International Innovative Research Journal of Engineering and Technology
Vol. 11, Issue. 4, Jun 2026, pp. 53~63
ISSN: 2456-1983, DOI: 10.32595/iirjet.org/v11i4.2026.250

Deep Learning-Based Intelligent Framework for Cloud-Native 5G Core Network Management and Optimization

Dr. B. Nagarajan[1], Dr. R. Mahalingam[2]

[1]Assistant Professor, Department of Computer Science, Manbhumigu Dr. Puratchithalaivar MGR Govt. Arts and Science College, Keelavaniyur, Kattumannarkoil, Tamilnadu, India.
E-mail: thilaknaga@gmail.com
[2]Assistant Professor/Programmer, Department of Computer and Information Science, Faculty of Science, Annamalai University, Tamilnadu, India.
E-mail: r.mahalingamphd@gmail.com

Article Info

Article History:

Received Apr 12, 2026
Revised May 13, 2026
Accepted Jun 16, 2026

Keywords:

Closed loop robotics,
Data analysis,
Machine learning,
Cloud native technologies,
5G and 6G networks,
Artificial intelligence (AI).

ABSTRACT

Cloud native technology, which offers previously unheard-of levels of operational robotics, scalability, and versatility, has completely transformed 5G and 6G communications networks. However, distributing resources for fluid cloud computing environments faces a new difficulty due to the wide range of cloud native services and apps. The proposed framework, based on the CygNet MaSoN architecture, integrates real-time monitoring, data aggregation, predictive analytics, and deep learning models to optimize resource utilization and detect anomalous network behavior. The system enables proactive identification of service degradation, network performance issues, and security threats while supporting self-organizing and closed-loop automation capabilities required for autonomous 5G networks. Furthermore, the framework incorporates sequence-aware learning algorithms and synthetic data generation techniques to improve model performance in dynamic and context-dependent network environments. The components of a system and architecture are described in detail. After that, three actual use cases that have been performed on this structure are explained. The features taken into consideration are discussed together with machine learning in general models created and synthetic data production techniques used. These findings support the significance of sequence-aware algorithms for protecting roaming environments, which frequently involve context-dependent and fleeting dangers. The suggested paradigm offers a route for robust security in networks beyond 5G and 6G as well as a basis for intelligent, adaptive security monitoring in 5G.

Corresponding Author:

Dr. B. Nagarajan,
Assistant Professor, Department of Computer Science,
Manbhumigu Dr. Puratchithalaivar MGR Govt. Arts and Science College, Kattumannarkoil, Tamilnadu, India.
E-mail: thilaknaga@gmail.com.

1. INTRODUCTION

By linking to worldwide partner cellular networks, mobile users can access voice, message, and data services outside of their home networks through global roaming. To guarantee interoperability across various infrastructures, it depends on inter-operator agreements and defined communication standards. Although roaming has historically been crucial for international travel, it is becoming more and more crucial for enterprise applications like IoT and private networks [1]. Use cases like asset tracking require seamless changes between public and private networks and interoperability with older systems, even though only 64% of businesses now require worldwide IoT access. The safeguarding of signalling data becomes crucial when roaming spreads into these new areas, especially because sensitive data is transferred between networks with different security requirements.

However, there have long been serious security issues with the roaming environment, many of which are related to the usage of outdated signaling protocols. Due to the SS7 protocols lack of encryption and authentication in 2G and 3G systems, attackers were able to track subscriber positions, listen in on conversations, and launch denial-of-service or call extraction efforts [2]. Despite the advancements brought forth by Diameter in 4G, many networks continued to be incompatible with SS7, and IPX vendors continued to use unsafe connection techniques [3]. Mobile core networks were exposed to cross-domain assaults as a result of these hybrid installations [4]. For instance, the Syniverse breach showed how hackers might continue to get illegal access to signaling equipment for years, impacting millions of users and hundreds of operator.

