

Instructions for Running the Day-Ahead Price Prediction Script

This script predicts the day-ahead electricity prices using historical and predicted data. It requires the following **5 data files**:

Data Files

1. **Actual Consumption 2015-2024-to dec 7 -11PM.xlsx**
 - Contains historical electricity consumption data from **January 1, 2015**, to **December 7, 2024**.
2. **Actual Generation 2015-2024-to dec 7 -11PM.xlsx**
 - Contains historical electricity generation data from **January 1, 2015**, to **December 7, 2024**.
3. **Weather_Data_2015_2024-to dec 7 -11PM.xlsx**
 - Contains historical weather data from **January 1, 2015**, to **December 7, 2024**.
4. **Price_2015_2024-to dec 7 -11__ADDED MISSING PRICES PM.xlsx**
 - Contains historical electricity prices from **January 1, 2015**, to **December 7, 2024**.
5. **Manual_Features.xlsx**
 - Contains **predicted data** for electricity generation, consumption, and weather for the day you want to predict prices.
 - This file must be manually prepared with **24 rows** (one for each hour of the day).
 - Data sources:
 - Electricity generation and consumption: [SMARD](#)
 - Weather data: [OpenWeatherMap](#)

N.B. We are working to process this data automatically from the script end so that we don't have to make a file and input the data manually.

Step-by-Step Instructions

Step 0: Install Libraries

Install the libraries which we have used in our code and they are:

Xgboost, sklearn, numpy, pandas and matplotlib. This can be done simply by running this line in cmd :

```
pip install xgboost scikit-learn matplotlib pandas numpy.
```

In order open xlsx files install openpyxl:

```
pip install openpyxl
```

Step 1: Update File Paths

Download and open the code file named “**electricity_prediction.ipynb**”. In **CELL 2**, update the file paths for the datasets to point to the location of the downloaded ZIP folder on your local storage.

Example:

```
data_dir = "folder path where dataset files are located"
```

Step 2: Preprocess and Merge Data

In **CELL 4**, the script will preprocess and merge all historical datasets into a single file. Specify the **output file name and location** where the merged data will be saved.

Example:

```
merged_file_path = 'file path/file_name.xlsx'
```

Step 3: Prepare Manual Features

In **CELL 9**, we need data for electricity generation, consumption, and weather for the day you want to predict prices. Currently, this file must be prepared manually with 24 rows (one for each hour of the day). ***We are working to process this data automatically from the script end so that we don't have to perform this step manually.***

After preparing the file please specify the following:

- The file path for the **Manual_Features.xlsx** file in the `manual_features_path` variable.

Example:

```
manual_features_path = '/home/Desktop/Manual_Features.xlsx'
```

- The **date for which you want to predict prices** in the `prediction_hours` variable.

Example:

```
# Define the prediction start date dynamically  
prediction_start_date = datetime(2025, 2, 18) # Change this  
date as needed
```

Step 4: Save Predictions

1. In **CELL 10**, specify the **output file name and location** where the predicted prices will be saved.

Example:

```
csv_filename = "/home/Desktop/DS/Predicted_Prices_Feb_13_2025.csv"
```

Step 5: Run the Script

1. Execute the script cell by cell, following the order of the steps above.
2. After running the script, the predicted day-ahead prices will be saved in the specified output file.