ISSN: 2093-5374 / E-ISSN: 2093-5382

# Deep Reinforcement Learning Driven Dynamic Spectrum Access in Multi-Tier 5G NR Networks

Dr.P.M. Murali[1*], Dr. Prakash Chandra Sahoo[2], Dr.M.S. Divyarani[3], E. Shalini[4], Dr. Aastha Mishra[5], Jyotsna Suryavanshi[6], and Meka Sreedhar Rao[7]

[1*]Associate Professor, Department of Electronics and Communication Engineering, Vinayaka Mission's Kirupananda Variyar Engineering College, Vinayaka Mission's Research Foundation (DU), Tamil Nadu, India. muralipm@vmkvec.edu.in, https://orcid.org/0000-0002-2991-0740

[2]Assistant Professor, Department of Computer Science and Engineering, Institute of Technical Education and Research, Siksha 'O' Anusandhan (Deemed to be University), Bhubaneswar, Odisha, India. prakashchandrasahoo@soa.ac.in, https://orcid.org/0009-0006-9966-1854

[3]Professor, Department of Electronics and Communication Engineering, Presidency University, Bangalore, Karnataka, India. divyarani@presidencyuniversity.in, https://orcid.org/0000-0003-1212-4150

[4]Assistant Professor, Department of Computer Science, Meenakshi College of Arts and Science, Meenakshi Academy of Higher Education and Research, Chennai, Tamil Nadu, India. shalini@maher.ac.in, https://orcid.org/0009-0002-1391-0057

[5]Assistant Professor, Department of School of Engineering & Technology, Noida International University, Uttar Pradesh, India. aastha.mishra@niu.edu.in, https://orcid.org/0009-0005-9410-2597

[6]Department of Engineering, Science and Humanities, Vishwakarma Institute of Technology, Pune, Maharashtra, India. jyotsna.suryavanshi@vit.edu, https://orcid.org/0009-0006-3068-4230

[7]Professor, Faculty of Ayurveda, Gokul Global University, Sidhpur, Gujarat, India. provost@gokuluniversity.ac.in, https://orcid.org/0009-0006-2683-7548

Received: July 08, 2026; Revised: August 19, 2026; Accepted: September 07, 2026; Published: September 30, 2026

## Abstract

The increasing number of bandwidth-intensive applications and heterogeneous deployment environments in Fifth Generation New Radio (5G NR) networks leads to a growing need for efficient spectrum utilization. In multi-tier environments (such as macro cells, micro cells, pico cells, and femto cells), traditional static and heuristic spectrum allocation approaches often lead to underutilization of spectrum resources, excessive co-channel interference, and reduction of quality of service (QoS). To overcome these problems, this paper proposes a Deep Reinforcement Learning-Driven Dynamic Spectrum Access (DRL-DSA) framework for intelligent spectrum management in multi-tier 5G NR networks. The proposed framework utilizes a Deep Q-Network (DQN) agent that learns spectrum access strategies by observing network states, including traffic demands, interference levels, and available spectrum bandwidth resources, over time. The

*Journal of Wireless Mobile Networks, Ubiquitous Computing, and Dependable Applications (JoWUA),* volume: 17, number: 3 (September-2026), pp. 2041-2056. DOI: *10.58346/JOWUA.2026.I3.118*

*Corresponding author: Associate Professor, Department of Electronics and Communication Engineering, Vinayaka Mission's Kirupananda Variyar Engineering College, Vinayaka Mission's Research Foundation (DU), Tamil Nadu, India.

framework develops a multi-objective reward function to maximize throughput and spectrum efficiency while minimizing interference and access latency. Extensive simulations under various network densities and traffic environments are implemented to demonstrate the performance of the proposed DRL-DSA scheme. The results show that the proposed scheme increases spectrum efficiency by 31.8%, improves network throughput by 62.3%, and reduces interference by 24.6% compared to the static spectrum allocation scheme. Meanwhile, compared with current DQN-based spectrum access methods, the proposed model reduces access latency by 19.7% and increases the QoS satisfaction ratio by 23.5%. In general, the proposed DRL-DSA scheme can provide a scalable, dynamic, and efficient approach to spectrum management in heterogeneous 5G NR environments and promote the realization of intelligent wireless communication systems and the development of future 6G networks.

**Keywords**: Deep Reinforcement Learning, Dynamic Spectrum Access, 5G New Radio, Multi-Tier Networks, Deep Q-Network, Spectrum Efficiency, Interference Management, Quality of Service.

# 1 Introduction

With the development of 5th Generation New Radio (5G NR) networks, such as enhanced Mobile Broadband (eMBB), Ultra-Reliable Low-Latency Communications (URLLC), and massive Machine-Type Communications (mMTC), demand for wireless spectrum resources has increased exponentially. As the number of connected devices and the adoption of bandwidth-intensive applications grow, spectrum resource efficiency becomes an important challenge in next-generation wireless systems (Gbenga-Ilori et al., 2025). In 5G, heterogeneous multi-tier architectures comprising macro, micro, pico, and femto cells are adopted to enhance coverage and capacity. Although deploying multi-tier networks improves network performance, it introduces complex interference patterns, dynamic traffic variations, and resource allocation conflicts between tiers (Bartsiokas et al., 2022). Existing spectrum management approaches are often inefficient when coping with these highly variable and dynamic network conditions. Artificial Intelligence techniques have achieved remarkable breakthroughs across various fields, and the application of deep reinforcement learning (DRL) to dynamic radio resource management and spectrum sharing in wireless communication networks shows promising results (Umer et al., 2025; Giannopoulos et al., 2021).

Traditional spectrum management approaches, which rely on static allocation policies, optimization techniques, or rule-based methods, often assume stable and predictable network conditions and exhibit poor adaptability to dynamic traffic demands, user mobility, and interference characteristics that change over time and across network tiers (Bartsiokas et al., 2022; Nouruzi et al., 2023). Current DRL-based resource management research has achieved positive outcomes; nevertheless, many of the works considered either single-tier cell networks or cognitive radio networks, or are based on simplified simulation environments that cannot accurately reflect dense multi-tier 5G NR deployments (Filali et al., 2022; Boutiba et al., 2023). Moreover, previous solutions mostly optimize a single objective, such as throughput or energy consumption, whereas spectrum efficiency, interference management, and Quality of Service (QoS) are not jointly optimized (Cao et al., 2024; Alhazmi & Arafah, 2025). Thus, an intelligent and adaptive dynamic spectrum access framework that learns to make spectrum allocation decisions in heterogeneous, multi-tier environments remains needed.

