Vol 6, Issue 6, June 2026 www.ijesti.com E-ISSN: 2582-9734
International Journal of Engineering, Science, Technology and Innovation (IJESTI)

Intelligent Machine Learning Approach for Effective Network Traffic Management

Wasim Ahmad Sheikh
M. Tech. in Computer Science Engineering, CBS Group of Institutions, Jhajjar, Haryana.

Amreesh Kumar Yadav
A.P CSE Department, CBS Group of Institutions, Jhajjar, Haryana.

ABSTRACT

This study focuses on machine learning-based network traffic classification for efficient network management. The main aim is to identify and classify different types of traffic such as browsing, streaming, file transfer, VoIP, gaming, and suspicious traffic. Network traffic data were processed through feature extraction, data cleaning, and model training. Machine learning algorithms such as Naïve Bayes, Decision Tree, Support Vector Machine, Random Forest, and Artificial Neural Network were applied and compared. The result showed that Random Forest achieved the highest accuracy, proving its effectiveness in traffic classification, bandwidth management, congestion control, and network security improvement.

Keywords: *Machine Learning, Network Traffic Classification, Network Management, Random Forest.*

I. INTRODUCTION

Machine Learning-Based Network Traffic Classification for Efficient Network Management has become an important area of study in modern computer networks because the volume, variety, and complexity of internet traffic are increasing rapidly with the growth of cloud computing, online video streaming, social media, Internet of Things devices, mobile applications, online gaming, e-commerce, and enterprise communication systems. In earlier network environments, traffic classification was mainly performed using traditional methods such as port-based classification, payload inspection, and rule-based filtering. However, these methods have become less effective because many applications use dynamic port numbers, encrypted communication, tunneling techniques, and frequently changing traffic patterns. As a result, network administrators face difficulties in accurately identifying different types of traffic and managing network resources efficiently. Machine learning offers a powerful solution to this problem by enabling systems to automatically learn patterns from network data and classify traffic based on features such as packet size, flow duration, protocol type, source and destination addresses, number of packets, inter-arrival time, byte count, and statistical behavior of network flows. Instead of depending only on fixed rules, machine learning models can analyze large datasets, identify hidden relationships, and improve classification accuracy over time. Network traffic classification plays a vital role in efficient network management because it helps in bandwidth allocation, congestion control, quality of service management, anomaly detection, intrusion prevention, and policy enforcement. For example, real-time applications such as video conferencing and voice over IP require low latency and stable bandwidth, while file downloads and background updates can tolerate delay. By classifying traffic correctly, network managers can prioritize important applications, reduce unnecessary congestion, and provide better service quality to users. Machine learning algorithms such as Decision Tree, Random Forest, Support Vector Machine, Naïve Bayes, K-Nearest Neighbour, Artificial Neural Network, and deep learning models can be applied to classify traffic into different categories such as normal traffic, streaming traffic, browsing traffic, peer-to-peer traffic, email traffic, gaming traffic, and malicious traffic. These models are trained using historical network traffic datasets and then tested to evaluate their ability to classify new and unseen traffic. The performance of such models is commonly measured using accuracy, precision, recall, F1-

score, and confusion matrix. Among different algorithms, ensemble and deep learning methods often provide better results because they can handle complex and nonlinear traffic patterns. Another important advantage of machine learning-based traffic classification is its usefulness in cybersecurity. Modern networks are continuously exposed to threats such as malware, denial-of-service attacks, phishing traffic, botnets, and unauthorized access attempts. By learning the difference between normal and abnormal traffic behavior, machine learning models can support early detection of suspicious activities and help administrators take preventive action before serious damage occurs. This makes traffic classification not only useful for performance improvement but also for network security and reliability. In enterprise networks, educational institutions, banking systems, smart cities, and data centers, efficient traffic management is essential for smooth communication and uninterrupted digital services. The increasing use of encrypted traffic has also created new challenges, as traditional deep packet inspection may not be suitable due to privacy and security concerns. Machine learning can address this issue by classifying encrypted traffic using flow-based and statistical features without directly examining the content of packets. However, the successful implementation of machine learning-based traffic classification depends on proper data collection, feature selection, preprocessing, model training, and continuous updating of the classification system. Challenges such as imbalanced datasets, high computational cost, changing traffic behavior, and real-time processing requirements must also be considered. Despite these challenges, machine learning provides a flexible, scalable, and intelligent approach for modern network traffic classification. Therefore, the study of machine learning-based network traffic classification is highly relevant for efficient network management, as it supports better resource utilization, improved quality of service, stronger security monitoring, and smarter decision-making in complex and dynamic network environments.

