08/06/2023

# Final Project for ECE 514

Time Series Forecasting for Energy Consumption in Greece

PANAGIOTIS TOLOUDIS
AEM: 2995
Email: ptoloudis@uth.gr

# Find the Dataset.

The dataset to use for the project is from the organization Open Power System Data. More info about the Open Power System Data. This is a free-of-charge data platform dedicated to electricity system researchers. The project is a service provider to the modeling community: a supplier of a public good. It collects, checks, processes, documents, and publishes data that is publicly available but currently inconvenient to use. The platform provides data on installed generation capacity by country/technology, individual power plants (conventional and renewable), and time series data. The latter includes electricity consumption, spot prices, and wind and solar generation, both measured and derived from weather models. The scripts used for downloading and processing are open-source, version-controlled, and available through GitHub. For machine readability and easy integration, it follows the Data Package convention and publishes data primarily as CSV files and metadata as JSON files (Excel and SQLite files are also available). To facilitate reproducibility, all data as well as all scripts are version-controlled and available through stable URLs.

I select the data for the package version 2020-10-06. This dataset has 32 different European countries and have 3 different dataset the 15 min, 30 min and 60 min.
I work with the 60 min model because this have the data for the Greece.

# Prepare the dataset.

The dataset finds in the link: Data Platform – Open Power System Data (open-power-system-data.org)

First remove all the information to not use.

The clean dataset is in the Kaggle: Timeseries Greece Power | Kaggle

Second the dataset is too big from my machine, I keep the data from the 1/1/2018-0:00 to 10/1/2020-1:00.

# ARIMA

## What is the Arima model?

The ARIMA model is a statistical tool that can be used to analyze and forecast time series data. ARIMA stands for Auto-Regressive Integrated Moving Average, which means that the model combines three components:

- Auto-regression (AR): The model uses past values of the variable to predict its future values.
- Integration (I): The model applies differencing to the data to make it stationary, meaning that it has a constant mean and variance over time.
- Moving average (MA): The model incorporates the errors or noise from past predictions into the current prediction.

The ARIMA model is represented by three parameters: p, d, and q, which correspond to the order of the AR, I, and MA components respectively. For example, an ARIMA (1,1,1) model means that it has one AR term, one differencing term, and one MA term. The ARIMA model can be used to fit various types of time series data that have different patterns or trends.

In this model I split the dataset to 4 categories':

- GR_load_actual_entsoe_transparency
- GR_load_forecast_entsoe_transparency
- GR_solar_generation_actual
- GR_wind_onshore_generation_actual

## The GR_load_actual_entsoe_transparency.

First check the data for the stationarity. The p-value is 0.000, this is low of the 0.05. This shows the data is stationary.

## The model results.

Best model: ARIMA (4,1,4) (0,0,0)[0]

Total fit time: 1722.682 seconds
SARIMAX Results:

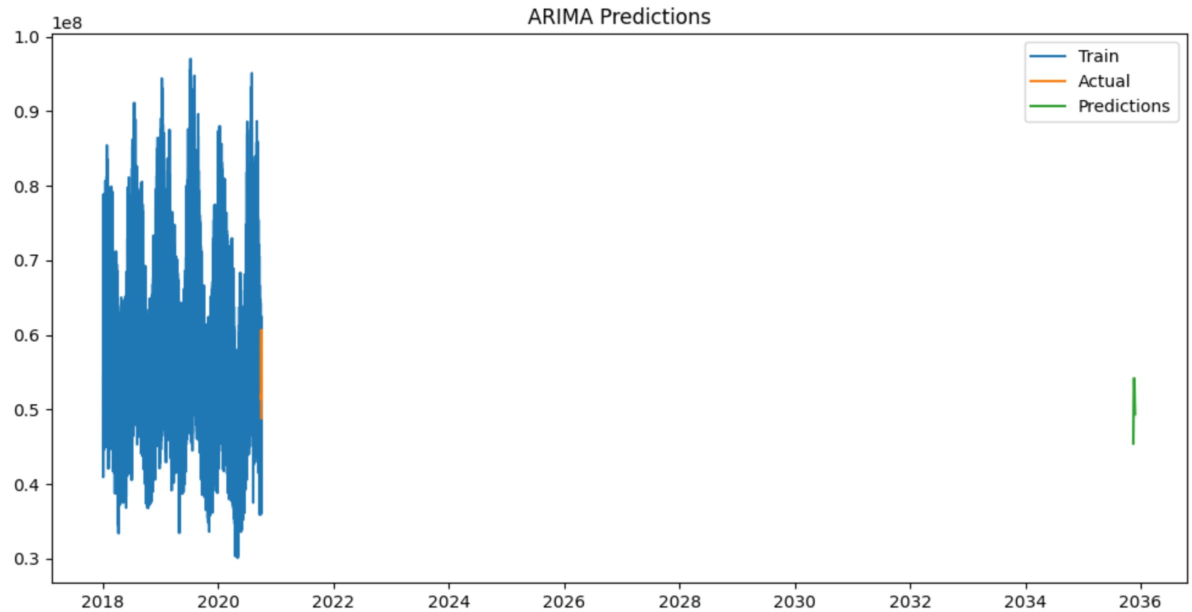
Dep. Variable: y
No. Observations: 24072
Model: SARIMAX (4, 1, 4)
Log Likelihood -381081.931
Date: Wed, 10 May 2023 AIC 762181.863
Time: 17:11:59 BIC 762254.662
Sample: 0 HQIC 762205.469 - 24072
Covariance Type: opg

