WWW.IJRPR.COM

International Journal of Research Publication and Reviews

Journal homepage: www.ijrpr.com ISSN 2582-7421

Revolutionizing 6G Networks: Deep Reinforcement Learning Approaches for Enhanced Performance

***V Jaya Sri* [a]**

[a] Btech Student, GMR Institute of Technology, Vizianagaram Dist, 535558, Andhra Pradesh, India
vangarajayasri2005@gmail.com

ABSTRACT

The advent of sixth-generation (6G) wireless networks is set to revolutionize global connectivity by offering unparalleled speed, reliability, and intelligent network management. This paper reviews the significant role of machine learning (ML) in realizing the full potential of 6G networks, focusing on recent advancements, key challenges, and future trends. ML techniques are critical for addressing complex tasks such as resource allocation, energy efficiency, and latency reduction in 6G networks, which will enable new applications like autonomous systems and holographic communication. Despite the promising progress, challenges such as developing scalable ML models, ensuring network security, and integrating AI-driven solutions into existing infrastructures persist. The review categorizes ML applications in 6G based on different network configurations and highlights open research issues, particularly in optimizing the balance between computational complexity and performance. This paper lays the groundwork for future research, emphasizing the importance of ML in shaping the next generation of wireless networks.

Keywords: Machine Learning, 6G Networks, Resource Management, Energy Efficiency, Latency Reduction, Power Optimization, Network Management

Introduction

The transition from 5G to 6G networks represents a groundbreaking step in communication technology, one that promises to redefine how the world connects, interacts, and operates. 6G will provide an extraordinary leap in performance, with data rates projected to reach an unprecedented 1 terabit per second (Tbps) and latency reduced to as low as 0.1 microseconds. This evolution is expected to enable applications and services that were previously unimaginable, positioning 6G as a core technology in the next wave of digital transformation. One of the most exciting aspects of 6G is its potential to power advanced applications requiring ultra-low latency and high bandwidth. Autonomous vehicles, for example, will benefit from near-instantaneous data transmission, allowing them to communicate with each other and with infrastructure in real time to improve safety and efficiency. This is crucial for reducing accidents and enabling the development of intelligent transportation systems. Additionally, immersive experiences such as holographic communication, augmented reality (AR), and virtual reality (VR) will reach new heights in 6G, offering fully interactive, three-dimensional experiences that feel life-like and are limited only by imagination.6G will also leverage the terahertz (THz) spectrum to deliver massive bandwidth, enhancing data throughput and allowing for unprecedented connection densities. The THz spectrum enables faster data transfer over short distances, which makes it ideal for applications in dense urban areas and smart cities, where vast numbers of devices require simultaneous high-speed connectivity. As a result, 6G is set to support the continued expansion of the Internet of Things (IoT), connecting billions of smart devices in real time. This vast network will enable a range of new services and applications, from intelligent lighting and energy management in smart cities to predictive maintenance in industry. Moreover, 6G's ability to manage enormous device densities and data volumes will enhance critical services such as real-time telemedicine, remote diagnostics, and even robotic surgery. With its near-zero latency and ultra-high reliability (up to 99.99999%), 6G will allow medical professionals to perform complex procedures from afar, broadening access to quality healthcare. This is particularly transformative for rural and remote areas where medical expertise and resources are limited. Beyond speed and connectivity, sustainability will be a central feature of 6G networks. These next-generation networks are designed to optimize energy efficiency by employing advanced energy-harvesting techniques and improved network management. Through intelligent resource allocation, 6G networks will minimize energy consumption across infrastructure and connected devices. This focus on sustainability aligns with global efforts to reduce environmental impact and build greener technologies, ensuring that the growth of digital connectivity does not come at the expense of the planet. Security and privacy enhancements will also be critical components of 6G, as the vast scale of connected devices and the depth of data generated require robust protections. Advanced encryption, decentralized network architectures, and AI-driven security measures will form the backbone of a secure 6G environment, addressing vulnerabilities that come with increased connectivity. In summary, 6G represents more than just faster and broader connectivity. It's the foundation for a hyper-connected, intelligent, and sustainable digital future. By enabling revolutionary applications in healthcare, transportation, industry, and more, 6G is set to play a critical role in shaping how societies operate, delivering both convenience and innovation while prioritizing environmental responsibility. This visionary technology will transform communication into an immersive, highly reliable, and energy-conscious experience, bridging distances and making futuristic concepts part of everyday reality.

