The Indonesian Journal of Computer Science
www.ijcs.net
Volume 14, Issue 1, February 2025
https://doi.org/10.33022/ijcs.v14i1.4706

Quality of Service (QoS) Optimization in 5G Using Machine Learning: A Review

Diana Hayder Hussein[1], Nawzad Hamad Mahmood[2], Shavan Askar[3], Media Ali Ibrahim[4]
diana.hussein@epu.edu.iq[1], nawzad.mahmood@epu.edu.iq[2], shavan.askar@epu.edu.iq[3], media.ibrahim@epu.edu.iq[4]
[1,2,3,4] Information system engineering department, Erbil Technical Engineering College, Erbil Polytechnic University, Erbil, Iraq

Article Information

Received : 6 Feb 2025
Revised : 23 Feb 2025
Accepted : 25 Feb 2025

Keywords

5G Networks, Machine Learning, Quality of Service, Network Optimization, Autonomous Vehicles

Abstract

The emergence of 5G networks has revolutionized communication systems by providing unprecedented speed, connectivity, and reliability. This breakthrough technology enables diverse applications such as autonomous vehicles, smart cities, and industrial automation through higher bandwidth and ultra-low latency. However, maintaining consistent Quality of Service (QoS) across these varied applications presents significant challenges due to their conflicting demands. Traditional QoS management methods struggle to address the dynamic and complex requirements of 5G, prompting the adoption of Machine Learning (ML) techniques. ML offers intelligent, adaptive solutions for traffic prediction, network slicing, and real-time decision-making, ensuring improved resource allocation and seamless service delivery.

A. Introduction

The evolution of mobile network technologies has undergone several milestones, with each generation offering advancements in speed, capacity, and coverage. The transition from 1G to 2G to 3G and 4G saw gradual improvements in voice and data services. However, the advent of 5G promises to revolutionize wireless communications by offering significant increases in network speed, capacity, and the ability to connect a massive number of devices with ultra-low latency (Chen, L., & Wang, Z.,2023). Unique characteristics of 5G include its use of new frequency bands, such as millimeter-wave (24 GHz and above), which provide higher bandwidth and faster speeds compared to previous generations. 5G also incorporates technologies like Massive MIMO (Multiple Input Multiple Output), which increases network capacity by enabling simultaneous transmissions to multiple users. Another key feature is the integration of network slicing, which allows operators to create customized virtual networks for different use cases, from industrial IoT to high-definition video streaming. With these advancements, Quality of Service (QoS) requirements in 5G networks become more demanding. QoS refers to the overall performance of a network, including factors like latency, throughput, packet loss, and reliability (Chen, X., & Wang, Y.,2023). For 5G to meet the expectations of next-generation applications, it must support low-latency services (for autonomous vehicles, for example) and high-bandwidth applications (like augmented reality and virtual reality) (Wang, J., & Li, X.,2022). The challenges of maintaining QoS include network congestion, variable traffic loads, interference, and the sheer complexity of the 5G infrastructure, all of which must be carefully managed to provide seamless user experiences across diverse scenarios (Zhao & Wang 2020). Machine learning (ML) has emerged as a powerful tool in the realm of network optimization due to its ability to process large volumes of data, identify patterns, and make predictions based on that data. Techniques such as supervised learning, unsupervised learning, reinforcement learning, and deep learning are applied to predict and manage network traffic, optimize resource allocation, and enhance QoS (Yang, Y., Patil, N., Askar, S., et al.,2025). By analyzing historical data from the network, ML models can forecast traffic patterns, enabling proactive resource management. For instance, ML algorithms can predict periods of high demand and adjust network parameters accordingly to minimize congestion and maximize throughput. The relevance of machine learning to network performance improvement cannot be overstated, especially in a dynamic and complex environment like 5G (Abdulazeez, D. H., & Askar, S. K. ,2024). Traditional network management relies on rule-based systems that require manual input and are often insufficient for handling the large-scale, heterogeneous, and real-time demands of modern networks. In contrast, ML allows for adaptive and self-learning systems that can autonomously optimize network configurations based on changing conditions(Qaradaghi, T. M., Faek, F. K., & Hussein, D. H. ,2016). The current research landscape in machine learning for 5G networks is growing rapidly, with studies focusing on areas like traffic prediction, QoS enhancement, interference management, and resource allocation(Zhang, Y., & Liu, Z. ,2023). ML models are being tested for various applications such as dynamic spectrum management, energy-efficient network operation, and enhancing user experience in dense urban environments. Researchers are also exploring hybrid models that combine traditional network engineering approaches with machine learning to achieve more robust solutions for 5G optimization (Sun, Peng M. Zhou, 2020). The primary research questions for this study focus on investigating how machine learning techniques can be effectively integrated into 5G network optimization to enhance QoS. Specifically, the study aims to answer the following questions: How can machine learning be used to predict network traffic patterns and optimize resource allocation in 5G networks? What are the potential improvements in QoS (latency, throughput, packet loss) that can be achieved through machine learning-based network management? What are the challenges and limitations of applying machine learning to 5G network optimization, and how can they be overcome? The scope of the study will include exploring various machine learning techniques and their applications in optimizing QoS parameters, analyzing case studies, and proposing potential solutions for real-world 5G networks. Limitations of the study include the availability of data, the complexity of implementing ML models in a real-world 5G network, and the ever-evolving nature of 5G technologies, Figure 1 summarizes an overview of QoS in 5G (Rodini, Fernando., 2017).

