National College of Ireland

Configuration Manual

MSc Research Project
Data Analytics

Neethu Elizhabeth Sebastian
Student ID: x23179996

School of Computing
National College of Ireland

Supervisor: Noel Cosgrave

National College of Ireland
Project Submission Sheet
School of Computing

National College of Ireland

Student Name:	Neethu Elizhabeth Sebastian
Student ID:	x23179996
Programme:	Data Analytics
Year:	2024
Module:	MSc Research Project
Supervisor:	Noel Cosgrave
Submission Due Date:	29/1/2025
Project Title:	Configuration Manual
Word Count:	644
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:	NEETHU ELIZHABETH SEBASTIAN
Date:	28th January 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

Neethu Elizhabeth Sebastian
x23179996

1 Introduction

This work comes as a configuration manual of the Electricity Demand Forecasting System using machine learning models such as ARIMA, LSTM, Regression, and OSI model for data processing. This setup is aimed to predict the future electricity demand of the UK by the historical data obtained from the National Grid ESO (2009–2024).

2 System Requirements

To run the Electricity Demand Forecasting System, the following hardware and software configurations are required:

2.1 Hardware

- **CPU**: Minimum of 4 cores, 2.5 GHz or higher.
- **RAM**: 16 GB or more for handling large datasets.
- **GPU** (for LSTM model training): NVIDIA GPU with CUDA support for accelerated computation (e.g., Tesla or RTX series).
- **Storage**: 100 GB or more of free space for storing the dataset and model outputs.

2.2 Software

- Operating System: Windows 10/Ubuntu 20.04 or later.
- Version 3.8 or higher.
- Required Libraries:

Library	Function
Pandas	Data manipulation
Numpy	Numerical computations
Matplotlib, Seaborn	Data visualization
Scikit-learn	Machine learning models and data preprocessing
Keras, TensorFlow	Building and training the LSTM model
Fbprophet	Forecasting

3 Data Preprocessing Configuration

3.1 Data Import

- Import the dataset provided by National Grid ESO, which includes hourly or half-hourly electricity demand and generation data.
- Ensure that the dataset includes variables such as National Demand (ND), Transmission System Demand (TSD), and Embedded Wind Generation. These features are essential for modeling and forecasting.

3.2 Missing Data Handling and Feature Engineering

- Check for missing values in the dataset. You can impute missing values using interpolation or backfilling depending on the nature of the data.
- Generate additional features like lag features, which are critical for time-series forecasting models:
- Add seasonal features such as hour of the day or day of the week to capture daily and weekly patterns

4 Model Configuration

LSTM Configuration : The LSTM model is suitable for capturing long-term dependencies and non-linear patterns in the data
Regression Model Configuration: Use linear or polynomial regression for baseline modeling. This is helpful for comparing with more complex models.

5 Integration with OSI Model

The OSI model gives an architectural design to dissect the stream and procedures of the information. Below is the mapping of the OSI layers to different steps in the forecasting process:
OSI Layer Mapping

- Layer 7 (Application Layer): Tasks that fall under sample acquisition and preparation such as cleaning, transformation as well as treatments of missing values.
- Layer 6 (Presentation Layer): Feature preprocessing, for example feature scaling including normalization in this specific case of the program, in order to make all the inputs to be used for training to be in the same range of values.
- Layer 5 (Session Layer): Training of models, check that data is split into train, validation and test data, check that data flows in batch.
- Layer 4 (Transport Layer): Cleverly splits the data to pass from one system to another (for example, for cross-validation or parallel computing).

- Layer 3 (Network Layer): Large data set or large models distributed over two or more machines, if working in a cloud environment or distributed system.
- Layer 2 (Data Link Layer): Preserves the data within the communication process, so that data will not be altered when being transferred from one point to another especially when using distributed or cloud based systems.
- Layer 1 (Physical Layer): Relates to the physical hardware on which the dataset and corresponding models have been stored as well as the physical infrastructure needed for model implementation.

6 Conclusion

This configuration manual outlines how to configure an electricity demand forecasting system with the use of LSTM and regression models. Organization and integration of OSI model makes the forecasting process more organized and relevant within the data process flow stages facilitating proper processing of data. With the help of outlined steps users correctly predict the intensity of electricity demand and can effectively manage consumption for strategic purposes.