AI-Native Resource Allocation and Predictive Optimization in Ultra-Dense 6G Communication Networks

Kameswaran Arunachalam
Principal Engineer, Systems Architecture
T-Mobile USA Inc
kameswaran7@ieee.org

Venkata Rama Uday Kiran Bokam
Solutions Architect
Nokia
uday_kiran.bokam@nokia.com

***Abstract*—Efficient congestion management, performance optimization, and resource allocation are critical challenges in next-generation 6G communication networks due to increasing traffic demands, massive IoT connectivity, and the growing integration of smart-grid communication infrastructures. This paper proposes an AI-driven framework for predictive congestion management, network performance optimization, and intelligent resource allocation using the Telecom Network Dataset. The framework incorporates temporal feature extraction, feature engineering, categorical encoding, SMOTE-based class balancing, and StandardScaler normalization to enhance data quality and model performance. Multiple deep learning models, including DNN, LSTM, BiGRU, CNN-LSTM, and Transformer-based architectures, are employed for congestion prediction, while regression-based analysis is utilized for network performance optimization and resource-allocation estimation. Experimental outcomes shows that the CNN-LSTM model achieves the highest accuracy (98.33%), recall (98.57%), and F1-score (98.45%), whereas the Transformer model attains the best AUC-ROC value of 99.18%. Comparative analysis confirms the superiority of deep learning approaches over conventional machine-learning methods. Furthermore, LIME and SHAP techniques improve model interpretability, making the proposed framework suitable for intelligent telecom and smart-grid communication systems.**

Keywords—6G Networks, Intelligent Resource Allocation, Predictive Network Optimization, Telecom Network Analysis, Deep Learning, Network Traffic Analysis, AI-Driven Communication Systems.

I. INTRODUCTION

The significant advancement that fifth-generation (5G) communication networks have made in delivering fast, low latency, and reliable wireless connectivity has raised global anticipations for further advancements in the field of communication infrastructure systems, which is the sixth-generation (6G) networks [1][2]. Despite the fact that 5G communication technologies have brought about significant improvements in terms of data communication quality and efficiency as opposed to previous technologies [3].

The development of 6G communications mark a new milestone in wireless communication technologies by providing data transfer rates exceeding 100 Gbps, autonomous intelligence, and harmonious coordination among terrestrial, aerial, space, and edge networks [4]. Contrary to the past communication technologies, 6G networks enable ultra-dense communication environments whereby millions of nodes interact [5]. While such ultra-dense architecture offers advantages in terms of coverage and efficiency, it poses various challenges including those of interference handling, spectrum allocation, resource management, mobility, and energy usage [6]. Beyond telecom applications, the rapid evolution of smart-grid systems has increased the demand for intelligent communication infrastructures capable of supporting smart meters, distributed energy resources, phasor measurement units, and real-time grid control. Therefore, AI-driven congestion prediction and resource allocation techniques can play an important role in improving communication reliability and operational efficiency in future smart-grid environments.

AI and ML techniques have become an attractive choice for managing 6G communication networks by enabling intelligent, autonomous network resource management systems [7]. The algorithms used in machine learning offer sophisticated methodologies to perform efficient allocation of network resources based on learning from massive amounts of network data while adjusting to varying environmental conditions in real-time [8]. They attempt to imitate human thought processes in making decisions and are capable of efficiently allocating bandwidth, optimizing power control, managing traffic, sharing spectrum, and associating users [9]. By analyzing massive amounts of data, artificial intelligence systems can predict congestion, user behavior, channel usage, and resource requirements.

A. Motivation and Contributions of the Study

The fast spread of ultra-dense 6G networks and the very changeable nature of traffic raise some big issues for congestion management, resource allocation, and the general network dependability. In real life, the solutions are to some degree constrained when it comes to processing data at scale in telecom, and lack reactiveness if the network state is constantly changing. To overcome all of this, this study suggests an AI-based framework for congestion prediction, more intelligent resource allocation, end-to-end delay reduction, and even proactive network steering. The following section identifies the value added by the recommended structure:

- It provides a friendly platform based on artificial intelligence to improve the efficiency of resources, in addition to intelligent congestion management in super dense 6G communication networks.
- It uses advanced data preparation techniques such temporal feature extraction, FeatureEngineering, SMOTE balancing, and feature scaling to improve prediction performance.
- Design and test different DL architectures such as DNN, LSTM, BiGRU, CNN-LSTM and Transformer-based attention network for the efficient network traffic analysis, congestion prediction and predictive optimization.

XXX-X-XXXX-XXXX-X/XX/$XX.00 ©20XX IEEE

- Enables more proactive event detection in the network and more intelligent resource allocation with dynamic bandwidth distribution and latency optimization.

B. *Novelty of the paper*

The novelty of this work lies in the development of a unified AI-native framework that simultaneously addresses congestion prediction, performance optimization, and intelligent resource allocation in next-generation communication networks. Unlike existing studies that primarily focus on classification-based traffic prediction, the proposed framework integrates DL-based congestion forecasting with regression-driven network performance optimization and explainable AI techniques. The combination of CNN-LSTM, Transformer architectures, LIME, and SHAP enables both accurate prediction and transparent decision-making. Additionally, the framework extends beyond telecom applications to smart-grid communication networks, supporting adaptive, reliable, and energy-efficient cyber-physical system management.

C. *Structure of the Paper*

Here is how the paper is structured: Section II provides a summary of the relevant papers. Section III describes suggested approach, which includes a predictive modeling framework and dataset preparation. Section IV discusses outcomes of experiment and performance evaluation. The study's key findings and directions for further research are summarised in Section V.

II. Literature Review

This section reviews ML, DL, and explainable AI techniques used for intelligent congestion prediction, resource allocation, and predictive network optimization in ultra-dense 6G communication systems.

