Exploring Machine Learning Applications in 5G Network Optimization

Code Folder Link-

Devesh Sharma 202111024
Didhit Patel 202111025
Prachi Gurav 202111064

Abstract *– The advancement of 5G networks has brought significant innovations in terms of services and technologies but this has also imposed challenges related to energy consumption. Over 70% of energy consumption was projected to be attributed to Radio Access Networks (RANs), specifically Base Stations (BSs), with data centers and fiber transport contributing to a lesser extent. The objective of this study was to optimize the parameters of BSs and energy-saving methods, providing a deep understanding of how these elements influence energy consumption. This study introduces a hybrid-boosted ensemble model tailored for predicting energy utilization in 5G base stations. The methodology merges ridge regression for linear trend analysis, XGBoost to tackle non-linear intricacies, and a final refinement strategy built upon a mixture of experts. Preliminary results demonstrate the model's capability to adjust to new and diverse data scenarios, achieving up to a 31.94% improvement in 5G energy forecasting compared to traditional models. This research constitutes a pivotal initiative in confronting adaptability hurdles in 5G energy predictions, setting a foundation for subsequent inquiries in this essential domain.*

Keywords – 5G, artificial intelligence, energy consumption, hybrid-boosted model, mixture of experts.

1. INTRODUCTION

Discussions on sustainable development, initiated in the United Nations agenda in 1972, culminated in the 2015 launch of the 2030 Agenda, which establishes 17 Sustainable Development Goals (SDGs) with 169 targets, integrating social, environmental, and economic dimensions to address critical global challenges [1,2]. However, progress has been insufficient. According to the SDG Progress Report, which assessed approximately 140 targets for which trend data is available, about half of the targets are moderately or severely off track, and over 30% have either seen no movement or regressed below the 2015 baseline, particularly affecting the most vulnerable populations [3,4].

The imperative for a fundamental shift is pressing, requiring a recalibration in commitment, solidarity, financing, and action to guide the world towards a more positive trajectory. Within this context, the crucial role played by the mobile industry in facilitating this essential transition is underscored. Certainly, mobile technology functions as a crucial gateway for billions, offering access to the Internet and essential services in education, healthcare, and finance. Moreover, it serves as a cornerstone in the digital economy, spearheading transformative changes across diverse sectors [5]. However, according to the GSMA's Mobile Impact Report, the industry is projected to achieve only 76% of its full potential impact on the 17 SDGs by 2030, based on its current trajectory. To enhance this impact, the report suggests expediting mobile connectivity for unconnected populations and addressing challenges, specifically focusing on incentives for infrastructure investment, affordability, the gender gap, digital skills, and the availability of local content and services. Additionally, the report emphasizes the importance of operators, governments, international organizations, and other industries collaborating to support the scaling of new and existing mobile solutions [3].

The introduction of the fifth generation of mobile networks (5G) will disrupt industrial organization and production models due to its technical features, including higher transmission speeds, ultra-reliable low latency, enhanced network security, massive machine-type communications, and improved device energy efficiency. Consequently, the deployment of these networks will enable the expansion of wireless broadband services beyond mobile Internet, reaching complex Internet of Things systems [6]. In addition, the development of advanced technologies like 5G, the Internet of Things, and artificial intelligence is expected to contribute to a potential 15% reduction in global carbon emissions. This reduction represents nearly one-third of the proposed 50% decrease by 2030. The achievement of this goal is envisioned through the development of innovative solutions spanning various sectors, including energy, manufacturing, agriculture, natural resource extraction, construction, services, transport, and traffic management [6, 7]. This can mitigate certain adverse effects stemming from the production and utilization of these technologies, characterized by significant energy consumption (1.4% of the global total), extensive generation of electronic waste, and the extraction of natural resources like copper and lithium [6, 8].

