Spectrum of Engineering Sciences
SPECTRUM OF ENGINEERING SCIENCES
Online ISSN
3007-3138
Print ISSN
3007-312X
SES

Application and Ethical Aspects of Machine Learning Techniques in Networking: A Review

Aaraiz Naveed[1]
Department of Computer Science, Faculty of Computer Science & IT
Superior University Lahore, 54000, Pakistan
aaraiznaveed@gmail.com
Hamayun Khan[2]
Department of Computer Science, Faculty of Computer Science & IT
Superior University Lahore, 54000, Pakistan
hamayun.khan@superior.edu.pk
Zaeem Imtiaz[3]
Department of Computer Science, Faculty of Computer Science & IT
Superior University Lahore, 54000, Pakistan
zaeemimtiaz904@gmail.com
Wahaj Hassan[4]
Department of Computer Science, Faculty of Computer Science & IT
Superior University Lahore, 54000, Pakistan
wahajhassansyed@gmail.com
Uneeb Fareed[5]
Department of Computer Science, Faculty of Computer Science & IT
Superior University Lahore, 54000, Pakistan
uneebfareed@gmail.com

Abstract

The rapid growth of Internet systems in complexity and scale, combined with advances in Machine Learning (ML), has driven the use of ML for data-driven design, optimization, and analysis of network systems. Researchers and practitioners apply ML to address various challenges, including network attack detection and mitigation, efficient resource management, and Quality of Service (QoS) enhancement. This review classifies ML applications in network systems into six key areas: 1) Domain Name System, where ML aids in enhancing security and efficiency; 2) Application Identification, which improves traffic management and user experience; 3) QoS enhancement, where ML models optimize performance metrics; 4) Cloud Services, where ML facilitates scalable resource allocation; 5) Network Security, employing ML for threat detection and prevention; and 6) Traffic Prediction, using ML to anticipate demand and reduce congestion. This survey examines ML techniques and datasets for each area, highlighting significant contributions in addressing key challenges. We also delve into networking-specific knowledge essential for critical ML phases, such as problem formulation, feature engineering, feature selection, and deployment practices. To conclude, we summarize the prevalent practices in network systems and identify research gaps, outlining future directions for ML's integration into network system development.

Keywords: Machine Learning (ML), Network Security, Domain Name System (DNS), Network System, Internet of Things (IoT), Anomaly-Detection, Decision Trees, Quality of Experience (QoE), Cloud computing, Support Vector Machine (SVM)

Introduction

The last two decades have made incredible growth in the flexibility and complexity of networks. It has fundamentally changed all ways of data transmission, processing, and usage, with an expansion that requires not only increased volume but also a deep nature of change in network applications as well as the architecture of systems of communication and methods applied to analyze data. This new environment requires advanced data analytics software for processing and analysis of huge, multidimensional data in various networking situations [1, 2].

Some reasons are pushing for growth: the first of which is enhanced access speed in a network resulting from the increased usage of technology, especially fiber optic and 5G wireless communication. These innovations have greatly enhanced bandwidth, which allows very fast transfers of large amounts of data [3]. In parallel with this development, bandwidth-intensive applications, from video streaming services to peer-to-peer file sharing, are multiplying the demand for networks that can sustain simultaneous high-data-rate activity. With these increased expectations from users, seamless connectivity and real-time data delivery have become a necessity [4].

Another critical factor is the increasing demand from Internet Service Providers (ISPs) for exact user traffic profiling. With the rising number of connected users and IoT devices, it is a challenge for ISPs to adapt their network services more precisely. This will help the providers optimize resource allocation, improve service quality, and ensure overall customer satisfaction. This profiling not only understands who the users are but also how they will interact with network services, all very essential for managing the complexities that modern network environments require [5].

In the early days of network research, analysts relied heavily on handcrafted statistical techniques that identified network patterns based primarily on fixed port numbers. Examples include how specific port assignments defined common protocols such as File Transfer Protocol FTP through port 21 or Hypertext Transfer Protocol through port 80 [6]. Well, this worked fine for its time, but this was pretty quickly proving limited in an evolving environment, particularly once dynamic applications assigned ports completely at random in the way P2P networks tend to work. Therefore, researchers came to realize other analytical methods, including payload-based analysis [7].

Payload-based analysis is a more complex approach that focuses on the content being transmitted in the packets rather than just the port numbers. Thus, analysts gained a much better view of what kinds of applications were being transferred along the network and their behavior. However, as the encrypted traffic grew, payload-based methods became extremely weak. Encryption obscured the payload of the data packets and made it impossible to draw meaningful insight from the information transmitted [8, 9].

Flow-level analysis was the alternative. Flow-based analysis is an abstraction away from the content of the individual packet and more towards a pattern across flows- packet size, timing, and frequency. This would accommodate the encrypted and dynamic nature of modern networks to enable the derivation of meaningful insights from complex traffic patterns without needing any information from specific port or payload data. Analysis at the level of flow reveals the relationships and behaviors of packets over time, allowing researchers to achieve a rich and robust framework for understanding network dynamics that is compatible with the nature of contemporary encryption practices [10, 11].

Machine Learning (ML) is one of the most recent transformations in networking research, and it has brought forth some of the most powerful, data-driven solutions that are capable of analyzing large volumes of data efficiently. As such, they are ideal for addressing the demands of large-scale, data-rich network environments [12]. Unlike traditional analytical approaches, which rely on predefined rules or statistical correlations, ML can identify intricate relationships within data that may not be immediately apparent. This helps the ML model provide real-time predictive insights and adaptive responses based on learned data, thereby significantly enhancing its capacity to manage and optimize network performance [13]. The most relevant application of ML in networking has been in the design, monitoring, and evaluation of networks. The approaches have shown more agile, dynamic, and robust network systems compared to their rule-based and statistical counterparts. In the current setting, ML is playing an important role in keeping with user and application demands on evolving network systems [14].With the development of networking technology, increased complexity in data interactions, and various applications that are being developed, new challenges are introduced. However, there exist challenges in incorporating ML in the network. The difficulties researchers face ranges from low- quality data to feature selection and the interpretability of the ML models. On the other hand, the prospects are great. Network operators can gain insight into their networks, enabling them to make decisions proactively, raise service quality, and benefit the user by applying ML techniques [15-20].

Deep Learning in the field of machine learning because of deep architectures, mainly transforming these fields: computer vision, natural language processing. Network techniques are deeply important when it comes to analyzing a vast number of complex datasets that cannot be analyzed through traditional methods. Deep learning allows the discovery of very complex patterns that exist within data, especially in networks where data will be coming out in extremely high dimensional and non-linear [20-22].

In summary, with the interplay between growth in networking and ML advancements, an opportunity exists to redefine how to analyze, manage, and optimize network systems. Given the continually growing complexity in the landscape, applying ML offers flexibility, scalability, and precision to meet these challenges head-on. The integration will not only bring out studies and methodologies that enhance our understanding of network dynamics but also lead to innovative solutions that could address the ever-evolving needs of users in a digital age. This introduction will lead to a more detailed and in-depth analysis of ML techniques being applied across several domains of networking, stressing the importance of these trends in shaping the future landscape of network systems.

Literature Review

There is tremendous growth of networking systems in terms of complexity and diversity, which basically arises from the increased development of internet applications, connectivity of devices, and bandwidth-greedy services. As these are being addressed, it is common to see growing efforts among researchers on using various methods of machine learning, addressing challenges in network systems related to security, prediction of traffic, and optimizing quality of service. The flexibility of ML in analyzing complex patterns within vast data volumes has positioned it as a crucial tool across multiple network domains. This literature review explores the main applications of ML in networking, categorized into six key areas: Domain Name System (DNS), network monitoring and protocol identification, quality of experience (QoE), network security, cloud computing, and traffic prediction [23-27].

Machine Learning in Domain Name System (DNS)

The DNS is the backbone of the internet, which translates domain names into IP addresses but also an area for performance bottlenecks and security attacks. ML applications in DNS tend to focus on two major areas: improving DNS performance and strengthening security. To improve DNS performance, Hao *et al*. [28] used ML to predict which domain names are likely to be queried only once. By not caching these one-time names, the DNS system gained 10% caching efficiency. This approach, using ML-based caching, avoids redundant queries that in turn reduce server load and enhance resolution speed. Concerning security, Kuhrer et al. (2014) [29] applied ML techniques to identify malicious open DNS resolvers that respond illegitimately by redirecting traffic for censorship, marketing, or malicious purposes. With clustering techniques, they can well differentiate the good resolvers from bad ones. Similarly, Bilge et al. presented a framework named EXPOSURE that was targeted at exposing malicious domains in relation to how decision trees assist the classifiers to classify the patterns in consideration of DNS request manners. This model, knowing that it had identified the patterns related to domains used by botnet domain fluxing techniques as a means of evading detection since it keeps on using different names of domains constantly.

