UDC 004
Original Scientific Paper
https://doi.org/10.62598/JVA.11.1.5.16

Received: October, 2024
Accepted for publishing: June 17, 2025

ANALYSIS OF COMPUTER NETWORK USER'S ACTIVITIES USING SUPPORT VECTOR MACHINE (SVM) AND LONG SHORT-TERM MEMORY (LSTM) NETWORK

Olajide, Olajide Blessing, *Engr. Dr., orcid: https://orcid.org/0000-0002-6498-2475 , Department of Computer Engineering, Federal University Wukari,Wukari, Nigeria, olajide@fuwukari.edu.ng*

Adeogun, Adetoyese Elijah, *Department of Computer Science, Ladoke Akintola University of Technology, Ogbomoso, Nigeria, adeogunadetoyese.aae@gmail.com*

Olayinka, Temitope Omotayo, *Department of Electrical and Computer Engineering, University of Maine, Orono, United State, temitope.olayinka@maine.edu*

Balogun, Emmanuel, *ICT Unit, Ajayi Crowther University, Oyo, Nigeria, eo.balogun@acu.edu.ng*

Areola, Raphael Ibukun, *Department of Electrical Power Engineering, Durban University of Technology, Durban, 4001, South Africa, 22494951@dut4life.ac.za*

Falade, Mutiu Bolarinwa, *Department of Computer Engineering, Federal University Wukari,Wukari, Nigeria, mutiufalade@gmail.com*

Kemi Ogunniyi Olufunke, *Department of Computer Science, Ladoke Akintola University of Technology, Ogbomoso, Nigeria, olufayokemi28@gmail.com*

Abstract:

The rapid growth of number of network users have led to a significant rise in network traffic. Analysing user activities within computer networks is essential for optimizing performance, enhancing security, and improving user experience. This study explores the application of machine learning techniques, specifically Support Vector Machine (SVM) and Long Short-Term Memory (LSTM) networks, to analyse computer network user activities. SVM is employed for its effectiveness in binary classification tasks and its ability to handle high-dimensional data, making it suitable for identifying distinct user activities based on network traffic *patterns. Conversely, LSTM networks was utilized to capture temporal dependencies in sequential data, allowing for the prediction of future user actions based on their historical activities. The precision, recall, FI-score and accuracy results for SVM model for analysing computer network user's activities are 96.00, 99.00, 98.00 and 95.40 respectively. While the precision, recall, FI-score and accuracy results for LSTM model for analysing computer network user's activities are 90.00, 91.00, 91.21 and 93.50 respectively. Trailing to this, the SVM has a better performance than the LSTM model. Therefore, this research contributes to the field of network analytics by offering insights that will improve network management strategies, resource allocation, and security measures.*

***Keywords**: network, user, activities, SVM, LSTM, management.*

1. Introduction

The dramatic surge in the demand for internet and multimedia services, alongside the exponential growth in the number of internet users throughout the years, has made the management of network resources increasingly complex. (Rodjas, 2020: 431). In other to meet the increasing demand to ensure a good quality of service, managing network resources has become a continuous struggle for network providers due to the aforementioned factors (Rodjas, 2020: 436). The outbreak of COVID-19 also contributed significantly to the explosion in digital media consumption all over the world in conjunction with both wired and wireless Internet connectivity speeds and bandwidth. Due to the restrictions on both indoor and outdoor activities imposed by COVID-19, there was a notable increase in the demand for video streaming services following the shutdown of movie theaters. This surge has provided consumers with greater ease of access to the media content they desire, available at any time and place, typically at reduced prices (Rodjas, 2020: 446). Over-the-top (OTT) applications have swiftly become a primary tool through which consumers now have access to these media contents. This has drastically influenced user's behaviour as users are now in high demand for services that are more user-friendly and respond to their desires. Users no longer depend on their television sets for entertainment rather, with OTT applications, a new era of binge-watching has become a preferred choice. OTT applications have supplanted traditional television as the primary means of delivering media content via the internet, utilizing the infrastructure established by network operators. (Elijah et al., 2024: 238). The services provided by these applications utilize a significant amount of network resources, resulting in a substantial impact on the operation and administration of the network (Jain, 2021: 260).

