National College of Ireland

Configuration Manual

MSc Research Project
MSc Data Analytics

Jismol Thomas -
Student ID: 23414952

School of Computing
National College of Ireland

Supervisor: Harshani Nagahamulla

National College of Ireland
Project Submission Sheet
School of Computing

National College of Ireland

Student Name:	Jismol Thomas -
Student ID:	23414952
Programme:	MSc Data Analytics
Year:	2025
Module:	MSc Research Project
Supervisor:	Harshani Nagahamulla
Submission Due Date:	28/01/2026
Project Title:	Multi-Source Deep Learning for TSO-Level Electricity Demand Forecasting
Word Count:	727
Page Count:	9

I hereby certify that the information contained in this (my submission) is information pertaining to research I conducted for this project. All information other than my own contribution will be fully referenced and listed in the relevant bibliography section at the rear of the project.

ALL internet material must be referenced in the bibliography section. Students are required to use the Referencing Standard specified in the report template. To use other author's written or electronic work is illegal (plagiarism) and may result in disciplinary action.

Signature:	Jismol.
Date:	27th January 2026

PLEASE READ THE FOLLOWING INSTRUCTIONS AND CHECKLIST:

Attach a completed copy of this sheet to each project (including multiple copies).	☐
Attach a Moodle submission receipt of the online project submission, to each project (including multiple copies).	☐
You must ensure that you retain a HARD COPY of the project, both for your own reference and in case a project is lost or mislaid. It is not sufficient to keep a copy on computer.	☐

Assignments that are submitted to the Programme Coordinator office must be placed into the assignment box located outside the office.

Office Use Only	
Signature:	
Date:	
Penalty Applied (if applicable):	

Configuration Manual

Jismol Thomas -
23414952

1 Introduction

The purpose of this configuration manual is to give all the directions required to replicate the Multi-Source Deep Learning for TSO-Level Electricity Demand Forecasting research.

2 Hardware Requirements

This research was implemented using the following hardware configuration.

2.1 Laptop Specifications

Laptop specifications (Device specifications & Windows Specifications) used in this research are described in Table 1. Figure 1 and Figure 2 to show the same.

Component	Specification
Processor	AMD Ryzen 5 7520U (2.80 GHz)
RAM	16 GB installed (13.7 GB usable)
System Type	64-bit Operating System, x64-based processor

Table 1: Laptop specifications

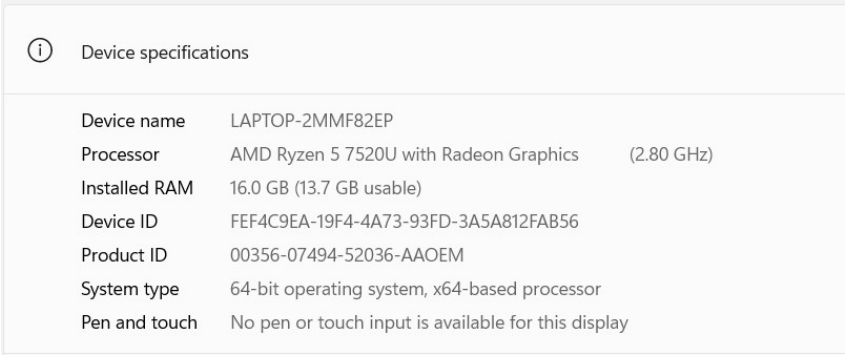

Figure 1: Device Specifications

Figure 2: Windows Specifications

3 Software Requirements

3.1 Jupyter Notebook Configuration

The stages like EDA, preprocessing, merging datasets, and creating the unified feature datasets were carried out using Jupyter Notebook installed on the local machine. The device specifications are shown in Table 2. Details of Jupyter are provided in Figures 3

Component	Configuration
Python Version	Python 3.10.18 (Conda-Forge build)
Platform	Windows 10 (Build 26100), 64-bit (AMD64)
Environment	Jupyter Notebook (local machine)
Usage	Preprocessing, EDA, dataset construction

Table 2: Configuration details of Jupyter Notebook

Figure 3: Python and platform information displayed in Jupyter Notebook

3.1.1 Libraries used in Jupyter Notebook

Libraries used include pandas, numpy, matplotlib, seaborn, plotly, scipy, statsmodels,requests, urllib, json, os, sys, glob, pathlib, datetime, joblib, geopandas, shapely, rapidfuzz, unidecode, sklearn, tqdm, gc.