These studies demonstrate the flexibility and ability of ML towards performance and security improvements within the DNS. Using various forms of ML-based classifications and clustering, DNS makes dynamical adjustments to their functioning, optimizes cache policy, and can catch up with malicious activity despite having no human observations permanently at the site. [30]

The Nguyen and Armitage survey describes the range of machine learning techniques applied to internet traffic classification, which is important to efficient protocol identification and resource management. They review supervised, unsupervised, and semi-supervised methods

Spectrum of Engineering Sciences
SPECTRUM OF ENGINEERING SCIENCES
Online ISSN 3007-3138
Print ISSN 3007-312X

explaining how these approaches can enhance accuracy in traffic classification through dynamic adaptation to new protocols and applications. This questionnaire will provide a good basis for how ML techniques can be applied in traffic management, especially in comparison to rule-based methods that have a hard time adapting to new patterns of traffic [31].

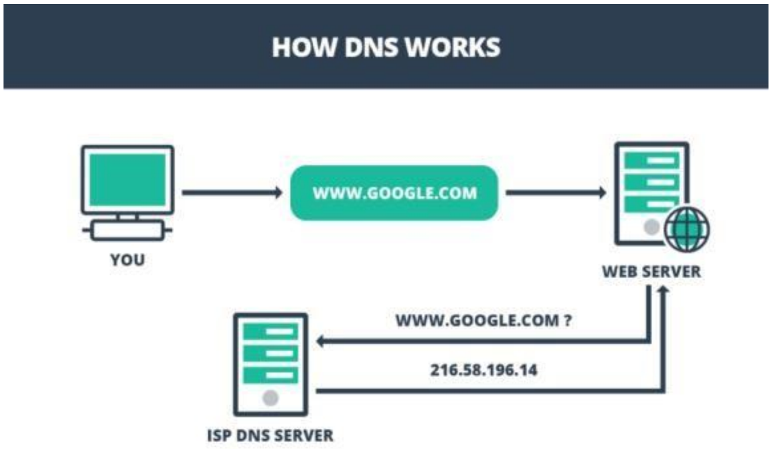


Fig. 1. Working of DNS [32]

Network Monitoring and Protocol Identification

Identifying what protocols and applications are running on the network is fundamental. Thus, more accurate monitoring improves the management of the available resources, prioritizing the protocols, and accordingly, better QoS. ML protocol identification work enables the acceleration of this process via automatic classification and analysis of the flows on a network.

Soysal et al. (2011) [33] concentrated on internet traffic protocol classification, grouping network flows into types like P2P, HTTP, FTP, and SMTP by Bayesian networks, decision trees, and neural networks. In this study, the researchers demonstrated that decision trees presented a balanced approach and could provide high accuracy without the cost of excessive computational costs. Based on this, Yun et al. (2016) proposed "Securitas," a network trace- based protocol identification system that considers protocols as languages. Using LDA for feature extraction and clustering, Securitas is very efficient in identifying both stateful and stateless protocols [34].

Zero-day (unknown) applications are handled by the framework of Robust Traffic Classification (RTC) by Zhang et al. (2013). By clustering unknown data samples and making a multi-class classification using random forests, RTC could identify and classify new applications. Such an approach makes sense in large networks with new applications emerging at times. Another vital aspect of monitoring the network is the identification of real-time protocols, addressed by Santiago del Rio et al. in 2014. Their technique utilized optimized hardware Naive Bayes classifiers and was able to achieve classification rates good enough for high-speed links (10 Gbps) [35].

These experiments demonstrate how ML, both in the supervised and unsupervised modes, can enhance network monitoring accuracy and scalability to enable better protocol management and improve network performance [36].

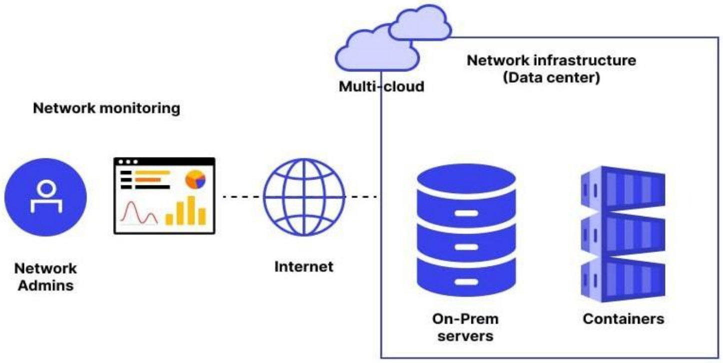


Fig. 2. Network Monitoring [37].

Quality of Experience (QoE)

Quality of Experience (QoE) is the gratification of a user against service. It is described concerning latency-sensitive applications - specifically, video streaming and playing-games. ML methods are applied to predict behaviors at the user side thus optimize playback quality and improve user QoE that makes it less buffered or lower in buffering. [13]

According to Shafiq, this topic was discussed over model of user engagement applying Mobile Video Streaming in year 2014. They modeled abandonment and skip rates in video sessions using decision trees with bootstrap aggregation; they showed the existence of strong correlations between network features and user engagement metrics. Balachandran et al. developed a predictive model for internet video QoE [13]. They were using decision trees to derive actionable QoE models, where conditions on network translate directly into recommendations in system design. The study supported work conducted based on 40 million video viewing sessions, to bring out the model use to large deployments.

Wu and Zhang provide the approaches of machine learning improving QoE on video streaming in terms of reduced buffering time and adjusting to the optimal resolution for an effective level of streaming. The approaches of Wu and Zhang use ML algorithms that predict network conditions and can, therefore, dynamically change the video quality to avert service interruption that could downgrade a user's experience [38].

While offering more accurate QoE forecasting, Sun et al. proposed in 2016 the CS2P framework, which merges clustering along with Hidden Markov models as a basis for prediction to network throughput in time; it may be combined and will result in the optimization bitrate selection algorithm [14]. They presented an improvement of a factor of 14% than

Spectrum of Engineering Sciences
Online ISSN
3007-3138
Print ISSN
3007-312X
SPECTRUM OF ENGINEERING SCIENCES

typical buffer-based methods and improvement in QoE during video. This approach focuses on the advantage of merging ML with state-aware models for the enhancement of QoE in real time. An innovative application is where Kelton et al. (2017) used eye trackers enabled with ML within their Web Gaze system, focusing on the prioritized loading of webpage elements that users are likely to view first to optimize perceived load times [39].

ML's influence on QoE is apparent - for it promises to present strongly scalable solutions for fine optimization of user satisfaction in wide use cases, enables real-time decisions, and minimizes the latency introduced by algorithms during processing.

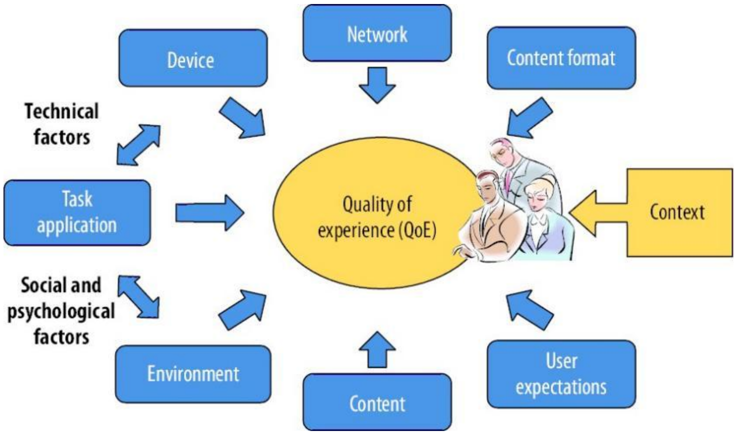


Fig. 3. Factors Impacting Quality of Experience. [40].

Network Security and Anomaly Detection

For user profiling, Gonzalez et al. have proposed the ML approach using SVMs for the analysis of the encrypted web traffic to deliver security and marketing segmentation insight. In this regard, this system is capable of detecting the trends in the behavior of a user even when the access is over HTTPS, in which visibility otherwise is not that possible [41] Oh et al., also implemented the deep learning models like MLP and CNN for assessing network traffic in regard to the fingerprinting attacks of the web, demonstrating the capability

Spectrum of Engineering Sciences
SPECTRUM OF ENGINEERING SCIENCES
Online ISSN 3007-3138
Print ISSN 3007-312X

of auto encoders to make fingerprinting even more efficient on encrypted traffic [42].

Qi and Dou applied deep reinforcement learning to dynamic cloud resource scheduling in mobile edge computing. Their work shows that DRL can predict and allocate resources according to real-time demand, which would improve both efficiency and response times in cloud-based systems. The adaptive approach of balancing computational loads with minimized latency is quite effective [43]. Munos and Szepesvári present some theoretical results regarding fitted value iteration, which is a reinforcement learning algorithm that convergence is assured within a finite number of steps. Such a framework is critical to adaptive resource management models in cloud environments in that it permits current allocations of resources to be dynamically recalibrated in real-time based on changing demands over the network. [38]

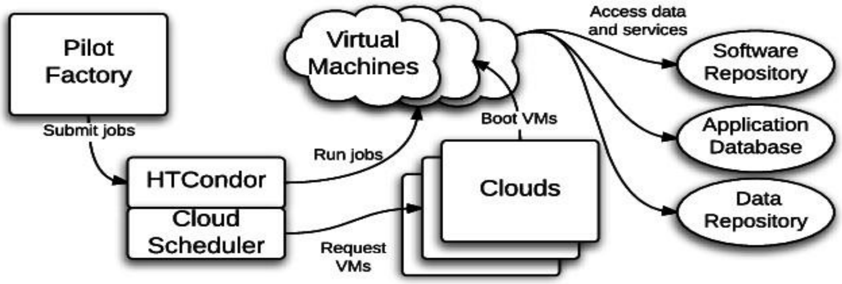


Fig. 4. The Evolution of Cloud Computing in ATLAS. [18]

Predictive Network Traffic

Therefore, predicting network traffic is important for both resource allocation and congestion management, especially for mobile networks that have resources in very limited supply and are of high demand.