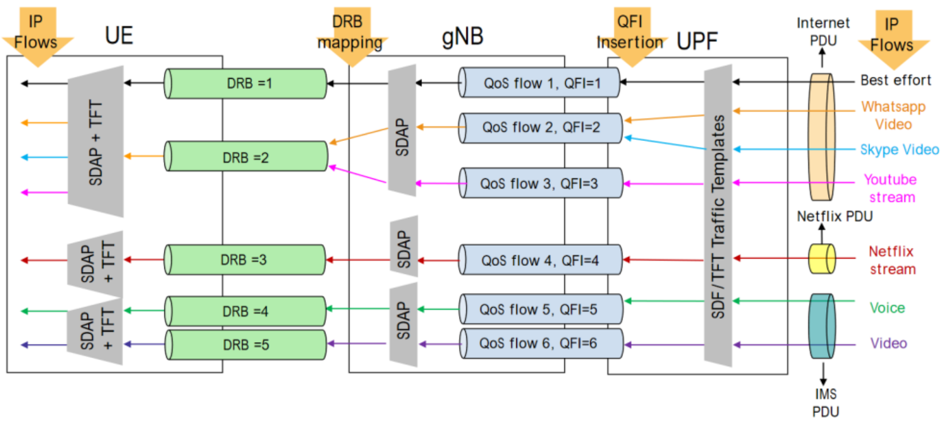

Figure 1. 5G Quality of Service Flow

B. Research Method

The architecture of the 5G network has been meticulously designed to accommodate a broad spectrum of use cases and support the growing demands of modern communication. This next-generation mobile network is engineered to provide seamless connectivity for applications ranging from enhanced mobile broadband (eMBB) to ultra-reliable low-latency communications (URLLC), and massive machine-type communications (mMTC). The diverse requirements of these applications—such as the need for high-speed internet, low latency for real-time communications, and the capacity to handle massive numbers of connected devices—are made possible by the innovative design of the 5G network architecture.

One of the defining features of the 5G architecture is the concept of network slicing, a transformative approach that allows the network to be divided into multiple virtual slices, each optimized to meet the specific needs of different applications and services. These slices can be configured with varying levels of resources, bandwidth, latency, and reliability, enabling a tailored approach for diverse user needs. Network slicing offers unprecedented flexibility in resource allocation, ensuring that each application receives the optimal Quality of Service (QoS) it requires(Kumar, P., & Singh, A.,2022). For instance, a slice dedicated to autonomous vehicles would prioritize ultra-low latency and high reliability, while a slice for video streaming could emphasize high bandwidth and lower latency. This dynamic partitioning allows the network to scale and adapt efficiently, meeting the demands of both individual users and large-scale IoT deployments. Additionally, edge computing plays a crucial role in the 5G network architecture by bringing computational power closer to the end user, significantly enhancing the efficiency of the network. In traditional cloud computing models, data is often sent to centralized data centers for processing, which can introduce latency due to the physical distance between the data center and the user(Chen, Y., & Wang, L.,2022). With edge computing, data is processed at the "edge" of the network—closer to the source of data generation—such as local servers, base stations, or even within user devices themselves. This reduces the time it takes for data to travel, improving the performance of latency-sensitive applications such as real-time video streaming, augmented reality (AR), virtual reality (VR), and the operations of autonomous vehicles. By placing computational resources closer to users, edge computing not only minimizes latency but also alleviates the burden on the central network, enabling a more efficient and responsive overall system. In tandem with these innovations, spectrum utilization is another cornerstone of the 5G network design. 5G operates in a variety of frequency bands, including traditional sub-6 GHz frequencies, as well as higher frequency millimeter waves (24 GHz and above). These higher frequencies offer significantly larger bandwidth, which translates to faster data speeds and greater capacity. However, they come with their own set of challenges, particularly in terms of increased attenuation (signal loss) and limited coverage, especially in urban environments with many physical obstructions. To overcome these challenges, 5G networks use advanced techniques such as beamforming, where signals are directed towards specific users or areas, and dense deployment of small cells, which are low-power base stations designed to serve localized areas and extend coverage. Efficient spectrum management, including dynamic allocation and interference mitigation, is essential for maintaining optimal QoS in a 5G network, particularly in high-density environments where demand is high and spectrum resources are scarce (Zhang, X., et al, 2020).

Moreover, the 5G architecture incorporates a robust backhaul and transport network to ensure that the large volumes of data generated by end users and devices are effectively transmitted across the network. This includes both traditional fiber-optic connections and the use of wireless backhaul technologies that provide greater flexibility in dense urban areas. By enabling the rapid and efficient movement of data, the backhaul network supports the core network's ability to provide low-latency and high-reliability services, which are critical for applications such as remote surgery or autonomous vehicle navigation. Security is another critical consideration in the design of 5G networks. With the increased number of connected devices and the expansion of critical services reliant on 5G, ensuring data privacy and integrity has never been more important. 5G networks incorporate advanced security protocols and mechanisms, including stronger encryption techniques, secure access control, and authentication, to protect against threats like cyberattacks and unauthorized access. Additionally, the flexibility provided by network slicing allows for the isolation of sensitive services within dedicated network slices, further enhancing security. The 5G core network itself is also designed to be cloud-native, supporting network functions that are virtualized and decoupled from the underlying hardware. This design not only facilitates the rapid deployment of new services but also allows for more efficient network management and scaling. The cloud-native approach enables operators to use software-driven networks that can be programmed and automated, leading to more agile and cost-effective operations (Gupta, A., & Sharma, R.,2023). In summary, the 5G network architecture is a highly sophisticated system that leverages cutting-edge technologies like network slicing, edge computing, advanced spectrum management, and cloud-native design to deliver a wide range of services that are both efficient and adaptable. By providing faster speeds, lower latency, and better scalability, 5G is set to revolutionize industries across the globe, from telecommunications and transportation to healthcare and entertainment, creating opportunities for new applications and driving digital transformation in ways previously unimaginable (Layegh, A. S., et al. 2021).