Telecommunication providers typically offer data plans with consumption limits, and service degradation is a common strategy employed to regulate the volume of data users can transfer over time. Upon exceeding their allocated data limit, the network provider will curtail the user's bandwidth based on their consumption to ensure the network operates efficiently (Rodjas, 2020: 440). This mechanism in itself does not apply limit with an exception, rather the limit applied affects the overall activity such a user might be performing on the network. And this makes this mechanism less efficient. The utilization of substantial network resources by OTT applications to provide services like audio, video, and various other functionalities significantly affects network operation and management due to the high volume of traffic generated (Rodjas, 2020: 438). Utilizing service degradation as a strategy to mitigate the excessive traffic produced by a specific OTT application impacts the performance of all other applications utilized by the user, without considering the user's behavior regarding which OTT application is causing the high traffic. Therefore, the execution of service degradation might lead to infringements of the service level agreements that the Internet Service Provider may have with other OTT service providers (Rodjas et al., 2019: 582). Due to the large consumption of network resources of OTT applications, efficient management of the available resources based on the data usage behaviour of users becomes a necessity (Cisco, 2017: 1). Studying the user's preference and usage behaviours trend on the computer network to decipher the OTT application which consumes beyond the limit of the allocated network resources to perform informed service degradation to the OTT application without affecting other OTT applications as the traditional service degradation would, will ensure that the service level agreement between OTT application providers and ISPs is not breached. Machine learning algorithms can predict future network demands, allowing for practive resources allocation and capacity planning, which minimizes overprovision and downtime. Aditionally, ML can automate tasks like anomaly detection and security treat identification, reducing manual effort and improving response time. Hence, this study performed a comparative profiling of network user's activities using Support Vector Machine (SVM) and Long-Short-Term Memory (LSTM) Network for effective selective OTT application network resources consumption management.

1.1 Support Vector Machine (SVM)

A support vector machine (SVM) is a type of supervised machine learning model that utilizes classification algorithms to solve problems involving two categories (Okpor et al., 2024: 218). The SVM algorithm is founded on the idea of decision planes, employing hyperplanes to differentiate between a set of objects. After being trained on a labeled dataset, SVM can classify new data effectively. The primary objective of the SVM algorithm is to establish a decision boundary that divides n-dimensional space into distinct classes, facilitating the accurate categorization of future data points. This optimal decision boundary is known as a hyperplane, which is determined by selecting the extreme points or vectors that contribute to its formation. These critical instances are termed support vectors, which is the origin of the algorithm's name, Support Vector Machine. The classification of two different categories using a decision boundary is illustrated in Figure 1. In contrast to more recent algorithms such as neural networks, SVM has two key advantages: faster processing speed and better performance with a limited number of labeled samples (Madugu et al., 2023:108).

