ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/25sep051

AI/ML Driven 5G Networks: From Deployment to Optimization

Manish Kumar[1]

[1]AMD, India

Publication Date: 2025/09/10

Abstract: Fifth-generation (5G) mobile networks introduce unprecedented performance requirements including enhanced mobile broadband (eMBB), ultra-reliable low-latency communications (URLLC), and massive machine-type communications (mMTC). These requirements demand highly adaptive, automated, and intelligent management. Artificial Intelligence (AI) and Machine Learning (ML) have emerged as essential enablers for 5G system deployment, installation, configuration, and optimization. This paper explores the integration of AI/ML with 3GPP-defined network functions and management frameworks, focusing on specifications 3GPP TS 28.104, 3GPP TS 28.105, and 3GPP TR 29.908, while also analyzing how AI-driven analytics enhance deployment, coverage, sustainability, energy efficiency, and user experience.

Keywords: *5G Networks, Artificial Intelligence (AI), Machine Learning (ML), Network Data Analytics Function (NWDAF), 3GPP Standards, Network Slicing, Coverage Optimization, Traffic and Mobility Optimization, Energy Efficiency*

How to Cite: Manish Kumar (2025) AI/ML Driven 5G Networks: From Deployment to Optimization. *International Journal of Innovative Science and Research Technology*, 10(9), 162-168.
https://doi.org/10.38124/ijisrt/25sep051

I. INTRODUCTION

Rolling out 5G networks brings forth issues related to intricacy, diversity, and expansion potential. Conventional oversight methods fall short in environments demanding instantaneous adjustments. AI/ML approaches foster decisions rooted in data, anticipatory capabilities, and independent functioning within 5G frameworks, aiding elements like the Network Data Analytics Function (NWDAF), coordination, and feedback-driven automation.

3GPP has formalized AI/ML handling in its 28-series documents, outlining designs and operational supports for embedding AI in 5G setups. Concurrently, scholarly investigations showcase AI/ML's role in refining radio asset handling, flow management, and segment coordination.

Incorporating AI/ML transcends mere improvements; it is vital for overseeing the magnitude of 5G implementations. With countless linked devices and applications requiring top-tier performance, mechanization is essential to lessen functional burdens and promote economic viability. This document reviews the present deployment landscape, details pertinent AI/ML methods for 5G, elaborates on 3GPP guidelines, and explores AI/ML uses in actual setups to foster superior installation, arrangement, coverage, and rollout, yielding enhanced user offerings.

II. CURRENT 5G DEPLOYMENT

By August 2025, 5G networks have achieved substantial worldwide progress, propelled by investments from providers and technological evolutions. Reports indicate over 354 commercial 5G networks operational globally as of March 2025, with population coverage reaching about 51% worldwide. Subscriptions to 5G are projected to surpass 2.9 billion by year-end 2025, comprising one-third of total mobile links, with 2.25 billion already recorded by March 2025 (Ericsson Mobility Report, 2025). This expansion is evident in North America, Europe, and Asia, where mid-band frequencies have attained 50% coverage in certain locales.

Primary rollout strategies encompass standalone (SA) and non-standalone (NSA) designs. NSA utilizes legacy 4G setups for quicker launches, whereas SA delivers complete 5G core advantages such as network slicing and enhanced QoS management. Fixed wireless access (FWA) has become a revenue generator, with 5G anticipated to manage 80% of mobile data by 2030. In the US, nearly half of mobile connections are expected to be 5G by 2025, amounting to around 180 million. Globally, 5G is set to encompass one-third of people by 2025, influencing sectors such as healthcare, logistics, and manufacturing through enhanced connectivity (GSMA, 2025).

Nevertheless, deployments encounter hurdles in frequency assignment, setup expenses, and energy efficiency. Providers are embracing 5G Advanced from 3GPP Release 18 to boost efficacy. Business uptake is increasing, with firms using 5G for automation, remote operations, and cybersecurity enhancements. Data volumes have escalated to 172 exabytes per month in early 2025, up 19% from the prior year.

