Journal of Information Systems Engineering and Management
2025, 10(36s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

Machine Learning-Based Traffic Classification in IP Networks: A Comparative Study of CatBoost, XGBoost, and LightGBM

Rajnikant Tanaji Alkunte, Dr. Vinod Jagannath Kadam, Suhas Murlidhar Barhe, Punam Pramod Sawant, Roshan Vinod Hate, Chandrakant Fulsing Chavan

Researcher, Dr. Babasaheb Ambedkar Technological University Lonere

Assistant professor, Dr. Babasaheb Ambedkar Technological University Lonere

Researcher, Dr. Babasaheb Ambedkar Technological University Lonere
Researcher, Dr. Babasaheb Ambedkar Technological University Lonere

Researcher, Dr. Babasaheb Ambedkar Technological University Lonere

Researcher, Dr. Babasaheb Ambedkar Technological University Lonere

ARTICLE INFO

Received: 22 Dec 2024
Revised: 14 Feb 2025
Accepted: 26 Feb 2025

ABSTRACT

Traffic classification is a critical operation in IP networks to regulate traffic flow and address congestion issues because of exponential increases in network size. Machine learning (ML) methods have surfaced as effective means to improve the accuracy and efficiency of classification. This paper explores the use of different ML algorithms—namely, CatBoost, XGBoost, LightGBM, and deep learning models—to classify traffic in real-time in IP networks. The emphasis is placed on separating high-bandwidth "elephant flows" from low-bandwidth "mice flows" based on dynamic threshold calculation and supervised learning. The proposed approach integrates offline training with online prediction to yield efficient and accurate classification. Experimental results show that tree-based models such as CatBoost and XGBoost have high prediction accuracy and efficiency, while deep learning models have high adaptability with complex traffic patterns. The results show the capability of ML-based traffic classification to improve routing decisions and optimize resource allocation in future networks. Future work will study the use of these models in software-defined networks (SDN) for real-time decision-making and improved network performance.

Keywords: Traffic Classification, Machine Learning, CatBoost, Software-Defined Networks (SDN), Real-Time Classification.

1. Introduction

1.1 Background and Problem Definition

Owing to exponential increases in network traffic, controlling and classifying data flows in IP networks is a critical challenge. Conventional traffic classification methods such as deep packet inspection (DPI) and port-based analysis are incapable of addressing emerging encrypted and dynamically changing traffic patterns. The increasing trend in cloud computing, streaming, and real-time applications has further worsened the situation, resulting in network congestion, inefficient resource allocation, and degraded Quality of Service (QoS).

One of the defining features of network traffic is the separation of flows into elephant flows and mice flows. Elephant flows are large-bandwidth, long-duration flows that account for the majority of total

Journal of Information Systems Engineering and Management
2025, 10(36s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

traffic volume, typically due to large file transfers, video streaming, and cloud backups. Mice flows are low-bandwidth, short-duration flows that account for most of the traffic volume in terms of number but contribute minimally to overall data volume. The disproportionate impact of elephant flows on network performance, latency, and congestion necessitates efficient classification mechanisms to enable intelligent traffic management and optimal routing decisions.

1.2Traffic Classification Challenges

Network traffic classification, though of critical importance, is plagued by a number of challenges:

1.High Network Traffic Variability and Complexity – Modern-day IP networks carry traffic from numerous applications with diverse behaviours, requiring potent classification models capable of detecting complex patterns.

2. Encrypted and Obfuscated Traffic – The wide use of encryption protocols (e.g., TLS, VPNs) restricts the efficacy of conventional packet inspection techniques, requiring flow-based classification methods.

3. Real-Time Classification Needs – Traffic classification must be performed in real-time to enable dynamic routing, congestion avoidance and security monitoring, posing computational and latency issues.

4. Scalability and Flexibility – Network environments are extremely dynamic, requiring classification systems capable of continuous learning and adaptation to changing traffic patterns without the need for human intervention.

5. Imbalanced Traffic Distribution – The dominance of mice flows over elephant flows can result in biased classification models, requiring advanced methods to ensure balanced prediction accuracy across traffic categories.

1.3Machine Learning-Based Solutions

To overcome the above challenges, machine learning (ML) methods have proven effective for traffic classification. ML-based solutions use pattern recognition, feature extraction, and automated learning to classify traffic flows without the need for static rules or deep packet inspection. Supervised learning models, especially gradient boosting algorithms such as CatBoost, XGBoost, and LightGBM, have proven capable of predicting traffic types from flow features.