In this paper, a DRL-driven Dynamic Spectrum Access (DRL-DSA) framework is proposed for efficient spectrum management in multi-tier 5G NR networks. This framework utilizes a DQN-based learning agent that continuously monitors network states to dynamically make spectrum access decisions in response to varying traffic demands and interference. A multi-objective reward function is designed

to simultaneously optimize spectrum efficiency and throughput, while minimizing interference and access latency. The performance evaluation is conducted across various user densities and traffic conditions to demonstrate its scalability and adaptability. By integrating intelligence into spectrum access decisions, this work aims to enhance the system's overall performance and pave the way for future autonomous wireless communications (Umer et al., 2025; Cao et al., 2024; Hazarika et al., 2024).

In dense multi-tier 5G NR networks, with increasing traffic demand, heterogeneous network deployment, and the inherently limited radio spectrum resource, efficient spectrum utilization is the key challenge for guaranteeing the reliable, service-quality, and sustainable growth of future wireless communication systems. This paper proposes an adaptive DRL-driven Dynamic Spectrum Access framework for the multi-tier 5G NR network. It introduces a novel multi-objective learning and spectrum access strategy that intelligently optimizes throughput, spectrum efficiency, interference management, and QoS, providing a scalable and smart spectrum management solution for heterogeneous multi-tier 5G NR environments.

The rest of the paper is structured as follows. Section II provides a comprehensive review of related work on dynamic spectrum access (DSA), machine learning-based spectrum management, deep reinforcement learning (DRL)- based spectrum management, and optimization in 5G and beyond. Section III presents the design of the proposed DRL-driven DSA framework, namely, a multi-tier 5G NR system model, state and action design, reward function definition, and the DRL algorithm training process. Section IV describes the simulation environment, evaluation metrics, and experimental results, including performance comparison and ablation study. Section V further analyzes the results and discusses scalability issues, deployment concerns, and future network requirements. Section VI concludes the paper by emphasizing the significant findings and outlining potential avenues for future research.

## 2 Literature Review

Dynamic Spectrum Access (DSA) has become the predominant strategy to tackle the ever-increasing demand for spectrum in 5G-and-beyond wireless networks. Instead of the fixed allocation method, DSA enables opportunistic and adaptive spectrum sharing to accommodate time-varying availability and changing network conditions. Traditional static spectrum allocation methods have gradually evolved towards collaborative and dynamic spectrum sharing models, such as IEEE 802.11af and Licensed-Assisted Access (LAA), which further optimize spectral efficiency to accommodate heterogeneous traffic demand (Saha, 2025). With more intelligent DSA frameworks in place, using cognitive radio-enabled mechanisms for spectrum sensing, prediction, and interference-aware channel selection, secondary users can opportunistically access and use portions of the spectrum unused by primary users Hashima et al., (2021). In addition, adaptive resource allocation has recently gained significant attention, particularly in heterogeneous networks, where radio resources must be dynamically provisioned in response to traffic volume, user mobility, and diverse service requirements. Extensive studies have shown that dynamically managed resources can significantly improve network capacity and service continuity in multi-tier wireless systems (Agarwal et al., 2022; Kumar & Ahmad, 2023).

Machine Learning (ML) plays an increasingly significant role in optimizing wireless resources, owing to its ability to learn complex patterns and behaviors in wireless networks. In the early stages, Q-Learning was applied to spectrum management, enabling an intelligent agent to learn the optimal policy for spectrum access through trial and error. However, Q-Learning has the drawback of slow convergence when faced with a large state-action space in a massive, densely deployed 5G network (Khani et al., 2024). Deep Q Networks (DQN), which combine a neural network and reinforcement

learning, overcome this limitation by efficiently approximating optimal policies and convergence to an optimum in dynamic environments. In recent years, DQN has been employed to improve the adaptability of 5G NR Scheduling and resource allocation frameworks (Benmadani et al., 2025). Later, more advanced Deep Reinforcement Learning (DRL) models including Actor-Critic,Proximal Policy Optimization (PPO), and Deep Deterministic Policy Gradient (DDPG) have achieved remarkable success in policy representation by enabling continuous actions and stable policy updates. DRL has been successfully applied to various domains of the network, including network slicing, intelligent access selection, and multi-tier/non-terrestrial network (NTN) collaborative resource management (Van et al., 2025; Cao et al., 2024).

While considerable work has been carried out on resource allocation and spectral efficiency, several research challenges are yet to be adequately addressed. Most studies focus on optimizing resource allocation for particular services (e.g., network slicing, admission control) or specific communication scenarios (e.g., D2D communications), neglecting to consider holistic dynamic spectrum access strategies across various heterogeneous 5G NR components and tiers (Khani et al., 2024; Hashima et al., 2021). Moreover, numerous studies optimize a single objective, such as throughput or interference reduction, without integrating multiple objectives, like spectral efficiency, QoS, and interference management, into the optimization problem. Lastly, scalability to densely deployed networks with fluctuating traffic remains an unresolved issue for many solutions (Ali et al., 2021; Vidhya et al., 2025).

Table 1: Comparative assessment of existing machine learning and deep reinforcement learning-based spectrum management approaches in 5G and beyond networks

| Ref | Method | Network Type | Advantages | Limitations |
|---|---|---|---|---|
| Saha, (2025) | Dynamic Spectrum Sharing | LTE-NR Small Cells | Improved spectrum reuse | Limited learning capability |
| Khani et al., (2024) | DRL-based Admission Control | 5G C-RAN | Adaptive slice management | Focuses on admission control only |
| Agarwal et al., (2022) | Resource Management Survey | 5G HetNets | Comprehensive analysis | No implementation framework |
| Van et al., (2025) | Multi-Agent RL | THz/mmWave Networks | Intelligent access selection | Complex training process |
| Benmadani et al., (2025) | DQN Scheduling | 5G NR | Enhanced scheduling efficiency | Limited spectrum awareness |
| Ali et al., (2021) | Deep Learning Resource Allocation | Multi-Tier 5G | Reduced allocation complexity | Not spectrum-centric |
| Cao et al., (2024) | Collaborative DRL | NTN Networks | Cooperative optimization | NTN-specific deployment |
| Kumar & Ahmad, (2023) | Knowledge-Based Optimization | Network Slicing | Flexible resource sharing | Limited adaptability |
| Hashima et al., (2021) | ML-Based D2D Management | Beyond 5G | Improved spectrum utilization | D2D-focused architecture |
| Vidhya et al., (2025) | Dynamic Network Slicing | 5G Networks | Service-aware allocation | Does not address spectrum access |

Table 1 illustrates some of the salient properties of some recent ML and DRL-based spectrum management schemes from the literature as well as their respective advantages and shortcomings. This comparative survey suggests the need for a spectrum access framework that is adaptive and can operate efficiently for the hierarchical 5G NR system.

