

Configuration Manual

MSc Research Project
MSc Data Analytics

Rishabh Deepak Saudagar
Student ID: X23288914

School of Computing
National College of Ireland

Supervisor: Sallar Khan

National College of Ireland
MSc Project Submission Sheet
School of Computing



Student Name: Rishabh Deepak Saudagar
Student ID: X23288914
Programme: Msc Data Analytics **Year:** 2024-2025
Module: Msc Research Practicum
Lecturer: Sallar Khan
Submission Due Date: 11/08/2025
Project Title: Enhancing CNN-LSTM Models for Financial Time-Series Forecasting through Hyperparameter Optimization and Regularization
Word Count: 1731 **Page Count:** 10

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: Rishabh Deepak Saudagar

Date: 11/08/2025

PLEASE READ THE FOLLOWING INSTRUCTIONS AND CHECKLIST

| | |
|---|--------------------------|
| Attach a completed copy of this sheet to each project (including multiple copies) | <input type="checkbox"/> |
| Attach a Moodle submission receipt of the online project submission, to each project (including multiple copies). | <input type="checkbox"/> |
| 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. | <input type="checkbox"/> |

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

Rishabh Deepak Saudagar
X23288914

1. Introduction

The Configuration Manual provides multidimensional and step-by-step guidelines on how to set up and implement the MSc research project, CNN-LSTM Financial Time-Series Forecasting. It is the aim of the research to increase the stability and accuracy of CNN-LSTM hybrid models that are used to predict the prices of stocks across multiple foreign markets (India, Ireland, USA) through rigorous data processing, hyperdrive tuning, and overfitting avoidance.

2. System Requirements

Below are the minimum system requirements to the project successfully.

- **Operating System:** Windows 10/11, macOS, or Ubuntu Linux
- **Processor:** Intel i5/Ryzen 5 or higher (recommended)
- **RAM:** 8 GB minimum (16 GB recommended)
- **Disk Space:** At least 5 GB free
- **Python Version:** 3.8 or higher

3. Project Structure

The repository organization will be presented in a logical way to offer simplicity, narrowing down, and usability in the development and evaluation procedures. Any data will be stored in the data/ folder, which is categorized into subfolders of raw, cleaned, and file result of prediction. Jupyter notebooks in the notebooks/ directory allow the fundamental implementation and analysis in a sequential fashion, exploration, modeling, and assessment. There are supporting files such as requirements.txt that provides a detailed list of all requirements to reproduce, and also README.md gives a brief description of the project, as well as a quick overview of how it can be set up. Using this framework, the user can seamlessly move between data, code and results and this ensures full transparency in the workflow that can be replicated easily.

