Vol 6, Issue 3, March 2026 www.ijesti.com E-ISSN: 2582-9734
International Journal of Engineering, Science, Technology and Innovation (IJESTI)

Machine Learning-Based Network Traffic Classification for Enhanced Security and Efficient Network Management: A Review

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

With the rapid growth of network traffic driven by various applications, traditional network traffic classification methods like port-based and deep packet inspection (DPI) have become ineffective. As encrypted traffic surges, machine learning (ML) and deep learning (DL) models have emerged as promising solutions. Techniques such as CNN, LSTM, and hybrid models have demonstrated significant improvements in traffic classification accuracy and network management. This study explores the use of ML in traffic classification for enhancing network security, scalability, and adaptability while addressing challenges such as encryption and concept drift in dynamic environments.

Keywords: *Network Traffic, Machine Learning, Deep Learning, Classification, Security.*

I. INTRODUCTION

The rapid expansion of computer networks and internet-based services has led to an unprecedented increase in network traffic volume, diversity, and complexity. Modern networks now carry heterogeneous traffic generated from applications such as video streaming, cloud computing, Internet of Things (IoT) devices, online gaming, and enterprise services. This exponential growth has made network traffic classification a fundamental requirement for efficient network management, security enforcement, and Quality of Service (QoS) provisioning. Traditional traffic classification methods such as port-based and Deep Packet Inspection (DPI) techniques are no longer effective due to the widespread use of encryption, dynamic port allocation, and evolving application behaviors. Rezaei and Liu (2018) emphasized that the surge in encrypted traffic has significantly reduced the effectiveness of conventional classification approaches, thereby necessitating more intelligent and adaptive solutions. Similarly, Shafiq et al. (2017) highlighted that increasing internet usage and diverse application environments demand accurate and scalable traffic identification techniques for Internet Service Providers (ISPs). In addition, the emergence of Software-Defined Networking (SDN) has further transformed network architectures by introducing centralized control and programmability, enabling more dynamic traffic management strategies. Mohammed et al. (2019) observed that although SDN provides improved control and monitoring capabilities, the massive volume of traffic data generated still presents significant challenges for real-time processing and decision-making. Furthermore, the integration of IoT devices into modern networks has introduced additional complexity, as highlighted by Kumar et al. (2021), who noted that IoT traffic requires continuous monitoring to ensure security and operational efficiency. Collectively, these developments underline the necessity for intelligent, scalable, and adaptive network traffic classification systems capable of handling modern network environments effectively.

In response to the limitations of traditional approaches, machine learning (ML) and deep learning (DL) techniques have emerged as powerful tools for network traffic classification. These techniques enable automated feature extraction, pattern recognition, and predictive modeling, making them highly suitable for complex and high-dimensional network data. Over the past decade, researchers have transitioned from

classical machine learning models such as Support Vector Machines (SVM), Decision Trees, and Artificial Neural Networks (ANN) to more advanced deep learning architectures. Labayen et al. (2020) demonstrated that hybrid machine learning systems combining clustering and supervised learning can achieve high classification accuracy for user activity detection in network traffic. Similarly, Izadi et al. (2022) proposed a deep learning-based framework integrating Convolutional Neural Networks (CNN), Deep Belief Networks (DBN), and Multi-Layer Perceptrons (MLP) with Bayesian decision fusion, achieving improved performance in encrypted traffic classification. Long and Jinsong (2023) further showed that deep neural networks trained on NetFlow data can achieve approximately 95% classification accuracy, outperforming traditional classifiers in real-world datasets. More recent advancements include hybrid deep learning architectures such as 1D-CNN and LSTM models, which combine spatial and temporal feature extraction capabilities. Zhou and Zhao (2025) demonstrated that a 1DCNN-LSTM hybrid model achieved an accuracy of 99.051%, significantly outperforming standalone models. Wang and Song (2025) also reported that CNN-based architectures achieved up to 98.7% accuracy in traffic classification and anomaly detection tasks. These advancements highlight the growing effectiveness of deep learning techniques in addressing the limitations of conventional methods while improving classification accuracy, robustness, and adaptability in dynamic network environments.