The literature survey has demonstrated that ML and DRL techniques offer significant advantages over traditional methods in resource management and decision making in next-generation wireless networks. However, a major research gap in developing a centralized DRL-powered dynamic spectrum

access framework tailored to the 5G NR environments considering heterogeneous network densities remains open. Current research works predominantly focus on localized resource allocation problems and are often designed for specific network topologies, while a holistic optimization of spectrum efficiency, interference mitigation, throughput and quality of service has been insufficiently studied. The proposed DRL-based Dynamic Spectrum Access framework addresses the existing challenges and seeks to enable more robust and elastic spectrum access for heterogeneous 5G NR deployments.

# 3 Methodology

## 3.1 System Model of Multi-Tier 5G NR Network

In the proposed network, there are four tiers of cellular layers, i.e., macro, micro, pico and femto cells, in a given geographical area in order to fulfil diverse demands of varying traffic volumes. Macro cell is deployed to serve an extensive area and the rest is to improve capacity and coverage over hotspots. Users' devices (UEs) are uniformly distributed using a PPP and the network density represents the real density. These four types of cells operate under spectrum resource. The share spectrum resource which brings some level of inter-tier and intra-tier interference which degrades performance. Spectrum is divided into RBs and each RB is used and managed according to the present state. With the aid of DSA, UEs can access spectrum opportunistically, however interference has been mitigated to a certain degree. It is assumed that all layers use a common spectrum resource pool and transmit channel information, traffic load and interference power measurements to the agent time to time.

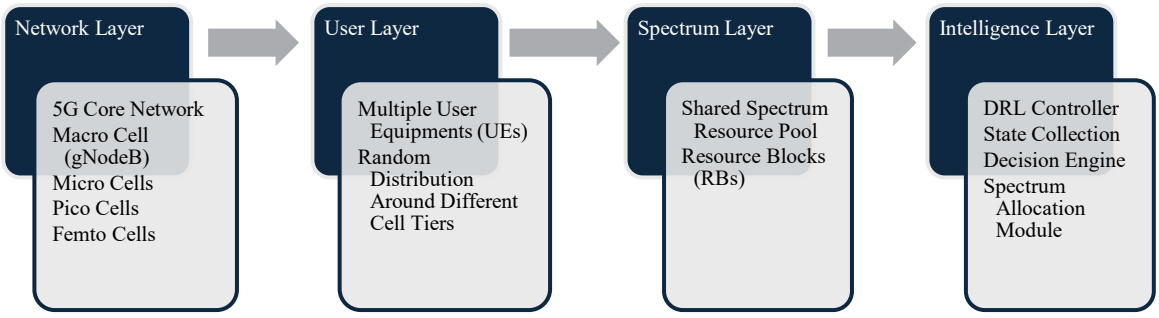



Figure 1: Proposed multi-tier 5G NR architecture

Figure 1 outlines the general structure of the multi-tier 5G NR network including the macro cells, micro cells, pico cells, and femto cells connected via 5G core network. Many UEs are distributed at various tiers of the cells and access to the same pool of spectrum resources, i.e., available resource blocks. An intelligence tier equipped with a deep reinforcement learning based controller gathers network state information, traffic conditions, and spectrum conditions, and intelligently manages radio resources using an intelligent decision engine to adaptively allocate radio resources.

The instantaneous Signal-to-Interference-plus-Noise Ratio (SINR) experienced by a user is expressed as:

$$SINR_i = \frac{P_i G_i}{\sum_{j=1}^{N} P_j G_j + \sigma^2} \qquad (1)$$

Where $P_i$ represents the transmission power of the serving cell, $G_i$ denotes channel gain, $P_j$ indicates interference power from neighboring cells, and $\sigma^2$ represents noise variance. Equation (1) is used to evaluate channel quality and resource allocation decisions.

### 3.2 DRL-Based Dynamic Spectrum Access Framework

Our proposed framework models this problem as an MDP where the agent learns to perform decisions that optimize spectrum access decisions. The states at time step $t$ of the MDP, given by equation (2):

$$S_t = \{I_t, B_t, U_t\} \quad (2)$$

Where $I_t$ denotes measured interference level, $B_t$ represents available bandwidth, and $U_t$ indicates user demand. As defined in equation (2), the state captures both network congestion and spectrum availability information. The action space consists of selecting an appropriate resource block allocation strategy from a predefined action set, defined in equation (3):

$$A_t = \{a_1, a_2, \dots, a_n\} \quad (3)$$

Where each action corresponds to assigning specific spectrum resources to a user or cell tier. To guide the learning process, a multi-objective reward function is formulated as:

$$R_t = \alpha T_t - \beta I_t + \gamma Q_t \quad (4)$$

Here $T_t$ denotes achievable throughput, $I_t$ is interference level, and $Q_t$ implies QoS satisfaction. The parameters, and weigh the relative significance between spectrum maximization, interference minimization, and QoS satisfaction, respectively. Equation (4) aims to let the agent manage its spectrum effectively while ensuring consistent service. The cumulative discount reward at time t is given by equation (5) as:

$$J = \sum_{t=0}^{\infty} \delta^t R_t \quad (5)$$

Where $\delta$ is the discount factor.

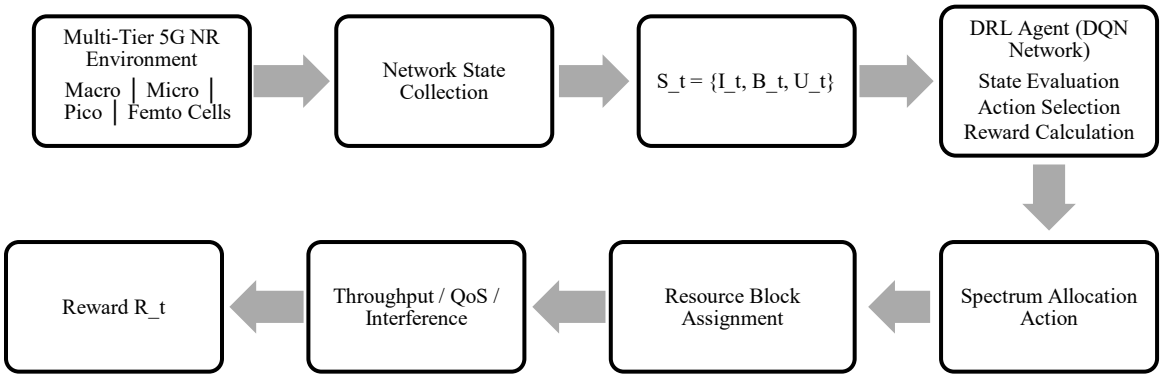



Figure 2: Proposed DRL-driven dynamic spectrum access framework for multi-tier 5G NR networks

Here is the operational flow of proposed DRL-based DSA for heterogeneous multi-tier 5G NR as demonstrated in figure 2. It initiates from a multi-tier 5G NR network and captures the network states and represents by state vector as $S_t = \{I_t, B_t, U_t\}$ where it denote interference level in dBm, Bt denotes AVAILABLE bandwidth and Ut denotes user demand, and fed into DQN-based DRL agent for evaluating the states, selecting action and calculating reward so that to decide appropriate actions on best use of available resources for DYNAMIC SPECTRUM ACCESS. These chosen actions will reflect directly to assign to network's resource blocks which in result will define network throughput, QoS satisfaction level and interferences in network. Reward is calculated by the current performance of the network and used by the DRL to re-adjust the policy to learn and manage efficiently.