II. RESEARCH BACKGROUND

Martínez Hernández et al. (2026) examined the growing significance of IT security in the context of increased public access to computers and expanding digital connectivity, which had simultaneously enabled cybercriminals to exploit the anonymity and privacy of the Internet for unlawful activities. The authors observed that artificial intelligence had emerged as one of the most innovative solutions for safeguarding systems and networks, although these same technologies had also become attractive targets for adversaries attempting to compromise organisational security. In their study, attacks directed at machine learning algorithms employed for the classification of messaging application traffic were analysed through the use of Generative Adversarial Networks (GANs). Three specific algorithms were evaluated and their performances under adversarial conditions were compared. The findings indicated that all the analysed algorithms exhibited a certain degree of vulnerability to malicious manipulation, thereby underscoring the necessity of strengthening defence mechanisms to enhance the resilience and security of AI-based traffic classification systems.

Pulido et al. (2026) proposed a novel synthetic data generation method for large-scale Wi-Fi deployments based on first-order auto-regressive noise statistics. The study highlighted that synthetic data had served as an effective and privacy-preserving approach for augmenting and enriching datasets in artificial intelligence (AI) and machine learning (ML), while also reducing storage complexity and cost. It was reported that the proposed approach required only minimal real data to generate statistically rich traffic patterns that closely resembled real Access Point (AP) behavior. The experimental findings indicated that ML models trained on the generated synthetic data had achieved Mean Absolute Error (MAE) values within 10–15 of those obtained from models trained on real data for the same APs, despite using substantially less training data. Furthermore, when model generalization was emphasized, the synthetic-

data-trained models were found to have improved prediction accuracy by up to 50% over real-data-trained baselines due to the increased variability and diversity of the generated traces. Overall, the study demonstrated that the proposed framework had offered a scalable, efficient, and real-time solution for practical Wi-Fi traffic forecasting in modern wireless networks.

Lohiya and Bamnote (2025) investigated the growing complexity and volume of internet traffic, which had prompted researchers to explore machine learning as a potential solution for effective traffic classification. They argued that integrating intelligence into network processes could enhance network management and optimization. The study examined four supervised learning techniques—Support Vector Machine (SVM), Random Forest (RF), K-Nearest Neighbors (KNN), and Decision Tree (DT)—to predict network traffic categorization. Through a comparative analysis, the authors evaluated the performance of these algorithms in terms of accuracy, precision, recall, and computational efficiency using a standardized dataset. Their findings suggested that although each algorithm exhibited specific strengths and limitations, Random Forest generally outperformed the others across most metrics. The study was concluded to provide meaningful insights into the applicability of these techniques for real-time internet traffic management and future network optimization strategies.

Chen et al. (2025) investigated the application of uncrewed aerial vehicles (UAVs) in real-time data collection, processing, and transmission within the rapidly advancing field of aerial computing and the expanding fifth-generation (5G) networks. They highlighted that the heterogeneity of UAV-generated traffic in different mission scenarios posed substantial challenges for traffic classification. To address this, they proposed a novel traffic classification model based on a spatial attention-enhanced convolutional neural network (SAE-CNN), which was reported to improve both classification accuracy and latency, particularly for 5G services such as enhanced mobile broadband (eMBB), ultrareliable low-latency communication (URLLC), and general Internet service. The study further described the development of a 5G heterogeneous network platform to collect UAV-related aerial computing data, on which extensive experiments were conducted. Their results indicated that the SAE-CNN model outperformed existing state-of-the-art methods and enabled more effective traffic management and classification for UAV applications in complex 5G environments.

