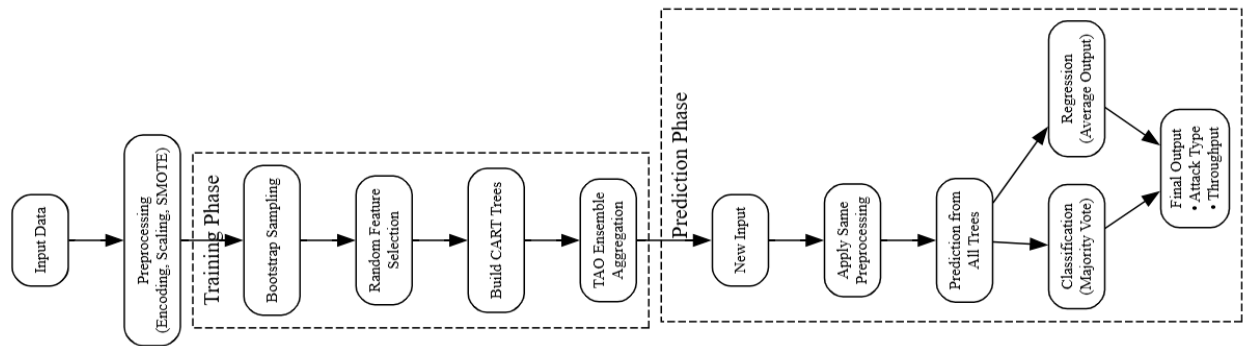


Fig. 3: Internal working of TAO Tree model.

**Data Sampling (Bootstrap Aggregation):** During training, multiple subsets of the dataset are created using random sampling with replacement. Each subset is used to train an individual decision tree. This process ensures diversity among trees and improves robustness.

**Feature Subset Selection:** For each tree, a random subset of features is selected instead of using all features. This reduces correlation between trees and allows the model to capture different patterns in the data.

**Tree Construction (CART Logic):** Each decision tree is built using CART principles:

- For classification: splits are chosen to reduce impurity (e.g., Gini index)
- For regression: splits are chosen to minimize prediction error (e.g., MSE)

The tree recursively splits the data until stopping criteria are met.

**Ensemble Learning (TAO Strategy):** The TAO model combines predictions from all trees:

- Classification: majority voting across trees determines attack type
- Regression: average of outputs from all trees predicts throughput

This aggregation improves stability and accuracy compared to single-tree models.

**Model Optimization:** By combining bootstrap sampling and feature randomness, the TAO model reduces overfitting and improves performance on unseen data. It balances bias and variance effectively.

**Prediction Process:** During inference, new input data is pre-processed and passed through all trained trees. Each tree independently produces classification and regression outputs, which are then aggregated to generate final predictions.

**Output Prediction:** For each input sample, the model produces:

- Predicted attack type (classification output)
- Predicted throughput value in Mbps (regression output)

These outputs are generated simultaneously, providing a comprehensive understanding of network security and performance.

#### 4. Results and Discussion

Fig. 4 shows EDA of network data. countplot of attack types (toprow-left). Boxplot of latency distribution per attack type (toprow - middle). Violin plot of signal strength by attack type (toprow-right). Boxplot of throughput by signal strength (bottom row - left). Scatter plot of throughput vs latency (bottom row - middle). Correlation heatmap with throughput (bottom row - right).

