A MACHINE LEARNING APPROACH FOR ENERGY CONSUMPTION IN 5G NETWORKS

Ilyasse Makroz [1], Ahmed Nafidi[1], Ramón Gutiérrez-Sánchez [2], Salaheddine Aourik[3]
[1]Hassan First University of Settat, National School of Applied Sciences, Department of Mathematics and Informatics, LAMSAD, Berrechid, Morocco, [2]Department of Statistics and Operational Research, Facultad de Ciencias, Campus de Fuentenueva, University of Granada, Granada, Spain, [3] MIET Laboratory, Faculty of Sciences and Techniques, Hassan First University of Settat, Settat 26002, Morocco

NOTE: Corresponding author: MAKROZ, ilyass.makroz@gmail.com

Abstract – *Energy consumption modeling in 5G networks is a complex task due to the variability in network configurations, traffic conditions, and the deployment of energy-saving techniques. Machine Learning (ML) offers a promising approach to address these challenges by leveraging data-driven insights for accurate predictions. This study aims to explore this idea by developing an ML-based model to predict energy consumption in 5G networks. We employ XGBoost, CatBoost, and Artificial Neural Networks (ANN), combined through a weighted average approach, to enhance prediction accuracy. Our findings indicate that the ensemble model significantly improves the estimation of energy consumption patterns, providing valuable insights for efficient energy management in 5G networks.*

Keywords – 5G, artificial neural network, energy consumption, machine learning

1. INTRODUCTION

5G, or fifth-generation, wireless technology represents a transformative leap forward in mobile communication, promising unprecedented speed, low latency, and enhanced connectivity. Operating on higher frequency bands, 5G networks deliver data speeds that surpass their predecessors, enabling users to download large files, stream high-definition content, and engage in real-time applications seamlessly [1]. The low latency of 5G facilitates responsive interactions critical for applications like augmented reality and virtual reality [2]. Moreover, 5G's increased capacity accommodates the burgeoning Internet of Things (IoT), connecting a myriad of devices simultaneously [3]. While heralding a new era of innovation with applications like autonomous vehicles and smart cities, 5G deployment faces challenges such as limited coverage, infrastructure requirements, and device compatibility issues [4, 5].

The deployment of 5G networks brings with it a noteworthy consideration in terms of energy consumption. While 5G technology is designed to be more energy-efficient compared to previous generations, the deployment of additional infrastructure, including a denser network of smaller cells and increased data processing capabilities, can still contribute to higher energy demands [6]. The deployment of massive Multiple Input Multiple Output (mMIMO) antennas and the need for increased processing power at the network edge may require additional power resources [7]. However, efforts are being made to address this concern through the development of more energy-efficient hardware, optimization of network architectures, and the exploration of renewable energy sources to power Base Stations (BSs) [8].

The significance of energy consumption in 5G networks becomes even more apparent when considering the financial implications for network operators. On average, OPerational EXpenditure (OPEX) related to network operations constitutes approximately 25 percent of the total costs borne by operators, with a substantial 90 percent of that expenditure allocated to covering substantial energy bills [9]. Within this context, the Radio Access Network (RAN) emerges as a major contributor, accounting for over 70 percent of the energy consumed, particularly by the energy-intensive BSs. In contrast, while still consuming energy, data centers and fiber transport components represent a smaller fraction of the overall energy footprint [10]. This underscores the critical need for continued efforts in optimizing the energy efficiency of 5G infrastructure to not only alleviate environmental concerns but also to manage the substantial operational costs associated with powering and maintaining these advanced networks.

In this work, we focus on developing a machine learning-based solution to estimate energy consumption in 5G networks, considering the engineering configurations, traffic conditions, and energy-saving methods. This research contributes to the telecom industry by providing a model that generalizes across various base station products and configurations, predicting energy consumption based on real network parameters.

The rest of the paper is organized as follows: Section 2 reviews related work, Section 3 details the datasets used, Section 4 presents our exploratory data analysis, Section 5 explains the feature engineering process, Section 6 introduces the evaluation metrics, Section 7 describes the modeling techniques, Section 8 discusses the results, and Section 9 concludes the paper.

2. RELATED WORK

Several works can be found in the literature addressing the problem of energy efficiency in 5G networks. [11] highlights the need for energy-efficient power control algorithms to reduce base station energy consumption, while [12] underscores the challenges of high power consumption in 5G networks. [13] further emphasizes the need for a balance between performance and energy consumption in 5G applications. [14] provides a broader perspective, discussing the impact of 5G on frequency spectrum and energy consumption, and the potential changes in mobile operators' strategies.

However, existing studies often focus on specific aspects of energy consumption, such as power control or network configurations, without providing a comprehensive model that generalizes across different base station products and configurations. For instance, in [15] the authors describe the power consumption model of a base station and formulate the optimization problem of network power consumption, focusing on specific configurations. Similarly, [16] presents an advanced power model supporting a broad range of network scenarios and base station types, yet it focuses on specific configurations. Additionally, [17] discusses various strategies and models for managing energy consumption in base stations, emphasizing the focus on specific configurations and control mechanisms. Furthermore, [18] discusses a model for energy consumption that can change from one base station to another, highlighting the lack of a comprehensive model that generalizes across different configurations. Our approach differs by integrating multiple machine learning algorithms to develop an ensemble model that predicts energy consumption based on a wide range of parameters, including engineering configurations, traffic conditions, and energy-saving methods. This comprehensive approach addresses the gap in the literature by providing a model that generalizes well across various scenarios.

