

---

---

# AI/ML for 5G-Energy Consumption Competition Solution

---

---

Ning Jia

---

---

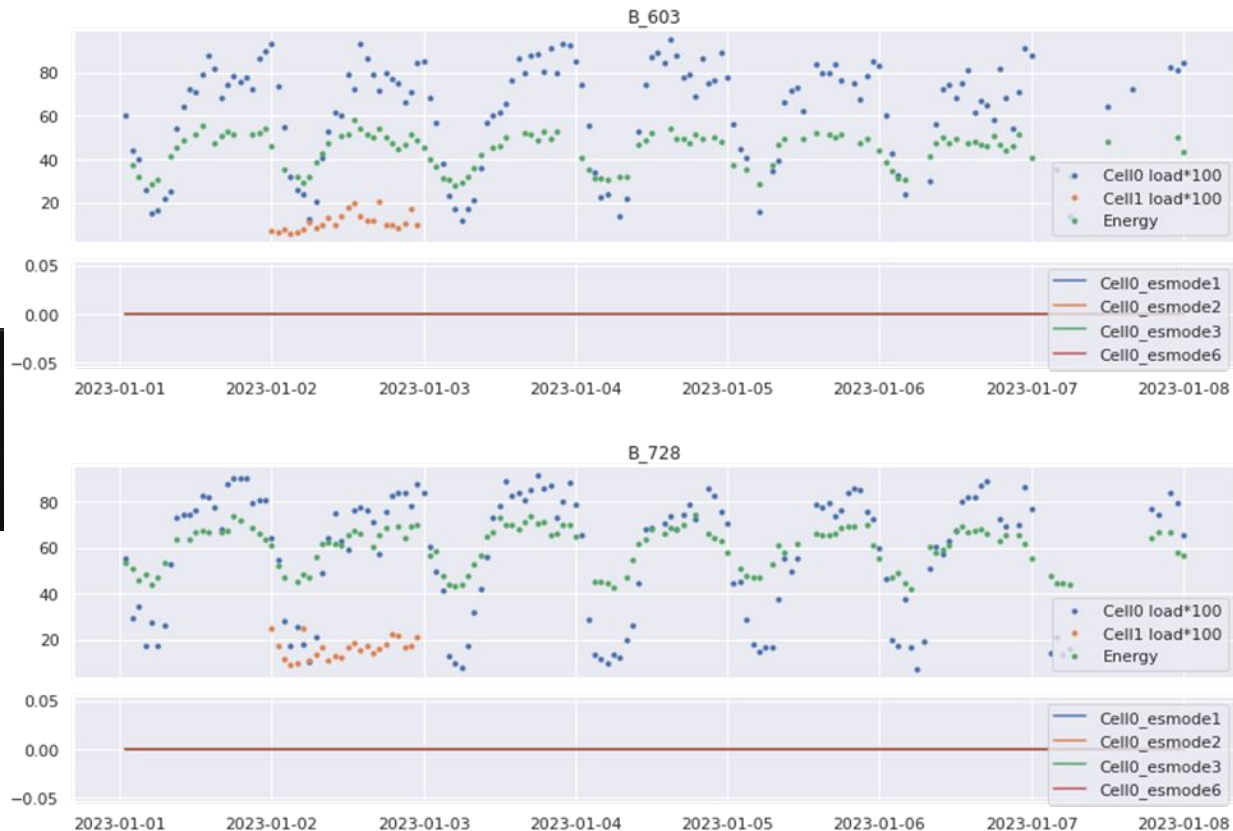
# Challenge: Unseen Stations and Unseen Configurations

About 40% of rows in the test dataset belong to unseen stations (weights of 5 is applied)

Rutype	Type 11	Type 12
Ratio of all test of unseen stations	28%	18%
Observations in Train	B_835, only 1 data point	B_854, only 1 data point
New Configurations	Frequency: 979.9980 Bandwidth: 8 Antennas: 64	Frequency: 715.9980 Bandwidth: 8

# Challenge: Hidden Factors Depend on Individual Stations

bs	cellname	rutype	mode	frequency	bandwidth	antennas	txpower
B_603	Cell0	Type1	Mode2	365.0000	20	2	6.8759
B_603	Cell1	Type1	Mode2	426.9800	2	2	7.3259
B_728	Cell0	Type1	Mode2	365.0000	20	2	6.8759
B_728	Cell1	Type1	Mode2	426.9800	2	2	7.3259



# Challenge: Not Use Any Future Data

Improper aggregation, smoothing, interpolation may introduce future data.