A. Radio Network Planning and Spectrum Fragmentation

5G functions over low-, mid-, and high-band ranges, each with distinct signal behaviors. Frequency division across licensed and unlicensed bands complicates design, requiring advanced tools for forecasting signal paths, curbing disruptions, and fine-tuning site positions. High-band mmWave yields immense speeds but limited distance and penetration, while low-band offers broad reach but constrained capacity. Without intelligent planning, reconciling these is challenging (Wang et al., 2020)

B. Configuration Management in Heterogeneous Networks

5G's mixed nature includes macro cells, small cells, massive MIMO, and Open RAN components. Rule-centric setups inadequately handle varied gear from assorted suppliers. Manual adjustments of myriad settings are laborious and mistake-prone. Automated self-configuration and optimization are necessary to minimize delays and boost efficiency (Pradhan et al., 2023).

C. Dynamic Traffic and Mobility Management

Urban zones face erratic data spikes from gatherings, IoT proliferation, and user concentrations. Static methods cannot cope. Predictive analytics and mobility-aware optimization enable proactive load balancing, smooth handovers, and adaptive mobility strategies. By discerning patterns, intelligent systems reduce disconnections in fast scenarios such as rail or roads (Agbon et al., 2024).

D. Continuous Software Updates and Integration

Virtualized (VNFs) and containerized (CNFs) elements demand regular refreshes. Unlike older systems, 5G requires phased updates to maintain uptime. Manual checks pose risks. Automated deployment systems scrutinize data, spot irregularities, and propose fixes for uninterrupted service, enabling hands-free provisioning and rollbacks (Panek et al., 2023).

E. Configuration Management in Heterogeneous Networks

5G's distributed structure amplifies threats across cloud, edge, and core. Privacy issues arise from data-driven optimization needs. Static defenses are insufficient. Intrusion detection powered by analytics and privacy-preserving approaches such as federated learning safeguard information and ensure compliance with regulations (Latha et al., 2024).

F. Energy Efficiency and Sustainability

Power use is a pressing concern, especially in dense deployments and massive MIMO. Optimized resource allocation techniques deactivate idle stations or tweak antenna parameters, aiding ecological targets and cost savings. With sustainability goals becoming a priority, efficiency is a key focus in both research and deployment strategies (Haidine et al., 2021).

III. AI/ML TECHNIQUES IN 5G SYSTEMS

A. Supervised Learning

Supervised learning is widely applied in 5G for traffic demand prediction, fault classification, and QoS estimation. Labeled datasets from network operations help train models such as decision trees, support vector machines (SVMs), and neural networks. For instance, supervised learning models predict user throughput and optimize spectrum allocation accordingly. Anomaly detection systems can also be trained to recognize known failure modes, reducing downtime.

B. Unsupervised Learning

Unsupervised learning is critical for tasks where labeled data is unavailable. Clustering techniques such as K-means and DBSCAN group users by mobility patterns or classify cell behaviors for anomaly detection. Dimensionality reduction methods like PCA help analyze large-scale telemetry data for network insights. For example, unsupervised models can identify unusual traffic patterns indicating potential security breaches or misconfigurations.

C. Reinforcement Learning (RL)

RL enables adaptive decision-making in dynamic 5G environments. Techniques such as Q-learning and deep reinforcement learning (DRL) optimize handover policies, power control, and radio resource allocation. RL agents learn optimal strategies by interacting with the environment and adjusting configurations to maximize reward metrics such as throughput or latency. Multi-agent RL is particularly useful in scenarios with multiple base stations coordinating resource allocation, reducing interference, and ensuring fair distribution of resources among users

D. Deep Learning and Neural Networks

Deep Neural Networks (DNNs) and Convolutional Neural Networks (CNNs) are used for channel estimation, beamforming optimization, and coverage prediction. Recurrent Neural Networks (RNNs) and Long Short-Term Memory (LSTM) models are effective in traffic forecasting and user mobility prediction, enabling proactive resource provisioning. Generative Adversarial Networks (GANs) are increasingly being used for synthetic data generation in cases where training data is scarce, enabling robust AI models for 5G scenarios

