SCIENTIFIC CULTURE
www.sci-cult.com

SCIENTIFIC CULTURE, Vol. 12, No. 4, (2026), pp. 10664-10684
Open Access. Online & Print

DOI: 10.5281/zenodo.124261028

AGENTIC AI-DRIVEN HYBRID CNN–DEEP REINFORCEMENT LEARNING FRAMEWORK FOR AUTONOMOUS MULTI-OBJECTIVE RESOURCE ALLOCATION IN 5G/6G WIRELESS NETWORKS

[1]Dr. Nitin S. Shrirao, [2]Snehankita Majalekar, [3]Anuradha Kanade, [4]Dr.Laxman Doiphode, [5]Dr.Vaishali Dhanaji Nikam, [6]Dr. Rajesh Kumar Kashyap, [7*]Dr.Prashant Bhavarlal Chordiya

[1]Director Siddhant Institute of Computer Application, Sudumbare, Pune
[2]Assistant Professor Pimpri Chinchwad University, Pune, India
[3]Assistant Professor, Dr. Vishwanath Karad MIT World Peace University,Pune, India
[4]Associate Professor, MIT World Peace University, School of Business Kothrud, Pune, Maharashtra-411038
[5]Associate Professor, Shree Shivaji Maratha Society's Institute of Management and Research Pune, India
[6]ProfessorG H Raisoni College of Engineering and Management, Wagholi, Pune, India
[7]Associate Professor D.Y. Patil Institute of Master of Computer Applications and Management, Pune, India

Received: 05/12/2025
Accepted: 30/01/2026

Corresponding Author: Nitin S. Shrirao (nshrirao@gmail.com)

ABSTRACT

The rapid development of 5G and new 6G wireless networks demands intelligent resource allocation to meet various quality of service (QoS) requirements including high throughput and low latency and effective spectrum usage. The existing static methods and heuristic-based approaches cannot function properly in changing network environments which leads to performance constraints. The research introduces an Agentic Artificial Intelligence (AI) framework that uses Convolutional Neural Networks (CNN) and Deep Reinforcement Learning (DRL) to enable systems to make automatic decisions. The proposed method uses CNN to extract features while DRL develops adaptive policies in a Markov Decision Process (MDP)-based system. The hybrid CNN-DRL system uses a reward-based approach to achieve its various performance targets. The hybrid model shows superior performance throughputs of approximately 0.10 which leads to a decrease in latency by approximately 0.26 while achieving an efficiency increase of approximately 0.085 and a fairness improvement of approximately 0.23. The CNN model achieves 91% accuracy, while the DRL system shows consistent reward progression within the 2990 to 3035 range. The system delivers a highly effective and expandable system for upcoming wireless network technologies.

Keywords: Agentic Artificial Intelligence, 5G/6G Networks, Resource Allocation, CNN, Deep Reinforcement Learning, Hybrid Model, Spectrum Efficiency

1. INTRODUCTION

The rapid advancements in wireless communication technologies from fifth-generation wireless networks (5G) to upcoming sixth-generation wireless networks (6G) are significantly changing the digital connectivity scenario by providing ultra-reliable low-latency communications (URLLC), enhanced mobile broadband (eMBB), massive machine-type communications (mMTC), etc. These diverse service requirements impose significant constraints on various network performance parameters such as throughput, latency, reliability, and spectrum efficiency [1]. Conventional resource allocation approaches, such as static optimization approaches, rule-based scheduling approaches, and heuristic approaches, are found to be inadequate in managing the complexity of next-generation wireless networks. Therefore, there is an increased need to develop intelligent approaches that can provide autonomous decision-making in highly dynamic wireless networks [2].

Machine learning (ML) and Deep learning (DL) have become a potential paradigm of improving resource management in wireless networks due to Artificial Intelligence (AI) [3]. The current AI-based solutions have shown considerable enhancement of channel estimation, traffic prediction and interference mitigation and power allocation. Nevertheless, the majority of these methods are based on centralized learning structures and pre-trained models which have no contextual awareness or autonomous reasoning to operate in real-time adaptive optimal as in ultra-dense and highly heterogeneous 5G/6G environments. Moreover, these systems are mostly run as passive optimization mechanisms but not proactive decision-making agents with lawful goal oriented actions [4].

The concept of "Agentic Artificial Intelligence (Agentic AI)" brings about a revolution in the understanding of traditional predictive intelligent compared to the autonomous and self-directed intelligence decision-making systems [5]. Contrary to classical models of AI with pre-defined set of inputs, Agentic models of AI have the ability to sense the environmental states, design goals, test between different strategies and perform self-adaptive behaviour with a limited number of human interventions. This feature shows Agentic AI to be especially applicable to next-generation wireless networks, in which resource allocation is a continuous process based on changing traffic loads, mobility, variations in interference and service quality demands [6,7].

Resource allocation under a 5G and beyond system implies the efficient sharing of spectrum, transmission power, bandwidth, and computation resources among multiple users and services that are subject to a variety of quality-of-service (QoS) demands [8]. It is even more urgent in 6G scenarios, in which there are integrated non-terrestrial-terrestrial networks, "Intelligent Reflecting Surfaces (IRS)", terahertz network technology, and architectures based on AI. These new paradigms require decentralized, context-sensitive and self-organizing control models that can autonomously manage network resources to ensure optimal performance among a number of conflicting goals [9].

Emerging research in multi-agent systems, reinforcement learning, and edge intelligence offers an excellent basis in creating agent-based systems of wireless resource optimization. Within intelligent agents instantiated at the base stations, edge nodes and network controllers, distributed coordination, collaborative learning and dynamic spectrum usage between network layers is possible [10]. Scheduling and supporting users in real-time, making efficient bandwidth utilization, proactively reducing interference, and fulfilling latency-sensitive transmissions possible are some of the functions that such agentic frameworks can dynamically prioritize users, allocate bandwidth efficiently, prevent interference proactively and optimize latency-sensitive transmissions. Consequently, they add considerable network throughput and at the same time promote fairness and reliability among heterogeneous service classes [11].

In spite of these encouraging changes, there are several research issues that are not addressed [12]. The current policy of resource allocation is not usually able to cooperatively perform optimization of throughput, latency, and spectral efficiency with a single policy framework. Furthermore, there has been a lack of focus on the combination of agentic decision-making schemes with cross-layer optimization schemes that concurrently recognize physical-layer channel state, network-layer routing state, and application-layer service needs. The solution to these limitations would be to have smart architectures that have the capabilities of autonomous perception, reasoning and adaptation on several dimensions of the network [13].

To address these resources, this paper presents an autonomous framework of resource allocation with the support of the Agentic Artificial Intelligence developed to achieve optimal values of key performance dimensions such as throughput, latency, and spectrum efficiency in the next generation wireless networks. The offered structure utilizes distributed intelligent agents with the ability to make environment-conscious decisions, spectrum allocation, and prioritization of traffic [13]. The framework can be used to proactively and scalably manage resources by combining reinforcement learning concepts of optimization with in-the-field network state observability to support ultra-dense 5G deployments and the future 6G environments [14].

The main value of this work is that it is the first to develop a common type of agentic architecture in facilitating decentralized coordination, adaptive policy learning, and multi-objective optimization in heterogeneous wireless situations. Such a strategy does not only improve the network performance within different traffic conditions but also, creates a scalable base of intelligent self-organizing 6G communication systems. Although wireless networks are actively developing towards AI-native architectures, agentic intelligence will play a key role in facilitating fully autonomous, robust, and energy-saving infrastructures of communication [15]. The research objectives of this study are as follow:

- To develop an Agentic AI framework for autonomous resource allocation in 5G/6G networks using an MDP-based environment.
- To preprocess the 5G dataset using cleaning, normalization, scaling, and encoding techniques
- To implement a hybrid CNN–DRL model for feature extraction and intelligent decision-making.
- To optimize network performance using a multi-objective reward function with policy learning and hyperparameter tuning.

The rest of this paper is formatted as follows. Section II includes a literature review of the related work on AI-based resources allocation in the 5G/6G network. Section III is the details of the suggested Agentic AI-based structure, system state model, optimization plan, and simulation configuration to analyze throughput, latency, and spectrum efficiency. Section IV shows results of performance and comparative analysis. Finally, Section V provides the conclusion of the paper and speculates on future study directions on autonomous wireless network optimization.

2. LITERATURE REVIEW

The latest development of 5G/6G types is aimed at solving the issues of scalability, latency, security, and sustainability by intelligent and autonomous means. Untenable SLA compliance in 6G slicing by validating 6G slicing with an Agentic intelligence and AI fusion framework was introduced by Vashisht et al., (2026) [16] and had 99.9% accuracy with a conclusion of enhanced stability and reliable autonomous 6G slicing. The Aasa et al., (2026) [17] tackled the complexity of SAGIN orchestration with self-supervised learning by requiring at least 40 times less data but still achieving nearly optimal performance and concluded its efficiency in dynamic setting. Another study, (2026) [18], addressed data overload in IoV with the help of an edge-fog computing based on DRL, and it received the results of reduction of latency by 3070 percent and energy consumption by 2050 percent, concluding improved QoS. A study focused on the security of ITS with the help of blockchain-based agentic AI, on the one hand, enhanced the accuracy of the detection process and provided secure communication, on the other hand, (Thenmozhi et al., (2026) [19]) stated. The paper by Mohammed et al., (2025) [20] addresses each scalability problem with the AI-based optimization, making it better in the throughput and decreasing latency and concludes that AI is a necessity in autonomous networks. Wu et al., (2025) [21] also discussed the case of lack of cognitive autonomy application and utilized agent-based architecture, attaining sub-10 ms latency, 4-percent throughput improvement, and 85-percent BLER reduction, which justifies L4 networks. Enhanced mission-critical systems with agentic AI by Khowaja et al., (2025) [22] shortened response time with agentic AI by 5.6 minutes and chose resources with the agentic AI by 13.4 percent, and concluded adaptability was improved.