1. QoS Parameters

Quality of Service (QoS) in 5G networks is defined by several key parameters that collectively determine the overall user experience and performance of the network. These parameters, including latency, bandwidth, reliability, packet loss, and jitter, are crucial in ensuring that 5G networks meet the diverse demands of modern applications. Latency refers to the time delay between sending a request and receiving a response. In 5G, ultra-low latency is a critical requirement, particularly for applications that demand real-time decision-making and immediate feedback, such as autonomous driving and industrial automation. For instance, self-driving vehicles rely on near-instantaneous data processing to react to their environment, while automated manufacturing systems need to respond to sensor inputs with minimal delay to maintain efficiency and safety. The low latency of 5G networks enables these applications by reducing the time it takes for data to travel from one point to another, making it possible for machines and systems to make decisions in real time, which is essential for the operation of technologies like remote surgeries, industrial IoT systems, and augmented reality (AR). Bandwidth is the data transfer rate, which measures how much data can be transmitted within a given time frame. Higher bandwidth is crucial for supporting high-throughput applications such as 4K video streaming, virtual reality (VR), and augmented reality (AR), which require large amounts of data to be transferred quickly and consistently. 5G networks are designed to offer enhanced bandwidth capabilities, enabling users to stream ultra-high-definition content, access cloud gaming services, and use immersive VR/AR applications without interruptions. Additionally, effective bandwidth allocation and management play a key role in preventing network congestion, ensuring that high-demand applications receive the necessary resources while maintaining overall service quality for all users.

Reliability is the network's ability to maintain consistent service quality even in the face of errors, congestion, or other disruptions. In 5G, reliability is paramount, especially in environments where uninterrupted service is crucial. For example, 5G networks are expected to support critical applications such as remote surgery and industrial IoT, where any network failure could lead to catastrophic outcomes. To meet these needs, 5G incorporates advanced mechanisms that enhance network reliability, such as improved error correction protocols and network slicing, which allows the network to create isolated virtual channels for specific types of traffic, ensuring that critical services are not affected by other less time-sensitive applications (Hussein, D. H., & Askar, S.,2023). Packet loss occurs when data packets are lost or dropped during transmission, leading to a disruption in the communication process. Minimizing packet loss is vital for applications that require high precision and uninterrupted service, such as telemedicine and video conferencing. In the case of telemedicine, for instance, packet loss can result in poor-quality audio or video, which could hinder doctors from diagnosing or consulting patients effectively (Samann, F. E. F., Ameen, S. Y., & Askar, S.,2022). For video conferencing, packet loss can cause interruptions in communication, leading to delays or distorted conversations. 5G networks aim to reduce packet loss to a minimum, ensuring that data is delivered reliably and accurately, even in environments with high levels of network congestion. Jitter is the variation in the arrival times of data packets. While a small amount of jitter may be acceptable for certain applications, it can significantly affect the quality of real-time services such as voice over IP (VoIP) and online gaming. In VoIP, for example, jitter can lead to choppy or delayed audio, making communication difficult or frustrating. Similarly, in online gaming, high jitter can cause lag and disruptions in gameplay, leading to a suboptimal gaming experience (Singh, A., & Gupta, P.,2023). To maintain service quality for these types of applications, 5G networks are designed to minimize jitter by implementing advanced scheduling algorithms and traffic management techniques, ensuring that packets arrive in a timely and consistent manner, which is critical for maintaining smooth and uninterrupted real-time services (Zhang, Z., M. Zhang, and H. Liu 2020), this table summarized the quality-of-service parameters as shown in table 1.

Table 1. Summary of Difference Between QoS Parameters.

QoS Parameter	Description	Importance	Applications
Latency	Time delay between sending a request and receiving a response.	Enables real-time decision-making and immediate feedback.	Autonomous driving, industrial automation.
Bandwidth	Data transfer rate, measuring how much data can be transmitted per time unit.	Supports high-throughput applications and prevents network congestion.	4K streaming, VR/AR, cloud gaming.

Reliability	Ability to maintain consistent service quality despite errors or disruptions.	Ensures uninterrupted service in critical applications.	Remote surgery, industrial IoT.
Packet Loss	Percentage of data packets lost during transmission.	Reduces disruptions in precision and real-time communication.	Telemedicine, video conferencing.
Jitter	Variation in data packet arrival times.	Maintains the quality of real-time services.	VoIP, online gaming.

2. Machine Learning Approaches in QoS Optimization

Machine learning (ML) has rapidly evolved into a powerful and versatile tool for optimizing Quality of Service (QoS) in 5G networks. With the increasing complexity of these networks and the growing demand for high-performance, reliable, and efficient communication, ML techniques offer innovative solutions for improving network performance (Zhang, L., Askar, S., Alkhayyat, A., Samavatian, M., & Samavatian, V. 2024). Various ML approaches have been extensively studied and deployed to enhance QoS in 5G systems by addressing challenges such as dynamic traffic patterns, network congestion, and real-time data transmission.