3. DATASETS

The datasets utilized in this study were provided by the International Telecommunication Union (ITU) and Zindi as part of a competition aimed at predicting energy consumption in 5G networks. The data represents real-world measurements collected from operational 5G base stations over a specified period, encompassing various hardware configurations, user activities, and network conditions.

3.1 Data collection

The data was collected from a network of one thousand two hundred seventeen base stations located across different regions. These base stations include a mix of urban, suburban, and rural deployments, ensuring a diverse representation of network scenarios. The collection period spanned eight days, from July the first to July the eighth, two thousand twenty three, capturing both weekdays and weekends to account for variations in user behavior and traffic patterns.

3.2 Dataset details

The datasets comprise three CSV files:

1. **Configuration parameters and hardware attributes**: This file contains detailed information about the hardware and configuration of each base station. Each row represents a unique base station and includes the following columns:
   - *BS_name*: Unique identifier for each base station.
   - *cell_name*: Identifier for each cell within the base station.
   - *RUType*: The type of Radio Unit, indicating the base station product type (e.g., Macro Cell, Small Cell).
   - *Mode*: Transmission mode of the base station (e.g., Frequency Division Duplex - FDD, Time Division Duplex - TDD).
   - *Frequency*: Operating frequency band in MHz.
   - *Bandwidth*: Allocated bandwidth in MHz for each cell.
   - *Antennas*: Number of antennas deployed at the base station.
   - *TXpower*: Maximum transmit power in dBm for each cell.
2. **Cell-level data**: This file includes hourly counters for each cell, capturing dynamic operational metrics. The columns are:
   - *Time*: Timestamp indicating the date and hour of the measurement.
   - *BS_name*: Identifier for the base station.
   - *cell_name*: Identifier for the cell.
   - *Load*: The fraction of utilized resources (ranging from 0 to 1), representing the cell's traffic load during that hour.
   - *ESM_1* to *ESM_6*: Activation duration of six different Energy-Saving Methods (ESMs). Each value ranges from 0 to 1, indicating the fraction of the hour during which the specific ESM was active.
3. **Energy consumption data**: This file records the total energy consumed by each base station on an hourly basis. The columns include:
   - *Time*: Timestamp of the measurement.

- *BS_name*: Identifier for the base station.
- *Energy_Consumption*: Measured in kilowatt-hours (kWh), representing the energy consumed by the base station during that hour.

3.3 Data representation

The datasets represent a diverse range of hardware configurations and operational scenarios:

- **Hardware and configurations**:
  - *RUType*: Includes 12 different base station product types, varying in design and energy efficiency characteristics.
  - *Antennas*: The number of antennas ranges from 4 to 64, accommodating both conventional and massive MIMO (Multiple Input Multiple Output) configurations.
  - *Frequency bands*: Base stations operate across multiple frequency bands, including sub-1 GHz (e.g., 700 MHz) for wide coverage and mid-band frequencies (e.g., 1800 MHz, 2600 MHz) for higher capacity.
- **User activity and traffic conditions**:
  - *Load*: Captures real user traffic, with load values fluctuating based on time of day, user behavior, and geographic location.
  - *Temporal Patterns*: The data spans peak hours with high network demand and off-peak periods, allowing analysis of diurnal energy consumption patterns.
- **Energy Saving Mode (ESM)**:
  - *ESM Activation*: The six ESMs represent various techniques deployed to reduce energy consumption, such as shutting down carriers, symbols, or antenna elements during low traffic periods.
  - *Operational impact*: The activation of ESMs depends on network conditions and policies, affecting the base station's energy efficiency.

3.4 Data relevance and scope

The data reflects real-world operations of a 5G network, capturing the interplay between hardware configurations, traffic loads, energy-saving strategies, and energy consumption. This comprehensive dataset provides a robust foundation for developing predictive models that can generalize across different base station products and configurations.

3.5 Ethical considerations and data privacy

All data used in this study has been anonymized to protect sensitive information. No Personally Identifiable Information (PII) is included. The data collection and usage comply with relevant data protection regulations and ethical guidelines.

3.6 Data limitations

While the dataset is extensive, it represents a snapshot of eight days. Seasonal variations and longer-term trends are not captured. Additionally, some base station products and configurations may be underrepresented, posing challenges for model generalization, which we address in our modeling approach.

4. EXPLORATORY DATA ANALYSIS

In this section, we aim to gain insights into the characteristics of the provided datasets through Exploratory Data Analysis (EDA). The analysis involves a comprehensive examination of the variables and their distributions, as well as the relationships between key features. Our first observation unveils a discernible cyclic pattern in energy consumption, as portrayed in Fig. 1, where we plotted the mean and standard deviation of hourly energy consumption. This leads us to conclude that temporal features will play a pivotal role in the prediction of energy consumption.

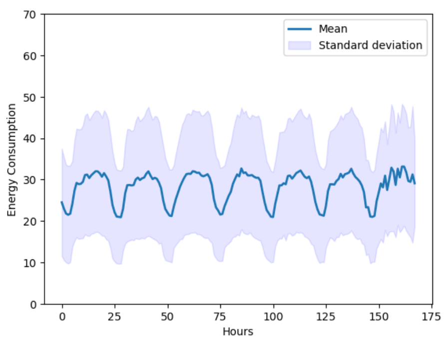

Figure 1 – Mean and standard deviation of hourly energy consumption

In our second observation, it becomes clear that energy consumption is significantly influenced by the configuration of the base station, including variables such as RU type, the number of antennas, and other relevant parameters. This relationship is visually depicted in Fig. 2 3. Differences among RU types impact energy usage due to varying coverage areas and power requirements. Additionally, we notice that base stations with four antennas tend to consume more energy on average than those with eight antennas. This is because 8-antenna configurations often utilize advanced signal processing techniques like beamforming, which improve energy efficiency by directing signals more precisely, whereas 4-antenna setups require higher power levels to maintain similar coverage and performance.

