Residual LightGBM: A Tree Based 5G Base Station Energy Consumption Prediction Algorithm

Licheng Zheng

Synchromedia Lab , École de technologie supérieure, University of Quebec
Montreal, Canada
Email: licheng.zheng.1@ens.etsmtl.ca

***Abstract*—Addressing the challenge of base station energy consumption is increasingly important for environmentally-friendly networks. Predicting energy consumption entails comprehending the influence of configuration parameters and energy-saving modes, a task that is often complex. In this report, our solution introduces novel features and structures for predicting base station energy consumption. The effectiveness of our solution is validated through submission scores, demonstrating its efficiency.**

I. Introduction

Nowadays, the telecommunications industry is experiencing a surge in energy consumption. It accounts for approximately 3% of the world's total energy consumption, as reported in a study [1]. What's more, the cost of electricity is a significant financial burden for network operators, constituting as much as 25% of the expenses for an average mobile network operator (MNO). Within the broader operational expenditure (OPEX) for a mobile network operator, electricity costs alone make up a substantial 90% [2]. With the rapid adoption of 5G technology, the energy requirements of base stations have surged and can be up to three times higher compared to the previous 4G networks [3]. This necessitates the implementation of energy-efficient practices, such as optimizing base station parameters and adopting energy-saving techniques. However, understanding the complex and time-sensitive effects of these parameters and methods presents a challenging task, particularly in the context of energy consumption prediction.

This competition involves data that is divided into three main sections: Basic information about Base Stations, information at the cell level, and energy consumption data. The accuracy of predictions in this competition hinges on three significant factors: the processing of the data, the creation of relevant features, and the selection of the simulation model. We will elaborate on each of these factors in the upcoming sections.

The key contributions of our research are outlined as follows:

- **Data Processing**: We provide the details of data processing, encompassing tasks such as data merging and feature generation.
- **Feature Engineering**: We introduce techniques related to data engineering, such as Previous energy and Percentage of Features.
- **Residual Model**:We introduce a novel prediction model called Residual LightGBM, which builds upon LightGBM and is designed to address the problem of overfitting due to low bias.

II. Data Preprocessing

The competition dataset comprises three files, each containing distinct types of information. These files include the fundamental details of base stations, cell-level data for each hour, and energy consumption data for each hour. To facilitate our analysis, we combine these three files based on timestamps and base station identifiers. The details can be seen in [4]. Table I provide the details of features.

We also create some new features to increase the accuracy, including the difference and product between features. The feature created are presented in II

There are also some methods we tried but doesn't work well, there are: difference between ESMode, division of features, Summation of features, number of cells.

III. Feature Engineering

A. What worked

We did some experiments for on the features of historical data in the same hours a day. The best performance on previous days are given as following: the energy of past 5 days in the same hours, the features of past 6 days in the same hours, the difference on features of past 6 days in the same hours, the percentage on features of past 6 days.

B. What did not work

There are also some features which are not helpful. They are listed as following: skew, std, min, max, mean of features, hours from weekend, hours from month, is weekend or not, feature lagging, energy lagging, number of cells.

IV. Model Prediction

In this subsection, we provide the models and input of data.

TABLE I
FEATURE MERGING

features and cell	cell0	cell1	cell2	cell2
load	load_cell0	load_cell1	load_cell2	load_cell3
esmode1	esmode1_cell0	esmode1_cell1	esmode1_cell2	esmode1_cell3
esmode2	esmode2_cell0	esmode2_cell1	esmode2_cell2	esmode2_cell3
esmode3	esmode3_cell0	esmode3_cell1	esmode3_cell2	esmode3_cell3
esmode4	esmode4_cell0	esmode4_cell1	esmode4_cell2	esmode4_cell3
esmode5	esmode5_cell0	esmode5_cell1	esmode5_cell2	esmode5_cell3
esmode6	esmode6_cell0	esmode6_cell1	esmode6_cell2	esmode6_cell3
frequency	frequency_cell0	frequency_cell1	frequency_cell2	frequency_cell3
bandwidth	bandwidth_cell0	bandwidth_cell1	bandwidth_cell2	bandwidth_cell3
antennas	antennas_cell0	antennas_cell1	antennas_cell2	antennas_cell3
txpower	txpower_cell0	txpower_cell1	txpower_cell2	txpower_cell3