Figure 1: Features of 6G

Literature Review

The evolution of 6G networks increasingly depends on machine learning (ML) and artificial intelligence (AI) for managing complex infrastructures, achieving high efficiency, and supporting low-latency communication in IoT environments. Noman et al. [1] delve into the use of supervised learning to optimize resource allocation, enhancing 6G's reliability and responsiveness. Similarly, Mahmood et al. [2] focus on reinforcement learning applications that address the dynamic requirements of wireless networks, making them more adaptable to real-time demands. Tshakwanda et al. [3] propose a Speed-optimized LSTM (SP-LSTM) model, which leverages predictive analytics to enhance network performance through dynamic routing, underscoring the critical role of ML in network management. Moreover, Puspitasari et al. [4] highlight ML applications in emerging 6G technologies, such as intelligent reflecting surfaces (IRS), to meet next-generation network standards, while Muscinelli et al. [5] discuss how non-orthogonal multiple access (NOMA) technology benefits from deep learning (DL) techniques to improve spectral efficiency. The transformational potential of AI in 6G networks is further explored by Chataut et al. [6], who illustrate its role in enabling self-optimizing networks through technologies like terahertz communication and ultra-massive MIMO. Alibraheemi et al. [7] examine device-to-device (D2D) communication, providing insights into how AI facilitates seamless connectivity and resource management. Salama et al. [8] propose a deep reinforcement learning (DRL) algorithm that optimizes throughput in IoT-based 6G environments, emphasizing its effectiveness in handling high-density connections. Khan et al. [9] provide perspectives on AI-enabled transceivers, focusing on the Internet of Everything (IoE) and the ability to enhance 6G networks' adaptive and secure frameworks. Chowdhury et al. [12] explore ML's role in 6G's network slicing and ultra-reliable low-latency communication (URLLC), while Hurtado Sánchez et al. [11] examine DRL and network slicing as solutions for quality of service (QoS) in 6G. Their work emphasizes the importance of ML in maintaining service consistency in applications requiring varied network specifications. Brik et al. [14] present deep learning as a key element in optimizing O-RAN (Open Radio Access Networks), enhancing adaptability for dynamic network environments. Gismalla et al. [15] focus on D2D communication within ultra-dense networks, identifying ML techniques that maximize spectral efficiency and connectivity. Da Silva et al. [19] examine Q-learning's applications in NOMA systems, showing how this method improves resource allocation and user experience in high-demand settings. The incorporation of AI and ML across different network layers, such as Multi-access Edge Computing (MEC) and Network Functions Virtualization (NFV), is further examined by Ahmad et al. [10] and Salh et al. [13]. These studies illustrate ML's role in optimizing operations in environments with high computational and connectivity demands. Kaur et al. [16] review reinforcement learning and federated learning methods to enhance wireless network scalability, an essential feature for 6G's wide application range. Rekkas et al. [17] address distributed ML's impact on creating intelligent, energy-efficient, and secure communication systems. Alhashimi et al. [18] emphasize adaptive resource management in decentralized networks, outlining the benefits of ML in building resilient and scalable networks. Alsharif et al. [20] offer a forward-looking perspective on 6G, focusing on ML-driven intelligent systems for smart cities and autonomous systems, while Montejo-Sánchez et al. [19] discuss advanced AI algorithms, including those supporting quantum communication and intelligent transceivers. These comprehensive studies collectively underscore ML's transformative potential in 6G, setting the foundation for a future of globally connected, highly efficient networks. These studies affirm the importance of DRL, Q-learning, LSTM, and various ML frameworks in enabling the next generation of wireless networks, setting the stage for transformative changes in global communication.