The Jupyter Libraries with versions are shown in Figure 4

3.2 Google Colab Configuration

Google Colaboratory is used for the development of all models, hyperparameter tuning, training, and evaluation, because of the computational complexity of deep learning models. The device configuration provided of Colab is summarised in Figures 5 and 6.

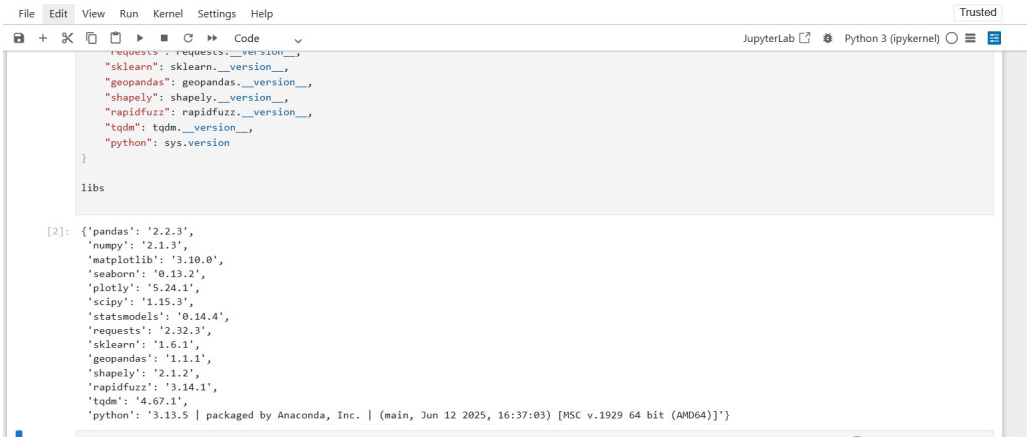

Figure 4: Library versions used in Jupyter Notebook

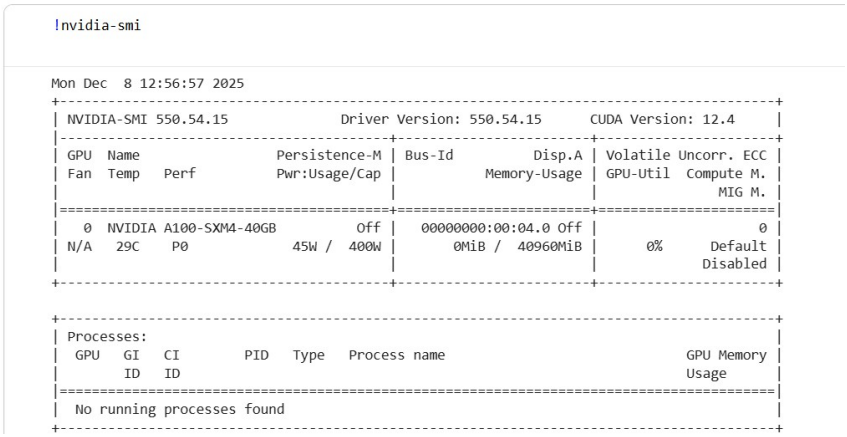

Figure 5: Google colab pro+ configuration

Figure 6: Python version in Google colab pro+

4 Dataset

Multiple sources of data are used in this research to construct a unified dataset for TSO-level electricity demand forecasting. The datasets include:

- TSO electricity consumption, generation and day ahead price data from SMARD.
- Open-meteo API weather data on parameters like temperature, relative humidity, wind speed, shortwave radiation, precipitation, surface pressure.
- Socio-economic indicators such as GDP, Population density from Eurostat.
- Mapping of districts to control zones of German Transmission System Operators (TSOs) dataset from Zenedo.