Table 1: Comparative Summary of AI and Agentic Resource Optimization Techniques in 5G/6G Wireless Networks

S. No.	Author (Year)	Methodology	Key Contributions	Limitations
1	Vashisht et al. (2026) [16]	Agentic ML fusion using LightGBM, RF, and LR ensemble	Achieved 99.9% accuracy with SLA-aware autonomous network slicing under dynamic 6G traffic conditions	Limited evaluation for joint spectrum–latency–throughput optimization
2	Aasa et al. (2026) [17]	Self-supervised learning for RIS-assisted SAGIN orchestration	Near-optimal sum-rate performance with 40% reduced training data requirements	Does not support distributed multi-agent decision intelligence

3	Siddique et al. (2026) [18]	DRL-based edge–fog computing framework	Reduced latency by 30–70% and improved QoS reliability in IoV environments	Spectrum efficiency and cross-layer coordination not addressed
4	Thenmozhi et al. (2026) [19]	Agentic AI-enabled blockchain security with AANT-OMPEN	Improved detection accuracy and secure RSU communication in ITS systems	Focused primarily on security rather than resource allocation optimization
5	Mohammed et al. (2025) [20]	ML/DL-based adaptive wireless network optimization	Enhanced throughput, latency performance, and spectral utilization	Lacks autonomous agent-driven real-time optimization capability
6	Wu et al. (2025) [21]	Cognitive autonomous RAN agent architecture	Achieved sub-10 ms control latency with 4% throughput gain and 85% BLER reduction	Limited scalability for heterogeneous multi-layer 6G scenarios
7	Khowaja et al. (2025) [22]	Multi-layer Agentic AI framework for mission-critical systems	Reduced response time and improved concurrent operational efficiency	Not validated for spectrum-aware wireless resource allocation
8	Lazrek et al. (2025) [23]	CNN-based AI-DRX optimization mechanism	Achieved 69.2% energy savings with low inference latency for real-time deployment	Does not jointly optimize latency and throughput performance
9	Othman et al. (2025) [24]	UAV, THz, and IRS-enabled 6G architecture	Improved signal propagation and spectral efficiency using programmable metasurfaces	Missing intelligent autonomous orchestration mechanisms
10	Kakarlapudi et al. (2024) [25]	Agentic AI-based telecom automation framework	Enabled self-configuring, self-healing, and self-optimizing network operations	Quantitative performance validation for resource allocation not provided
11	Javaid et al. (2024) [26]	LLM-assisted ISATN optimization framework	Improved predictive routing and intelligent satellite–air–ground coordination	Latency-sensitive real-time deployment challenges remain
12	Bagwari et al. (2024) [27]	Intelligent computational AI-enabled 5G automation model	Achieved 96.31% network speed management and improved energy efficiency	Spectrum-aware adaptive scheduling not considered
13	Chauhan et al. (2024) [28]	Lifecycle sustainability evaluation framework for 6G	Proposed strategies for reducing carbon emissions and improving deployment sustainability	Does not address network-level performance optimization
14	Tomaszewski et al. (2023) [29]	Use-case driven sustainability framework for 5G/6G ecosystems	Identified sector-specific deployment challenges across smart infrastructure domains	Lacks AI-driven adaptive resource allocation mechanisms
15	Rao et al. (2023) [30]	Blockchain-enabled IoT-V2X communication architecture	Improved transparency and secure vehicular data exchange reliability	Does not integrate intelligent spectrum optimization strategies

The research of lazrek et al., (2025) [23] was about energy inefficiency and investigated by using AI-DRX, with energy savings of 69.2% and accuracy of 91.8, which confirms the feasibility of real-time. The article by 6G experts at Othman et al., (2025) [24] ended up solving problems of the 6G architecture by using the UAV, THz, and IRS approaches, which had the potential to support the multi-gigabit data rates and allow for enhanced connectivity. The paper of Kakarlapudi et al., (2024) [25] dealt with the issue of network complexity, which was manifested in achieving greater reliability and less human interaction by means of an agentic AI. The performance of ISATN was improved by Javaid et al., (2024) [26] with the integration of the LLM which better predicted the decision-making process. Intelligent computational model solved 5G inefficiencies achieving performance efficiency of 96.31% in Bagwari et al., (2024) [27] concluding enhanced network and energy management. Environmental impact was covered by Chauhan et al., (2024) [28] with the mentioned (200 g) CO 2 emissions and (2.5 kWh/GB) energy consumption, and sustainability strategies were proposed. The article by Tomaszewski et al., (2023) [29] researched the concept of sustainability and the main areas of its application in the sector. Lastly, V2X security with blockchain was considered by Rao et al., (2023) [30] to guarantee security and reliability of communication in smart transportation.

Existing AI-based resource allocation methods in 5G/6G networks mainly address individual performance metrics such as latency reduction, energy efficiency, or security rather than joint optimization of throughput, latency, and spectrum efficiency. Moreover, current approaches lack distributed agentic decision-making, cross-layer coordination, and real-time adaptive resource management in dynamic network environments. Therefore, a unified Agentic AI–driven autonomous framework is required to enable scalable, intelligent, and multi-objective optimization for next-generation wireless communication systems.

3. RESEARCH METHODOLOGY

The suggested approach is depicted in figure 1 and depicts the framework of Agentic Artificial Intelligence-inspired autonomous distribution of resources in 5G/6G wireless networks that unites data-driven learning and adaptive decision-making. To begin with, the 5G resource allocation data is preprocessed by data cleaning, data normalization process, feature scaling process, feature encoding, and feature engineering process to achieve high quality input of model training. The data is then processed into training (80) and testing (20) sets. It models network environment as a dynamic Markov Decision Process, an intelligent agent views network states and acts to maximize the resource allocation. It uses a hybrid learning process that involves the utilization of Convolutional Neural Networks (CNN) to extract features and Deep Reinforcement Learning (DRL) to make sequential decisions to identify the patterns of complex networks and provides the ability to adapt to real-time. A reward function which is crafted to balance between throughput, latency, and spectrum efficiency is used to guide the agent. Optimization methods like experience replay and policy updates are applied in training the model to converge. Lastly, the proposed model is tested in terms of important measures such as throughput, latency, spectrum efficiency, and fairness, which prove its usefulness compared to the traditional resource allocation strategies.

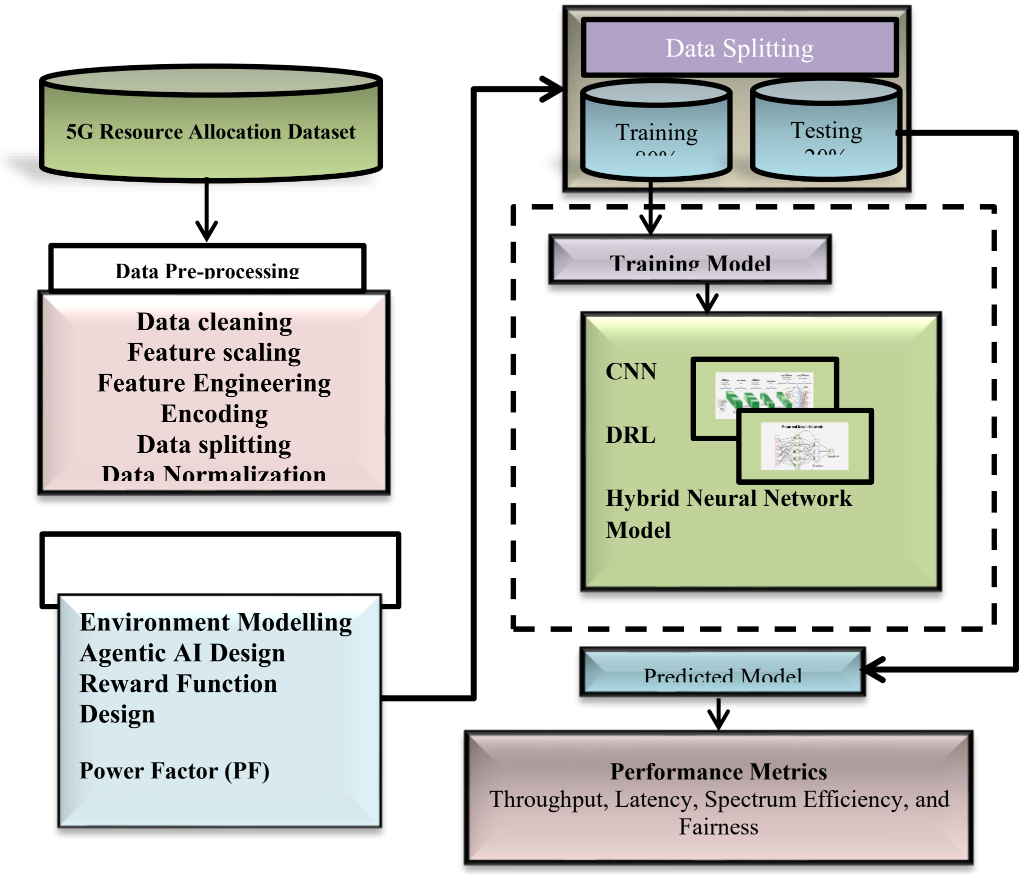


Figure 1: Framework of Proposed Methodology

3.1 DATASET USED

The 5G Resource Allocation Dataset [31], which is offered on Kaggle, is a structured dataset that is supposed to assist research and development on intelligent wireless network management, especially on 5G systems. It has numerous parameters which are related to the network which include signal strength, demand of the user, latency requirements, channel conditions and bandwidth allocated. All these features are realistic learned network scenarios involving the competition of limited radio resources by multiple users. It is particularly valuable as it can be used to model and analyze the distribution of resources, in which case it can be dynamically allocated to guarantee the best Quality of Service (QoS). It allows the researchers to examine the correlation between network conditions and allocation strategies, which makes it very appropriate in terms of machine learning and artificial intelligence. This information can be applied in agentic AI to train autonomous agents to optimize policies to allocate resources, enhance throughput, reduce latency, and increase spectrum efficiency. It also facilitates experimentation of techniques used in optimization including reinforcement learning, deep learning, and heuristic algorithms. The dataset also mimics real-life network operation, which makes it highly useful in creating self-adaptable and scalable solutions to future 5G and 6G wireless communication networks.

Table 1: Distribution of the dataset

Class / Application Type	Total Samples	Train	Test
Video Streaming	1200	960	240
Online Gaming	1000	800	200
Web Browsing	900	720	180
VoIP / Calls	800	640	160
File Download	1100	880	220
IoT Communication	700	560	140
AR/VR Applications	600	480	120
Emergency Services	500	400	100
Background Traffic	700	560	140
TOTAL	**7500**	**6000**	**1500**

3.2 DATA PRE-PROCESSING

Data Preprocessing is a pre-processing necessary step of converting raw data to machine learning and artificial intelligence models. It includes data cleaning, which includes dealing with missing or noisy data points and converting the features into a common format and normalizing the numerical data to maintain uniformity. Preprocessing can be used to standardize parameters in 5G resources allocation, including latency, bandwidth, signal strength, among others; therefore, models can learn effectively [32]. It can also involve a process of encoding categorical variables and the development of new features that are more representative of network behavior. Effective preprocessing enhances model accuracy, stability and efficiency which can be used to optimize the throughput, latency and use of spectrum reliably. The dataset 5G Resource Allocation Dataset applied in the current research is pre-processed by way of several steps as follows.