IV. 3GPP SPECIFICATIONS FOR AI/ML IN 5G

A. 3GPP TS 28.104 (Management Data Analytics - MDA)

TS 28.104 specifies the architecture and services for data analytics in 5G. NWDAF (Network Data Analytics Function) provides data collection, analysis, and distribution to other network functions (NFs). Examples include predicting network congestion, optimizing resource utilization, and delivering tailored QoE guarantees. NWDAF thus acts as a central intelligence engine for network management.

In terms of energy efficiency and sustainability, TS 28.104 enables energy-aware analytics by monitoring base station utilization and suggesting power-saving strategies. For instance, NWDAF can recommend switching small cells into low-power modes during low traffic periods, reducing unnecessary energy consumption. The specification also defines analytics for traffic load balancing, which indirectly enhances sustainability by minimizing over-provisioning.

For user experience, NWDAF supports analytics for Quality of Experience (QoE) estimation and prediction. By analyzing data such as packet loss, jitter, and latency, NWDAF allows proactive adjustments to maintain service quality for applications such as video streaming, gaming, or remote healthcare.

B. 3GPP TS 28.105 (AI/ML Management)

TS 28.105 focuses on AI/ML model lifecycle management. It defines functions for model training, validation, deployment, monitoring, and retraining. This ensures that AI/ML models in 5G systems remain accurate and adaptive in changing environments.

From an energy efficiency perspective, 28.105 supports the deployment of adaptive models that optimize network resource utilization with minimal power consumption. For example, models can be trained to predict when base stations should scale down transmit power or enter energy-saving modes. Lifecycle management ensures these models evolve with traffic changes, maximizing long-term sustainability.

In terms of user experience, 28.105 emphasizes monitoring AI/ML models that support QoS and QoE. If a mobility prediction model begins to degrade in accuracy, the specification ensures it is retrained or replaced, preventing poor service continuity. It also supports federated learning approaches, where user data remains private while still contributing to model improvement, safeguarding both performance and privacy.

C. 3GPP TR 29.908 (Study on AI/ML Management)

TR 29.908 presents a study of AI/ML management frameworks and potential use cases. It highlights scenarios such as dynamic spectrum allocation, intelligent slicing, energy-efficient resource control, and anomaly detection.

For energy efficiency and sustainability, the study outlines AI/ML-driven optimization strategies where networks dynamically adjust spectrum allocation and cell activity based on traffic demand. It emphasizes cross-domain AI coordination, ensuring that RAN, core, and transport domains collectively optimize power consumption, reducing the carbon footprint of large-scale deployments.

For user experience, TR 29.908 explores advanced use cases such as adaptive video streaming, low-latency assurance for gaming and industrial automation, and personalized services. By integrating AI/ML into network management functions, operators can ensure consistent QoE across a wide range of applications. The study also stresses the importance of explainable AI (XAI) to increase user trust in network-driven decisions.

V. APPLICATIONS OF AI/ML IN 5G DEPLOYMENT

AI/ML integration into 5G deployment, configuration, and optimization has matured beyond theoretical concepts, with live operator deployments demonstrating measurable benefits. This section elaborates on the key applications.

A. Intelligent Installation and Configuration

AI-powered planning tools automate site selection, antenna alignment, and parameter configuration. By analyzing geospatial datasets, propagation models, and historical deployment performance, these tools minimize installation time while maximizing coverage. For example, reinforcement learning agents can select optimal antenna tilt angles based on real-time feedback from drive-test data (Zhang et al., 2023).

Vodafone has adopted AI-assisted site acquisition systems that factor in traffic density, zoning regulations, and construction costs to accelerate rollout timelines (Vodafone, 2024).

Automated configuration reduces operator dependence on manual intervention, which is critical in multi-vendor Open RAN deployments.