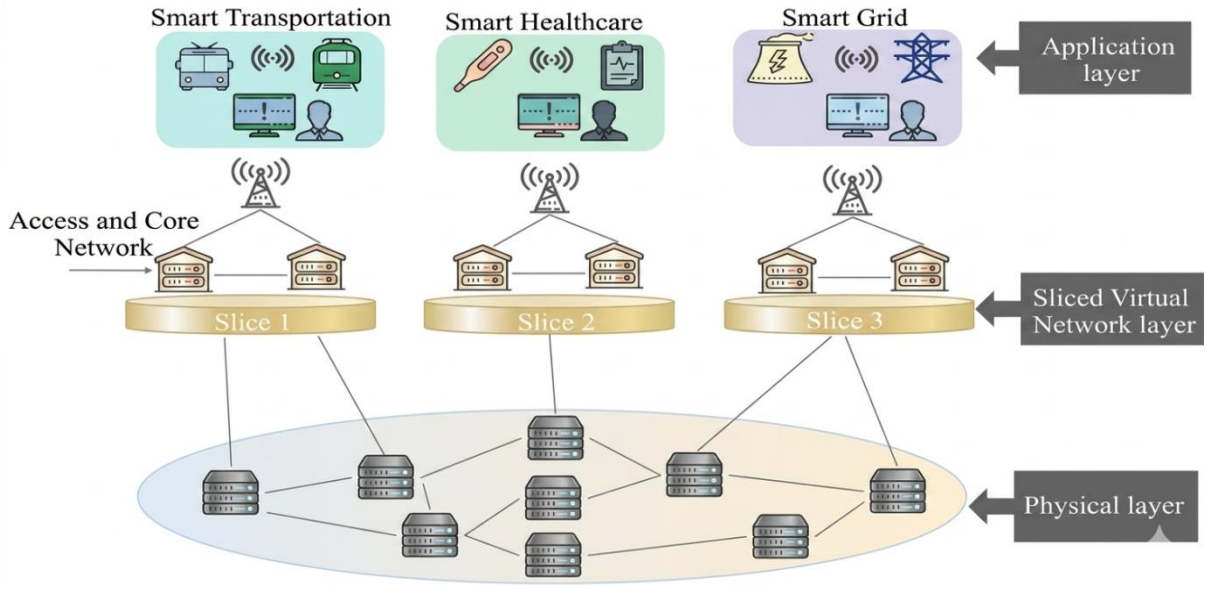

Figure 1. Slicing Architecture for 5G Networks

Figure 1 illustrates the current 5G network design, which enables a physical node to offer services to different tenants according to their needs. Depending on their setup, these tenants offer customized services [5]. Software solutions like NFV and Software-defined networking (SDN) are used to achieve network cutting. Programmability and modularity are two characteristics of the SDN and NFV approaches. Resources are used to establish new numerous logical connections [6]. The material layer is a standardized layer that uses the radio access network to serve a variety of different logical slicing networks, including core networks like data networks and mobile communication.

While the goal of network slices in a long-term 5G network is to lessen the strain on one physical cloud-based resource, the current architecture cannot satisfy the needs of contemporary applications. Applications based on smart cities, such as Smart Manufacturing, Smart Automobiles, and Intelligent Grid, require a variety of devices with various network setups and demands. For instance [7], hospitals have a variety of low-processing devices for cancer therapy, like blood pressure devices and weight scales. Network setups with low latency and high dependability are necessary for these devices. Databases and device safety systems, on the other hand [8], are very computationally demanding systems that need a lot of bandwidth to keep an eye on numerous devices dispersed throughout several regions. The performance required to control the infrastructure and gadgets in the best possible way could not be provided by a single network slice based on a common setup for every hospital.

The remainder of this essay is structured as follows: Section 2 examines related studies that use network slicing as a critical factor and seminal contribution for the effective distribution of resources; The suggested deep learning-based Network Sub-slicing Framework in a sustainable 5G environment is described in Section 3, along with the methodological flow ; in Section 4, we design a prototype and present an analysis of the proposed structure that outperforms previous studies, based on the simulated evaluation; The conclusion is presented in Section 5.

2. LITERATURE REVIEW

Although roaming is still an essential part of mobile communication systems, it creates serious security risks that are present in all network technology generations. Even under revised 5G protocols, the study shows how attackers with possession of roaming agreements can use authentication processes to install covert Rogue Base Stations (RBSes) [9], allowing for undetectable traffic surveillance and modification. In a similar vein, the study examines signaling-based DoS attacks in 5G traveling settings and demonstrates that, in comparison to non-roaming scenarios, these assaults can increase their effect on PLMNs by up to 2.69. There are extra dangers associated with legacy signalling protocols like SS7, which are still largely utilized for integration in regional architecture. The paper offers a thorough analysis of SS7 weaknesses, describing how hackers might leverage its trust-based architecture to intercept e-mails, track users, take over calls, or perpetrate financial fraud. Expanding upon these issues, the article offers an extensive empirical investigation of DL-based SS7-related recognizing an attack. It assesses eight guided and five semi-supervised ML models, showing that DL techniques show more promise, especially in identifying attacks attacking location update features in visiting networks, while rule-based detection mechanisms are mainly useless.