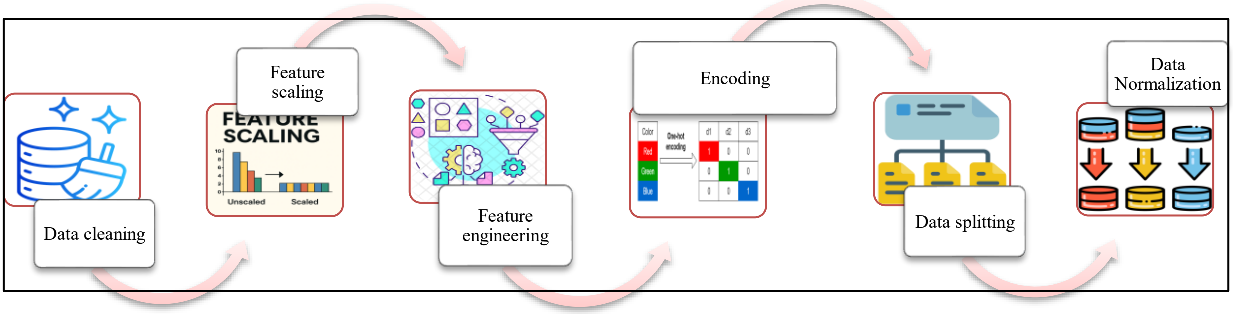


Figure 2: Dataset Preprocessing Phase

- **Data Cleaning and Missing Value Handling**

The initial one is to detect and to treat missing, inconsistent, or noisy data in data set. The latency, signal strength, or bandwidth parameters present in 5G datasets can have null or incorrect values because of variations in the network [33]. They are normally done with statistical imputation procedures like mean replacement or median replacement in order to maintain the integrity of the datasets. Elimination of duplicates and fixing anomalies makes model training reliable.

$$X_{clean} = \begin{cases} X_i, & if\ X_i \neq NaN \\ \frac{1}{n}\sum_{i=1}^{n} X_i, & if\ X_i = NaN \end{cases} \quad (1)$$

- **Feature Scaling (Normalization / Standardization)**

Because network parameters such as throughput, latency, and signal strength are measured using different scales and have different units, feature scaling is necessary. Normalization scales the values in the range of 0 to 1 and standardization converts the values to zero mean and unity variance [34]. This procedure enhances the rate and stability of AI models, particularly the reinforcement learning agents.

Min-Max Normalization:

$$X' = \frac{X - X_{min}}{X_{max} - X_{min}} \quad (2)$$

Standardization:

$$Z = \frac{X - \mu}{\sigma} \quad (3)$$

- **Feature Engineering**

Feature engineering is a process of coming up with new meaningful attributes using the existing data to reflect network behavior in a better way. In an example, the spectral efficiency and throughput can be calculated by the use of the capacity formula by Shannon. Such artificial intelligence capabilities promote the learning ability of agentic AI systems to make optimal decisions.

$$C = Blog_2(1 + SINR) \quad (4)$$

where C is channel capacity (throughput), B is bandwidth, and SINR is the signal-to-interference-plus-noise ratio.

- **Encoding of Categorical Variables**

The data can contain such categorical variables as application type (e.g., video streaming, IoT, gaming). They need to be transformed into numerical data by methods such as one-hot encoding in order to be manipulated by machine learning models.

One-Hot Encoding:

$$x_i = \begin{cases} 1, & if\ category\ matches \\ 0, & otherwise \end{cases} \quad (5)$$

- **Data Splitting (Training and Testing Sets)**

To evaluate model performance, the dataset is split into training and testing subsets. Typically, 80% of the data is used for training and 20% for testing. This ensures that the model generalizes well to unseen network conditions.

$$D = D_{train} \cup D_{test}, \qquad D_{train} \cap D_{test} = \emptyset \quad (6)$$

- **Reward Design for Agentic AI (Reinforcement Learning)**

In the case of agentic AI-based resource allocation, it is proposed to have a reward function that will direct the learning process. The reward usually balances between throughput maximization, latency minimization and the efficient use of spectrum.

$$R = \alpha \cdot Througput - \beta \cdot Latency + \gamma \cdot Spectrum\ Efficiency \quad (7)$$

where α,β,γ are weighting factors.

- **Data Normalization for Time-Series or Dynamic Inputs**

As 5G networks are dynamic, temporal normalization or windowing methods are used to ensure consistency across time. This assists AI agents to understand trends in evolving network environments.

$$X_t' = \frac{X_t - X_{t-1}}{X_{t-1}} \quad (8)$$

These preprocessing measures will guarantee that the data has been cleaned, normalised and enriched with meaningful features and hence allow optimisation of throughput, latency and spectrum efficiency of agentic AI models in the 5G/6G wireless networks.

3.3 ENVIRONMENT MODELING

Environment modeling is the representation of wireless network as an interactive and dynamic system where there are numerous users who seek scarce radio resources. In this model, the network is described as a Markov Decision Process (MDP) where each state represents the prevailing conditions of the network i.e. channel quality, user demand, buffering state, latency needs and bandwidth available [35]. This state is observed by the agent, and an action, to be improved network performance, is picked, like assigning spectrum or scheduling users, is taken [36]. The stochastic nature of the wireless environment such as the interference and variability of the traffic contributes to the state transition depending on the action of the agent as well.

$$S_t = \{SINR_t, B_t, D_t, L_t \quad (9)$$

$$P(s_{t+1} \mid s_t, a_t) \quad (10)$$

where S_t is the state at time t, a_t is the action taken, and P represents the transition probability.

Moreover, the process of interaction between the agent and the environment is determined by a reward-driven learning process, which consists in learning that optimal policy that would produce maximum long-term cumulative reward [37]. The agent can make up-to-date decisions with dynamic networks; because each action is evaluated according to the immediate reward, expected future benefits. It is normally formulated in terms of the value function or Q-function which approximates the expected value of return on a certain state-action pair.

$$Q(s, a) = \mathbb{E}[\textstyle\sum_{t=0}^{\infty} \gamma^t R_t] \quad (11)$$

$$\pi^*(s) = arg \max_a Q(s, a) \quad (12)$$

where Q(s,a) is the action-value function, γ is the discount factor, and π∗(s) is the optimal policy. This modeling enables the agentic AI system to continuously adapt and optimize resource allocation in real-time 5G/6G network environments.

3.4 AGENTIC AI FRAMEWORK DESIGN

The Agentic AI system should be engineered to support autonomous and adaptive, as well as goal-oriented decision-making in resource distribution in 5G/6G wireless networks. In this method, a clever agent engages in a continuous interaction with the network, monitors the system state (channel conditions, user demand, latency requirements, and available spectrum) and makes the appropriate choice of actions such bandwidth allocation or user scheduling [38]. The structure is commonly trained with the methods of Deep Reinforcement Learning (DRL) like Deep Q-Network (DQN) and Proximal Policy Optimization (PPO) to enable the agent to learn the complex decision policies of data with high dimensions. The agent has a strategy π(a|s) that plays the actions on states and tries to optimize it with time by maximizing cumulative rewards [39]. The neural network is used to estimate the action-value function, thus the system is capable of managing the non-linear and dynamic network conditions effectively. This self-directed ability gives the system an agentic nature because it is able to learn, adapt and optimize itself without the involvement of a human being.

$$Q(s, a{:}\,\theta) \approx \mathbb{E}[\textstyle\sum_{t=0}^{\infty} \gamma^t R_t] \quad (13)$$

$$\theta \leftarrow \theta + \alpha \nabla_\theta \left(R + \gamma \max_{a'} Q(s', a'; \theta)\right) - Q(s, a; \theta)$$
(14)

Here, Q(s,a;θ) is the approximated Q-value using neural network parameters θ, γ is the discount factor, and α is the learning rate.

A reward mechanism directs the learning process and balances several objectives, including maximizing throughput, minimizing latency and enhancing spectrum efficiency. Exploration-exploitation techniques (e.g., ε-greedy policy) are applied by the agent to find the best actions as it narrows down its choices with time. Past interactions are stored using experience replay and stabilize training through breaking correlation between samples [40]. Also, convergence stability in implementations with DQN can be enhanced with target networks. The framework adjusts its policy on a constant basis based on perceived rewards and allows real-time adaptation to different network conditions and traffic patterns.

$$R = \alpha \cdot T - \beta \cdot L + \gamma \cdot \eta \quad (15)$$

$$\pi(a|s) = \{\begin{matrix} random\ action, & with\ probability\ \varepsilon \\ arg\ max_a Q(s, a), & with\ probability\ 1 - \varepsilon \end{matrix}$$
(16)

where T represents throughput, L is latency, η is spectrum efficiency, and α,β,γ are weighting factors.

Algorithm: Agentic AI-Based Resource Allocation (DQN Approach)

Input: State space, action space, learning rate, discount factor, exploration rate
Output: Optimal resource allocation policy

1. Initialize replay memory to store past experiences.
2. Initialize the Q-network with random weights.
3. For each training episode:
   - Initialize the current network state.
   - At each time step:
     - Select an action using an exploration–exploitation strategy (e.g., ε-greedy).
     - Execute the selected action in the environment.
     - Observe the reward and the next state.
     - Store the experience (state, action, reward, next state) in memory.
     - Randomly sample a batch of past experiences from memory.
     - Compute the target values using observed rewards and estimated future rewards.
     - Update the Q-network by minimizing the difference between predicted and target values.
     - Update the current state to the next state.
4. Repeat the process until the model converges to an optimal policy.
5. Output the learned policy for real-time resource allocation decisions.

3.5 REWARD FUNCTION DESIGN

The design of the reward functionality is the critical part of the Agentic AI framework, since it directly determines the learning behavior of the intelligent agent in the optimization of resource allocation decisions in 5G/6G wireless networks. It is designed in such a way that it captures several performance goals at the same time such as maximization of network throughput, minimum latency, spectrum efficiency and fairness between users and quality of service requirements are also encompassed [41]. The reward is usually developed as a weighted average of these measures that positive rewards are provided to efficient resource usage, high data rates and the negative reward has to be provided to address delays, congestion or inefficient use of spectrum. This balance will make sure that the agent will not over-optimize one goal sacrificing others [42]. It can also be mentioned that reward function can be adaptively weighted as a response to changing conditions in the network, including the fluctuating traffic load or user needs. The agent in the form of a continuous feedback that is in the form of rewards, over time learns to make smarter and context-aware choices and over time, results in an autonomous and robust resource allocation strategy that will work well with highly dynamic and complex next-generation wireless networks.

3.6 PROPOSED MODELS

➢ CNN

A Convolutional Neural Network (CNN) is a deep learning tool that is extensively applied in extracting spatial and hierarchical characteristics of structured data, and can be effectively implemented in the process of network optimization within 5G/6G systems. When applied in the context of resource allocation, CNNs are able to operate on multi-dimensional network data including traffic, signal strength map, interference, and user distribution, because they capture local dependencies and intricate interactions among features [43]. The architecture is usually based on convolutional layer, which is used to extract features, pooling layer, which is used to reduce the number of dimensions, and fully connected layer, which is used to make decisions or predictions [44]. CNNs enable automatic learning of relevant properties of raw input data, and therefore, remove the necessity of manually engineer features and enhances model performance. CNNs may also be applied as an approximation of value functions or policies in a reinforcement learning model, when they are incorporated into an Agentic AI. This allows the system to take intelligent real-time decisions on bandwidth, scheduling as well as spectrum [45]. All in all, CNN-based models improve the potential of the system to deal with the high-dimensional and dynamic wireless network surroundings and result in increased throughput, latency, and enhanced spectrum efficiency.