Methodology

In advancing 6G networks, various machine learning methodologies have been employed to tackle complex problems in connectivity, resource management, and latency reduction. Q-learning and Deep Q-learning are widely used reinforcement learning techniques that facilitate optimal decision-making by learning the most effective actions through interaction with the environment. These methods are instrumental in dynamically adjusting network parameters to enhance performance and efficiency. Long Short-Term Memory (LSTM) networks, along with the Speed-optimized variant SP-LSTM, are effective in predicting network traffic patterns, enabling proactive adjustments in real-time communication scenarios. Multi-Agent Reinforcement Learning (MARL) extends these capabilities by using multiple agents that collaborate within a shared environment, a critical approach for handling complex, distributed systems and achieving self-organizing capabilities in ultra-dense networks. Together, these methodologies form a foundational toolkit for realizing robust and intelligent 6G systems.

1. Long Short Term Memory (LSTM):

Long Short-Term Memory (LSTM) networks are designed to handle sequential time-series data, such as network traffic patterns, by processing information over multiple time steps.

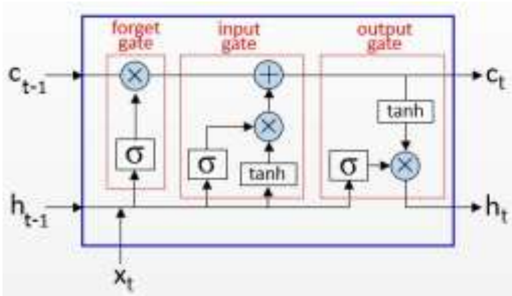

Figure 2. Block Diagram of LSTM

The workflow begins with the input sequence handling, where each input at a given time is combined with the hidden state from the previous step to provide context. The forget gate operation then decides what past information to retain or discard, allowing the network to forget irrelevant data and avoid outdated dependencies. Next, the input gate activation selects which new information from the current input should be added to the memory, ensuring that only useful data is stored while filtering out noise. The memory state update combines the retained past information with the relevant new input, maintaining a continuous memory of important features and long-term dependencies. The output gate control determines which parts of the memory state should influence the output at the current time step, generating the next hidden state that carries important information forward.

$$\begin{aligned}
\text{Forget gate:}\quad & f_t = \sigma(W_f \cdot [h_{t-1}, x_t] + b_f) \\
\text{Input gate:}\quad & i_t = \sigma(W_i \cdot [h_{t-1}, x_t] + b_i) \\
& \hat{C}_t = \tanh(W_C \cdot [h_{t-1}, x_t] + b_C) \\
\text{Cell state update:}\quad & C_t = f_t * C_{t-1} + i_t * \hat{C}_t \\
\text{Output gate:}\quad & o_t = \sigma(W_o \cdot [h_{t-1}, x_t] + b_o) \\
& h_t = o_t * \tanh(C_t)
\end{aligned}$$

Here, **σ** is the sigmoid function that determines how much information should be gated, while **tanh** is used to scale the candidate cell state. The ***** operator indicates element-wise multiplication, and **[.,.]** represents concatenation of the previous hidden state (**ht-1**) and the current input (**xt**). The weight matrices (**Wf**, **Wi**, **WC**, **Wo**) transform the concatenated input, and the bias vectors (**bf**, **bi**, **bC**, **bo**) are added before applying activation functions.The final output is generated from the hidden state of the last time step, which contains the learned patterns and dependencies. This output is used to predict network conditions, such as traffic load or congestion. Prediction and feedback occur as the LSTM outputs predictions for future network behavior, enabling proactive actions like rerouting data. The continuous feedback loop improves the model's accuracy over time, enhancing overall network performance, especially for complex tasks in 6G networks.