The advancement of 5G networks has brought significant innovations in terms of services and technologies but has also imposed challenges related to energy consumption. According to estimates, 5G networks are about 4 times more energy efficient than their 4G predecessors, yet their energy consumption is approximately three times higher due to the need for a greater number of cells and the additional processing required for broader bandwidths [9]. Notably, network Operational Expenditure (OPEX) constitutes around 25% of the total cost for network operators, with 90% of this allocation directed toward elevated energy bills. Over 70% of the energy consumption is projected to be attributed to Radio Access Networks (RANs), specifically Base Stations (BSs), with data centers and fiber transport contributing to a lesser extent. The energy consumption of base stations is influenced by various factors, including architecture, configuration parameters, traffic conditions, and the implementation of energy-saving methods. In order to minimize network energy usage, it is vital to fine-tune both base station parameters and energy-saving techniques. This requires a deep understanding of how these parameters and methods impact the energy consumption of different base stations.

While previous studies have explored various approaches to energy modeling, there is a significant gap regarding the adaptation of these models to new datasets, especially in dynamic environments such as those found in 5G networks. The need for models that can quickly adjust to new data and traffic conditions is crucial for effective energy management. In this context, the focus of this paper is to develop a precise and adaptable model for predicting the energy consumption of Base Stations (BSs) in 5G networks.

To achieve this, a hybrid-boosted ensemble model [10] is introduced that combines the feature transformation power of ridge regression with the target transformation flexibility of XGBoost. Inspired by the mixture of experts approach, a final refinement model is applied to enhance prediction accuracy [11]. The objective of this study was to optimize the parameters of BSs and energy-saving methods, providing a deep understanding of how these elements influence energy consumption. The model was tested in scenarios with significant variations in data characteristics, demonstrating great adaptability and accuracy in estimates.

2. RELATED WORK

A study conducted by Ardabili et al. [12] examined the evolution and trajectory of hybrid machine learning models, asserting their ongoing growth and development across various scientific domains. As computational power continues to advance through the emergence of new technologies enabling faster and more intricate computations, the significance of hybrid models is on the rise. These models not only facilitate the integration of advanced techniques spanning artificial intelligence and machine learning fields but also enhance accuracy and predictive capabilities. It is highly probable that hybrid models will become the new standard for generating precise and reliable models within the realm of data science in the future [12].

The research methodology will draw upon background research conducted in the field of hybrid machine learning. Specifically, it will be informed by studies conducted by Khan et al. [1315], which implement various hybrid machine learning techniques to predict energy demand using features extracted from time series data.

In addition to studies on hybrid machine learning models, a notable approach that exemplifies the evolution and applicability of these models is the "Mixture of Experts" (MoE) technique. This methodology, introduced by Jacobs et al. [16], is based on the idea of dividing a complex problem into simpler subproblems, allowing individual "experts" to focus on specific parts of the problem. The integration of these experts' results is done by a "gate" that learns to weigh each expert's contribution to the final solution. This approach allows for the construction of models that are not only highly adaptable to different types of data but also capable of capturing complex nuances that single models might not achieve.

Expanding upon the flexibility and adaptability discussed in the mixture of experts methodology, the literature review reveals that, although various energy prediction approaches have been explored, few studies explicitly address the adaptability of models to new data conditions, a critical requirement for rapidly evolving technologies like 5G. Previous studies tend to focus on static models that, although efficient under tested conditions, fail to adapt when exposed to significant variations in traffic and technology. This study advances in this aspect by integrating the flexibility of the 'Mixture of Experts' approach, a methodology that allows the model to dynamically adjust, improving precision and effectiveness under different operational scenarios.

3. DATASET

The data used in this study come from a machine learning competition, and the training dataset includes cell-level traffic statistics of 4G/5G sites collected on different days [17]. The dataset was categorized into three main groups, as summarized below [17]:

- **Base station basic information:** Include configuration parameters and hardware attributes, which consist of predominantly categorical data such as "RUType" and "Mode", along with continuous numerical data like "Frequency" and "TXpower" (Table 1).
- **Cell-level data:** Include hour-level counters, including service compliance counters (e.g., load) and energy-saving methods' counters (e.g., duration of energy saving mode activation). The continuous numerical data was normalized for "Load" and energysaving modes ("ESMode1" to "ESMode6"; Table 2).
- **Energy consumption data:** Include hourlevel energy consumption specifications (e.g., total energy consumption of the base stations). The continuous numerical data was represented in relative units of energy (Table 3).

Table 1 – Data of base station basic information.