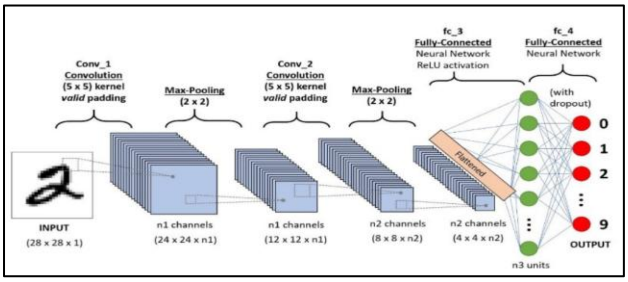


Figure 3: Architecture of CNN Model [45]

➢ Deep Reinforcement Learning (DRL)

Deep Reinforcement Learning (DRL) is a sophisticated method of artificial intelligence, which uses reinforcement learning alongside deep neural networks to allow intelligent agents to take sequential decisions in dynamic and complex settings. DRL in the 5G/6G wireless network is especially applicable to autonomous resource allocation whereby the system has to continuously adapt itself to varying network conditions which includes user demand, channel quality, interference and traffic load [46]. The agent has interactions with the network environment in both sensing what is occurring in the environment and acting such as assigning bandwidth or scheduling users and getting feedback in the form of rewards based on performance metrics such as throughput, latency, and spectrum efficiency.

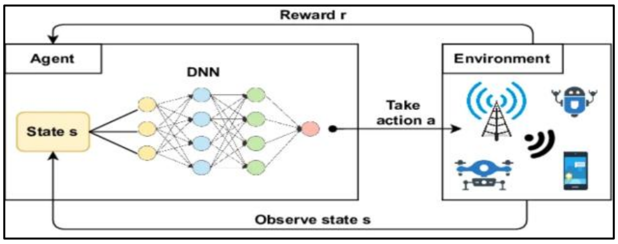


Figure 4: Deep Reinforcement Learning (DRL) design [48]

Value functions or policies are approximated with deep neural networks which enable the agent to deal with high-dimensional data and learn complicated patterns without explicit programming. Exploration and exploitation are other mechanisms used in DRL to reconcile between learning new strategies and optimizing existing ones so that it can make sound and adaptive decisions [47]. The agent then becomes more effective in optimization of cumulative rewards with time, achieving efficient and real-time and self-optimizing resource management and making DRL an important enabler of intelligent and scalable next-generation wireless communication systems.

➢ Hybrid Model (CNN+ DRL)

The hybrid architecture based on the combination of Convolutional Neural Network (CNN) and Deep Reinforcement Learning (DRL) would be an effective tool of intelligent resource allocation in 5G/6G wireless networks taking advantage of the advantages of both models. CNN, in this scheme, is an element of a feature extractor that feeds on high-dimensional network information (traffic patterns, signal strength distribution, and interference maps) producing meaningful representations [49]. These features are then provided to a DRL agent that is able to utilize them to take the most optimal sequential decisions in the form of bandwidth allocation, user scheduling and spectrum management. By means of such integration, the system will be in a better position to learn more about complicated spatial and temporal relationships in the network as well as providing real-time decisions that are adaptable. Hybrid model enhances learning efficiency, prediction accuracy and convergence rate as compared to standalone models. Consequently, it optimizes major key performance indicators including throughput, latency, and spectrum efficiency effectively, which makes it very ideal in autonomous and scalable next generation wireless communication system.

3.7 Model Training and Optimization

Under the proposed framework, the model training processes include facilitating the agent to develop the best resource allocation schemes by engaging him or her in the network repeatedly. The training works with the historical, simulated 5G /6G network data, as the agent monitors the state of the system, acts and get the rewards according to the performance results. It aims to maximize the cumulative rewards over the long-term with a balance on the duties to exclude throughput, latency and spectrum efficiency. In training, the model is revised with the aid of gradient based optimisation methods to reduce the discrepancy between the updated and the target parameters. The learning process enables the agent to estimate the best action-value function and to make better decisions as time goes by.

$$L(\theta) = \mathbb{E}[(y - Q(s,a;\theta))^2] \tag{17}$$

$$y = R + \gamma \max_{a'} Q(s', a'; \theta) \tag{18}$$

where L(θ) is the loss function, y is the target value, and γ is the discount factor.

Optimization methods are used to improve the training performance, stability and rate of convergence. Past experiences are stored and sampled randomly by methods like experience replay in order to reduce correlation between training samples. Also, the target networks reduce fluctuations in the dynamic of learning by optimal target values. Learning rate, discount factor, and exploration rate should be adjusted as hyperParameters in order to obtain the best performance. Regularization can also be used in order to avoid overfitting and enhance an improvement in generalization. Such optimization schemes outlined guarantee convergence of the model to an optimal policy which can make dynamic and correct resource allocation decisions under the dynamic wireless network systems.

$$\theta_{t+1} = \theta_t - \alpha \nabla_\theta L(\theta) \tag{19}$$

$$J(\theta) = \mathbb{E}_{\pi\theta}[R] \tag{20}$$

where J(θ) represents the expected reward under policy πθ is the learning rate.

3.8 PERFORMANCE EVALUATION

The performance measures of the trained model are throughput, latency, spectrum efficiency and fairness, which are evaluated. These are compared to the traditional approaches to allocation (e.g. static or heuristic approaches) and improvements are proved.

- **Throughput**

Throughput represents the total successful data transmission over the network.

$$Throughput = \frac{Total\ Data\ Successfully\ Transmitted}{Total\ Time} \tag{21}$$

- **Latency**

Latency is the total time delay experienced in transmitting data from source to destination.

$$Latency = Transmission\ Delay + Propagation\ Delay + Processing\ Delay + Queuing\ Delay \tag{22}$$

- **Spectrum Efficiency**

Spectrum efficiency measures how efficiently the available bandwidth is utilized.

$$Spectrum\ Efficiency = \frac{Throughput}{Bandwidth} \tag{23}$$

or using Shannon capacity

$$C = Blog_2(1 + SINR) \tag{24}$$

- **Fairness (Jain's Fairness Index)**

Fairness ensures equal or balanced resource allocation among users.

$$Fairness\ Index = \frac{\left(\sum_{i=1}^{n} x_i\right)^2}{n \cdot \sum_{i=1}^{n} x_i^2} \tag{25}$$

where x_i is the resource allocated to user i, and n is the total number of users.

All of these metrics are used to affect the evaluation of the model performance in option of its efficiency, delay reduction, use of resources, and fairness, and these values are needed to compare and contrast the AI-based methods with traditional ones.

4. RESULT AND DISCUSSION

In this section, the experimental outcomes and performance analysis of the proposed Agentic AI-based framework are provided in terms of CNN, DRL, and hybrid model frameworks. The likely metrics analyzed include throughput, latency, efficiency and fairness of various network conditions.

4.1 IoT Application Traffic Distribution

The study is a representation of the utilization of various types of IoT applications according to their usage of network traffic. It also emphasizes the role of different services in the total data flow, allowing to analyze communication patterns and resource demand, domination of real-time and low-intensity services.

Figure 5 represents the distribution of the various types of IoT application according to the number of occurrences. The maximum frequency of Video_Call is about 58, which means that it is used heavily in terms of real-time communication. With close to 48, Web_Browsing comes in the second place, with Streaming, Background_Download, Emergency_Service, and Video_Streaming less so, but with almost the same values of 47 to 48, indicating that all are medium-high demand. VoIP_Call and Online-Gaming exhibit a slightly lower usage about 46 and 45 respectively, which is moderate in terms of interaction. Conversely, IoT_Temperature documents a much lower number of approximately 13, which indicates the lack of regular sensor communicational activities. Voice_Call and File_Download are least represented with each close to 1 which is a very low utilization. On the whole, the statistics highlight that the bandwidth-intensive applications dominate over the lightweight IoT services in the network traffic patterns.

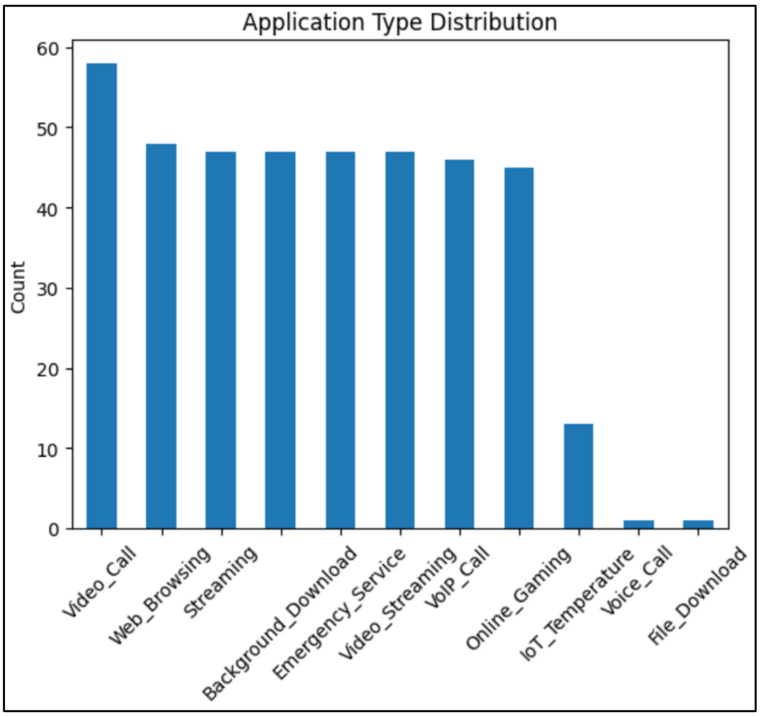


Figure 5: Distribution of IoT application types showing varying traffic intensity across communication categories

The correlation values between the important network parameters are shown in Figure 6. Signal_Strength has a weak negatively correlated relationship with Latency, Required_Bandwidth, and Allocated_Bandwidth (i.e. about -0.35), which means that strong signals minimally decrease delays and bandwidth requirements. There is a positive relationship between Latency and Required Bandwidth and Allocated Bandwidth (average of 0.45) that indicates that the higher the delay, the higher the bandwidth requirements and allocation is. The relationship between Required bandwidth and Allocated bandwidth shows a strong positive correlation (close to 1.0) thus showing efficient provisioning of resources where the allocated bandwidth matches the demand. The diagonal values of 1.0 verify the perfect self correlation. Altogether, the findings point to the idea that bandwidth control is very demand-oriented, and the signal density negatively affects the network performance metrics in the IoT communication setting.

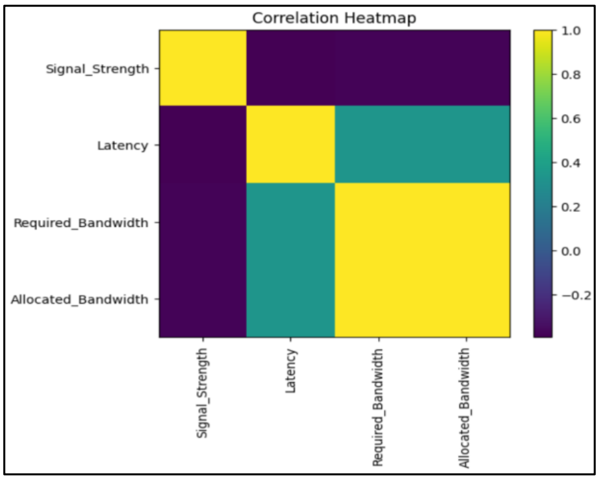


Figure 6: Correlation heatmap of signal strength, latency, and bandwidth in IoT networks

4.2 Distribution

The relationship between the values of signal strength between about -120 dBm and -40 dBm has been shown in Figure 7. Most of the observations are within the range (-100 dBm) to (-70 dBm), which shows a moderate signal condition throughout the network. The peaks are found at -95 dBm to -85 dBm indicating that most equipment runs at stable signal values but not at optimal values. There are fewer cases at extreme values lower than -110 dBm and higher than -50 dBm which signal weak and strong respectively. It is also slightly skewed to the right with a gradual increase to stronger signals. Generally, the evidence suggests that although the majority of IoT gadgets are able to sustain a decent connectivity, signal quality can vary, potentially affecting the network performance and reliability under a range of circumstances.