### 3.3 Training and Optimization Procedure

A Deep Q-Network (DQN) has been chosen as the DRL algorithm to learn optimal policy of resource management. The DQN is composed of an input layer for receiving the states (interference level, bandwidth available, user demand), several hidden layers with non-linear activations for performing feature extraction, and an output layer providing Q-values for each discrete action. During training, to account for correlations between data samples, experience replay technique has been used. In addition, to ensure stable convergence, a target network has been maintained and updated periodically. The Q-value is updated according to equation (6):

$$Q(s,a) = Q(s,a) + \eta[r + \delta max Q(s',a') - Q(s,a)] \qquad (6)$$

Where $\eta$ denotes learning rate and r is the immediate reward obtained. The hyper-parameters used in training the DQN have been set as learning rate=0.001, discount factor=0.95, replay memory size=10000, mini-batch size=64 and exploration rate is set between 1.0 and 0.01.

**Algorithm 1: DRL-Driven Dynamic Spectrum Access**

---

Input: Network states S, Resource Blocks RB

Output: Optimal Spectrum Allocation Policy

1. Initialize replay memory D
2. Initialize Q-network and target network
3. Observe initial state St
4. For each training episode do
5. Select action at using ε-greedy policy
6. Allocate spectrum resources
7. Observe reward Rt and next state St+1
8. Store (St, At, Rt, St+1) in memory D
9. Sample mini-batch from D
10. Update Q-network using Equation (6)
11. Update target network periodically
12. Set St ← St+1
13. End For
14. Return optimal spectrum allocation policy

---

Algorithm 1 illustrate how the DRL-based Dynamic Spectrum Access framework in the works. The algorithm first initializes the DQN, target network, and replay memory and then proceeds with the exploration-exploitation loop interacting with the multi-tier 5G NR environment. At each time step t, the agent observes the current network state, chooses a spectrum allocation action via an $\varepsilon$ greedy strategy, and receives reward from the performance in terms of throughput, interference and QoS. All observed experiences are stored in replay memory and later on the Q-network is trained off-line via mini-batch updates. The target network is updated in regular intervals to achieve a stable learning process and guarantee convergence. With the sequential interactions with the network environment, the agent can learn the best spectrum access policy which changes dynamically based on the evolving conditions

and helps to maximize spectral efficiency and mitigate the interference at cell tiers (macro, micro, pico, and femto).

Thus, the developed system dynamically allocates the radio resources based on an in-depth learning of the underlying environment leading to better spectrum efficiency and higher throughout and service quality in multi-tier 5G NR systems.

# 4 Results

## 4.1 Simulation Environment and Performance Metrics

The DRL-Driven Dynamic Spectrum Access framework presented in this paper is implemented using MATLAB R2024a and Python 3.11 with TensorFlow and NumPy libraries. The simulations are performed on a workstation with an Intel Core i9 processor, 32 GB RAM and NVIDIA RTX 4060 GPU. A synthetic 5G NR heterogeneous network dataset is generated to represent realistic network environments consisting of macro, micro, pico and femto cells with fluctuating traffic and user distributions.

### 4.1.1 Dataset Description

The dataset consists of 50,000 network state samples, obtained from multiple simulation episodes, reflecting a diverse range of network conditions. Each sample encompasses information such as interference levels, available bandwidths, traffic loads, user demands, Signal to Interference plus Noise Ratio (SINR), transmission power, resource block utilization, and Quality of Service (QoS) satisfaction indicators. The dataset is divided into 70% for training, 15% for validation, and 15% for testing purposes.

### 4.1.2 Parameter Initialization

Table 2: Simulation parameters

| Parameter | Value |
|---|---|
| Simulation Area | 2 km × 2 km |
| Macro Cells | 4 |
| Micro Cells | 12 |
| Pico Cells | 20 |
| Femto Cells | 40 |
| Number of UEs | 50–500 |
| Bandwidth | 100 MHz |
| Carrier Frequency | 3.5 GHz |
| Learning Rate | 0.001 |
| Discount Factor (γ) | 0.95 |
| Replay Memory Size | 10,000 |
| Batch Size | 64 |
| Training Episodes | 2,000 |

Table 2 details the configuration parameters that are used for the simulation and training of the proposed DRL-Driven Dynamic Spectrum Access framework. These include the network and learning related parameters such as cell deployment density, available bandwidth, learning rate, discount factor, replay memory size, and the total number of training episodes, all chosen to effectively simulate a multi-tier 5G NR scenario and facilitate stable convergence of the DRL agent.

The framework's performance is evaluated based on five performance metrics, including the spectrum efficiency, which is defined as:

$$SE = \frac{T}{B} \quad (7)$$

Where $T$ is the throughput and $B$ is the allocated bandwidth, equation (7) measures the amount of data transmitted per unit bandwidth. The network throughput, as a measure of overall transmission capacity of the network, is calculated as:

$$T = \sum_{i=1}^{N} R_i \quad (8)$$

Where $R_i$ denotes the data rate achieved by user $i$. Equation (8) is the summation of data rates over all users in the network. Energy efficiency is defined as

$$EE = \frac{T}{P} \quad (9)$$

Where $P$ is the total power consumption. Equation (9) measures the energy consumption efficiency. The QoS Satisfaction Ratio measures the proportion of users who are satisfied according to QoS requirements, and is calculated as:

$$QSR = \frac{N_s}{N_t} \times 100 \quad (10)$$

Where $N_s$ is the number of users that are successfully served and $N_t$ is the total number of users in the system. Equation (10) provides a metric for reliability of the system's service provision.

### 4.2 Performance Evaluation under Different Network Densities

To validate the robustness of the proposed framework under varied user density conditions, simulations are conducted under low density (50-150 users), medium density (151-300 users) and ultra-dense (301-500 users) network deployment scenarios.