2. Speed Optimized Long Short-Term Memory (SP-LSTM):

SP-LSTM (Simplified Processing Long Short-Term Memory) improves upon traditional LSTM models by streamlining the architecture for better efficiency and performance. Unlike LSTM, which uses separate gates for input, forget, and output operations, SP-LSTM merges the input and forget gates into a single unified update gate. This simplified approach reduces the computational complexity, making the model more efficient. In terms of processing speed, SP-LSTM integrates attention mechanisms that allow for parallel processing of time steps, significantly speeding up computations compared to the sequential nature of LSTM. This enhancement is particularly beneficial for tasks that require faster processing of large volumes of data.In addition to speed improvements, SP-LSTM introduces advanced memory management capabilities. It incorporates an explicit memory mechanism similar to Neural Turing Machines, which enables the model to store and recall important patterns over long sequences. This allows SP-LSTM to better handle long-term dependencies, addressing a key limitation of traditional LSTM, which sometimes struggles with remembering information over extended time periods.

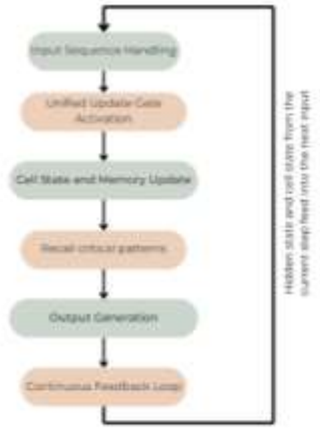

Figure 3. Flowchart of SP-LSTM

The workflow in SP-LSTM begins with the unified update gate, which handles both the input and forget operations simultaneously, reducing the model's complexity. This gate efficiently updates the cell state by retaining only the most relevant information from the input, ensuring that important temporal dependencies are maintained while discarding unnecessary data. This improvement in memory handling enables SP-LSTM to process long sequences more effectively than LSTM, without the complexity of multiple gates. The model also benefits from its attention mechanisms, which allow the network to focus on the most important time steps, further improving the model's ability to handle complex tasks. In summary, SP-LSTM outperforms traditional LSTM by offering enhanced efficiency, faster processing, and better memory management, making it a suitable choice for applications involving long sequences and large-scale data processing.

3. Q-Learning:

Q-learning is a model-free reinforcement learning algorithm used to find the optimal action-selection policy for an agent interacting with an environment. It operates by learning from the consequences of its actions to maximize cumulative reward over time. The algorithm uses a Q-table to store Q-values for each state-action pair, which represent the expected future reward for taking a specific action in a given state. The state space is typically small and discrete, making it suitable for environments where all states can be explicitly defined. However, Q-learning is most effective in environments with limited state spaces, as its reliance on a Q-table becomes inefficient for large, continuous state spaces.

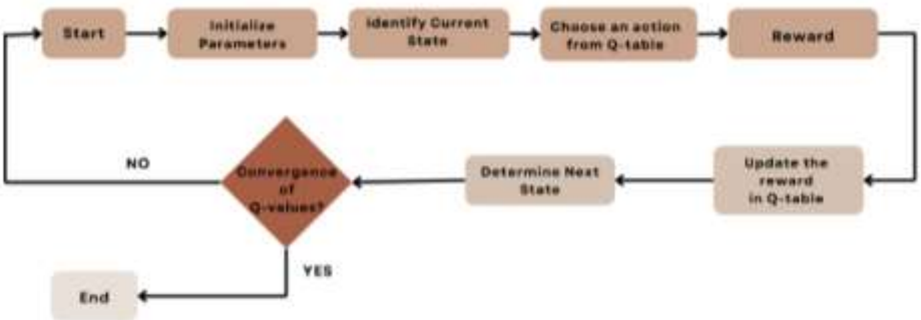

Figure 4. Flowchart of Q-Learning

The workflow of Q-learning begins by initializing key parameters, including the Q-table, the learning rate (α), the discount factor (γ), and the exploration-exploitation trade-off (ε). The agent then starts interacting with the environment. The process begins by identifying the current state, after which an action is selected from the Q-table using an exploration-exploitation strategy. Once the agent takes the action, the environment generates a reward. This reward is then used to update the Q-table using the Bellman equation:

$$Q(s,a) = r(s,a) + \gamma \max_{a} Q(s',a)$$

In the above equation, Q(s, a) represents the expected reward for taking action ‘a’ in state ‘s’. r(s, a) is the actual reward received after performing action ‘a’ in state ‘s’. s′ denotes the next state after taking action ‘a’. α is the learning rate, controlling how much new information overrides the old, while γ is the discount factor, determining the importance of future rewards in the learning process.Next, the agent determines the next state based on the environment's response. The loop continues until a stopping condition is met. These conditions can include when the Q-values converge within a threshold (ε), when the maximum number of iterations or episodes is reached, or when the desired network performance, such as low latency or high energy efficiency, is achieved. Despite its effectiveness, Q-learning has some limitations. For instance, it updates Q-values immediately after each action, which can cause fluctuations and instability, particularly in environments with high variance. Additionally, Q-learning struggles with large state spaces, as the

Q-table becomes increasingly difficult to manage, making it less practical for environments with continuous or high-dimensional state spaces. Lastly, its performance can degrade if the exploration-exploitation balance is not properly tuned, resulting in suboptimal learning.

4. Deep Q-Learning:

Deep Q-Learning (DQN) is an advanced reinforcement learning algorithm that combines Q-learning with deep neural networks, enabling it to handle complex, high-dimensional state spaces. In the context of a 6G network, the input layer includes key network components such as bandwidth, user mobility, and energy consumption, which directly influence network performance. These parameters, along with use cases like autonomous systems and holographic communication, form the foundation for optimizing the network. The dataset collection process involves gathering real-time data on network parameters and user behavior, using sensors and monitoring tools to ensure accuracy. This data is then cleaned and normalized to create a suitable dataset for training machine learning models.

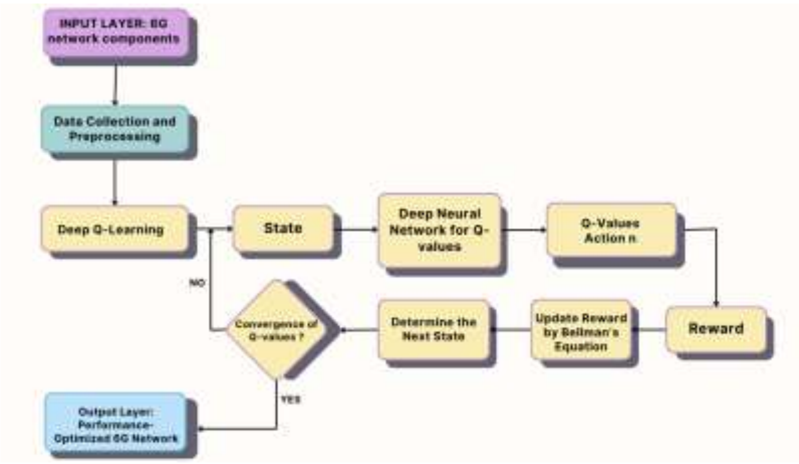

Figure 5. Flowchart of Deep Q-Learning

The workflow of Deep Q-Learning in this context begins with defining the state, which represents the current network configuration, such as bandwidth usage and user distribution. The agent then selects an action, which involves adjusting parameters like resource allocation and network configurations. The Q-network, a deep neural network, approximates the Q-values for state-action pairs, outputting the expected reward for each possible action. To enhance learning stability, experience replay is employed, storing past experiences (state, action, reward, next state) in a replay buffer. Mini-batches are randomly sampled from this buffer to break correlations in the training data. A target network is also used to stabilize learning, where its weights are periodically updated with the main Q-network's weights. The reward signal, such as lower latency or energy savings, provides feedback to the model, which is used to update the Q-values using the Bellman equation. The feedback loop continues until a stopping condition is met: Q-values converge within a threshold, the maximum number of iterations or episodes is reached, or the desired network performance, such as low latency or high energy efficiency, is achieved. Finally, the output layer focuses on performance-optimized 6G networks. Intelligent network management through automated adjustments leads to reduced latency and higher reliability. Energy-efficient resource allocation ensures optimal use of network resources, while seamless user experiences are achieved by enhancing performance for next-gen applications. This approach showcases how Deep Q-Learning can dynamically optimize the 6G network by addressing the challenges of traditional Q-learning, with mechanisms like experience replay and target networks providing stable and effective training.