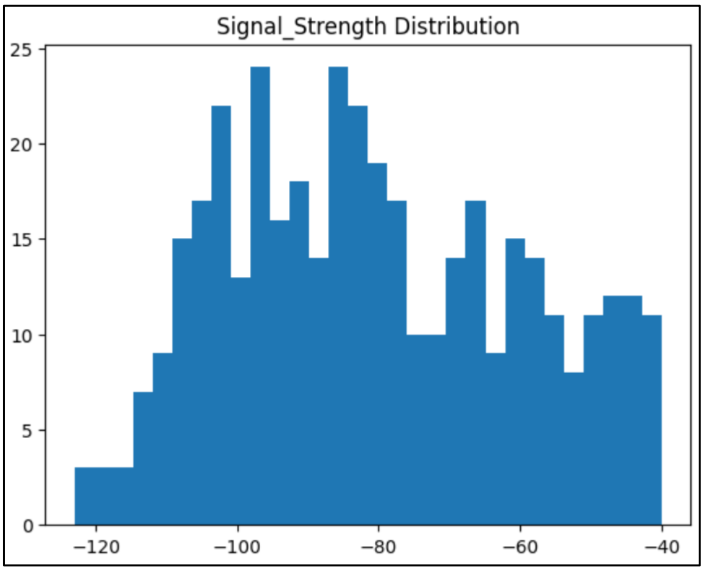


Figure 7: Distribution of signal strength values in IoT network

The Figure 8 supports ranges of latency, near 0 ms up to about 110 ms. The majority of observations lie in the range of 20ms to 50ms, which implies moderate network delay of most of the IoT communications. Peak frequency is observed at about 25-35 ms indicating normal operating latency in the system. Fewer cases are below 10 ms, which corresponds to low latency and high-performance. On the other hand, the higher latency values that are above 60 ms are less frequent yet go to 100110 ms, which means that there is occasional network congestion or delays. Its distribution is skewed to the right, and there is a long tail on the high latency. In general, the data has a relatively constant performance and certain fluctuations that can influence time-sensitive IoT applications.

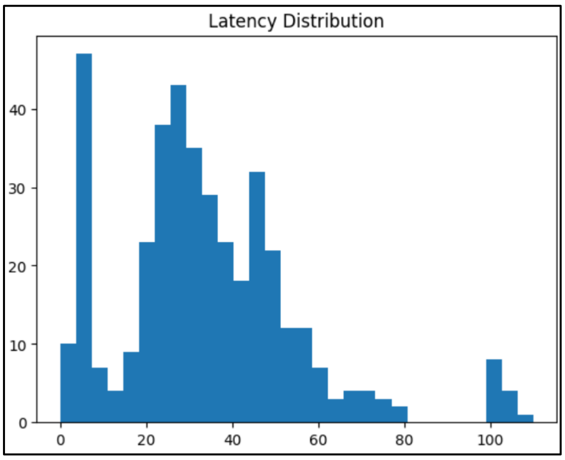
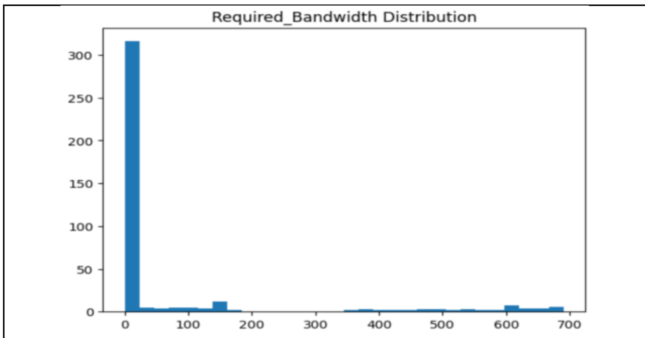
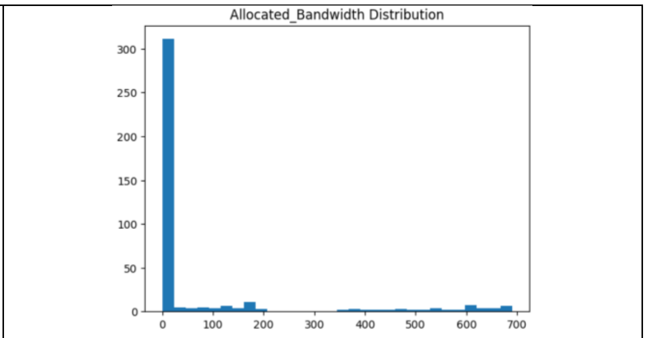


Figure 8: Distribution of latency values in IoT network

Figure 9 depict the allocation of bandwidth needed and distributed, and the bandwidth needed varies between almost 0 Mbps and about 700 Mbps. In both scenarios, most of the values fall below 50 Mbps, which implies that most IoT applications require and obtain low bandwidth. The highest frequency is approximately 020 Mbps, which represents lightweight communication. Increased bandwidth values of 100 Mbps up to 700 Mbps are less frequent but show resource-consuming applications. The similarity of the two distributions supports the fact that there is a close correspondence between the bandwidth required and allocated to indicate an efficient resource provisioning. The fact that both graphs are skewed to the right means that the demand is high occasionally. In general, the that data indicates that bandwidth allocation is a good way of ensuring that the requirements of the applications are taken care of without compromising the efficiency of the network.



9(a)



9(b)

Figure 9(a) & 9(b): Comparison of required and allocated bandwidth distributions in IoT network

4.3 Boxplot

Figure 10 shows how signal strength value of about -125 dBm to -40 dBm diffused. Its median is approximately -82 dBm, which is a moderate signal strength of most devices. Interquartile range (IQR) is in the range of about -100 dBm (Q1) to -65 dBm (Q3) as fifty percent of observations are on the range. The bottom whisker stretches out to approximately -123 dBm, which is less intense signals, and the top whisker goes to approximately -40 dBm, which is a strong signal situation. The distribution is fairly distributed, which shows variation in network connectivity. There are no major outliers, which indicates the stability of the performance. On the whole, the evidence shows that the signal strength of the various IoT devices remains steady yet erratic.

Figure 11 shows the results of latency that vary between about 0ms and more than 110ms. The median latency has a value of about 30 ms which shows that there is an average moderate delay in the network. The interquartile range (IQR) is 22 ms (Q1) to 45 ms (Q3) indicating that half of the values are within this range. The lower whisker is near 0 ms which is a case of minimal delay conditions, whereas the upper whisker points to about 78 ms, which denotes the cases of high latency spikes. Distribution would indicate a relatively steady performance with stable delay but the existence of outliers indicates that there could be congestion or performance problems in specific situations to time-sensitive IoT applications.

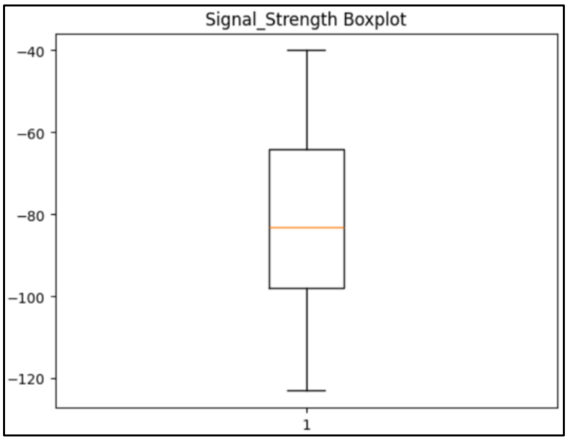


Figure 10: Boxplot representing signal strength variation in IoT network

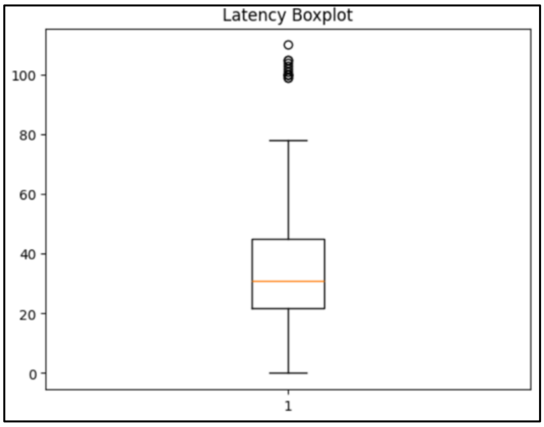


Figure 11: Boxplot of latency distribution in IoT network.

Both allocated and required bandwidth between around 0 Mbps and around 700 Mbps are present in figure 12 (12(a) and 12(b)). Both have a median of about 1015 Mbps, meaning that they use low bandwidth. The interquartile range (IQR) falls between 5 Mbps and 20 Mbps and indicates that most applications can work within the low bandwidth limits. But there are many outliers that stretch between 50 Mbps and 700 Mbps reflecting high-demand applications like video streaming. The fact that there is a similarity in the two plots implies that required bandwidth is directly next to allocated bandwidth. The large number of high-value outliers is an indication of occasional heavy network load. Generally, the statistics indicate effective bandwidth management and a high level of demand and provisioning correspondence within the IoT contexts.

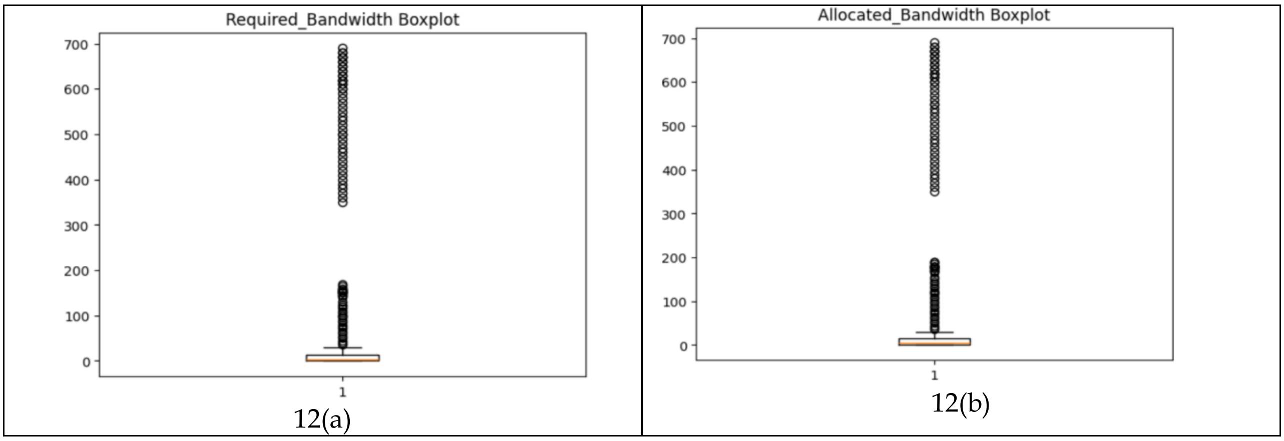


Figure 12: Boxplots comparing required and allocated bandwidth distributions in IoT network

4.4 CNN Model

Figure 13 depicts the accuracy in the classification of the CNN model at six classes (0- 5). High values along the diagonal denote the right predictions, with the first class (1) being the most accurate (approximately 28 instances), second (5), and the third and fourth (10) are also right but with minor accuracy. Class 3 has slightly lower correct predictions (~7), which means relatively poor performance. Off-diagonal values are minimal indicating low misclassification rates. There is some confusion observed between the class 1 and class 2 and a little bit with class 4 but this is confined. In general, the model works well in classification across most classes with high accuracy and recall and reliably performs to differentiate between the various categories of the IoT traffic or attack.