L. Zhang et al. (2026) proposed a PPO-assisted multi-slot matching algorithm to enhance solution efficiency in large-scale dynamic scenarios. The suggested algorithm yields an ideal impact of 94.6% in quasi-static circumstances, according to the simulation findings, with a complexity reduced to 3.2%. In dynamic scenarios, the results are 87.2% and 28.7%, respectively [10]. X. Liu et al. (2026) tested FAML in simulations with different distributions of environmental variables, and it outperformed the state-of-the-art approaches in terms of optimisation performance and adaption efficiency. Specifically, as compared to DRL and TL, FAML reduces training episodes by an average of 80% and 57%, respectively [11].

X. Wang et al. (2025) in both small-scale and large-scale networks, the suggested HoMa method performs better than benchmark algorithms, reducing the cost of multi-objective optimization, ComputationalCost, and SpectrumResource use by 20.29%, 38.96%, and 28.24%, respectively. The efficiency of the ILP model in optimizing ResourceAllocation is demonstrated by the near performance match [12]. Y. Sanjalawe et al. (2025) explores the possibility of using AI with 6G to improve the Smart Grid environment. This study presents SAFES-6G, a new conceptual framework for addressing latency, trust, and interoperability issues via the integration of edge intelligence, scalable AI, and explainable cybersecurity solutions. The results provide promising directions for privacy-preserving analytics in the areas of edge-native intelligence, quantum-safe encryption, and federated learning [13].

J. Cai et al. (2024) suggested approach significantly enhances network performance by boosting resource efficiency by up to 33.17% and 41.55%, respectively, and by improving throughput by more than 70% when compared to random algorithms and Q-Learning algorithms, respectively [14]. H. Muhammad Fahad Noman et al. (2024) proposed a method that performs 11.65%, 24.78%, and 47.29% better than distributed deep-deterministic policy gradient, multi-agent actor-critic, and deep Q network scheduling, respectively, and increases the total system Energy Efficiency by 41.52%. Additionally, when compared to MAAC, D3PG, and DQN scheduling, the suggested system reduces the likelihood of cellular outages by 5.88%, 15.79%, and 27.27%, respectively [15].

Existing studies have concentrated on specific 6G optimization problems, including energy consumption, throughput, and spectrum management. There is very little work on predictive congestion management, resource allocation, explainable AI, and traffic optimization in ultra-dense 6G networks in smart grid.

III. Methodology

The suggested method employs preprocessing, feature selection, SMOTE resampling, and DL-based predictive models for intelligent congestion prediction and optimization in ultra-dense 6G networks. Several DL models, including DNN, LSTM, BiGRU, CNN-LSTM, and transformer models, are evaluated using conventional metrics and explainable AI methods. Figure. 1 below shows the suggested method's entire workflow:

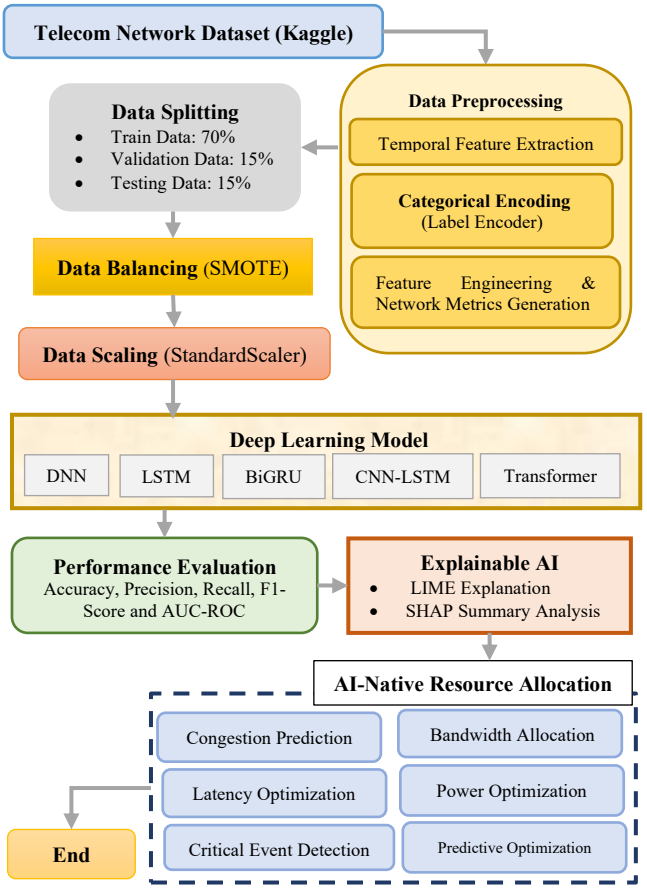

Fig. 1. Proposed Flowchart of the Proposed Predictive 6G System

A. Data Collection and Analysis

The Telecom Network Dataset [16] from Kaggle is used in the study, it comprises 3,605 records and 8 features related to telecom network operations and traffic analysis. The data includes network performance and usage metrics suitable for data preprocessing, feature analysis, and predictive modeling. It is suitable for applying network analysis, anomaly detection, and traffic prediction in intelligent telecommunication systems.