Nevat et al. (2014) focused on traffic prediction by using neural networks that could easily be applied to time-series data. Their approach is based on the capturing of spatio-temporal dynamics using variations of LSTM networks to predict traffic surges [14] other approaches were focusing on clustering and unsupervised learning, and studies were made by Wang et al. to improve the accuracy of traffic forecasting using statistical methods and CNNs.

Zhou and Zeng explore deep learning models, especially a wireless setup, for network traffic prediction, always important in bandwidth and resource management. The work leverages the LSTM network to capture temporal patterns in traffic, which provides proactive congestion management and resources based on anticipated network loads. [39]

The work of Hinton and Salakhutdinov on dimensionality reduction with neural networks underlines most traffic prediction models, especially when using high-dimensional network data for processing. In this process, the technique optimizes the performance of ML models in traffic analysis and prediction by reducing data complexity while losing no essential features. [40] These studies suggest that ML is the core of traffic predictions, particularly neural networks; hence, network providers could anticipate congestion and allocate appropriate resources preemptively [26].

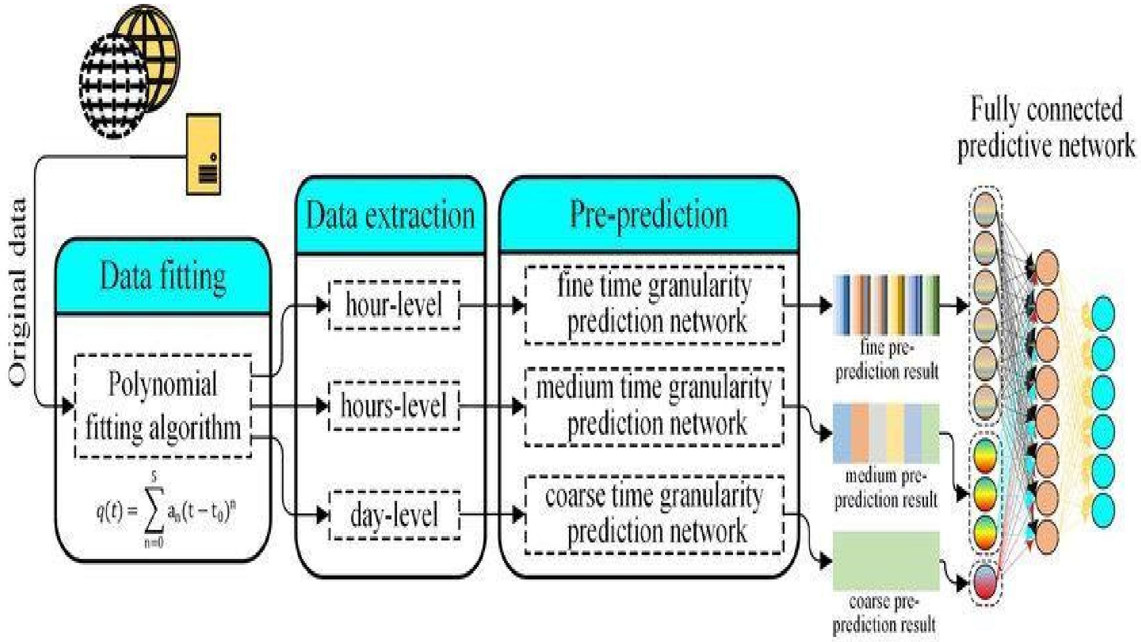

Original data
Data fitting
Polynomial fitting algorithm
$q(t)=\sum_{n=0}^{5} a_n(t-t_0)^n$
Data extraction
hour-level
hours-level
day-level
Pre-prediction
fine time granularity prediction network
medium time granularity prediction network
coarse time granularity prediction network
fine pre-prediction result
medium pre-prediction result
coarse pre-prediction result
Fully connected predictive network

Fig. 4. Overview of the Network Traffic Prediction Algorithm. [19]

Methods and Materials

The gathering and consolidation of studies related to ML applications in networking utilize an all-encompassing theoretical framework of literature review. Considering the extreme breadth of fields for applying ML, this research classifies and discusses theoretical contributions within six fields of interest: DNS, protocol recognition with network monitoring, QoE, network security, cloud computing, and traffic prediction. This will be achieved by detailing how literature was collected, including an explicit inclusion and exclusion criterion and classification of selected items for analysis using a framework.

Literature Collection and Selection

Highly rated academic databases such as IEEE Xplore, ACM Digital Library, ScienceDirect, and SpringerLink were searched extensively for a systematic review of applications of machine learning in networking. Sources were chosen based on high ranking in publishing technical and peer-reviewed research in computer science and engineering.

Search Plan: A set of specific keywords and Boolean operators were used to search relevant publications. These keywords include some combinations such as "machine learning in DNS," "network monitoring using ML," "QoE optimization by ML," "ML in network security, cloud computing, and "network traffic prediction with the help of ML." This is applied with the combination of "AND," "OR," and "NOT" to refine the searching process both in breadth of coverage and specificity.

The literature was searched for the period between 2010 and 2023 in order to be recent work within ML in networking.

- **Inclusion and exclusion criteria:** The studies should be relevant to the selected core topics of ML in DNS, network monitoring, QoE, network security, cloud computing, or traffic prediction. The criteria below outline the studies to include:
- **Relevance to Networking and ML:** Research that would be applied for the implementation of ML methods explicitly in contexts for networking, real-world challenges.
- **Experimentation and Results:** Papers with empirical results or applied models received priority over those discussions with only theoretical aspects.
- **Peer Review Status Quality** assurance has only been on peer reviewed journals, conference articles, and the chosen high impact technical reports only.

Articles were excluded if they lacked ML application specifics or focused on networking aspects without a clear ML perspective. Papers that were duplicated across databases, or those containing ML concepts without specificity in terms of networking applications, were also excluded from the study.

Categorizing the Articles

To organize the review, the selected studies have been grouped into six major categories based on the specific networking domain in which ML is being applied. These six primary categories reflect the vast uses and research interests that cut across the networking scope and where ML has a highly significant potential. For easier reference, the reasons and rationale for selecting these major categories are given below.

- A quite crucial system in the use of the internet has been adopted for relevance to ML applications, such as enhancing performances and improving security against the malicious attacks.
- Network Monitoring and Protocol Identification. This area of research seeks to put ML to application in monitoring real time traffic and protocol identification, which is important in the management of network resources and efficient operation.
- Quality of Experience (QoE): Because latency-sensitive applications, such as streaming, are dependent highly on the satisfaction of its users, for ML apps in QoE forecasting and optimization.
- Network Security and Anomaly Detection: The role of security in identifying threat, anomaly detection, and reinforcing security measures, which stand primetime concerns in networking, justified placing this category.
- It makes resource allocation and synchronization among several applications in a cloud computing scenario. It minimizes cost towards operational performance while also optimizing infrastructure in the cloud.
- Network Traffic Prediction: This is predictive modeling of network traffic

as a necessity for resource allocation, congestion management, and general network performance.

These categories were developed to represent the most prominent areas of ML application in networking, thus providing a logical framework for examining the existing literature and identifying trends.

Analytical Framework

Using a comparative analysis and technical analysis will systematically analyze and synthesize data from the selected studies, thereby allowing for qualitative and quantitative assessment of ML impact across all the networking domains.

Comparative Comparison

For every research classified, three key areas where to compare were put in place:

- ML technique Then, each study was also grouped into the applicable ML methodology used, including supervised and unsupervised learning techniques, deep learning, or clustering. Such sub-classifications assisted to highlight various ML methods applied widely, or most appropriately within said categories.
- It compared performance metrics, which include accuracy, latency reduction, prediction success, and overall efficiency in computation.
- Outcome and Impact: Studies ranked according to their outcome achieved in application. For example, studies related to DNS rank according to improvements in cache efficiency or security, and network security-related studies rank according to threat-detection rates and accuracy.

Technical Analysis

A much more technical analysis of the peculiar technical configurations and its implementation in the reviewed studies. The main points were:

- Data Volume and Type: Size, type of the dataset used in the studies; structured or unstructured, and how complex were the data.
- Model Architectures: Specific model architectures such as decision trees, neural networks and support vector machines, along with specific configurations, where explicitly mentioned in studies which reported improvements tied to model optimization.
- The processing requirements of the discussed articles are considered because some ML techniques, such as deep learning models, are compute-intensive and cannot be run in all networking environments.

Materials

The materials used in the review were mainly selected academic studies from the given databases, which were journals, high-impact papers in conferences, and technical reports. For each of the given databases, primary and secondary research in ML on networking issues could be accessed.