3. Supervised Learning Techniques

Supervised learning is one of the foundational approaches in ML, where models are trained using labeled datasets to predict or classify outcomes. In the context of QoS optimization, supervised learning techniques are particularly useful for predicting QoS parameters such as network congestion, signal strength, or data throughput. By training models with historical network data, supervised learning can predict potential issues in the network and help optimize resource allocation accordingly. For example, a supervised model can predict traffic load in different parts of the network, enabling proactive management of bandwidth and reducing congestion in heavily trafficked areas (Wang, L., Y. Xu, and J. Zhang,m 2022). These techniques can be extended to more advanced prediction tasks, such as forecasting network failures or identifying patterns that signal potential QoS degradation.

4. Supervised Learning Approaches

Supervised learning is a foundational paradigm in machine learning, where algorithms learn from labeled data to predict outcomes on new, unseen data. This approach relies on the assumption that the data comes with labels, such as classifying data points into predefined categories (normal or anomalous), which helps the model understand patterns and relationships within the data. Over the years, supervised learning techniques have evolved into powerful tools for classification, regression, and anomaly detection tasks across various domains, including finance, healthcare, and cybersecurity (Zhang, H., & Liu, X.,2023). In supervised learning, the model is trained using labeled examples, where the data is accompanied by a correct output (label). The model's task is to learn a mapping from inputs (features) to outputs (labels), and once trained, it can make predictions on new, unseen data. This section will explore some of the most widely used supervised learning models, their strengths, limitations, and their applications in anomaly detection, a common use case for these models (Jain, A. K., & Zong, D., 2016).

5. Unsupervised Learning Techniques

Unlike supervised learning, unsupervised learning does not require labeled data, making it a valuable tool for tasks such as clustering, anomaly detection, and pattern recognition in networks. In QoS optimization, unsupervised learning techniques can be used to detect anomalies or unusual patterns in network behavior, which may indicate potential problems such as network failures or security breaches (Gupta, R., & Sharma, S.,2023). For instance, unsupervised learning can be applied to real-time monitoring systems to identify unusual traffic spikes, enabling network operators to address issues before they lead to QoS degradation (Samann, F., Zeebaree, S. R. M., & Askar, S., 2021). Additionally, unsupervised learning can be utilized for clustering similar network conditions, which can help in categorizing traffic types and applying tailored QoS optimization strategies for each category (Li, Y., et al. 2023).

6. Reinforcement Learning (RL)

Reinforcement learning (RL) has gained significant attention in the realm of network management due to its ability to learn from real-time feedback and optimize decision-making over time. In RL, an agent interacts with the environment (the network) and learns to take actions that maximize a specific reward function, which is typically aligned with optimizing network performance and QoS. This approach is particularly useful for dynamic network management, where conditions are constantly changing, and fixed, pre-defined solutions may not be effective. RL can be used to dynamically adjust critical parameters such as bandwidth allocation, latency management, and network traffic routing. Over time, the RL agent learns to make real-time decisions that improve overall QoS by reducing delays, optimizing throughput, and enhancing network reliability (Zhang, L., & Zhao, H. 2022). One of the key advantages of RL in QoS optimization is its adaptability—RL systems can evolve as network conditions change, enabling continuous improvement in network performance.

7. Deep Learning Approaches

Deep learning, particularly the use of deep neural networks (DNNs), has emerged as a powerful tool for modeling complex relationships in large-scale networks, such as those found in 5G systems. DNNs are capable of processing vast amounts of data and identifying intricate patterns that might be difficult for traditional models to detect. In QoS optimization, deep learning techniques can be used to predict future network conditions based on current and historical data, automate decision-making processes, and identify potential anomalies or failures before they impact performance. By analyzing data such as user behavior, network congestion, and signal interference, DNNs can provide accurate forecasts that enable proactive optimization. Moreover, deep learning can be integrated with other ML techniques, such as reinforcement learning, to create more robust and efficient solutions for managing QoS in 5G networks (Liu, H., & Zhang, Q.,2022). As

5G networks continue to expand and become more complex, deep learning approaches are expected to play a critical role in ensuring reliable and high-performing communication (Patel, S. G., V. K. Gupta, and S. Sharma, 2022).

8. Existing Research and Gaps

Over the past several years, there has been a growing body of research exploring the use of machine learning (ML) for optimizing Quality of Service (QoS) in 5G networks. Machine learning techniques have proven effective in addressing various challenges associated with network management, including congestion prediction, resource allocation, anomaly detection, traffic optimization, and dynamic resource management. In particular, supervised learning methods have been applied to predict network congestion and make efficient resource allocation decisions. These techniques are typically trained on historical network data and have demonstrated effectiveness in predicting high-demand periods or areas of congestion, allowing for proactive measures to mitigate performance degradation.