The application of machine learning-based traffic classification extends beyond performance improvement and plays a crucial role in enhancing network security, anomaly detection, and intelligent network management. In Software-Defined Networking (SDN), ML-based traffic classification enables real-time monitoring and adaptive control of network flows. Serag et al. (2024) emphasized that integrating ML with SDN improves QoS, intrusion detection, and attack mitigation by enabling dynamic and intelligent decision-making. Similarly, Eldhai et al. (2024) demonstrated that machine learning techniques such as Hoeffding Adaptive Trees (HAT) and Adaptive Random Forest (ARF) can effectively handle streaming traffic data and concept drift in SDN environments, achieving up to 95% classification accuracy. Gómez et al. (2023) further illustrated the use of decision tree-based models for distinguishing elephant and mice flows in IP networks, enabling efficient traffic routing and resource allocation. In the context of cybersecurity, Kirubavathi et al. (2026) proposed an ensemble-based ML framework for Android ransomware detection using network traffic metadata, achieving high adaptability against evolving threats and concept drift. Pekar et al. (2026) provided a comprehensive workflow for flow-based traffic classification, emphasizing real-world deployment considerations, feature engineering, and reproducibility in network environments. These studies collectively demonstrate that ML-based traffic classification is not only a tool for categorizing network flows but also a critical component for proactive security management, anomaly detection, and optimization of network resources in modern digital infrastructures.

Despite significant advancements, several challenges remain in the field of machine learning-based network traffic classification. One of the primary challenges is the dynamic and evolving nature of network traffic, which leads to concept drift and reduces model performance over time. Additionally, the increasing use of encryption limits the availability of payload-based features, making classification more dependent on metadata and statistical flow features. Another major challenge is the scarcity of high-quality, labeled datasets that accurately represent real-world network conditions. Getman and Ikonnikova (2022) emphasized that dataset availability, feature selection, and model evaluation remain critical issues in developing reliable classification systems. Furthermore, deep learning models often require high computational resources, which may limit their deployment in resource-constrained environments such as edge and IoT devices. Scalability is also a concern, particularly in large-scale SDN deployments where real-time processing is essential. Serag et al. (2024) highlighted that integrating ML into SDN

architectures introduces performance and scalability constraints that must be addressed for practical implementation. Therefore, future research should focus on developing lightweight, adaptive, and interpretable machine learning models capable of operating efficiently in dynamic and large-scale network environments. The present study aims to explore machine learning-based network traffic classification techniques for efficient network management, focusing on improving accuracy, adaptability, and scalability while addressing existing limitations in real-world applications.

II. RESEARCH BACKGROUND

Kirubavathi et al. (2026) examined the growing threat of Android ransomware in mobile ecosystems, noting that such attacks increasingly exploited obfuscated payloads and dynamic command-and-control channels to bypass conventional detection systems. The authors observed that existing static and batch-trained models often lacked adaptability to evolving threat behaviors, which led to performance degradation over time due to concept drift. To address this limitation, they proposed a robust ensemble-based machine learning framework for proactive Android ransomware detection using network traffic metadata. The framework was reported to integrate advanced classifiers such as Light Gradient Boosting Machine (LightGBM), eXtreme Gradient Boosting Machine (XGBoost), and Random Forest, along with Synthetic Minority Oversampling Technique (SMOTE)-enhanced stratified cross-validation to handle class imbalance and improve generalizability. Furthermore, SHapley Additive exPlanations (SHAP) and Local Interpretable Model-Agnostic Explanations (LIME) were employed to enhance interpretability and analyst trust. The study also conducted a concept drift evaluation using an incremental LightGBM model across chronologically partitioned traffic data, demonstrating that LightGBM achieved the highest classification performance and exhibited strong adaptability for real-time Android ransomware mitigation in dynamic network environments.

Pekar et al. (2026) presented a practical tutorial on machine-learning-based network traffic flow classification for modern networks carrying diverse and encrypted traffic, where traditional port-based and payload-based approaches had become less effective. The authors explained an end-to-end workflow beginning with flow metering and dataset creation, followed by ground-truth labeling, feature engineering, leakage-resistant experimental design, model training, evaluation, explainability, and deployment considerations. Their study particularly emphasized supervised flow-based classification methods, which were considered suitable for encrypted traffic environments. They also discussed algorithm selection, appropriate performance metrics, and realistic data partitioning strategies, while highlighting common real-world measurement artifacts and methodological pitfalls that could affect classification accuracy. Furthermore, the tutorial was supported by a companion set of five Jupyter notebooks made available on GitHub, which demonstrated the complete data-to-model pipeline using real traffic captures and enabled reproducibility. Overall, the work was intended to guide researchers and practitioners in designing and deploying robust and practical traffic classification systems in operational network environments.