In order to implement MTD methods and improve safety in cloud-native network slicer settings, this study proposes a novel method that makes use of DRL. Our method dynamically chooses and implements MTD operations, such shuffling, immigration, and redundant operations, to proactively reduce possible vulnerabilities by incorporating Bayesian Attack Graphing (BAGs) for a thorough risk assessment [10]. The suggested Multi-Agent Deep Q-Learning architecture minimizes resource cost and efficiency loss while enabling each network slice to independently learn the best protection tactics.

[11] Suggests a dynamic security measures setting framework for micro-services as a solution to this problem. The system, which is based on the Multi-Agent Proximal Policy Optimization (MAPPO) algorithm, uses cooperative gaming and state sharing among intelligent agents to address high-dimensional collective decision-making challenges in microservice

situations. Our approach facilitates more effective security cooperation in multi-agent contexts when compared to current static or single-agent solutions. The framework can dynamically create security protection and QoS guarantee techniques based on the current system state, as shown by simulation studies. This effectively resists ongoing adversarial harm while reducing the impact on customer service.

[12] Suggested CausaLM-Net, a single diagnosis system that combines causal dependence modeling and semantic reasoning, as a solution to this problem. To find fault-relevant indicators automatically, a multi-large LLMs semantics feature selection module first aligns KPI definitions with operational logs. Second, a score-based causal graph creation module creates a sparse, comprehensible causal graph by learning the directional dependence structure of KPIs. Lastly, a specialized state-space-enhanced graph learning component adaptively controls edge dependencies and node dynamics, allowing the model to minimize noise and highlight fault-critical causal networks.

[13] Follows Kitchenham's criterion and the PRISMA technique, including papers published between 2015 and 2025. In this research, HO is categorized into three main groups: conventional, AI/ML-based, and blended approaches. Rule-based, fuzzy logic-based, velocity-based, and weighted-based tactics are traditional approaches that yield reliable and efficient results. Nevertheless, these approaches have limited flexibility in dynamic network scenarios and are limited by static standards. AI/ML-based methods that increase throughput and decrease handover mistakes and ping-pong effects by incorporating adaptive and prescriptive tuning of HCPs parameters include supervisory studying, reinforcement based teaching, and deep learning. However, these methods require a lot of data and processing power. Hybrid techniques have disadvantages like additional variables and lack of standardization, even if they beat single technique solutions by integrating the stability of conventional procedures with the flexibility of AI.

To address these issues, [14] propose RFD-R, an AI-driven runtime system that continuously improves the architecture of RAN CNFs. RFD-R optimizes a multi-objective cost function that combines latency, CPU utilization, and emigration overhead by using a DRL agent with Proximal Policy Optimization (PPO) to make realtime options. We derive formal analytic limits for the cost-benefit trade-off of function repacking and the worst-case service interruptions period. Our investigation, which combines computational modeling with artificial trace generation, shows that RFD-R outperforms dynamic and deterministic baselines, offering up to a 15% boost in CPU efficiency and a 35% drop in 95% percentile speed while retaining SLA delivery.

3. METHODS AND MATERIALS

3.1 Cloud Native Architecture

Figure 1 show the design of a cloud native wifi network, which is presently being researched for 5G beyond and 6G. It is made up of a RAN, a cloud native Core, and a transport network that links the RAN and the Core [15]. The cloud native system enables the development of highly portable and isolated setting that can be swiftly and simply deployed by utilizing modern innovations like containerizing.

A cloud native network's hardware elements can be divided into several virtual sections and distributed among NFs thanks to this design [16]. This increases the system's flexibility and efficiency by enabling the virtualization and dynamic distribution of physical resources, such as networks, storage space, and computing resources, among various NFs.