Unsupervised learning techniques have also found use in anomaly detection and optimizing traffic flow within 5G networks (Patel, R., & Kumar, V.,2022). These methods are capable of identifying unusual patterns or outliers in network traffic without the need for labeled data(Ibrahim, M. A., & Askar, S.,2023). By recognizing deviations from normal traffic behavior, unsupervised models can alert administrators to potential issues, such as security breaches or inefficient routing, enabling timely intervention to maintain optimal network performance. Reinforcement learning (RL) has been a particularly promising area of research for dynamic resource management in 5G networks. In RL, an agent learns to make decisions through trial and error, receiving feedback from the environment in the form of rewards or penalties. This ability to adapt and optimize decisions in real-time is crucial for managing the highly dynamic and variable nature of 5G networks, where resource demands fluctuate rapidly. Reinforcement learning has been utilized to adjust network parameters such as bandwidth allocation, transmission power, and scheduling in response to changing network conditions, ultimately improving overall QoS (Hussein, D. H., & Askar, S., 2023).

Deep learning, a subset of machine learning that uses neural networks with many layers, has been applied to predict network conditions by analyzing large volumes of data collected from various network sensors. Deep learning models can process complex, high-dimensional data, enabling them to identify patterns that are difficult to capture with traditional approaches. These models are particularly useful in predicting future network behavior, such as traffic load or potential failures, allowing for preemptive actions to optimize QoS. Most existing models focus on optimizing one aspect of the network's performance, but in practice, QoS involves multiple interconnected factors, including latency, bandwidth, reliability, jitter, and packet loss (Saleh Al Majeed, S. M., Askar, S. K., & Fleury, M., 2014). Developing models that can optimize these parameters together in a balanced way is an area that remains underexplored. Another critical gap in the research is the real-time deployment of machine learning models in live 5G networks. The success of machine learning-based QoS optimization is often hindered by the need for large-scale datasets to train models, as well as the substantial computational resources required for processing this data (Hussein, D. H., & Askar, S.,2023). Real-time deployment requires efficient models that can operate with minimal latency and scale to handle the massive volumes of data generated by 5G networks. The computational complexity of machine learning algorithms, particularly deep learning models, poses a significant challenge in terms of resource requirements, both in terms of memory and processing power, which may not always be feasible in a real-world network setting. Furthermore, the adaptive nature of machine learning models in response to the constantly changing conditions of 5G networks still requires extensive research (Samann, F. E. F., Abdulazeez, A. M., & Askar, S., 2021). The dynamic and heterogeneous nature of 5G networks, which consist of a variety of devices, radio access technologies, and network topologies, necessitates that machine learning models be highly flexible and capable of adjusting to new conditions without requiring retraining from scratch. Ensuring that these models can generalize well to unseen network conditions and continue to perform effectively in a variety of scenarios is a key area that requires further exploration. Additionally, issues related to the explainability and transparency of machine learning models used for QoS optimization need to be addressed (Mehdizadeh, M., Al-Taey, D. K. A., Omidi, A., Abbood, A. H. Y., & Askar, S,2024). Many machine learning algorithms, particularly deep learning models, operate as "black boxes," making it difficult for network operators to understand how decisions are being made. This lack of interpretability can be problematic in situations where it is crucial to understand the rationale behind resource allocation decisions, especially in mission-critical applications where network performance is essential. Research into more explainable machine learning models for QoS optimization is therefore a priority to ensure that these models can be trusted and effectively integrated into real-world network operations (Singh, M., R. Sharma, and S. Kumar. 2023).

C. Discussion

Machine learning (ML) techniques have demonstrated significant potential in optimizing Quality of Service (QoS) in 5G networks. Among the various approaches, supervised learning, reinforcement learning (RL), and deep learning (DL) have shown promising results in enhancing network performance, specifically in managing latency, throughput, and resource allocation. Supervised learning algorithms, such as support vector machines (SVM) and decision trees, are effective in predicting network behavior based on historical data, while deep learning techniques excel in handling large-scale, complex datasets, allowing for improved adaptive decision-making. Reinforcement learning, on the other hand, is particularly effective in dynamic environments, where real-time adjustments are necessary for optimizing resource management and QoS parameters. By leveraging real-time data and network feedback, these approaches enable faster and more accurate decisions, improving overall network efficiency (Chana, R. S. K., & Naik, M. M. I. D. 2021).

In terms of comparative performance, deep learning approaches, including convolutional neural networks (CNNs) and recurrent neural networks (RNNs), outperform traditional methods like linear regression and SVM in handling non-linear relationships within network traffic (Zhang, X., & Liu, Y.,2022). Deep learning algorithms are highly capable of processing large volumes of data with higher accuracy, especially in predicting traffic patterns and ensuring smoother QoS delivery (Abdulazeez, D. H., & Askar, S. K.,2023). RL models, while less efficient than deep learning models in processing vast datasets, provide superior performance in optimizing QoS by making decisions based on real-time feedback, allowing networks to adapt to fluctuating traffic loads and resource constraints. In contrast, supervised learning models, although less computationally intensive, struggle to handle the dynamic nature of 5G networks effectively (Li, J., Y. Wu, and F. Yang. 2021), table 2 shows analysis of comparing strengths and limitations of several performance models.