==================================================================

| | coef | std | err | z | P>\|z\| | [0.025 0.975] |
|---|---|---|---|---|---|---|
| ar.L1 | 0.2531 | 0.005 | 53.451 | 0.000 | 0.244 | 0.262 |
| ar.L2 | 0.5698 | 0.003 | 197.920 | 0.000 | 0.564 | 0.575 |
| ar.L3 | 0.5487 | 0.003 | 169.169 | 0.000 | 0.542 | 0.555 |
| ar.L4 | -0.7740 | 0.005 | -153.919 | 0.000 | -0.784 | -0.764 |
| ma.L1 | 0.6229 | 0.006 | 97.791 | 0.000 | 0.610 | 0.635 |
| ma.L2 | -0.2955 | 0.005 | -56.519 | 0.000 | -0.306 | -0.285 |
| ma.L3 | -1.0125 | 0.005 | -200.536 | 0.000 | -1.022 | -1.003 |
| ma.L4 | -0.1897 | 0.007 | -27.201 | 0.000 | -0.203 | -0.176 |
| sigma2 | 3.81e+12 | 1.51e -15 | 2.52e+27 | 0.000 | 3.81e+12 | 3.81e+12 |

==================================================================

Ljung-Box (L1) (Q): 7.57
Jarque-Bera (JB): 132785.25
Prob(Q): 0.01
Prob (JB): 0.00
Heteroskedasticity (H): 0.83
Skew: -0.32
Prob(H) (two-sided): 0.00
Kurtosis: 14.49




## The GR_load_forecast_entsoe_transparency.

First check the data for the stationarity. The p-value is 0.000, this is low of the 0.05. This shows the data is stationary.

## The model results.

Best model: ARIMA (2,1,5) (0,0,0) [0]
Total fit time: 940.954 seconds
SARIMAX Results
Dep. Variable: y
No. Observations: 24072
Model: SARIMAX (2, 1, 5)
Log Likelihood -381716.491
Date: Wed, 10 May 2023
AIC 763448.981
Time: 17:28:59 BIC 763513.691
Sample: 0 HQIC 763469.964 - 24072
Covariance Type: opg

==============================================================================

| | coef | std err | z | P>\|z\| | [0.025 0.975] |
|---|---|---|---|---|---|
| ar.L1 | 1.8561 | 0.003 | 735.650 | 0.000 1.851 | 1.861 |
| ar.L2 | -0.9258 | 0.002 | -401.057 | 0.000 -0.930 | -0.921 |
| ma.L1 | -1.1370 | 0.005 | -232.837 | 0.000 -1.147 | -1.127 |

| | | | |
|---|---|---|---|
| ma.L2 -0.1368 | 0.008 -16.100 | 0.000 -0.153 | -0.120 |
| ma.L3 0.1390 | 0.010 13.863 | 0.000 0.119 | 0.159 |
| ma.L4 -0.0374 | 0.010 -3.833 | 0.000 -0.056 | -0.018 |
| ma.L5 0.2180 | 0.007 29.409 | 0.000 0.203 | 0.233 |
| sigma2 4.029e+12 | 1.59e-15 2.53e+27 | 0.000 4.03e+12 4.03e+12 | |