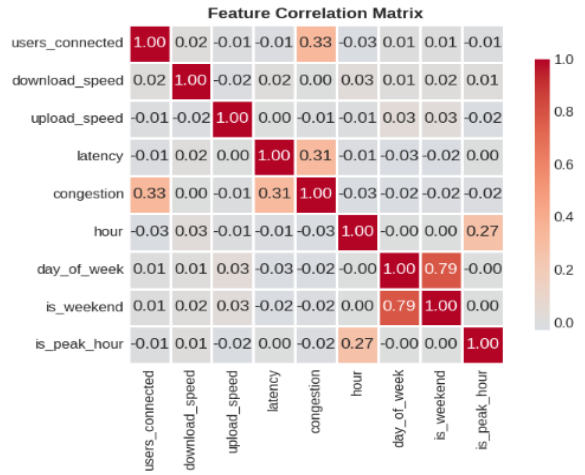

Fig. 2. Feature Correlation Matrix

Figure. 2 presents correlation matrix of selected features used in model. The plot reveals that there is a moderate positive correlation between the two variables (users_connected and latency) with network_congestion, implying that these metrics have a significant impact on the prediction of network congestion.

B. Data Preprocessing

The raw telecom network data is preprocessed using temporal feature extraction, Label Encoder-based categorical encoding, SMOTE balancing, StandardScaler normalization, and feature engineering approach. It significantly enhances traffic behavior analysis and the performance of predictive models with temporal patterns, categorical labeling, and a more balanced class distribution, as well as feature scaling.

- **Temporal Feature Extraction:** The time stamp values are converted to a datetime, and then the hour, day, month and day of week attributes are extracted. Then, it includes a few more variables, like if it's a weekend day, if it's a high-traffic hour, if it's a night time etc, to make sure that the traffic patterns are being picked up more effectively. Then the whole set is sorted by tower ID and date, thus maintaining the chronological order of the sequence.
- **Categorical Feature Encoding using Label Encoder:** In order to be used by machine learning methods, the categorical meteorological data was converted to numerical format before training the model. A unique number was assigned to each weather category using the Label Encoder class from the Scikit-learn library for the weather feature. The output was then saved as an additional variable, weather_encoded, for further analysis and model construction.
- **Feature Engineering and Network Metric Generation:** The records are sorted by date and tower ID, making the data ready to be modeled. After that, it may create some summary statistics for network metrics like the number of users connected, upload and download speeds, and latency. This makes it simpler to identify patterns and temporal relationships across time. when there are missing values, they are managed with forward filling and backward filling methods. Additionally, network efficiency metrics such as speed ratio, users per speed, and network load are created to better characterize congestion patterns and overall network behavior.

C. Data Splitting

A stratified split was used to divide the dataset into subsets with a balanced class distribution. The process generated about 70% Training Data, 15% Validation Data, and 15% Testing Data for the model's development and evaluation for classification and regression.

D. SMOTE-Based Oversampling Technique

By linearly interpolating between preexisting minority-class samples, the SMOTE algorithm produces new minority-class samples. This method restricts new samples to reside along linear pathways between existing ones, which may not precisely reflect the underlying data distribution, even though it somewhat reduces overfitting. In a similar vein, a random quantity is added to create a sample that is closer to the neighbor when the base sample is closer to the first minority sample.

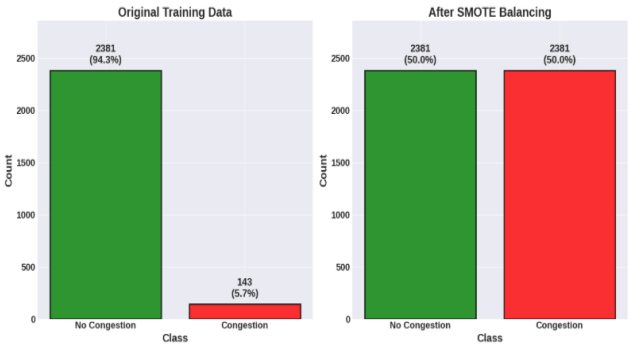

Fig. 3. Class Distribution Before and After SMOTE Balancing

Figure 3 shows the Class Distribution both before and after SMOTE balancing. The original data set was highly skewed, with majority of the data belonging to the No Congestion class. Both classes, No Congestion and Congestion, are also balanced after applying SMOTE, thus enhancing the suitability of the datasets for model training and classification performance.

E. Data Scaling using StandardScaler

The process of standardization involves changing a dataset's properties so that they follow a typical normal distribution with a standard deviation of 1 and a mean of 0. The z-score scaling standardization transformation, which was employed in this procedure, is represented mathematically by Equation (1). The StandardScaler from scikit-learn was used to scale dataset's features, as shown in Equation (1).

$$Z = \frac{(xi - mean(x))}{stdev(x)} \quad (1)$$

F. Deep Learning Models for Ultra-Dense 6G Networks

The proposed framework uses DNN, LSTM, BiGRU, CNN-LSTM and Transformer models for prediction and

optimization. These models work well in learning complex spatial and temporal patterns, which help to enhance network performance and resource management.

1) Deep Neural Network (DNN)

DNNs are one of the deep learning algorithms that are utilized the most. They have a fully connected structure, with numerous layers, and all the neurons in each layer before and after are connected to all the other neurons in that layer [17]. A DNN's architecture allows for increasingly more complex representations to be extracted as the input moves through each layer. Since DNNs learn hierarchically, they can detect the complex relationships and patterns in the data. Dropout, batch normalization, and many dense layers with the ReLU activation function are all used in the DNN model. Acc, prec, rec, and AUC are employed as performance measurements, while during training, a Binary cross-entropy loss is combined with a learning rate of the Adam optimizer is 0.001. The training process uses 50 epochs and a batch size of 32.

2) Long Short-Term Memory (LSTM)

LSTM was very successful at simulating input data's long-range temporal dependencies. By accurately simulating intricate patterns and long-term temporal connections in time series data, LSTMs improve traffic forecasts [18]. The weights of three NNs are used to determine hidden state h_t as a function of input tensor X_t, prior hidden state h_t-1, and cell state c_t, as shown in Equation (2).

$$h_t = f(W_o, W_i, W_f)(h_{t-1}, X_t, c_t) \quad (2)$$

Where:

- The output gate's associated weights are denoted by W_o.
- The Forget gate's associated weights are denoted by W_f.
- The input gate weights are denoted by W_i.