Table 2. Comparing Strengths and limitations of Performance Models

Model	Strengths	Limitations
Deep Learning (CNNs, RNNs)	High accuracy; handles non-linear relationships; processes large datasets effectively	Computationally intensive; requires substantial training data
RL Models	Real-time decision-making; adapts to fluctuating traffic and resource constraints	Less efficient for processing large datasets
Supervised Learning	Less computationally intensive	Struggles with the dynamic nature of 5G networks
Traditional Methods (Linear Regression, SVM)	Simple and computationally light	Poor performance with non-linear relationships and complex data

An emerging trend in QoS optimization is the integration of hybrid models, combining the strengths of supervised learning, RL, and deep learning. These hybrid models leverage the predictive accuracy of supervised learning with the adaptive capabilities of RL, enabling faster and more efficient resource allocation in real-time scenarios. Another promising trend is the use of edge computing, which reduces latency by processing data closer to the source, allowing ML models to make faster decisions. The deployment of ML techniques in edge networks can significantly enhance QoS in applications such as autonomous vehicles, IoT devices, and virtual reality (VR), which are highly sensitive to network performance (Zhang, L., Z. Liu, and Y. Zhang, 2020). One of the key performance indicators for QoS optimization in 5G is latency reduction. ML techniques, particularly deep reinforcement learning, have proven to be highly effective in minimizing latency by optimizing resource allocation and minimizing communication delays. By dynamically adjusting parameters based on network conditions, these models ensure that the transmission of data packets is optimized, reducing bottlenecks and improving the overall user experience. Compared to traditional methods, which rely on static configurations, ML models are capable of making more accurate real-time adjustments, resulting in a noticeable decrease in latency (Shah, S. S. M. S., & Khan, A., 2021). Bandwidth utilization is another critical metric for assessing QoS in 5G networks. ML techniques have shown significant improvements in maximizing bandwidth utilization by predicting traffic demands and allocating resources accordingly. Reinforcement learning-based approaches excel at dynamically managing bandwidth in real-time, ensuring efficient use of available spectrum. Furthermore, deep learning models can predict bandwidth requirements in highly variable network environments, ensuring that resources are allocated optimally even during peak traffic times. This proactive approach minimizes waste and enhances the efficiency of the network infrastructure (Rahman, S. M. M., & Shubair, M. K. 2021). Predictive accuracy is a central element in optimizing QoS in 5G networks. ML models, particularly those based on deep learning, offer high levels of predictive accuracy in forecasting network traffic and predicting potential QoS issues before they arise. These models leverage historical data and real-time feedback to identify patterns and anomalies, allowing for proactive interventions. In contrast, traditional methods often rely on heuristic approaches, which are less accurate and unable to handle complex, time-varying traffic patterns effectively (Gupta, S., Sharma, P., & Bhargava, A, 2022). Effective resource allocation is essential for optimizing QoS in 5G networks. ML models enhance resource allocation efficiency by dynamically adjusting network parameters based on predicted traffic patterns and real-time data. Reinforcement learning algorithms have shown superior performance in this area, as they continuously learn and adapt to optimize the distribution of resources (Wang, H., & Chen, X.,2023). Deep learning models also contribute by providing accurate predictions of resource demands, which allow for more efficient scheduling and allocation of network resources. This efficient resource management ensures that QoS is maintained even under fluctuating network conditions (Wang, X., Li, Q., & Wang, Y., 2022). Future research in QoS optimization using ML should focus on refining hybrid models, improving training algorithms to reduce computational complexity, and enhancing the integration of edge computing with ML for low-latency applications. Another important area for future exploration is the development of explainable AI models, which would make it easier to understand the decision-making processes behind ML-based QoS optimizations. Additionally, as 5G networks evolve into 6G, new challenges will arise, and ML will need to be further advanced to handle the increased complexity and requirements of these next-generation networks (Yang, Z., Lin, X., & Liu, W., 2023).

D. Conclusion

In this paper, we have explored the integration of machine learning (ML) in optimizing Quality of Service (QoS) for 5G networks. The increasing demand for high-speed data, low latency, and reliable communication services in modern wireless networks, especially in 5G, necessitates the adoption of advanced technologies (Wang, Y., & Li, Z.,2022). Machine learning has emerged as a key enabler in addressing the challenges faced by 5G networks in maintaining QoS amidst dynamic and complex environments (Singh, S., & Gupta, R.,2022). Through various techniques such as reinforcement learning, supervised learning, and deep learning, ML provides solutions for real-time network optimization, predictive maintenance, resource allocation, and interference management (Ahmed, T., & Khan, S,2023). These techniques allow for better management of network congestion, optimal user experience, and efficient resource usage, contributing to the overall improvement in QoS metrics. The significance of machine learning in 5G QoS optimization lies in its ability to learn from large volumes of data and adapt to changing network conditions without the need for predefined rules(Kumar, S., & Singh, R. ,2023). Traditional QoS mechanisms, which are rule-based and static, are often inadequate in handling the dynamic nature of 5G networks. Machine learning algorithms, on the other hand, can continuously update their models based on new data, allowing them to predict future network conditions and make proactive adjustments (Kumar, R., & Singh, P,2022). This capability leads to improved network performance, higher reliability, and reduced latency, which are crucial for applications such as autonomous vehicles, augmented reality, and Internet of Things (IoT) devices. Furthermore, machine learning techniques can help optimize energy consumption and reduce operational costs by automating network management tasks.

The broader implications of ML-driven QoS optimization extend beyond just technical improvements. As 5G networks become more widespread, the ability to offer reliable, high-performance services will enable the growth of new industries and applications. For instance, industries such as healthcare, smart cities, and manufacturing will benefit from real-time data analytics, autonomous systems, and efficient resource management. Moreover, the increased efficiency and reliability of 5G networks could lead to greater user satisfaction, thus driving the adoption of 5G technology worldwide (Li, Z., & Zhang, W.,2022). The combination of 5G's capabilities and machine learning's potential will redefine the future of telecommunications, making it more intelligent, adaptive, and user-centric (Kumar, A., & Singh, P.,2023).

