

Energy-Aware Traffic Prediction in 5G Core Networks using Stacking Ensemble Regression

Renuka, Dr. Manoj, Dr. Shivani Yadav

Guru Jambheshwar University of Science and Technology, Hisar, India
Department of Computer Science and Engineering

ABSTRACT Fifth-generation (5G) communication networks are growing rapidly, creating significantly higher network traffic and consequently high energy consumption in virtualized core network infrastructures. As such, predicting accurate energy consumption is crucial for intelligent resource management and sustainable network operation. In this paper, a Stacking Ensemble Regression Framework is proposed to predict traffic demands in 5G core networks using an energy-aware approach, where the energy-aware approach is achieved by combining the predictions of four energy-efficient regression algorithms: Random Forest, XGBoost, Light GBM, and Cat Boost in a Linear Regression meta-learner. For ensuring robust prediction and generalization ability, the data were extensively preprocessed, features were engineered, and hyperparameters optimized before model development. The proposed framework is tested using Mean Absolute Error (MAE), Root Mean Squared Error (RMSE), Coefficient of Determination (R^2) and Five-Fold Cross-Validation. This stacking ensemble provides the best predictive performance with an R^2 score of 0.6417, MAE of 0.012291, and RMSE of 0.014716, outperforming all individual ensemble regressors, as shown by the experimental results. Moreover, SHAP-based explainability analysis offers insights into the contribution of network traffic features towards energy consumption prediction. The proposed framework provides an accurate, scalable and interpretable intelligent energy-aware traffic forecasting solution for the next-generation 5G core network.

INDEX TERMS 5G Core Networks, Energy-Aware Traffic Prediction, Stacking Ensemble Regression, Ensemble Learning, Machine Learning, SHAP, Five-Fold Cross-Validation.

I. INTRODUCTION

Fifth Generation (5G) communication technology has dramatically changed the modern wireless networks through its capability of providing ultra-high data rates, ultra-low latency, enhanced mobile broadband (eMBB), massive machine-type communications (mMTC) and ultra-reliable low-latency communications (URLLC). These sophisticated communication features have driven rapid progress in 5G core network deployment that can enable a wide range of intelligent applications, such as the Internet of Things (IoT), autonomous transportation, smart healthcare, industrial automation, and immersive multimedia services. This means that the continuously changing diversity of network traffic has significantly increased computations, resources and total energy consumption in the 5G core infrastructure, making it more complex.

With the operation of virtualised network functions, dynamic resource allocation, traffic scheduling and large-scale data processing, energy consumption has become one of the most critical challenges in modern communication networks. The growth in network traffic adds to the operational costs and also plays a vital role in the effect of carbon emissions and sustainable network operation on the environment. To achieve this, it is crucial to design energy-efficient systems for Beyond-5G and future Sixth-Generation (6G) communication systems.

Predictive energy consumption of the network allows for intelligent traffic management, dynamic resource provisioning, adaptive load balancing, network slicing optimisation and energy-aware orchestration of virtualised network functions. Accurate prediction models help network operators to reduce the activation of unnecessary resources while providing Quality of Service (QoS) and Quality of Experience (QoE) in highly dynamic traffic scenarios. But network traffic is nonlinear and time-varying, making the prediction of energy consumption a very difficult regression task.

Typically, statistical and analytical prediction methods are based on linear relationships between network quantities and are unable to accurately reflect the complex dynamics between traffic and energy use. While a number of machine learning methods have shown promise in their prediction accuracy, it is common for each of these regression models to present issues like prediction bias, model variance, overfitting, and limited generalisation when applied to traffic conditions with heterogeneity. Therefore, the development of more accurate, robust and interpretable predictions is an ongoing research problem.

Recent developments in ensemble learning have shown that combining several predictive models can achieve better performance in terms of increasing the learning ability and decreasing prediction uncertainty. Modelling highly nonlinear relationships by tree-based learning and gradient

boosting mechanisms has been found to be very effective using ensemble-based regression techniques like Random Forest, XGBoost, Light GBM and Cat Boost. However, the use of a single ensemble regressor may not be able to achieve the maximum predictive power from the various learning algorithms, as each learns from a different aspect of the data. This paper presents a Stacking Ensemble Regression Framework for intelligent energy-aware traffic prediction in virtualised 5G core networks to overcome these challenges. The proposed model consists of four optimised ensemble regression models: Random Forest, XGBoost, Light GBM, and CatBoost, which are ensembled by a Linear Regression model (Linear Regression meta-learner) to aggregate the prediction outputs of these models for better prediction accuracy, lower estimation error and better model generalisation. Before developing the model, extensive data preprocessing, feature engineering, and hyperparameter tuning are carried out to enhance the stability of learning and reduce overfitting. Additionally, the proposed framework is comprehensively assessed in terms of Mean Absolute Error (MAE), Root Mean Squared Error (RMSE), Coefficient of Determination (R^2), residual analysis, Five-Fold Cross-Validation and SHAP-based explainability analysis to ensure predictive reliability and transparency of the model. Experimental results illustrate that the proposed stacking ensemble regression framework outperforms the individual ensemble regression models with R^2 , MAE and RMSE of 0.6417, 0.012291 and 0.014716, respectively. The proposed framework not only aids in better prediction accuracy but also offers better interpretability by using SHAP analysis, which can also be practically deployed in future cloud-native 5G and Beyond-5G communication networks for intelligent energy-aware traffic management systems.