5. Multi-agent Reinforcement Learning(MARL):

Multi-Agent Reinforcement Learning (MARL) is a subfield of reinforcement learning where multiple agents interact within an environment to achieve their respective goals, which may be cooperative, competitive, or a combination of both.

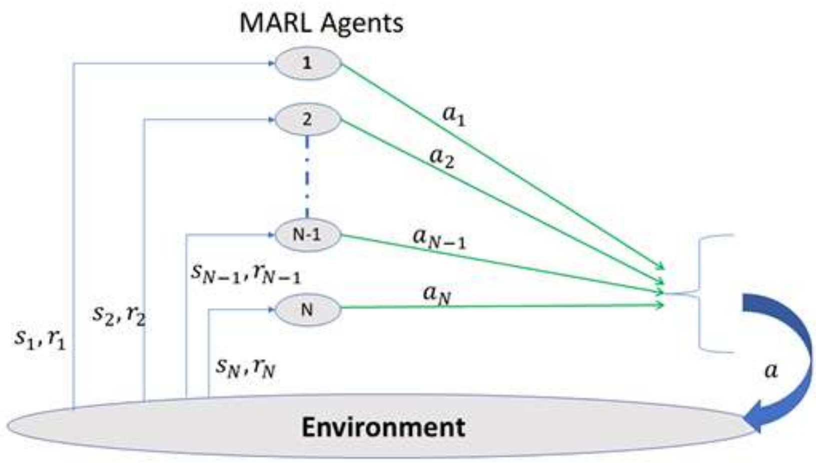

Figure 6. Multi-agent reinforcement learning framework

In the context of 6G networks, each agent observes a portion of the network state, such as bandwidth usage or user traffic, based on its local environment. These observations form the basis of their decision-making process. Once an agent has gathered the necessary information, it selects an action, such as resource allocation or routing decisions, aimed at optimizing its individual objective. However, the collective actions of all agents influence the network's next state, demonstrating the interdependence of agents in the system. The interaction between agents can take different forms: cooperative agents work towards a common goal, competitive agents aim to maximize their own rewards, and hybrid agents combine both strategies to achieve optimal performance. After each action is taken, the environment transitions to a new state, and the agents receive feedback in the form of rewards, which could reflect reduced latency, energy savings, or improved throughput. This reward feedback is crucial for guiding the agents' learning processes. In the learning phase, agents refine their strategies by updating Q-values or policies based on the received rewards. In complex scenarios, neural networks are often employed to approximate these values, enabling the agents to handle high-dimensional state spaces. During real-time operation, agents continuously adapt to dynamic network conditions using the policies they've learned, optimizing the performance of the 6G network. Through this process, MARL enables decentralized control, where multiple agents independently make decisions that collectively improve network efficiency and performance.

Results and Discussion

The performance of various algorithms was assessed, including LSTM, SP-LSTM, and Q-Learning, Deep Q-Learning, and Multi-agent RL. LSTM and SP-LSTM demonstrated high training and prediction accuracies, with LSTM achieving 0.9822 and 0.99999, respectively, and SP-LSTM showing 0.984 and 0.99870, alongside very low MSE values. In contrast, reinforcement learning models showed good accuracy ranges, with Q-Learning reaching a peak of 88% and a low of 65%, Deep Q-Learning achieving 92% and 84.2%, and Multi-agent RL attaining 90% and 65%.