Kalwar and Bhatti (2024) examined the rapid expansion of the Internet of Things (IoT) and the consequent surge of heterogeneous network traffic from interconnected devices. They highlighted that effective classification of such traffic was critical for optimizing resource allocation, improving security, and ensuring efficient network management in IoT systems. The authors noted that deep learning had emerged as a prominent technique for network traffic classification due to its capacity to autonomously learn complex patterns and representations from raw data. Their survey systematically reviewed and categorized existing deep learning approaches applied to IoT network traffic, evaluating the strengths and limitations of various models in addressing the distinctive challenges of IoT environments. They emphasized that the study provided researchers and practitioners with insights into current methodologies, identified gaps in the literature, and suggested directions for future research aimed at enhancing the efficiency and effectiveness of deep learning-based IoT traffic classification.

Najm et al. (2024, May) investigated the challenges in network traffic classification, highlighting its critical role in ensuring the security and efficiency of modern computer networks. They noted that existing techniques often struggled to effectively identify and characterize network traffic patterns. To address this gap, they proposed an updated approach leveraging machine learning algorithms, extending previous studies by emphasizing robust feature engineering and the application of various algorithms, including decision trees, random forests, support vector machines, and recurrent neural networks. Their

methodology involved training on a comprehensive dataset representing diverse network traffic scenarios. The study reported promising results, with decision trees achieving 93% accuracy, random forests 97.89%, support vector machines 91%, and recurrent neural networks 89.49%. Their findings were interpreted to indicate that machine learning approaches held significant potential for improving real-world network traffic classification.

Chauhan and Jain (2023) examined the growing risks associated with technological networks, emphasizing that continuous information sharing often exposed systems to cyber-attacks, particularly through open ports on networking devices. They highlighted that in real-world scenarios, users did not always have access to secure private networks or VPNs, making networks increasingly vulnerable to evolving threats. The authors reviewed efforts to develop networking tools, including network profiling, vulnerability scanning, and network mapping, aimed at supporting intrusion detection systems (IDS). They noted that machine learning (ML) had recently gained prominence for identifying fraudulent network traffic, with the performance of ML models largely dependent on the quality of training datasets. In their study, they used profiling network data to train a classification model capable of distinguishing between normal and anomalous traffic. Various approaches, such as Naive Bayes, Bayes Net, Naive Bayes Multinomial Text, and Naive Bayes Updateable, were employed using Weka 3.8.5, and the resulting models were tested across multiple simulations to evaluate their effectiveness.

Gupta et al. (2023) examined the necessity of detecting applications traversing networks within the internet community to perform specific tasks. They highlighted that Internet Service Providers (ISPs) typically applied network traffic classification to identify connection prerequisites, which in turn influenced the efficiency of cable networks. The study discussed various Internet Protocol (IP) methods, including bandwidth-based, stream-based, and machine learning (ML) approaches, noting the distinct advantages and limitations of each. The authors reported that the ML approach had gained popularity due to its widespread application across disciplines and the growing familiarity of researchers with its methodology. Their work involved comparing the performance of Naive Bayes and K-Nearest Neighbor (KNN) algorithms using a network-specific dataset obtained from live stream feeds and Ethernet software. Tools such as Python's pandas, numpy, and sklearn modules were employed to develop the algorithms. Findings indicated that the KNN algorithm outperformed Support Vector Machine, Naive Bayes, and Decision Tree methods in terms of overall efficiency.