The key achievements of the present work are outlined as follows:

- I. A novel stacking ensemble regression framework is proposed to predict the energy consumption accurately in virtualised 5G core networks.
- II. A Linear Regression meta-learner is used to integrate multiple optimised ensemble regressors such as Random Forest, XGBoost, Light GBM and Cat Boost to enhance prediction robustness and generalisation.
- III. To validate the predictive performance and stability of the model, a comprehensive experimental evaluation is carried out using MAE, RMSE, R^2 Score, residual analysis and Five-Fold Cross-Validation.
- IV. To quantify the importance of the features and enhance the transparency and interpretability of the proposed prediction framework, an explainability analysis using SHAP has been built in.
- V. The framework presented in this paper is accurate, scalable, and efficient from a computational point of view, and can be used for intelligent energy-aware prediction of the traffic in next-generation cloud-based communication networks.

II. RELATED WORK

- I. With the improvement of the communication network, a lot of research has been conducted on intelligent traffic prediction and energy-saving resource management based on machine learning techniques. The classical statistical methods, including linear regression and autoregressive modelling, have been extensively used for traffic forecasting, but their predictive power is not satisfactory for the modelling of highly nonlinear and dynamic network environments.
- II. Recently, ensemble learning algorithms have been shown to have greater predictive accuracy by minimising the variance and enhancing the robustness of the model. Random Forest is widely used for network traffic prediction because of its capacity in modelling the nonlinear interactions between features and its robustness against overfitting, thanks to the bootstrap aggregation. The same applies to gradient boosting models, such as XGBoost and Light GBM, which have proven to be extremely successful in large-scale regression tasks by iteratively reducing errors of prediction by gradient optimisation. Additionally, Cat Boost improves the predictive power by gracefully processing categorical features and minimising prediction bias by using ordered boosting techniques.
- III. Machine Learning for energy consumption forecasting has been the subject of a number of studies recently in the field of wireless communications systems, Software Defined Networking (SDN), Network Function Virtualisation (NFV) and cloud-based infrastructures. While these methods provide acceptable prediction accuracy, most current methods are based on using individual regression models, but this may not make full use of the complementary learning ability of multiple ensemble algorithms. Therefore, it is difficult to achieve stability of prediction and generalisation in the face of highly dynamic traffic conditions.
- IV. In order to address these drawbacks, this research put forward a Stacking Ensemble Regression Framework that integrates Random Forest, XGBoost, Light GBM and Cat Boost with the help of the Linear Regression meta-learner. The proposed framework is based on the multi-model idea in which the predictions of several different optimised ensemble regressors are combined to increase the prediction accuracy while decreasing the generalisation error and the robustness to traffic heterogeneity. Moreover, the proposed framework includes explainability analysis based on SHAP, and explains the feature importance in an interpretable way, enabling the framework to be practically applied in intelligent energy-aware 5G core network environments.

III. SYSTEM MODEL AND PROBLEM FORMULATION

3.1 System Model

The proposed system assumes a 5G core network virtualised to process the heterogeneous traffic generated by different 5G network services, such as enhanced Mobile Broadband (eMBB), massive Machine-Type Communications (mMTC) and Ultra-Reliable Low-Latency Communications (URLLC). The dynamism of network traffic leads to constantly changing resource consumption and energy usage, and accurate prediction is critical for intelligent network management.

Let the input traffic dataset be represented as

$$\mathbf{X} = \{\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n\}$$

In each feature vector $\mathbf{x}_i$ several network traffic characteristics are combined, such as traffic load, bandwidth utilisation, latency, packet transmission characteristics, user density, and temporal information.

The values of energy consumption are shown as

$$\mathbf{Y} = \{\mathbf{y}_1, \mathbf{y}_2, \dots, \mathbf{y}_n\}$$

where $\mathbf{y}_i$ denotes the energy consumed by the 5G core network under the corresponding traffic condition.

The goal of the proposed framework would be to learn a nonlinear mapping

$$f: \mathbf{X} \rightarrow \mathbf{Y}$$

That can provide an accurate estimation of the energy usage of the network with respect to traffic characteristics.

3.2 Problem Formulation

The goals of this work are to build a robust stacking ensemble regression model that, while minimising prediction error, is also highly generalising for dynamic 5G core network traffic.