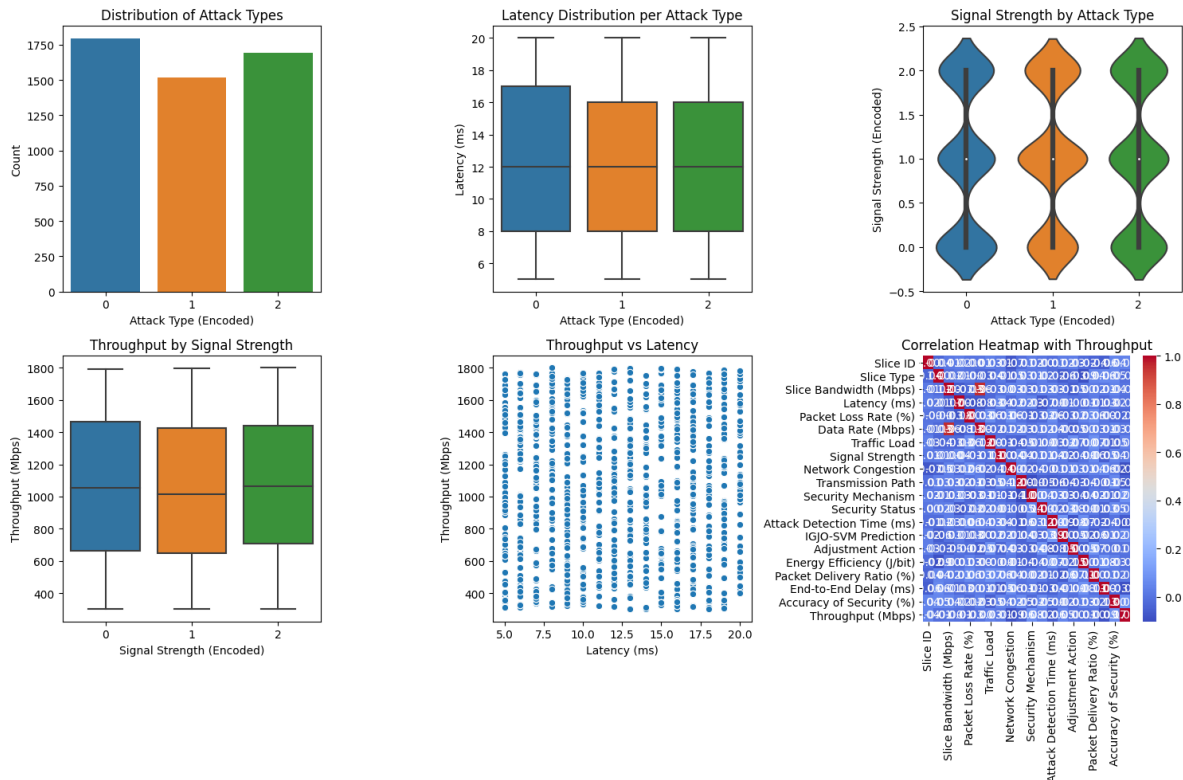


Fig. 4: EDA of Network Dataset with Various Plots.

The confusion matrices in Fig. 5 illustrate the performance of four models (a) SVM, (b) KNN, and (c) TAO Tree in classifying DDoS, Eavesdropping, and Spoofing. The SVM matrix shows poor differentiation, with high misclassifications (e.g., 198 Spoofing predicted as Eavesdropping). The KNN matrix performs better, with 264 correct Eavesdropping predictions but 316 DDoS misclassifications. The TAO Tree matrix excels, with 299 correct Eavesdropping, 335 correct Spoofing, and 356 correct DDoS predictions, showing minimal errors.