Jmila et al. (2022) examined the security challenges of smart home IoT devices, noting that their lack of proper protection raised significant safety and privacy concerns. They argued that conventional one-size-fits-all network administration was inadequate due to the diverse QoS requirements of IoT devices, and proposed that device classification could enhance both administration and security by identifying vulnerable or rogue devices and enabling automated management based on device type or function. The study emphasized the emerging role of Machine Learning (ML)-based traffic analysis to reveal hidden patterns in IoT traffic and support automatic device classification. The authors outlined a generic workflow for IoT device classification and reviewed existing methods and solutions at each stage, including traffic data acquisition methodologies, public dataset categorization, feature extraction techniques, open-source tools, and ML approaches with their evaluation strategies. Their analysis presented findings in taxonomies with statistical insights into literature trends and identified potential research directions that remained underexplored.

Jonathan et al. (2021, February) examined network traffic classification, which was defined as the process of assigning appropriate identification to each traffic flow within a network. They noted that traditional approaches, such as port-based, payload-based, and behavior-based methods, had been applied

previously but were found to possess various limitations. The study highlighted the growing focus on machine learning (ML) techniques, which leveraged the statistical properties of traffic flows; however, these methods were reported to underperform when handling large-scale traffic data with numerous features and instances. To address this, the researchers employed feature selection to remove irrelevant and redundant features prior to applying ML classifiers. Their investigation compared classification performance with and without feature selection, considering metrics such as runtime, accuracy, recall, precision, and F-score. The results were reported to show that classification without feature selection achieved an average accuracy of 94.14% and a runtime of 0.52 seconds, whereas the feature-selected approach improved accuracy to 95.61% and reduced runtime to 0.25 seconds, underscoring the benefit of including only relevant features in network traffic classification.

Trang and Nguyen (2021) examined the rapid increase of Internet usage and its integration into everyday life, highlighting the corresponding growth in network infrastructure. They emphasized that protecting users' confidential information had become a critical concern, as existing encryption algorithms and techniques, despite being widely applied by internet providers and web hosting services, still left networks vulnerable to anonymous hacker attacks. The authors noted that classifying network data streams had gained attention as a means to enhance system quality and security. Their study introduced a machine learning-based approach aimed at identifying the most suitable model for network traffic classification. They reported that data pre-processing was first conducted to normalize feature types, after which various techniques, including k-Nearest Neighbors (KNN), Artificial Neural Networks (ANN), and Random Forest (RF), were applied. Using the open-access ISCXVPN2016 dataset, which included VPN and non-VPN traffic across seven categories, they found that the proposed models achieved high classification rates, with the Random Forest model performing the best.

Raikar et al. (2020) examined the critical role of traffic classification in network operations, emphasizing its importance for security monitoring, traffic engineering, fault detection, usage accounting, billing, and differentiating Quality of Service (QoS) among network services. They noted that with the rapid growth in internet users, traditional primitive techniques for network traffic classification had struggled to maintain reliable accuracy due to the exponential increase in devices and data flows. To address these limitations, they proposed the integration of Software Defined Network (SDN) architecture with machine learning methods. The study applied three supervised learning models—Support Vector Machine (SVM), nearest centroid, and Naïve Bayes (NB)—to classify network traffic based on application types within an SDN platform. Network traffic traces were captured, and flow features were extracted and fed into the classifiers for prediction. The reported accuracies were 92.3% for SVM, 96.79% for NB, and 91.02% for nearest centroid. They highlighted challenges related to live network traffic capture and accurate application classification within the SDN environment.

Bakker et al. (2019, April) reported on their experience with deploying network traffic classifiers in a live Software Defined Network (SDN). They selected five simple machine learning (ML) algorithms and implemented them for Distributed Denial of Service (DDoS) detection. Using publicly available datasets, they established a standard reference for the performance of each classifier in terms of accuracy, precision, and detection rate. Their subsequent experiments over a live SDN revealed that the classifiers performed significantly worse compared to the reference standards, showing reductions of up to 11.2% in accuracy, 30.2% in precision, and detection rates falling below 15% versus 98% in the reference. They argued that interactions between network elements, such as the switch and controller, substantially affected the performance of ML algorithms in a live environment, highlighting the need to account for such operational factors in real-world deployments.