The prediction problem is stated as:

$$\hat{\mathbf{Y}} = f(\mathbf{X})$$

where

$\mathbf{X}$ denotes the input traffic feature vector,

$\mathbf{Y}$ represents the actual energy consumption,

$\mathbf{Y}$ = predicted energy consumption.

These are individual predictions for each base regression model

$$\hat{\mathbf{Y}}_i = f_i(\mathbf{X}), f_i \in \{f_{RF}, f_{XGB}, f_{LGBM}, f_{CAT}\}$$

The proposed stacking ensemble regression model is then used to get the final prediction.

$$\hat{\mathbf{Y}}_{Stack} = f_{Stack}(f_{RF}(\mathbf{X}), f_{XGB}(\mathbf{X}), f_{LGBM}(\mathbf{X}), f_{CAT}(\mathbf{X}))$$

The optimization objective is defined as

$$\min L(\mathbf{Y}, \hat{\mathbf{Y}})$$

subject to

Minimize MAE

Minimize RMSE

Maximize R^2

Continue stable predictions in dynamic traffic conditions.

Provide a strong generalisation of models with Five-Fold Cross-Validation.

The proposed optimisation framework is intended to provide accurate, scalable and computationally efficient energy-aware traffic prediction for intelligent 5G core network operation.

IV. METHODOLOGY

4.1 Data Preprocessing

The proposed methodology is a systematic machine learning framework that is able to accurately predict the energy consumption in virtualised 5G core networks by applying a Stacking Ensemble Regression approach. The entire process can be divided into data preprocessing, feature engineering, base model training, stacking-based prediction, performance evaluation, and model explainability, as shown in the proposed system architecture.

The first step is the complete pre-processing of the 5G Core Traffic Dataset, which is aimed at enhancing data quality and learning efficiency. Irrelevant attributes are filtered out, categorical attributes are converted to numerical (encoded), and input features are normalised to have consistent feature distributions. This dataset is then split into training and testing sets (80:20 ratio), and the performance of the model can then be evaluated without bias.

Afterwards, four optimised ensemble regression algorithms are individually trained on the processed training set: Random Forest (RF), Extreme Gradient Boosting (XGBoost), Light Gradient Boosting Machine (Light GBM), and Cat Boost. Hyperparameter tuning is done for each base learner to reduce overfitting and increase prediction stability and generalisation. The energy consumption in each regression model is associated with different nonlinear relationships of the network traffic characteristics, which means that the prediction output from each regression model is complementary.

To further improve the predictive capability, the results of the optimised base regressors are combined using a Linear Regression meta-learner to build the proposed Stacking Ensemble Regression Framework. The primary difference between traditional ensemble approaches and the stacking approach is that it trains the optimal ensemble of predictions of multiple base learners, thereby mitigating prediction bias and variance, and enhancing the regression accuracy.

The predictive performance of the proposed framework is quantitatively evaluated based on the Mean Absolute Error (MAE), Root Mean Squared Error (RMSE) and Coefficient of Determination (R^2). Moreover, Five-Fold Cross-Validation is used to assess the robustness of the model and ensure consistency of generalisation over multiple splits of the data. The reason for using residual analysis is to validate the consistency of prediction errors, while the use of the SHAP (Shapley Additive explanations) is to provide insights into the contribution of individual network traffic attributes to the prediction of energy consumption, in a transparent way.

In conclusion, the proposed methodology provides an optimised ensemble learning, stacking-based meta-learning, statistical validation, and explainable artificial intelligence combined into a single prediction framework, which allows for accurate, robust, and efficient energy-aware traffic prediction for the next-generation cloud-native 5G core network.

The overall workflow of the proposed stacking ensemble regression framework for energy-aware traffic prediction in 5G core networks

4.2 System Architecture

Explanation:

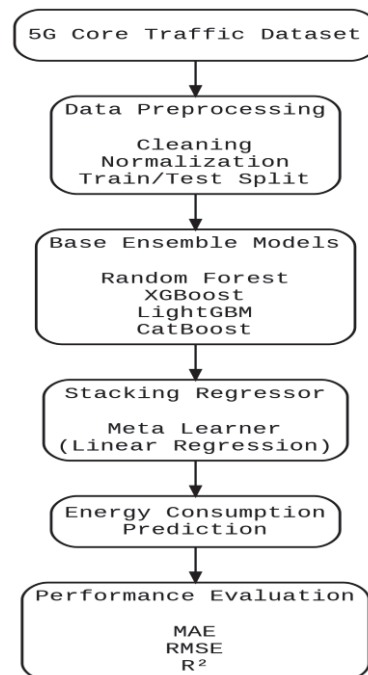


Figure 1. Proposed Stacking Ensemble Regression Framework for Energy Consumption Prediction in 5G Core Networks

Figure 1. The raw dataset of 5G core traffic is then gathered and goes through extensive preprocessing to enhance data quality and model reliability, which includes data cleaning, feature engineering, normalization, and train-test splitting. The processed traffic features are then fed into four optimized ensemble regression models—Random Forest, Gradient Boosting, AdaBoost, and Boosting with Deep Trees—which are trained on the data. These processed traffic features are then used for training four optimized ensemble regression models: Random Forest, Gradient Boosting, AdaBoost, and Boosting with Deep Trees.