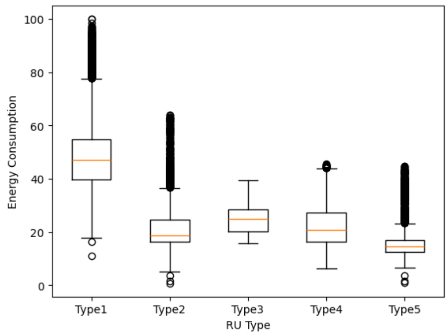

Figure 2 – The distribution of energy consumption values based on RU type

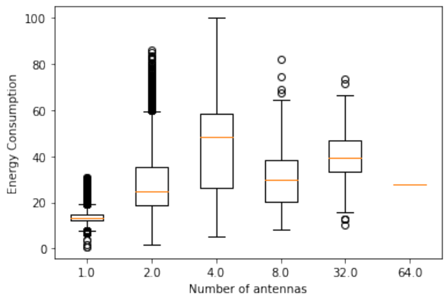

Figure 3 – The distribution of energy consumption values based on number of antennas

Our third observation is the clear connection between the amount of activity on the cellular network, known as "load," and the energy used by the base stations. Fig. 4 shows a noticeable pattern that suggests when the network is busier (higher load), the base stations tend to consume more energy. Moreover, we applied min-max scaling, which involves transforming data to a specific range (usually between 0 and 1), to normalize the energy consumption and cell load. The resulting time series for selected base stations are visualized in Fig. 5. This scaling process ensures that both variables are on a comparable scale, facilitating a more meaningful comparison of their trends over time.

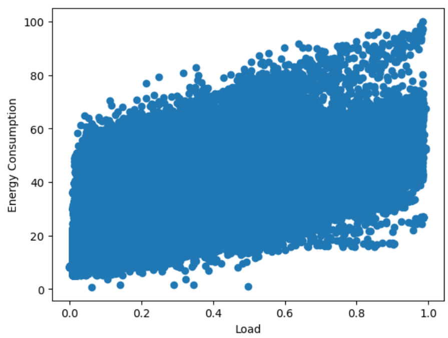

Figure 4 – Scatter plot of correlation between cell load and energy consumption

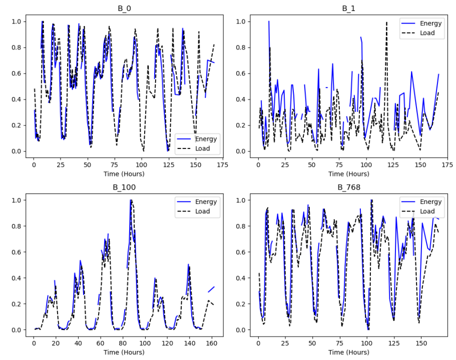

Figure 5 – Scaled energy consumption and cell load for some base stations against time

5. FEATURE ENGINEERING

In our feature engineering process, we strategically augment the cell-level data by employing a combination of historical features and statistical measures, enhancing the dataset's descriptive power. Specifically, historical features are incorporated through cumulative summation of energy-saving modes, denoting the evolving durations of each mode over time. Concurrently, statistical features such as minimum, maximum, mean, and standard deviation are computed, providing insights into the central tendencies and variability of relevant metrics. Additionally, we use the difference in load to capture the variation in load values within each base station and cell group. This amalgamation of historical and statistical features, along with the incorporation of load differences, aims to capture nuanced patterns and trends within the cellular network, fostering a more comprehensive understanding of its operational dynamics.

6. EVALUATION METRICS

6.1 Model performance metrics

We evaluated the models using several metrics to assess performance comprehensively:

- **Mean Absolute Percentage Error (MAPE)**: Measures the average magnitude of errors as a percentage of actual values, providing an intuitive understanding of prediction accuracy. It is defined as:

$$\mathrm{MAPE} = \frac{100\%}{n}\sum_{i=1}^{n}\left|\frac{y_i - \hat{y}_i}{y_i}\right|, \tag{1}$$

  where:

  - y_i is the actual value at observation i.
  - $\hat{y}_i$ is the predicted value at observation i.
  - n is the total number of observations.

  MAPE provides a percentage error, making it easy to interpret how far off predictions are on average in relation to the actual values. It is particularly useful for comparing the accuracy across different datasets or models. However, it can be sensitive to instances where y_i is close to zero, potentially inflating the error percentage.

- **Root Mean Squared Error (RMSE)**: Provides a quadratic scoring of errors, penalizing larger deviations more than smaller ones. It is sensitive to outliers and is useful when large errors are particularly undesirable. RMSE is defined as:

$$\mathrm{RMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}. \tag{2}$$

  RMSE gives a higher weight to large errors due to the squaring of the residuals. It is expressed in the same units as the target variable (in this case, energy consumption), facilitating direct interpretation of the error magnitude. RMSE is particularly appropriate when significant errors are especially problematic and need to be minimized.