==============================================================================

Ljung-Box (L1) (Q): 10.48

Jarque-Bera (JB): 86014.36

Prob(Q): 0.00

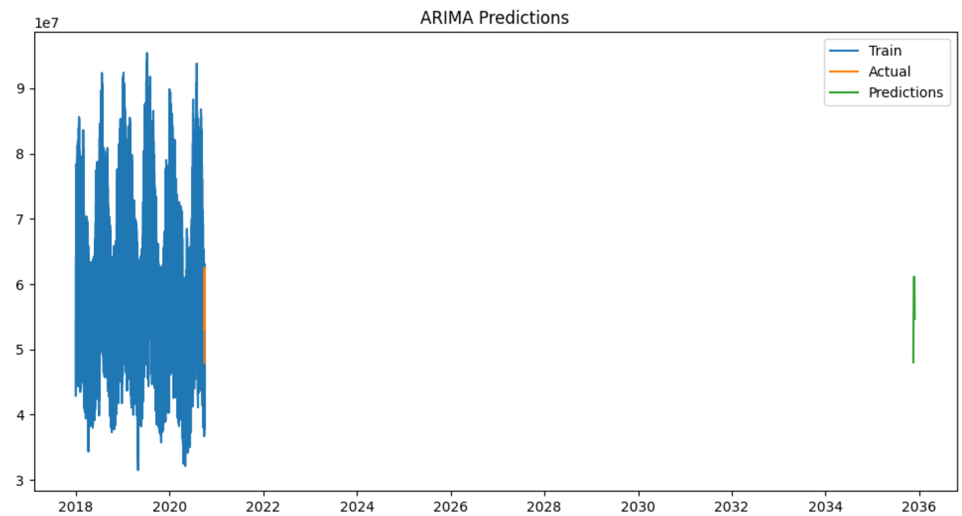
Prob (JB): 0.00

Heteroskedasticity (H): 0.94

Skew: -0.13

Prob(H) (two-sided): 0.01

Kurtosis: 12.26




## The GR_solar_genaration_actual.

First check the data for the stationarity. The p-value is 0.000, this is low of the 0.05. This shows the data is stationary.

## The model results.

Best model: ARIMA (5,1,0) (0,0,0) [0]

Total fit time: 312.852 seconds

SARIMAX Results:

Dep. Variable: y

No. Observations: 24072

Model: SARIMAX (5, 1, 0)

Log Likelihood -361085.348

Date: Wed, 10 May 2023 AIC 722182.695

Time: 17:35:01 BIC 722231.228

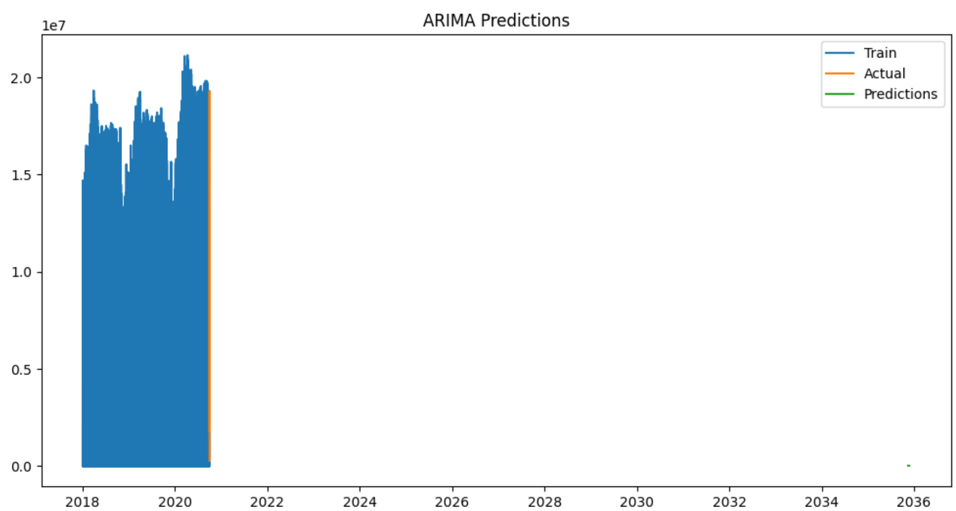
Sample: 0 HQIC 722198.432 - 24072

Covariance Type: opg
================================================================================== coef
std err z P>|z| [0.025 0.975]