Wang and Song (2025) examined the growing complexity and diversity of internet traffic and observed that traditional traffic classification and anomaly detection methods had become insufficient for addressing modern network threats. To overcome this limitation, they proposed a deep learning-based approach for network traffic classification and anomaly detection. Their study had constructed a convolutional neural network (CNN) model using extensive network traffic datasets to accurately distinguish between normal and abnormal traffic patterns. The experimental findings indicated that the proposed model had achieved a high accuracy of 98.7% on the test dataset. It had also demonstrated excellent performance on the CIC-IDS2017 and ISCX VPN-NOVPN datasets, attaining accuracies of

98.5% and 99.2%, respectively. Furthermore, the model had significantly improved recall and F1-score while reducing error rates, thereby outperforming conventional methods. Through comparative analysis of different network architectures, the authors had further optimized the model and reduced the false alarm rate to 1.5%. Overall, the study had provided effective technical support for enhancing network security and had demonstrated strong robustness and adaptability across diverse real-world network environments.

Zhou and Zhao (2025) reported that malicious Internet attacks had continued to increase, while traditional network architectures had limited network managers from obtaining a complete network view, thereby creating significant challenges in network security maintenance and management. Their study had focused on network traffic identification, classification, and anomaly detection for improving network security control and enhancing overall protection capabilities. The authors had proposed an anomaly traffic detection algorithm based on the integration of 1D Convolutional Neural Network (1DCNN) and Long Short-Term Memory (LSTM), which had utilized the temporal characteristics of network data. It had effectively combined the spatial feature extraction capability of CNN with the temporal sequence learning strength of LSTM. The experimental findings had demonstrated that the proposed 1DCNN-LSTM model had significantly reduced the volume of data requiring processing and had achieved an accuracy of 99.051% on the NSL-KDD dataset, which had been approximately 20% and 10% higher than models using CNN or LSTM individually.

Eldhai et al. (2024) investigated the use of machine learning (ML) for traffic classification (TC) in software-defined networks (SDN) and reported it as a promising approach to enhance network management. They noted that TC could improve SDN operability, while SDN facilitated faster feature selection (FS), particularly when ML was applied to extract measurements and related information from incoming data at the SDN controller. Despite these benefits, they highlighted challenges arising from the frequent similarity of traffic profiles, which complicated classification, and the additional difficulties posed when combining TC with stream learning (SL). To address these issues, they proposed robust statistical flow features capable of extracting online features, handling concept drift, and processing continuous data streams within limited time and memory resources. The study introduced the Boruta FS mechanism and evaluated three streaming-based TC methods—Hoeffding adaptive trees (HAT), adaptive random forest (ARF), and k-nearest neighbor with adaptive sliding window detector (KNN-ADWIN)—using real and synthetic traffic traces. Their results demonstrated that Boruta achieved up to 95% average accuracy, while SL methods maintained up to 85% accuracy with reduced memory and time overhead, particularly for HAT.

Serag et al. (2024) examined the ongoing efforts to upgrade legacy communication networks to support modern services and applications, emphasizing the adaptability of smart networks to emerging technologies and traffic trends. They investigated the role of software-defined networking (SDN) in separating the control plane from the data plane to enable centralized programmability, highlighting how SDN, combined with machine learning (ML), could enhance network performance and quality-of-service (QoS). The study analyzed ML classification techniques in traffic classification (TC), demonstrating their advantages over traditional methods and underscoring the effectiveness of ML algorithms for accurately classifying SDN traffic flows. It further explored the benefits of dynamic and adaptive TC through ML, including improvements in network security via anomaly detection, intrusion detection, and attack mitigation, stressing proactive threat management. Additionally, Serag et al. discussed the challenges of integrating ML with SDN for QoS, identified scalability and performance issues in large-scale SDN deployments, and highlighted key research gaps for future investigation.