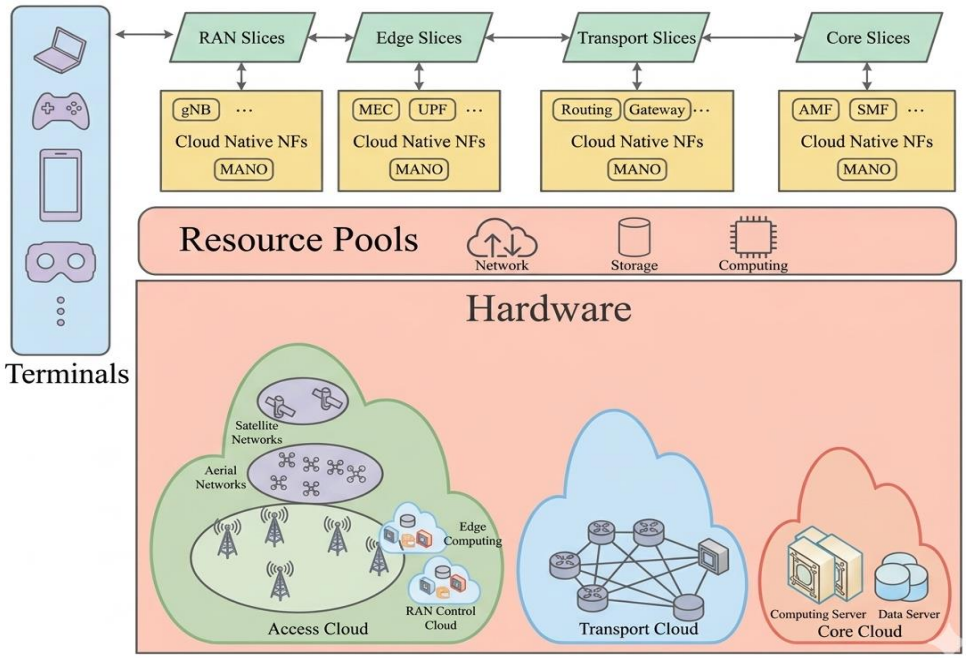

Figure 2. Architecture of the Proposed Cloud-Native Wireless Communication Framework

The proposed framework integrates cloud-native network functions, deep learning analytics, and automated orchestration to enable intelligent resource management and real-time optimization in 5G core networks in Figure 2. It supports anomaly detection, predictive decision-making and closed-loop automation to enhance network performance, scalability, and service reliability.

For instance [17], the UPF is in charge of directing traffic to the intended location. The UPF can be viewed as a decentralized and flexible information plane from the standpoint of MEC systems. Additionally, the SMF is in charge of choosing and managing the UPFs, and the Access and AMF manages mobility-related processes. Each NF is carried out in a cloud native manner by integrating Management and Orchestration (MANO) into each component to guarantee independent monitoring. Infrastructure resources can therefore be shared thanks to the cloud native construction, allowing for their dynamic allocation to satisfy a range of use cases. In order to determine the best resource allocation techniques, we concentrate on network segmentation and MEC in this work.

One key technology that was developed to give 5G networks the required degree of flexibility is network slicers. Network operators can separate the whole network into several virtual components to satisfy various user needs by creating network slices. However, NFs were smallest units for allocating resources in the Core prior to the implementation of the cloud's native structure. It is now feasible to distribute capabilities to an additional smaller unit, where the virtualized assets of computers can be distributed dynamically, thanks to the cloud native architecture in the Core. With the help of this feature, network administrators can create an evolving allocation of resources policy that maximizes the usefulness of the system. The cloud native structure improves resource utilization efficiency by allowing resources to be allocated to smaller units, which results in cost reductions, scalable performance, and more network agility.

4. PERFORMANCE OUTCOMES

This section presents the results of evaluating the effectiveness of the model implementation mentioned for all three use cases. The first use case's F1 score, the subsequent use case's MAE, MAPE, and MSE, and the third use case's MAE are the measures used in performance assessment.

4.1 UPF Anomalous behaviour Identification

Four distinct synthetic information sets were created in order to test the built XGBoost model for identifying UPF abnormal behavior. UPF PM KPI data records with a precision of fifteen minutes for a total of 90 days were included in each batch; a portion of these entries corresponded to abnormal information. This resulted in a total of approximately 25,000 data entries for each UPF instance considered during synthetic data generation. The 4 sets of synthetic data were generated such that every data set contained 8 different UPF instances across 2 networks where Network-1 contained 5 UPF instances and Network-2 contained 3 UPF instances. The 4 data sets were generated such that they contained 10%, 15%, 20% and 30% anomalous data entries for each UPF instance considered during synthetic data generation.