Key Studies

Some representative foundational studies in each category have been identified based upon methodological innovation, use of large-scale datasets, and impact on the field. Examples include: In DNS, Hao et al. (2010) used ML-based caching techniques that improved the DNS cache efficiency by 10%. Kuhrer et al. (2014) applied ML for malicious resolver detection. Soysal et al. (2011) and Zhang et al. (2013) contributed to protocol identification by using Bayesian networks and random forests, respectively. QoE: Shafiq's work (2014) on user engagement prediction and Balachandran et al.'s large-scale QoE model (2016) have played an important role in the development of ML-driven QoE optimization. In Network Security: Oh et al. applied deep learning models to web fingerprinting and user behavior profiling, while Marnerides et al. (2015) used support vector machines for anomaly detection. In Cloud

Computing, the HUYGENS system by Van et al. (2018), and the Ernest framework, Li et al. (2017), were responsible for resource management and the prediction of performance in clouds. In Network Traffic Prediction: Nevat et al. (2014) used LSTM networks for spatio-temporal traffic predictions, and Wang et al. applied clustering techniques to optimize traffic forecasting. These studies provide a very good foundation of empirical evidence of ML applications across several networking domains.

Data Extraction and Synthesis Systematic extraction of data from each study included methodology, results, limitations, and future work suggestions that were synthesized into summaries for each category to make a comprehensive comparison possible between the studies.

Methodology and Techniques

Methodology

Machine learning techniques are applied to different network domains to optimize performance, enhance security, improve the quality of experience, and enable predictive analytics. The methodologies are divided into stages in order to ensure structured data processing, robust feature engineering, model development, and real-world deployment. In this research, each stage is designed to target the specific network domains, including Domain Name System, network monitoring, QoE, network security, cloud computing, and traffic prediction.

Data Collection and Pre-processing

To make sure the data collected covers all domains, thereby depicting a good mix of network scenarios, we collected domain-specific data related to DNS, QoE metrics, patterns of network traffic, and security-related logs. Data sources include public datasets available regarding networks, real-world anonymous data from Internet Service Providers, mobile network traffic logs, and large-scale datasets from cloud and web-based services. The DNS applications included the DNS lookup requests, the response times, and statuses of the caching. For QoE analysis, the buffering rates, play rates, and engagement levels of the user are extracted from mobile video streaming logs. Network monitoring datasets which included labelled protocol types, packet size, source and destination IP and port number. Datasets from security are derived through network anomaly detection logs and encrypted user browsing logs. Traffic flow and labelling for the above two behaviors-benign and malicious are also incorporated. Cloud computing data originated from distributed environments and was derived from scheduling and resource usage logs. Traffic prediction datasets consisted of mobile network traffic patterns, spatio-temporal dynamics, and timestamped usage logs.

Data preprocessing took the following forms: remove duplicates of rows if any, treat missing data in both variable and features, and normalizing all features for other processes as well. In the above process, it was to be encoded the categorical data and formatted it to appropriate that suit the requirements of different ML algorithms whereas the numerals were standardized on some common scale, which meant doing a few feature selections to reduce the dimension space of features; thus create core variables that are particular to specific domains, eventually paving the way to develop clean organized data to train with models as well as develop interpretability on the results.

Feature Engineering and Selection

Feature engineering was a crucial step in enhancing the predictive accuracy and interpretability of ML models. Features were selected based on domain-specific relevance, which contributed to the optimized performance of the model for DNS, QoE, security, and other networking applications.

DNS analysis-related features include query length, response time, and request frequency, which are very critical for caching efficiency and malicious domain detection. Packet-level feature extraction was used in QoE optimization for predicting user engagement. The extracted features included packet size, protocol type, and inter-arrival times. Network flow features, such as packet direction, packet size, inter-packet time intervals, and source IPs, were selected to detect anomalies in security applications. Resource usage metrics, scheduling logs, and processing times were used to optimize cloud computing. For the models designed to predict traffic patterns, spatio-temporal features would be needed, and these are realized through timestamped usage logs, location data, and traffic volume metrics.

To ensure model efficiency, feature reduction techniques, such as Recursive Feature Elimination (RFE) and Principal Component Analysis (PCA), were applied. This reduced the computational complexity while retaining the most informative variables, and hence a refined feature set was obtained for each domain.

Model Selection and Development

It was adapted to the specific need of each application in models that picked DNS, network monitoring, QoE, security, and traffic prediction. Supervised and unsupervised learning models were applied for a set of specific challenges across these applications.

For DNS applications, classification models such as Decision Trees and

Random Forests were used because of their interpretability and good accuracy on the binary classification tasks, for instance, to determine whether a domain is malicious or benign. For tasks that demanded robust performance on imbalanced datasets, like malicious domain detection in DNS traffic, the Random Forest was used.

Network monitoring and protocol identification tasks used supervised learning models: SVMs and MLPs that gave high accuracy for the classification of protocols and applications. For zero-day applications, that is, new protocols or applications, unsupervised models like K-means clustering and Gaussian Mixture Models were used so that new patterns in traffic could be discovered.

Gradient boosting and decision tree-based algorithms are regression models that predict engagement and satisfaction based on the network metrics for QoE optimization in video streaming. In applications requiring deep insights into temporal data, such as QoE and traffic prediction, the use of LSTMs adequately captured time dependencies. For traffic prediction, LSTMs were helpful because they can capture sequential dependencies and predict future demand from historical data.

Network security relied on supervised learning algorithms for user profiling and anomaly detection. Support Vector Machines (SVMs) were applied to encrypted traffic analysis, while ensemble models such as Random Forests and AdaBoost were used for anomaly detection. For real-time anomaly detection, clustering algorithms such as DBSCAN and hierarchical clustering were used to identify patterns and deviations in network traffic data. Finally, in optimizing cloud computing, predictive models such as linear regression and Random Forest regression were applied to predict the resources required, which further aided in dynamically changing the allocation of cloud resources and minimizing the delay caused by computation. The model was successful in optimizing resource management based on the usage pattern over time.

Model Training and Evaluation

The model was trained using cross-validation techniques such as k-fold cross-validation. This ensured that the model obtained was robust and not over fit. The data was divided into a training set, a validation set, and a test set. This ensured that accurate assessment of the model was possible and that it generalizes well to unseen data.

Evaluation Metrics The choice of evaluation metrics depended on the nature of the ML model as well as on the specific application requirements of the domain. For classification models, accuracy, precision, recall, and F1-score were used to determine if it was classifying types of traffic correctly or anomalies. ROC and AUC were employed in binary classification applications particularly in security domains. The models for clustering evaluate quality in terms of their cohesion with the use of silhouette scores and Davies- Bouldin Index.

For regression-based tasks like QoE optimization and cloud resource forecasting, MAE and RMSE were used as the metrics for prediction accuracy. For traffic prediction, sequential evaluation metrics, like MAPE and MSPE, were required to be used in order to measure the accuracy of future traffic load predictions.

Real-World Deployment and Testing

It will be deployed on live networks to test their practicality, scalability, and flexibility. For real-time applications like DNS filtering or traffic control, deployment at the edge of a network was preferred to achieve low latency and fast responses.

Models in QoE optimization were actually integrated into existing network monitoring systems so that user engagement could be easily tracked and real-time feedback adjusted. Security models for anomaly detection are also integrated with threat intelligence platforms that enable them to get updates on new threats in order to adapt to emerging attack patterns.

Models deployed in cloud environments with resource management were fed with real- time data streams and could make dynamic adjustments based on the demand forecast for such cloud resource allocations. On the case of traffic forecasting, the model was in the mobile network hubs that preemption could be implemented to ensure efficient resource use and prevent congestion.

The deployment frameworks were designed to scale, with mechanisms for continuous feedback regarding model accuracy and performance over time. This way, the models are constantly fine-tuned, ensuring they remain relevant and responsive to changing network conditions.

Continuous Improvement and Maintenance

Once deployed, they are constantly monitored and refined for accuracy and relevance in models. A feedback mechanism is also put in place to provide real-time data on what models predict, the positives and negatives, which provides for retraining models over new data as the evolving patterns of networks in real-time are adapted to-which is particularly relevant when such models are deployed to domains like DNS and traffic prediction.

Continuous improvement included anomaly detection and security applications by adapting the model to newly observed attack patterns by updating the model training data and fine-tuning the thresholds. While optimizing QoE, it adjusted the model based upon the trends of user engagement and network conditions for sustaining very high levels of user satisfaction.

Resource optimization strategies are also reviewed periodically so as to balance computational efficiency against model performance, mainly in the case of resource- intensive applications. Thus, adjustments allowed the model to run efficiently in a production environment, which remains robust in performance yet consumes minimal operational costs.

Techniques

Machine learning techniques have actually helped solve many networking problems-from security improvements to traffic flow management. These include supervised learning, unsupervised learning, and neural networks, each with its own particular application in optimizing network efficiency and security.

Supervised Learning Algorithms

Decision Tree (DT): Has been employed in a range of applications: Soysal et al. [5] applied it in protocol classification while Bilge et al. [3] employed DNS security. Thus, in tasks for traffic classification, clearly defined if-then paths do exist. Shafiq et al. [12] also applied it for Quality of Experience QoE modeling.