- **Mean Absolute Error (MAE)**: Represents the average absolute differences between predictions and actual values, treating all errors equally regardless of their direction or magnitude. MAE is defined as:

$$\mathrm{MAE} = \frac{1}{n}\sum_{i=1}^{n}|y_i - \hat{y}_i|. \tag{3}$$

  MAE provides a straightforward interpretation of the average error magnitude, making it useful for understanding the typical size of the prediction errors. It is less sensitive to outliers compared to RMSE, as it does not square the residuals. MAE is expressed in the same units as the target variable, which aids in understanding the practical implications of the errors.

In these equations:
- y_i denotes the actual energy consumption at observation i. - $\hat{y}_i$ denotes the predicted energy consumption at observation i. - n is the total number of observations in the dataset.
By evaluating the models using these metrics, we gain a comprehensive understanding of their performance from different perspectives:
- MAPE allows us to understand the error relative to the size of the actual values, providing insight into the percentage accuracy of our predictions. - RMSE emphasizes larger errors, which is important in contexts where significant deviations are particularly detrimental. - MAE gives an average of all errors, offering a general sense of prediction accuracy without disproportionately weighting larger errors.

6.2 Feature importance

Feature importance is a technique used to quantify the contribution of each input feature to the predictions made by a model. Understanding feature importance helps in interpreting the model's decisions and identifying which variables have the most significant impact on the target variable, in this case, energy consumption.
In tree-based models like XGBoost and CatBoost, feature importance can be calculated based on how often and how effectively a feature is used to split the data across all trees in the ensemble. Common methods include:

- **Gain**: Measures the improvement in accuracy brought by a feature to the branches it is involved in. It quantifies the total reduction of the loss function attributed to a feature.
- **Frequency**: Counts the number of times a feature is used in all trees. A higher count indicates that the feature plays a more significant role in the model.

By analyzing the feature importances derived from our models, we observed that different algorithms prioritize different features. Specifically, the XGBoost models highlighted the importance of features such as cell load and hour, while CatBoost models emphasized categorical features like RU type and the number of antennas. This diversity in feature importance suggests that each model captures distinct aspects of the data due to differences in their learning mechanisms and handling of feature types.
According to ensemble learning theory, combining models that make uncorrelated errors can lead to a reduction in generalization error [19]. By aggregating the predictions of models that focus on different features, we can leverage their complementary strengths to mitigate individual biases and enhance predictive accuracy [20]. This approach aligns with established findings in ensemble learning literature, where integrating diverse models often results in superior performance compared to single models.

7. MODELING

In this study, we pose the problem as a regression task. The goal is to predict the continuous value of energy consumption based on the input features derived from the datasets. We employed three algorithms: XGBoost, CatBoost, and Artificial Neural Networks. The flowchart detailing our model is depicted in Fig. 6.

7.1 XGBoost

XGBoost is a robust and distributed ML system crafted to amplify the scalability of tree-boosting algorithms. Noteworthy features include optimized parallel tree construction and fault tolerance in distributed settings.

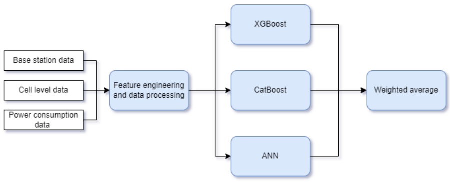

Figure 6 – Model flowchart

XGBoost excels in handling tens of millions of samples on a single node and seamlessly scales to manage billions of samples through distributed computing [21]. XGBoost has been successfully applied to a range of problems, demonstrating its versatility and effectiveness. [22] used XGBoost to detect suspicious network traffic events, achieving high classification accuracy. [23] applied XGBoost to personal credit evaluation, outperforming other models in feature selection and classification performance. [24] utilized XGBoost for bearing fault diagnosis, demonstrating its superior accuracy and speed compared to other tree algorithms. [25] employed XGBoost as a predictive tool in a CNC turning process, achieving excellent accuracy in predicting response values. In [26], XGBoost is used to construct electricity prediction models for different user categories and then obtain the overall electricity consumption in a region.

In this paper, we used XGBoost to predict energy consumption in a 5G network. We used a 5-fold cross-validation approach with a squared error loss. The results obtained are summarized in Table 1.

Table 1 – Training and validation set metrics for XGBoost models

Metric	MAPE (%)	RMSE	MAE
Training Set	2.8317	1.0932	0.7340
Validation Set	4.0088	1.7157	1.0980

Furthermore, in our feature analysis, the examination of feature importances revealed the most influential features in our models, including cell load and hour features. Fig. 7 depicts the features by importance in XGBoost models.

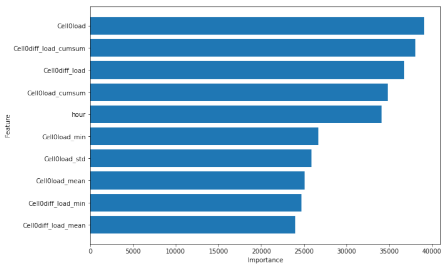

Figure 7 – Top 10 features by importance in XGBoost models

7.2 CatBoost

CatBoost is a gradient-boosting algorithm that has been shown to outperform other implementations in terms of quality and speed, particularly in handling categorical features [27]. This is due to its innovative algorithmic techniques, including ordered boosting and a method for processing categorical features, which address prediction shifts and target leakage issues [28]. CatBoost has been successfully used in machine learning studies involving big data, particularly in tasks with categorical, heterogeneous data [29]. In [30], the authors used CatBoost to predict housing prices. In [31], CatBoost was utilized to detect fraud in financial transactions.