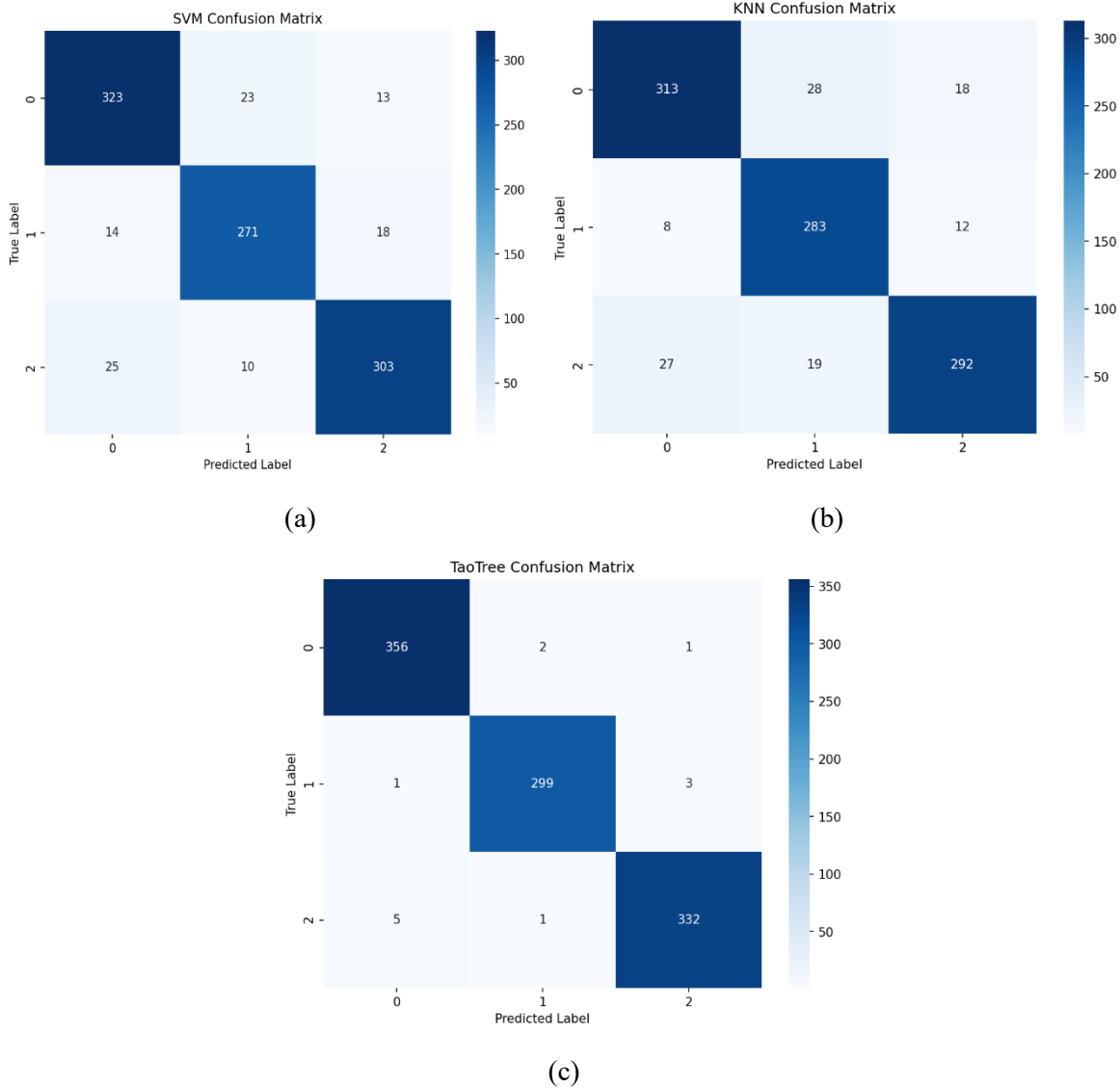


Fig. 5 Confusion matrix obtained using (a) SVM model. (b) KNN model. (c) Proposed TAO Tree Classifier

Table 1 presents performance metrics for four different regression models evaluated on a dataset. The metrics included are Mean Absolute Error (MAE), Mean Squared Error (MSE), Root Mean Squared Error (RMSE), and R2 Score. The models listed are:

- SVR Model:** MAE is 370.851, MSE is 18.8609, RMSE is 431.828, and R2 Score is -0.009, indicating poor predictive performance with a negative R2 score suggesting the model is not better than a mean baseline.
- KNN Regressor:** MAE is 133.998, MSE is 49133.987, RMSE is 221.662, and R2 Score is 0.739, indicating a significant improvement in predictive accuracy and a substantial positive R2 score.