Gómez et al. (2023) investigated the challenges faced by data centers in higher education institutions and large corporations regarding traffic flow management. They observed that limitations in hardware resources or memory constraints could cause network performance to lag, particularly under exponential traffic growth. The study contributed by implementing a machine learning-based system for classifying elephant and mice flows, enabling early detection through the dynamic calculation of thresholds based on input parameters. Initially, training algorithms were employed to identify optimal performance, followed by the selection of the model with the highest predictive accuracy from a supervised learning algorithm trained offline. In the subsequent online prediction phase, the algorithm was reported to predict traffic types with high precision and to dynamically update thresholds for flow characterization, allowing network hardware to route traffic accordingly. Their results indicated that the decision tree model achieved the best predictive performance, reaching a confidence level of 100%.

Long and Jinsong (2023) examined the critical role of network traffic classification in supporting activities such as security monitoring, accounting, quality of service, and long-term provisioning. They noted that conventional approaches struggled to handle the growing prevalence of encrypted traffic, which had led to the increasing adoption of machine learning methods. Despite the widespread use of NetFlow as a network monitoring tool, the authors observed that machine learning applications for traffic classification based on sampled NetFlow data were underdeveloped. To address this gap, they proposed a network traffic classification module that combined NetFlow data with a deep neural network. The study reported that the module achieved an average classification accuracy of 95% across approximately 1.4 million test cases from two real-world datasets and outperformed three other state-of-the-art classifiers. The authors concluded that their proposed module offered a promising and effective alternative for network traffic classification.

Getman and Ikonnikova (2022) presented a survey focused on network traffic classification, emphasizing the application of machine learning algorithms in this domain. They began by outlining the task, its various formulations, and potential real-world applications, before reviewing the historical methods used for network traffic classification along with their limitations and the evolution that led to the adoption of machine learning as the primary solution. The authors described the most widely used machine learning algorithms, illustrating their advantages and disadvantages with references to relevant studies. They also addressed the challenge of feature selection and discussed the broader issue of obtaining suitable datasets for effective network traffic classification, providing examples of commonly used datasets. Finally, the survey highlighted ongoing challenges in the field, including model training and evaluation, protection of user data, and the inherent volatility of network traffic.

Izadi et al. (2022) investigated the increasing significance of network traffic classification due to the rapid expansion of computer networks and their applications. They noted that deep learning had emerged as a prominent approach in this field, but its requirement for large amounts of training data posed a challenge, particularly when sufficient data for diverse network traffic types was unavailable, adversely affecting classification accuracy. To address this limitation, they proposed a network traffic classification method combining deep learning and decision-level data fusion techniques. The approach involved preprocessing the dataset, applying three deep learning models—Deep Belief Network, Convolutional Neural Network, and Multi-Layer Perceptron—to classify traffic, and then integrating the outputs using Bayesian decision fusion. The method was capable of identifying encrypted traffic and differentiating between VPN and non-VPN traffic. Their experiments on the ISCX VPN-nonVPN dataset demonstrated that the proposed method enhanced classification performance across various traffic types, achieving an average accuracy of 97%.

Kumar et al. (2021) examined the role of the Internet of Things (IoT) in enabling interconnected smart environments through embedded devices that transmit and share information. They highlighted that regular monitoring of IoT network traffic was essential for ensuring proper device functioning and detecting malicious activities, with particular emphasis on IoT device classification. It was noted that effective classification could assist administrators in monitoring device activities, implementing Quality of Service, and identifying malicious devices. The authors reviewed literature on various machine learning methods for IoT traffic classification, emphasizing that their accuracy depended on factors such as data sources, extracted features, and deployment environments, while also noting that manual feature and algorithm selection could introduce errors. In their study, they conducted a comparative analysis of popular machine learning algorithms using features extracted from a public dataset comprising 20 days of network traces from 20 IoT devices. They processed the traces, applied state-of-the-art algorithms, evaluated performance in terms of accuracy, speed, and training time, and finally suggested algorithm selection strategies for different use cases based on the results.

Labayen et al. (2020) investigated the challenges arising from the daily deployment of new applications and the exponential increase in network traffic, which had led to heightened complexity in network analysis and monitoring. They noted that the growing availability and reduced cost of computational resources had facilitated the wider adoption of machine learning algorithms. The authors proposed a system for classifying user activities from network traffic by leveraging both supervised and unsupervised learning methods. The system analysed user behaviour across all traffic generated within a specified time window, extracting features from network and transport layer headers. A three-layer classification model was introduced, in which the first two layers applied K-Means clustering and the final layer used a Random Forest to assign activity labels. The study reported an average accuracy of 97.37%, alongside high precision and recall, demonstrating the system's capability to perform online network traffic classification for Quality of Service (QoS) and user profiling, surpassing prior approaches.