This not only applies to the target energy column, but also other columns such as the load, energy saving mode, etc.

# Algorithm and Preprocessing

Algorithm: LightGBM with KFold 10 folds

Preprocessing:

- Remove esmode4 and esmode5 because esmode4 is never enabled, and esmode5 is enabled for 3 data points.
- Remove the load of cell2 and cell3 because they are only available for stations 'B\_105' and 'B\_745'.
- Generate configuration groups. There are 84 unique groups.
- Pseudo-label 2 test stations of Type 11 and 1 test station of Type 12 by analytic models.

# Input Features: No Historical and Future Data

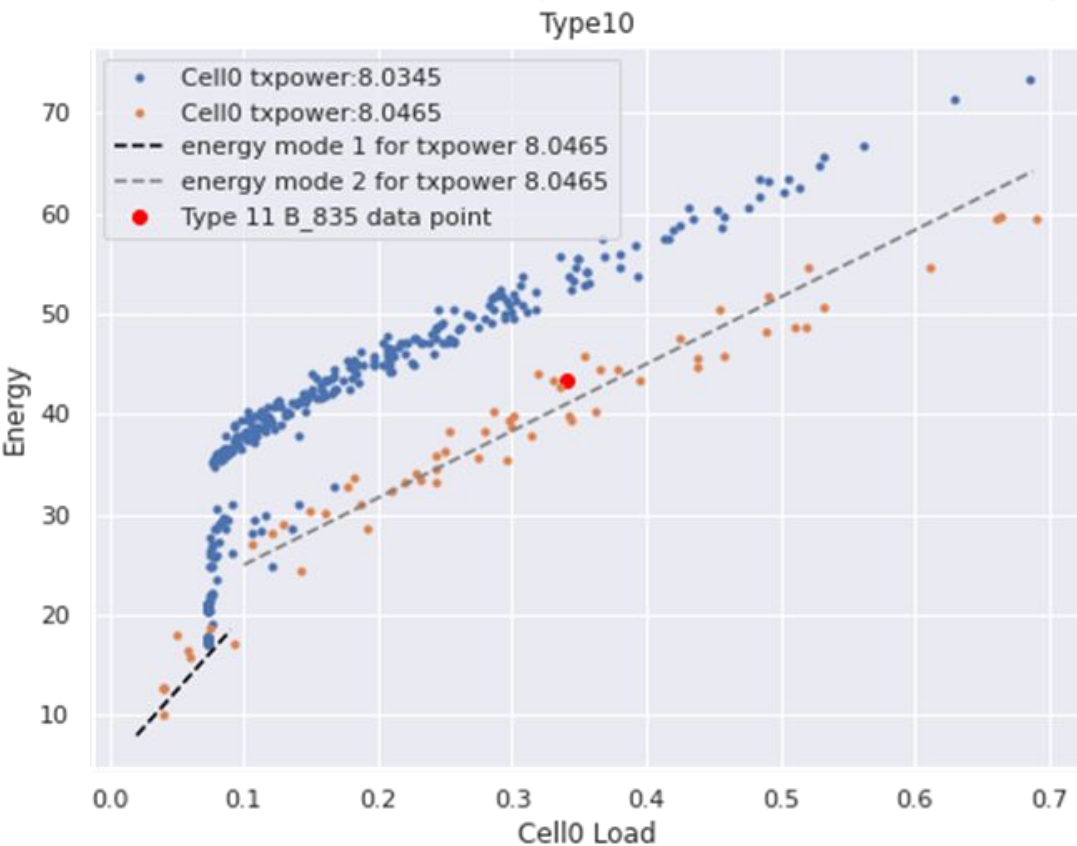
Category	Features
Load and energy-saving modes	Cell0_load,Cell0_esmode1,Cell0_esmode2,Cell0_esmode3,Cell0_esmode6 Cell1_load,Cell1_esmode1,Cell1_esmode2,Cell1_esmode3,Cell1_esmode6
Station configuration	rutype, cell_count, Cell0 antenna-txpower, Cell0 freq-band, mode, Cell0_antennas, Cell0_frequency, Cell0_txpower, Cell0_bandwidth Cell1_antennas, Cell1_frequency, Cell1_txpower, Cell1_bandwidth
Time	hour, day
Special configuration	configuration group, station Id