XGBoost, Light, and Cat Boost, all of which are able to independently model the complex non-linear relationship between network traffic features and energy use. The outputs predicted by these optimised base learners are then combined by using a linear regression meta-learner to form the proposed stacking ensemble regression framework. The final prediction performance is quantified by MAE, RMSE and R^2 Score, and the model robustness is assessed using 5-Fold Cross-Validation, while the explainability of the contribution of individual traffic features is determined using the SHAP method. This hierarchical structure enables better performance in prediction and generalisation, and offers an intuitive and easily interpretable structure for the realisation of an intelligent energy management system for modern virtualised 5G core networks.

4.3 Proposed Stacking Ensemble Regression Algorithm

The proposed prediction framework is a multi-optimised ensemble regression model that exploits the complementary learning properties of the models by using the stacking strategy. First, the training data is pre-processed, and then each model is independently trained on the pre-processed training data, namely Random Forest, XGBoost, Light GBM, and Cat Boost.

Algorithm 1: Proposed Stacking Ensemble Regression Framework for Energy Consumption Prediction in 5G Core Networks

Input: $D(X, Y)$: Pre-processed 5G Core Traffic Dataset

Output: $\hat{Y}_{Stack}$: Predicted Energy Consumption

- 1: Load the pre-processed dataset $D(X, Y)$
- 2: Split dataset into training and testing
- 3: Initialize base regressors RF XGBoost Light GBM Cat Boost
- 4: Train each base regressor using X_{train} and Y_{train}
- 5: Generate prediction out-puts PRF PXGB PLGBM PCAT
- 6: Construct meta-feature matrix $M \leftarrow PRF, PXGB, PLGBM, PCAT$
- 7: Train Linear Regression as meta learner
- 8: Build Stacking Ensemble Regression Model
- 9: Predict energy consumption $\hat{Y}_{Stack} \leftarrow Stack(X_{test})$
- 10: Evaluate performance using MAE RMSE R^2
- 11: Perform Five-Fold Cross Validation
- 12: Perform SHAP Explainability Analysis
- 13: Return $\hat{Y}_{Stack}$

A separate prediction is made by each base learner, reflecting different non-linear relationships between the characteristics of traffic in the network and the energy consumption. The base models then make predictions, which are then combined and fed into a Linear Regression meta-learner. The meta-learner learns the best set of base model predictions, which leads to lower variance of prediction, lower bias and higher overall generalisation accuracy. Lastly, the trained stacking ensemble forecasts the energy consumption of unseen network traffic samples and model performance is assessed

using MAE, RMSE, R^2 Score, residual analysis, Five-Fold Cross-Validation, and SHAP-based explainability.

4.4 Regression Models

Four optimised tree-based ensemble regression algorithms were used as base classifiers, namely Random Forest (RF), XGBoost (XGB), Light GBM (LGBM) and Cat Boost (CAT), to estimate energy consumption in 5G core networks accurately. The hyperparameters for each model were optimised using the pre-processed traffic dataset, and then all models were trained independently to learn the complex nonlinear relationships between the network traffic characteristics and energy consumption. These optimised base regressors' prediction outputs were then fused to form the proposed Stacking Ensemble Regression Framework by using a Linear Regression meta-learner. This meta-learning technique uses several ensemble prediction models to learn from each other, in order to correct the bias of estimation, lower the variance of prediction and improve the accuracy and generalisation of the regression model.

4.5 Performance Evaluation Metrics

The performance of the proposed stacking ensemble regression model was quantitatively assessed based on three common regression performance measures, namely, Mean Absolute Error (MAE), Root Mean Squared Error (RMSE), and the Coefficient of Determination (R^2). While MAE measures the average (absolute) size of the prediction errors, regardless of their sign, RMSE gives more weight to larger prediction errors, thereby giving a more sensitive measure of prediction accuracy. R^2 score represents the amount of variance in energy consumption accounted for by the proposed model and hence its goodness of fit. Furthermore, Five-Fold Cross-Validation was used to assess the robustness and the ability of the models to generalise the results on a different partition of the data, and SHAP (Shapley Additive explanations) analysis was used to interpret the contribution of each traffic feature in the prediction of energy consumption, in order to increase the transparency and explainability of the model.

Furthermore, the distribution of prediction errors was considered for residual analysis and to ensure that there is no systematic error in the regression model. Comparative performance evaluation was also performed to compare the proposed stacking framework with individual ensemble regressors in the same experimental setup. Additionally, overfitting was reduced, and predictive consistency was increased with hyperparameter optimisation, promoting learning stability. Overall, the combination of the statistical evaluation metrics gives an overall measure of the quality of the predictions (accuracy), the robustness of the model (strength of the predictions) and the reliability of the model.