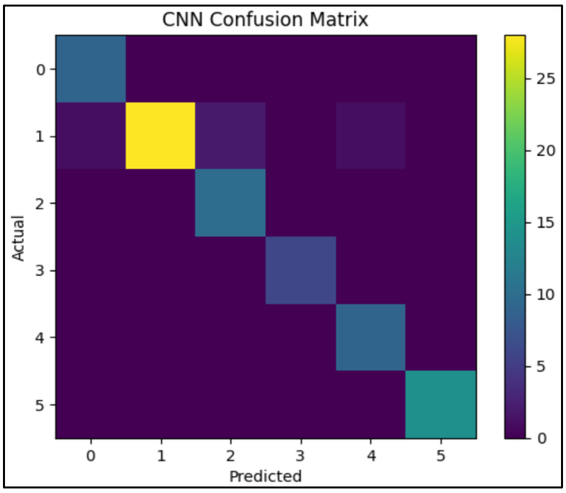


Figure 13: CNN confusion matrix showing classification performance across six classes

Figure 14 demonstrates the training and validation accuracy change with ten epochs. The accuracy of training steadily increases starting at the epoch 0 of about 0.47, reaching about 0.91 at epoch 4, and then it remains constant. The validation accuracy begins to be larger at 0.64 levels, although it decreases a little to 0.60 at epoch 1, and then it increases sharply to an average of 0.95 at epoch 3 and continues to be similar. The weakness between training and validation accuracy is small, and this implies good generalization with less overfitting. The fast convergence of the first 3-4 epochs is an indication of effective learning through the CNN model. Possessing a high accuracy (more than 90% on training and validation), in general, the model meets the requirements of high performance in classification tasks.

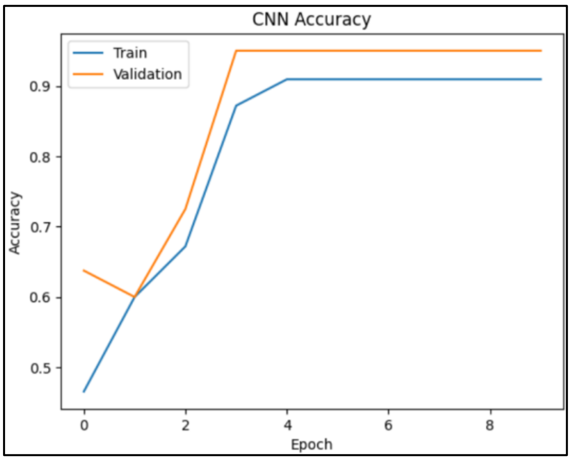


Figure 14: CNN training and validation accuracy over epochs

Figure 15 demonstrates how the training and validation loss are decreasing after 10 epochs. The initial loss is high at about 2.05 but gradually decreases to about 0.39 towards epoch 9, which shows that it is in a continuous learning. The validation loss starts at the point of approximately 1.85 and decreases faster until epoch 5 when it reaches almost 0.30 before leveling off with minor variances. Validation loss is always lower than training loss indicating good generalization and no overfitting. The steep reduction in the first 3 4 epochs is an indicator of effective model convergence. Minor variation since epoch 6 signals the stabilization of the model. All in all, the successful optimization was achieved in both curves, and low end loss values prove the presence of better predictions and stable work of the CNN model.

Figure 16 shows the key performance indicators of CNN model, such as throughput, latency, efficiency and fairness. Latency exhibits the largest value of about 0.29 meaning that it has the greatest influence on the network performance. Following fairness with approximately 0.19 which indicates equal allocation of resources among applications. The throughput is not very high, and it is approximately 0.085 which implies that there is a moderate level of data processing. Efficiency is the lowest of the three values, at 0.075, which indicates a relatively poor degree of optimization in the use of resources. The difference between these metrics points to trade-offs of network performance, in which the improvement of one metric can affect other factors. On the whole, CNN model has high latency processing and fairness and can be enhanced with throughput and efficiency in IoT network settings.

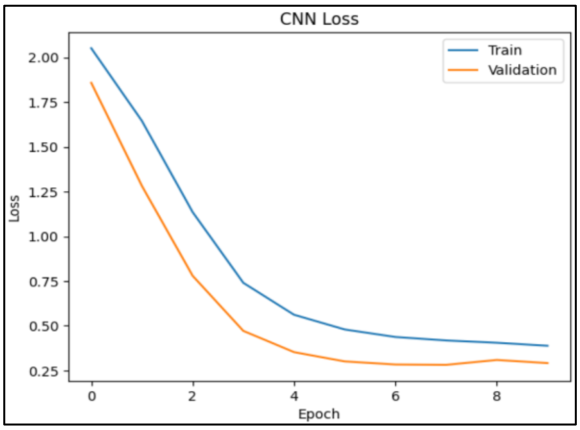


Figure 15: CNN training and validation loss over epochs

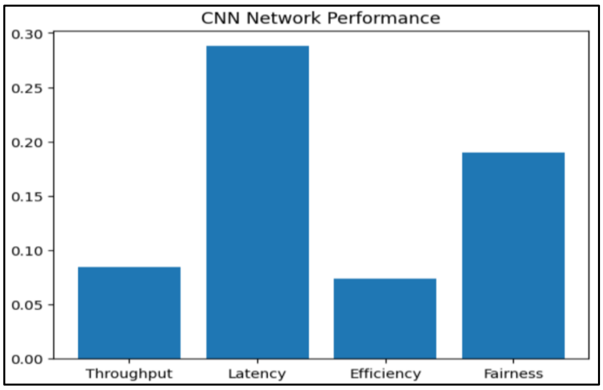


Figure 16: CNN network performance metrics comparison

4.5 Deep Reinforcement Learning (DRL)

The proposed framework makes use of Deep Reinforcement Learning (DRL) in making autonomous decisions. A Deep Q-Network (DQN) is used as the DRL algorithm to acquire the optimal policies of resource allocation in this implementation. Figure 17 depicts the reward development of the DQN model during 25 episodes. The reward begins at a level of about 2995 and rapidly rises to a high of about 3035 towards the episode 2, which shows that it learns at a very fast rate. Following this high, the reward oscillates between 2980 and 3010 indicating exploration and policy changes. The minimum reward is recorded to be around 2979 and the majority of the reward is concentrated towards 2990-3005 in the later episodes. The oscillations show continuous learning and adaptation and not the fluctuation. The range of variation is rather limited, indicating that performance is uniform once they start converging. On the whole, the DQN model exhibits consistent learning behavior in the situations with high reward value, which proves efficient policy optimization and enhanced decision-making throughout the network management.

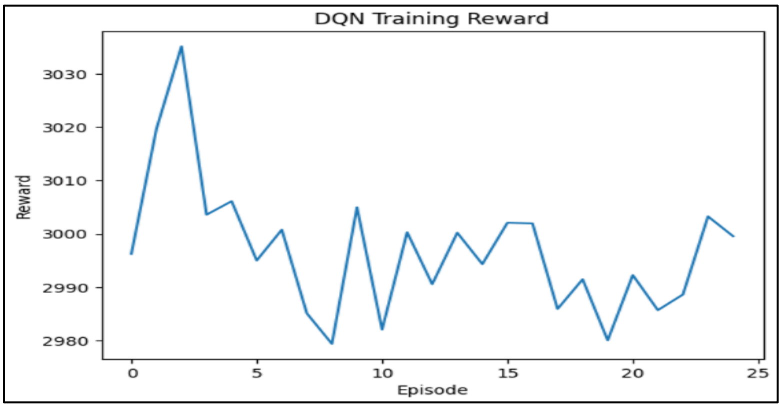


Figure 17: DQN training reward variation across episodes

The figure 18 shows the important network performance indicators such as throughput, latency, efficiency, and fairness. The highest value of latency is about 0.29, which means that it is more important in performance and could cause delays. The next one is fairness with a value of approximately 0.19, which implies that resource allocation to users or applications is relatively balanced. The throughput is approximately 0.085 which shows a moderate capability to transmit data. The least value of efficiency is close to 0.075, which means that it could be improved in the aspect of optimal resource utilization. These differences between these values are used to reflect the trade-offs in system performance, in which an improvement in latency or fairness can affect throughput and efficiency. The network in general is very fair and has manageable latency, but throughput and efficiency are to be optimized.

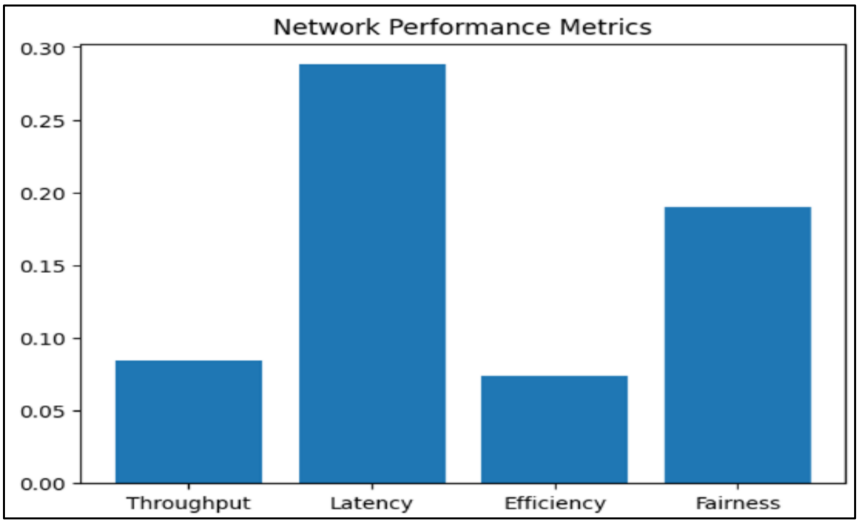


Figure 18: Network performance metrics comparison

Figure 19 shows the convergence of rewards of the DQN model with 25 episodes. There is early rapid learning as the reward increases at a rate of about 2995 with a peak of around 3035 at episode number 2. Subsequently, rewards vary within a slimmer spectrum of the range of 2980 to 3010, which indicates the stabilization of the process of learning. The minimum value is nearly 2979 with most of the values being near 2990–3005 in subsequent episodes. These changes are indicative of further exploration and still overall performance. The convergence pattern implies that the model attains a consistent policy in a few episodes. In general, the DQN shows stable and predictable behavior with low variance, which proves efficient learning and optimization of network decision-making problems.