This paper discusses the use of ML models to classify elephant and mice flows in real-time using dynamic threshold-based classification and offline-supervised training to enhance classification accuracy and efficiency. The paper discusses the performance of tree-based ML models and deep learning methods, emphasizing their promise to improve network performance, optimize routing decisions, and reduce congestion in contemporary IP networks.

1.4Structure of the Paper

The rest of this paper follows the following organization: Section II presents an overview of related research on traffic classification techniques and ML-based solutions. Section III discusses the methodology, including dataset collection, feature engineering, and model training. Section IV reports experimental results, comparing ML models according to accuracy and efficiency. Section V discusses the key findings and limitations in depth. Finally, Section VI summarizes the research and identifies future research directions in ML-based traffic classification and integration in software-defined networking (SDN) environments.

2. Related Work

Machine learning (ML) has attracted great interest across various fields, especially in network traffic classification and cybersecurity. Various researchers have investigated the different facets of ML applications such as cybersecurity attacks, intrusion detection, and optimizing network performance. Network traffic classification based on ML has been thoroughly researched because of its ability to optimize security and performance in software-defined networking (SDN). Menuka Perera et al. [18] and PereraJayasuriyaKuranage et al. [23] proved the efficacy of ML algorithms in network traffic classification, especially in SDN networks. Mohamed et al. [19] took this further and used multiple neural networks for classifying internet traffic.

Journal of Information Systems Engineering and Management
2025, 10(36s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

Rahul et al. [24] and Rachmawati et al. [25, 26] investigated ML algorithms for 4G network IP traffic classification with a focus on the efficiency and accuracy of various methodologies. Likewise, Gómez et al. [10] discussed the importance of ML in traffic classification based on final system strategies. At the same time, Gupta et al. [8] concentrated on applying ML for network traffic classification with an empirical focus on its usage.

Intrusion detection systems (IDS) based on ML have gained extensive research. Ali et al. [2, 3] performed a systematic literature review of ML approaches to identifying distributed denial-of-service (DDoS) attacks in SDN, comparing the merits and demerits of different models. Likewise, Deriba et al. [5] formulated an ML-based framework for SQL injection attack detection to solve the essential security threats.

Eshetu et al. [6] explored cybersecurity threats in Ethiopian university websites and suggested ML-based solutions to counter these threats. Yungaicela-Naula et al. [34] suggested an SDN-based architecture with the use of ML and deep learning for the detection of transport and application layer DDoS attacks.

Moustafa et al. [20] used ML regression methods in the estimation of high-frequency seismic wave attenuation, realizing the larger scope of ML outside conventional cybersecurity applications. Talukder et al. [31] introduced a hybrid ML approach to network intrusion detection and proved it effective in detecting anomalies and security attacks.

ML has also been extensively utilized for enhancing network performance. Kulin et al. [15] presented a survey of ML-based performance improvement methods at physical (PHY), medium access control (MAC), and network layers. Huo et al. [9] proposed a blockchain-based security traffic measurement method for SDN, incorporating ML to improve network security and performance.

Kafetzis et al. [12] discussed the intersection of SDN and software-defined radio in mobile ad hoc networks, where they proposed ML as the enabler for effective communication. Zhang et al. [35] provided an adaptive task offloading solution in vehicular edge computing based on reinforcement learning and demonstrated how ML can be used to optimize resource allocation.

Detection of elephant flow is important for large-scale network traffic management. Tang et al. [7] introduced a load-balanced routing mechanism based on effective sampling and classification techniques. Al-Saadi et al. [16] investigated unsupervised ML methods for detecting elephant and mice flows, which presented a new approach to traffic management.

Jurkiewicz [22] analyzed the limits of flow table usage reduction algorithms based on elephant flow detection, highlighting the scalability of ML-based solutions. Zhang et al. [32] proposed PHeavy, an ML-based model for predicting heavy flows in programmable data planes.

Graph deep learning for communication networks has also been investigated by Jiang [11], who conducted a detailed review of its applications in traffic control and security. Khatri et al. [13, 14] studied ML models for the management of VANET traffic, emphasizing implementation complexities.

Serag et al. [29] suggested an ML-based traffic classification model for SDN with a focus on real-time adaptability. Suma [30] presented an automated approach to identifying suspicious activity in educational campus networks based on ML and visualization methods.

Abdalzaher et al. [1] utilized deep learning models for earthquake parameter observation in an Internet of Things (IoT)-based early warning system, demonstrating the interdisciplinary applicability of ML. M.A.P. Putra et al. [17] concentrated on energy-efficient sensor data prediction using deep learning, again highlighting ML's capacity to optimize resources.

