

Configuration Manual

MSc Research Project
MSc in Data Analytics

Serap BATIR
Student ID: x22205179

School of Computing
National College of Ireland

Supervisor: Furqan Rustam

**National College of Ireland
Project Submission Sheet
School of Computing**



Student Name:	Serap BATIR
Student ID:	X22205179
Programme:	Msc in Data Analytics
Year:	2025
Module:	MSc Research Project
Supervisor:	Furqan Rustam
Submission Due Date:	15/09/2025
Project Title:	Configuration Manual
Word Count:	497
Page Count:	3

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:	
Date:	15th September 2025

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

Serap BATIR
X22205179

1 Required Software and Tools

The experimental setup for this project was implemented using the following software and tools:

- **Python 3.10**
- **Anaconda Jupyter Lab**
- **Deep Learning Frameworks:**
- **TensorFlow 2.15** – LSTM, GRU, LSTM(a few shot) implementations..
- **Machine Learning Libraries:**
- **scikit-learn 1.4** – Random Forest, SVM, KNN, XGBoost.
- **xgboost 2.0** – Gradient boosting model implementation.
- **Data Processing & Visualization:**
- **pandas 2.2, numpy 1.26** – Data manipulation.
- **matplotlib 3.8, seaborn 0.13** – Visualization.

2 Dataset and File Paths

Due to space limitations in the submission link (500 MB), the dataset file was removed from the ZIP file. It must be downloaded manually from the above link. All notebooks, scripts, and data files should be placed in the **same working directory**. The code uses **relative paths** assuming that the dataset file is located in the same folder as the Jupyter notebooks.

- **Dataset:** *ElectricityLoadDiagrams2011–2014* (UCI Machine Learning Repository)
- **Download link:** <https://archive.ics.uci.edu/dataset/321/electricityloadaddiagrams20112014>

3 Execution Order

To reproduce the experiments, run the Jupyter notebooks in the following sequence:

1. **EDA.ipynb** – Perform exploratory data analysis on the original dataset.
2. **modelprocessing.ipynb** – Prepare and preprocess the dataset for training, including scenario generation (theftgenerator.py) and sliding window creation (windowslide.py).
3. **GRU.ipynb / LSTM-V1.ipynb / LSTM-fewshot.ipynb** – Train deep learning models.
4. **MLmodels.ipynb** – Train and evaluate machine learning models (Random Forest, XGBoost, SVM, KNN).
5. **evaluation.ipynb** – Perform final evaluation and generate performance comparison results.
6. **Save outputs** – The final results will be saved in modelresults.csv and allmodelsoutputs.npz.