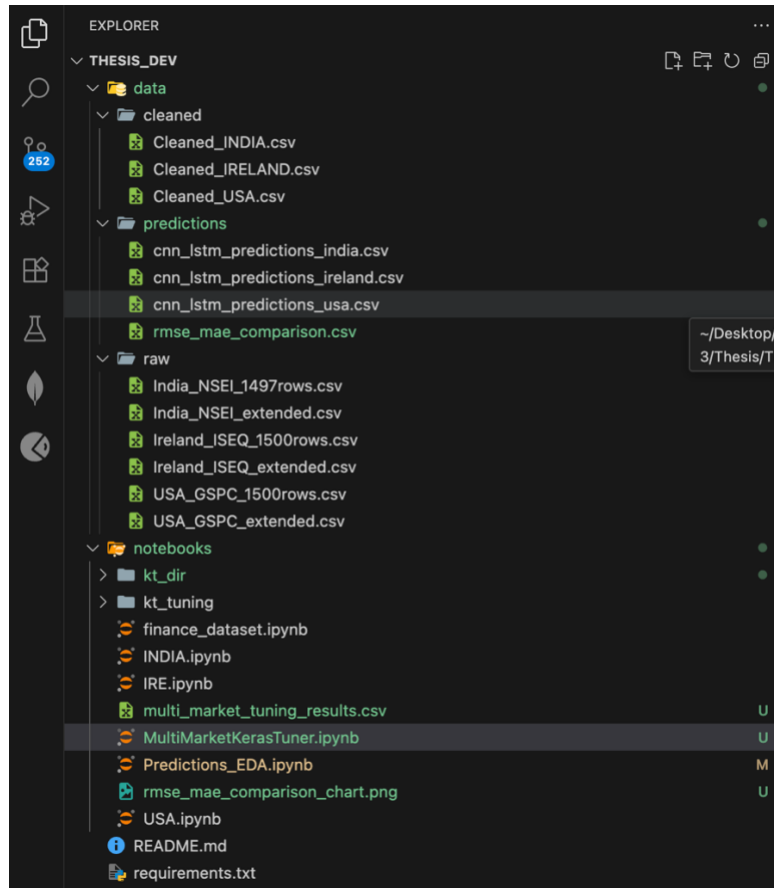


Figure 1: Project Directory Structure

4. Environment Setup

So there will be no conflict between dependencies, it is recommended to create a clean Python environment before installing project prerequisites. You can do this via built-in venv module or the Anaconda distribution, whichever you prefer.

4.1 Create a Virtual Environment

- **Using venv:**
`python -m venv cnn_lstm_env`
- **Activate the environments:**
 - **On Windows:**
`cnn_lstm_env\Scripts\activate`
 - **On macOS/Linux:**
`source cnn_lstm_env/bin/activate`

- **Using Anaconda:**
conda create -n cnn_lstm_env python=3.8
conda activate cnn_lstm_env

4.2 Install Project Dependencies

Once the environment is active, install all required packages using the provided “*requirements.txt*” file:

```
pip install -r requirements.txt
```

This will automatically install all required libraries ones- TensorFlow, Keras, Pandas, NumPy and others- in their proper versions needed to process the project. Such an environment will allow perfect compatibility with the project as well as permit reproducibility of results.

5. Key Python Libraries

Now, when you have your environment, all the dependencies installed, you can begin working with the project code and notebooks. It aims to be easy-to-use to enable you to explore the data, construct and test models and see the results interactively.

5.1 Launch Jupyter Notebook

To start the Jupyter Notebook in the activated environment run:

```
jupyter notebook
```

This command will launch Jupyter inside your default web browser where it is possible to navigate into the notebooks/ directory and open the notebook files.

5.2 Execute Notebooks Step-by-Step

Choose the notebook that you want to execute (e.g., *model_training.ipynb*), then execute the code cells one after another. The notebooks are kept structured so as to lead you through:

- Data exploration and preprocessing
- Model building and hyperparameter tuning
- Generating predictions and visualizations



Figure 2: Jupyter Notebook interface open in web browser

5.3 Running Python Script

If the project includes Python scripts (e.g., for automated or batch processing), you can run them from the command line as follows:

```
python src/[script_name].py
```

All outputs—including prediction files and generated plots—will be saved in the appropriate output directories, such as “data/predictions/” or “outputs/”.

6. Running the Project

This project involves easy installation and does not require too much tinkering of most users. All of the critical file paths and settings are pre-set in the notebooks and scripts so that they follow the recommended project structure.

- Data Paths:**

All files containing the dataset and the output belong in data/, and there should also be subdirectories inside data/, namely raw/ cleaned/ and predictions/. In case you desire to redocument your files, then feel free to reconfigure the file paths within the relevant notebooks or scripts thereof.
- Output Locations:**

Model predictions, processed datasets, and visualizations are saved to “data/predictions/” or “outputs/”, as indicated in each notebook or script.
- API Keys and Environment Variables:**

Regular usage does not require API keys and environment variable. The yfinance will be used to acquire all data used in this project; it does not require authentication. In case you would wish to add some external sources of data in future, do give relevant configuration guidelines here.

- **Custom Settings:**

If you would want to change model parameters, data range or other hyperparameters you can change it directly in the notebook cells prior to executing the code. All the adjustable parameters are marked on the frontal side with marks by which they are easily identified.