- **TAO Tree Regressor:** MAE is 29.042, MSE is 5787.744, RMSE is 76.007, and R2 Score is 0.969, demonstrating the best performance among the models with the lowest errors and a very high R2 score, suggesting excellent predictive capability.

Table 2 presents performance metrics for four different classification algorithms evaluated on a dataset. The metrics included are Accuracy, Precision, Recall, and F1-Score.

- **SVC Model:** Accuracy is 89.7%, Precision is 89.7%, Recall is 89.7%, and F1-Score is 89.7%, indicating relatively average performance across all metrics, suggesting limited ability to correctly classify instances.
- **KNN Classifier:** Accuracy is 88.9%, Precision is 88.907%, Recall is 88.857%, and F1-Score is 88.876%, demonstrating strong performance with high values across all metrics, indicating effective classification.
- **TAO Tree Classifier:** Accuracy is 98.7%, Precision is 98.7%, Recall is 98.7%, and F1-Score is 98.7%, showcasing the best performance among the algorithms with near-perfect scores, suggesting excellent classification capability.

Table 1: Performance Evaluation of Various Regressor Models.

Model/Metric	MAE	MSE	RMSE	R2-score
SVR Model	370.871	18.8609	431.828	0.009
KNN Model	133.998	49133.987	221.662	0.739
TAO Tree Regressor	0.0297	5787.744	76.007	0.969

Table 2: Performance evaluation of Various Classification Models.

Algorithm	Accuracy	Precision	Recall	F1-Score
SVC Model	89.7	89.7	89.7	89.7
KNN Model	88.9	88.907	88.857	88.876
TAO Tree Classifier	98.7	98.7	98.7	98.7

## 5. Conclusion

In this study, presents a web-based ML framework for intelligent network monitoring that performs both attack classification and throughput prediction simultaneously. The system integrates multiple ML models,

including SVM, KNN, and the proposed TAO Tree model, to analyze network data efficiently and provide meaningful insights into both security and performance aspects. The implementation covers the complete pipeline, including dataset upload, preprocessing, EDA, model training, evaluation, and real-time prediction through a user-friendly Flask interface. Techniques such as label encoding, standardization, and SMOTE are applied to improve data quality and model performance. The dual-task approach enables the system to classify different types of network attacks while predicting throughput, offering a comprehensive understanding of network behavior. Among the implemented models, the TAO model demonstrates improved performance due to its ensemble nature and ability to handle non-linear relationships effectively. The system successfully achieves its objective of providing a scalable, efficient, and intelligent solution for next-generation network environments. So, the work highlights the importance of ML-based approaches in enhancing both network security and performance management.

### References

- [1]. Fernandes, G.; Rodrigues, J.J.; Carvalho, L.F.; Al-Muhtadi, J.F.; Proença, M.L. A Comprehensive Survey on Network Anomaly Detection. *Telecommun. Syst.* **2019**, *70*, 447–489.
- [2]. Cisco Annual Internet Report (2018–2023) White Paper; Technical Report; Cisco: San Jose, CA, USA, 2023.
- [3]. 2022 Global Networking Trends Report; Technical Report; Cisco: San Jose, CA, USA, 2022.
- [4]. 2023 Global Internet Phenomena Report; Technical Report, Sandvine Intelligent Broadband Networks; Sandvine Inc.: Waterloo, ON, Canada, 2022.
- [5]. Ericsson Mobility Report; Technical Report; Ericsson: Stockholm, Sweden, 2022.
- [6]. Saeed, M.M.; Saeed, R.A.; Abdelhaq, M.; Alsaqour, R.; Hasan, M.K.; Mokhtar, R.A. Anomaly Detection in 6G Networks Using Machine Learning Methods. *Electronics* **2023**, *12*, 3300. <https://doi.org/10.3390/electronics12153300>.
- [7]. Rekkas, V.P.; Sotiroudis, S.; Sarigiannidis, P.; Wan, S.; Karagiannidis, G.K.; Goudos, S.K. Machine Learning in Beyond 5G/6G Networks—State-of-the-Art and Future Trends. *Electronics* **2021**, *10*, 2786. <https://doi.org/10.3390/electronics10222786>.
- [8]. Ismail, W.N. A Novel Metaheuristic-Based Methodology for Attack Detection in Wireless Communication Networks. *Mathematics* **2025**, *13*, 1736. <https://doi.org/10.3390/math13111736>.
- [9]. Zahid, M.U.; Nisar, M.D.; Fazil, A.; Ryu, J.; Shah, M.H. Composite Ensemble Learning Framework for Passive Drone Radio Frequency Fingerprinting in Sixth-Generation Networks. *Sensors* **2024**, *24*, 5618. <https://doi.org/10.3390/s24175618>.
- [10]. Puspitasari, A.A.; An, T.T.; Alsharif, M.H.; Lee, B.M. Emerging Technologies for 6G Communication Networks: Machine Learning Approaches. *Sensors* **2023**, *23*, 7709. <https://doi.org/10.3390/s23187709>.
- [11]. Rzym, G.; Masny, A.; Chołda, P. Dynamic Telemetry and Deep Neural Networks for Anomaly Detection in 6G Software-Defined Networks. *Electronics* **2024**, *13*, 382. <https://doi.org/10.3390/electronics13020382>.