# Separate Models for Seen and Unseen Stations

Public Score: 0.081 Private Score: 0.081

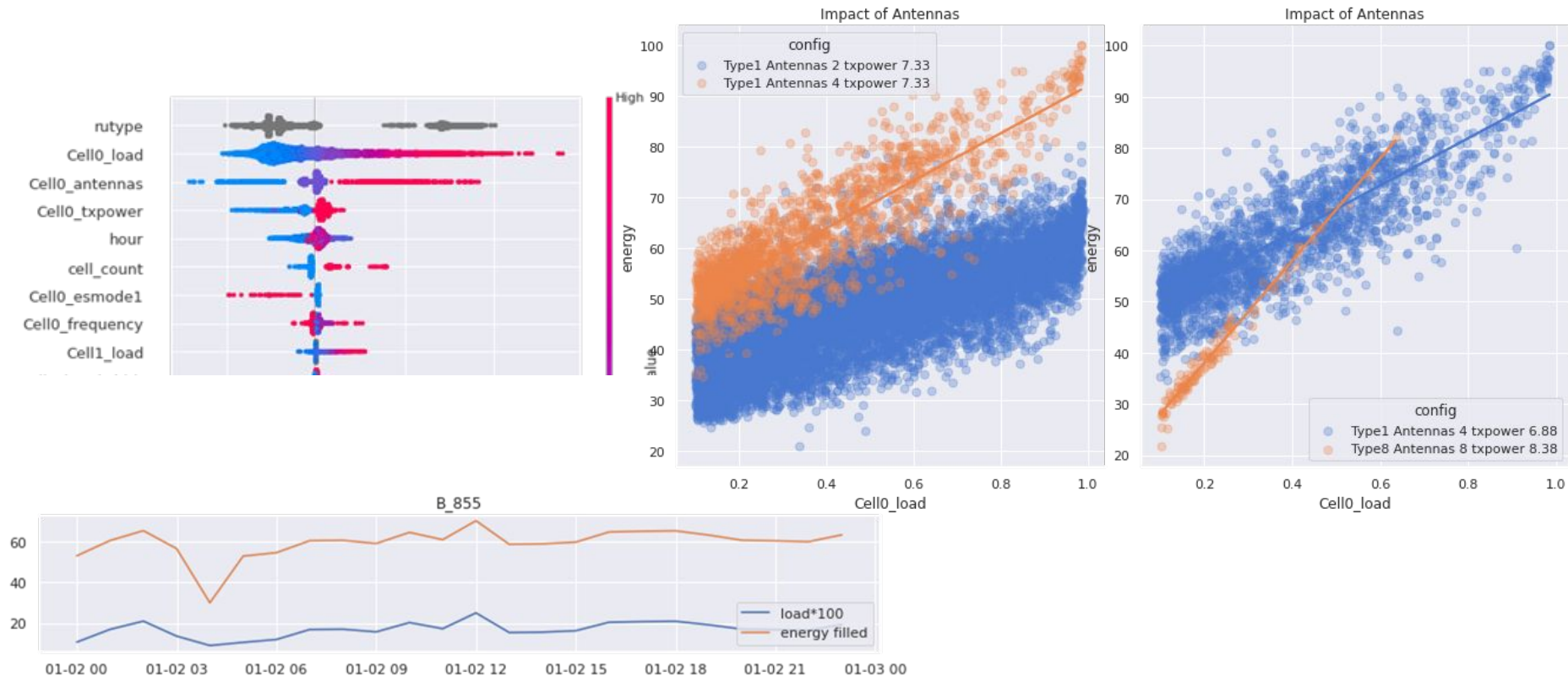
	Seen Test Stations	Unseen Test Stations
Input Features	27	25, Not using configuration group, station Id
Training Data	All train stations without B_835 and B_855	All train stations + 3 test stations(Type 11 and 12) with pseudo-labelled energy
Most Important Features	configuration group station Id rctype, Cell0 Load	rctype Cell0 load Cell0 antennas and txpower

# Pseudo-label Type 11 with analytic models





# Pseudo-label Type 12 with analytic models



Q & A

Thank you!