B. Coverage Optimization

Coverage optimization is crucial in India, Africa, and dense urban environments. AI/ML models dynamically tune antenna tilt, power levels, and beamforming weights, reducing coverage holes and improving spectral efficiency.

Deep reinforcement learning (DRL) models have been applied for adaptive beamforming in massive MIMO, delivering up to 30% capacity improvements in dense deployments (Sun et al., 2022).

NTT Docomo applies AI-driven beam adjustments during high-density events (Tokyo Marathon) to maintain QoE under fluctuating traffic demand (NTT Docomo, 2024).

Ericsson reports AI-based tilt optimization has reduced dropped calls and boosted throughput by 10–15% in commercial deployments (Ericsson Mobility Report, 2025).

C. Network Slicing and Resource Management

Dynamic slice orchestration is a cornerstone of 5G monetization. AI/ML models predict slice-specific demands and allocate resources in real time.

For eMBB slices, AI predicts peak demand hours and pre-allocates bandwidth. For URLLC slices, AI optimizes edge resources to meet latency targets <1 ms (Polese et al., 2022).

Telefónica uses AI-powered slicing management in Spain and Brazil, allowing enterprises to dynamically request slices for manufacturing and automotive use cases (Telefónica, 2024).

NWDAF-defined slice QoS prediction (3GPP TS 28.104) ensures automated adaptation of slice parameters to meet SLA commitments.

D. Predictive Maintenance and Fault Management

AI/ML enables proactive monitoring by detecting anomalies in KPIs and forecasting equipment failures before they occur.

ML models trained on time-series data from base stations identify hardware degradation patterns, reducing downtime by enabling predictive repairs (Panek et al., 2023).

AT&T applies predictive fault detection in its U.S. 5G core to anticipate node failures, reducing mean-time-to-repair (MTTR) by up to 25% (AT&T, 2024).

In Open RAN, predictive maintenance ensures that software components (VNFs and CNFs) remain resilient, with AI-enabled rollbacks and self-healing minimizing service disruptions.

E. Traffic and Mobility Optimization

AI/ML enhances traffic steering by predicting congestion and proactively rerouting flows. Mobility optimization ensures seamless experiences in high-speed environments such as rail networks and highways.

Federated learning models deployed across RAN elements predict cell congestion without centralizing sensitive data, aligning with privacy requirements (Latha et al., 2024).

SK Telecom applies AI-driven handover optimization in Seoul's subway systems, reducing call drops by over 20% in fast-moving mobility environments (SK Telecom, 2024).

Jio in India is experimenting with AI-based mobility prediction for rural FWA users, ensuring stable connectivity in regions with fluctuating backhaul availability (Economic Times, 2024).

VI. DEPLOYING 5G ON AMD ARCHITECTURE

AMD's EPYC processors and Instinct GPUs, based on the Turin Zen5 architecture, are engineered for high throughput, energy efficiency, and scalability in cloud-native telecom environments. The Zen5 design introduces improvements in IPC (instructions per cycle), cache hierarchy, and memory bandwidth, which are essential for handling 5G's real-time workloads.

A. Virtualized and Containerized Network Functions

AMD EPYC processors offer large core counts and NUMA-aware performance, enabling operators to host VNFs and CNFs with minimal latency. Zen5-based EPYC CPUs support simultaneous multi-threading and extended memory bandwidth, ideal for Open RAN deployments where compute requirements are distributed across edge and central sites.

B. AI/ML Integration for NWDAF and Edge Analytics

With AMD Instinct GPUs, AI/ML workloads such as anomaly detection, traffic forecasting, and slice performance analytics can be accelerated. These accelerators enable real-time inference at the edge, reducing dependence on centralized cores and improving latency-sensitive applications such as URLLC.

C. Energy Efficiency and Sustainability

Zen5 architecture introduces advanced power management features, including per-core dynamic voltage and frequency scaling (DVFS). This reduces power draw during low-load periods, aligning with sustainability goals. Compared to older architectures, Zen5 delivers higher performance per watt, critical for operators aiming to reduce their carbon footprint while scaling their 5G infrastructure.