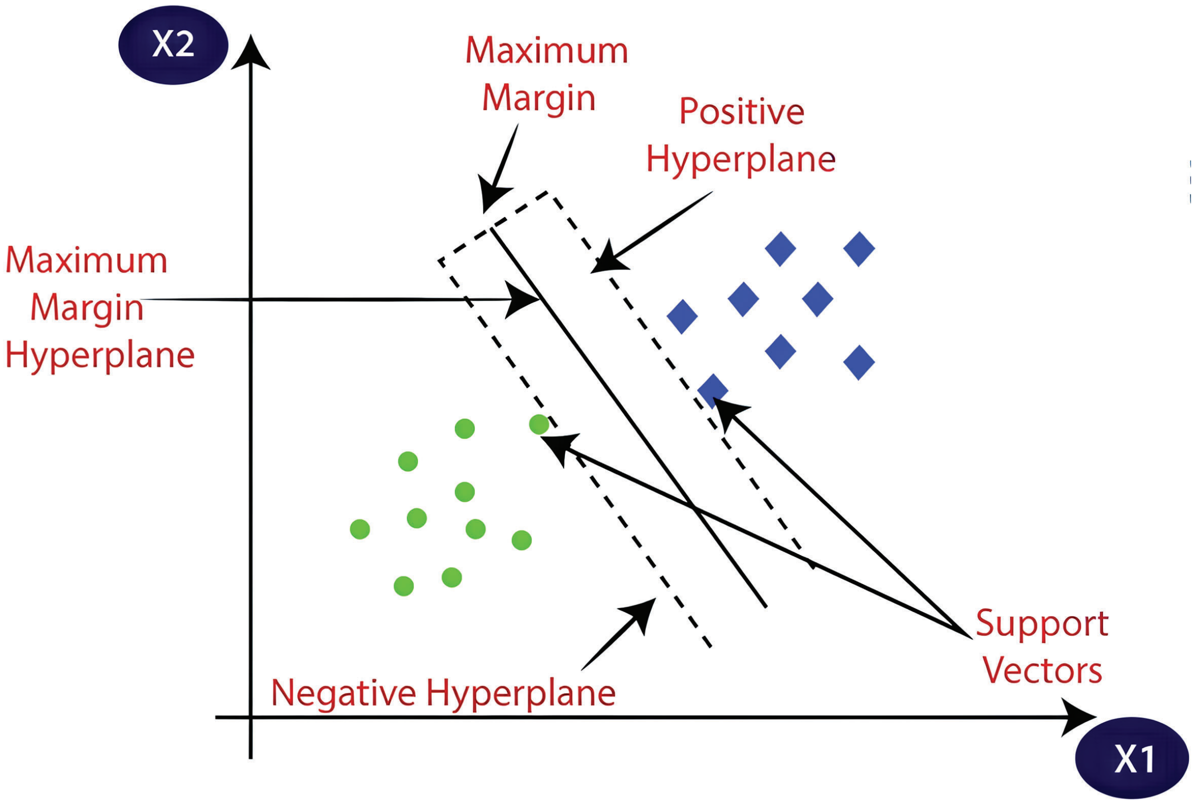

Figure 1: Classification using a decision boundary (Madugu *et al.,* 2023:108)

1.2 Long Short-term memory (LSTM) Network

The Long Short-Term Memory is an extension of the Recurrent Neural Networks (RNNs) in which the output processed from the previous layer is fed back to the current time stage layer enabling the model to learn while maintaining long-term dependencies. Hence, recalling past information for long periods is the default behaviour of Long Short-Term Memory (Olajide *et al.,* 2024: 755). In order to boost performance, the RNN utilizes the Long Short-Term Memory (LSTM) concept, which is a refined RNN network that incorporates memory cells to ensure the retention of cell state when processing sequential data (Okpor et al., 2023: 60). Therefore, the LSTM model proposed is crafted to process dataset elements one at a time while preserving the information state through its memory state (Oluranti et al., 2023: 8). This mechanism successfully mitigates the vanishing gradient problem. The behavior of an LSTM cell is defined by the following equations:

$$c_{in}^{t} = \tanh(W_{xc}x^{t} + W_{hc}h^{t-1} + b_{c}) \tag{1}$$

$$i^{t} = \text{sigmoid}(W_{xi}x^{t} + W_{hi}h^{t-1} + b_{i}) \tag{2}$$

$$o^{t} = \text{sigmoid}\left(W_{xo}x^{t} + W_{ho}h^{t-1} + b_{o} + b_{forget}\right) \tag{3}$$

$$f^{t} = \text{sigmoid}\left(W_{xf}x^{t} + W_{hf}h^{t-1} + b_{f}\right) \tag{4}$$

$$c^{t} = f^{t}c^{t-1} + i^{t}c_{in}^{t} \tag{5}$$

$$h^{t} = o^{t}\tanh(c^{t}) \tag{6}$$

where $\mathrm{W} \in \mathbb{R}^{ls \times ls}, x^{t}, h^{t}, o^{t}, f^{t}, c^{t}, \mathrm{b} \in \mathbb{R}^{ls}$. ls is a hyperparameter, called the LSTM size, and is defined upfront by design as constant among all cells.