Shafiq et al. (2018) highlighted that class imbalance had emerged as a significant challenge causing inaccurate traffic classification, which is critical for tasks such as security monitoring, IP management, and intrusion detection. They noted that machine learning (ML) approaches had been widely applied in the literature to address this problem. In their study, they proposed an ML-based hybrid feature selection algorithm, WMI_AUC, which utilized two metrics—weighted mutual information (WMI) and area under the ROC curve (AUC)—to identify effective features from traffic flows. Additionally, they introduced a robust feature selection algorithm to further refine the selected features. Their approach was reported to enhance the accuracy of ML classifiers and improve detection of malicious traffic. The study evaluated the proposed algorithms across 11 well-known ML classifiers using various network environment trace datasets, and the experimental results were observed to achieve flow accuracy exceeding 95%, demonstrating the effectiveness of their methodology.

Vlăduţu et al. (2017) examined the application of machine learning in network traffic classification as an alternative to traditional deep packet inspection techniques. They highlighted that machine learning offered both unsupervised and supervised algorithms capable of distinguishing similar traffic types or identifying Internet protocols from pre-labeled training samples. In their study, they proposed a novel approach whereby unidirectional and bidirectional flows were first extracted from captured network traffic, with each flow defined as a set of packets sharing the same sender and receiver IP addresses and ports. Subsequently, relevant statistical properties of these flows were selected, and an unsupervised learning mechanism was applied to cluster flows based on similarity. The resulting clusters were then utilized as training input for a supervised learning engine, which was intended to accurately classify new, previously unseen network traffic flows.

III. METHODOLOGY

The methodology of this study was designed to classify network traffic using machine learning techniques for efficient network management. First, network traffic data were collected from a simulated or real network environment, including different traffic types such as web browsing, video streaming, file transfer, VoIP, gaming, email, and suspicious traffic. After data collection, preprocessing was performed to remove missing values, duplicate records, and irrelevant features. Important flow-based features such as packet size, flow duration, protocol type, number of packets, byte count, source port, destination port, and inter-arrival time were selected for classification. The dataset was then divided into training and testing sets, where 80% of the data were used for training and 20% for testing. Different machine learning algorithms such as Naïve Bayes, Decision Tree, Support Vector Machine, Random Forest, and Artificial Neural Network were applied to classify traffic patterns. The trained models were evaluated using accuracy, precision, recall, F1-score, and processing time. Finally, the model with the best performance was selected for improving network monitoring, bandwidth allocation, congestion control, and security management.

IV. RESULT

The result of the study showed that machine learning-based network traffic classification improved the efficiency and accuracy of network management compared to traditional traffic identification methods. The collected network traffic data were processed using feature extraction, data cleaning, normalization, and classification techniques. Different traffic categories such as web browsing, video streaming, file transfer, VoIP, gaming, email, and suspicious traffic were classified using machine learning algorithms. The performance of the models was evaluated on the basis of accuracy, precision, recall, F1-score, and processing time. The results indicated that ensemble-based algorithms such as Random Forest performed better than basic classifiers because they handled complex traffic patterns more effectively. Support