A stack of LSTM layers with batch normalization and dropout is used in the model architecture to detect sequences. The cost function was binary cross-entropy, and training was done using AdamOptimizer (LearningRate = 0.001). Acc, Prec, Rec, and AUC are used as performance metrics. Training was carried on for about 100 epochs, with Early Stopping , in batches of 32.

3) Bidirectional Gated Recurrent Unit (BiGRU) Network

A forward-facing GRU layer plus a backwards-facing GRU layer basically forms the Bi-GRU. The Bi-GRU's bidirectional encoding method helps it, to constantly blend earlier information and later information at the same time, so it can grab the evolving changes and those short-term correlations within the sequence more effectively [19]. The BiGRU architecture uses a few stacked layers of Bidirectional GRUs along with batch normalization and Dropout, to help with the sequential learning part and also to limit overfitting a bit too much. Adam's optimization with a LearningRate of 0.001 is combined with a binary cross-entropy loss during the training phase. For evaluation, also look at acc, prec, rec, and AUC, as the usual assessment metrics. During training, also apply early stopping and callbacks so it doesn't always run too long, generally 100 epochs, with a batch size of 32.

4) CNN-LSTM Hybrid

The CNN-LSTMmodel was formerly known as the Long-term Recurrent ConvolutionalNetwork (LRCN). In this context, LRCN is often referred to as broader phrase CNN-LSTM [20]. Basically, CNN-LSTM model predicts output data by applying an LSTM only after features are taken from the input data via a CNN. So, the CNN-LSTM architecture can be seen as a hybrid ensemble of Conv1D and LSTM components, where the Conv1D part extracts features from input, while LSTM component handles sequential processing. Overfitting is avoided by using batch normalization and dropout. In this instance, binary cross-entropy is the loss function and Adam, an optimizer with a LearningRate of 0.01, is employed.

5) Transformer-Based Multi-Head Attention Network

The Transformer module learns word-level representations of aspects and contexts, respectively, in addition to using the Tree Transformer module. Recursive and convolutional components are absent from the Transformer, a novel sequence-to-sequence model [21]. The self-attention mechanism, which calculates characteristics for each pair of words in the sentence, is its fundamental component. Multi-Head Self-Attention, Layer Normalization, skip connections, and a feed-forward network make up transformer encoder structure used in the Transformer-based Multi-Head Attention technique. The Adam optimizer (learning rate = 0.001) with binary cross-entropy as its objective function is used to assess the model's accuracy, precision, recall, and AUC. In addition, the model has 4 heads, head dimensionality equals 32, and there are 128 units for the feed-forward layer, plus three transformer layers.

G. Performance Metrics

The performance of these proposed models is kind of checked using a confusion matrix, but only after the preprocessing and modeling steps are done. For the training dataset, the acc, prec, rec, F1, and AUC-ROC are computed from that ConfusionMatrix. When True Negative (TN) and True Positive (TP) scores are greater, model often appears to be better. True Negative indicates that negative instances are accurately forecasted as 0, whereas True Positive indicates that positive situations are accurately anticipated as 1. The formulas for acc, prec, rec, F1, and AUC-ROC are shown in Equations (3)–(6), respectively.

Accuracy: This statistic is defined as the ratio of accurately predicted samples to all samples in a certain set of data. It is expressed as Equation (3).

$$Accuracy = \frac{TP+TN}{TP+FP+TN+FN} \quad (3)$$

Precision: A measure of a model's accuracy in making positive predictions is its precision. The statistic calculates ratio of successfully detected positive instances to all anticipated positive cases. It is expressed as Equation (4):

$$Precision = \frac{TP}{TP+FP} \quad (4)$$

Recall: This metric evaluates model's ability to recognize positive samples among all actual positive samples. It's sometimes called sensitivity or TPR. In mathematical form, it is given as Equation (5):

$$Recall = \frac{TP}{TP+FN} \quad (5)$$

F-score: F1 score is a reasonable parameter for evaluating the model's efficacy as it integrates recall and accuracy. Mathematically, it is given as Equation (6):

$$F1 - Score = 2 * \frac{(Precision*Recall)}{Precision+Recall} \quad (6)$$

Collectively, these measures provide a comprehensive evaluation of the model's accuracy and overall efficacy in forecasting the dependent variable.

IV. Result Analysis and Discussion

The experimental analysis was performed to assess deep learning models for intelligent resource allocation and predictive network optimization using a system equipped with an IntelCore i7-6500U CPU, 8 GB RAM, and 512 GBstorage [22]. A Critical Event is defined as a network condition where the predicted congestion probability exceeds 0.80, indicating significant performance degradation risk. A High-Risk Event is defined as a severe congestion condition where the congestion probability exceeds 0.90 and immediate resource allocation intervention is required. As shown in Table I, CNN-LSTM Hybrid model outperformed among all models in terms of acc (98.33%) and F1score (98.45%), while Transformer model performed well in terms of AUC-ROC value (99.18%). The DNN, LSTM and Bidirectional GRU models showed good performance, with accuracies above 97%. The results show that hybrid and attention-based deep learning models are very effective in optimizing next-generation communication networks.