2. Literature Review

Vinupaul et al. (2016: 4) presented the notion of flow-bundle-level features, which can be derived from packet-level and flow-level characteristics typically gathered by flow probes, based on the assessment of flow features. They identified a set of flow-bundle-level attributes capable of accurately recognizing users. This set comprises two distinct types of features: user-features, which pertain to the specific activities of a user, and host-features, which relate to the characteristics of the user's host platform. They further argue that as the use of personal mobile devices increases, one host, one-user systems will become more prevalent, thereby enhancing the user identification model through the integration of host features. The model was validated through the application of four distinct supervised learning techniques on a dataset comprising 65 user flow data, achieving a peak accuracy of 83%. Furthermore, Rojas et al. (2019: 590) presented a performance comparison of traditional and incremental machine learning algorithms applied to data regarding users' Over-The-Top consumption to see which approach is capable of a continuous model adaptation while maintaining their usefulness over time. For the tests, two datasets are used: the first is made up of 1,581 instances from a genuine network experiment, while the second has 150,000 instances that were created artificially. Upon reviewing the results, it was concluded that the Support Vector Machine excelled in the traditional method, whereas the optimal classifier for the incremental approach was an ensemble method combining the K-Nearest Neighbor algorithm with Oza Bagging.

(Radha, 2013: 10) proposed a method for quantifying YouTube usage, primarily focusing on the analysis of encrypted network traffic. An Android application named YouQ was developed to facilitate the creation of a training dataset derived from monitored client-side application layer KPIs. The dataset, which comprised 1060 video streaming instances, was subsequently divided into two subsets: the reduced dataset and the full dataset. Five machine learning algorithms were employed to construct a model, including OneR, Naive Bayes, SMO, J48, and Random Forest. The reduced dataset emphasized realistic network conditions without significant fluctuations in bandwidth. The Random Forest algorithm achieved an accuracy of 80.18% on the full dataset, while the Naive Bayes algorithm yielded the highest accuracy of 83.94% on the reduced dataset. (Goeffery *et al.,* 2014: 26) Employed an unsupervised machine learning method using a clustering algorithm based on Expectation Maximization (EM) to identify Internet traffic, comparing it with a supervised Naïve Bayes classifier. The unsupervised method achieved 91% accuracy, exceeding the supervised method by 9%. (Oluranti *et al,* 2021: 8) The study illustrated the classification of network traffic through machine learning techniques from two distinct viewpoints: one incorporating feature selection and the other excluding it. The experimental findings reveal that the classification method without feature selection achieved an average accuracy of 94.14% and a runtime of 0.52 seconds. Conversely, the approach utilizing feature selection recorded an accuracy of 95.61% with an average runtime of 0.25 seconds. (Juan *et al.,* 2020: 25) established an OTT consumption analysis model utilizing Incremental Learning algorithms, which include Naive Bayes, K Nearest Neighbour, Adaptive Random Forest, Leverage Bagging, Oza Bagging, Learn++, and Multilayer Perceptron. The results indicate that the Adaptive Random Forest and a combination of Leverage Bagging and Adaptive Random Forest deliver the highest performance, achieving classification precision and recall exceeding 90%. Based on these findings, they proposed personalized service degradation policies to aid decision-making in mission-critical systems. (Jeffery *et al.,* 2007: 28) utilized two unsupervised clustering techniques, specifically K-Means and Density-Based Spatial Clustering of Applications with Noise (DBSCAN), to categorize network traffic. The performance of these algorithms was assessed and compared against the established AutoClass algorithm found in existing literature, using empirical Internet data. The results indicated that their methods outperformed AutoClass in both speed and efficiency. Additionally, the research highlighted that while DBSCAN exhibited lower accuracy than K-Means and AutoClass, it produced superior clusters. Conversely, AutoClass was noted for its significant time consumption during model construction, which discourages system developers from adopting this algorithm, despite the infrequent need for model retention.

3. Research Methodology

The SVM and LSTM models for analyzing user activities on a computer network are developed on a 64-bit Windows OS with an Intel Core i5-3630QM CPU at 2.40GHz and 4GB RAM. The development environment is Anaconda, using Python 3.8, the Sklearn API, and essential libraries like NumPy, pandas, TensorFlow, and Matplotlib. The research consists of three phases: data pre-processing, model implementation, and model evaluation. In the data pre-processing phase, the Universidad Del Cauca network dataset from Kaggle, containing packet captures from April to June 2019, is utilized. A total of 2,704,839 flow instances dataset having 50 features detailing IP address flows from a network device, such as source and destination IPs, ports, flow durations, inter-arrival times, packet sizes, and the layer 7 protocol for classification was used. Data filtration addressed missing values, and scaling converted non-numeric data to numeric format. The Spearman rank correlation coefficient was used for feature extraction, suitable for non-linear correlations in large datasets. The implementation phase involved splitting the pre-processed dataset into a 70:30 training and test set. The dataset was then used in two deep learning algorithms: Support Vector Machine (SVM) and Long Short-term Memory (LSTM) network. Parameter tuning was performed for both algorithms to optimize performance, with configurations detailed in Tables 2 and 3.