All datasets were downloaded separately and processed using Jupyter Notebook. The folder structure of the raw datasets is shown in Table 3. Screenshots of Jupyter details are provided in Figures 7 `Electricity_Demand/data`. Below are the files and folders inside the data folder.

File / Folder	**Description**
`weather_raw_tso`	Raw Open-Meteo weather data
`SMARD_DATA`	Consumption, generation, and day ahead price data in seperate folders
`demo_r_d3dens_page_linear_2_0.csv`	Eurostat population density dataset for NUTS3 regions
`nama_10r_3gdp_custom_18511051_linear_2_0.csv`	Eurostat GDP dataset for NUTS3 regions
`GER_NUTS3_TSOs.*`	NUTS3-to-TSO spatial mapping shapefiles

Table 3: Description of dataset files included in the project

After data cleaning and transformation, all processed datasets and final datasets are stored inside processed folder in data folder.

All code files are available in Electricity_Demand/Code folder as shown in Figure 8

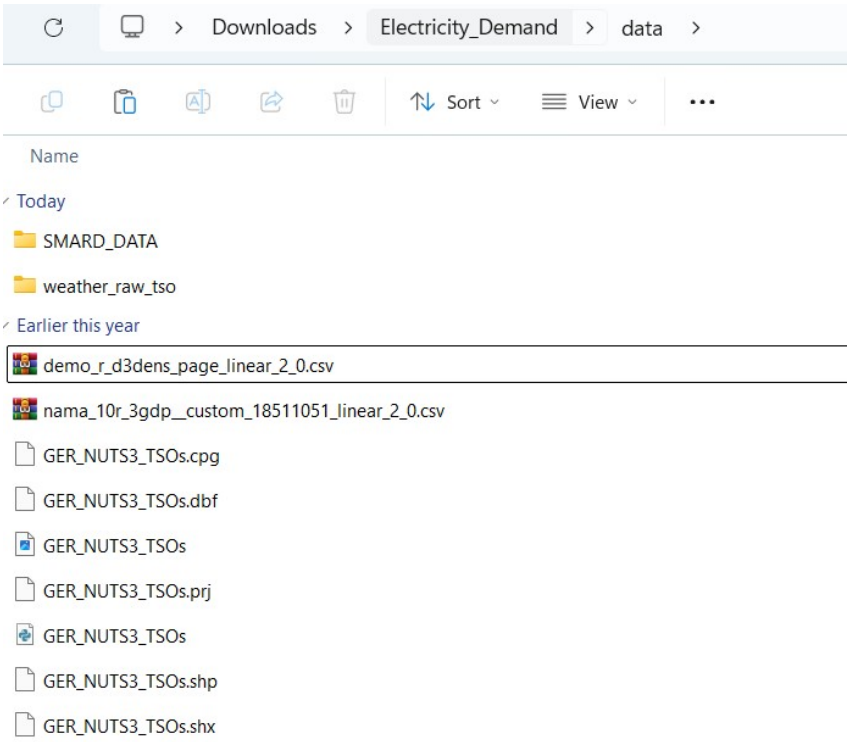

Figure 7: Structure of Data Folder

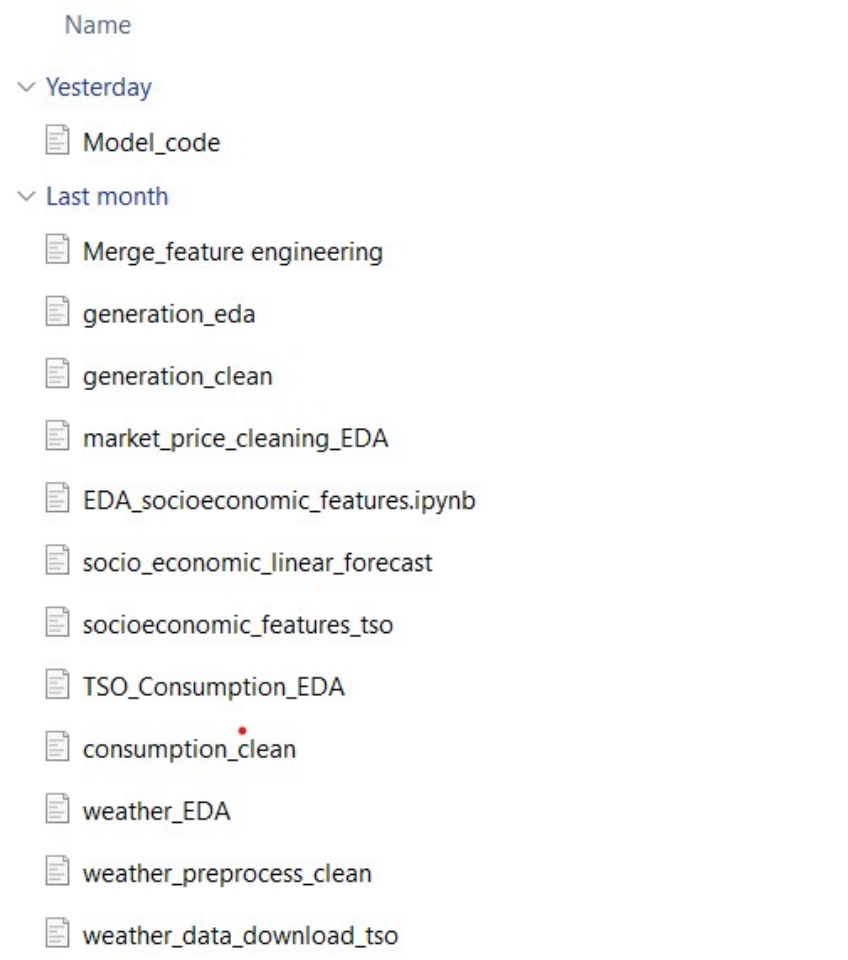

Figure 8: Structure of Code Folder

Using the notebook `Merge_feature_engineering.ipynb`, all processed datasets were merged into a final unified dataset:

This merged dataset is then uploaded to Google Drive to use in Google Colaboratory for modeling purpose.

4.1 Required Libraries

The following libraries must be installed in Colab is shown in Figures 11 and Figures 12

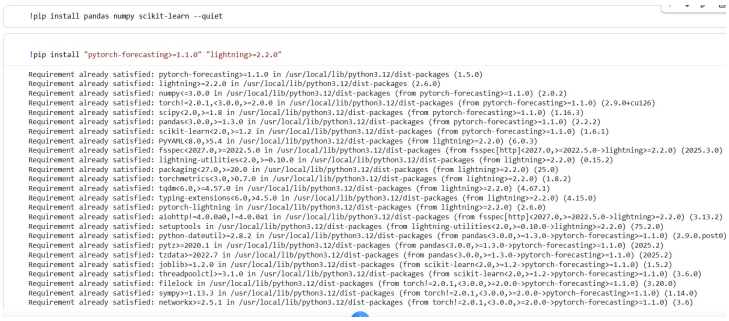

Figure 9: Pytorch package in Colab

Figure 10: Optuna package installed in Colab

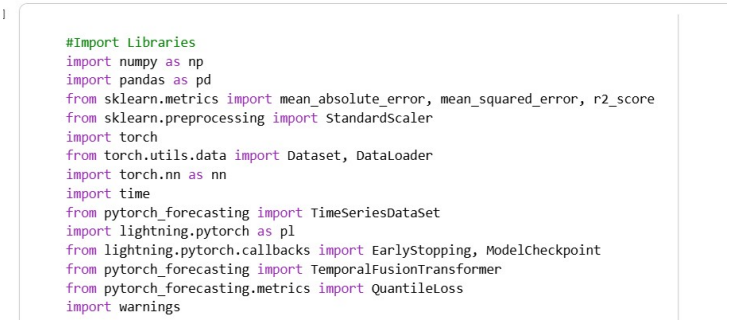

Figure 11: Libraries needed in Colab

4.2 Colab GPU Configuration

Enable GPU before training:

`Runtime → Change runtime type → Hardware accelerator → GPU`

```python
import optuna
from google.colab import drive
import matplotlib.pyplot as plt
import seaborn as sns

plt.style.use("default")
sns.set(style="whitegrid")
```

Figure 12: Libraries needed in Colab

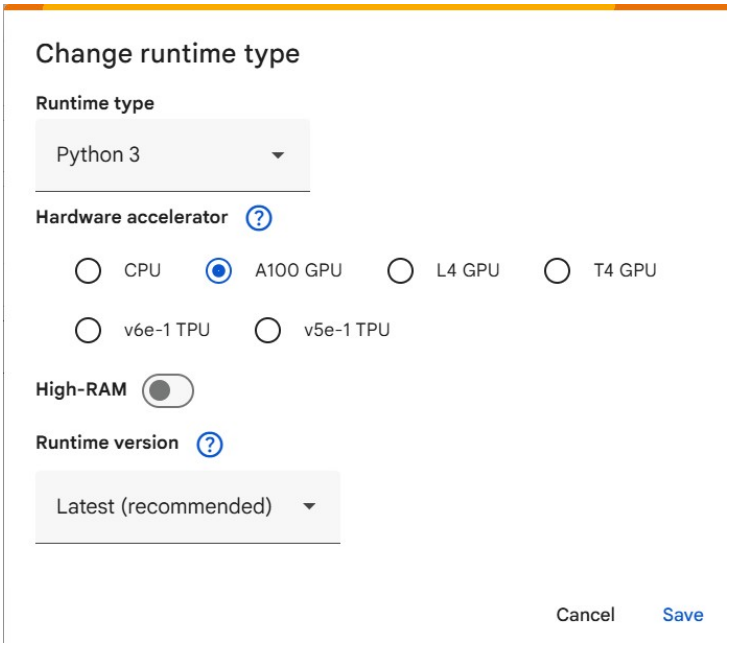

Figure 13: Change Runtime to A100

4.3 Mount Google Drive

The Dataset needs to be uploaded in Google drive. Then mount the google drive with Google colab as shown in Figure 14.

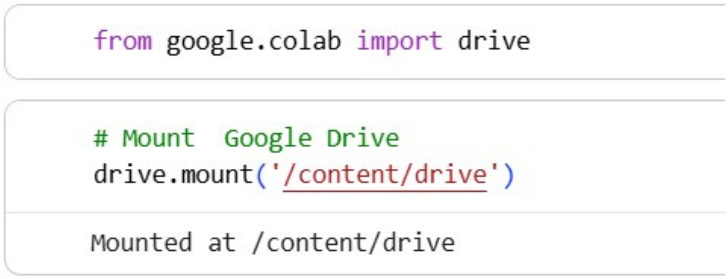

Figure 14: Mount Colab

4.4 Loading the Dataset in Colab

The dataset is loaded in Google Colab as shown in Figure 15.