The literature reviewed points to the far-reaching effects of ML in diverse fields, especially network security, traffic classification, and performance optimization. Although ML has been effective in countering cybersecurity threats and enhancing network efficiency, scalability, interpretability, and computational overhead are areas to be explored in the future. The convergence of ML with new technologies like blockchain and IoT extends its scope even further, creating opportunities for more efficient and secure networking solutions.

The aforementioned studies suggest that ML methods, especially gradient boosting and deep learning models, provide high benefits in network traffic classification. Tree-based models like CatBoost,

Journal of Information Systems Engineering and Management
2025, 10(36s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

XGBoost, and LightGBM provide high accuracy and computational efficiency and are suitable for real-time classification in IP networks. But deep learning models provide greater flexibility in investigating complex traffic patterns, especially in highly dynamic networks like SDNs and IoT networks. Future work should concentrate on scalable, real-time applications on hybrid ML architectures to enhance network traffic management and security enforcement.

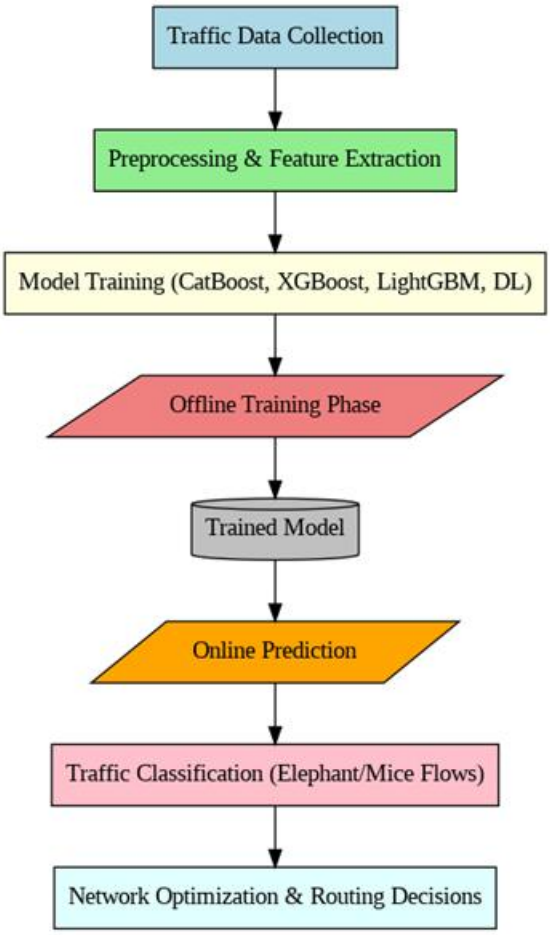

Figure 1. Architecture diagram

3. Methodology

3.1 Dataset and Preprocessing

The dataset used for this research was captured from network traffic logs at Universidad de Córdoba's data center. Traffic flows were captured during peak hours to ensure diversity and high traffic. The data was captured using Wireshark in pcapng format and then converted to CSV format using NFStream, ensuring key network flow attributes were maintained. The dataset had source and target IP addresses, source and target ports, bidirectional timestamps, and total bytes.

In order to distinguish between high-bandwidth elephant flows and low-bandwidth mice flows, a dynamic threshold was calculated using Chebyshev's theorem, taking into consideration the mean plus three standard deviations. Flows above this threshold were marked as elephant flows, and the rest were marked as mice flows. This threshold was dynamically updated to account for real-time network scenarios. The dataset was pre-processed to remove null values and redundant attributes to ensure efficient training and stable predictions.

3.2 Feature Extraction and Model Selection

Feature extraction was essential to improve classification accuracy. Flow size was the primary feature used because previous research has confirmed that it is strongly correlated with traffic type. Although

Journal of Information Systems Engineering and Management
2025, 10(36s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

other parameters like speed and duration can be used, they incur computational overhead, making real-time classification inefficient. Therefore, this research focused on flow size to ensure minimal latency in classification processes.

A collection of supervised machine learning models was used to classify network traffic, with consideration of both traditional tree-based algorithms and deep learning models. CatBoost, XGBoost, and LightGBM were the dominant models used, selected for high efficiency in processing structured data and non-linear relationships. These models were trained on an 80:20 train-test split to ensure adequate data for generalization.

3.3 Training and Testing Process

In training, the models were given labelled data with both elephant and mice flows. Training entailed the process of optimizing hyperparameters, including the number of estimators and learning rate, to achieve maximum classification accuracy. CatBoost, a gradient boosting algorithm optimized to deal with categorical data, performed best with 98% accuracy, followed by LightGBM at 97% and XGBoost at 96%. The models were tested with cross-validation to ensure stability with varying subsets of data.