Table 1: Dataset Feature descriptions

Dataset feature name	Feature description	
flow_key	Flow identifier through a hash algorithm	
src_ip_numeric	Source IP in decimal format	
src_ip, dst_ip	Source and destination IP in network format	
src_port, dst_port	Source and destination port number	
Proto	Transport protocol number according to IANA (e.g., 1 for ICMP, 6 for TCP, 17 for UDP)	
pktTotalCount	Total number of packets in both directions	
octetTotalCount	Total number of bytes exchanged in both directions, focusing on the IP payload only	
min_ps, max_ps	Minimum and maximum packet size on the flow, in both directions	
avg_ps	Average packet size on the flow, in both directions	
std_dev_ps	Packet size standard deviation, in both directions	
flowStart	Flow start time in seconds using UNIX time format	
flowStart, flowEnd, flowDuration	Flow start time, flow end time, and total flow duration time in seconds using UNIX time format	
min_piat, max_piat, avg_piat	Minimum packet inter-arrival time, maximum packet inter-arrival time, and average packet inter-arrival time on the flow, in both directions	
std_dev_piat	Standard deviation of packet inter-arrival times, in both directions	
f_pktTotalCount	Total number of packets, in the forward direction	
f_octetTotalCount	Total of bytes exchanged in the forward direction, focusing on the IP payload only	
f_min_ps, f_max_ps, f_avg_ps, f_std_dev_ps	Minimum packet size, maximum packet size, average packet size and packet size standard deviation on the flow, in the forward direction	
f_flowStart, f_flowEnd, f_flowDuration	Flow start, flow end, and total flow time in seconds using UNIX time format, in the forward direction	

Table 2: Parameter Set for SVM model for User's Network Resources Use Analysis

SVM Parameter	Value
Kernel	Linear
Probability	True
Degree	3
Cache-size	200

Table 3: Parameter Set for LSTM model for User's Network Resources Use Analysis

LSTM Parameters	Value
Max-Batch Size	5
Max epochs	20
Initial learn rate	0.0001
Optimizer	Adam
Drop-out	0.2
Activation	Relu and Softmax

The maximum batch size indicates how many input instances are processed per layer. Maximum epochs define the total number of complete passes through the training dataset. The learning rate controls the size of updates to the model's parameters during training. The optimizer adjusts the model's weights and biases to improve accuracy and performance. The activation function determines when an input is activated. For LSTM and MLP algorithms, RELU (Rectified Linear Unit) is used in input and hidden layers, while SoftMax is used in the output layer for multiple class labels. The Kernel Cache Size has a strong impact on run times for the applied Support Vector Machine (SVM) algorithm. It determines the cache size needed for the program in the RAM (Random Access Memory), the cache size used for SVM was set to 200. Considering that the variables from the dataset are linear, the linear kernel was used. In the final stage of model evaluation, this study employed three performance metrics to assess the two developed models: accuracy, precision, F-score, and recall. Accuracy indicates the classifier's recognition rate, while precision measures the exactness of the classifier's predictions. Recall reflects the sensitivity of the classifier. The F-score represents the harmonic mean of precision and recall, thereby integrating both values into a single score. The equations for accuracy, precision, recall, and F-score are presented as equations 7, 8, 9, and 10, respectively (Okpor *et al.,* 2024: 222).

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN\ (total\ sample)} \tag{7}$$

$$precision = \frac{TP}{TP+FP} \tag{8}$$

$$Recall = \frac{TP}{TP+FN} \tag{9}$$

$$F1_{-Score} = 2 \times \frac{Precision \times Recall}{Precision + Recall} \tag{10}$$

Where; P is the total positive predictions, N is the total negative predictions. True Positive (TP) are correctly identified positives, True Negative (TN) are correctly identified negatives, and False Positive (FP) are incorrectly classified positives, which are negative tuples that are incorrectly labelled as positive, and False Negative (FN), which are positive tuples that are incorrectly labelled as negative (Olajide and Andrew, 2023: 63).