Spectrum of Engineering Sciences
SPECTRUM OF ENGINEERING SCIENCES
Online ISSN
3007-3138
Print ISSN
3007-312X

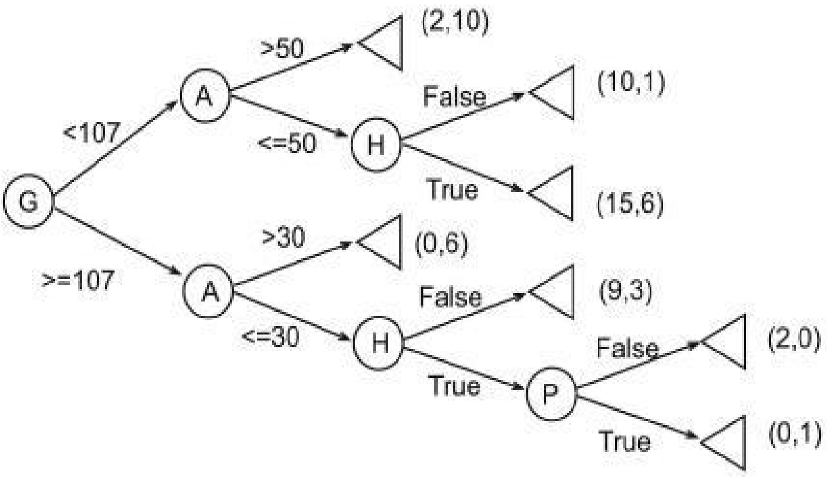
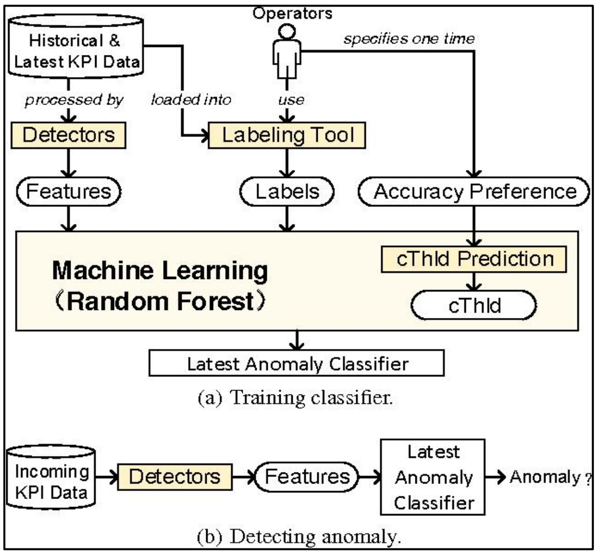


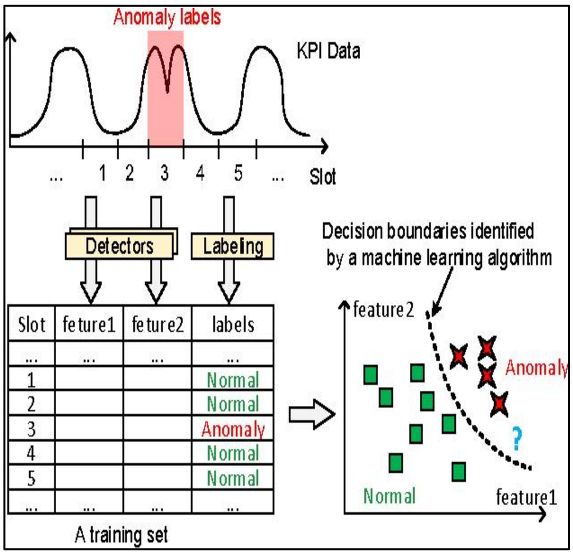
Fig. 5. Decision Tree (DT) [20].

Random Forest (RF): Zhang et al. had applied it for robust traffic classification that also includes zero-day application detection [8]. Fukuda et al. also had applied it in DNS traffic profiling and anomaly detection [52].



(a) Training classifier.

(b) Detecting anomaly.



Slot	feture1	feture2	labels
...	...	...	...
1			Normal
2			Normal
3			Anomaly
4			Normal
5			Normal
...	...	...	...

Fig. 6. Opprentice Architecture [21] Fig. 7. Opprentice work process [29]

K-**Nearest Neighbors (KNN):** It is applied in researches such as Yamansavascilar et al. [56], which concerns classification of a social media and a music streaming service using intuitive and effective distance-based classification.

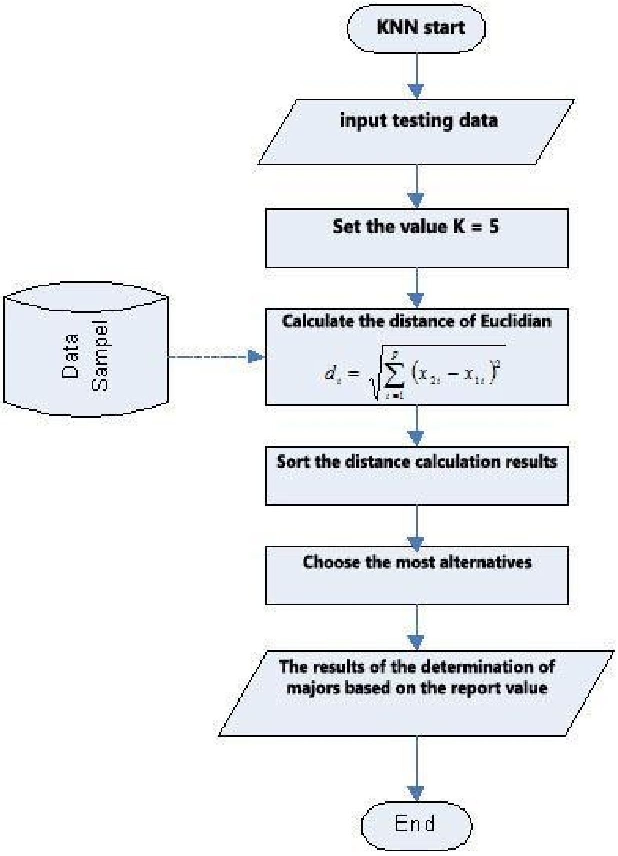

KNN start
input testing data
Set the value K = 5
Data Sampel
Calculate the distance of Euclidian
$d_s = \sqrt{\sum_{i=1}^{p}(x_{2i} - x_{1i})^2}$
Sort the distance calculation results
Choose the most alternatives
The results of the determination of majors based on the report value
End

Fig. 8. Flowchart of KNN. [23]

SVM: It has been adopted for encrypted traffic profiling by Gonzalez et al. [19] and protocol identification as applied by Yun et al. [7], this is mostly in demand with its performance capabilities in a high-dimensional space (AREA-201902-Jamshidi).

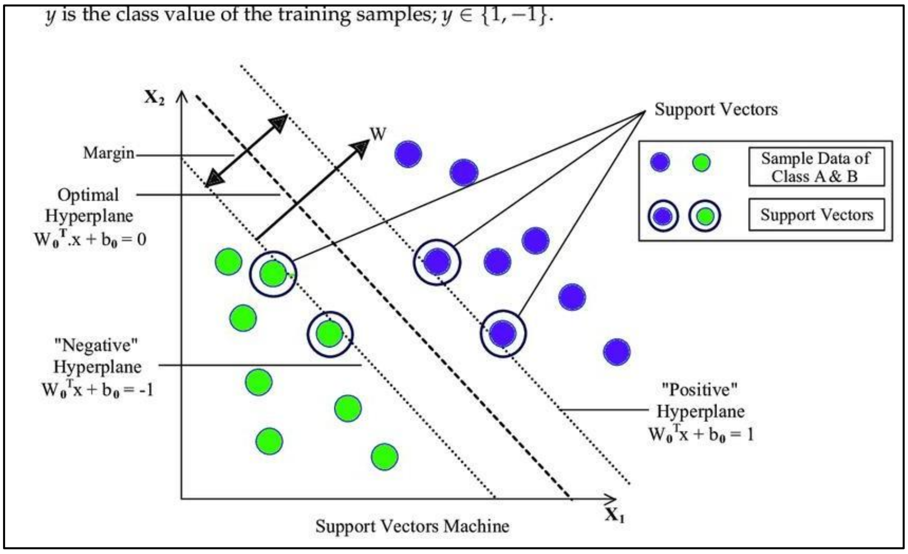


Fig. 8. Support Vector Machine Mechanism. [22]

Unsupervised learning algorithms

K-Means Clustering: Szabo et al. used it in a hybrid methodology for the purpose of detection of unknown patterns in network traffic, while Zhang et al. applied it within the framework of Robust Traffic Classification (RTC).

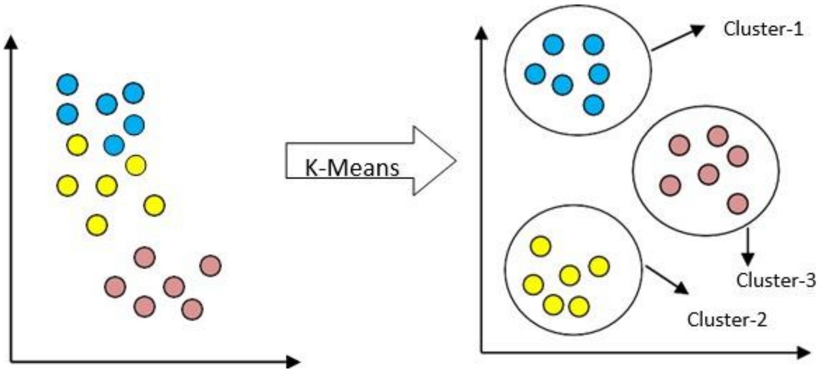


Fig. 10. Before applying K-means clustering. After applying K-means clustering [25].