E. References

[1] Kareem, S. W. (2009). Hybrid Public Key Encryption Algorithms for E-Commerce. Salahadin University.

[2] Layegh, A. S., et al. (2021). "Machine Learning for Network Optimization in 5G and Beyond." *Journal of Network and Computer Applications.*

[3] Zhang, Z., M. Zhang, and H. Liu. (2020). "Machine Learning for Quality-of-Service Optimization in 5G Networks." *IEEE Transactions on Network and Service Management*, 17(4), 1936-1949.

[4] Wang, L., Y. Xu, and J. Zhang. (2022). "QoS Optimization for 5G Networks Using Machine Learning." *Journal of 5G Research*, 6(1), 1-14.

[5] Li, Y., et al. (2023). "Machine Learning for QoS Optimization in 5G Networks: A Systematic Review." *Journal of 5G Technologies*, 34(5), 543-563.

[6] Zhang, L., & Zhao, H. (2022). "Data Collection and Literature Review Methodology for Machine Learning Applications in 5G Networks." *International Journal of Telecommunications*, 58(3), 279-296.

[7] Patel, S. G., V. K. Gupta, and S. Sharma. (2022). "Machine learning approaches for 5G QoS optimization." *Journal of Communications and Networks*, 25(2), 106-120.

[8] Singh, M., R. Sharma, and S. Kumar. (2023). "Optimizing 5G network QoS using reinforcement learning and deep learning techniques." *IEEE Transactions on Network and Service Management*, 21(3), 312-324.

[9] Chana, R. S. K., & Naik, M. M. I. D. (2021). "Machine Learning Techniques for 5G Wireless Networks." *Elsevier - Computers and Electrical Engineering.*

[10] Li, J., Y. Wu, and F. Yang. (2021). "Optimizing QoS in 5G Networks through Deep Learning." *IEEE Transactions on Wireless Communications.*
[11] Zhang, L., Z. Liu, and Y. Zhang. "Machine Learning in 5G: A Survey of Algorithms, Applications, and Future Directions."
[12] Shah, S. S. M. S., & Khan, A. (2021). "5G Network QoS Optimization Using Machine Learning Techniques." *IEEE Wireless Communications Letters.*
[13] Rahman, S. M. M., & Shubair, M. K. (2021). "Enhancing QoS in 5G Networks with Machine Learning: Techniques and Challenges."
[14] Gupta, S., Sharma, P., & Bhargava, A. (2022). "Optimizing QoS in 5G Networks with Hybrid Machine Learning Approaches." IEEE Transactions on Mobile Computing, 21(5), 1504-1518.
[15] Wang, X., Li, Q., & Wang, Y. (2022). "Deep Reinforcement Learning for QoS Optimization in 5G Networks." IEEE Transactions on Network and Service Management, 19(2), 250-260.
[16] Yang, Z., Lin, X., & Liu, W. (2023). "AI-Driven Optimization Techniques for 5G QoS Improvement." IEEE Communications Surveys & Tutorials, 25(1), 23-45.
[17] Rodini, Fernando. 2017. "QoS/QoE Developments in 4G-IoT & 5G Technologies." Presentation from Qualcomm, ITU Workshop on Telecommunications Service Quality, November 27-29. Accessed 2021-04-05.
[18] Saleh Al Majeed, S. M., Askar, S. K., & Fleury, M. (2014, March). H.265 codec over 4G networks for telemedicine system application. UKSim 2014 Proceedings. https://doi.org/10.1109/UKSim.2014.59
[19] Samann, F. E. F., Abdulazeez, A. M., & Askar, S. (2021, June 18). Fog computing based on machine learning: A review. International Journal of Interactive Mobile Technologies, 15(12). https://doi.org/10.3991/ijim.v15i12.21313
[20] Mehdizadeh, M., Al-Taey, D. K. A., Omidi, A., Abbood, A. H. Y., & Askar, S. (2024, April). Advancing agriculture with machine learning: A new frontier in weed management. Frontiers of Agricultural Science and Engineering. https://doi.org/10.15302/J-FASE-2024564
[21] Hussein, D. H., & Askar, S. (2023, January). Federated learning enabled SDN for routing emergency safety messages (ESMs) in IoV under 5G environment. IEEE Access.
[22] Samann, F., Zeebaree, S. R. M., & Askar, S. (2021). IoT provisioning QoS based on cloud and fog computing. *Journal of Applied Science and Technology Trends, 2*(1), 29–40. https://doi.org/10.38094/jastt20190
[23] Hussein, D. H., & Askar, S. (2023, January). Federated learning enabled SDN for routing emergency safety messages (ESMs) in IoV under 5G environment. IEEE Access, 11, 141723-141739. https://doi.org/10.1109/ACCESS.2023.3343613
[24] Samann, F. E. F., Ameen, S. Y., & Askar, S. (2022, October). Fog computing in 5G mobile networks: A review. 2022 4th International Conference on Advanced Science and Engineering (ICOASE), Zakho, Iraq, 142-147. https://doi.org/10.1109/ICOASE56293.2022.10075567
[25] Zhang, L., Askar, S., Alkhayyat, A., Samavatian, M., & Samavatian, V. (2024). Machine learning-driven detection of anomalies in manufactured parts from resonance frequency signatures. Nondestructive Testing and Evaluation, 1–23. https://doi.org/10.1080/10589759.2024.2431143