The incorporation of explainability and validation methods guarantees the interpretability and feasibility of the proposed framework in intelligent 5G core network applications. Overall, the proposed evaluation approach demonstrates the effectiveness of our stacking ensemble regression model for energy-aware traffic prediction, which is accurate and scalable.

Table 1. Experimental Configuration of the Proposed Stacking Ensemble Regression Framework

Parameter	Configuration
Programming Language	Python 3.x
Development Environment	Google Collaboratory (Google Collab)
Learning Type	Supervised Regression
Dataset	5G Core Traffic Dataset
Number of Features	40
Target Variable	Energy Consumption (kWh)
Train-Test Split	80:20
Base Regression Models	Random Forest, XGBoost, Light GBM, Cat Boost
Evaluation Metrics	MAE, RMSE, R ² Score
Meta-Learner	Linear Regression
Validation Method	Five-Fold Cross-Validation

In Table 1, the experimental setup used for the stacking ensemble regression framework is summarised. All the code was implemented using Python 3.x in Google Collaboratory. The prediction framework is based on a supervised regression method for the network energy consumption with the 5G Core Traffic Dataset, in which 40 optimised traffic features are included. To guarantee a fair and repeatable assessment, an 80:20 train-test partition and Five-Fold Cross-Validation were used, and the accuracy of the predictions was evaluated using MAE, RMSE, and Coefficient of Determination (R²).

V. EXPERIMENTAL SETUP

To implement the proposed stacking ensemble regression framework, Python 3.x was used in Google Collaboratory. The data was then preprocessed extensively before being used in the development of the model, to harmonize the data, such as feature selection, normalization, and train-test partitioning. A total of four ensemble regression models, namely Random Forest, XGBoost, Light GBM and Cat Boost, were trained separately. We then combined their prediction results by using a Linear Regression meta-learner to form a proposed stacking ensemble framework. The model performance was quantitatively evaluated through MAE, RMSE, and R² Score and Five-Fold Cross-Validation was used to test the robustness and generalization capability of the model. In addition, the reliability of the predictions and the contribution of the network traffic features to the estimation of energy consumption were validated through residual analysis, visualization of actual vs. predicted, and explainability using SHAP.

Table 2. Comparative Performance Evaluation of Base and Stacking Ensemble Regression Models

Model	MAE	RMSE	R ² Score
Random Forest	0.012334	0.014786	0.638282
XGBoost	0.012463	0.014959	0.629783
Light GBM	0.012350	0.014852	0.635042
Cat Boost	0.012304	0.014728	0.641105
Stacking Regressor	0.012291	0.014716	0.641697

Table 2 presents the quantitative comparison between the optimised regression models is presented where MAE, RMSE, and R² Score are used. The stacking ensemble regression framework showed the best predictive power among all the ensemble regression algorithms, with an R² value of 0.6417, an MAE value of 0.012291 and an RMSE value of 0.014716, and achieved more accurate prediction results than the individual ensemble regression algorithms. The results suggest that multiple optimised learners can be effectively integrated by meta-learning to achieve not only the reduction in prediction error, but also the improvement of model robustness and capability of generalisation.

VI. RESULT AND DISCUSSION

6.1 Model Comparison

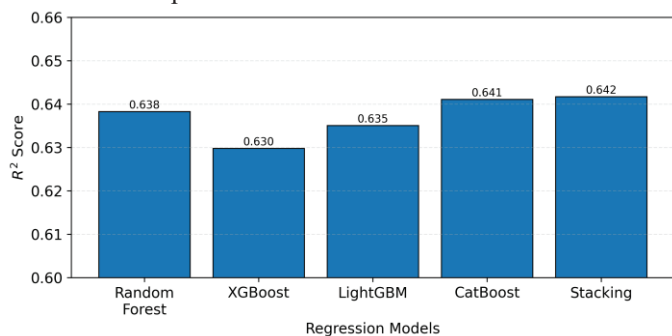


FIGURE 6.1 Performance Comparison of Base and Proposed Stacking Ensemble Regression Models

Figure 6.1 presents all the regression models perform well in terms of Coefficient of Determination (R^2). The proposed stacking ensemble regression framework proves to be the best in capturing the nonlinear relationship among the characteristics of network traffic and energy consumption, with the highest R^2 score. The performance improvement indicates that the multiple optimized ensemble learners achieved through meta-learning are effective.