D. Comparative Advantage

Deploying 5G on AMD architecture provides a balance between cost efficiency and performance scalability. Studies show AMD's Zen5 architecture achieves significant TCO reductions compared to traditional Intel-based deployments, especially when combined with GPU acceleration for AI/ML integration (AMD Whitepaper, 2024).

VII. REAL-WORLD DEPLOYMENT CASE STUDIES

The practical application of AI/ML in 5G networks is already transforming the telecommunications landscape. Operators worldwide are leveraging intelligent automation, predictive analytics, and energy-efficient strategies to meet user demand and improve operational performance. Below are key real-world deployments.

A. Vodafone

Vodafone has integrated AI/ML across its European and African networks for predictive maintenance, traffic forecasting, and energy optimization. Through AI-driven spectrum allocation and cell on/off switching, Vodafone reported up to 18% energy savings in certain markets. Additionally, its collaboration with Google Cloud has enabled cloud-native AI platforms for customer experience optimization and network troubleshooting (Vodafone, 2024).

B. AT&T

AT&T applies AI for traffic prediction, anomaly detection, and handover optimization across its U.S. network. Its AI-powered Fixed Wireless Access (FWA) management reduces congestion during peak periods, enabling reliable broadband in rural and suburban areas. Moreover, AT&T has been experimenting with closed-loop automation using NWDAF, achieving faster incident resolution and improving customer QoE (AT&T, 2024).

E. NTT Docomo

In Japan, NTT Docomo leverages AI for real-time beamforming, adaptive coverage, and QoE monitoring. During high-density events such as the Tokyo Marathon, AI dynamically optimized radio beams to maintain service quality. Additionally, AI-based predictive load balancing reduces energy usage during off-peak periods, aligning with sustainability goals (NTT Docomo, 2024).

ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/25sep051

F. China Mobile

China Mobile has scaled AI integration across its 5G deployments, particularly in smart city and industrial IoT use cases. Using AI-driven optimization, China Mobile has enhanced spectrum utilization in urban hotspots while reducing operational energy consumption. In collaboration with Huawei, China Mobile launched AI-assisted RAN slicing, enabling differentiated services for industries such as manufacturing, mining, and logistics.

G. SK Telecom

SK Telecom in South Korea employs AI for end-to-end service automation. Its AI-powered "TANGO" platform predicts traffic demand and dynamically reallocates resources across radio, core, and transport layers. The operator also focuses on AI-based security monitoring, applying anomaly detection in 5G core functions to counter cyber threats.

H. Deutsche Telekom

Deutsche Telekom has deployed AI in multi-vendor Open RAN environments, where configuration management and interoperability challenges are high. AI is used to harmonize parameters across different vendor equipment and to improve self-optimizing network (SON) functions. In addition, DT applies AI analytics to reduce downtime during network upgrades by predicting integration failures in advance.

I. MTN (Africa)

MTN Group, one of Africa's largest operators, has adopted AI/ML for coverage expansion and energy efficiency across rural and peri-urban regions. AI-driven analytics optimize tower placement and power management, particularly in solar-powered base stations. MTN also applies predictive algorithms to balance data traffic across urban hotspots like Lagos and Johannesburg, reducing service outages and improving user QoE (MTN, 2024).

J. Telefonica (Europe & Latin America)

Telefónica has embedded AI into its UNICA Next cloud-native platform to manage virtualized and containerized 5G functions. In Spain and Brazil, Telefónica uses AI to automate slice management for enterprise services, ensuring differentiated QoS for verticals such as automotive and manufacturing. AI-enabled energy-saving initiatives have resulted in measurable reductions in CO_2 emissions across its European operations (Telefonica, 2024).