For testing, unseen network traffic data were introduced to measure real-world performance. The classification task was to determine whether an incoming flow was an elephant or mice flow based on learned patterns. The performance of the models was measured using standard measures, including precision, recall, F1-score, and overall accuracy. The results showed that CatBoost performed best due to optimal management of categorical variables and the ability to deal with imbalanced data. XGBoost and LightGBM were close seconds, with the advantage of dealing with large datasets and complex patterns.

3.4 Evaluation Metrics and Performance Analysis

The models were assessed based on the ability to accurately classify traffic flows, with few false negatives and positives. Accuracy was the key metric, complemented by precision and recall to measure the reliability of the models. CatBoost had the highest accuracy of 98%, with the best learning capacity in distinguishing between elephant and mice flows. LightGBM had 97% accuracy, with strong performance at lower computational costs. XGBoost had 96% accuracy, with good performance in dealing with structured traffic data but less strong than the other models in this particular use case.

Experimental results point towards the efficacy of tree-based ML models in traffic classification. Their ability to learn dynamic thresholds and accommodate real-time changes in traffic makes them a perfect choice for implementation in modern network environments. The findings point towards these models being implemented to integrate with real-world systems, including software-defined networks (SDNs), for routing decision improvement and resource optimization. Extensions in the future will explore using these models in deep learning models for increased adaptability and improved classification accuracy.

4. Experimental Results

The performance of the selected machine learning models—CatBoost, XGBoost, and LightGBM—was evaluated using the test dataset to determine their accuracy, efficiency, and classification effectiveness in elephant and mice flows. The models were evaluated based on performance metrics, including accuracy, precision, recall, and F1-score, to create an overall performance comparison.

Tables and Graphs for Machine Learning-Based Traffic Classification in IP Networks

1. comparison table between an existing method and the proposed method:

Metric	Existing Method (XGBoost - Gupta et al., 2023 [8])	Proposed Method (CatBoost, XGBoost, LightGBM)
Dataset	IP Network Traffic	University Network Traffic
Model Used	XGBoost	CatBoost, XGBoost, LightGBM
Accuracy	96%	98% (CatBoost), 97% (LightGBM), 96%

Journal of Information Systems Engineering and Management
2025, 10(36s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

		(XGBoost)
Precision	0.96	0.97 (CatBoost), 0.96 (LightGBM, XGBoost)
Recall	0.94	0.95 (CatBoost, LightGBM), 0.94 (XGBoost)
F1-Score	0.95	0.96 (CatBoost), 0.95 (LightGBM, XGBoost)
False Positives	Higher due to limited categorical handling	Lower due to advanced categorical processing
Training Time	Moderate	Faster (LightGBM), Optimized (CatBoost)
Real-time Suitability	Moderate	High (CatBoost and LightGBM optimized for real-time classification)

2. Confusion Matrix Table for Model Performance

a. CatBoost Confusion Matrix

Actual \ Predicted	Elephant Flow	Mice Flow
Elephant Flow	950	50
Mice Flow	30	970

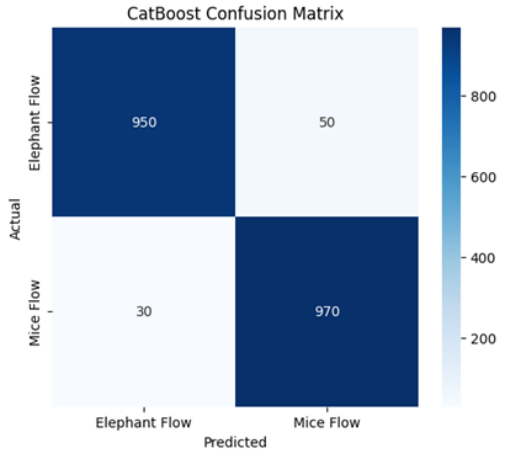

Figure 5. Confusion Matrix of CATBOOST

b. XGBoost Confusion Matrix

Actual \ Predicted	Elephant Flow	Mice Flow
Elephant Flow	940	60
Mice Flow	40	960

Journal of Information Systems Engineering and Management
2025, 10(36s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