S.NO	ALGORITHMS	ACCURACY	MEAN SQUARE ERROR(MSE)
1.	Long Short Term Memory (LSTM)	Training Accuracy=0.9822, Prediction accuracy=0.99999	4.95e-07
2.	Speed Optimized Long Short-Term Memory (SP-LSTM)	Training Accuracy=0.984, Prediction accuracy=0.99870	4.09e-05
3	Q-Learning	Highest Accuracy= 88% Lowest Accuracy=65%	-
4	Deep Q-Learning	Highest Accuracy=92% Lowest Accuracy= 84.2%	-
5	Multi-agent reinforcement learning	Highest Accuracy=90% Lowest Accuracy= 65%	-

TABLE 1. Comparison of various algorithms

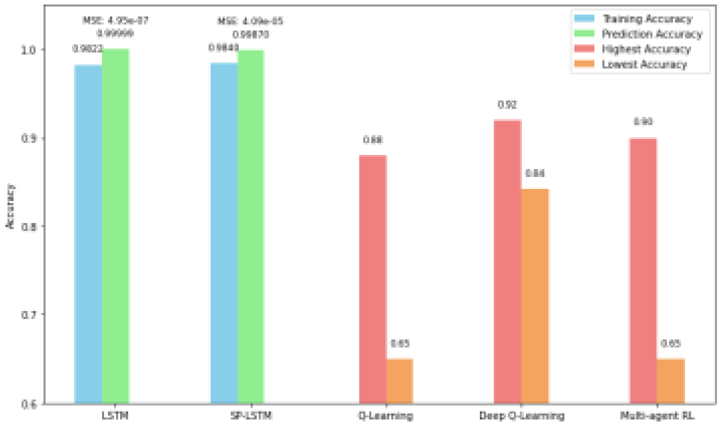

Figure 7. Graphical representation of performance metrics of various algorithms

Conclusion

In conclusion, machine learning (ML) plays a pivotal role in the development of 6G networks, addressing key challenges related to resource management, energy efficiency, and latency reduction. The integration of advanced techniques such as Q-learning, Deep Q-learning, LSTM, SP-LSTM, and Multi-Agent Reinforcement Learning (MARL) is essential for optimizing network performance, enabling dynamic adaptability, and fostering intelligent decision-making in ultra-dense network environments. However, challenges remain in ensuring the scalability of ML models, maintaining network security, and effectively integrating AI-driven solutions within existing infrastructure. Looking forward, the future scope of ML in 6G will focus on enhancing model efficiency, improving real-time adaptability, and advancing cross-domain collaboration for complex systems. Additionally, future research must address issues such as reducing computational complexity, developing secure AI frameworks, and enabling seamless interoperability between emerging technologies. As 6G networks evolve, ML will be instrumental in shaping next-generation wireless communication systems that support new applications like autonomous vehicles, smart cities, and immersive holographic communication.

References

[1] Noman, H. M. F., Hanafi, E., Noordin, K. A., Dimyati, K., Hindia, M. N., Abdrabou, A., & Qamar, F. (2023). Machine Learning Empowered Emerging Wireless Networks in 6G: Recent Advancements, Challenges & Future Trends. IEEE Access.

[2] Mahmood, M. R., Matin, M. A., Sarigiannidis, P., & Goudos, S. K. (2022). A comprehensive review on artificial intelligence/machine learning algorithms for empowering the future IoT toward 6G era. *IEEE Access*, *10*, 87535-87562.

[3] Tshakwanda, P. M., Arzo, S. T., & Devetsikiotis, M. (2024). Advancing 6G Network Performance: AI/ML Framework for Proactive Management and Dynamic Optimal Routing. *IEEE Open Journal of the Computer Society*.

[4] Puspitasari, A. A., An, T. T., Alsharif, M. H., & Lee, B. M. (2023). Emerging technologies for 6G communication networks: Machine learning approaches. *Sensors*, *23*(18), 7709.

[5] Muscinelli, E., Shinde, S. S., & Tarchi, D. (2022). Overview of distributed machine learning techniques for 6G networks. *Algorithms*, *15*(6), 210.