| | | | | | | |
|---|---|---|---|---|---|---|
| ar.L1 | 1.3561 | 0.002 | 634.320 | 0.000 | 1.352 | 1.360 |
| ar.L2 | -0.6724 | 0.002 | -295.743 | 0.000 | -0.677 | -0.668 |
| ar.L3 | 0.0998 | 0.003 | 36.040 | 0.000 | 0.094 | 0.105 |
| ar.L4 | 0.0515 | 0.004 | 14.056 | 0.000 | 0.044 | 0.059 |
| ar.L5 | -0.1322 | 0.004 | -35.872 | 0.000 | -0.139 | -0.125 |
| sigma2 | 6.269e+11 | 3.28e-15 | 1.91e+26 | 0.000 | 6.27e+11 | 6.27e+11 |

==================================================================================

Ljung-Box (L1) (Q): 4.23
Jarque-Bera (JB): 1220839.20
Prob(Q): 0.04
Prob (JB): 0.00
Heteroskedasticity (H): 1.14
Skew: 0.67
Prob(H) (two-sided): 0.00
Kurtosis: 37.86




## The GR_wind_onshore_generation_actual.

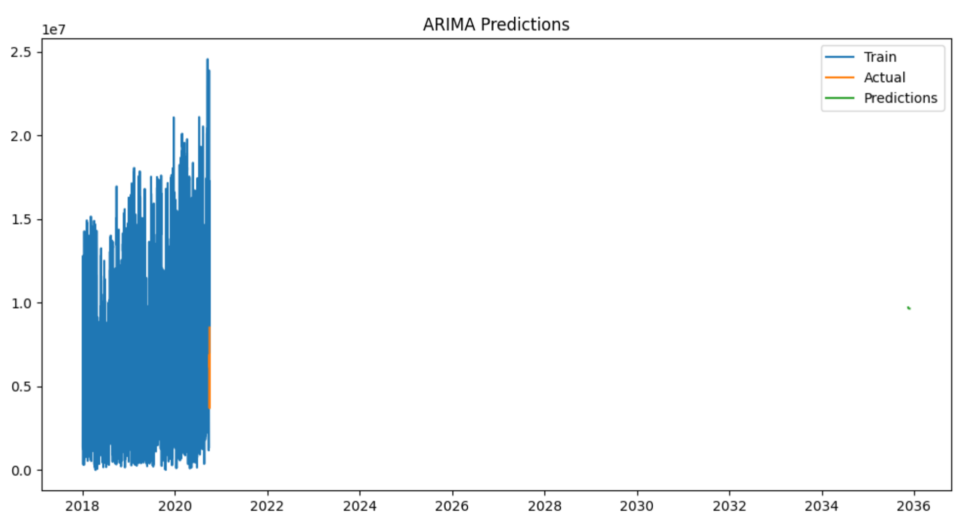
First check the data for the stationarity. The p-value is 0.000, this is low of the 0.05. This shows the data is stationary.

## The model results.

Best model: ARIMA (1,1,2) (0,0,0) [0]
Total fit time: 208.810 seconds

SARIMAX Results
Dep. Variable: y
No. Observations: 24072
Model: SARIMAX (1, 1, 2)
Log Likelihood -347855.364
Date: Wed, 10 May 2023 AIC 695718.728
Time: 17:38:38 BIC 695751.083
Sample: 0 HQIC 695729.219 - 24072
Covariance Type: opg

============================================================================== coef
std err z P>|z| [0.025 0.975]
ar.L1 0.6329 0.012 51.964 0.000 0.609 0.657
ma.L1 -0.0156 0.013 -1.199 0.231 -0.041 0.010
ma.L2 -0.0740 0.009 -8.239 0.000 -0.092 -0.056
sigma2 2.088e+11 1.34e-14 1.56e+25 0.000 2.09e+11 2.09e+11
==============================================================================
Ljung-Box (L1) (Q): 0.00
Jarque-Bera (JB): 7833.46
Prob(Q): 0.98
Prob (JB): 0.00
Heteroskedasticity (H): 1.62
Skew: -0.01
Prob(H) (two-sided): 0.00
Kurtosis: 5.79



# LSTM

## What is the LSTM model?

The LSTM model is a type of recurrent neural network (RNN) that can process sequential data, such as speech, text, or video. Unlike standard RNNs, LSTM has a special structure that allows it to remember information over long periods of time. An LSTM cell consists of four components: a cell state, an input gate, an output gate, and a forget gate. The cell state stores the long-term memory of the network, while the gates control how much information to keep or discard from the cell state and the inputs. LSTM can learn complex patterns and dependencies in sequential data and is widely used for tasks such as natural language processing, speech recognition, machine translation, and more.