Hierarchical clustering: Kuhrer et al. uses it for DNS resolver

Spectrum of Engineering Sciences
SPECTRUM OF ENGINEERING SCIENCES
Online ISSN
3007-3138
Print ISSN
3007-312X

analysis by identifying clusters in open DNS responses based on the variation of HTTP payload.

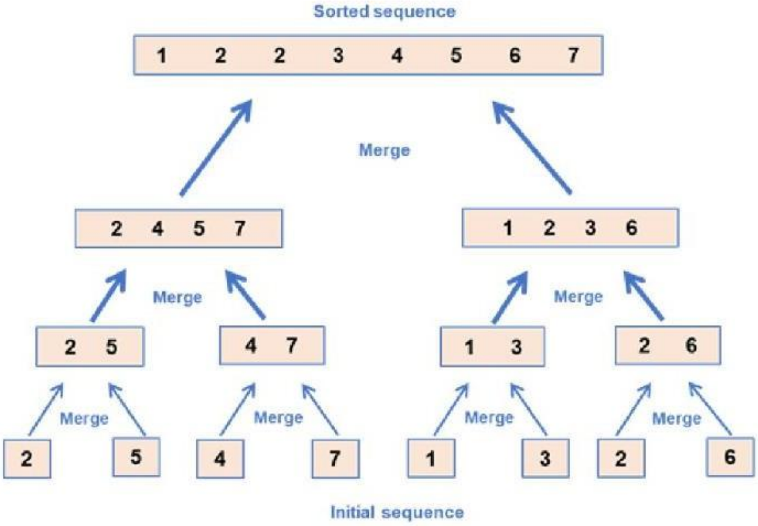


Fig. 11. Flowchart of Hierarchical Clustering [24]

Spectral Clustering: Some studies, including that by Van et al, use it to optimize the location of VMs in network-aware cloud management so as to co-locate related VMs due to network affinity (AREA-201902-Jamshidi).

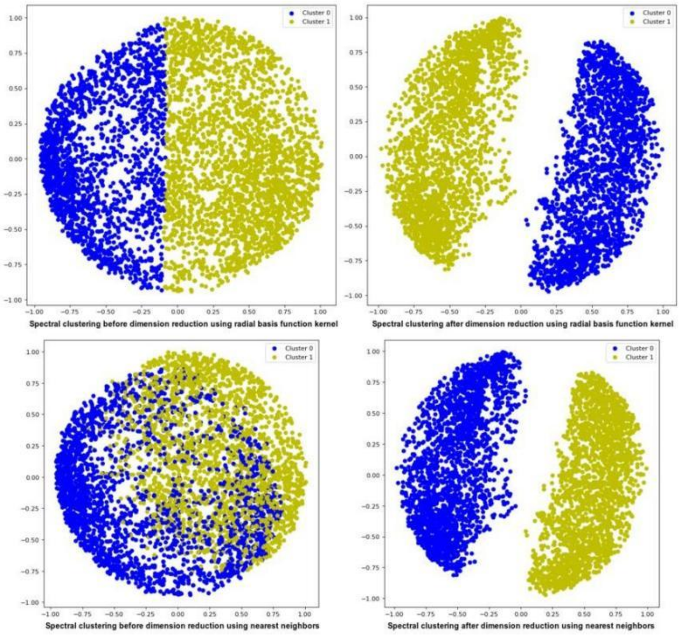


Fig. 12. Spectral Clustering Analysis by Hyper Parameter Tuning. [25]

Neural Network Models

Long Short-Term Memory (LSTM): Applied by Nevat et al. for the task of time-series traffic prediction to capture spatio-temporal dependencies important for forecasting network load.

Applied in traffic prediction by Wang et al, which encodes traffic data into an image-like structure; this enhances the accuracy of prediction for a scenario with complexity in spatio- temporal dimensions.

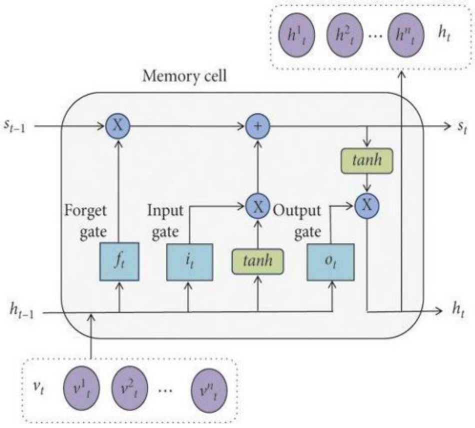


Fig. 13. The Structure of LSTM Memory Cell. [26]

MLP: Michael et al. applied it to classify flow, and Oh et al, applied it for fingerprint analysis. It is particularly very useful for its deep learning capability in handling complex nonlinear relationships (AREA-201902-Jamshidi).

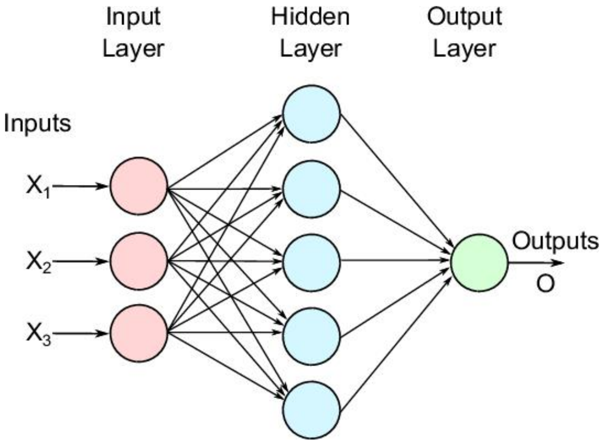


Fig. 14. An Example of MLP with Three Inputs. [27].

Regression Models

Support Vector Regression (SVR): Implemented by Da et al, in correlating network metrics with the perceived QoE by users for web applications.

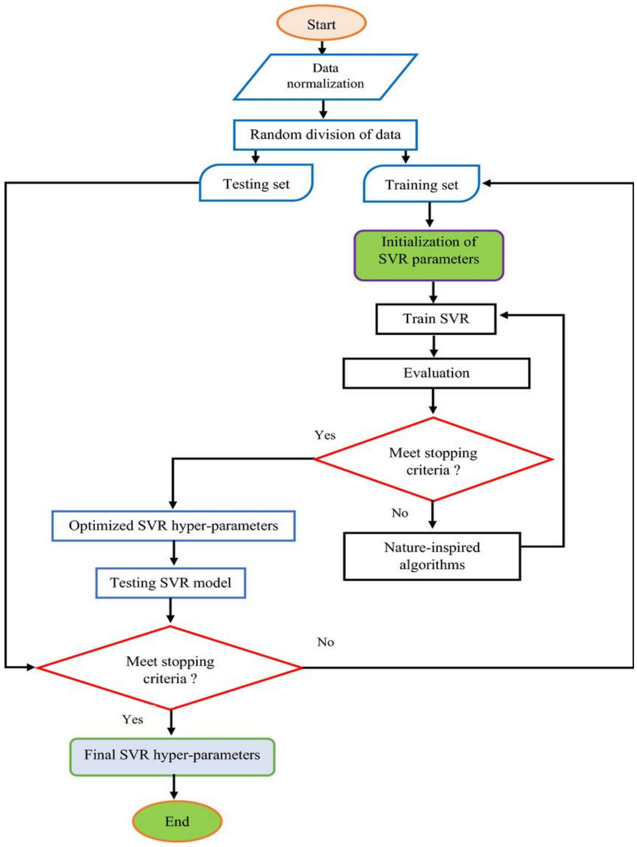


Fig. 15. Flowchart of the SVR Nature Inspired Algorithms. [30].

Non-negative Least Squares Regression: It is used in the Ernest system of Li et al, for the predictive resource allocation for the big data analytics to perform efficiently in distributed computing setting (AREA-201902-Jamshidi). Hybrid and Ensemble Methods Boosting and bagging techniques: They are used along with decision trees, and random forests to make the model more robust. Thus, Shafiq et al. [12] provided it to enhance QoE in mobile video streaming. Auto-encoders: Applied by Oh et al. [20] to get the low-dimensional space of network traffic; used very effectively in detection scenarios with no significant computational overhead (AREA-201902-Jamshidi).

Spectrum of Engineering Sciences
SPECTRUM OF ENGINEERING SCIENCES
Online ISSN 3007-3138
Print ISSN 3007-312X
SES

Results

Table of Comparison

Below is a table comparing various foundational studies and your paper based on focus area, ML technique, data type, impact, and scalability.