Mohammed et al. (2019) highlighted that the Internet had been continuously expanding and increasing in complexity, prompting ongoing advancements in networking to manage the growing network traffic. They observed that while Software Defined Networking (SDN) provided a centralized mechanism for traffic measurement, control, and prediction, the SDN controller still faced challenges in handling the massive volumes of data it received. To address this, it was suggested that Machine Learning (ML) could be employed for more efficient data processing. Their review examined existing proposals for integrating ML within an SDN framework, particularly focusing on traffic classification and prediction. The study noted that the application of Deep Learning (DL) in traffic prediction had remained largely unexplored in prior surveys. Additionally, they discussed persistent challenges in this domain and outlined potential directions for future research to enhance the performance and scalability of ML-driven SDN solutions.

Rezaei and Liu (2018) examined the evolution of traffic classification over two decades, highlighting its applications in quality of service provisioning, billing in ISPs, and security measures such as firewalls and intrusion detection systems. They noted that traditional methods, including port-based techniques, data packet inspection, and classical machine learning approaches, had been widely employed, but their accuracy had declined due to significant changes in Internet traffic patterns, especially the surge in encrypted traffic. The study indicated that with the emergence of deep learning methods, researchers had increasingly explored these techniques for traffic classification and reported notable improvements in accuracy. Rezaei and Liu proposed a general framework for deep-learning-based traffic classification, reviewed commonly adopted deep learning methods and their specific applications, and further analyzed open problems, associated challenges, and potential opportunities within the field, emphasizing the growing relevance of advanced machine learning approaches in modern network management.

Shafiq et al. (2017) highlighted the increasing importance of Network Traffic Classification due to the rapid growth in internet users and diverse application usage. They emphasized that monitoring network traffic was critical for Internet Service Providers (ISPs). Previous studies on network traffic classification using machine learning for traffic identification, often based on single-campus network datasets, were reported to yield limited accuracy. In their study, they attempted to achieve higher precision by employing two datasets, namely HIT and NIMS. They captured online traffic from seven different applications, including DNS, FTP, TELNET, P2P, WWW, IM, and MAIL, and extracted packet features using the NetMate tool. Subsequently, they applied three machine learning algorithms—Artificial Neural Network, C4.5 Decision Tree, and Support Vector Machine—to evaluate performance. The experimental findings indicated that all algorithms produced high accuracy, with the C4.5 Decision Tree algorithm achieving the highest precision of 97.57%, outperforming the other two approaches.

III. KEY FINDINGS FROM STUDY

Author (Year)	Study Focus	Methodology	Key Findings	Limitation/Gap
Kirubavathi et al. (2026)	Android ransomware detection using network traffic	Ensemble ML (LightGBM, XGBoost, Random Forest), SMOTE, SHAP, LIME	High detection accuracy; LightGBM best under concept drift	High model complexity in real-time deployment
Pekar et al. (2026)	Flow-based traffic classification workflow	Supervised ML pipeline with full lifecycle design	Provided end-to-end reproducible framework for traffic classification	Limited focus on real-time adaptive systems
Wang & Song (2025)	Traffic classification & anomaly detection	CNN-based deep learning model	Achieved up to 98.7% accuracy; strong anomaly detection	Requires high computational resources
Zhou & Zhao (2025)	Hybrid traffic classification system	1DCNN-LSTM model	Achieved 99.051% accuracy; strong spatial-temporal learning	Limited generalization across datasets
Eldhai et al. (2024)	Streaming traffic classification in SDN	HAT, ARF, KNN-ADWIN, Boruta feature selection	Up to 95% accuracy; effective concept drift handling	Accuracy slightly lower than deep learning models
Serag et al. (2024)	ML in SDN traffic classification	ML-based classification in SDN architecture	Improved QoS, security, and traffic control	Scalability and integration issues in SDN
Gómez et al. (2023)	Flow classification in IP networks	Decision Tree-based ML model	Achieved very high accuracy; effective flow classification	Limited adaptability to encrypted traffic
Long & Jinsong (2023)	NetFlow-based traffic classification	Deep Neural Network	~95% accuracy; strong performance on real datasets	Needs large labeled datasets