TABLE I. Performance Evaluation of Deep Learning Models in 6G Networks

Matrix	Accuracy	Precision	Recall	F1-Score	AUC-ROC
DNN	97.96	80.64	83.33	81.96	99.04
LSTM	97.78	82.14	76.66	79.31	99.12
Bidirectional GRU	97.59	79.31	76.66	77.96	98.91
CNN-LSTM Hybrid	98.33	81.81	90	85.71	93.79
Transformer	97.59	75.75	83.33	79.36	99.18

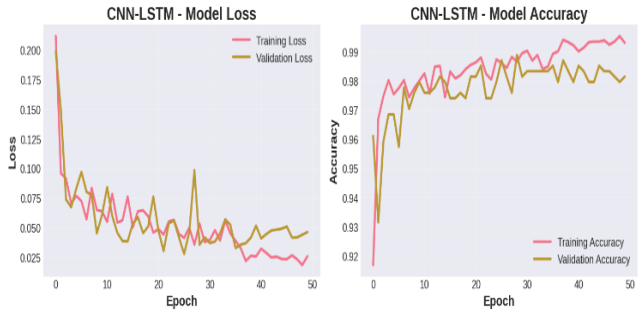

Fig. 4. Training and Validation Performance of the CNN-LSTM Model

Figure. 4 shows the CNN-LSTMmodel's training and validation loss and accuracy curves over epochs. The loss gradually decreases while the accuracy steadily increases for both training and validation datasets, indicating effective learning and stable model performance with minimal overfitting.

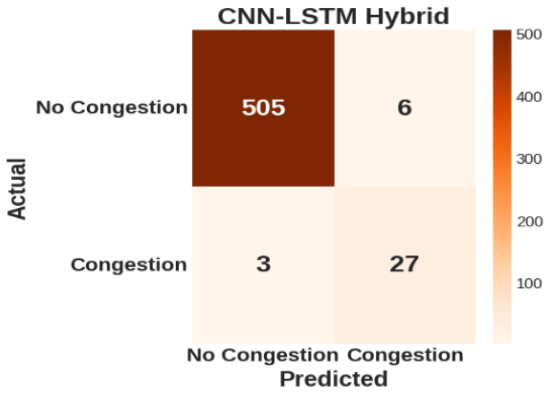

Fig. 5. Confusion Matrix of the CNN-LSTM Model

The confusion matrix for CNN-LSTMmodel's prediction of congestion is displayed in Figure. 5. The CNN-LSTM model successfully identified 505 cases of No Congestion and 27 cases of Congestion, with very few misclassified cases, indicating high accuracy and robust performance of CNN-LSTM model.

Figure. 6 presents an ROC curve comparison of different models, in the CNN-LSTM model's AUC was 0.9380. This curve shows model's ability to accurately predict congestion and non-congestion classes, and its performance is compared to the random classifier as a baseline.

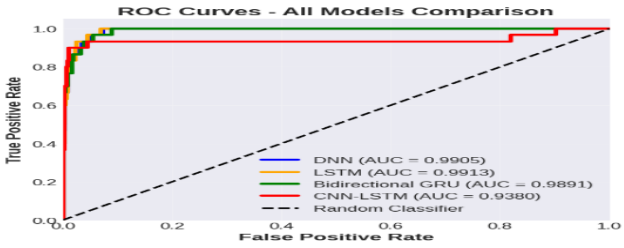

Fig. 6. ROC Curve of the CNN-LSTM Model

Table II demonstrates the optimized performance achieved by optimizing the 6G towers directly in the AI domain. The framework is dynamically optimized, to deliver a power on the order of 21.91–22.38 W, along with efficiency that kind of hovers around 66.49–69.46. After that, the system can detect 231 critical events, and it can even detect 25 high-risk events, so it seems as if it is pretty good at resource stewardship and network stability in these very dense 6G environments.

The resource allocation statistics are generated by the proposed AI-native resource allocation engine using predicted congestion risk, network load, latency, and bandwidth demand indicators. Critical and High-Risk Events correspond to congestion probabilities exceeding 0.80 and 0.90, respectively.

TABLE II. AI-Native Tower Performance Evaluation in 6G Simulation

Tower ID	Power Allocation	Critical Events	High-Risk Events	Total Records	Efficiency Score
1	22.19	47	6	721	68.85
2	22.38	50	8	721	66.49
3	21.91	40	5	721	69.46
4	22.36	52	3	721	68.44

A. Predictive Optimization Using Bidirectional GRU Regression

Table III presents the predictive optimization results for download speed and latency forecasting using Bi-GRU. The framework achieved low prediction errors, with download speed MAE and RMSE of 1.26 and 1.62, respectively, and latency MAE and RMSE of 1.73 and 2.31. The high R^2 values of 0.9965 and 0.9983 demonstrate excellent forecasting accuracy and reliability.

TABLE III. NETWORK PERFORMANCE FORECASTING RESULTS

Metric	Mean	Std
Download Speed MAE	1.26	±0.08
Download Speed RMSE	1.62	±0.11
Download Speed R^2	0.9965	±0.0004
Latency MAE	1.73	±0.10
Latency RMSE	2.31	±0.14
Latency R^2	0.9983	±0.0003

B. 6G IMT-2030 Benchmark Analysis and AI-Optimized Network Performance Evaluation

Table III is a comparison of the simulated behavior of a 6G network to the performance targets set by ITU-R IMT-2030, while at the same time examining the amount of efficiency improvement that can be obtained through the use of AI-based network optimization. Overall, it appears the simulation is pretty good at capturing some of the key aspects of 6G networks such as ultra-dense deployment, fast data paths and massive user connectivity, though it's more of a scaled-down proxy, which is to be expected. In the AI optimization framework (shown in Table IV), there is a small latency decrease, from 105.10 ms to 105.00 ms. it dynamically allocates 450.4 Mbps of bandwidth and around 22.2 watts of power per tower, also it proactively identifies 231 critical events and 25 high-risk events. The findings in this paper indicate that the dependability and overall performance of next-generation 6G wireless communication systems can be enhanced by aggregating resources and performing preemptive congestion monitoring.