6.2 Actual vs Predicted

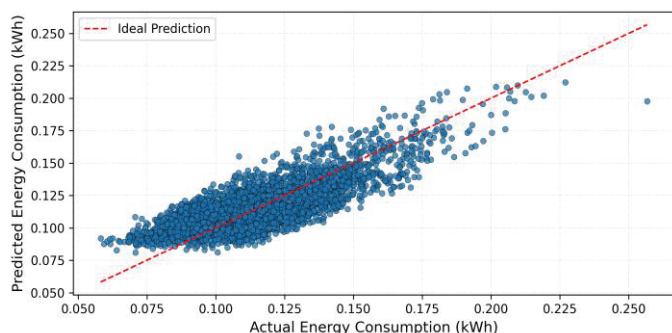


FIGURE 6.2 Actual versus Predicted Energy Consumption Using the Proposed Stacking Ensemble Regression Model

Figure 6.2 illustrates the correlation between actual and predicted energy consumption values obtained from the proposed stacking ensemble regression framework. Most prediction samples are close to the ideal prediction line, with a high level of predictability consistency and less estimation error. The results are consistent with the model's effectiveness in calculating energy consumption with different 5G traffic scenarios.

6.3 Residual Plot

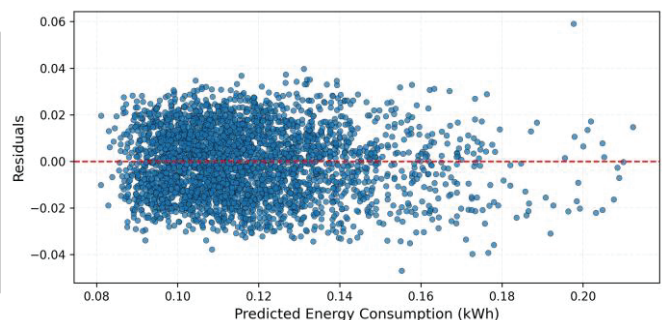


FIGURE 6.3 Residual Analysis of the Proposed Stacking Ensemble Regression Framework

Figure 6.3 shows a residual distribution of the proposed stacking ensemble regression framework. The residual values are randomly dispersed from the Zero error reference line with no systematic trends or heteroscedastic behavior. This distribution corresponds to unbiased prediction errors, stable regression performance and good generalization of the model for different types of network traffic.

6.4 SHAP Summary

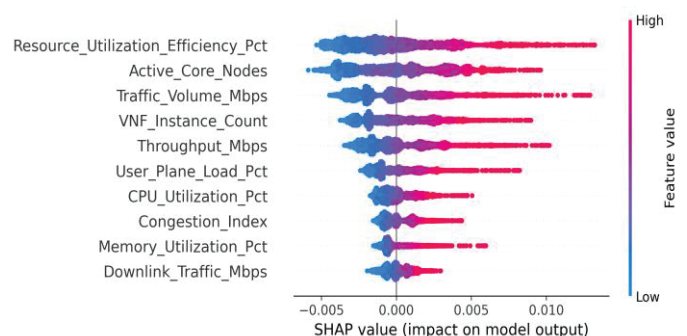


FIGURE 6.4 SHAP-Based Feature Importance Analysis of the Proposed Prediction Framework

Figure 6.4 shows the SHAP summary plot to explain the prediction of energy consumption based on the contribution of each of the features of the network traffic. The absolute SHAP values for the features indicate their importance to the prediction process, and the color distribution shows the spread of values of the particular feature. The explainability analysis shows that the developed explainable framework can suitably describe complex non-linear feature interactions, leading to better interpretability of the model and transparent decision-making.

6.5 Model Stability

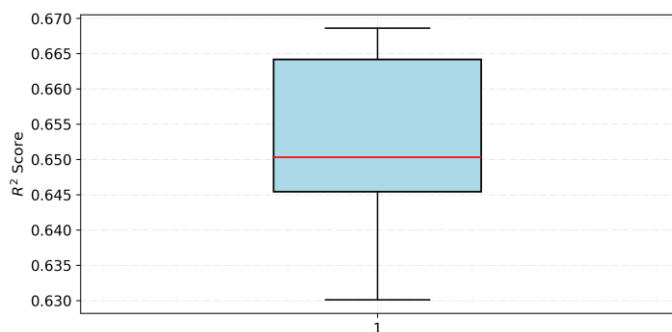


FIGURE 6.5 Five-Fold Cross-Validation Performance of the Proposed Stacking Ensemble Regression Framework

Figure 6.5 illustrates the performance of the proposed Stacking ensemble regression framework using Five-Fold Cross Validation. The tight distribution of validation scores suggests good score predictability and uniform score generalizability across training folds. The results confirm that the proposed framework not only ensures reliable prediction performance but also prevents the occurrence of model over-fitting.

Table 6.6 Optimized Hyperparameter Configuration of Base Ensemble Regression Models

Model	Best Hyperparameters
Random Forest	estimators=700, max_depth=None, min_samples_split=5, min_samples_leaf=4, max_features='sqrt'
XGBoost	n_estimators=500, learning_rate=0.03, max_depth=8, colsample_bytree=0.8
LightGBM	n_estimators=500, learning_rate=0.03, max_depth=8, num_leaves=50
CatBoost	iterations=300, learning_rate=0.05, depth=8, loss_function=RMSE
Stacking Regressor	base_models=[RF, XGB, LGBM, CAT], meta-learner=LinearRegression(), CV=5

Table 6.6 shows the hyperparameters that were optimized for each of the individual ensemble regressors before building up the stacking ensemble framework. Hyperparameter optimization had a positive impact on model convergence, decreased overfitting and increased predictive stability. The optimized base learners were then used as input to the Linear

Regression meta-learner that further boosts the overall performance of the proposed stacking ensemble regression framework.