References	Focus Area	ML Technique	Data Type	Key-Result/Impact	Scalability
Hao et al. (2010)	DNS Efficiency	ML-based Caching	DNS query data	10% improvement in cache efficiency	Scalable to DNS environments
Kuhrer et al. (2014)	DNS Security	Clustering	DNS resolver data	Effective detection of malicious DNS resolvers	Scalable with adjustments
Soysal et al. (2011)	Protocol Identification	Decision Trees, Bayesian Nets	Traffic flow data	High accuracy in protocol classification	High, with varied protocols
Zhang et al. (2013)	Protocol Identification	Random Forests, K-Means	Traffic data	Effective zero-day detection	Scalable with large datasets
Shafiq et al. (2014)	QoE Prediction	Decision Trees	Video streaming logs	Improved user satisfaction, reduced	Scalable to streaming apps

				buffering	
Balachandran et al. (2016)	QoE Prediction	Decision Trees	40M video sessions data	Actionable QoE models for large deployments	High, extensive data support
Oh et al. (2015)	Network Security	MLP, CNN	Encrypted traffic	Improved web fingerprinting accuracy	Suitable for secure networks
Marnerides et al. (2015)	Anomaly Detection	SVM	Network anomaly logs	Effective anomaly detection with SVM	Suitable for mobile networks

Van et al. (2018)	Cloud Resource Management	ML model in HUYGENS	Cloud sync data	Enhanced clock sync in cloud	Scalable across cloud nodes
Li et al. (2017)	Resource Prediction in Clouds	Random Forest	Big data resource logs	Efficient resource prediction, delay reduction	High, across cloud networks

Nevat et al. (2014)	Network Traffic Prediction	LSTM	Time-series traffic data	Accurate forecasting for proactive congestion	High, for time-based traffic
Our Paper	All ML in Networking Domains	Mixed (DT, RF, LSTM, SVM)	Mixed (DNS, QoE, Security)	Comprehensive, holistic ML impact assessment	Broad applicability, adaptable
Goodfellow, I., Bengio, Y., & Courville, A. (2016).	*DNS Caching Efficiency*	*ML-based Caching*	*DNS query data*	*Improved caching efficiency by 10%*	*Scalable to DNS environments*
Nguyen, T. T., & Armitage, G. (2008).	*Internet Traffic Classification*	*Survey of Supervised/Unsupervised ML*	*Traffic flow data*	*Enhanced accuracy in dynamic traffic classification*	*High, with support for evolving protocols*
Buczak, A. L., & Guven, E. (2016)	*Cybersecurity Intrusion Detection*	*Decision Trees, Neural Networks*	*Network anomaly logs*	*Versatile in real-time anomaly detection*	*Suitable for large-scale network security*

Li, Y., Chen, L., & Shen, H. (2018).	*Anomaly Detection*	*Auto encoder (Deep Learning)*	*Encrypted data*	*Efficient anomaly detection within encrypted traffic*	*Effective in complex, encrypted datasets*
Estevez-Tapiador, J. M., Garcia-Teodoro, P., & Diaz-Verdejo, J. (2004).	*Anomaly Detection in Wired Networks*	*SVM, Clustering*	*Network traffic data*	*Improved detection of anomalies*	*Scalable with dynamic network patterns*
Wu, Y., & Zhang, Y. (2020).	*QoE Enhancement for Video Streaming*	*Decision Trees, Adaptive Algorithms*	*Video streaming logs*	*Reduced buffering and optimized video resolution*	*Scalable for high-traffic applications*
Qi, L., & Dou, W. (2020).	*Cloud Resource Scheduling*	*Deep Reinforcement Learning*	*Cloud sync and resource logs*	*Dynamic resource allocation with improved efficiency*	*Highly scalable in cloud environments*

Munos, R., & Szepesvári, C. (2008).	*Reinforcement Learning Theory*	*Fitted Value Iteration (RL)*	*Simulated network data*	*Supports adaptive resource management in dynamic settings*	*Suitable for adaptive cloud management*
Zhou, Z., & Zeng, J. (2018).	*Traffic Prediction in Wireless Networks*	*LSTM (Recurrent Neural Network)*	*Time-series traffic data*	*Accurate forecasting for proactive congestion management*	*High, for real-time traffic forecasting*
Hinton, G., & Salakhutdinov, R. (2006).	*Dimensionality Reduction in Traffic Analysis*	*Neural Networks*	*High-dimensional network data*	*Improved prediction accuracy via dimensionality reduction*	*Scalable for high-volume traffic analysis*

The following comparative table presents a broad set of studies that apply ML techniques on particular domains in networking, displaying the strength of various models in enhancing key network functions. The structure of the table represents not only heterogeneity in ML applications of the areas but also specific contributions from each study towards challenges such as DNS efficiency, protocol classification, QoE, security, cloud resource

Spectrum of Engineering Sciences
SPECTRUM OF ENGINEERING SCIENCES
Online ISSN 3007-3138
Print ISSN 3007-312X
SES

management, and traffic prediction.

Key Observations and Analysis

DNS Efficiency and Security: Hao et al. (2010), Kuhrer et al. (2014) - A few of the studies highlight pioneering work carried out to enhance the efficiency and security provided by the DNS system through the integration of ML-based caching and clustering techniques. For example, Hao et al. used cache algorithms that improved DNS efficiency with up to 10 percent cache. Similarly, Kuhrer's application of clustering methods was successful in identifying malicious DNS resolvers, a critical improvement given DNS's vulnerability to security breaches. These efforts underline ML's potential to both optimize performance and safeguard DNS operations.

Protocol Identification: This is the traditional task in network management where the studies conducted by Soysal et al. (2011) and Zhang et al. (2013) was based on the applications of Bayesian networks, decision trees, and random forests, to classify different types of network traffic. These issues address the increased complexity in the network traffic with automation when the emerging applications have changed protocols dynamically like zero day applications. Zhang's implementation of random forests for clustering and zero-day detection is one of the scalable approaches that monitor evolving network traffic patterns, thus showing the applicability and robustness of both supervised and unsupervised methods in handling different types of traffic scenarios.

Quality of Experience (QoE): Optimization of QoE, especially video streaming, is critical in terms of user satisfaction. Latency-sensitive applications are dependent on user satisfaction. Shafiq et al. (2014) and Balachandran et al. (2016) mention that decision trees and other predictive models can be used for maintaining a high QoE since the time spent in buffering is minimized, and video bit rates are changed in real- time. Such studies highlight how ML helps networks become adaptable in real-time and directly enhance user experiences and satisfaction. For example, large-scale QoE models based on data from millions of video sessions are developed in Balachandran that support the model's scalability and relevance to high-traffic environments.

This also shows that ML plays an important role in the reinforcement of network security by employing neural networks and SVMs in the detection of anomalies and the analysis of encrypted traffic in studies. Oh et al. applied multi-layer perceptrons (MLPs) and convolutional neural networks (CNNs) to effectively fingerprint web traffic, even when encrypted, thereby reinforcing security protocols. Similarly, Marnerides et al. (2015) used SVMs for anomaly detection in mobile networks, showcasing how ML aids in adaptive security measures and threat detection, a key feature for safeguarding networks against sophisticated, evolving attacks.

Resource management in cloud computing: In this approach, resource optimization in cloud computing is critical, which decreases the operating costs while enhancing performance. For example, Van et al. (2018) and Li et al. (2017) focused on the frameworks of resource prediction, whereas HUYGENS enhanced the capabilities of clock synchronization and Ernest framework utilized the predictive allocation of resources through random forests. These examples illustrate how the ML models make cloud infrastructure efficient by dynamically allocating resources according to demand, in large-scale distributed systems by balancing cost and performance. Traffic Forecasting: Nevat et al. (2014) and Wang et al. were utilizing LSTM networks along with clustering techniques to predict traffic patterns. This thus makes advanced congestion management and effective use of resources possible. Time-series-based traffic models by network operators help pre-allocation of resources while ensuring smooth network operations because usage spikes can be easily forecasted.

Conclusion

The paper addresses the exponentially increasing complexity of modern network systems and the resulting growth demand for data-driven and adaptive solutions. Approaches like rule-based, predefined sets or simple statistical models have failed to cope with networks that are expanding at the fast pace of growing traffic, new applications, and increased security risks. The processing and analysis of gigantic amounts of network data using machine learning have become a transforming tool in networking, further helping to improve system performance, enhance security, and achieve optimal Quality of Experience (QoE) for the users. This review will be based on the most critical applications of ML in six networking domains, such as DNS, protocol identification, QoE optimization, network security, cloud resource management, and traffic prediction. This paper demonstrates that ML methods are addressing specific network challenges and driving efficiency, security, and user satisfaction through the systematic analysis of recent research in this field.

This review broadly covers a type of customized ML models fitting into specified needs of a network. For the reasons of security in DNS protocol classification and QoE prediction purposes, some commonly used and frequently applied techniques of supervised learning are: decision tree, random forests, SVMs. Such models are preferable for their interpretability and adaptability to network environments. Unsupervised learning methods, particularly clustering algorithms such as K-means and hierarchical clustering, are well-suited for discovery of unknown or emerging network patterns, including zero-day applications in protocol identification. Deep learning models, which include neural networks and LSTM networks, are central applications requiring time-series analysis or handling encrypted traffic. Other than that, these models have outperformed other models in predicting network traffic, anomaly detection, and resource management. Complex nonlinear patterns within large data sets can be captured through these models, and a hybrid and ensemble technique could combine several methods of ML to enhance their robustness and accuracy for real-time applications and dynamic adjustments.