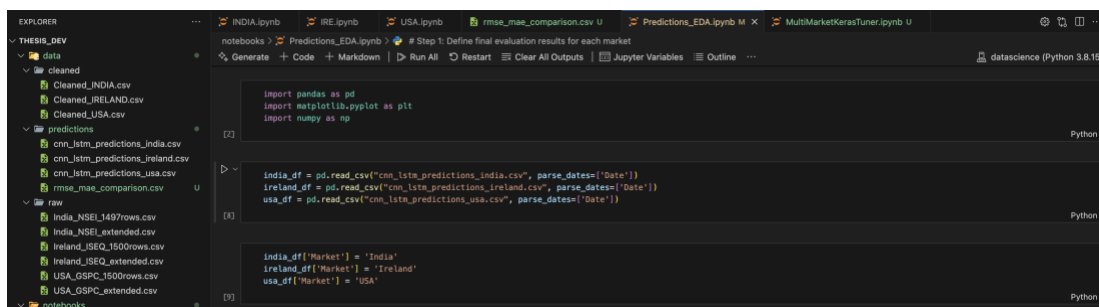


Figure 3: Example notebook cell showing configurable file paths or parameters

7. Configuration Details

The project is designed to be deployed without much effort and, as such, can be configured by most users, with minimal effort. The notebooks and scripts have all necessary file locations and settings that will match the stipulated project structure.

- **Data Paths:**

Any file with the data and output files should be located within the data/ directory and shall have a sub-directory within the directory entitled clean/. In case you choose to restructure your files, you should make the appropriate modifications on the file paths in the relevant notebooks or scripts.

- **Output Locations:**

The processed data, outputs and plots are stored in data/predictions/ or outputs/, depending on the notebook or script.

- **API Keys & Environment Variables:**

There are no API keys required or environment variables needed to use normally. The information used in this project is received with the help of yfinance which is not required to enter authentication. In the event that you will include some further data sources, please add relevant configuration guidelines at the bottom.

- **Custom Settings:**

To tune model parameters, data range, or other hyper-parameters you can do it directly in cells prior to execution in the notebook. Any variable that is possible to adjust has a prominent label to allow easy identification.

8. Outputs

Upon the successful execution of the notebooks or scripts, the project will produce a range of outputs vital for analysis and reporting.

- **Prediction Files:**

Each market (e.g. India, Ireland, USA) has their "predicted" stock prices stored as CSV files in the data/predictions/ directory. The files are named after their market or a model, e.g. cnn_lstm_predictions_india.csv.

```
data > predictions > cnn_lstm_predictions_india.csv
1 Date,Actual_Close,Predicted_Close,Market
2 2024-05-02,22475.849609375,22719.648,India
3 2024-05-03,22442.69921875,22728.562,India
4 2024-05-06,22302.5,22707.32,India
5 2024-05-07,22302.5,22708.535,India
6 2024-05-08,21957.5,22725.361,India
7 2024-05-09,22055.19921875,22673.59,India
8 2024-05-10,22104.05078125,22686.494,India
9 2024-05-13,22217.849609375,22642.09,India
10 2024-05-14,22200.55078125,22592.188,India
11 2024-05-15,22403.849609375,22575.564,India
12 2024-05-16,22466.099609375,22564.156,India
13 2024-05-17,22529.05078125,22537.615,India
14 2024-05-21,22597.80078125,22568.62,India
15 2024-05-22,22967.650390624996,22562.787,India
16 2024-05-23,22957.099609375,22602.775,India
17 2024-05-24,22932.44921875,22654.67,India
18 2024-05-27,22888.150390625,22714.414,India
19 2024-05-28,22704.69921875,22787.066,India
20 2024-05-29,22488.650390625004,22825.957,India
```

Figure 4: Prediction CSV File sample

- **Visualizations:**

All major analysis and model evaluation notebooks produce graphs and charts, including:

- **Time series plots of actual vs. predicted prices**
- **Loss curves and model performance metrics**
- **Feature importance or attention maps (if included)**

These figures can be exported as image files for use in your thesis or Power BI dashboards.

- **Processed Datasets:**

The converted and cleaned data used as a model is already saved in the “data/cleaned/” folder, which helps to track the every step of the data pipeline.

9. Troubleshooting and Tips

This section offers resolutions to prevalent problems that may occur during the setup or execution of the project. If you have issues not mentioned here, verify that all directions in the preceding sections have been meticulously adhered to.

- **Dependency Errors:**

When failing to encounter any of the mentioned errors (ModuleNotFoundError or ImportError), be sure that your virtual environment is activated and all required dependencies are installed via:

```
pip install -r requirements.txt
```

If issues persist, check that your Python version matches the one specified in the system requirements.

- **Data Loading Issues:**

In a case where the notebooks did not find a data file, do make sure that all the required CSV files are named and placed in the right folders under data/ subdirectories. Adapt any pathnames in the notebook to your directory structure in case it is not the same as in the template.