TABLE IV. 6G IMT-2030 TARGET VERSUS SIMULATED NETWORK PERFORMANCE

Metric	6G IMT-2030 Target	Simulated Dataset Performance	Description
Peak Download Speed	100,000+ Mbps	~100 Mbps	Simulation proxy for 6G throughput
Average Latency (ms)	< 1 ms	105.10 ms	Scaled latency model
Congestion Rate (%)	< 1%	5.63%	AI-reduced congestion target
Ultra-Dense Cells	>> 1000 cells/km²	5 towers	Dense network deployment proxy
User Connections per Tower	100,000+ users	531 users	Scaled user connectivity model

TABLE V. AI OPTIMIZATION IMPACT ON SIMULATED 6G NETWORK PERFORMANCE

Performance Metric	Baseline Network	AI-Optimized Network	Observation / Impact
Average Latency (ms)	105.10 ms	105.00 ms	Slight latency reduction achieved
Latency Reduction (%)	—	0.10%	AI-assisted optimization
Average Download Speed (Mbps)	52.09 Mbps	51.37 Mbps	Minor variation after optimization
Speed Change (%)	—	-1.37%	Slight decrease in throughput
Average Bandwidth Allocation	—	450.4 Mbps per tower	Dynamic AI-based allocation
Average Power Allocation	—	22.2 W per tower	Energy-aware optimization
Critical Events Detected	—	231	Proactive AI intervention enabled
High-Risk Events Detected	—	25	Predictive congestion monitoring

Table V provides a breakdown of what has happened per network tower in the effect of the resource allocation by the AI, sort of like a little break down of the outcome. It also provides performance indicators like predicted speed latency, allocation of bandwidth and then the congestion risk factor. From the analysis, it can be observed that all towers are under OPTIMAL status category, with low congestion risk ranging from 0.064 to 0.079, which is expected to result in good network speeds over 50 Mbps, and the latency values are close to 105 ms, which suggests no major changes in the allocation settings are needed to ensure a stable and efficient network.

TABLE VI. AI-NATIVE RESOURCE ALLOCATION AND PER-TOWER ACTION

Tower ID	Network Status	Action Priority	Congestion Risk	Predicted Speed (Mbps)	Predicted Latency (ms)	Bandwidth Allocation (Mbps)	Recommended Action
Tower 1	OPTIMAL	P4 - Routine	0.073	52.1	105.3	439	Stable maintain current allocation parameters
Tower 2	OPTIMAL	P4 - Routine	0.079	50.1	106.5	464	Stable — maintain current allocation parameters
Tower 3	OPTIMAL	P4 - Routine	0.064	51.7	102.5	443	Stable — maintain current allocation parameters
Tower 4	OPTIMAL	P4 - Routine	0.079	51.7	105.0	456	Stable — maintain current allocation parameters
Tower 5	OPTIMAL	P4 - Routine	0.067	51.4	105.7	449	Stable — maintain current allocation parameters

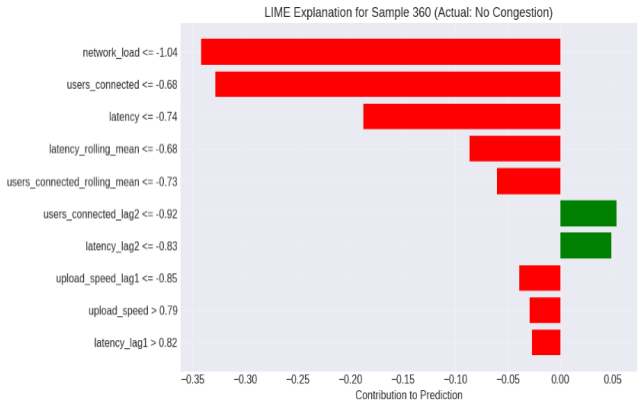

Fig. 7. LIME Explanation of the CNN-LSTM Model

The CNN-LSTM model's LIME explanation is shown in Figure. 7, where each feature's contribution to the prediction of a No Congestion event is highlighted. The interpretation of the model highlights its decision-making process, as the features with a negative impact on the prediction, including network_load, users_connected, and latency, are dominant, while a few lag features have positive values.

Figure. 8 displays a SHAP summary graphic for the CNN-LSTM model, showing how different characteristics affect its predictions. The most important features for model's predictions are features network_load, users_connected and latency. The prediction output increases with features that have positive SHAP values and decreases with features that have negative SHAP values.

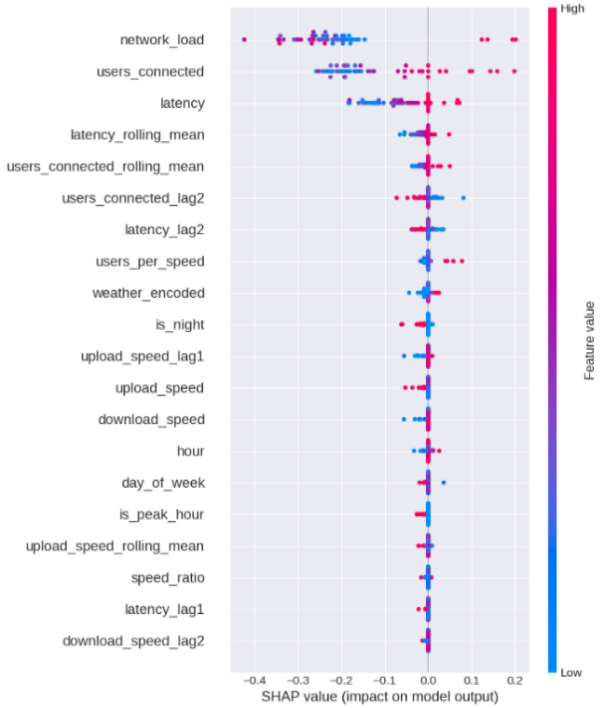

Fig. 8. SHAP Summary Plot for the CNN-LSTM Model

C. Comparison and Discussion

Several ML and DL models were compared for tasks related to network optimisation and prediction in Table VI. The CNN-LSTMmodel achieved highest acc of 98.33%, outperforming DNN model with 97.96% accuracy and traditional approaches including Logistic Regression, Gaussian SVM, and Gradient Boosting. The results demonstrate that DL-based hybrid architectures provide better learning capability and improved predictive performance compared to conventional ML methods.