In this paper, we trained the CatBoost model using a 5-fold cross-validation scheme and mean squared error (MSE) loss. The results of our model's accuracy are shown in Table 2.

Table 2 – Training and validation set metrics for CatBoost models

Metric	MAPE (%)	RMSE	MAE
Training Set	3.8392	1.4935	1.0128
Validation Set	4.2880	1.7557	1.1528

Furthermore, we examined the feature importances within our models. The analysis uncovered that the CatBoost models placed greater emphasis on categorical features, particularly RU type and the number of antennas. The outcomes are presented in Fig. 8. The difference in feature importance between XGBoost and CatBoost models may be due to the way each algorithm handles categorical features and constructs decision trees, impacting the interpretation of variable significance and overall model performance metrics (MAPE, RMSE, and MAE).

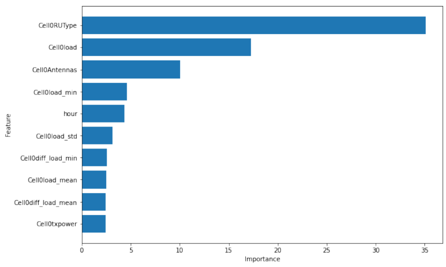

Figure 8 – Top 10 features by importance in CatBoost models

7.3 Artificial Neural Network

Artificial Neural Networks (ANNs), inspired by the human brain, are powerful tools for nonlinear mapping and learning from empirical data [32][33]. They consist of interconnected simulated neurons, which process information in parallel and communicate through activation signals [34]. These networks are highly task-efficient and learn from experience, following specific learning laws and receiving feedback from the environment [35].

Input data is fed into the network through the input layer, and information is progressively transformed as it passes through hidden layers, with each layer contributing to the extraction of features and patterns. The output layer produces the final result or prediction. ANNs are particularly adept at learning complex relationships and patterns in data through a training process that involves adjusting the weights of connections between neurons. This adaptability makes ANNs versatile for various tasks, including classification, regression, and pattern recognition, contributing to their widespread application across diverse domains of machine learning and artificial intelligence. These applications include automatic speech processing [36], nonlinear process forecasting [37], predictive modeling of complex data [38], and image processing and visualization [39].

The ANN employed in this study was specifically configured with five layers and nonlinear activation functions. This architecture was tailored to address the unique characteristics of our dataset and optimize performance for the task at hand. Fig. 9 shows our model configuration. This model was trained using 5-fold cross-validation and mean absolute error loss. The results obtained are depicted in Table 3.

Table 3 – Training and validation set metrics for ANN models

Metric	MAPE (%)	RMSE	MAE
Training Set	3.0849	1.9875	1.0098
Validation Set	3.8363	1.2048	1.0673

7.4 Ensemble approach

Ensemble learning is a powerful machine learning paradigm where multiple models are combined to produce a single predictive model with improved performance. The fundamental idea is that by aggregating the predictions of diverse models, we can reduce the variance and bias inherent in individual models, leading to better generalization on unseen data [20]. In this study, we employ an ensemble approach by combining the predictions of XGBoost, CatBoost, and ANN models. Each of these models contributes unique strengths to the prediction task: XGBoost excels at capturing complex nonlinear relationships through gradient boosting of decision trees [21]; CatBoost effectively handles categorical features and mitigates overfitting using ordered boosting and permutation techniques [28]; and ANN is capable of modeling intricate patterns by learning hierarchical feature representations [40].

To combine these models, we adopted a weighted averaging method. Let $\hat{y}_{\text{XGB}}$, $\hat{y}_{\text{CAT}}$, and $\hat{y}_{\text{ANN}}$ denote the predictions from the XGBoost, CatBoost, and ANN models, respectively. The final ensemble prediction $\hat{y}_{\text{ensemble}}$ is computed as:

$$\hat{y}_{\text{ensemble}} = w_{\text{XGB}} \cdot \hat{y}_{\text{XGB}} + w_{\text{CAT}} \cdot \hat{y}_{\text{CAT}} + w_{\text{ANN}} \cdot \hat{y}_{\text{ANN}},$$

where $w_{\text{XGB}} = 0.3$, $w_{\text{CAT}} = 0.3$, and $w_{\text{ANN}} = 0.4$, satisfying $w_{\text{XGB}} + w_{\text{CAT}} + w_{\text{ANN}} = 1$. These weights were determined through experimentation and cross-validation on the validation set. The ANN received a slightly higher weight due to its superior performance in capturing nonlinear relationships, as indicated by lower error metrics during validation.