Label	Description
BS	Name of the base station
Cell Name	Name of the cell
RUType	Name of the radio unit type
Mode	Transmission mode
Frequency	Frequency of the cell
Bandwidth	Bandwidth of the cell
Antennas	Number of antennas of the base station
TXpower	Maximum transmit power of the cell

Table 2 – Cell-level data.

Label	Description
Time	Data and time in which the measurement was collected
BS	Name of the base station
Cell Name	Name of the cell
Load	Load of the cell (values in 0-1)
ESMode (1-6)	Intensity of the activation of different energy-saving modes (values in 0-1)

Table 3 – Energy consumption data.

Label	Description
Time	Data and time in which the measurement was collected

BS	Name of the base station
Energy	Energy consumption measurement

4. PROPOSED MODEL

4.1 Introduction

In this study, three central objectives were addressed, which have significantly influenced both the conceptualization and the adopted methodology. A hybrid-boosted model was employed, combining residual fitting techniques similar to gradient boosting methods with an ensemble approach influenced by a mixture of experts. This algorithm was specifically designed to meet the three main objectives of this study, identified as Objectives A, B, and C.

4.1.1 Objective A: Estimating energy consumption in specific base station products

To meet this objective, a data sample comprising 816 BSs was meticulously analyzed. Each BS in the sample provided a complete seven-day history of energy consumption, along with statistics related to traffic and the activation of energy-saving systems. Fig. 1 illustrates a potential seasonality in both "Energy" and "Load" variables, while also implying a possible correlation between them.

4.1.2 Objective B: Generalization across different base station products

In this objective, the issue arose of the absence of BS information in the training dataset, which had the potencial to limit the consideration of seasonality and historical traffic of that particular 5G network region in the predictions. However, the specific configurations of these stations were available in the dataset. This provides the opportunity to prioritize these configurations in an attempt to generalize energy consumption, without the need to anchor it to the historical traffic of the 5G bases. This approach was explored and elucidated in subsequent sections.

4.1.3 Objective C: Generalization across different base station configurations

In Objective C, an exacerbated challenge is encountered due to the absence of both BSs and their associated configurations in the training dataset. This scenario implied a high degree of generalization required by the algorithm, given the need to operate under conditions considerably different from those observed during training. It is also noteworthy that this dataset had a specific limitation: the available data is restricted to day 02.