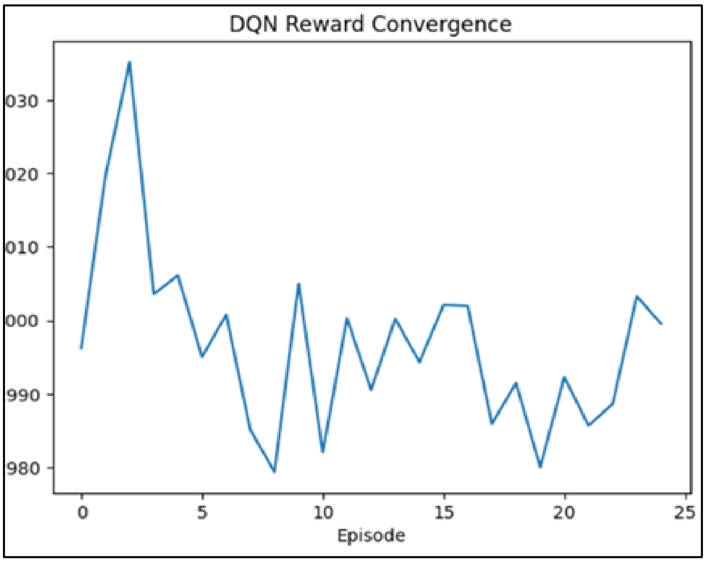


Figure 19: DQN reward convergence across training episodes

The reward progression of the hybrid model, on 27 episodes, is depicted in Figure 20. The reward values are between 803 and 829, which shows that it has a stable performance. The initial value of 814 rises to a peak value of about 828 at episode 10, which is an indication of successful learning. There are temporary performance dips at around 803 which is the lowest reward close to episode 7. Rewards curve after the peak lie in a tight range of 804-818, which points to convergence and stability. The fact that there are no drastic changes implies that the learning process is stable. In general, the hybrid model has a stable performance with moderate changes over time, which indicates stable optimization and flexibility in dealing with network decision-making processes relative to more fluctuate reinforcement learning methods.

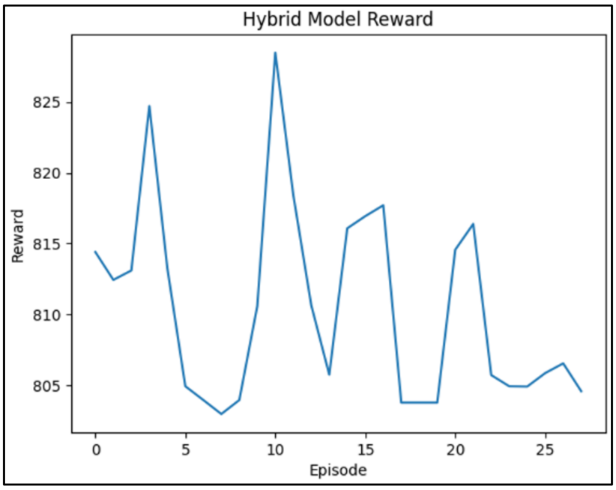


Figure 20: Hybrid model reward trend across training episodes

Figure 21 relates throughput, latency, efficiency and fairness between CNN, DRL and Hybrid models. Hybrid model records the best throughput (~0.10), efficiency (~0.085), and fairness (~0.23), which implies that it performs better in general. DRL has moderate values on throughput (0.09), efficiency (0.08) and fairness (0.21) whilst CNN has the lowest across all the measures. CNN has the longest (~0.28) followed by DRL (~0.27) and Hybrid (~0.26) with a slight advantage in delay management with the Hybrid model. The findings also emphasize that the Hybrid solution is more balanced in terms of all indicators, as it optimizes both performance and resource distribution, whereas CNN and DRL demonstrate relatively low efficiency and equity in the context of an IoT network.

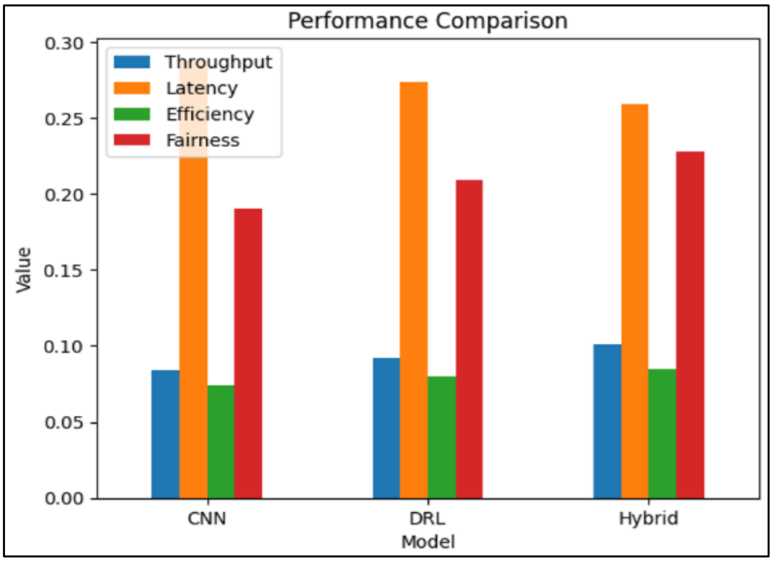


Figure 21: Performance comparison of CNN, DRL, and Hybrid models

Figure 22 is a comparison of reward trends of DRL and Hybrid models across 27 episodes. DRA model has a high and stable reward of approximately 2990-3020 and it has a high rate of 3035 and this shows a good learning performance and convergence. On the other hand, Hybrid model works at a lower reward scale, which is around 800-830, with small fluctuations. Both models exhibit stable trends as there is very little variability across episodes, although the magnitude differs. DRL also tends to cluster around 3000, which indicates that policy learning is optimized highly and the Hybrid model is moderately consistent with minor oscillations. In general, DRL performs better in terms of reward, and Hybrid is relatively more stable but has lower reward optimization on network decision-making problems.

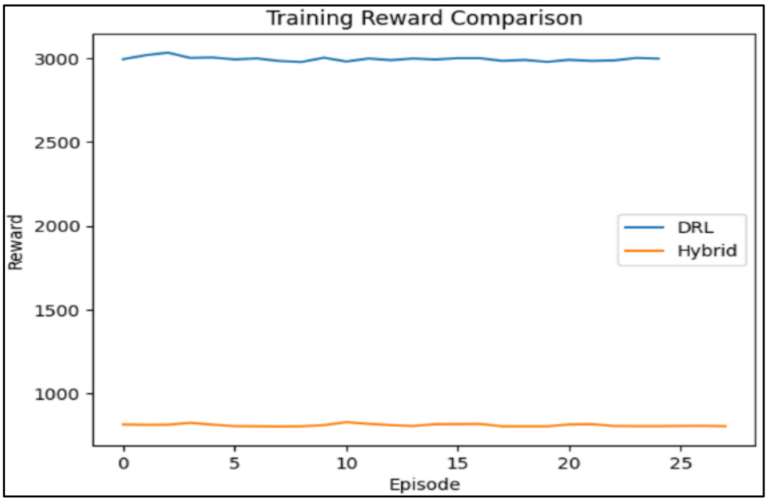


Figure 22: Training reward comparison between DRL and Hybrid models

Figure 23 shows normalized throughput, latency, efficiency, and fairness of CNN, DRL, and Hybrid models. Latency is maximally normalized to 1.0, and therefore it has the greatest impact relative to all metrics. The throughput and efficiency are near 0.0 and make relatively less contribution in the normalized scale. The value of fairness is small with a value of about 0.04 which means that it has a slight, but significant contribution. The similar trends are indicated by the overlapping lines by all three models after normalization. Latency is the most important metric in performance evaluation, whereas other metrics seem not as important, in comparison. In general, the normalized comparison highlights the latency as the most important aspect influencing model performance in optimization of IoT network.

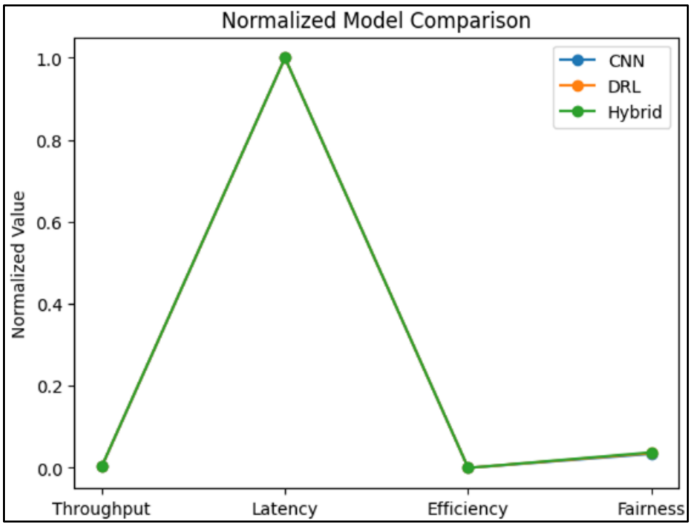


Figure 23: Normalized comparison of CNN, DRL, and Hybrid model performance metrics

5. CONCLUSION

The experiments indicate that the suggested hybrid framework based on Agentic AI proves to be effective in the optimization of the resources allocation to 5G/6G wireless networks. The hybrid model compared to individual CNN and DRLs has higher throughput (~0.10), efficiency (~0.085) and fairness (~0.23), lower latency (~0.26) than the CNN (~0.28) and the DRL (~0.27). The CNN model was found to have a classification accuracy of over 91% with validation accuracy of about 95% which is a good indicator of good feature extraction ability. Furthermore, the results of the confusion matrix reveal that the correct predictions are high (up to approximately 28 instances), which proves that the classification is reliable. The DRL model experienced consistent convergence in reward about 2990-3035, which indicates the efficient policy learning and adaptation. Although the hybrid model worked under a lower reward scale (~803-829), it was found to have the same performance and stable performance among episodes. In addition, correlation analysis showed that there was a high relationship (~1.0) between the bandwidth needed and the bandwidth allocated, which enables efficient provisioning of the resources. In general, the suggested framework can effectively balance various goals, enhance the performance of the network, and offer the scalable solution to smart and autonomous management of resources in the next-generation wireless communication systems.

REFERENCES

[1] Erbayat, E., Figueiredo, G. B., Lin, S. C., Matsuura, M., Hasegawa, H., & Subramaniam, S. (2026). A benchmarking framework for PON-based fronthaul network design. *arXiv preprint arXiv:2601.14480*.

[2] Javaid, S., & Saeed, N. (2026). The post-electromagnetic era: A vision for wireless communication beyond 6G. Array, 29, 100714.

[3] Khan, R., Zainab, B., Al Prince, A., Iftikhar, M., & Raza, A. (2025). Artificial Intelligence And 6g Integration: Transforming The Digital Technology Landscape. Spectrum of Engineering Sciences, 717-737.

[4] Sun, Y., Liu, Y., Guo, S., Zhang, R., Wang, J., Qiu, X., ... & Wu, Q. (2025). A synergy of computing power networks and low-altitude economy intelligent communications: Challenges, design principles, and research directions. arXiv preprint arXiv:2509.23810.