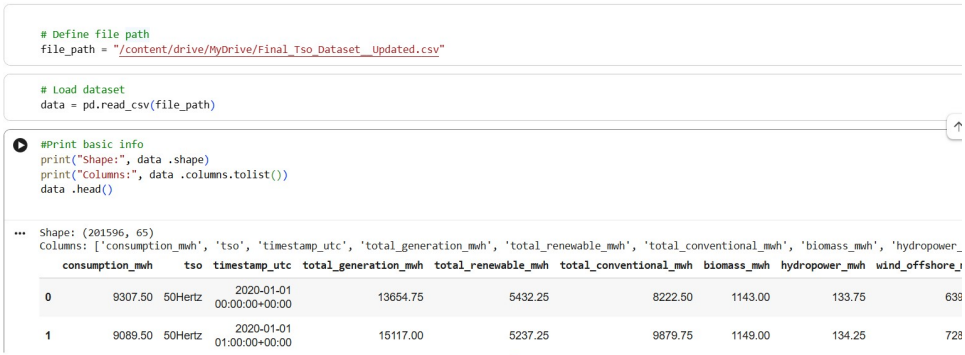

Figure 15: File Loaded in Colab

5 Execution Steps

1. Unzip the ICT Solution Artefact-23414952 Folder and extract the `Electricity_Demand` folder.
2. Run all notebooks **except** `Model_code.ipynb` in the `code` folder (inside the `Electricity_Demand` directory) directly in Jupyter Notebook by clicking *Restart Kernel and Run All*.
3. Run the files in the following order:
   (a) `weather_data_download_tso.ipynb`
   (b) `consumption_clean.ipynb`
   (c) `generation_clean.ipynb`
   (d) `weather_preprocess_clean.ipynb`
   (e) `socioeconomic_features_tso.ipynb`
   (f) `socio_economic_linear_forecast.ipynb`

(g) `TSO_Consumption_EDA.ipynb`

(h) `generation_eda.ipynb`

(i) `weather_EDA.ipynb`

(j) `market_price_cleaning_EDA.ipynb`

(k) `EDA_socioeconomic_features.ipynb`

(l) `Merge_feature_engineering.ipynb`

(m) `Model_code.ipynb` (Google Colab Pro+)

4. Execute `Model_code.ipynb` in Google Colab Pro+. Prior to the execution of the code, update dataset path to match your Google Drive directory.

5. Ensure GPU runtime is enabled to avoid slow training.

6. Click *Run all*.

After this, open the research_dashboard.pbix in Power BI. Then upload the combined result CSV file generated from Collab, then click Home → Refresh. Figure 16 is the Power BI Desktop Version used in this research.

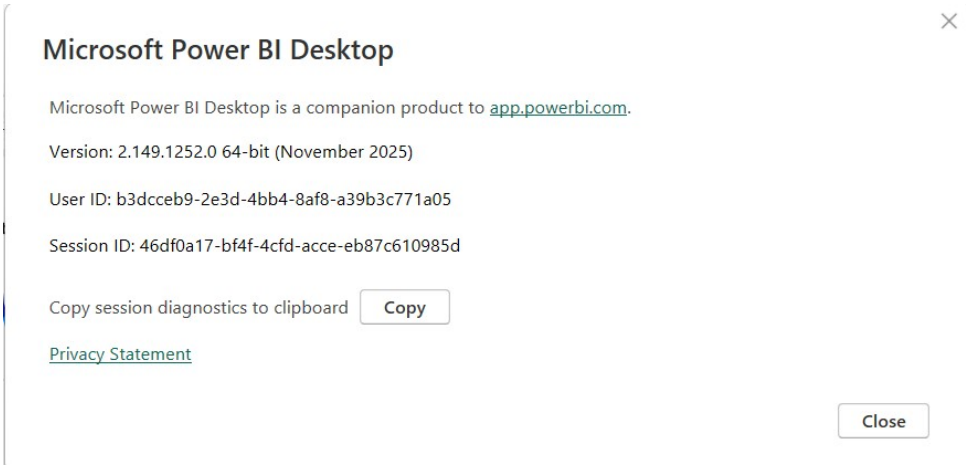

Figure 16: PowerBI Version

These steps will make sure that the entire process of working with raw data until the final model prediction can be reproduced successfully.

References

SMARD Electricity Data: `https://www.smard.de/en`

Open -meteo: `https://open-meteo.com/`

Eurostat Socio-Economic Indicators: `https://ec.europa.eu/eurostat`

Frysztacki, M.M. (2023) "Mapping of districts to control zones of German Transmission System Operators (TSOs)". Zenodo. doi:10.5281/zenodo.7530196.