The objective is to build a machine-learning model capable of predicting the Actual Price of Electricity at a given time.

Time series problem

Make submission without creating any feature

Have all combined

Regression challenge

Metric we will use = MAE

The error metric for this competition is the Root Mean Squared Error.

For every row in the dataset, submission files should contain 2 columns: ID and "actual price".

If it has above 60% null values…. Remove it

Remove the 2 columns

Look for anomalies and outliers

Error of multicollineroity

Use uncorrelated features to form a new feature that is correlation with price actual

EDUCATION

pip install pandas numpy seaborn

C:\Users\USER>pip install scikit-learn

C:\Users\USER>python.exe -m pip install --upgrade pip

C:\Users\USER>pip install warnings

C:\Users\USER>pip install warning

C:\Users\USER>pip install catboost lightgbm xgboost

C:\Users\USER>Pip install sklearn