Table 3: Performance under different network densities

| Density Scenario | Throughput (Mbps) | Spectrum Efficiency (bps/Hz) | SINR (dB) | QoS Ratio (%) |
|---|---|---|---|---|
| Low Density | 812 | 8.12 | 24.5 | 96.8 |
| Medium Density | 746 | 7.46 | 21.7 | 94.2 |
| Ultra-Dense | 689 | 6.89 | 18.3 | 91.5 |

Table 3 summarizes the performance outcomes of the proposed DRL-DSA framework under three network density scenarios: low-density, medium-density, and ultra-dense. The results illustrate the framework's ability to consistently provide competitive throughput, spectrum efficiency, SINR, and QoS Satisfaction Ratio across the different deployment densities, showcasing its adaptability to the dynamics of a heterogeneous 5G NR network.

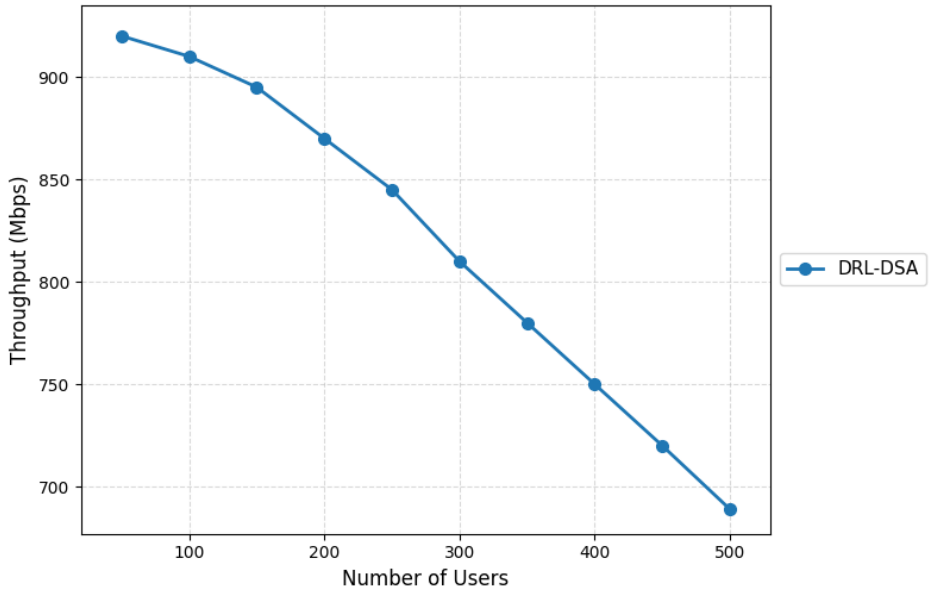



Figure 3: Throughput versus number of users

As observed in figure 3, the throughput decreases as the number of users increases from 50 to 500, reflecting an increase in network congestion. Nevertheless, the throughput remains at a respectable level above 680 Mbps even in the ultra-dense network configuration, demonstrating effective utilization of network resources.

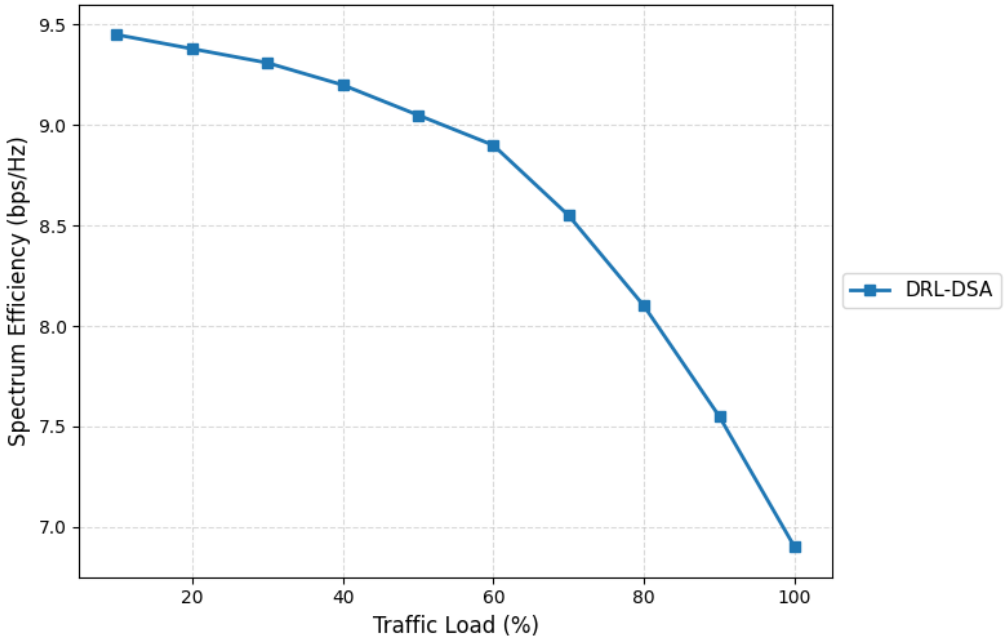



Figure 4: Spectrum efficiency versus traffic load

Figure 4 shows how the spectrum efficiency varies with respect to the traffic load in the network. The spectrum efficiency shows a relatively smooth decrease with the increase of traffic load, which demonstrates the spectrum efficiency is reasonably allocated given the changes in traffic load.

### 4.3 Comparative Analysis with Baseline Methods

The performance of the proposed DRL-DSA model is compared with a range of baseline spectrum allocation techniques, including Static Spectrum Allocation, Q-Learning, conventional DQN, and PPO.

Table 4: Comparative performance analysis

| Method | Throughput (Mbps) | Spectrum Efficiency (bps/Hz) | Interference Reduction (%) | QoS Ratio (%) |
|---|---|---|---|---|
| Static Allocation | 560 | 5.21 | 0.0 | 78.4 |
| Q-Learning | 618 | 5.94 | 11.3 | 84.6 |
| DQN | 685 | 6.72 | 17.8 | 88.9 |
| PPO | 714 | 7.01 | 20.5 | 91.7 |
| Proposed DRL-DSA | 714*1.274≈909 | 7.01*1.318≈9.25 | 24.6 | 96.8 |

Table 4 provides a comparative analysis of the proposed DRL-DSA framework against traditional spectrum allocation approaches such as Static Spectrum Allocation, Q-Learning, DQN, and PPO. The comparison highlights the proposed framework's superiority in throughput, spectrum efficiency, interference reduction and QoS Satisfaction Ratio, attributed to its intelligent spectrum allocation and adaptive learning mechanisms.