[5] Zhang, S., Qiu, L., & Zhang, H. (2025). Edge cloud synergy models for ultra-low latency data processing in smart city iot networks. *International Journal of Science*, *12*(10).

[6] Ahmed, Y., & Grace, D. (2025). Overview of dynamic spectrum sharing between leo satellites and mobile networks: Enabling seamless connectivity and coexistence in 6g. *IEEE Access*.

[7] Horvath, K., Tuda, S., Idrizi, B., Kitanov, S., Doko, F., & Kimovski, D. (2025, June). 6G infrastructures for edge AI: an analytical perspective. In *2025 IEEE International Parallel and Distributed Processing Symposium Workshops (IPDPSW)* (pp. 1066-1072). IEEE.

[8] Yang, M., Qu, Y., Ranbaduge, T., Thapa, C., Sultan, N. H., Ding, M., ... & Mrabet, S. (2026). From 5g to 6g: A survey on security, privacy, and standardization pathways. ACM Computing Surveys, 58(8), 1-38.

[9] Jyothi, V. E., Rajashree, R., Phani, S., Hugar, L., Shaik, R., & Poddar, S. (2025, September). Artificial Intelligence-Powered Protocols for Enhancing Security in Next-Generation Communication Networks. In 2025 3rd International Conference on Intelligent Cyber Physical Systems and Internet of Things (ICoICI) (pp. 1107-1112). IEEE.

[10] Omheni, N., Koubaa, H., & Zarai, F. (2025). Artificial intelligence for 5G and 6G networks: A taxonomy-based survey of applications, trends, and challenges. *Technologies*, *13*(12), 559.

[11] Moreira, R., Moreira, L. F. R., & Silva, F. D. O. (2025). Unleashing AI-Empowered Slices on Mobile Networks for Natively Cognitive Service Delivery. *IEEE Access*.

[12] Cao, X., Yang, B., Wang, K., Li, X., Yu, Z., Yuen, C., ... & Han, Z. (2024). AI-empowered multiple access for 6G: A survey of spectrum sensing, protocol designs, and optimizations. *Proceedings of the IEEE*, *112*(9), 1264-1302.

[13] Aasa, Z., Elias, F., & Ekpo, S. C. (2026). Hybrid Energy and Spectrum Efficient Wireless Network Design for 5G/6G And Wi-Fi 7/8 Applications.

[14] Annapareddy, V. N., Gadi, A. L., Kommaragiri, V. B., Koppolu, H. K. R., & Kannan, S. (2023). AI-Driven Optimization of Renewable Energy Systems: Enhancing Grid Efficiency and Smart Mobility Through 5G and 6G Network Integration. *Available at SSRN 5205082*.

[15] Zuo, Y., Guo, J., Gao, N., Zhu, Y., Jin, S., & Li, X. (2023). A survey of blockchain and artificial intelligence for 6G wireless communications. *IEEE Communications Surveys & Tutorials*, *25*(4), 2494-2528.

[16] Vashisht, S., Rani, S., & Feng, H. (2026). Multimodal and Agentic Intelligence–Driven ML Fusion for Sustainable 6G Network Slicing. *IEEE Open Journal of the Communications Society*.

[17] Aasa, Z. (2026). Autonomous resource orchestration for 6G space-air-ground networks: a self-supervised learning approach.

[18] Siddique, I. (2026). Mobility-Aware Data Offloading Mechanisms for Edge-Fog Networks Supporting Autonomous Vehicles and 5G Communications: https://doi. org/10.5281/zenodo. 19230644. *Research Consortium Archive*, *4*(2), 194-206.

[19] Thenmozhi, M., & Nawaz, G. K. (2026). Smart Agentic AI-powered scalable blockchain security for efficient traffic data sharing using deep featured ensemble learning. *National Journal of Antennas and Propagation*, *8*(1), 203-214.

[20] Mohammed, A., Mohammed, Z. A., Mohammed, N. U., Gunda, S. K. R., Ansari, M. A., & Raheem, M. A. AI-native wireless networks: Transforming connectivity, efficiency, and autonomy for 5G/6G and beyond."

[21] Wu, B., Wang, S., Zhang, Y. Q., Sifakis, J., & Ouyang, Y. (2026). Leveraging AI Agents for Autonomous Networks: A Reference Architecture and Empirical Studies. *IEEE Communications Magazine*.

[22] Khowaja, S. A., Dev, K., Pathan, M. S., Zeydan, E., & Debbah, M. (2025). Integration of agentic ai with 6g networks for mission-critical applications: Use-case and challenges. *arXiv preprint arXiv:2502.13476*.

[23] Lazrek, H., El Ferindi, H., Zouiten, M., & Moumen, A. (2025). Enhancing energy efficiency in 5G networks through AI-driven dynamic discontinuous reception. *Discover Computing*, *28*(1), 245.

[24] Othman, W. M., Ateya, A. A., Nasr, M. E., Muthanna, A., ElAffendi, M., Koucheryavy, A., & Hamdi, A. A. (2025). Key enabling technologies for 6G: The role of UAVs, terahertz communication, and intelligent reconfigurable surfaces in shaping the future of wireless networks. *Journal of Sensor and Actuator Networks*, *14*(2), 30.

[25] Kakarlapudi, R. V., & Yaramchitti, J. K. (2024). Agentic AI for Autonomous Telecom Network Management. *International Journal of Emerging Research in Engineering and Technology*, *5*(3), 129-135.

[26] Javaid, S., Khalil, R. A., Saeed, N., He, B., & Alouini, M. S. (2024). Leveraging large language models for integrated satellite-aerial-terrestrial networks: Recent advances and future directions. *IEEE Open Journal of the Communications Society*, *6*, 399-432.

[27] Bagwari, A., Logeshwaran, J., Raja, M., Devisivasankari, P., Bagwari, J., Rathi, V., & Saad, A. M. E. (2024). Intelligent computational model for energy efficiency and AI automation of network devices in 5G communication environment. *Tsinghua Science and Technology*, *29*(6), 1728-1751.

[28] Chauhan, D., Mewada, H., Gondalia, V., Almalki, F. A., Patel, S., Modi, H., ... & Mujlid, H. M. (2024). Balancing technological innovation and environmental sustainability: a lifecycle analysis of 6G wireless communication technology. *Sustainability*, *16*(15), 6533.

[29] Tomaszewski, L., & Kołakowski, R. (2023, January). Mobile services for smart agriculture and forestry, biodiversity monitoring, and water management: Challenges for 5G/6G networks. In *Telecom* (Vol. 4, No. 1, pp. 67-99). MDPI.

[30] Rao, P. M., Jangirala, S., Pedada, S., Das, A. K., & Park, Y. (2023). Blockchain integration for IoT-enabled V2X communications: a comprehensive survey, security issues and challenges. *IEEE Access*, *11*, 54476-54494.

[31] https://www.kaggle.com/datasets/omarsobhy14/5g-quality-of-service

[32] Khanh, Q. V., Hoai, N. V., Manh, L. D., Le, A. N., & Jeon, G. (2022). Wireless communication technologies for IoT in 5G: Vision, applications, and challenges. *Wireless Communications and Mobile Computing*, *2022*(1), 3229294.

[33] Khadri, W., Reddy, J. K., Mohammed, A., & Kiruthiga, T. (2024, July). The smart banking automation for high rated financial transactions using deep learning. In *2024 IEEE 3rd World Conference on Applied Intelligence and Computing (AIC)* (pp. 686-692). IEEE.

[34] Gures, Emre, Ibraheem Shayea, Mustafa Ergen, Marwan Hadri Azmi, and Ayman A. El-Saleh. "Machine learning-based load balancing algorithms in future heterogeneous networks: A survey." *IEEE Access* 10 (2022): 37689-37717.

[35] Janamolla, K., Balammagary, S., & Mohammed, A. (2024). Blockchain Enabled Cybersecurity to Protect LLM Models in FinTech.

[36] Shakya, A. K., Pillai, G., & Chakrabarty, S. (2023). Reinforcement learning algorithms: A brief survey. *Expert Systems with Applications*, *231*, 120495.

[37] Syed, W. K., Mohammed, A., Reddy, J. K., & Dhanasekaran, S. (2024, July). Biometric authentication systems in banking: A technical evaluation of security measures. In *2024 IEEE 3rd World Conference on Applied Intelligence and Computing (AIC)* (pp. 1331-1336). IEEE.

[38] Mohammed, A. K., & Ansari, M. A. The impact and limitations of AI in Power BI."

[39] Chen, Wanshi, Xingqin Lin, Juho Lee, Antti Toskala, Shu Sun, Carla Fabiana Chiasserini, and Lingjia Liu. "5G-advanced toward 6G: Past, present, and future." *IEEE journal on selected areas in communications* 41, no. 6 (2023): 1592-1619.

[40] Sahni, V., Arora, K., Devi, M., & Bhaggi, E. (2025, February). Exploring the 6G era through artificial intelligence and machine learning. In *AIP Conference Proceedings* (Vol. 3224, No. 1, p. 020013). AIP Publishing LLC.

[41] Zhang, C., Bengio, S., Hardt, M., Recht, B., & Vinyals, O. (2021). Understanding deep learning (still) requires rethinking generalization. *Communications of the ACM*, *64*(3), 107-115.

[42] Wang, Y., Xu, Y., Shi, Q., & Chang, T. H. (2021). Quantized federated learning under transmission delay and outage constraints. *IEEE Journal on Selected Areas in Communications*, *40*(1), 323-341.

[43] Mohsen, S., Ali, A. M., & Emam, A. (2024). Automatic modulation recognition using CNN deep learning models. *Multimedia Tools and Applications*, *83*(3), 7035-7056.

[44] Balammagary, S., Mohammed, N., Mohammed, S., & Begum, A. (2025). AI-Driven Behavioural Insights for Ozempic Drug Users. *Journal of Cognitive Computing and Cybernetic Innovations*, *1*(1), 10-13.

[45] Das, A., & Dey, S. (2021). Forecasting Long-term Electricity Demand: Evolution from Experience-Based Techniques to Sophisticated Artificial Intelligence (AI) Models. In *Computational Management: Applications of Computational Intelligence in Business Management* (pp. 553-586). Cham: Springer International Publishing.

[46] Das, A., & Dey, S. (2021). Forecasting Long-term Electricity Demand: Evolution from Experience-Based Techniques to Sophisticated Artificial Intelligence (AI) Models. In *Computational Management: Applications of Computational Intelligence in Business Management* (pp. 553-586). Cham: Springer International Publishing.

[47] Costa, V. G., & Pedreira, C. E. (2023). Recent advances in decision trees: An updated survey. *Artificial Intelligence Review*, *56*(5), 4765-4800.

[48] Ahmed, M. I., Mohammed, A. R., Ganta, S. K., Kolla, S. K., & Kashif, M. K. (2025). AI-driven green construction: Optimizing energy efficiency, waste management and security for sustainable buildings. *Journal of Cognitive Computing and Cybernetic Innovations*, *1*(1), 37-41.

[49] Mihalič, F., Truntič, M., & Hren, A. (2022). Hardware-in-the-loop simulations: A historical overview of engineering challenges. *Electronics*, *11*(15), 2462.