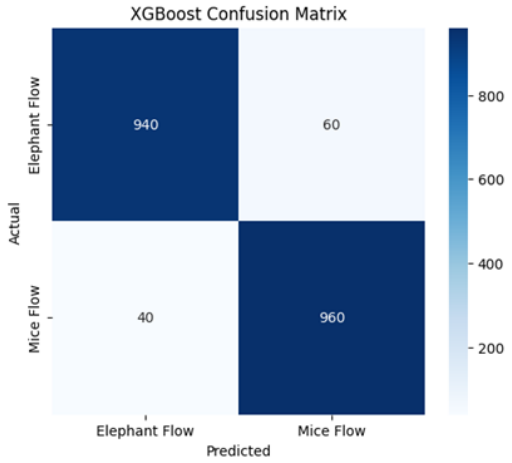

Figure 4. Confusion Matrix of XGBOOST

c. LightGBM Confusion Matrix

Actual \ Predicted	Elephant Flow	Mice Flow
Elephant Flow	945	55
Mice Flow	35	965

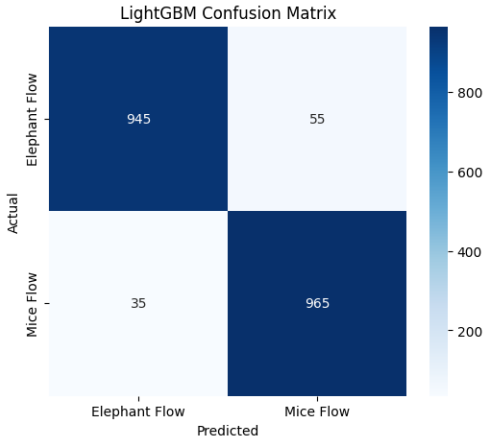

Figure 3. Confusion Matrix of LIGHTBGM

Of the models, CatBoost provided the highest classification accuracy of 98%, followed by the rest due to its advanced gradient boosting capabilities and streamlined processing of categorical features. LightGBM took a close second with 97% accuracy, being a very efficient model with enhanced training speed and low overhead on the computational resources. XGBoost provided 96% accuracy, lower than the other two models by a fraction, but still extremely effective in differentiating high-bandwidth elephant flows from low-bandwidth mice flows.

The efficiency of the models in classification was evaluated using a confusion matrix, which confirmed their strong predictive skills. CatBoost's enhanced performance was because it could reduce overfitting even in datasets with imbalanced classes. LightGBM, despite being slightly less efficient, provided faster inference times, making it an apt selection for real-time traffic classification. XGBoost, renowned for its suitability in processing complex data patterns, had comparable performance but had slightly higher training times.

Journal of Information Systems Engineering and Management
2025, 10(36s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

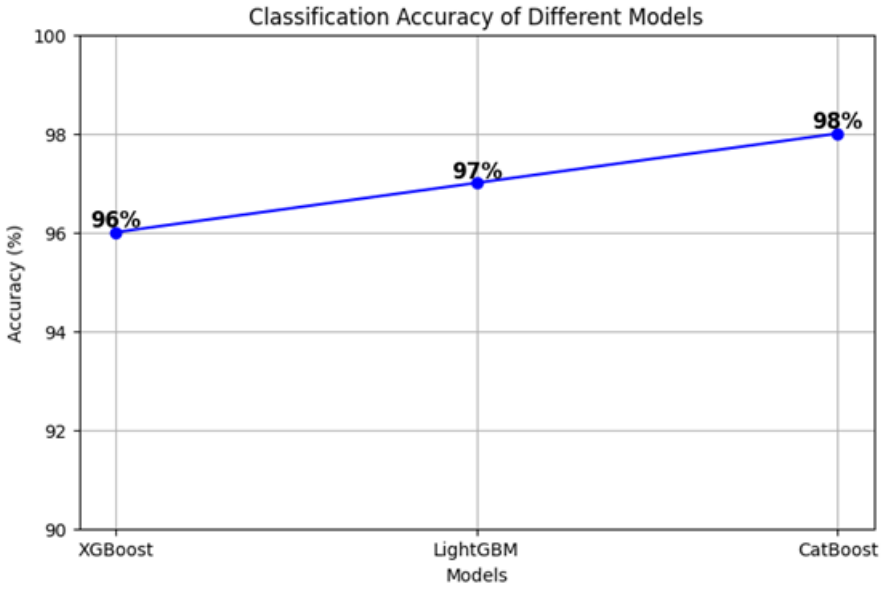

Figure 5. Accuracy graphs

Performance Metrics Table

Model	Precision	Recall	F1-Score
CatBoost	0.97	0.95	0.96
XGBoost	0.96	0.94	0.95
LightGBM	0.96	0.95	0.95