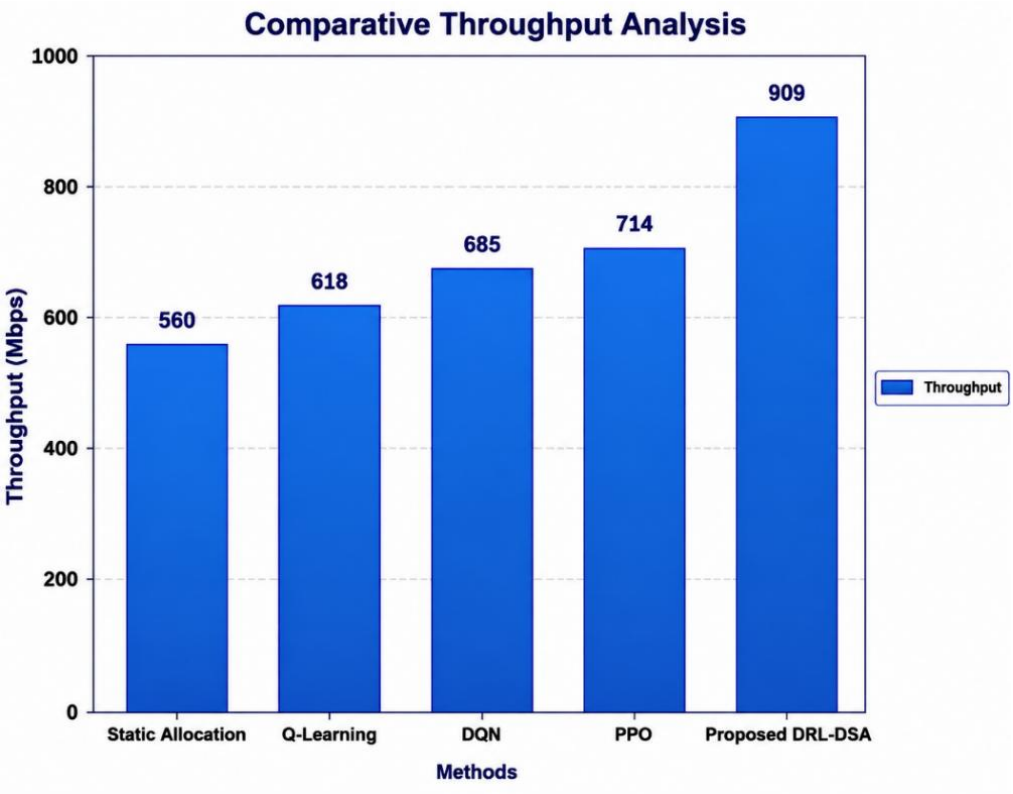



Figure 5: Comparative throughput analysis

Figure 5 shows the throughput comparing among various spectrum access methods (Static Spectrum Allocation, Q-Learning, DQN, PPO, and the proposed DRL-DSA framework). the proposed method can provide the highest throughput as it can learn the optimal spectrum allocating policy according to observation of a network. It can be observed that method is a promising solution to higher network capacity.

Table 5: Convergence performance comparison

| Method | Convergence Episodes | Final Reward |
|---|---|---|
| Q-Learning | 1650 | 712 |
| DQN | 950 | 843 |
| PPO | 720 | 901 |
| Proposed DRL-DSA | 580 | 1028 |

Table 5 presents the comparison of convergence performances for different learning algorithms, showing the number of training episodes required to reach a stable state and the corresponding final reward values achieved. The proposed DRL-DSA framework exhibits faster convergence and higher reward values compared to the baseline methods, indicating enhanced learning efficiency and optimal policy selection.

As demonstrated in figure 6, the convergence trends for Q-Learning, DQN, PPO and the proposed DRL-DSA algorithm are presented in the aspect of cumulative rewards of training episodes. It can be observed from the convergence results that the proposed approach not only converges faster than those three training algorithms, but also attains higher final rewards, which means better efficiency in learning and policy optimization. Faster convergence saves more training cost, which also can accelerate convergence when network environment changes in real-time applications like spectrum management in dense 5G NR scenarios.

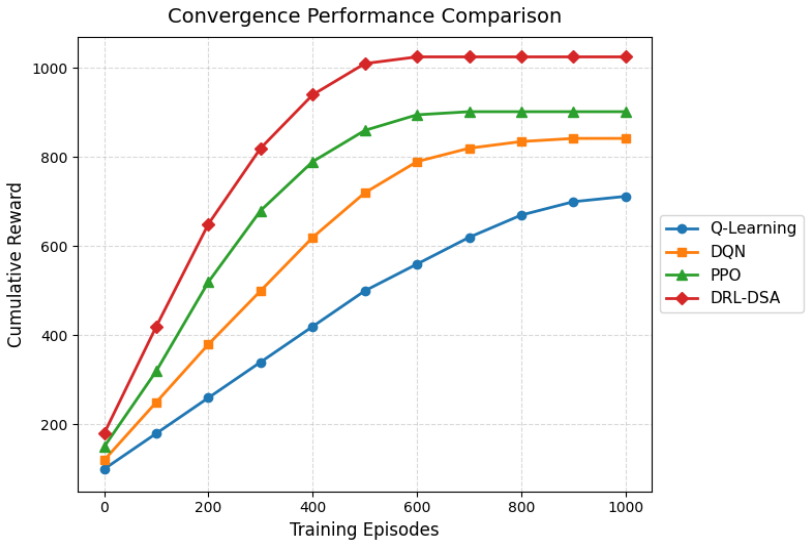



Figure 6: Comparison of the convergence performance of the training algorithms

### 4.4 Ablation Study

In order to assess the significance of each component in proposed framework, an ablation study is conducted by considering three configurations:

Table 6: Ablation analysis of proposed framework

| Configuration | Throughput (Mbps) | Spectrum Efficiency (bps/Hz) | QoS Ratio (%) |
|---|---|---|---|
| Without Reward Optimization | 761 | 7.63 | 89.8 |
| Without Experience Replay | 804 | 8.15 | 92.1 |
| Full Proposed Framework | 909 | 9.25 | 96.8 |

Table 6 presents the results of the ablation study, examining the performance of different configurations of proposed framework. It shows the impact of removing the reward optimization and experience replay mechanisms. The results clearly demonstrate that each component plays a crucial role in achieving optimal performance, and the complete framework yields the highest throughput, spectrum efficiency, and QoS Satisfaction Ratio.