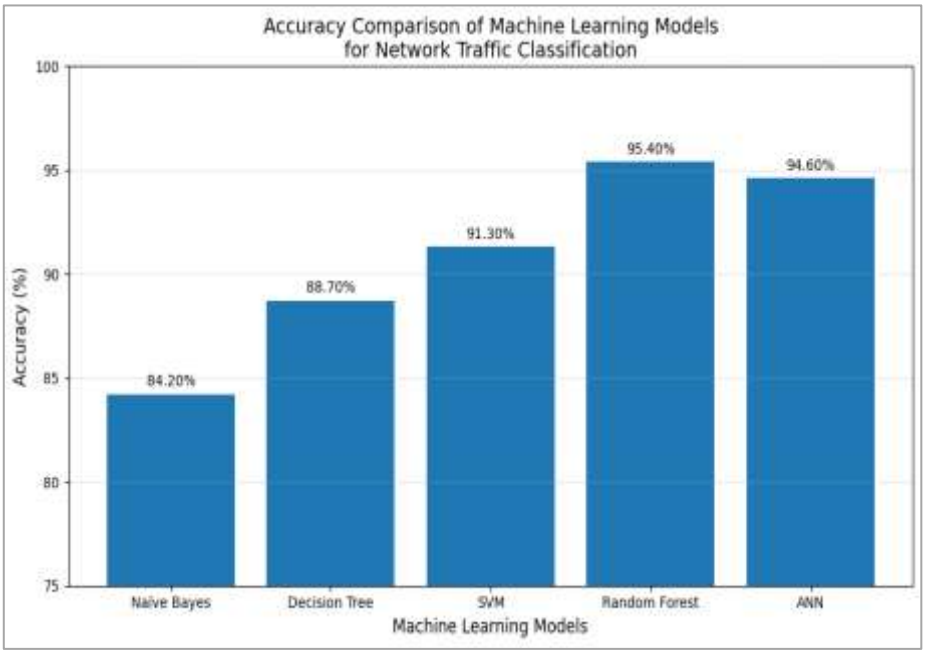
Vector Machine also produced good classification accuracy but required comparatively higher computation time. Decision Tree gave faster results, but its accuracy was slightly lower due to overfitting in some cases. Naïve Bayes showed moderate performance and was suitable for simple traffic classification tasks. The Artificial Neural Network achieved high accuracy in identifying complex and encrypted traffic patterns, but it required more training time and computational resources.

Table 1: Performance Comparison of Machine Learning Models for Network Traffic Classification

Machine Learning Model	Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)	Processing Time (Sec)
Naïve Bayes	84.20	82.60	81.90	82.25	1.8
Decision Tree	88.70	87.40	86.80	87.10	2.1
Support Vector Machine	91.30	90.50	89.80	90.15	4.6
Random Forest	95.40	94.80	94.10	94.45	3.2
Artificial Neural Network	94.60	93.90	93.50	93.70	5.4

The overall findings revealed that Random Forest achieved the highest accuracy of 95.40%, followed by Artificial Neural Network with 94.60% and Support Vector Machine with 91.30%. This indicates that Random Forest was more suitable for efficient network traffic classification because it provided a strong balance between accuracy and processing time. The classification results also showed that machine learning models could identify traffic types more accurately by analyzing flow-based features instead of depending only on port numbers or packet payloads. This is especially useful in modern networks where encrypted and dynamic traffic patterns are common. Therefore, the result confirms that machine learning techniques can significantly support efficient network management by improving traffic monitoring, bandwidth allocation, quality of service, and early detection of suspicious network activities.

Bar Graph



The bar graph shows the accuracy comparison of different machine learning models used for network traffic classification. Random Forest achieved the highest accuracy of 95.40%, showing that it performed best in identifying different types of network traffic. Artificial Neural Network also gave strong performance with 94.60% accuracy because it can learn complex traffic patterns. Support Vector Machine achieved 91.30%, indicating reliable classification ability. Decision Tree recorded 88.70%, while Naïve

Bayes showed the lowest accuracy of 84.20%. Overall, the graph proves that advanced machine learning models improve traffic classification accuracy and support efficient network monitoring, bandwidth management, and security analysis.

V. CONCLUSION

The study concluded that machine learning-based network traffic classification is an effective approach for improving modern network management. Traditional traffic classification methods are not sufficient for present network environments because network traffic has become more dynamic, encrypted, and complex. Machine learning models can identify traffic patterns using flow-based features such as packet size, protocol type, flow duration, byte count, and packet frequency. The result showed that Random Forest achieved the best performance with the highest accuracy, followed by Artificial Neural Network and Support Vector Machine. This proves that advanced models can classify different types of traffic more accurately than simple techniques. Effective traffic classification helps network administrators manage bandwidth, reduce congestion, improve quality of service, detect suspicious activities, and strengthen network security. Therefore, machine learning provides a smart, scalable, and reliable solution for efficient network monitoring and management.