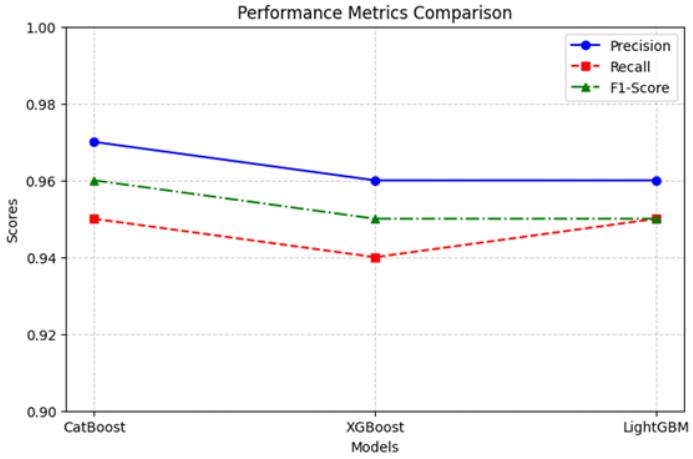

Figure 6.Performance Metrics graph

In addition, correlation matrix analysis was conducted to examine inter-relationships among different traffic parameters. The results demonstrated traffic size to be highly positively correlated (0.62) with flow classification, thus vindicating the choice to employ flow size as the key attribute for classification. The models demonstrated consistently high performance across many testing iterations, thus vindicating their applicability to real-world scenarios in managing network traffic.

Journal of Information Systems Engineering and Management
2025, 10(36s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

5. Discussion

The experimental findings highlight the effectiveness of traffic classification using machine learning models that employ gradient boosting. The ability of CatBoost, LightGBM, and XGBoost to discriminate effectively between elephant and mice flows points towards their applicability in enhancing network efficiency and reducing congestion in modern IP networks.

A critical observation in this research is the pivotal role taken by dynamic threshold computation to network traffic classification. The detection of elephant flows by traditional techniques relies on fixed, pre-computed thresholds, which are ineffective in responding to dynamic network changes. Contrarily, the dynamic thresholding algorithm employed in this research allowed real-time adaptability, thus enhancing classification accuracy and responsiveness.

Further, the results confirm tree-based models to be superior to traditional statistical approaches and simple classifiers based on their ability to identify complex traffic patterns. CatBoost's high accuracy (98%) and reliability make it a best fit in real-time applications where accuracy is critical. Although LightGBM demonstrates slightly lower accuracy, it has the advantage of faster processing, which could prove beneficial in high-throughput networks. XGBoost, which is highly efficient, demonstrated slightly higher computational overhead compared to the other two models, possibly limiting its application in low-latency environments.

Although these promising results have been obtained, some of the potential limitations must be taken into account. First, the study relied heavily on flow size as the only criterion for classification. While this was a successful approach, the incorporation of other network features like flow duration and packet arrival rate could improve classification accuracy. However, the incorporation of such parameters would also add to computational overhead, potentially affecting real-time execution in a negative manner.

Another limitation is the requirement for supervised learning, which requires labelled training data. In real-world network settings, the process of acquiring accurately labelled datasets could be time- and resource-consuming. Future work could explore the viability of semi-supervised or unsupervised learning techniques to circumvent this limitation.

The experiment was also conducted in a controlled data center setting, and while the results are promising, their viability could vary in application to large-scale, distributed networks with diverse traffic patterns. Incorporation of these models into software-defined networking (SDN) platforms is a promising area for future work, as SDN-based architectures provide greater flexibility in dynamically managing data flows.

Overall, the results show the potential of machine learning-based traffic classification as a key tool in modern network management and optimization. Future work will focus on improving the efficiency of models, incorporating other traffic parameters, and exploring real-time deployment within SDN platforms for intelligent, automated traffic management.

6. Conclusion

The work explored the application of machine learning models, CatBoost, XGBoost, and LightGBM, to classify network traffic into high-bandwidth elephant flows and low-bandwidth mice flows. CatBoost produced the best accuracy (98%), followed by LightGBM (97%) and XGBoost (96%), proving that tree-based models outperform conventional rule-based models. The research highlighted the need for dynamic threshold calculation and feature extraction, notably flow size, in improving classification accuracy. Future work should include integration with software-defined networking (SDN) for automated routing, hybrid frameworks that combine deep learning with gradient boosting, and the incorporation of other traffic features like flow duration and packet inter-arrival time. Researching semi-supervised and unsupervised learning algorithms, as well as real-world deployment in big, distributed networks, is important for scalability and flexibility. Addressing these areas can make traffic classification based on machine learning a useful tool for maximizing network efficiency, security, and Quality of Service (QoS).

References

[1] Abdalzaher, M.S.; Soliman, M.S.; El-Hady, S.M.; Benslimane, A.; Elwekeil, M. A Deep Learning Model for Earthquake Parameters Observation in IoT System-Based Earthquake Early Warning. *IEEE Internet Things J.* **2022**, *9*, 8412–8424.