4. Results and Discussion

During the training of the developed models aimed at analyzing user activities within computer networks, this study divided the dataset into two distinct segments: training and testing. The training segment was specifically used for the purpose of model training, while the testing segment was utilized to evaluate the performance of the models. Out of a comprehensive total of 26,014 records, 30%—which equates to 7,805 records—were set aside for testing and validating the performance of each model through various evaluation metrics. Conversely, the remaining 70%, amounting to 18,209 records, was allocated for the training of the models. To enhance the feature selection process, the Spearman rank correlation coefficient technique was applied to the dataset, with the outcomes depicted in Figure 2

```
Out[153]: Index(['src_ip', 'dst_ip', 'proto', 'pktTotalCount', 'octetTotalCount',
             'min_ps', 'max_ps', 'avg_ps', 'std_dev_ps', 'flowStart', 'flowEnd',
             'min_piat', 'max_piat', 'avg_piat', 'std_dev_piat', 'f_pktTotalCount',
             'f_octetTotalCount', 'f_min_ps', 'f_max_ps', 'f_avg_ps', 'f_std_dev_ps',
             'f_flowStart', 'f_flowEnd', 'f_flowDuration', 'f_min_piat',
             'f_max_piat', 'f_avg_piat', 'f_std_dev_piat', 'b_pktTotalCount',
             'b_octetTotalCount', 'b_min_ps', 'b_max_ps', 'b_avg_ps', 'b_std_dev_ps',
             'b_flowStart', 'b_flowEnd', 'b_min_piat', 'b_max_piat', 'b_avg_piat',
             'b_std_dev_piat', 'flowEndReason', 'category', 'application_protocol',
             'web_service'],
            dtype='object')
```

Figure 2: Selected Features from the Dataset

Table 4 illustrates the results obtained from the SVM and LSTM models employed to assess user activities within computer networks. Both algorithms delivered commendable outcomes, with performance metrics exceeding 90%. However, the SVM model outperformed the LSTM models in several key areas, including precision, recall, F1-score, and overall accuracy. To be specific, the SVM model achieved impressive values of 96.00 for precision, 99.00 for recall, 98.00 for F1-score, and 95.40 for accuracy. In contrast, the LSTM model recorded lower performance metrics, with precision at 90.00, recall at 91.00, F1-score at 91.21, and accuracy at 93.50.

Table 4: Results of the SVM and LSTM Models for Analysing Computer Network User's Activities

Algorithm	Precision (%)	Recall (%)	Fl-score (%)	Accuracy
Support Vector Machine (SVM)	96.00	99.00	98.00	95.40
Long Short-Term Memory (LSTM)	90.00	91.0	91.21	93.50

5. Conclusion and Recommendation

This study has performed a comparative analysis of the performance of two developed models for analysing computer network user's activities. The two models were developed using
Support Vector Machine and Long-Short-Term Memory. The study classified the OTT application of a particular user according to the data consumption rate of each application so that the service degradation mechanism can be applied to the particular application generating the traffic instead of a generalized degradation approach which is not the most efficient. The dataset employed in this research was sourced from the Kaggle machine learning repository, a well-known platform for data science and machine learning resources. Specifically, the data was collected from the network of Universidad Del Cauca, located in Popayán, Colombia. This collection involved capturing packets from real users who were utilizing Over-The-Top (OTT) applications, leading to a substantial dataset comprising 2,704,839 individual data instances. These instances were gathered at various times throughout the day to ensure a comprehensive representation of user activity. To evaluate the performance of each machine learning model developed in this study, several metrics were utilized, including accuracy, recall, precision, and the F-measure, which collectively provide a robust assessment of model effectiveness.. The experimental result showed that the SVM model performed as the best model to analyzing computer network user's activities with an accuracy of 95.4%. It is therefore recommended that this work be implemented on live computer networks to evaluates its credibility for carrying out smart service degradation on user-by-application basis.