TABLE II
NEW FEATURES CREATED

Type	Features
difference	load_minus_frequency load_minus_bandwidth load_minus_txpower frequency_minus_bandwidth frequency_minus_antennas frequency_minus_txpower bandwidth_minus_antennas bandwidth_minus_txpower antennas_minus_txpower
product	load_product_frequency load_product_bandwidth load_product_txpower frequency_product_bandwidth frequency_product_antennas frequency_product_txpower bandwidth_product_antennas bandwidth_product_txpower antennas_product_txpower

A. Structure of Model

In this competition, we encounter two major challenges. Firstly, the data contains varying numbers of cells, which presents a challenge because it results in inputs with inconsistent shapes for Deep Neural Networks. Decision Tree-based algorithms are well-suited to addressing such issues. LightGBM, a highly efficient Gradient Boosting Decision Tree algorithm, falls into this category. However, boosting algorithms like LightGBM are iterative in nature. They adjust the sample weights based on previous iteration predictions. As a result, with each iteration, the errors tend to decrease, and the model's bias decreases accordingly [5]. Unfortunately, in this competition, the data is characterized by substantial noise, which can be problematic for LightGBM.To mitigate the impact of noise, we've introduced a solution called Residual LightGBM. This approach involves using predictions as inputs for LightGBM, effectively reducing the noise's influence and enhancing the model's robustness to noisy data.

The Residual LightGBM model comprises two LightGBM structures. Initially, a portion of the data is directed to the first LightGBM model. Subsequently, the output of the first LightGBM model is combined with the remaining data and passed to the second LightGBM model. The details are depicted in Fig. 1.

B. Data Input

The Input data contains the features in Section II, the energy in past 5 days in the same hours and the percentage of features in past 6 days. We will call the data of features in Section II as basic inputs, the energy in past 5 days as previous energy and the percentage of features in past 6 days as Precentage of Features. The Basic inputs will be send to both LightGBM1 and LightGBM2. The Percentage of Feature will be used as the input of LightGBM1. The output of LightGBM1 and previous energy are also the input of LightGBM2.

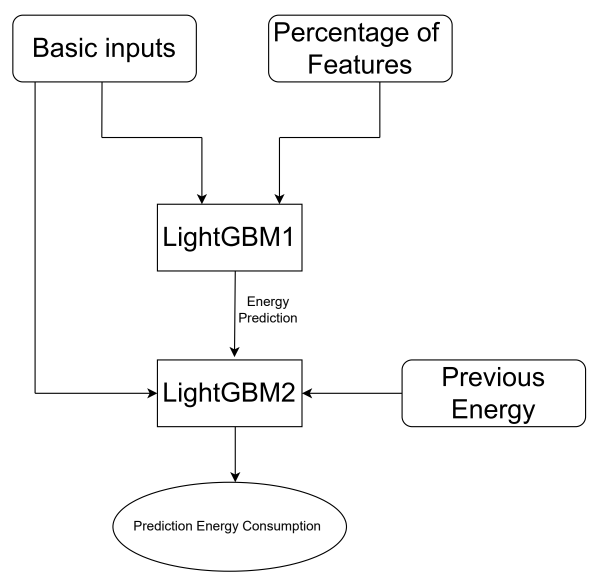

Fig. 1. An illustration of Residual LightGBM

V. SUBMISSION SCORES

We conducted experiments to demonstrate the effectiveness of our approach. This section is divided into two subsections. The first subsection illustrates how feature engineering contributes to improved scores, while the second one assesses the performance of Residual LightGBM.

A. *feature engineering*

In this part of our analysis, we conducted experiments to measure how the scores are enhanced by incorporating features related to differences, product, percentages of features, and previous energy values. Additionally, we evaluated the performance of these different features using the LightGBM model. The results of these experiments are presented in Table III.

TABLE III
SUBMISSION SCORES OF DIFFERENT FEATURES

features	Public Scores	Private Scores
Base features	0.124709515	0.123136091
Differences	0.113376462	0.111411138
Product	0.118879751	0.117040871
Percentages of Features	0.111652422	0.109514682
Previous Energy	0.110787858	0.109594089

B. *Residual LightGBM*

- **Residual LightGBM**: This is the Residual LightGBM we proposed in Section IV.