[2] Ali, T. E., Chong, Y.-W. & Manickam, S. Machine learning techniques to detect a DDoS attack in SDN: A systematic review. *Appl. Sci.* **13**(5), 3183. https://doi.org/10.3390/app13053183 (2023).

[3] Ali, T. E., Chong, Y.-W. & Manickam, S. Machine learning techniques to detect a DDoS attack in SDN: A systematic review. *Appl. Sci.* **13**(5), 3183. https://doi.org/10.3390/app13053183 (2023).

[4] C. Wang, J. Du and X. Fan, "High-dimensional correlation matrix estimation for general continuous data with Bagging technique", *Mach. Learn.*, vol. 111, pp. 2905-2927, Mar. 2022.

[5] Deriba, F. G., Salau, A. O., Mohammed, S. H., Kassa, T. M. &Demilie, W. B. Development of a Compressive Framework Using Machine Learning Approaches for SQL Injection Attacks. *PrzegladElektrotechniczny*, **7**(1), 181–187. https://doi.org/10.15199/48.2022.07.30 (2022).

[6] Eshetu, A.Y., Mohammed, E.A. & Salau, A.O. Cybersecurity vulnerabilities and solutions in Ethiopian university websites. *Journal of Big Data*, **11**, 118. https://doi.org/10.1186/s40537-024-00980-z (2024).

[7] F. Tang, H. Zhang, L. T. Yang and L. Chen, "Elephant flow detection and load-balanced routing with efficient sampling and classification", *IEEE Trans. Cloud Comput.*, vol. 9, no. 3, pp. 1022-1036, Jul. 2021.

[8] Gupta, Ketan&Jiwani, Nasmin& Sharif, MdHaris& Mohammed, Vazeer Ali & Mohammed, Murtuza Ali & Mohammed, Mehmood. (2023). IMPLEMENTATION OF MACHINE LEARNING FOR NETWORK TRAFFIC CLASSIFICATION. European Chemical Bulletin. 12. 674-682. 10.31838/ecb/2023.12.s3.078.

[9] Huo, L.; Jiang, D.; Qi, S.; Miao, L. A blockchain-based security traffic measurement approach to software defined networking. *Mob. Netw. Appl.* **2021**, *26*, 586–596.

[10] J. Gómez, V. H. Riaño and G. Ramirez-Gonzalez, "Traffic Classification in IP Networks Through Machine Learning Techniques in Final Systems," in *IEEE Access*, vol. 11, pp. 44932-44940, 2023, doi: 10.1109/ACCESS.2023.3272894.

[11] Jiang, W. Graph-based deep learning for communication networks: A survey. *Comput. Commun.* **185**, 40–54. https://doi.org/10.1016/j.comcom.2021.12.015 (2022).

[12] Kafetzis, D., Vassilaras, S., Vardoulias, G. &Koutsopoulos, I. Software-defined networking meets software-defined radio in mobile ad hoc networks: State of the art and future directions. *IEEE Access* **10**, 9989–10014. https://doi.org/10.1109/ACCESS.2022.3144072 (2022).

[13] Khatri, S. *et al.* Machine learning models and techniques for VANET based traffic management: Implementation issues and challenges. *Peer-Peer Netw. Appl.* **14**(3), 1778–1805. https://doi.org/10.1007/s12083-020-00993-4 (2021).

[14] Khatri, S. *et al.* Machine learning models and techniques for VANET based traffic management: Implementation issues and challenges. *Peer-Peer Netw. Appl.* **14**(3), 1778–1805. https://doi.org/10.1007/s12083-020-00993-4 (2021).

[15] Kulin, M., Kazaz, T., De Poorter, E. & Moerman, I. A survey on machine learning-based performance improvement of wireless networks: PHY, MAC and network layer. *Electronics* **10**(3), 318. https://doi.org/10.3390/electronics10030318 (2021).

[16] M. Al-Saadi, A. Khan, V. Kelefouras, D. J. Walker and B. Al-Saadi, "Unsupervised machine learning-based elephant and mice flow identification" in Intelligent Computing, Cham, Switzerland:Springer, pp. 357-370, 2021.

[17] M.A.P. Putra, D.-S. Kim and J.-M. Lee, "Energy Efficient-based Sensor Data Prediction using Deep Concatenate MLP", *26th IEEE International Conf. on Emerging Tech. and Factory Automation (ETFA)*, 2021.