Getman & Ikonnikova (2022)	Survey of ML methods in traffic classification	Literature review	Identified key ML techniques and dataset challenges	Lack of unified evaluation framework
Izadi et al. (2022)	Encrypted traffic classification	CNN, DBN, MLP + Bayesian fusion	Achieved ~97% accuracy; improved encrypted traffic detection	High computational cost
Kumar et al. (2021)	IoT traffic classification	Comparative ML algorithms	Trade-off between accuracy, speed, and training time	Manual feature selection limitations
Labayen et al. (2020)	User activity classification	K-Means + Random Forest hybrid model	Achieved 97.37% accuracy; effective online classification	Limited scalability for large networks
Mohammed et al. (2019)	ML in SDN traffic prediction	ML/DL-based SDN framework	Improved traffic prediction and control	DL integration still underexplored
Rezaei & Liu (2018)	Deep learning for encrypted traffic	Deep learning framework review	DL significantly improves classification accuracy	Dataset dependency issues
Shafiq et al. (2017)	ML-based traffic classification	ANN, SVM, Decision Tree	C4.5 achieved highest accuracy (97.57%)	Limited performance on modern encrypted traffic

IV. CONCLUSION

Machine learning-based network traffic classification has emerged as a fundamental approach for addressing the increasing complexity, volume, and diversity of modern network environments. Traditional traffic classification techniques such as port-based identification and Deep Packet Inspection (DPI) have become less effective due to widespread encryption, dynamic application behaviors, and evolving cyber threats. The reviewed literature clearly demonstrates that machine learning (ML) and deep learning (DL) techniques provide significantly improved performance in terms of accuracy, scalability, and adaptability for network traffic analysis and management. The studies indicate that both classical machine learning models such as Decision Trees, Support Vector Machines (SVM), and Random Forest, as well as advanced deep learning architectures like Convolutional Neural Networks (CNN), Long Short-Term Memory (LSTM), and hybrid models, have been successfully applied in traffic classification tasks. Recent research highlights that hybrid approaches, such as 1DCNN-LSTM and ensemble-based models (LightGBM, XGBoost), achieve superior performance by capturing both spatial and temporal traffic features effectively. These models not only improve classification accuracy but also enhance robustness in dynamic and encrypted network environments. Furthermore, the integration of machine learning with modern networking paradigms such as Software-Defined Networking (SDN) and Internet of Things (IoT) has significantly improved real-time traffic monitoring, anomaly detection, and Quality of Service (QoS) management. ML-based systems enable adaptive decision-making and automated network optimization,

thereby improving overall network efficiency and security. Studies also show that explainable AI techniques such as SHAP and LIME are increasingly being used to improve model transparency and trust in practical deployments. Despite these advancements, several challenges remain, including concept drift, high computational cost of deep learning models, lack of standardized datasets, and difficulties in handling encrypted traffic effectively. These issues highlight the need for more lightweight, adaptive, and scalable models capable of real-time deployment in large-scale and resource-constrained environments. In conclusion, machine learning-based network traffic classification represents a highly effective and evolving solution for modern network management. Continued research focusing on hybrid models, real-time learning, federated learning, and explainable AI will further enhance the efficiency, reliability, and security of future intelligent networking systems.

V. FUTURE SCOPE

- **Development of Real-Time Adaptive Learning Models:** Future research can focus on designing real-time machine learning systems capable of continuously learning from evolving network traffic patterns. This will help address issues like concept drift and ensure sustained classification accuracy in dynamic environments.
- **Integration of Federated Learning Approaches:** Federated learning can be used to enable distributed training of traffic classification models across multiple devices without sharing raw data. This will enhance privacy, reduce communication overhead, and support large-scale network deployments.
- **Explainable Artificial Intelligence (XAI) for Network Security:** The use of XAI techniques such as SHAP and LIME can be expanded to improve transparency in traffic classification models. This will help network administrators understand model decisions and increase trust in automated systems.
- **Lightweight Models for Edge and IoT Networks:** Future work should focus on developing computationally efficient and lightweight ML models suitable for edge devices and IoT networks, where processing power and memory are limited.
- **Advanced Hybrid Deep Learning Architectures:** New hybrid models combining CNN, LSTM, Transformer networks, and graph-based learning can be developed to improve feature extraction and classification accuracy for complex traffic patterns.
- **Encrypted Traffic Classification Techniques:** With increasing encryption, future systems should focus on metadata-based and behavior-based classification techniques that do not rely on payload inspection.
- **Integration with Software-Defined Networking (SDN) and 5G/6G Networks:** ML-based traffic classification can be integrated with SDN controllers and next-generation networks to enable intelligent routing, dynamic resource allocation, and improved Quality of Service (QoS).
- **Use of Reinforcement Learning for Autonomous Network Management:** Reinforcement learning can be explored for self-optimizing networks that automatically adjust traffic policies based on real-time conditions.
- **Development of Standardized and High-Quality Datasets:** Future research should focus on creating diverse, labeled, and realistic datasets that represent modern encrypted and heterogeneous traffic scenarios.
- **Energy-Efficient Machine Learning Models:** There is a need to design low-power ML algorithms to support sustainable computing in large-scale cloud and data center environments.