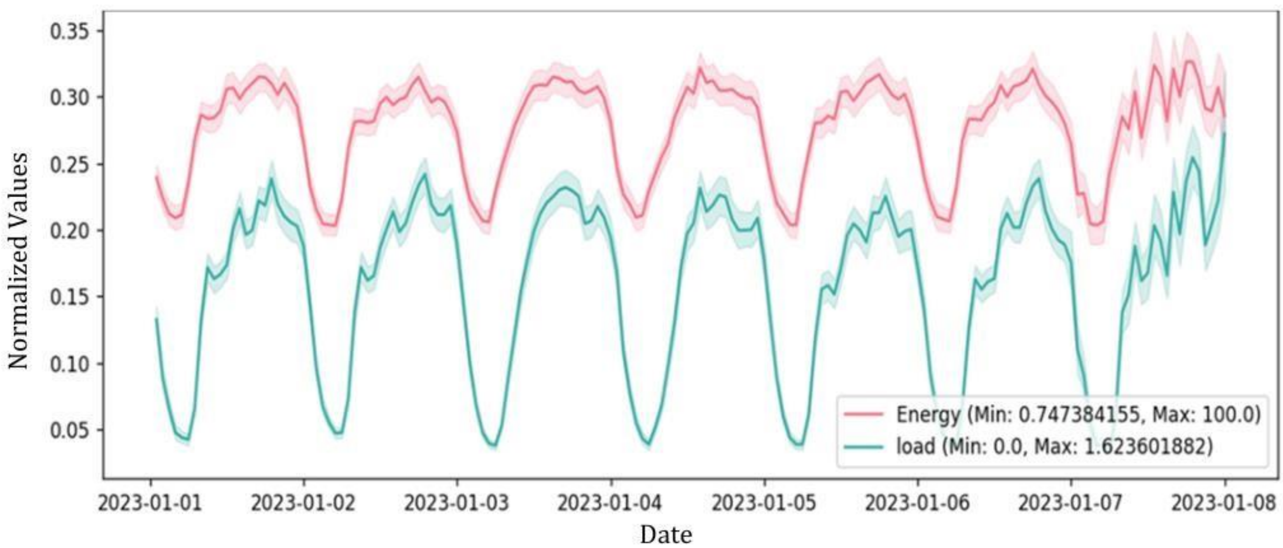

Figure 1 – This figure displays a normalized hourly graph of "Energy" and "Load" over a 7-day period, where the values have been scaled to fit within the same plot for easy comparison. The red line represents "Energy" and the green line represents "Load", with their value ranges detailed in the legend.

The peaks and valleys in the 5G base station's load graph directly reflect high and low traffic times, indicating increased energy consumption during periods of intense data activity and reduction during times of lower demand.

.

4.2 Inputs of the model

In the implementation of this approach, a structure based on a "Mixture of Experts" was employed. This strategic decision allowed for the creation of three specialized experts: Expert A, Expert B, and Expert C, each designated to specifically address one of the proposed objectives. In this section, specific inputs for each of these experts were outlined, emphasizing the importance of feature engineering and selection to optimize the performance of the algorithms and achieve the desired outcomes.

4.2.1 Expert A

The focus of "Expert A" was placed on temporal feature engineering. Attributes such as lags, differences, and purpose-specific attributes like "Load", "ESMode", and "Time" were meticulously developed. A notable addition was the target encoding on the hour variable, a critical feature for capturing the historical hourly seasonality of the BS. It was imperative to highlight that these temporal features were customized for each BS, ensuring accuracy while preventing data leakage.

4.2.2 Expert B

The focus of "Expert B" was on the implementation of advanced feature engineering and meticulous selection. Robust polynomial features were developed on attributes such as "Load", "ESMode6", "Antennas", "TXpower", and "Frequency".
Additionally, k-means clustering was applied to the BS configurations. The primary intent was to enhance the algorithm's ability to generalize over new stations, minimizing reliance on historical BS data and accentuating the significance of station settings and load. This method provided a hybrid approach, merging time-series features with robust BS configurations.

4.2.3 Expert C

The approach of "Expert C" was notably streamlined. Despite initial efforts to implement specialized features, such attempts proved ineffective. Optimized analyses consistently indicated that simpler approaches outperformed complex ones. Therefore, the judicious selection of features emerged as the crucial component in this context, eliminating attributes that led to overfitting or that could introduce contamination from the history of the BS or its robust settings.

In the engineering process, all features were formulated in a singular stage. However, for each expert, specific pre-processing was carried out, along with a judicious selection of attributes based on theoretical expertise and algorithmic optimization. This level of detail became essential in defining the input attributes for each specialized model.

4.3 Outputs of the model

The primary output of our hybrid-boosted ensemble model is a quantitative projection of energy consumption. This projection stems from the synthesis of insights derived from ridge regression and XGBoost, complemented by the appropriate model selection via an approach inspired by the mixture of experts. Through this composite methodology, the aim is to provide a rigorous and robust estimate of energy consumption, considering the intrinsic particularities of the provided data.

5. MODEL ARCHITECTURE

In the field of energy-consumption forecasting, unique challenges were faced concerning generalization with respect to the key objectives of energy prediction. To address these challenges, a boosted-hybrid ensemble model was developed, combining the feature transformation capabilities of ridge regression with the target transformation flexibility of XGBoost. This culminated in an application inspired by the mixture of experts for final refinement (Fig. 2).

5.1 Boosted ensemble hybrid model

In the boosted ensemble hybrid model, ridge regression was chosen for its ability to handle multicollinearity and extrapolate trends based on the characteristics of the input data. Indeed, this base model of the ensemble, particularly effective in identifying and mapping linear trends, belongs to the set of algorithms for feature transformation, which also includes linear regression and neural networks. This type of algorithm is notable for its ability to perform extrapolative inferences beyond the scope of the initial training set.