IMPORTANT: This model not work for me to split the data in categories.

## The model results.

The evaluate of the model is 0.03626 in the test data.

# PROPHET

## What is the Prophet model?

The Prophet model is a procedure for forecasting time series data based on an additive model where non-linear trends are fit with yearly, weekly, and daily seasonality, plus holiday effects. It works best with time series that have strong seasonal effects and several seasons of historical data. The Prophet model consists of four main components: a piecewise linear or logistic growth curve trend, a yearly seasonal component modeled using Fourier series, a weekly seasonal component modeled using dummy variables, and holiday effects modeled as dummy variables. The model is estimated using a Bayesian approach and is robust to missing data, shifts in the trend, and outliers. The Prophet model is designed to be simple for non-experts to use, yet flexible and powerful.

In this model I split the dataset to 4 categories':

- GR_load_actual_entsoe_transparency
- GR_load_forecast_entsoe_transparency

- GR_solar_generation_actual
- GR_wind_onshore_generation_actual

Show the first 5 hours of the predict.

## The GR_load_actual_entsoe_transparency.

```
      ds                  trend       yhat_lower yhat_upper    trend_lower
0 2018-01-01 00:00:00 5.115150e+07 4.289080e+07 5.383322e+07 5.115150e+07
1 2018-01-01 01:00:00 5.115819e+07 3.897426e+07 5.118129e+07 5.115819e+07
2 2018-01-01 02:00:00 5.116489e+07 3.780183e+07 4.940105e+07 5.116489e+07
3 2018-01-01 03:00:00 5.117158e+07 3.630325e+07 4.830509e+07 5.117158e+07
4 2018-01-01 04:00:00 5.117827e+07 3.619573e+07 4.747289e+07 5.117827e+07

   trend_upper additive_terms additive_terms_lower additive_terms_upper
0 5.115150e+07 -2.996816e+06 -2.996816e+06        -2.996816e+06
1 5.115819e+07 -5.902448e+06 -5.902448e+06        -5.902448e+06
2 5.116489e+07 -7.720492e+06 -7.720492e+06        -7.720492e+06
3 5.117158e+07 -8.874031e+06 -8.874031e+06        -8.874031e+06
4 5.117827e+07 -9.290518e+06 -9.290518e+06        -9.290518e+06

      daily     ...     weekly      weekly_lower weekly_upper     yearly
0 -6.314450e+06 ... -3.336074e+06 -3.336074e+06 -3.336074e+06 6.653709e+06
1 -9.470504e+06 ... -3.106071e+06 -3.106071e+06 -3.106071e+06 6.674127e+06
2 -1.154859e+07 ... -2.866484e+06 -2.866484e+06 -2.866484e+06 6.694581e+06
3 -1.297007e+07 ... -2.619027e+06 -2.619027e+06 -2.619027e+06 6.715070e+06
4 -1.366065e+07 ... -2.365456e+06 -2.365456e+06 -2.365456e+06 6.735593e+06

  yearly_lower yearly_upper multiplicative_terms
0 6.653709e+06 6.653709e+06 0.0
1 6.674127e+06 6.674127e+06 0.0
2 6.694581e+06 6.694581e+06 0.0
3 6.715070e+06 6.715070e+06 0.0
4 6.735593e+06 6.735593e+06 0.0

  multiplicative_terms_lower multiplicative_terms_upper yhat
0 0.0                          0.0                        4.815469e+07
1 0.0                          0.0                        4.525575e+07
2 0.0                          0.0                        4.344439e+07
3 0.0                          0.0                        4.229754e+07
4 0.0                          0.0                        4.188775e+07
```