6. References

1. Rojas, J. S.; Pekar, A.; Rendón, Á.; Corrales, J. C. (2020). Smart user consumption profiling: Incremental learning-based OTT service degradation. *IEEE access*, *8*, 207426-207442.
2. Elijah A. A.; Odeniyi O. A.; Olajide B. O.; Temitope O. O.; Paul O. U.; Emmanuel B.; Atiba T. J.; Adeosun T. H. (2024). Application of Enhanced Triple Modular Redundancy (ETMR) Architecture and Markov Processes to Eliminate Operational Downtime in Traffic Light System, Asian Research Journal of Current Sciences, 6(1): 236-248.
3. Jain M. K. (2021). The Rise of OTT Platform: Changing Consumer Preferences. *EPRA International Journal of Multidisciplinary Research (IJMR)*, *7*(6), 257-261.
4. Rojas J. S.; Rendon, A.; Corrales, J. C. (2019). Consumption behavior analysis of over the top services: Incremental learning or traditional methods? *IEEE Access*, *7*, 136581-136591.
5. Cisco, (2017). Cisco Visual Networking Index: Global Mobile Data Traffic Forecast Update, 2016–2021 White Paper. Retrieved March 11, 2018, from from https://www.cisco.com/c/en/us/solutions/collateral/service-provider/visual-networking-index-vni/mobile-white-paper-c11-520862.html.
6. Okpor J.; Olajide B. O.; Tsokwa I. D.; Ehizuenlen E. P. (2024). An ECG-Based Model for Heart Arrhythmia Classification using Deep Learning, FUW Trend in Science and Technology Journal, 2024; 9(2): 218-224.
7. Madugu J. O.; Olajide B. O.; Okpor J. (2023). Smart User Consumption Analysis Using Support Vector Machine and Multilayer Perceptron, Asian Research Journal of Current Science, 5(1): 102-114.
8. Olajide B. O.; James O.; Chukuwudi J. I.; Temitope O. O.; Adeogun A. E. (2024). A Child Dietary Prescription Model Using Artificial Neural Network and Fuzzy Logic, International Journal of Research Publication and Reviews, doi:https//doi.org/10.55248/gengpi.5.0924.2423, 5(9): 750-758.

9. Okpor J.; Olajide B. O.; Madugu J. O. (2023). Over-the-top Applications Traffic Analysis Model for Networks Using Multilayer Perceptron (MLP) and Long Short-Term Memory (LSTM), Asian Journal of Pure and Applied Mathematics, 7(1): 55-66.
10. Oluranti J; Sanjay M; Victor O. (2021). Comparative Analysis of Machine Learning Techniques for Network Traffic Classification, IOP Conference series: Earth and environmental Science, 4th International Conference on Science and Sustainable Development (ICSSD2020). 1-15. IOP Publishing doi:10.1088/1755-1315/655/1/012025.
11. Vinupaul M. V.; Bhattacharjee, R.; Rajesh R; Kumar G. S. (2016). User characterization through network flow analysis. International Conference on Data Science and Engineering (ICDSE). 1-6.
12. Radha R. (2013). Characterization of youtube video streaming traffic, a master's thesis submitted to Department of Electrical Engineering, School of Computing Bliekinge Institute of Technology, Kariskrona, Sweden. 1- 50.
13. Goeffery E.; Adrian M.; Aizoman A. (2014). Development of unsupervised machine learning Model for Internet Traffic Identification, Conference paper submitted to Department of Computer Science University of Calgary Calgary, AB, Canada T2N 1N4. 22-30. DOI: 10.1109/GLOCOM.2006.443
14. Juan S. R.; Adrian P.; Alvaro R.; Juan C. C. (2020). Smart User Consumption Profiling: Incremental Learning Based OTT Service Degradation, IEEE Access. 8: 23-30.
15. Jeffery E.; Anirban M.; Martin A. (2007). Internet Traffic Identification using Machine Learning, Conference paper submitted to Department of Computer Science University of Calgary Calgary, AB, Canada T2N1N4, 22-30.DOI: 10.1109/GLOCOM.2006.443 · Source: IEEE Xplore
16. Olajide B. O.; Andrew I. W. (2023): "Bank Term Deposit Service Patronage Forecasting Using Machine Learning", in International Journal Vallis Aurea (Journal of Sustainable Development and Innovation), 9(2): 55-64.