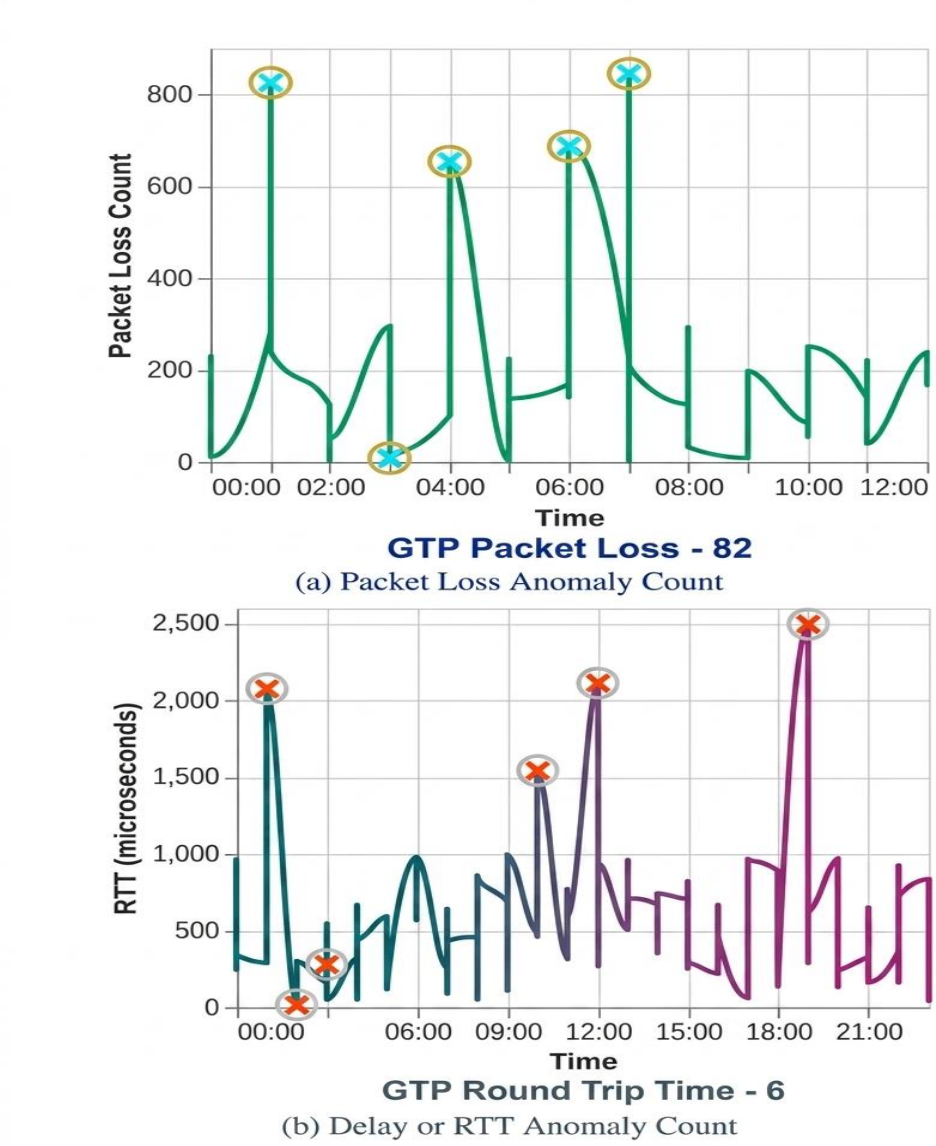

Figure 3. An Example of UPF Anomalous Detection Results

In Figure 3, 80% of the entire set of data entries was utilized for training, while the remaining 20% were used for testing. Each entry includes information for the 4 PM KPIs, assuming a N3 connection speed of 100 Gbps with 50 ms travel time and 1% packet loss. Typical data generating ranges were 0 to 80 Gbps for data volume, 0 to 10,000 packets for packet loss, 5 to fifty milliseconds for a round trip time, and 0 to 80 for link usage. A value was deemed abnormal if it exceeded the greatest value in the aforementioned ranges.

4.2 Forecasting Traffic Demand in 5G Cells

As previously stated, three machine learning models—RNN, LSTM, and LR—were taken into consideration. Ten epochs are used for both RNN and LSTM training models. As will be seen later, there was far less variation in the loss rate regardless of the count of epochs were raised to 50, 100, and 500. The activation methods examined were gaussian and relu, while the optimizers employed were Adam and Adamax. The total number of concealed layers was changed between 20 and 50.