## The GR_load_forecast_entsoe_transparency.

```
      ds                  trend       yhat_lower yhat_upper trend_lower
```

```
0 2018-01-01 00:00:00 4.998724e+07 4.280840e+07 5.390222e+07 4.998724e+07
1 2018-01-01 01:00:00 4.999596e+07 3.897755e+07 5.060132e+07 4.999596e+07
2 2018-01-01 02:00:00 5.000469e+07 3.736401e+07 4.865905e+07 5.000469e+07

3 2018-01-01 03:00:00 5.001341e+07 3.630386e+07 4.834796e+07 5.001341e+07
4 2018-01-01 04:00:00 5.002213e+07 3.604450e+07 4.705833e+07 5.002213e+07

    trend_upper additive_terms additive_terms_lower additive_terms_upper
0 4.998724e+07 -1.729728e+06 -1.729728e+06            -1.729728e+06
1 4.999596e+07 -4.843984e+06 -4.843984e+06            -4.843984e+06
2 5.000469e+07 -6.777526e+06 -6.777526e+06            -6.777526e+06
3 5.001341e+07 -7.955646e+06 -7.955646e+06            -7.955646e+06
4 5.002213e+07 -8.335098e+06 -8.335098e+06            -8.335098e+06

      daily     ... weekly        weekly_lower   weekly_upper yearly
0 -6.435641e+06 ... -3.141942e+06 -3.141942e+06 -3.141942e+06 7.847854e+06
1 -9.802059e+06 ... -2.909800e+06 -2.909800e+06 -2.909800e+06 7.867874e+06
2 -1.199596e+07 ... -2.669486e+06 -2.669486e+06 -2.669486e+06 7.887922e+06
3 -1.344094e+07 ... -2.422706e+06 -2.422706e+06 -2.422706e+06 7.907997e+06
4 -1.409200e+07 ... -2.171191e+06 -2.171191e+06 -2.171191e+06 7.928098e+06

  yearly_lower yearly_upper multiplicative_terms
0 7.847854e+06 7.847854e+06 0.0
1 7.867874e+06 7.867874e+06 0.0
2 7.887922e+06 7.887922e+06 0.0
3 7.907997e+06 7.907997e+06 0.0
4 7.928098e+06 7.928098e+06 0.0

   multiplicative_terms_lower multiplicative_terms_upper yhat
0 0.0                               0.0                4.825751e+07
1 0.0                               0.0                4.515198e+07
2 0.0                               0.0                4.322716e+07
3 0.0                               0.0                4.205776e+07
4 0.0                               0.0                4.168704e+07
```

## The GR_solar_generation_actual.

```
      ds                  trend        yhat_lower yhat_upper trend_lower
0 2018-01-01 00:00:00 3.812957e+06 -5.327350e+06 3.778284e+05 3.812957e+06
1 2018-01-01 01:00:00 3.812985e+06 -5.193556e+06 3.874846e+05 3.812985e+06
2 2018-01-01 02:00:00 3.813012e+06 -5.114559e+06 5.456537e+05 3.813012e+06
3 2018-01-01 03:00:00 3.813040e+06 -4.996994e+06 1.014039e+06 3.813040e+06
4 2018-01-01 04:00:00 3.813067e+06 -4.986908e+06 6.374663e+05 3.813067e+06

   trend_upper additive_terms additive_terms_lower additive_terms_upper
0 3.812957e+06 -6.235219e+06 -6.235219e+06            -6.235219e+06
```

```
1 3.812985e+06 -6.269811e+06 -6.269811e+06             -6.269811e+06
2 3.813012e+06 -6.070745e+06 -6.070745e+06             -6.070745e+06
3 3.813040e+06 -5.905784e+06 -5.905784e+06             -5.905784e+06
4 3.813067e+06 -6.056603e+06 -6.056603e+06             -6.056603e+06

      daily     ...       weekly     weekly_lower weekly_upper yearly
0 -4.399489e+06 ... -49700.446557 -49700.446557 -49700.446557 -1.786029e+06
1 -4.429838e+06 ... -54457.352429 -54457.352429 -54457.352429 -1.785516e+06
2 -4.226793e+06 ... -58949.774992 -58949.774992 -58949.774992 -1.785003e+06
3 -4.058160e+06 ... -63135.109986 -63135.109986 -63135.109986 -1.784489e+06
4 -4.205653e+06 ... -66973.909894 -66973.909894 -66973.909894 -1.783976e+06
     yearly_lower yearly_upper multiplicative_terms
0 -1.786029e+06 -1.786029e+06 0.0
1 -1.785516e+06 -1.785516e+06 0.0
2 -1.785003e+06 -1.785003e+06 0.0
3 -1.784489e+06 -1.784489e+06 0.0
4 -1.783976e+06 -1.783976e+06 0.0

  multiplicative_terms_lower multiplicative_terms_upper yhat
0 0.0                          0.0                        -2.422262e+06
1 0.0                          0.0                        -2.456826e+06
2 0.0                          0.0                        -2.257733e+06
3 0.0                          0.0                        -2.092745e+06
4 0.0                          0.0                        -2.243536e+06
```