After obtaining predictions from ridge regression, predictions were processed by calculating residuals, and XGBoost was trained on these residuals. This allows XGBoost to understand and capture nonlinear nuances and trends that the ridge regression may have missed. In fact, XGBoost is a target transformation-focused algorithm that employs feature clustering to optimize outputs based on the mean of these clusters. Its predictive capability is confined to the range of the training set, lacking the ability for extrapolation. Algorithms such as decision trees and k-Nearest Neighbors (k-NN) algorithms also fall into this category.

Summarily, estimates from ridge regression were utilized and subsequently adjusted by XGBoost predictions on the residuals, forming the final estimate.

5.2 Mixture of experts

The mixture of experts is a refined approach that segments the dataset and trains specialists, or individual models, for specific portions of the data. The gating mechanism, in this case, functions as a conditional model based on the "hot" or "cold" features of the data relative to the training set, determining which specialist is most appropriate to consult, considering the specificities of the input data to be forecasted.

Inspired by the mixture of experts method, the examined test data was segmented based on specific features, determining whether they are "hot" (similar to the training data) or "cold" (different from the training data). The adaptive ability of this strategy excels in accurately differentiating features in the data, allowing for the selection of the most suitable expert for each context. As new patterns emerge, the technique has the potential to recognize and proactively adjust its forecasting approaches.

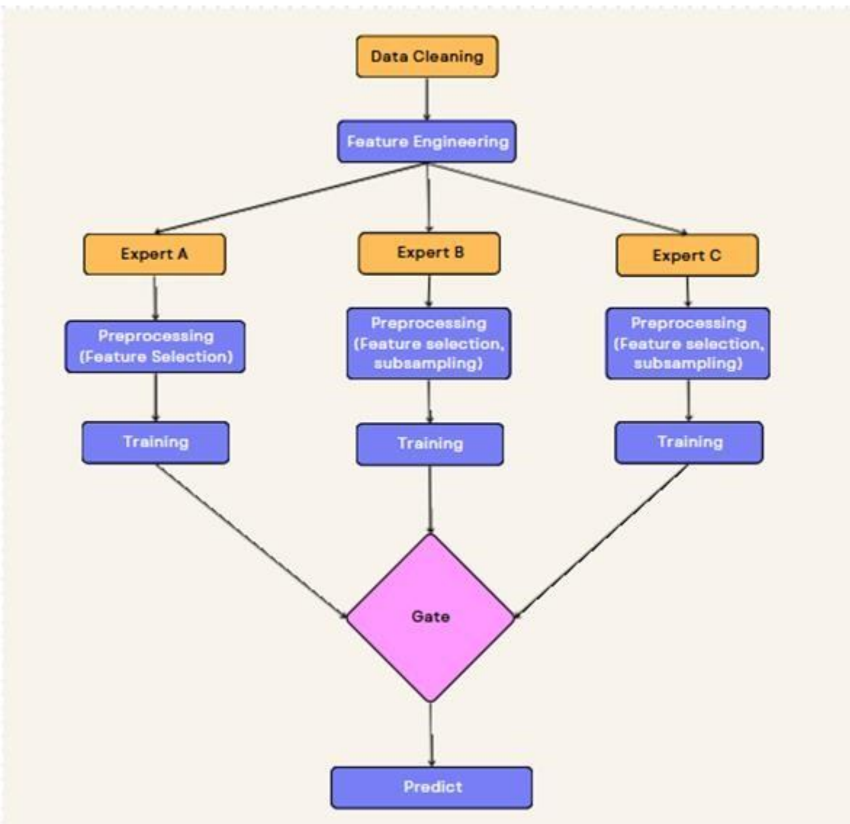

Figure 2 – Schematic representation of the model's workflow, detailing the steps from data cleansing to forecasting. The diagram illustrates the differentiation in preprocessing approaches for experts A, B, and C, culminating in the decision-making stage through the "Gate" for the generation of the final forecast.

6. TRAINING OF THE MODEL

The model was trained with the aim of optimizing three specific experts: Expert A, Expert B, and Expert C. Each of these experts has its own characteristics and nuances, making a robust, as well as meticulously optimized, validation strategy essential.