K. America Movil (Latin America)

América Móvil, the parent company of Claro, has employed AI to optimize spectrum allocation and manage rural connectivity projects across Latin America. In Mexico and Brazil, AI-based traffic prediction enhances FWA performance, providing broadband access to underserved regions. Additionally, América Móvil applies AI-enabled cybersecurity tools to monitor 5G core traffic for anomalies and prevent fraud in mobile financial services (America Movil, 2024).

L. Global Trends

Automation of 5G Core Functions: Operators increasingly deploy NWDAF to support closed-loop automation and real-time analytics.

- *Energy-Efficient AI:* AI-based cell sleep modes and antenna optimization reduce carbon footprints.
- AI for Private 5G: Enterprises deploying private 5G networks use AI for tailored QoS, predictive maintenance, and secure slicing.
- AI at the Edge: Operators globally are shifting towards edge-based AI inference for ultra-low-latency applications like AR/VR and autonomous vehicles.

M. Lessons Learned from Deployments

The experiences of operators across regions highlight several key lessons about the integration of AI/ML into 5G networks:

➢ *Cost Reduction and Efficiency*

AI-driven automation lowers operational expenses, as seen in Vodafone, Telefónica, and América Móvil deployments, where predictive maintenance and traffic forecasting cut both energy and maintenance costs.

➢ *Sustainability as a Core Goal*

MTN's use of AI for solar-powered base station management and Telefónica's CO_2 reduction initiatives confirm that AI-enabled energy efficiency directly supports global sustainability objectives.

➢ *User-Centric Improvements*

NTT Docomo, AT&T, and América Móvil demonstrate how AI-powered forecasting and dynamic coverage ensure higher service quality and improved broadband access, even in rural or event-heavy scenarios.

➢ *Ecosystem Collaboration*

Partnerships between operators, cloud providers, and vendors (e.g., Vodafone–Google Cloud, Telefónica–Nokia/VMware, China Mobile–Huawei) demonstrate that collaboration accelerates innovation and improves network automation.

➢ *Scalability and Flexibility*

Deutsche Telekom's Open RAN AI strategies and América Móvil's spectrum optimization illustrate that AI/ML is crucial for scaling heterogeneous networks with diverse vendors and spectrum environments.

➢ *Security and Trust*

AI-powered intrusion detection and fraud prevention solutions from AT&T and América Móvil highlight the importance of trust and resilience in AI-driven 5G deployments.

In summary, operators across Europe, Asia, Africa, and Latin America show that AI/ML in 5G is central not only for technical optimization but also for advancing sustainability, service equity, and resilience in diverse deployment contexts.

VIII. LEVERAGING AI/ML WITH NWDAF FOR INDIAN OPERATORS

India's rapid 5G rollout, spearheaded by operators such as Reliance Jio and Bharti Airtel, is creating unprecedented opportunities to apply AI/ML in tandem with Network Data Analytics Function (NWDAF) as defined by 3GPP. With a massive subscriber base exceeding 1.1 billion mobile connections and the largest single 5G deployments globally, Indian operators face unique challenges in scale, affordability, and energy efficiency.

A. Scale and Subscriber Diversity

Indian operators manage networks serving hundreds of millions of users with diverse needs—from urban high-speed demand in cities like Delhi and Mumbai to rural broadband connectivity across remote states. NWDAF-enabled AI/ML can process data from heterogeneous traffic patterns, helping optimize QoS prediction, cell load balancing, and spectrum allocation at scale (GSMA, 2025).

B. Energy Efficiency for Cost-Effective Rollouts

Energy consumption is a major cost driver for Indian operators, particularly given the reliance on diesel-powered base stations in rural areas. AI/ML integrated with NWDAF can dynamically power down underutilized cells and optimize MIMO antenna usage. Reliance Jio has already piloted AI-driven energy savings across its rural networks, reporting measurable reductions in power usage per cell (Economic Times, 2024).

C. Rural and Urban Differentiation

Unlike many developed markets, India requires simultaneous optimization of dense urban high-speed networks and rural low-density networks. AI/ML models running within NWDAF can differentiate resource allocation policies by geography, ensuring metro users receive ultra-low latency while rural sites prioritize cost-efficient coverage (TRAI, 2024).