- [12]. Patel, S., & Patyrykin, K. (2025). Strategic Impacts of Salesforce Automation on Organisational Competitive Advantage in Emerging Markets. *Journal of Posthumanism*, 5(12), 357–372. <https://doi.org/10.63332/joph.v5i12.3782>
- [13]. Okere, E.E.; Balyan, V. Sixth Generation Enabling Technologies and Machine Learning Intersection: A Performance Optimization Perspective. *Future Internet* **2025**, 17, 50. <https://doi.org/10.3390/fi17020050>.
- [14]. Kalae, U. K. (2025). Optimizing cost-effective cloud data pipeline orchestration across multiple cloud providers. *Journal of Information Systems Engineering and Management*, 10(63s), e726–e741.
- [15]. Damaševičius, R.; Venčkauskas, A.; Toldinas, J.; Grigaliūnas, Š. Ensemble-Based Classification Using Neural Networks and Machine Learning Models for Windows PE Malware Detection. *Electronics* **2021**, 10, 485. <https://doi.org/10.3390/electronics10040485>.
- [16]. Poojari, R. Frameworks for Data Management and Lineage in Large-Scale Healthcare Data Systems.
- [17]. Mahmoud, H.; Ismail, T.; Baiyekusi, T.; Idrissi, M. Advanced Security Framework for 6G Networks: Integrating Deep Learning and Physical Layer Security. *Network* **2024**, 4, 453-467. <https://doi.org/10.3390/network4040023>.
- [18]. Purmani, S. S. R. (2024). Aligning IT investment decisions with overall business strategy from an enterprise program management perspective, focusing on the integration of IT leadership in strategic decision-making processes. *International Journal of Communication Networks and Information Security*, 16(5), 1213–1219.
- [19]. Rekha Gangula, Murali Mohan Vutukuru, M. Ranjeeth Kumar. Intrusion Attack Detection Using Firefly Optimization Algorithm and Ensemble Classification Model. *Wireless Pers Commun* 132, 1899–1916 (2023). <https://doi.org/10.1007/s11277-023-10687-8>
- [20]. Rekha Gangula, Sreenivas Pratapagiri, Sridhara Murthy Bejugama, Sudharshan Ray, Gayatri Nandam, Swapna Saturi, "A Novel Intelligent Intrusion Prevention Framework for Network Applications". *Wireless Pers Commun* 131, 1833–1858(2023).
- [21]. Rekha Gangula, Murali Mohan Vutukuru, "A Deep Learning and Optimization Method for Detecting Network Intrusion in IOT," 2022 Second International Conference on Advanced Technologies in Intelligent Control, Environment, Computing & Communication Engineering (ICATIECE).