Table 1. Comparison of Performance Measures for Load Prediction

Model	Parameters	MAE	MAPE	MSE
RNN	(20, R, AM)	17.7894	2.4440	828.1244
	(20, R, AX)	17.8897	2.4829	541.9406
	(20, S, AM)	25.7015	3.8997	885.0872
	(20, S, AX)	25.5824	4.0983	877.3069
	(50, R, AM)	18.5583	2.6435	709.2011
	(50, R, AX)	18.3474	2.4360	792.4941
	(50, S, AM)	25.7153	3.9035	886.8465
	(50, S, AX)	25.5813	4.0978	876.8651
LSTM	(20, R, AM)	17.7709	2.4544	480.6453
	(20, R, AX)	17.8967	2.4431	693.7534
	(20, S, AM)	25.7382	3.8624	888.1511
	(20, S, AX)	25.5816	4.0882	876.9000
	(50, R, AM)	18.9260	2.5900	537.8674
	(50, R, AX)	18.2297	2.5493	532.2069
	(50, S, AM)	25.7074	3.9050	885.7953
	(50, S, AX)	25.5840	4.0886	876.8872
LR	–	17.5036	2.4057	1122.7104

RNN, LSTM, and LR were evaluated for MAE, MAPE, and MSE in order to assess the load forecasting method used. Table 1 displays the MAE, MAPE, and MSE values for each machine learning model along with the relevant model variable permutations. It is evident that, in comparison to sigmoid, the activation process yields lower MAE, MAPE, and MSE indices for RNN and LSTM. Aside from this, there is little variation in the MAE, MAPE, and MSE values between the outcomes with 20 and 50 hiding layers and the results for optimization algorithms

Adam and Adamax. While the MSE for LR proves to be greater than that of RNN and LSTM, the MAE and MAPE values for LR nearly match those of RNN and LSTM.

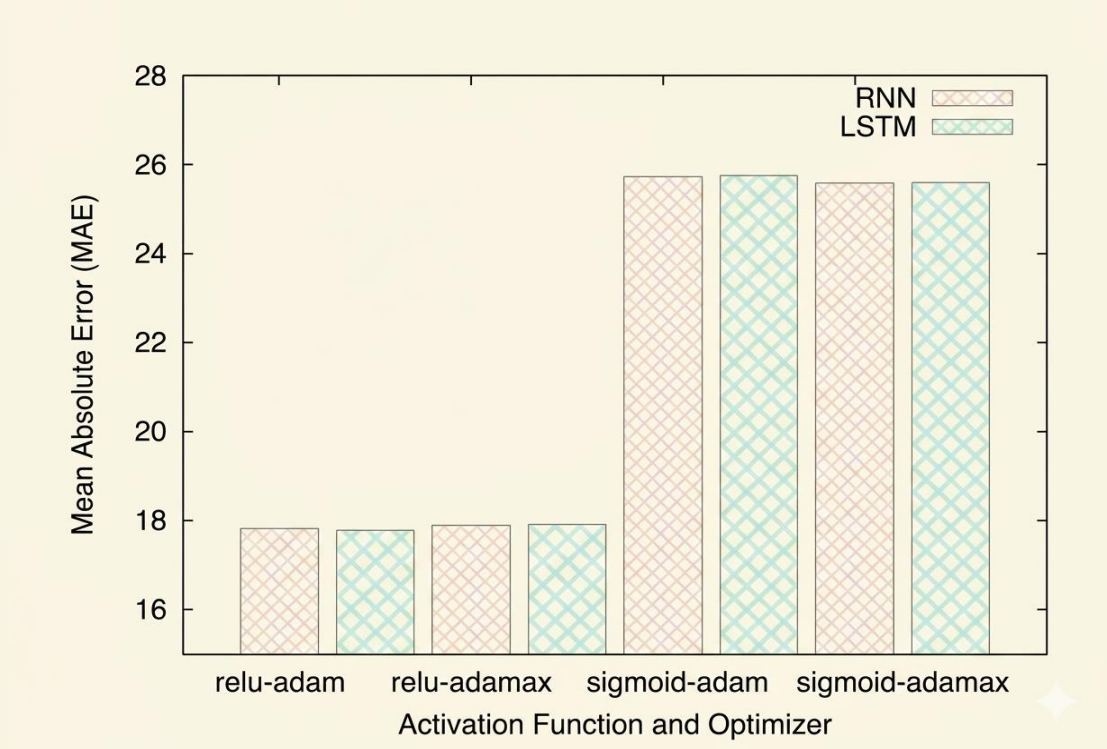

Figure 4. Findings for the Load Forecasting Use Case

The aforementioned study shows that all saturated cells were accurately identified by the applied method. By doing this, the network might be proactively configured to handle future traffic loads in Figure 4. In order to provide possible suggestions and ensure that the network is efficiently employed with the resources at hand, traffic volume monitoring would be consistently evaluated.