D. Network Slicing for Enterprises

With India's growing Industry 4.0 ecosystem, enterprises in sectors like manufacturing, logistics, and agriculture are demanding private 5G solutions. NWDAF-powered AI/ML allows operators to automate slice lifecycle management, tailoring bandwidth, latency, and reliability to enterprise use cases while ensuring security and regulatory compliance (Jio Platforms, 2024).

E. Monetization and Customer Experience

Given India's price-sensitive market, AI/ML analytics can support dynamic tariffing models, predicting demand surges and optimizing data plans. Bharti Airtel has been testing AI-driven customer retention models to reduce churn, integrating NWDAF insights into its CRM systems (Airtel, 2024).

F. Security and Fraud Detection

As India expands digital financial services and IoT ecosystems over 5G, fraud and cyberattacks are increasing. AI-powered anomaly detection within NWDAF helps operators spot suspicious activities in real time. Both Airtel and Jio have deployed AI-based fraud detection systems for mobile payments and enterprise 5G services (NASSCOM, 2024).

G. Collaboration with Government Initiatives

Indian operators are aligning with government programs such as Digital India and Make in India. AI/ML integrated with NWDAF supports smart city deployments, e-health, and e-governance services by ensuring reliable, scalable, and secure connectivity across regions (Government of India, 2025).

H. Comparative Perspective: India, China, and the U.S.

- *India*: Focuses on scale, affordability, and rural connectivity. Operators apply NWDAF for energy efficiency, cost reduction, and customer retention in a price-sensitive market.
- *China:* China Mobile and China Unicom emphasize industrial applications and smart cities, using AI for spectrum slicing and ultra-dense deployments.
- *U.S.:* AT&T and Verizon prioritize enterprise services and FWA (Fixed Wireless Access), leveraging AI/ML primarily for traffic prediction and QoE enhancement.

India's strategy is distinctive because it blends mass-market affordability with enterprise innovation. While China focuses on urban industrial automation and the U.S. emphasizes enterprise monetization, India is uniquely positioned to demonstrate how AI/ML with NWDAF can deliver scale, sustainability, and inclusivity.

IX. CHALLENGES AND FUTURE DIRECTIONS

Despite considerable progress, several challenges remain in the adoption of AI/ML for 5G networks.

A. Data Availability and Quality

Training robust AI/ML models requires large, high-quality datasets. In telecom environments, much of this data is sensitive, making data-sharing difficult. Privacy constraints and limited availability of labeled datasets can reduce model accuracy and generalization.

B. Model Transparency and Trustworthiness

Operators are cautious about adopting "black-box" AI systems in mission-critical environments. Lack of explainability in ML models limits trust, particularly in applications involving security, QoS assurance, or regulatory compliance. Research into explainable AI (XAI) is vital for wider acceptance.

C. Interoperability in Multi-Vendor Environments

5G networks are inherently heterogeneous, incorporating hardware and software from multiple vendors. AI/ML models trained in one vendor's environment may not perform consistently across others. Standardized interfaces for model deployment and retraining are needed to support interoperability.

D. Security Risks in AI/ML Models

AI/ML systems themselves can be targeted by adversarial attacks, poisoning data pipelines or misleading model outputs. Ensuring secure and resilient AI/ML integration is critical for protecting 5G networks against new forms of cyber threats.

ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/25sep051

E. Resource Constraints and Edge Deployment

While cloud-based training offers scalability, inference must often occur at the edge to meet latency requirements. Deploying AI models on constrained edge devices presents challenges in balancing accuracy, energy efficiency, and real-time performance.

F. Future Research Directions

- *Federated Learning:* Enables collaborative training across distributed nodes without centralizing data, addressing privacy and scalability concerns.
- *Explainable and Responsible AI:* Ensures transparent, auditable decision-making in critical 5G operations.
- *Quantum and Neuromorphic Computing:* Emerging paradigms that could accelerate model training and inference for 6G networks.
- *Integration with 6G Visions*: Many of these techniques will provide the foundation for 6G intelligent networks, where autonomy and sustainability will be even more critical.