The technique of "MultilabelStratifiedKFold" with ten folds was utilized for the crucial task of validation, selected after demonstrating the best results in multiple optimizations. The subsequent challenge was to identify the columns that offered the best stratification of the data for each expert. After several rounds of optimization, the decision was made to use the same columns that were employed in the adversarial subsampling process for each expert.

The training process of the model involved the use of ten algorithms derived from Stratified Kfold. The complexity of the model and the diversity of metrics to be optimized resulted in variable training durations. However, on average, each fold iteration required approximately 110 seconds for its complete execution.

7. MODEL PERFORMANCE EVALUATION

To assess the performance of the proposed models, meticulous comparisons were conducted between the forecasted results and the actual metrics obtained during the testing phase. The metric employed was Weighted Mean Absolute Percentage Error (WMAPE). Indeed, to illustrate the model's efficacy, a comprehensive Table 4 was prepared with complete results. This table details how the model behaves under different conditions and configurations of the 5G network, offering an extensive view of its performance.

Analyzing the final cross-validation score (Cv) and private LeaderBoard (LB) score of three models revealed some peculiarities, as described below (Table 4):

- **Only XGBoost model:** This model exhibited a Cv final of 0.05730 and a private LB of 0.0968. Notably, XGBoost demonstrated superior performance in terms of Cv final when compared to ridge regression. However, it did not perform as well on the private LB. This observation suggests that while XGBoost proves to be efficient on the training data, its generalization and extrapolation to unknown data might be less effective.
- **Only ridge regression:** This model had a Cv final of 0.08373 and a private LB of 0.0955. Despite having a lower Cv final than XGBoost, ridge regression outperformed the latter on the private LB. This highlights the importance of feature transformation, which appears to be a crucial factor in improving the model's generalization on test data.
- **Boosted hybrid:** This is clearly the most robust model, with a Cv final of 0.05721 and a private LB of 0.0699. It not only excelled on the training data but also demonstrated significant outperformance on the actual

Table 4 - Performance of simple and hybrid models.

Strategy	Cv Expert A	Cv Expert B	Cv Expert C	Cv Final	Public LB	Private LB
Only XGBoost	0.04256	0.11232	0.04229	0.05730	0.0978	0.0968
Only Ridge Regression	0.06154	0.11986	0.11712	0.08373	0.0958	0.0955
Boosted Hybrid	0.04083	0.11743	0.04165	0.05721	0.0690	0.0699

The metrics CV experts A, B, and C assess cross-validation on distinct subsets using the WMAPE metric, demonstrating the robustness of each expert model. The CV final is the average of these values across the respective subsets. Public LB and Private LB indicate the model's performance on public and private test sets, respectively, evaluating the model's generalization capability and helping to prevent overfitting.

test data. This suggests that the combination of feature transformations with target adjustments resulted in a powerful hybrid model that exhibited superior performance across all aspects.

In conclusion, it was evident that feature transformation played a significant role in the efficacy of the model in a real-world environment. Furthermore, the boosted hybrid approach demonstrated itself to be the most effective, surpassing other models in both training and test metrics.

Table 5 - Performance with and without Mixture of Experts (MoE).

Strategy	Cv Expert A	Cv Expert B	Cv Expert C	Cv Final	Public LB	Private LB
Without MoE	0.04083	-	-	0.04083	0.1026	0.1027
With MoE	0.04083	0.11743	0.04165	0.05721	0.0690	0.0699

The metrics CV experts A, B, and C assess cross-validation on distinct subsets using the WMAPE metric, demonstrating the robustness of each expert model. The CV final is the average of these values across the respective subsets. Public LB and private LB indicate the model's performance on public and private test sets, respectively, evaluating the model's generalization capability and helping to prevent overfitting.

In Table 5, the efficacy of models with and without the implementation of the mixture of experts was evaluated. The key point of that analysis lay in the comparison between the generalization power of the two strategies, as assessed through the Cv final and private LB metrics.

It was observed that the "With MoE" model displayed a Cv final of 0.05721, which was higher than the "Without MoE" model that recorded a Cv final of 0.04083. However, when the private LB scores were evaluated, the "With MoE" model demonstrated significantly better performance, with a score of 0.0699, compared to the score of 0.1027 for the "Without MoE" model. This observation was especially relevant as it indicated that, despite a slightly higher Cv final, the "With MoE" model exhibited a considerably greater generalization power on the test set. It was noted that the Cv final and private LB scores for the "With MoE" model are closer to each other, corroborating its higher efficacy in adapting to new conditions and unseen variables during the training phase.