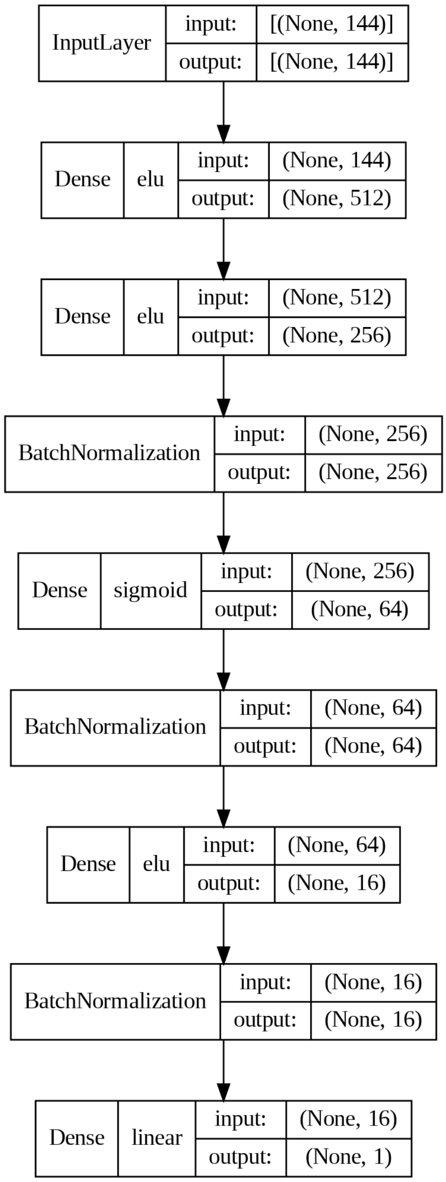

Figure 9 – ANN model architecture

The rationale behind choosing these specific models and weights includes leveraging the diversity of model architectures to capture a broader spectrum of patterns within the data. By integrating these models, we aim to mitigate individual model biases and reduce the overall generalization error [19].

In implementing the ensemble, each base model was trained independently using the same training data and cross-validation splits to ensure consistency. We evaluated each model's performance on the validation set using the metrics defined in Section 6. Based on these evaluations, we experimented with different weight combinations and selected the ones that minimized the overall validation error of the ensemble.

Ensemble methods have been successfully applied in energy consumption forecasting. For instance, Zhang et al. [41] utilized an ensemble of neural networks for short-term load forecasting, demonstrating improved accuracy over single models. Similarly, Li et al. [42] combined multiple machine learning models to predict building energy consumption, achieving better performance through ensemble learning. Our approach aligns with these studies, affirming the effectiveness of ensemble methods in enhancing predictive performance in energy-related applications.

8. RESULTS

8.1 Performance comparison

Table 4 presents the performance metrics of the individual models and the ensemble on the validation set.

Table 4 – Performance metrics of individual models and the ensemble

Model	MAPE (%)	RMSE	MAE
XGBoost	4.0088	1.7157	1.0980
CatBoost	4.2880	1.7557	1.1528
ANN	3.8363	1.2048	1.0673
Ensemble	**3.5620**	**1.1524**	**1.0245**

The ensemble model outperformed the individual models, achieving the lowest MAPE, RMSE, and MAE. This demonstrates the effectiveness of the ensemble approach in enhancing prediction accuracy.

8.2 Computational efficiency

The model training was performed on an NVIDIA P100 GPU. The training time for the ensemble was approximately 20 minutes, indicating computational efficiency suitable for practical applications. The scalability of the models allows for deployment in real-time network management systems.

8.3 Implications and applications

The improved accuracy of the ensemble model provides valuable insights for network operators:

- **Energy management**: Accurate predictions enable proactive adjustments to network configurations, optimizing energy consumption without compromising service quality.
- **Policy planning**: Understanding energy usage patterns assists in developing effective energy-saving policies and deploying energy-saving methods strategically.
- **Scalability**: The model's ability to generalize across different products and configurations facilitates its application in diverse network environments.

9. CONCLUSIONS

In conclusion, our study demonstrates that machine learning models can effectively predict energy consumption in 5G networks. The use of XGBoost, CatBoost, and Artificial Neural Network (ANN), combined through an ensemble approach, significantly improves prediction accuracy, addressing the gap in existing literature and offering valuable insights for energy-efficient management of 5G networks. Our findings highlight the importance of considering a wide range of parameters, including engineering configurations, traffic conditions, and energy-saving methods, to enhance the generalizability of the models across various scenarios. Future work may explore the integration of additional data sources and advanced machine learning techniques to further refine the predictions and support the ongoing efforts to optimize energy consumption in 5G infrastructure.

ACKNOWLEDGEMENT

We express our gratitude to the International Telecommunication Union (ITU) and Zindi for graciously providing the dataset and offering us the opportunity to participate in this challenging competition. We also extend our heartfelt thanks to the reviewers for their constructive comments, insightful feedback, and valuable suggestions, which have significantly contributed to the improvement and refinement of our work.

REFERENCES

[1] Marat Nuriev, Anna Kalyashina, Yuri Smirnov, Guzel Gumerova, and Gulaybat Gadzhieva. "The 5G revolution transforming connectivity and powering innovations". In: *E3S Web of Conferences*. Vol. 515. EDP Sciences. 2024, p. 04008.

[2] Ante Mihaljević, Ana Kešelj, and Adriana Lipovac. "Impact of 5G Network Performance on Augmented Reality Application QoE". In: *2021 International Conference on Software, Telecommunications and Computer Networks (SoftCOM)*. 2021, pp. 1–4. DOI: 10.23919/SoftCOM52868.2021.9559067.

[3] Georgios Gkagkas, Dimitrios J Vergados, Angelos Michalas, and Michael Dossis. "The Advantage of the 5G Network for Enhancing the Internet of Things and the Evolution of the 6G Network". In: *Sensors* 24.8 (2024), p. 2455.

[4] Md Maruf Ahamed and Saleh Faruque. "5G network coverage planning and analysis of the de-

ployment challenges". In: *Sensors* 21.19 (2021), p. 6608.