Results of these methods manifest major development in networking spaces.

In DNS, efficiency in ML-based caching and good malicious resolver identification have been established using clustering techniques. Supervised learning in protocol identification has made classification accuracy high in traffic identification and flexibility in the adaptation of new traffic patterns with respect to clustering. Video streaming applications have seen optimization models that are successful; predictive models minimize buffering time and improve the satisfaction levels of users. DL-based models have increased the improvement of network security in traffic analysis while going encrypted; SVMs proved to be reliable anomaly detectors. In cloud computing, ML brought sophisticated resource management that helps optimize performance distributed and costing. Lastly, the network operators could implement an early warning system regarding congestion as LSTM networks are applied together with methods in clustering and traffic prediction.

Looking ahead, the opportunities in networking with ML are numerous. When 5G, Internet of Things (IoT), and edge computing arrived and continue to develop as per their pattern, demand is likely to be there for scalable solutions and large-scale ML models; more particularly, these shall require much larger volumes with ever-complex traffic scenarios. Future research would be on developing more interpretable and transparent deep learning models that are so far considered "black boxes." Another promising solution in networks that require both the real- time insights and preserve privacy, such as that involving IoT devices, would be federated learning: train ML models across decentralized devices without centralizing data. Mechanisms of continuous improvement will be very vital in allowing ML models the ability to adapt autonomously to emerging network patterns, threats, and performance demands. Thus, highly adaptive and robust networking solutions can arise from enhanced model retraining with real- time data integration. The future of ML within networking is promising toward offering robust, self-optimizing networks that can look into the future and understand beforehand what the user's demand is, what environmental and security challenges are on ground, and thus enable this era of highly intelligent network management.

References

[1]. Hao, S., and H. Wang. "Exploring Domain Name-Based Features on the Effectiveness of DNS Caching." ACM SIGCOMM Computer Communication Review, vol. 47, no. 1, 2017, pp. 36–42.

[2]. Kührer, M., T. Hupperich, J. Bushart, C. Rossow, and T. Holz. "Going Wild: Large-Scale Classification of Open DNS Resolvers." Proceedings of the Internet Measurement Conference (IMC), 2015, pp. 355–368.

[3]. Bilge, L., E. Kirda, C. Kruegel, and M. Balduzzi. "Exposure: Finding Malicious Domains Using Passive DNS Analysis." NDSS, 2011.

[4]. Yadav, S., A. K. K. Reddy, A. Reddy, and S. Ranjan. "Detecting Algorithmically Generated Malicious Domain Names." Proceedings of the Internet Measurement Conference (IMC), 2010, pp. 48–61.

[5]. Soysal, M., and E. G. Schmidt. "Machine Learning Algorithms for Accurate Flow-Based Network Traffic Classification: Evaluation and Comparison." Performance Evaluation, vol. 67, no. 6, 2010, pp. 451–467.

[6]. Michael, A. K. J., E. Valla, N. S. Neggatu, and A. W. Moore. Network Traffic Classification via Neural Networks. University of Cambridge, Computer Laboratory, Technical Report, 2017.

[7]. Yun, X., Y. Wang, Y. Zhang, and Y. Zhou. "A Semantics-Aware Approach to the Automated Network Protocol Identification." IEEE/ACM Transactions on Networking, vol. 24, no. 1, 2016, pp. 583–595.

[8]. Zhang, J., X. Chen, Y. Xiang, W. Zhou, and J. Wu. "Robust Network Traffic Classification." IEEE/ACM Transactions on Networking, vol. 23, no. 4, 2015, pp. 1257–1270.

Spectrum of Engineering Sciences
SPECTRUM OF ENGINEERING SCIENCES
Online ISSN
3007-3138
Print ISSN
3007-312X

[9]. Santiago del Rio, P. M., D. Rossi, F. Gringoli, L. Nava, L. Salgarelli, and J. Aracil. "Wire-Speed Statistical Classification of Network Traffic on Commodity Hardware." Proceedings of the Internet Measurement Conference (IMC), 2012, pp. 65–72.

[10]. Wang, Y., and R. Nelson. "Identifying Network Application Layer Protocol with Machine Learning." PAM- International Conference on Passive and Active Network Measurement, 2009.

[11]. Gonzalez, R., et al. "Behavioral Profiling of Web Users for Improved Marketing." IEEE Transactions on Knowledge and Data Engineering, 2018.

[12]. Li, J., et al. "Ernest: Efficient Performance Prediction for Large-Scale Analytics." Proceedings of the VLDB Endowment, 2017.

[13]. Shafiq, M. Z., et al. "Understanding the Impact of Network Dynamics on Mobile Video User Engagement." ACM SIGMETRICS, 2014.

[14]. Zhu, L., and N. Laptev. "Deep and Confident Prediction for Time Series at Uber." Data Mining Workshops (ICDMW), IEEE, 2017.

[15]. How DNS Works. Pixel Privacy, www.pixelprivacy.com/resources/speeding-up-your-internet/.

[16]. What Is Network Monitoring? Definition, Benefits, Tools. Wallarm, www.wallarm.com/what/what-is-network- monitoring-definition-benefits-tools.

[17]. Factors Impacting Quality of Experience. ResearchGate, www.researchgate.net/figure/Users-quality-of- experience-QoE.

[18]. The Evolution of Cloud Computing in ATLAS. ResearchGate, www.researchgate.net/figure/A-high-level-view- of-the-distributed-cloud-computing.

Spectrum of Engineering Sciences
SPECTRUM OF ENGINEERING SCIENCES
Online ISSN 3007-3138
Print ISSN 3007-312X

[19]. Overview of the Network Traffic Prediction Algorithm. ResearchGate, www.researchgate.net/figure/Overview- of-the-network-traffic-prediction-algorithm.

[20]. Decision Tree Process. ResearchGate, www.researchgate.net/figure/fig2_225237661.

[21]. Liu, D., et al. "Opprentice Architecture." Opprentice: Towards Practical and Automatic Anomaly Detection, IMC.

[22]. Support Vector Machine Mechanism. ResearchGate, www.researchgate.net/figure/Support-Vector-Machine-Mechanism_fig1_344827087.

[23]. Flowchart of KNN. ResearchGate, www.researchgate.net/figure/Flowchart-of-KNN-Method.

[24]. Hierarchical Clustering. ResearchGate, www.researchgate.net/figure/Flowchart-of-hierarchical-clustering.

[25]. Spectral Clustering Analysis by Hyper Parameter Tuning. ResearchGate, www.researchgate.net/figure/Spectral-clustering-analysis-by-hyperparameter-tuning.

[26]. LSTM Memory Cell Structure. ResearchGate, www.researchgate.net/figure/The-structure-of-LSTM-memory- cell.

[27]. Example of MLP with Three Inputs. ResearchGate, www.researchgate.net/figure/An-Example-of-MLP-with- three-inputs.

[28]. SVR Nature-Inspired Algorithms. ResearchGate, www.researchgate.net/figure/Flowchart-of-the-SVR-nature-inspired-algorithms.

[29]. Liu, D., et al. "Opprentice Work Process." Opprentice: Towards Practical and Automatic Anomaly Detection, IMC.

[30]. Khandelwal, Manoj, and K. M. R. Rao. "Evolving Support Vector Regression Using Grey Wolf Optimization; Forecasting the Geomechanical Properties of Rock." Measurement".

[31]. Goodfellow, I., Y. Bengio, and A. Courville. Deep Learning. MIT Press, 2016.

[32]. Nguyen, T. T., and G. Armitage. "A Survey of Techniques for Internet Traffic Classification Using Machine Learning." IEEE Communications Surveys & Tutorials, 2008.

[33]. Buczak, A. L., and E. Guven. "A Survey of Data Mining and Machine Learning Methods for Cybersecurity Intrusion Detection." IEEE Communications Surveys & Tutorials, 2016.

[34]. Li, Y., L. Chen, and H. Shen. "A Deep Learning Method for Anomaly Detection Based on Auto Encoder." IEEE Transactions on Networks and Service Management, 2018.

[35]. Estevez-Tapiador, J. M., P. Garcia-Teodoro, and J. Diaz-Verdejo. "Anomaly Detection Methods in Wired Networks: A Survey and Taxonomy." Computer Communications, 2004.

[36]. Wu, Y., and Y. Zhang. "Enhancing QoE for Video Streaming Using Machine Learning." IEEE Transactions on Multimedia, 2020.

[37]. Qi, L., and W. Dou. "Cloud Resource Scheduling in Mobile Edge Computing Using Deep Reinforcement Learning." IEEE Transactions on Parallel and Distributed Systems, 2020.

[38]. Munos, R., and C. Szepesvári. "Finite-Time Bounds for Fitted Value Iteration." Journal of Machine Learning Research, 2008.

[39]. Zhou, Z., and J. Zeng. "Traffic Prediction in Wireless Networks Using Deep Learning." IEEE Transactions on Wireless Communications, 2018.

Spectrum of Engineering Sciences
SPECTRUM OF ENGINEERING SCIENCES
Online ISSN
3007-3138
Print ISSN
3007-312X

[40]. Hinton, G., and R. Salakhutdinov. "Reducing the Dimensionality of Data with Neural Networks." Science, vol. 313, no. 5786, 2006, pp. 504–50