REFERENCES

1. Martínez Hernández, L. A., Pérez Arteaga, S., Sandoval Orozco, A. L., & García Villalba, L. J. (2026). Adversarial Attacks on Machine Learning Models for Network Traffic Filtering. *Engineering Proceedings*, *123*(1), 23.
2. Pulido, J., Wilhelmi, F., Fortes, S., Fernández-Durán, A., Giordano, L. G., & Barco, R. (2026). Studying the Role of Synthetic Data for Machine Learning-based Wireless Networks Traffic Forecasting. *arXiv preprint arXiv:2601.07646*.
3. LOHIYA, P. B., & Bamnote, G. R. (2025). Internet Traffic Classification through Supervised Learning: Exploring Machine Learning Techniques. *Intelligent Methods In Engineering Sciences*, *4*(1), 8-14.
4. Chen, C., Liu, Z., Yu, Y., Jin, F., Han, W., Berretti, S., ... & Pei, Q. (2025). A Deep-Learning-Based traffic classification method for 5G aerial computing networks. *IEEE Internet of Things Journal*, *12*(9), 11244-11257.
5. Kalwar, J. H., & Bhatti, S. (2024). Deep learning approaches for network traffic classification in the internet of things (iot): A survey. *arXiv preprint arXiv:2402.00920*.
6. Najm, I. A., Saeed, A. H., Ahmad, B. A., Ahmed, S. R., Sekhar, R., Shah, P., & Veena, B. S. (2024, May). Enhanced network traffic classification with machine learning algorithms. In *Proceedings of the cognitive models and artificial intelligence conference* (pp. 322-327).
7. Chauhan, D., & Jain, J. K. (2023, January). Profiling network traffic by using classification techniques in machine learning. In *International Conference on Smart Trends in Computing and Communications* (pp. 113-123). Singapore: Springer Nature Singapore.
8. Gupta, K., Jiwani, N., Sharif, M. H., Mohammed, V. A., Mohammed, M. A., & Mohammed, M. (2023). Implementation of machine learning for network traffic classification. *Eur. Chem. Bull.*, *12*, 674-682.
9. Jmila, H., Blanc, G., Shahid, M. R., & Lazrag, M. (2022). A survey of smart home iot device classification using machine learning-based network traffic analysis. *IEEe Access*, *10*, 97117-97141.

10. Jonathan, O., Misra, S., & Osamor, V. (2021, February). Comparative analysis of machine learning techniques for network traffic classification. In *IOP Conference Series: Earth and Environmental Science* (Vol. 655, No. 1, p. 012025). IOP Publishing.
11. Trang, K., & Nguyen, A. H. (2021). A comparative study of machine learning-based approach for network traffic classification. *Knowledge Engineering and Data Science*, *4*(2), 6.
12. Raikar, M. M., Meena, S. M., Mulla, M. M., Shetti, N. S., & Karanandi, M. (2020). Data traffic classification in software defined networks (SDN) using supervised-learning. *Procedia Computer Science*, *171*, 2750-2759.
13. Bakker, J., Ng, B., Seah, W. K., & Pekar, A. (2019, April). Traffic classification with machine learning in a live network. In *2019 IFIP/IEEE Symposium on Integrated Network and Service Management (IM)* (pp. 488-493). IEEE.
14. Shafiq, M., Yu, X., Bashir, A. K., Chaudhry, H. N., & Wang, D. (2018). A machine learning approach for feature selection traffic classification using security analysis. *The Journal of Supercomputing*, *74*(10), 4867-4892.
15. Vlăduţu, A., Comăneci, D., & Dobre, C. (2017). Internet traffic classification based on flows' statistical properties with machine learning. *International Journal of Network Management*, *27*(3), e1929.