- **Residual LightGBM2**:We also provide another Residual LightGBM structure. In this structure, the Previous Energy switched to the inputs of LightGBM1. Meanwhile, the Percentage of Features become the inputs of LightGBM2. The structure is shown in Fig. 2.

- **LightGBM**: This is baseline of Residual LightGBM, in which Basic inputs, Percentage of Features and Previous Energy are the inputs of a single LightGBM structure. The structure is shown in Fig. 3.

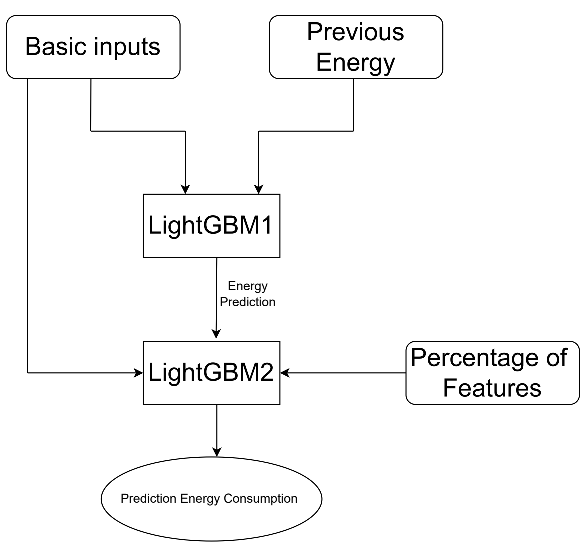

Fig. 2. Another Structure of Residual LightGBM

The results of these experiments are presented in Table IV

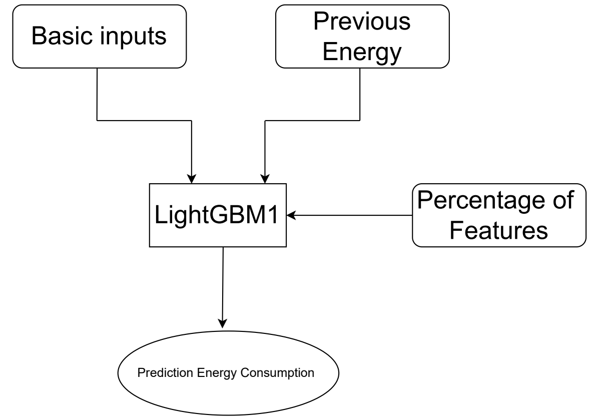

Fig. 3. An illustration of LightGBM

TABLE IV
SUBMISSION SCORES OF DIFFERENT STUCTURES

Structure	Public Scores	Private Scores
Residual LightGBM	0.093204536	0.092276470
Residual LightGBM2	0.101277934	0.100985862
LightGBM	0.105140196	0.104014936

VI. CONCLUSION

In this competition, we have machine learning techniques to forecast energy consumption, integrating new features and a novel structure into our approach. These features encompass Difference, Product, Feature Percentage, and Previous Energy. Our innovative structure is known as Residual LightGBM. The submission scores serve as a testament to the effectiveness of these algorithms.

References

[1] E. Obiodu, “The 5g guide: a reference for operators,” in *GSM Assocation*, 2019.

[2] T. Hatt and E. Kolta, “5g energy efficiencies: Green is the new black,” *GSMA Intelligence, White Paper*, vol. 11, 2020.

[3] N. Piovesan, D. Lopez-Perez, A. De Domenico, X. Geng, and H. Bao, “Power consumption modeling of 5g multi-carrier base stations: A machine learning approach,” *arXiv preprint arXiv:2212.04318*, 2022.

[4] G. Durgaprasad, “Sample solution - youtube video,” 2023. [Online]. Available: https://zindi.africa/competitions/aiml-for-5g-energy-consumption-modelling/discussions/18491

[5] G. Ke, Q. Meng, T. Finley, T. Wang, W. Chen, W. Ma, Q. Ye, and T.-Y. Liu, “Lightgbm: A highly efficient gradient boosting decision tree,” *Advances in neural information processing systems*, vol. 30, 2017.