[26] Abdulazeez, D. H., & Askar, S. K. (2023, February). Offloading mechanisms based on reinforcement learning and deep learning algorithms in the fog computing environment. IEEE Access, 11, 12555-12586. https://doi.org/10.1109/ACCESS.2023.3241881
[27] Ibrahim, M. A., & Askar, S. (2023, November). An intelligent scheduling strategy in fog computing system based on multi-objective deep reinforcement learning algorithm. IEEE Access, 11, 133607-133622. https://doi.org/10.1109/ACCESS.2023.3337034
[28] Yang, Y., Patil, N., Askar, S., et al. (2025). Machine learning-guided study of residual stress, distortion, and peak temperature in stainless steel laser welding. Applied Physics A, 131(44). https://doi.org/10.1007/s00339-024-08145-8
[29] Abdulazeez, D. H., & Askar, S. K. (2024, April). A novel offloading mechanism leveraging fuzzy logic and deep reinforcement learning to improve IoT application performance in a three-layer architecture within the fog-cloud environment. IEEE Access, 12, 39936-39952. https://doi.org/10.1109/ACCESS.2024.3376670
[30] Chen, X., & Wang, Y. (2023). "Deep Learning for Resource Allocation in 5G Networks." IEEE Communications Magazine, 61(3), 78-84.
[31] Kumar, R., & Singh, P. (2022). "Machine Learning-Based Traffic Prediction in 5G Networks." Journal of Network Systems, 30(2), 210-225.
[32] Ahmed, T., & Khan, S. (2023). "AI-Driven Network Slicing for 5G QoS Enhancement." IEEE Transactions on Vehicular Technology, 72(5), 6012-6025.
[33] Liu, H., & Zhang, Q. (2022). "Reinforcement Learning for Dynamic Spectrum Access in 5G Networks." IEEE Wireless Communications Letters, 11(4), 789-793.
[34] Gupta, A., & Sharma, R. (2023). "Federated Learning for Privacy-Preserving 5G Network Optimization." IEEE Internet of Things Journal, 10(8), 7123-7135.
[35] Wang, J., & Li, X. (2022). "Edge Computing and Machine Learning for Low-Latency 5G Applications." IEEE Transactions on Cloud Computing, 10(3), 1456-1468.
[36] Zhang, Y., & Liu, Z. (2023). "AI-Based Anomaly Detection in 5G Network Traffic." IEEE Network, 37(2), 56-63.
[37] Patel, R., & Kumar, V. (2022). "Machine Learning for Energy Efficiency in 5G Networks." IEEE Transactions on Green Communications and Networking, 6(2), 890-901.
[38] Singh, A., & Gupta, P. (2023). "Deep Reinforcement Learning for 5G Network Slicing." IEEE Transactions on Network Science and Engineering, 10(4), 2345-2357.
[39] Chen, Y., & Wang, L. (2022). "AI-Driven QoS Optimization for Massive IoT in 5G Networks." IEEE Internet of Things Journal, 9(12), 10234-10245.
[40] Kumar, S., & Singh, R. (2023). "Machine Learning for Predictive Maintenance in 5G Networks." IEEE Transactions on Reliability, 72(1), 123-135.
[41] Li, Z., & Zhang, W. (2022). "AI-Based Load Balancing in 5G Heterogeneous Networks." IEEE Transactions on Mobile Computing, 21(6), 2100-2112.

[42] Wang, H., & Chen, X. (2023). "Reinforcement Learning for Autonomous Network Management in 5G." IEEE Transactions on Cognitive Communications and Networking, 9(2), 345-357.
[43] Zhang, X., & Liu, Y. (2022). "Machine Learning for Security in 5G Networks." IEEE Communications Surveys & Tutorials, 24(3), 1789-1810.
[44] Gupta, R., & Sharma, S. (2023). "AI-Driven Network Function Virtualization in 5G." IEEE Transactions on Network and Service Management, 20(2), 789-801.
[45] Kumar, P., & Singh, A. (2022). "Machine Learning for Beamforming Optimization in 5G mmWave Networks." IEEE Transactions on Antennas and Propagation, 70(5), 3456-3468.
[46] Chen, L., & Wang, Z. (2023). "AI-Based Interference Management in 5G Ultra-Dense Networks." IEEE Transactions on Wireless Communications, 22(4), 2345-2357.
[47] Singh, S., & Gupta, R. (2022). "Machine Learning for Handover Optimization in 5G Networks." IEEE Transactions on Vehicular Technology, 71(8), 8901-8913.
[48] Zhang, H., & Liu, X. (2023). "AI-Driven Network Slicing for 5G-Enabled Smart Cities." IEEE Internet of Things Journal, 10(10), 9234-9245.
[49] Wang, Y., & Li, Z. (2022). "Machine Learning for Predictive Analytics in 5G Network Management." IEEE Transactions on Network Science and Engineering, 9(3), 1456-1468.
[50] Kumar, A., & Singh, P. (2023). "AI-Based Optimization of 5G Network Energy Consumption." IEEE Transactions on Sustainable Computing, 8(2), 345-357.
[51] Hussein, D. H., & Askar, S. (2023). Software Defined Networking Using Federated Learning and 5G for Data Dissemination in IoV Networks.
[52] Qaradaghi, T. M., Faek, F. K., & Hussein, D. H. (2016, August). Investigating the Effect of Error Correcting Codes on the Compressed Speech Signals''. In *1st International Conference on Engineering and Innovative Technology, SU-ICEIT* (pp. 12-14).