TABLE VII. COMPARATIVE ACCURACY ANALYSIS OF MACHINE LEARNING AND DEEP LEARNING MODELS

Model	Accuracy
CNN-LSTM	98.33
DNN	97.96
LR [23]	95%
Gaussian SVM [24]	92.3
GB[25]	92%

D. Discussion

Although the observed latency reduction is numerically modest, the primary objective of the proposed framework is proactive congestion prediction and intelligent resource allocation rather than direct latency minimization. The framework achieves 98.33% prediction accuracy, enabling early identification of congestion events and supporting preventive resource-allocation decisions before severe network degradation occurs. The CNN-LSTM model effectively captures both temporal and spatial traffic patterns, improving prediction and optimization capabilities, while the DNN model provides robust feature learning and reliable classification performance. The Transformer model further enhances discrimination ability with superior AUC-ROC performance. Together, these AI-driven models demonstrate improved adaptability, predictive intelligence, and autonomous network management compared with conventional resource-allocation approaches, making them suitable for future AI-native communication and smart-grid environments.

V. CONCLUSION AND FUTURE SCOPE

AI-driven congestion prediction and resource optimization have emerged as essential components of next-generation communication networks, particularly in the context of intelligent 6G-enabled services, large-scale IoT connectivity, and smart-grid communications. The proposed framework integrates advanced preprocessing, deep learning models, and explainable AI techniques to support proactive network management and intelligent resource allocation. Experimental results demonstrated strong predictive performance, with the CNN-LSTM model achieving 98.33% accuracy, 90% recall, and an F1score of 85.71%, while the Transformer model attained the highest AUC-ROC value of 99.18%. Furthermore, the framework enabled dynamic bandwidth allocation averaging 450.4 Mbps per tower, maintained an average power budget of 22.2 W, and proactively identified 231 critical and 25 high-risk congestion events. These outcomes highlight the capability of AI-native optimization to enhance network reliability, operational efficiency, and intelligent decision-making. Despite these encouraging outcomes, there are still a few restrictions. The experimental evaluation was performed using a Telecom Network Dataset containing 3,605 records collected from five communication towers, representing a scaled experimental environment rather than a fully operational ultra-dense 6G deployment. Consequently, the reported latency values reflect telecom-network conditions and should not be interpreted as compliance with IMT-2030 6G performance requirements. Therefore, the findings should be considered a proof-of-concept validation of the proposed framework. Future work will focus on validation using large-scale telecom datasets, realistic 6G testbeds, synthetic 6G simulations, and

operational smart-grid communication datasets involving smart meters, PMUs, microgrids, renewable-energy systems, and distributed energy resources. Additionally, federated learning, reinforcement learning, edge AI, Graph Neural Networks (GNNs), Temporal Fusion Transformers (TFTs), and Generative AI-based optimization techniques can be integrated to further improve congestion forecasting, scalability, energy efficiency, communication reliability, and intelligent resource allocation across large-scale communication and smart-grid cyber-physical systems.

REFERENCES

[1] S. Suyama, T. Okuyama, Y. Kishiyama, S. Nagata, and T. Asai, "A Study on Extreme Wideband 6G Radio Access Technologies for Achieving 100Gbps Data Rate in Higher Frequency Bands," *IEICE Trans. Commun.*, vol. E104.B, no. 9, p. 2020FGI0002, Sep. 2021, doi: 10.1587/transcom.2020FGI0002.

[2] V. Poirot, M. Ericson, M. Nordberg, and K. Andersson, "Energy efficient multi-connectivity algorithms for ultra-dense 5G networks," *Wirel. Networks*, vol. 26, no. 3, pp. 2207–2222, Apr. 2020, doi: 10.1007/s11276-019-02056-w.

[3] M. R. Cheriya Mukkollakkal, "IntelliStore: An Intelligent AI Agent Framework for Autonomous Storage and Database Optimization in Cloud-Native Microservices," *Int. J. Sci. Res. Mod. Technol.*, vol. 3, no. 12, pp. 243–250, 2024, doi: 10.38124/ijsrmt.v3i12.1024.

[4] R. Sinha, R. Gulati, A. Muttayane, K. Arunachalam, and J. Smith, "Network latency time measurement using DNS and web server messages," 11245606, 2022

[5] S. Singh, "Performance Evaluation of Machine Learning Regression Models for 5G Network Resource Allocation Optimization," in *2025 International Conference on Multimedia Computing, Networking and Applications (MCNA)*, 2025, pp. 47–52. doi: 10.1109/MCNA65829.2025.11124374.

[6] S. Singh, S. A. Pahune, P. Chatterjee, and R. Sura, "Advanced Machine Learning Methods for Churn Prediction and Classification in Telecom Sector," in *2025 IEEE 6th India Council International Subsections Conference (INDISCON)*, 2025, pp. 1–7. doi: 10.1109/INDISCON66021.2025.11252233.

[7] V. Sharma, "Cloud-Native 5G Deployments: Kubernetes and Microservices in Telco Networks," *Int. J. Innov. Res. Eng. Multidiscip. Phys. Sci.*, vol. 10, no. 3, pp. 1–8, May 2022, doi: 10.37082/IJIRMPS.v10.i3.232706.