8. DISCUSSION AND FUTURE WORK

One of the most crucial findings of this study lay in the need for algorithms with high generalization power to tackle the specificities of the energy consumption scenario in 5G networks. Particularly, it was observed that the test dataset includes what was termed as "cold data," which was substantially different from the "hot data" used during the training phase.

The discrepancy between the two sets of data necessitated that the employed models were capable of interpreting and adapting to new information effectively. In this study, it was highlighted that the hybrid-boosted model with an approach inspired by the mixture of experts exhibits substantially more robust generalization power in test environments. It was able to interpret and extrapolate the "cold data" in an organized and intelligent manner, resulting in more precise and efficient generalization. This result not only reaffirms the efficacy of the hybrid-boosted model with a mixture of experts in practical situations but also signals a promising path for future research. The next logical step would be to delve deeper into how replacing the ridge regression algorithm with neural networks [9] could optimize the precision of predictions in the hybrid-boosted model. As for future enhancements inspired by the mixture of experts paradigm, the focus could be on [18]:

- Adaptive masks: for dynamic adjustment based on the reliability of predictions.
- Iterative feedback: for continuous refinement of the masks.

In summary, an algorithm's ability to adapt to new data while maintaining accuracy and efficiency is of paramount importance in the practical application for predicting energy consumption in 5G networks. This study serves as an initial milestone for future investigations aiming to develop even more robust and adaptable models.

REFERENCES

[1] United Nations, "Resolution Adopted by the General Assembly on 25 September 2015. Transforming Our World: The 2030 Agenda for Sustainable Development", 2015. Available: https://www.un.org/en/development/desa/ population/migration/generalassembly/doc s/globalcompact/A_RES_70_1_E.pdf.

[2] Organização das Nações Unidas, "Declaração Política da Cúpula dos ODS", 2023. Available: https://brasil.un.org/sites/default/files/202 3-09/ONU_DeclaracaoPolitica_CupulaDosODS_20230918.pdf.

[3] United Nations, "The Sustainable Development Goals Report 2023: Special Edition", 2023. Available: https://unstats.un.org/sdgs/report/2023/T he-Sustainable-Development-Goals-Report2023.pdf.

[4] GSMA, "Mobile Industry Impact Report: Sustainable Development Goals", 2023. Available: https://www.gsma.com/betterfuture/wpcontent/uploads/2023/09/2023-MobileIndustry-Impact-Report-Summary.pdf

[5] Sachs, J.D., Lafortune, G., Fuller, G., Drumm, E., "Implementing the SDG Stimulus. Sustainable Development Report 2023", Dublin: Dublin University Press, 2023. Available: https://s3.amazonaws.com/sustainabledeve lopment.report/2023/sustainabledevelopment-report-2023.pdf.

[6] Economic Commission for Latin America and the Caribbean (ECLAC), "Digital technologies for a new future (LC/TS.2021/43)", Santiago, 2021. Available: https://www.cepal.org/sites/default/files/publication/files/46817/S2000960_en.pdf.

[7] Ekholm, B., J. Rockström, “Digital technology can cut global emissions by 15%. Here’s how”, Cologny, World Economic Forum (WEF), 15 January [online], 2019. Available: https://www.weforum.org/agenda/2019/01/why-digitalization-is-the-key-toexponential-climate-action/.

[8] Malmodin, J., D. Lunden, “The energy and carbon footprint of the global ICT and E&M sectors 2010–2015”, Sustainability, 2018, 10(9), 1-31. Available: https://doi.org/10.3390/su10093027.

[9] Piovesan, N., López-Pérez, De Domenico, A., Geng, X., Bao, H., “Power Consumption Modeling of 5G Multi-Carrier Base Stations: A Machine Learning Approach,” IEEE International Conference on Communications (ICC), 28 May-1 June, 2023. Available: https://doi.org/10.48550/arXiv.2212.04318.