## The GR_wind_onshore_generation_actual.

```
      ds                  trend        yhat_lower   yhat_upper
0 2018-01-01 00:00:00 4.361669e+06 -230469.052197 1.077036e+07
1 2018-01-01 01:00:00 4.362051e+06 -470565.292179 1.023173e+07
2 2018-01-01 02:00:00 4.362432e+06 -5299.662021   1.066581e+07
3 2018-01-01 03:00:00 4.362814e+06 -328813.833307 1.073347e+07
4 2018-01-01 04:00:00 4.363196e+06 -120597.679962 1.104952e+07

   trend_lower trend_upper additive_terms additive_terms_lower
0 4.361669e+06 4.361669e+06 7.685618e+05 7.685618e+05
1 4.362051e+06 4.362051e+06 7.956175e+05 7.956175e+05
2 4.362432e+06 4.362432e+06 8.543475e+05 8.543475e+05
3 4.362814e+06 4.362814e+06 9.488131e+05 9.488131e+05
4 4.363196e+06 4.363196e+06 1.059700e+06 1.059700e+06

  additive_terms_upper         daily    ...      weekly       weekly_lower
0 7.685618e+05           -131857.178234... -371711.009690 -371711.009690
1 7.956175e+05           -97225.030560 ... -378656.403487 -378656.403487
2 8.543475e+05           -33073.443969 ... -383438.959196 -383438.959196
3 9.488131e+05           64620.384804  ... -386020.132445 -386020.132445
```

```
4 1.059700e+06          176525.371365 ... -386383.242825 -386383.242825

    weekly_upper     yearly   yearly_lower yearly_upper
0 -371711.009690 1.272130e+06 1.272130e+06 1.272130e+06
1 -378656.403487 1.271499e+06 1.271499e+06 1.271499e+06
2 -383438.959196 1.270860e+06 1.270860e+06 1.270860e+06
3 -386020.132445 1.270213e+06 1.270213e+06 1.270213e+06
4 -386383.242825 1.269558e+06 1.269558e+06 1.269558e+06


  multiplicative_terms multiplicative_terms_lower
0 0.0                        0.0
1 0.0                        0.0
2 0.0                        0.0
3 0.0                        0.0
4 0.0                        0.0

multiplicative_terms_upper         yhat
0 0.0                        5.130231e+06
1 0.0                        5.157668e+06
2 0.0                        5.216780e+06
3 0.0                        5.311627e+06
4 0.0                        5.422896e+06
```

# XGBOOST

## What is the Xgboost model?

XGBoost is a machine learning framework that uses gradient boosting to create accurate and efficient predictive models. Gradient boosting is a technique that combines multiple weak learners (such as decision trees) into a strong learner by iteratively adding new learners that correct the errors of the previous ones. XGBoost stands for extreme Gradient Boosting, and it is designed to be scalable, fast, and robust to various types of data and problems.

In this model I split the dataset to 4 categories':

- GR_load_actual_entsoe_transparency
- GR_load_forecast_entsoe_transparency
- GR_solar_generation_actual
- GR_wind_onshore_generation_actual

## The GR_load_actual_entsoe_transparency.

| Prediction | Actual |
|---|---|
| 66201820 | 66171500 |
| 59507024 | 59467600 |
| 41082864 | 41086100 |
| 48974040 | 48956900 |
| 66303360 | 66414500 |
| 63670044 | 63641800 |

The Mean Squared Error: 4179604559.307996

## The GR_load_forecast_entsoe_transparency.

| Prediction | Actual |
|---|---|
| 51595084 | 51611600 |
| 40720100 | 40674700 |
| 76391744 | 76450000 |
| 60933364 | 60950300 |
| 55472148 | 55570000 |
| 63189272 | 63150000 |

The Mean Squared Error: 3971697205.2060227

## The GR_solar_generation_actual.

| Prediction | Actual |
|---|---|
| 1.6871027e+02 | 0.0 |
| 1.6562167e+07 | 16590000.0 |
| 1.5606325e+07 | 15590000.0 |
| 1.6871027e+02 | 0.0 |
| 1.5405820e+07 | 15390000.0 |
| 1.6871027e+02 | 0.0 |

The Mean Squared Error: 178525867.4920639

## The GR_wind_onshone_generation_actual.

| Prediction | Actual |
|---|---|
| 17332950 | 17390000.0 |

| | |
|---|---|
| 1454575.5 | 1460000.0 |
| 11891878 | 11890000.0 |
| 4290031.5 | 4290000.0 |
| 13257710 | 13260000.0 |
| 19425660 | 19420000.0 |

The Mean Squared Error: 227438983.1501955.