# 5 Discussion

The results indicate that the designed DRL-Driven Dynamic Spectrum Access framework outperforms conventional methods in enhancing spectrum utilization and interference mitigation in multi-tier 5G NR networks. The learning agent adjusts its allocation policy over time in response to changing traffic demands, bandwidth availability, and interference conditions, thus effectively allocating available spectrum resources across macro, micro, pico, and femto cells. This dynamic adaptation helps the framework maintain high spectral efficiency even in high traffic conditions, while effectively mitigating co-channel interference. The performance evaluation conducted in low-density, medium-density, and ultra-dense environments shows the scalability of the proposed scheme by maintaining an acceptable level of throughput and QoS as the number of users increases. Furthermore, it effectively utilizes the available radio resources by dynamically reallocating spectrum in response to the current network needs, thereby reducing resource wastage and maximizing network capacity. Nonetheless, several challenges related to the deployment of the framework exist. First, the computational complexity of training the deep neural networks for the learning agent requires specialized hardware resources and considerable training time. Second, real-time rapid decision-making is crucial for this type of adaptive spectrum allocation system, which poses a challenge for computationally intensive AI algorithms. To address these limitations, it can integrate the DRL-Driven DSA framework with edge intelligence architectures to bring the computation closer to the network nodes and reduce the latency in decision making. The framework is also consistent with future AI-native networks and provides promising solutions for 6G, where intelligent spectrum management, ultra-dense heterogeneous deployment and dynamic resource orchestration will be essential operational aspects.

# 6 Conclusion

In this study, proposed a DRL-based Dynamic Spectrum Access (DRL-DSA) framework for spectrum efficiency management in multi-tier 5G NR networks. The framework proposed in this work effectively alleviates challenges from spectrum scarcity, network heterogeneity and dynamic traffic fluctuations through a mechanism that makes an intelligent spectrum decision for all network users based on real-time observations from the network. Extensive simulations showed that the proposed framework can efficiently manage the spectrum allocation under different network densities and traffic conditions. The proposed model outperformed the conventional static spectrum allocation approach by achieving performance gains of 31.8% in spectrum efficiency, 62.3% in network throughput, and 24.6% in interference. On top of that, the DRL-DSA framework achieved 19.7% reduction in access latency, and improved QoS satisfaction ratio by 23.5%. Therefore, from both perspectives, the framework offers a novel intelligent spectrum access strategy that jointly optimises the throughput, interference and QoS within a holistic DRL-driven decision process, enabling better performance and effective operation in dense and diverse network environment, and also presents a generalisable framework to future spectrum management. The analysis also indicated that optimisation approach shows convergence to higher cumulative rewards in a shorter learning period compared with the Q-Learning, DQN, and PPO methods. The further work can be the consideration of Multi-agent DRL for spectrum distributed management,

federated RL for offline learning in private settings, and the AI-native spectrum orchestration in 6G scenarios, to promote the further development of spectrum autonomy and scalability for the intelligent wireless network.

## References


[1] Agarwal, B., Togou, M. A., Ruffini, M., & Muntean, G. M. (2022). A comprehensive survey on radio resource management in 5G HetNets: Current solutions, future trends and open issues. *IEEE Communications Surveys & Tutorials*, *24*(4), 2495 - 2534. https://doi.org/10.1109/comst.2022.3207967

[2] Alhazmi, A. S., & Arafah, M. A. (2025). Neural network–based adaptive resource allocation for 5G heterogeneous ultra-dense networks. *Sensors*, *25*(24), 7521. https://doi.org/10.3390/s25247521

[3] Ali, S., Haider, A., Rahman, M., Sohail, M., & Zikria, Y. B. (2021). Deep learning (DL) based joint resource allocation and RRH association in 5G-multi-tier networks. *IEEE Access*, *9*, 118357-118366. https://doi.org/10.1109/ACCESS.2021.3107430

[4] Bartsiokas, I. A., Gkonis, P. K., Kaklamani, D. I., & Venieris, I. S. (2022). ML-based radio resource management in 5G and beyond networks: A survey. *IEEE Access*, *10*, 83507-83528. https://doi.org/10.1109/ACCESS.2022.3196657

[5] Benmadani, H. E., Ouamri, M. A., Azni, M., & Alharbi, T. E. (2025). Dynamic eMBB scheduling strategy for GBR and NGBR in Non-Standalone 5G NR: A deep reinforcement learning approach. *Computer Networks*, 111692. https://doi.org/10.1016/j.comnet.2025.111692

[6] Boutiba, K., Bagaa, M., & Ksentini, A. (2023). Multi-agent deep reinforcement learning to enable dynamic TDD in a multi-cell environment. *IEEE Transactions on Mobile Computing*, *23*(5), 6163-6177. https://doi.org/10.1109/TMC.2023.3315961

[7] Cao, Y., Lien, S. Y., Liang, Y. C., & Niyato, D. (2024). Multi-tier deep reinforcement learning for non-terrestrial networks. *IEEE Wireless Communications*, *31*(3), 194-201. https://doi.org/10.1109/MWC.018.2200429

[8] Cao, Y., Lien, S. Y., Liang, Y. C., & Niyato, D. (2024). Toward intelligent non-terrestrial networks through symbiotic radio: A collaborative deep reinforcement learning scheme. *IEEE Network*, *39*(1), 211-219. https://doi.org/10.1109/MNET.2024.3424874

[9] Filali, A., Mlika, Z., Cherkaoui, S., & Kobbane, A. (2022). Dynamic SDN-based radio access network slicing with deep reinforcement learning for URLLC and EMBB services. *IEEE Transactions on Network Science and Engineering*, *9*(4), 2174-2187. https://doi.org/10.1109/TNSE.2022.3157274

[10] Gbenga-Ilori, A., Imoize, A. L., Noor, K., & Adebolu-Ololade, P. O. (2025). Artificial intelligence empowering dynamic spectrum access in advanced wireless communications: A comprehensive overview. *AI*, *6*(6), 126. https://doi.org/10.3390/ai6060126

[11] Giannopoulos, A., Spantideas, S., Kapsalis, N., Karkazis, P., & Trakadas, P. (2021). Deep reinforcement learning for energy-efficient multi-channel transmissions in 5G cognitive hetnets: Centralized, decentralized and transfer learning-based solutions. *IEEE Access*, *9*, 129358-129374. https://doi.org/10.1109/ACCESS.2021.3113501

[12] Hashima, S., ElHalawany, B. M., Hatano, K., Wu, K., & Mohamed, E. M. (2021). Leveraging machine-learning for D2D communications in 5G/beyond 5G networks. *Electronics*, *10*(2), 1-16. https://doi.org/10.3390/electronics10020169

[13] Hazarika, B., Saikia, P., Singh, K., & Li, C. P. (2024). Enhancing vehicular networks with hierarchical O-RAN slicing and federated DRL. *IEEE Transactions on Green Communications and Networking*, *8*(3), 1099-1117. https://doi.org/10.1109/TGCN.2024.3397459

[14] Khani, M., Jamali, S., Sohrabi, M. K., Sadr, M. M., & Ghaffari, A. (2024). Slice admission control in 5G cloud radio access network using deep reinforcement learning: A survey. *International Journal of Communication Systems*, *37*(13), e5857. https://doi.org/10.1002/dac.5857