4.3 The Automation of Closed Loops for SMF Resource Use

The quantity of PDU sessions formation request and QoS flows for each SMF are taken into account when assessing the effectiveness of SMF resource utilization. For a few prior intervals, the mean total number of PDU sessions formation request received and the percentage of PDU session requests that were successfully completed were taken into consideration. For a few prior intervals, the mean overall number of QoS flow requests submitted and the percentage of successfully completed QoS flow requests have also been taken into account. In addition, variations in the overall quantity of PDU sessions demands and QoS flow demands received, as well as the proportion of properly processed PDU request requests and QoS flow requests, have been taken into account separately for performance assessment. The related processor and memory use are also taken into account as features for every evaluation. In every instance, the last twelve, twenty-four, and forty-eight hours are taken into account.

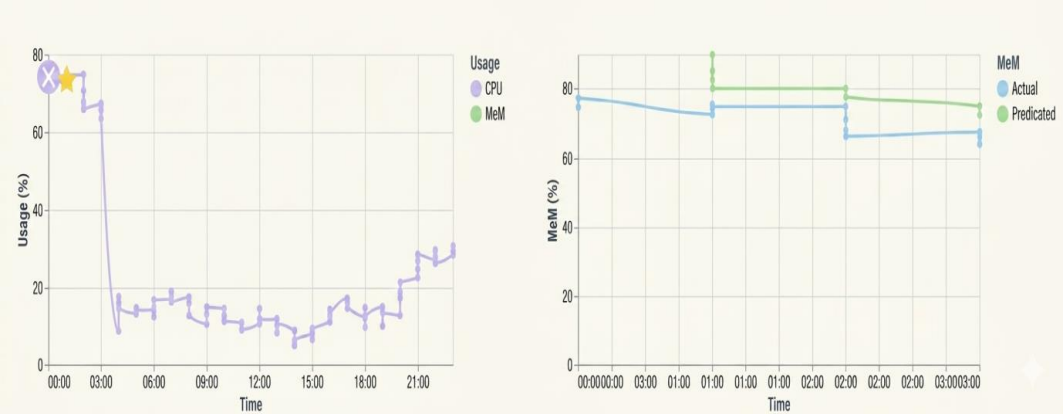

Figure 4. SMF Resource Usage Forecast Sample

This is demonstrated by the anticipated output sample snapshot from the system's client in Figure 4. A simulator was utilized to produce data in accordance with the stated details, and the proposed approach was applied in a test MaSoN system [18]. After a few weeks of system behavior observation, it became apparent that the MaSoN system constructed new SMF instances whenever excessive resource use was noted based on the most lately occurring in memory usage. Additionally, the resource use in the accessible SMF instances decreased considerably following each new SMF instantiation, demonstrating the efficacy of the adopted strategy [19]. The time moment at which a new SMF instantiate was initiated is indicated by the cross-mark sign on the left side graph. This picture unequivocally demonstrates that the SMF resource consumption dramatically dropped with the new SMF instantiation, and the projected SMF consumption of resources closely resembles the actual measured utilization.

5. CONCLUSION

In order to promote autonomous networking abilities in management of 5G networks, this study explains the architectural elements of the deployed management platform that is empowered with machine learning and analytics capabilities. Additionally discussed are the system's architecture, its components, and the specifics of these three use cases that were put into practice. Identification of UPF abnormal behavior, traffic load predictions in 5G cells, and SMF resource utilization predictions are the use cases that have been put into practice. Closed loop automated technology has been created for the SMF resource utilization use case, and it has been found to be successful in autonomously generating SMF instances by predicting high resource demand based on recent energy usage trends. Without the need for human intervention, this aids in preventing or mitigating possible traffic disturbances in the future. Because of its adaptability, the system may be expanded to accommodate more use cases in a network using 5G.

REFERENCES

[1] Wang, L., Wu, J., Gao, Y., & Zhang, J. (2023). Deep reinforcement learning based resource allocation for cloud native wireless network. *arXiv preprint arXiv:2305.06249*.

[2] Wang, L., Wu, J., Zhang, J., & Gao, Y. (2025). Efficient deep reinforcement learning based resource allocation for cloud native wireless network. *IEEE Transactions on Green Communications and Networking*.

[3] Ramachandran, M., Archana, T., Deepika, V., Kumar, A. A., & Sivalingam, K. M. (2022). 5g network management system with machine learning based analytics. *IEEe Access*, *10*, 73610-73622.

