AI/ML for 5G-Energy Consumption Modelling by ITU AI/ML in 5G Challenge

Table of Contents

Introduction

- **5G networks offer significant advantages over previous generations of cellular networks, but they also consume more energy.**
- **Modeling energy consumption of 5G products is challenging due to the variety of factors that affect it.**
- **Accurate modeling of energy consumption is essential for optimizing energy efficiency in 5G networks.**

The machine learning-based model for predicting energy consumption of 5G products was developed using a dataset of cell-level traffic statistics, base station data, and energy consumption data. The model achieved good performance on a holdout dataset, and it can be used by network operators to optimize energy efficiency in their 5G networks by:

- ❖ Identifying base stations that are consuming more energy than expected.
- ❖ Evaluating the impact of different energy-saving measures.
- ❖ Designing new 5G products that are more energy efficient.

The model can help network operators to reduce their energy costs and environmental impact.

Proposed solution

Data cleaning and preprocessing

Step 1: Merging the cell_level and base_station DataFrames on the BS and CellName columns

Step 2: Pivoting the merged DataFrame on the Time and BS columns, with the CellName column

Step 3: Resetting the index of the pivoted DataFrame

Step 4: Joining the pivoted DataFrame with a DataFrame containing the unique Mode and RUType

Step 5: Merging the pv DataFrame with the train and test DataFrame on the Time and BS columns

Feature engineering

Adding and counting features:

- ❖ Sum of base station features in row wise
- ❖ Count of unique values and how many values are appeared (non-NaN values)
- ❖ Clustered features
- ❖ Used PCA
- ❖ Grouped by BS and computed the max value
- ❖ Merged the resulting features on BS for count columns and sum columns

Feature engineering

Moving statics and lag features:

- Computing differences between consecutive values
- Computing ratios between consecutive values
- Computing rolling mean, standard deviation, maximum, minimum, quantiles, and forecast values of features
- Shifting features forward and backward to create lagged features

Use of a statistical model to predict load values

The Prophet model formula is `y(t) = g(t) + s(t) + h(t) + e(t)`

Feature engineering

Feature engineering in BS inside:

1. **Grouping the data by base station.**
2. **Calculating a variety of features for each base station, including:**
   - **Cosine similarity between the load time series and the energy time series.**
   - **Coefficients of the load, hour, and count features in a linear regression model predicting energy consumption.**
   - **Mean absolute error of the linear regression model predicting energy consumption.**
   - **Correlation coefficient between the first-order differences of the load and energy time series.**
3. **Training a random forest regressor model to predict energy consumption for each base station.**
4. **Calculating the mean absolute error of each model on the cross-validation set.**

Feature engineering

Target and mean encoding:

The following encoding techniques are used:

- **Encoding the BS and BS_hour features using the mean of the Energy target variable (target encoding)**
- **Encoding the load feature using the mean of the entire dataset (mean encoding)**
- **Encoding the diff1_eng and diff1_load_t+1 features (first-order differences of the Energy and load features) using target encoding**
- **Encoding the label_eng and label_load features (binary variables indicating whether the energy consumption or load increased) using target encoding**

Feature engineering

Probability of increasing with load and clustering :

- The probability of energy consumption and load increasing in the next hour is calculated based on the training data.
- The probability is calculated for each base station and for each hour class.
- cluster: The cluster number assigned to the base station.
- cluster_without_ru: The cluster number assigned to the base station, without considering the RUType feature.

get_prob_static(): Features that capture the historical energy consumption patterns of each base station.

get_cluster(): Features that capture the similarities between base stations.

Feature engineering

One hot encoding the quantile value:

- Use the `get_quantile()` function to calculate the quantiles of a column in a DataFrame.
- Create new columns in the DataFrame with the quantile bins, using the `_quantile` suffix.
- Use the `one_hot_encode()` function to create new columns for each unique value in the quantile columns.
- Assign each row to 1 for the column corresponding to the value of the original column in the input DataFrame.
- If the value of the original column is not present in the unique values for the column name, then the value of the new column is set to 0.

Model result and feature importance

WMAPE is 0.0745

diff1_eng_BS_hour_enc_	1329
pred_rf_bsc	1315
Energy_BS_hour_enc_	1314
diff1_load_t+1_BS_hour_enc_	1109
trend_forcast	897
dayofweekhour	861
corr_bsc	852
load_t+1_BS_hour_enc_	851
daydayofweekhour	840
diff1_hour_std	824
load_quantile_6.0	801
coef_load_bsc	792
label_eng_BS_hour_enc_	771
mae_cv_rf_bsc	768
coef_hour_bsc	755
trend_lower_forcast	748
all_sum_cols	737
label_load_BS_hour_enc_	715

Discussion

To improve the performance of the model and reduce the risk of data leakage, the following can be done:

- Use a deep learning model, such as a LSTM or GRU, which can often achieve better performance than traditional machine learning models.
- Use deep analysis to identify patterns and trends in the data, and create new features that are more informative and predictive of energy consumption.

Conclusion

In conclusion, this report has presented a machine learning-based model for predicting energy consumption of 5G products. The model was developed using a dataset of cell-level traffic statistics, base station data, and energy consumption data. The model achieved good performance on a holdout dataset, and it can be used by network operators to optimize energy efficiency in their 5G networks.

The model can be improved by using a larger dataset, using deep learning models, and using deep analysis and feature engineering to create more informative and predictive features.

The development of machine learning models for predicting energy consumption in 5G networks is a promising area of research with the potential to help network operators reduce their energy costs and environmental impact.