# AutoML

## What is the AutoML model?

The AutoML model is a type of machine learning model that can automatically select the best architecture, hyperparameters, and data preprocessing steps for a given task. The AutoML model can save time and resources for developers and researchers who want to build and deploy machine learning solutions without requiring extensive expertise or manual tuning. The AutoML model can also discover novel and efficient architectures that outperform human-designed ones in some cases. The AutoML model is not a single algorithm, but a general framework that can be applied to different domains such as computer vision, natural language processing, and tabular data analysis.

## The suggestions model.

### H2O.ai

The h2o.ai model is a machine learning platform that enables users to build and deploy scalable and accurate predictive models. The platform offers a variety of tools and features, such as automated data preparation, feature engineering, model selection, explain ability, and deployment. The h2o.ai model can handle structured and unstructured data, as well as supervised and unsupervised learning tasks. The h2o.ai model is designed to be user-friendly, fast, and flexible, making it suitable for various domains and applications.

### DataRobot

The model DataRobot is a software that automates the process of building and deploying machine learning models. It uses a variety of algorithms, mostly based on open source, to train and evaluate thousands of models on a given dataset. It also provides tools for data quality assessment, feature engineering, model interpretation, and MLOps. The model DataRobot can help users of different skill levels to leverage the power of artificial intelligence and make better decisions based on data.

### Google Cloud AutoML

Google Cloud AutoML is a service that allows users to create custom machine learning models without coding. Users can upload their own data, choose a model type, and train the model using Google's infrastructure. The service also provides tools for evaluating, deploying, and managing the models. Google Cloud AutoML supports various types of data, such as images, text, tabular, and video.

## The different between the models.

The H2O.ai and DataRobot offer more specialized features and capabilities for time series analysis compared to Google Cloud AutoML. H2O.ai provides the H2O AutoML framework with specific support for time series modeling, while DataRobot offers automated time series modeling as part of its comprehensive platform. On the other hand, Google Cloud AutoML provides a general-purpose machine learning platform that can be used for time series analysis by utilizing its broader set of tools and infrastructure.
I use the H20.ai model.

IMPORTANT: This model not work for me to split the data in categories.

## The result:

```
** Reported on train data. **
MSE: 3517418035650.899
RMSE: 1875478.0818902948
MAE: 1386162.5148232258
RMSLE: 0.031496644655704585
Mean Residual Deviance: 3517418035650.899
R^2: 0.973268898953974
Null degrees of freedom: 10039
Residual degrees of freedom: 10032
Null deviance: 1.3212865205156608e+18
Residual deviance: 3.5314877077935024e+16
AIC: 318553.3189175752

** Reported on cross-validation data. **
MSE: 4453587215909.751
RMSE: 2110352.391405225
MAE: 1547279.8532015597
RMSLE: 0.03528657985361276
Mean Residual Deviance: 4453587215909.751
R^2: 0.9659232204704427
```

```
Null degrees of freedom: 19375
Residual degrees of freedom: 19368
Null deviance: 2.5325041424212966e+18
Residual deviance: 8.629270589546734e+16
AIC: 619325.493958776

** Reported on test data. **
MSE: 4528498171569.509
RMSE: 2128026.8258575853
MAE: 1557603.0190972523
RMSLE: 0.03565571506703599
Mean Residual Deviance: 4528498171569.509
R^2: 0.9645925630870953
Null degrees of freedom: 4720
Residual degrees of freedom: 4713
Null deviance: 6.038496502126353e+17
Residual deviance: 2.1379039867979652e+16
AIC: 150992.22118511234
```

## Summary

Its excessive size (e+7) is the cause of the errors.
This project would benefit from having location, weather, and user power habits in addition to the power dataset.
XGBoost and LSTM are the models that fit this dataset the best.