[8] V. Sharma, "5G-Enabled Mission-Critical Networks Design and Performance Analysis," *Int. J. Sci. Technol.*, vol. 14, no. 4, pp. 1–11, Oct. 2023, doi: 10.71097/IJSAT.v14.i4.7998.

[9] M. Patel and U. Korat, "Swarm Optimization Algorithm-Enhanced Clustering Techniques for Reliable Wireless Sensor Networks Communication," in *2026 IEEE 5th International Conference on AI in Cybersecurity (ICAIC)*, IEEE, Feb. 2026, pp. 1–6. doi: 10.1109/ICAIC67076.2026.11395724.

[10] L. Zhang, J. Du, C. Jiang, X. Hou, Y. Ren, and Z. Han, "Dynamic Resource Allocation in Maritime Unmanned Networks: A Hybrid Approach of Three-Sided Matching and Reinforcement Learning," *IEEE Trans. Wirel. Commun.*, vol. 25, pp. 15377–15391, 2026, doi: 10.1109/TWC.2026.3682752.

[11] X. Liu, S. Yin, M. Cai, X. Fan, S. Wang, and S. Huang, "Fast Adaptive Meta-Learning Framework for Resource Allocation in Optical Networks," *IEEE Photonics J.*, vol. 18, no. 3, pp. 8600113–8600113, Jun. 2026, doi: 10.1109/JPHOT.2026.3690755.

[12] X. Wang *et al.*, "Flexible-Unit Deployment and Lightpath Provision Over 6G AI-RAN for Holographic Type Communications," *J. Light. Technol.*, vol. 44, no. 4, pp. 1342–1356, Feb. 2026, doi: 10.1109/JLT.2025.3640080.

[13] Y. Sanjalawe, S. Al-E'Mari, S. Fraihat, S. Naser Makhadmeh, and E. Alzubi, "AI-Powered Smart Grids in the 6G Era: A Comprehensive Survey on Security and Intelligent Energy Systems," *IEEE Open J. Commun. Soc.*, vol. 6, pp. 7677–7719, 2025, doi: 10.1109/OJCOMS.2025.3609144.

[14] J. Cai, M. Song, C. Yu, and T. Yu, "Resource Allocation Algorithm Based on D3QN for Ultra-Large-Scale Connection Scenarios in the 6G Context," in *2024 6th International Conference on Communications, Information System and Computer Engineering, CISCE 2024*, 2024. doi: 10.1109/CISCE62493.2024.10653416.

[15] H. M. F. Noman, K. Dimyati, K. A. Noordin, E. Hanafi, and A. Abdrabou, "FeDRL-D2D: Federated Deep Reinforcement Learning- Empowered Resource Allocation Scheme for Energy Efficiency Maximization in D2D-Assisted 6G Networks," *IEEE Access*, vol. 12, pp. 109775–109792, 2024, doi: 10.1109/ACCESS.2024.3434619.

[16] P. Parimi, "Telecom Network Dataset," Kaggle.

[17] A. A. Puspitasari, T. T. An, M. H. Alsharif, and B. M. Lee, "Emerging Technologies for 6G Communication Networks: Machine Learning Approaches," *Sensors*, vol. 23, no. 18, p. 7709, Sep. 2023, doi: 10.3390/s23187709.

[18] F. Alhaqui, I. Lahsen-Cherif, M. Elkhechafi, and A. Elkhadimi, "Deep Learning for Telecom Self-Optimized Networks: Benefits and Implications," *IEEE Access*, vol. 12, pp. 195422–195435, 2024, doi: 10.1109/ACCESS.2024.3521282.

[19] C. Fei and L. XiuMei, "Research on ETF prediction based on Bi-LSTM-Bi-GRU," *Financ. Res. Open*, vol. 2, no. 1, p. 100087, Mar. 2026, doi: 10.1016/j.finr.2025.100087.

[20] N. Xiao and Z. Song, "Interference Signal Suppression Algorithm Based on CNN-LSTM Model," *Sensors*, vol. 25, no. 16, p. 5048, Aug. 2025, doi: 10.3390/s25165048.

[21] J. Sun, P. Han, Z. Cheng, E. Wu, and W. Wang, "Transformer Based Multi-Grained Attention Network for Aspect-Based Sentiment Analysis," *IEEE Access*, vol. 8, pp. 211152–211163, 2020, doi: 10.1109/ACCESS.2020.3039470.

[22] U. Korat and A. Alimohammad, "Efficient Hardware Implementation of Eigen Solver," in *2026 IEEE 16th Annual Computing and Communication Workshop and Conference (CCWC)*, IEEE, Jan. 2026, pp. 1376–1385. doi: 10.1109/CCWC67433.2026.11393888.

[23] J. Zhang, N. Xie, X. Zhang, K. Yue, W. Li, and D. Kumar, "Machine learning based resource allocation of cloud computing in auction," *Comput. Mater. Contin.*, 2018, doi: 10.3970/cmc.2018.03728.

[24] E.V.N. Jyothi *et al.*, "Machine Learning-Based Optimization for 5G Resource Allocation Using Classification and Regression Techniques," *Int. J. Comput. Exp. Sci. Eng.*, vol. 11, no. 2, pp. 2525–2540, Apr. 2025, doi: 10.22399/ijcesen.1657.

[25] S. D. Pasham, "Dynamic Resource Provisioning in Cloud Environments Using Predictive Analytics," *Int. J. Eng. Comput. Sci.*, vol. 7, no. 01, pp. 23517–23520, Jan. 2018, doi: 10.18535/ijecs/v7i1.09.