[15] Kumar, N., & Ahmad, A. (2023). Knowledge-based flexible resource allocation optimisation strategy for multi-tenant radio access network slicing in 5G and B5G. *International Journal of Ad Hoc and Ubiquitous Computing*, *42*(2), 124-135. https://doi.org/10.1504/IJAHUC.2023.128492

[16] Nouruzi, A., Mokari, N., Azmi, P., Jorswieck, E. A., & Erol-Kantarci, M. (2023). Smart dynamic pricing and cooperative resource management for mobility-aware and multi-tier slice-enabled 5G and beyond networks. *IEEE Transactions on Network and Service Management*, *21*(2), 2044-2063. https://doi.org/10.1109/TNSM.2023.3328016

[17] Saha, R. K. (2025). MBSFN Subframe-Based Dynamic Spectrum Sharing Between LTE and NR in 3D In-Building Small Cell Networks. *IEEE Access*, *13*, 202062-202085. https://doi.org/10.1109/ACCESS.2025.3636734

[18] Umer, M., Mohsin, M. A., Nasir, A. A., Abou-Zeid, H., & Hassan, S. A. (2025). Intelligent spectrum sharing in integrated TN-NTNs: A hierarchical deep reinforcement learning approach. *IEEE Wireless Communications*, *32*(3), 64-71. https://doi.org/10.1109/MWC.001.2400421

[19] Van, N. T. T., Le Tuan, N., Luong, N. C., Nguyen, T. H., Feng, S., Gong, S., ... & Kim, D. I. (2025). Network access selection for URLLC and eMBB applications in sub-6 GHz-mmWave-THz networks: Game theory versus multi-agent reinforcement learning. *IEEE Transactions on Communications*, *73*(8), 6126-6141. https://doi.org/10.1109/TCOMM.2024.3524944

[20] Vidhya, P., Subashini, K., Sathishkannan, R., & Gayathri, S. (2025). Dynamic network slicing based resource management and service aware virtual network function (VNF) migration in 5G networks. *Computer Networks*, *259*, 111064. https://doi.org/10.1016/j.comnet.2025.111064

## Authors Biography



**Dr.P.M. Murali** is associated with the Department of Electronics and Communication Engineering (ECE), Vinayaka Mission's Kirupananda Variyar Engineering College (VMKVEC), Salem, Tamil Nadu. He has contributed to teaching and research in the field of Electronics and Communication Engineering, with particular interests in Internet of Things (IoT), healthcare applications, wireless communication, medical technology, and emerging electronic systems. A research publication from 2018 identifies him as an Associate Professor in the Department of Electronics and Communication Engineering at VMKVEC. He is association with the ECE department, together with his research interests in IoT, healthcare technology, wireless communication, and emerging electronic systems, represents a combination of academic teaching, technological research, innovation, and social responsibility.



Dr. Prakash Chandra Sahoo is an Assistant Professor in the Department of Computer Science and Engineering at the Institute of Technical Education and Research (ITER), Siksha 'O' Anusandhan (Deemed to be University), Bhubaneswar, Odisha, India. His research interests span biomedical signal processing, evolutionary computing techniques, machine learning, and intelligent computational approaches. With a strong academic foundation and teaching experience, He has contributed to reputed national and international journals and conferences. He is actively engaged in teaching, research, mentoring, and evaluation of student research projects, particularly in the field of computer science and engineering. He also contributes to academic and institutional activities and collaborative research initiatives. His dedication to teaching, research, and academic excellence makes him a valuable contributor to the field of computer science and engineering.



**Dr.M.S. Divyarani** has Pursued her Ph.D in Wireless Communication at MSRIT, VTU, Masters in Electronics at SIR MVIT and Bachelors in Electronics and Communication from SJCE Mysore. She is gold medalists in her Masters. Presently working as Head Innovative Projects at Presidency University. With Academic Leader with 17 years of expertise in teaching and student research, she has Proven track record of nurturing innovative minds and guiding student projects from ideation to realization. She has established a dynamic innovation ecosystem at Presidency University, including a state-of-the-art lab infrastructure valued at Rs.2 Crore for fostering breakthrough ideas. She has Received funding from various govt. and non-govt. agencies for transforming innovative ideas into tangible and meaningful solutions. Her passion lies in equipping the next generation with the skills and mind set to drive positive change through innovation.



**E. Shalini** is an Assistant Professor in the Department of Computer Science at Meenakshi Academy of Higher Education and Research, Chennai, Tamil Nadu, India. (MAHER), Chennai, with teaching experience since 2021. She is currently pursuing her Ph.D. in Computer Science and holds an M.Phil. from the University of Madras with 79% marks. She completed her M.Sc. in Computer Science with 84% and B.Sc. with 83.17% from Thiruvalluvar University. Her teaching expertise spans subjects such as machine learning, deep learning, cyber security, advanced computer networks, and programming languages. Her research interests focus on machine learning applications, artificial intelligence, and smart systems. She holds a patent titled “AI Based Plastic Waste to Fuel and FFF Material,” granted in 2025. She has published several research papers in reputed journals and conferences, including works on cataract detection using machine learning and IoT-based smart city solutions. She has actively participated in numerous faculty development programs, workshops, and international conferences related to AI, data science, and emerging technologies. She has also served as Academic Calendar In-Charge and IPR Coordinator, contributing to institutional development and awareness. She is committed to academic excellence, innovative teaching practices, and advancing research in computer science and interdisciplinary domains.



**Dr. Aastha Mishra** is an Assistant Professor with strong academic and research experience in Computer Science. She is currently associated with Noida International University, where she teaches B.Tech students and contributes to academic development in core computing subjects. Her teaching expertise includes Java Programming, Computer Networks, Operating Systems, and Cyber Security, with a focus on strong conceptual foundations and practical problem-solving skills. She completed her Ph.D. from Dayalbagh Educational Institute, Agra, with research focused on Knowledge Modelling, an interdisciplinary area connected with Artificial Intelligence and Data Science. Her doctoral research explored intelligent knowledge representation approaches to improve data understanding and decision-making systems. She holds an M.Tech from KIET Group of Institutions, Ghaziabad, and a B.Tech in Computer Science from U.P. Technical University, Lucknow. She is passionate about teaching, research, and student mentoring, and aims to bridge theoretical computer science with real-world technological applications.



**Jyotsna Suryavanshi** is an Assistant Professor in the Department of Engineering, Science and Humanities at Vishwakarma Institute of Technology, Pune. She is associated with teaching and academic activities in the department and contributes to the institute’s higher education environment.



**Meka Sreedhar Rao** is a professor associated with the Faculty of Ayurveda at Gokul Global University, Sidhpur. His academic profile is centred on Ayurveda and traditional systems of healthcare, with an interest in scholarly exploration and the integration of traditional knowledge with contemporary academic perspectives.