- **Jupyter Notebook Not Launching:**

Ensure Jupyter is installed in your current environment:

```
pip install jupyter
```

If the notebook server does not open in your browser, try copying the provided local URL into your browser manually.

- **CUDA/GPU Errors:**

If running on a GPU and encountering TensorFlow errors, confirm that the correct CUDA and CuDNN versions are installed as per [TensorFlow's official installation guide](#).

- **Permission Issues:**

On some systems, you may need to run your terminal or IDE as an administrator, especially when accessing or saving files.

10. References and Further Reading

- TensorFlow Installation Guide: <https://www.tensorflow.org/install>
- Keras Tuner Documentation: https://keras.io/keras_tuner/
- yfinance Python Package: <https://github.com/ranaroussi/yfinance>
- Official Python Documentation: <https://docs.python.org/3/>
- Project README for workflow and quick start: <https://github.com/WhiteHorse101/MSc-Research-Practicum-Rishabh>

11. Appendix

Requisite Python libraries and their versions are exhaustively listed in the requirements.txt file located in the root directory of the project.

To install all dependencies by running the following command in a working environment:

```
pip install -r requirements.txt
```

```
requirements.txt
1  absl-py==2.2.2
2  ace_tools==0.0
3  aiobotocore @ file:///private/var/folders/k1/30mswbxs7r1g6zwn8y4fyt500000gp/T/abs_1cl06d5vjc/croot/aiobotocore_1714464399334/work
4  aiohappyeyeballs @ file:///private/var/folders/nz/j6p8yfhx1mv_0grj5x14650h0000gp/T/abs_24gdb3bp5u/croot/aiohappyeyeballs_1732662351219/work
5  aiohttp @ file:///private/var/folders/nz/j6p8yfhx1mv_0grj5x14650h0000gp/T/abs_4dbk4qcpk1/croot/aiohttp_1725527761629/work
6  aioitertools @ file:///tmp/build/80754af9/aioitertools_1607109665762/work
7  aiosignal @ file:///tmp/build/80754af9/aiosignal_1637843061372/work
8 aiosqlite==0.21.0
9  alabaster @ file:///private/var/folders/nz/j6p8yfhx1mv_0grj5x14650h0000gp/T/abs_39uesgct45/croot/alabaster_1718201495024/work
10 alembic==1.15.1
11 altair @ file:///private/var/folders/nz/j6p8yfhx1mv_0grj5x14650h0000gp/T/abs_a8x4081_4h/croot/altair_1687526044471/work
12 anaconda-anon-usage @ file:///private/var/folders/nz/j6p8yfhx1mv_0grj5x14650h0000gp/T/abs_60q98_n2ty/croot/anaconda-anon-usage_1732732446690/work
13 anaconda-catalogs @ file:///private/var/folders/nz/j6p8yfhx1mv_0grj5x14650h0000gp/T/abs_e8tmw882qa/croot/anaconda-catalogs_1685727305051/work
14 anaconda-cli-base @ file:///private/var/folders/nz/j6p8yfhx1mv_0grj5x14650h0000gp/T/abs_4950slpzfw/croot/anaconda-cli-base_1731510911315/work
15 anaconda-client @ file:///private/var/folders/k1/30mswbxs7r1g6zwn8y4fyt500000gp/T/abs_82twczuten/croot/anaconda-client_1731432687858/work
16 anaconda-cloud-auth @ file:///private/var/folders/nz/j6p8yfhx1mv_0grj5x14650h0000gp/T/abs_66qoffz18/croot/anaconda-cloud-auth_1731432562188/work
17 anaconda-navigator @ file:///private/var/folders/k1/30mswbxs7r1g6zwn8y4fyt500000gp/T/abs_cac7cu5xww/croot/anaconda-navigator_1733149480396/work
18 anaconda-project @ file:///Users/cbousseau/work/recipes/ci_py311/anaconda-project_1677964558977/work
19 annotated-types @ file:///private/var/folders/nz/j6p8yfhx1mv_0grj5x14650h0000gp/T/abs_1fa2djiwhb/croot/annotated-types_1709542925772/work
20 anyio @ file:///private/var/folders/k1/30mswbxs7r1g6zwn8y4fyt500000gp/T/abs_68kdsx8iyd/croot/anyio_1729121281958/work
21 apache-airflow==2.10.5
```

Figure 5: Example Snippet of requirements.txt