X. CONCLUSION

The deployment of 5G networks worldwide has highlighted the need for adaptive, data-driven approaches to installation, configuration, and optimization. Traditional static management methods are insufficient to cope with the complexity of diverse service categories, traffic dynamics, and energy efficiency requirements. AI/ML integration, guided by 3GPP specifications, provides the foundation for intelligent network management.

This paper has shown how AI/ML techniques support coverage optimization, predictive maintenance, traffic management, and energy-efficient operations. It has also highlighted how 3GPP's TS 28.104, TS 28.105, and TR 29.908 formalize these capabilities within the framework of NWDAF and related functions. Real-world operator deployments demonstrate measurable improvements in performance, sustainability, and customer experience.

The adoption of AI/ML is not without challenges, including data privacy, model transparency, and interoperability. However, advancements in federated learning, explainable AI, and emerging computing paradigms point towards solutions. Looking ahead, AI/ML will not only enable the full realization of 5G but also serve as the foundation for 6G networks, where intelligent automation will be essential for global connectivity and digital transformation.

REFERENCES

[1]. Ericsson Mobility Report, June 2025.
[2]. GSMA, Mobile Economy 2025.
[3]. Wang, C.X., et al. (2020). Artificial Intelligence Enabled Wireless Networking for 5G and Beyond. IEEE Wireless Communications.
[4]. Pradhan, D., et al. (2023). Integration of AI/ML in 5G Technology. Taylor & Francis.
[5]. Agbon, E.E., et al. (2024). AI-driven Traffic Optimization in 5G and Beyond. Springer.
[6]. Panek, M., et al. (2023). 5G/5G+ Network Management Employing AI-based Continuous Deployment. Applied Soft Computing.
[7]. Latha, Y.L.M., et al. (2024). The Role of AI and Machine Learning in the Evolution of 5G and Beyond Networks. IEEE Access.
[8]. Haidine, A., et al. (2021). Artificial Intelligence and Machine Learning in 5G and Beyond: A Survey. Springer.
[9]. AMD Whitepaper, 2024. AMD EPYC and Instinct GPUs for 5G and AI Workloads.
[10]. Zhang, L., et al. (2023). AI-Driven Antenna Configuration for 5G NR Networks. IEEE Transactions on Mobile Computing.
[11]. Sun, Y., et al. (2022). Deep Reinforcement Learning for Adaptive Beamforming in 5G Massive MIMO. ACM MobiHoc.
[12]. Polese, M., et al. (2022). AI for 5G Network Slicing: Concepts, Architectures, and Challenges. IEEE Communications Magazine.
[13]. Vodafone Group, 2024. AI for Energy Efficient 5G
[14]. Telefonica, 2024. UNICA Next and AI in 5G
[15]. AT&T, 2024. 5G and AI Network Optimization
[16]. NTT Docomo, 2024. AI/ML in 5G Networks
[17]. China Mobile, 2024. AI-Driven 5G Innovation.
[18]. SK Telecom, 2024. TANGO AI Platform for 5G.
[19]. Deutsche Telekom, 2024. AI for Open RAN and Network Optimization.
[20]. MTN Group, 2024. AI for Network Expansion and Energy Efficiency in Africa.
[21]. America Movil, 2024. AI for Spectrum and Rural 5G.
[22]. Economic Times, 2024. Reliance Jio pilots AI for Energy Efficiency in 5G.
[23]. TRAI, 2024. Recommendations on Leveraging AI/ML for 5G Networks in India.
[24]. Jio Platforms, 2024. Enterprise 5G Solutions and AI Integration
[25]. Airtel, 2024. AI and Customer Experience in 5G.
[26]. NASSCOM, 2024. AI for Secure 5G Ecosystems in India.
[27]. Government of India, 2025. Digital India Initiative.