[5] Ashikur Rahman, Salsabil Arabi, and Raqeebir Rab. "Feasibility and challenges of 5g network deployment in least developed countries (ldc)". In: *Wireless Sensor Network* 13.1 (2021), pp. 1–16.

[6] Adil Israr, Qiang Yang, and Ali Israr. "Power consumption analysis of access network in 5G mobile communication infrastructures—An analytical quantification model". In: *Pervasive and Mobile Computing* 80 (2022), p. 101544.

[7] Poonam Tiwari, Vishant Gahlaut, Meenu Kaushik, Preeti Rani, Anshuman Shastri, and Bhupender Singh. "Advancing 5G connectivity: a comprehensive review of MIMO antennas for 5G applications". In: *International Journal of Antennas and Propagation* 2023.1 (2023), p. 5906721.

[8] Adil Israr, Qiang Yang, Wei Li, and Albert Y Zomaya. "Renewable energy powered sustainable 5G network infrastructure: Opportunities, challenges and perspectives". In: *Journal of Network and Computer Applications* 175 (2021), p. 102910.

[9] Ltd. Huawei Technologies Co. *Green 5G: Building a Sustainable World," Tech. Rep.* Aug. 2020. URL: `https://www.huawei.com/en/%20public-policy/green-5g-building-a-sustainable-world`.

[10] Josip Lorincz, Amar Kukuruzović, and Zoran Blažević. "A Comprehensive Overview of Network Slicing for Improving the Energy Efficiency of Fifth-Generation Networks". In: *Sensors* 24.10 (2024), p. 3242.

[11] Xiong Yu, Geng Li, and Wenbing Lu. "Power consumption based on 5G communication". In: *2021 IEEE 5th Information Technology, Networking, Electronic and Automation Control Conference (ITNEC)*. Vol. 5. IEEE. 2021, pp. 910–914.

[12] Dongzhu Xu, Anfu Zhou, Xinyu Zhang, Guixian Wang, Xi Liu, Congkai An, Yiming Shi, Liang Liu, and Huadong Ma. "Understanding operational 5G: A first measurement study on its coverage, performance and energy consumption". In: *Proceedings of the Annual conference of the ACM Special Interest Group on Data Communication on the applications, technologies, architectures, and protocols for computer communication.* 2020, pp. 479–494.

[13] Arvind Narayanan, Xumiao Zhang, Ruiyang Zhu, Ahmad Hassan, Shuowei Jin, Xiao Zhu, Xiaoxuan Zhang, Denis Rybkin, Zhengxuan Yang, Zhuoqing Morley Mao, et al. "A variegated look at 5G in the wild: performance, power, and QoE implications". In: *Proceedings of the 2021 ACM SIGCOMM 2021 Conference.* 2021, pp. 610–625.

[14] Zsolt Temesvári. "Review of Frequency Spectrum and Energy Consumption in 5G". In: *2019 International IEEE Conference and Workshop in Óbuda on Electrical and Power Engineering (CANDO-EPE)*. IEEE. 2019, pp. 21–26.

[15] Yong-Hoon Choi, Jungerl Lee, Juhoon Back, Suwon Park, Young-uk Chung, and Hyukjoon Lee. "Energy efficient operation of cellular network using on/off base stations". In: *International Journal of Distributed Sensor Networks* 11.8 (2015), p. 108210.

[16] Bjorn Debaillie, Claude Desset, and Filip Louagie. "A flexible and future-proof power model for cellular base stations". In: *2015 IEEE 81st Vehicular Technology Conference (VTC Spring)*. IEEE. 2015, pp. 1–7.

[17] Mohammed H Alsharif, Rosdiadee Nordin, and Mahamod Ismail. "A review on intelligent base stations cooperation management techniques for greener lte cellular networks". In: *Journal of Communications* 9.12 (2014), pp. 937–945.

[18] Jad Nasreddine, Khalil Fakih, Bassem Haidar, Dina Serhal, and Siba Haidar. "Estimating Base Station Power Consumption Using Regression". In: *2019 3rd International Conference on Bioengineering for Smart Technologies (BioSMART)*. IEEE. 2019, pp. 1–4.

[19] Thomas G Dietterich. "Ensemble methods in machine learning". In: *International workshop on multiple classifier systems*. Springer. 2000, pp. 1–15.

[20] Zhi-Hua Zhou. *Ensemble methods: foundations and algorithms*. CRC press, 2012.

[21] Tianqi Chen and Carlos Guestrin. "Xgboost: A scalable tree boosting system". In: *Proceedings of the 22nd acm sigkdd international conference on knowledge discovery and data mining*. 2016, pp. 785–794.

[22] Łukasz Podlodowski and Marek Kozłowski. "Application of XGBoost to the cyber-security problem of detecting suspicious network traffic events". In: *2019 IEEE International Conference on Big Data (Big Data)*. IEEE. 2019, pp. 5902–5907.

[23] Hua Li, Yumeng Cao, Siwen Li, Jianbin Zhao, and Yutong Sun. "XGBoost model and its application to personal credit evaluation". In: *IEEE Intelligent Systems* 35.3 (2020), pp. 52–61.

[24] Rongtao Zhang, Binbin Li, and Bin Jiao. "Application of XGboost algorithm in bearing fault diagnosis". In: *IOP Conference Series: Materials Science and Engineering*. Vol. 490. IOP Publishing. 2019, p. 072062.