[4] Vázquez-Rodríguez, Á., García-Santaclara, P., Pontón-Rodríguez, J., Fernández-Castro, B., & Giraldo-Rodríguez, C. (2025, October). AI-Based Resource Management for Network Slicing in Cloud-Native Mobile Networks. In *2025 IEEE 50th Conference on Local Computer Networks (LCN)* (pp. 1-7). IEEE.

[5] Saravanan, S., & Thanigasalam, T. (2025, September). Leveraging AI/ML for Self-Optimizing 5G RAN: A Cloud-Integrated Framework. In *2025 2nd Asia Pacific Conference on Innovation in Technology (APCIT)* (pp. 1-8). IEEE.

[6] Valkama, M. (2025). INTELLIGENT SCALING OF CLOUD-NATIVE NETWORK FUNCTIONS USING REINFORCEMENT LEARNING.

[7] M.Raghu, & Dr. S. Ganesh Kumar. (2024). A Conditional Tabular Generative Adversarial Network (CTGAN)-based approach to safeguarding artificially created smart IoT settings. International Innovative Research Journal of Engineering and Technology (IIRJET) , *9*(4). https://doi.org/10.32595/iirjet.org/v9i4.2024.194.

[8] Sharma, V. K. (2026, January). Cloud-Native Network Slicing in 5G-Enabled Iot Systems: Architectures, Mechanisms, and Applications. In *2026 5th International Conference on Communication, Computing and Electronics Systems (ICCCES)* (pp. 230-237). IEEE.

[9] Ramakrishna, S. (2023). Cloud-Native AI Platform for Real-Time Resource Optimization in Governance-Driven Project and Network Operations. *International Journal of Engineering & Extended Technologies Research (IJEETR)*, *5*(2), 6282-6291.

[10] Hirai, S., Baba, H., Matsumoto, M., Hamano, T., & Noguchi, K. (2022, April). Machine learning based performance prediction for cloud-native 5G mobile core network. In *2022 IEEE Wireless Communications and Networking Conference (WCNC)* (pp. 1230-1235). IEEE.

[11] Huang, S. Y., Chen, C. Y., Chen, J. Y., & Chao, H. C. (2023). A survey on resource management for cloud native mobile computing: opportunities and challenges. *Symmetry*, *15*(2), 538.

[12] Sharma, V. K. (2025). Cloud Computing & IoT: 5G Focused IoT with Cloud Solutions. *International Journal of AI, BigData, Computational and Management Studies*, *6*(3), 21-25.

[13] Zeydan, E., Arslan, S., & Turk, Y. (2026). A Cloud Native Journey for Telecommunication Networks: Components, Applications and Open Challenges. *ACM Computing Surveys*, *58*(10), 1-38.

[14] Liu, G., Li, N., Yuan, C., Chen, S., & Liu, X. (2025). Service-Based Architecture for 6G RAN: A Cloud Native Platform That Provides Everything as a Service. *Sensors*, *25*(14), 4428.

[15] Shih-Yun, H., Cheng-Yu, C., Chen, J. Y., & Chao, H. C. (2023). A Survey on Resource Management for Cloud Native Mobile Computing: Opportunities and Challenges. *Symmetry*, *15*(2), 538.

[16] Motamary, S. (2025). A Deep Dive into CI/CD Pipelines Tailored for Telecom: Orchestrating Cloud-Native 5G Services with DevOps and Infrastructure Automation. *CD Pipelines Tailored for Telecom: Orchestrating Cloud-Native 5G Services with DevOps and Infrastructure Automation (May 04, 2025)*.

[17] Pawana, I. W. A. J., Abella, V., Lastre, J. K., Ko, Y., & You, I. (2025). Enhancing Roaming Security in Cloud-Native 5G Core Network through Deep Learning-Based Intrusion Detection System. *Computer Modeling in Engineering & Sciences*, *145*(2), 2733.

[18] Abella, V., Lastre, J. K., Pawana, I. W. A. J., Lee, D., Kim, B., & You, I. (2026). Benchmarking Deep Learning Architectures for Real-Time Intrusion Detection in Kubernetes-Orchestrated 5G Core Networks. *Research Briefs on Information and Communication Technology Evolution*, *12*, 57-66.

[19] Devineni, M., & Kaliappan, V. K. (2025). Enhancing Cloud Security: The Role of Artificial Intelligence in Real-time and Proactive Cyber Threat Detection. *Journal of Wireless Networks and Communication Systems*.