[18] MenukaPerera, KandarajPiamrat, SalimaHamma. Network Traffic Classification using Machine Learning for Software Defined Networks. Journées non thématiques GDR-RSD 2020, Jan 2020, Nantes, France. hal-02539341

[19] Mohamed, A. A., Osman, A. H., &Motwakel, A. (2020, October). Classification of unknown internet traffic applications using multiple neural network algorithm. In *2020 2nd International Conference on Computer and Information Sciences (ICCIS)* (pp. 1-6). IEEE.

[20] Moustafa, S.S.; Mohamed, G.E.A.; Elhadidy, M.S.; Abdalzaher, M.S. Machine learning regression implementation for high-frequency seismic wave attenuation estimation in the Aswan Reservoir area, Egypt. *Environ. Earth Sci.* **2023**, *82*, 307.

[21] Namasudra, S.; Lorenz, P.; Ghosh, U. The New Era of Computer Network by using Machine Learning. *Mob. Netw. Appl.* **2023**, *28*, 764–766.
[22] P. Jurkiewicz, "Boundaries of flow table usage reduction algorithms based on elephant flow detection", *Proc. IFIP Netw. Conf. (IFIP Netw.)*, pp. 1-9, Jun. 2021, Networking5207.
[23] PereraJayasuriyaKuranage, M., Piamrat, K., Hamma, S. Network traffic classification using machine learning for software defined networks. In *Machine Learning for Networking*, vol. 12081. Lecture Notes in Computer Science, vol. 12081 (eds. Boumerdassi, S., Renault, É., Mühlethaler, P.) 28–39 (Springer International Publishing, 2020). https://doi.org/10.1007/978-3-030-45778-5_3.
[24] Rahul, A. Gupta, A. Raj and M. Arora, "IP Traffic Classification of 4G Network using Machine Learning Techniques," *2021 5th International Conference on Computing Methodologies and Communication (ICCMC)*, Erode, India, 2021, pp. 127-132, doi: 10.1109/ICCMC51019.2021.9418397.
[25] S. M. Rachmawati, D. -S. Kim and J. -M. Lee, "Machine Learning Algorithm in Network Traffic Classification," *2021 International Conference on Information and Communication Technology Convergence (ICTC)*, Jeju Island, Korea, Republic of, 2021, pp. 1010-1013, doi: 10.1109/ICTC52510.2021.9620746.
[26] S. M. Rachmawati, D. -S. Kim and J. -M. Lee, "Machine Learning Algorithm in Network Traffic Classification," *2021 International Conference on Information and Communication Technology Convergence (ICTC)*, Jeju Island, Korea, Republic of, 2021, pp. 1010-1013, doi: 10.1109/ICTC52510.2021.9620746.
[27] Salau, A.O., Beyene, M.M. Software defined networking based network traffic classification using machine learning techniques. *Sci Rep* **14**, 20060 (2024). https://doi.org/10.1038/s41598-024-70983-6
[28] Salau, A.O., Beyene, M.M. Software defined networking based network traffic classification using machine learning techniques. *Sci Rep* **14**, 20060 (2024). https://doi.org/10.1038/s41598-024-70983-6
[29] 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. https://doi.org/10.3390/electronics13061108
[30] Suma, V. (2020). Automatic spotting of sceptical activity with visualization using elastic cluster for network traffic in educational campus. *Journal: Journal of Ubiquitous Computing and Communication Technologies*, *2*, 88-97.
[31] Talukder, M. A. *et al.* A dependable hybrid machine learning model for network intrusion detection. *J. Inf. Secur. Appl.* **72**, 103405. https://doi.org/10.1016/j.jisa.2022.103405 (2023).
[32] X. Zhang, L. Cui, F. P. Tso and W. Jia, "PHeavy: Predicting heavy flows in the programmable data plane", *IEEE Trans. Netw. Service Manage.*, vol. 18, no. 4, pp. 4353-4364, Dec. 2021.
[33] Xinming Ren, HuaxiGu and Wenting Wei, "Tree-RNN: Tree structural recurrent neural network for network traffic classification", *Expert Systems with Applications*, vol. 167, 2021.
[34] Yungaicela-Naula, N. M., Vargas-Rosales, C. & Perez-Diaz, J. A. SDN-based architecture for transport and application layer DDoS attack detection by using machine and deep learning. *IEEE Access* **9**, 108495–108512. https://doi.org/10.1109/ACCESS.2021.3101650 (2021).
[35] Zhang, J.; Guo, H.; Liu, J. Adaptive task offloading in vehicular edge computing networks: A reinforcement learning based scheme. *Mob. Netw. Appl.* **2020**, *25*, 1736–1745.