[25] Shankar Chakraborty and Shibaprasad Bhattacharya. “Application of XGBoost algorithm as a predictive tool in a CNC turning process”. In: *Reports in Mechanical Engineering* 2.1 (2021), pp. 190–201.

[26] Danhuang Dong, Fan Wen, Yihong Zhang, and Wenxin Qiu. “Application of XGboost in Electricity Consumption Prediction”. In: *2023 IEEE 3rd International Conference on Electronic Technology, Communication and Information (ICETCI)*. IEEE. 2023, pp. 1260–1264.

[27] Anna Veronika Dorogush, Vasily Ershov, and Andrey Gulin. “CatBoost: gradient boosting with categorical features support”. In: *arXiv preprint arXiv:1810.11363* (2018).

[28] Liudmila Prokhorenkova, Gleb Gusev, Aleksandr Vorobev, Anna Veronika Dorogush, and Andrey Gulin. “CatBoost: unbiased boosting with categorical features”. In: *Advances in neural information processing systems* 31 (2018).

[29] John T Hancock and Taghi M Khoshgoftaar. “CatBoost for big data: an interdisciplinary review”. In: *Journal of big data* 7.1 (2020), pp. 1–45.

[30] Tie Wang, Yutian Wang, and Mingzhe Liu. “A price prediction method based on CatBoost”. In: *2021 International Conference on Culture-oriented Science & Technology (ICCST)*. IEEE. 2021, pp. 403–408.

[31] Yeming Chen and Xinyuan Han. “CatBoost for fraud detection in financial transactions”. In: *2021 IEEE International Conference on Consumer Electronics and Computer Engineering (ICCECE)*. IEEE. 2021, pp. 176–179.

[32] Jitendra R Raol and Sunilkumar S Mankame. “Artificial neural networks”. In: *Resonance* 1.2 (1996), pp. 47–54.

[33] S Lek and YS Park. “Artificial neural networks”. In: *Encyclopedia of Ecology, Five-Volume Set*. Elsevier Inc., 2008, pp. 237–245.

[34] Rudolf Kruse, Christian Borgelt, Christian Braune, Sanaz Mostaghim, Matthias Steinbrecher, Rudolf Kruse, Christian Borgelt, Christian Braune, Sanaz Mostaghim, and Matthias Steinbrecher. “Introduction to Neural Networks”. In: *Computational Intelligence: A Methodological Introduction* (2016), pp. 9–13.

[35] Mohammad Ehteshamullah. “Artificial Neural Network”. In: *Journal of Mathematical Sciences & Computational Mathematics* 1.1 (2020), pp. 50–54.

[36] Sabato Marco Siniscalchi, Torbjørn Svendsen, and Chin-Hui Lee. “An artificial neural network approach to automatic speech processing”. In: *Neurocomputing* 140 (2014), pp. 326–338.

[37] Manoochehr Ghiassi and Hassine Saidane. “A dynamic architecture for artificial neural networks”. In: *Neurocomputing* 63 (2005), pp. 397–413.

[38] Jonas S Almeida. “Predictive non-linear modeling of complex data by artificial neural networks”. In: *Current opinion in biotechnology* 13.1 (2002), pp. 72–76.

[39] Paul Dan Cristea. “Application of neural networks in image processing and visualization”. In: *GeoSpatial Visual Analytics: Geographical Information Processing and Visual Analytics for Environmental Security*. Springer. 2009, pp. 59–71.

[40] Ian Goodfellow. *Deep learning*. 2016.

[41] Xiaobo Zhang, Jianzhou Wang, and Kequan Zhang. “Short-term electric load forecasting based on singular spectrum analysis and support vector machine optimized by Cuckoo search algorithm”. In: *Electric Power Systems Research* 146 (2017), pp. 270–285.

[42] Chengdong Li, Zixiang Ding, Dongbin Zhao, Jianqiang Yi, and Guiqing Zhang. “Building energy consumption prediction: An extreme deep learning approach”. In: *Energies* 10.10 (2017), p. 1525.

AUTHORS

Ilyasse Makroz received a master’s degree in air traffic management from the International Academy Mohamed VI of Civil Aviation, Morocco. He obtained a Ph.D. in applied mathematics and statistics at University Hassan First, Settat, Morocco. His current research interests include diffusion processes and machine learning.

Ahmed Nafidi received his Ph.D. in mathematics from Granada University, Spain, in 1997. Now, he is a professor in the Computer Science and Mathematics Department of the National School of Applied Sciences at Hassan First University of Settat. A member of the Laboratory of Analysis and Modeling Systems for Decision Support (LAMSAD), Berrechid, Morocco. He is also a member of Grupo de Investigación (FQM-147) Análisis Estadístico de Datos Multivariantes y Procesos Estocásticos, Department of Statistics and Operations Research at University of Granada, Spain. His research interests include stochastic diffusion processes and their applications, and machine learning.

Ramón Gutiérrez-Sánchez joined the School of Science at Granada University in October 2000. He received his Ph.D. in statistics from the University of Granada, Spain where he is currently a full professor in statistics. He served as chair and co-chair for many international and domestic academic conferences. His current research interests include diffusion processes, data mining, and computer applications.

Salaheddine Aourik received a master's degree in electronic aviation safety from the International Academy Mohamed VI of Civil Aviation, Morocco. He obtained a Ph.D. in physics and engineering sciences at FST University Hassan First, Settat, Morocco. His research interests include the design and analysis of millimeter antenna arrays for 5G applications.