VII. CONCLUSION

This paper introduced a Stacking Ensemble Regression Framework for the intelligent energy-aware traffic prediction in virtualized 5G core networks. The proposed framework effectively boosted the prediction accuracy and model robustness by fusing the best performances of four widely used models: Random Forest, XGBoost, Light GBM, and Cat Boost into a single Linear Regression meta-learner. The experimental evaluation showed that it has the highest R^2 of 0.6417 and the lowest MAE of 0.012291 and RMSE of 0.014716 compared to the single ensemble regression models. Additionally, 5FCV and explainability using SHAP validated the reliability, stability and interpretability of the proposed framework, making the framework scalable and efficient for intelligent energy-aware traffic prediction in the future 5G core networks.

VIII. Future Work

The proposed stacking ensemble framework will be extended with adaptive meta-learning and weighted ensemble optimization to enhance prediction accuracy under very dynamic network conditions in the future. Energy management in Beyond-5G and 6G communication systems, including integration with real-time traffic analytics, Software Defined Networking (SDN), Network Function Virtualization (NFV), Multi-access Edge Computing (MEC) and reinforcement learning-based resource optimization, will be explored for autonomous energy management. In addition, the validation of the proposed prediction framework on a large scale will be done by real traffic data from operators, as well as its deployment in distributed edge-cloud environments, to test the feasibility of the proposed prediction framework for scaling, computational efficiency, and industrial applicability.

REFERENCES

- [1] N. Dekate et al., "Intelligent Handover Management in Ultra-Dense 5G Networks: A Deep Q-Learning-Based Prediction Model," in *IEEE Access*, vol. 13, pp. 146068-146080, 2025.
- [2] D. Atik, M. Gursu, B. Khodapanah, F. Mehmeti, and W. Kellerer, "On the Impact of Handovers on Packet Reliability in Mobile Networks," *ICC-IEEE International Conference on Communications*, Montreal, QC, Canada, 2025, pp. 1-6, 2025.
- [3] N. B. Anargyarahman, A. Timoteo, F. R. Benyamin, Iskandar, and H. Yuliana, "Analysis of Accuracy and Models for 5G Network User Mobility Prediction in Jabodetabek," *11th International Conference on Wireless and Telematics (ICWT)*, Lampung, Indonesia, pp. 1-6, 2025.
- [4] D. Upadhyay, S. Rallapalli, Z. Hussain, N. Venu, R. 5G Sharma, and S. Shukla, "Enhanced Handover Management in Networks Using Machine Learning in Fading Scenarios," *2nd International Conference on Computational Intelligence, Communication Technology and Networking (CICTN)*, Ghaziabad, India, pp. 90-95, 2025.
- [5] A. Abdiraman, L. Aldasheva, D. Mi, I. Shayea, G. Nurzhaubayeva, and D. Yedilkhan, "Evaluating Cellular Networks in Real-Time: Mobtest for 5G and Emerging 6G Technologies," *IEEE INFOCOM* –