- **Digital Twin-Based Network Simulation:** The use of digital twin technology can help simulate and test traffic classification models before real-world deployment, reducing risks and improving system reliability.
- **Cross-Domain Security Applications:** Future systems can extend ML-based traffic classification to cybersecurity domains such as intrusion detection, malware detection, and ransomware prevention for holistic network protection.

REFERENCES

1. Kirubavathi, G., Padma Mayuri, B., Pranathasree, S., Alagappan, R., Ajayan, A., Ismael, W. M., & Rehman, A. U. (2026). Ensemble machine learning for proactive android ransomware detection using network traffic. *Scientific Reports*.
2. Pekar, A., Plny, R., & Hynek, K. (2026). Tutorial on Flow-Based Network Traffic Classification Using Machine Learning. *arXiv preprint arXiv:2601.04089*.
3. Wang, Y., & Song, L. (2025). Application and optimization of convolutional neural networks based on deep learning in network traffic classification and anomaly detection. *Informatica*, *49*(14).
4. Zhou, L., & Zhao, D. (2025). Design and Implementation of an Network Traffic Classification System Based on Machine Learning. *Procedia Computer Science*, *259*, 969-976.
5. Eldhai, A. M., Hamdan, M., Abdelaziz, A., Hashem, I. A. T., Babiker, S. F., Marsono, M. N., ... & Jhanjhi, N. Z. (2024). Improved feature selection and stream traffic classification based on machine learning in software-defined networks. *IEEE access*, *12*, 34141-34159.
6. Serag, R. H., Abdalzaher, M. S., Elsayed, H. A. E. A., Sobh, M., Krichen, M., & Salim, M. M. (2024). Machine-learning-based traffic classification in software-defined networks. *Electronics*, *13*(6), 1108.
7. Gómez, J., Riaño, V. H., & Ramirez-Gonzalez, G. (2023). Traffic classification in IP networks through Machine Learning techniques in final systems. *IEEE Access*, *11*, 44932-44940.
8. Long, Z., & Jinsong, W. (2023). Network traffic classification based on a deep learning approach using netflow data. *The Computer Journal*, *66*(8), 1882-1892.
9. Getman, A. I., & Ikonnikova, M. K. (2022). A survey of network traffic classification methods using machine learning. *Programming and Computer Software*, *48*(7), 413-423.
10. Izadi, S., Ahmadi, M., & Rajabzadeh, A. (2022). Network traffic classification using deep learning networks and Bayesian data fusion. *Journal of Network and Systems Management*, *30*(2), 25.
11. Kumar, R., Swarnkar, M., Singal, G., & Kumar, N. (2021). IoT network traffic classification using machine learning algorithms: An experimental analysis. *IEEE Internet of Things Journal*, *9*(2), 989-1008.
12. Labayen, V., Magana, E., Morato, D., & Izal, M. (2020). Online classification of user activities using machine learning on network traffic. *Computer Networks*, *181*, 107557.
13. Mohammed, A. R., Mohammed, S. A., & Shirmohammadi, S. (2019, July). Machine learning and deep learning based traffic classification and prediction in software defined networking. In *2019 IEEE International Symposium on Measurements & Networking (M&N)* (pp. 1-6). IEEE.
14. Rezaei, S., & Liu, X. (2018). Deep learning for encrypted traffic classification: An overview. *arXiv preprint arXiv:1810.07906*.
15. Shafiq, M., Yu, X., & Wang, D. (2017, June). Network traffic classification using machine learning algorithms. In *International conference on intelligent and interactive systems and applications* (pp. 621-627). Cham: Springer International Publishing.