[6] Chataut, R., Nankya, M., & Akl, R. (2024). 6G networks and the AI revolution—Exploring technologies, applications, and emerging challenges. *Sensors*, *24*(6), 1888.

[7] Alibraheemi, A. M. H., Hindia, M. N., Dimyati, K., Izam, T. F. T. M. N., Yahaya, J., Qamar, F., & Abdullah, Z. H. (2023). A survey of resource management in D2D communication for B5G networks. *IEEE Access*, *11*, 7892-7923.

[8] Salama, G. M., Metwly, S. S., Shehata, E. G., & Abd El-Haleem, A. M. (2023). Deep reinforcement learning based algorithm for symbiotic radio iot throughput optimization in 6g network. IEEE Access, 11, 42331-42342.

[9] Khan, L. U., Yaqoob, I., Imran, M., Han, Z., & Hong, C. S. (2020). 6G wireless systems: A vision, architectural elements, and future directions. IEEE access, 8, 147029-147044.

[10] Ahmad, I., Shahabuddin, S., Malik, H., Harjula, E., Leppänen, T., Loven, L., ... & Riekki, J. (2020). Machine learning meets communication networks: Current trends and future challenges. IEEE Access, 8, 223418-223460.

[11] Hurtado Sánchez, J. A., Casilimas, K., & Caicedo Rendon, O. M. (2022). Deep reinforcement learning for resource management on network slicing: A survey. Sensors, 22(8), 3031.

[12] Chowdhury, M. Z., Shahjalal, M., Ahmed, S., & Jang, Y. M. (2020). 6G wireless communication systems: Applications, requirements, technologies, challenges, and research directions. IEEE Open Journal of the Communications Society, 1, 957-975.

[13] Salh, A., Audah, L., Shah, N. S. M., Alhammadi, A., Abdullah, Q., Kim, Y. H., ... & Almohammedi, A. A. (2021). A survey on deep learning for ultra-reliable and low-latency communications challenges on 6G wireless systems. IEEE Access, 9, 55098-55131.

[14] Brik, B., Boutiba, K., & Ksentini, A. (2022). Deep learning for B5G open radio access network: Evolution, survey, case studies, and challenges. IEEE Open Journal of the Communications Society, 3, 228-250.

[15] Gismalla, M. S. M., Azmi, A. I., Salim, M. R. B., Abdullah, M. F. L., Iqbal, F., Mabrouk, W. A., ... & Supa'at, A. S. M. (2022). Survey on device to device (D2D) communication for 5GB/6G networks: Concept, applications, challenges, and future directions. IEEE Access, 10, 30792-30821.

[16] Kaur, J., Khan, M. A., Iftikhar, M., Imran, M., & Haq, Q. E. U. (2021). Machine learning techniques for 5G and beyond. *IEEE Access*, *9*, 23472-23488.

[17] Rekkas, V. P., Sotiroudis, S., Sarigiannidis, P., Wan, S., Karagiannidis, G. K., & Goudos, S. K. (2021). Machine learning in beyond 5G/6G networks—State-of-the-art and future trends. *Electronics*, *10*(22), 2786.

[18] Alhashimi, H. F., Hindia, M. N., Dimyati, K., Hanafi, E. B., Safie, N., Qamar, F., ... & Nguyen, Q. N. (2023). A survey on resource management for 6G heterogeneous networks: current research, future trends, and challenges. *Electronics*, *12*(3), 647.

[19] da Silva, M. V., Montejo-Sánchez, S., Souza, R. D., Alves, H., & Abrão, T. (2022). D2d assisted q-learning random access for noma-based mtc networks. *IEEE Access*, *10*, 30694-30706.

[20] Alsharif, M. H., Kelechi, A. H., Albreem, M. A., Chaudhry, S. A., Zia, M. S., & Kim, S. (2020). Sixth generation (6G) wireless networks: Vision, research activities, challenges and potential solutions. Symmetry, 12 (4), 676.