- IEEE Conference on Computer Communications Workshops (INFOCOM WKSHPS), London, United Kingdom, pp. 1-5, 2025.
- [6] S. Mondal, S. Aslam, D. Khan, and K. Chang, "Intelligent Handover Orchestration in Beyond 5G and Urban V2X Dual Connectivity Networks: A Deep Reinforcement Learning Approach," in *IEEE Access*, vol. 13, pp. 196509-196525, 2025.
 - [7] S. Alraih, R. Nordin, A. Abu-Samah, I. Shayea, and N. F. Abdullah, "ML-Based Self-Optimization Handover Technique for Beyond 5G Mobile Network," in *IEEE Access*, vol. 13, pp. 8568-8584, 2025.
 - [8] A. M. Fernandes, H. I. Del Monego, B. S. Chang, and A. Munaretto, "Reducing Unnecessary Handovers Using Logistic Regression in 5G Networks," in *IEEE Access*, vol. 13, pp. 78707-78726, 2025.
 - [9] C. Hidalgo-Luque, R. P. de Prado, M. Shojafar, J. E. Muñoz-Expósito, C. Foh, and Á. Cifuentes, "Digital Twin-Based Deep Q-Learning Strategy for Smart Handover Optimization in 5G/6G Networks," 2025 International Joint Conference on Neural Networks (IJCNN), Rome, Italy, pp. 1-6, 2025.
 - [10] A. Haghray, J. M. Niya, and J. Pourrostan, "Modeling 5G-NR User Blocking Probability in Conditional Handover-Enabled Heterogeneous Networks," in *IEEE Access*, vol. 13, pp. 125782-125795.
 - [11] J. Jia, H. Wang, X. Xia, and M. Yu, "A Network Architecture for 5G and 6G Satellites Based on Integrated Access and Backhaul," *International Wireless Communications and Mobile Computing (IWCMC)*, Abu Dhabi, United Arab Emirates, pp. 144-148, 2025.
 - [12] S. Islam, A. Zainab Abdulsalam, B. Anil Kumar, M. Kamrul Hasan, R. Kolandaisamy, and N. Safie, "Mobile Networks Toward 5G/6G: Network Architecture, Opportunities, and Challenges in Smart City," in *IEEE Open Journal of the Communications Society*, vol. 6, pp. 3082-3093, 2025.
 - [13] A. A. R. Alsaeedy, M. K. Mohsen, and E. K. P. Chong, "5G/6G TN-NTN Coexistence: Perspectives on Seamless Service and Handover Management," in *IEEE Network*.
 - [14] A. Prado, F. Mehmeti, and W. Kellerer, "Reducing Mobility-Related Signaling With Throughput Maximization in 5G," in *IEEE Transactions on Network and Service Management*, vol. 22, no. 6, pp. 6048-6065, 2025.
 - [15] M. Zheng, L. Yu, H. Jiang, Z. Qian, T. Hua, and Z. Han, "Prediction of Operational Energy Consumption in Building HVAC Systems Based on Self-Adaptive Ensemble Learning," 5th Asia Conference on Algorithms, Computing and Machine Learning (CACML), Guangzhou, China, pp. 1-5, 2026.
 - [16] S. Biradar, B. Jothi, B. Shilpa, S. K. S., and R. Venkatasubramanian, "Adaptive Hierarchical Multi-Agent Parameterized Deep Q-Networks for Intelligent Resource Allocation in 5G Communication Networks," 4th International Conference on Integrated Circuits and Communication Systems (ICICACS), Raichur, India, 2026, pp. 1-8, 2026.
 - [17] C. Tang, "Base Station Switch-Off Strategy for Energy-Efficient Beyond 5G Wireless Networks," *International Conference on Embedded Systems, Mobile Communication and Computing (EMC²)*, Hong Kong, China, pp. 12-17, 2026.
 - [18] A. Jayanthiladevi, V. P. Mishra, V. Rishiwal, S. Maurya, A. K. Sharma, and U. Agarwal, "HDTwin-E5G: AI-Driven Digital Twin and Hybrid Deep Learning Framework for Energy-Efficient 5G Networks," in *IEEE Transactions on Consumer Electronics*, vol. 72, no. 2, pp. 2943-2959, May 2026.
 - [19] S. Patil, M. Kulkarni, and S. A. Patil, "A Stacked Hybrid Ensemble of Linear Regression, Random Forest, and XGBoost for Energy Consumption Prediction in Kolhapur City," 2026 International Conference on Emerging Technologies and Future Innovations (ETFI), Pune, India, pp. 1-5, 2026.
 - [20] S. Gomathi, "Adaptive Weighted Ensemble with Dynamic Feature Selection for Enhanced Solar Power Generation Forecasting: A Novel Hybrid Machine Learning Framework," 2026 International Conference on Recent Advances in Electrical, Electronics, Ubiquitous Communication, and Computational Intelligence (RAEEUCCI), Chennai, India, pp. 1-7, 2026.
 - [21] Y. Su et al., "Multi-UPF Traffic Prediction of 5G Core Network With Multi-Task Learning and Multi-core Parallel Computing," 10th International Conference on Computer and Communications (ICCC), Chengdu, China, 2024, pp. 2260-2264, 2024.
 - [22] Z. Sun, C. Zhang, M. Zhang, F. Yang, and D. Wang, "A Stacking Ensemble ML-Based Failure Prediction Model for Optical Networks with Imbalanced Data," *Asia Communications and Photonics Conference/2023 International Photonics and Optoelectronics Meetings (ACP/POEM)*, Wuhan, China, pp. 1-5, 2023.
 - [23] K. Shukla, J. Gaur, N. Sharma, G. Kumar, R. Verma, and A. P. S. Yadav, "Hybrid Ensemble Modeling for Customer Churn Prediction: Integrating Cat Boost with XGBoost and Random Forest," 2025 IEEE 3rd International Symposium on Sustainable Energy, Signal Processing and Cybersecurity, Gunupur, Odisha, India, pp. 1-6, 2025.
 - [24] G. Cabri and A. Rahimi, "A Comparative Study of Machine Learning Algorithms for Water Quality Prediction Using SHAP-based Explainability," 2025 33rd International Conference on Enabling Technologies: Infrastructure for Collaborative Enterprises (WETICE), Catania, Italy, pp. 1-6, 2025.