National College of Ireland

Configuration Manual

MSc Research Project
FinTech

Won Il Kang
Student ID: x20174675

School of Computing
National College of Ireland

Supervisor: Victor Del Rosal

National College of Ireland

MSc Project Submission Sheet

School of Computing

Student Name: ……Won Il Kang……

Student ID: ……x20174675……..……

Programme: ……MSc FinTech…… **Year:** ……1……..

Module: ……MSc Research Project……..……

Lecturer: ……Victor Del Rosal……..……

Submission Due Date: ……15.08.2022…… ……..……

Project Title: ……Predicting Asian Stock Market Index Using U.S. Financial Market Indexes and Machine Learning Techniques ……

Word Count: ……4…… **Page Count:** ……980……..……..……

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: ……Won Il Kang……

Date: ……15.08.2022……

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

Won Il Kang
Student ID: x20174675

1 Environment

In this project, data was analyzed using Python, and the program used was PyCharm. Python version 3.9 was used. The development environment is local. The installation method is as follows. Open a project with Python version 3.9 installed. Then, add the submitted program Python files to the project. The submitted files are shown in the table below. Next, install the libraries listed in section 2. After that, run the main file as in section 3 to run the program. Alternatively, you can install it through the github address below.

fileIO.py; load_result.py; loadData.py; Loger.py; LR.py; LSTM.py; main.py; Math.py; NN.py; readCSV.py; RF.py; RNN.py ; SVM.py
https://github.com/Wonil-kang/ ResearchProject_

2 Libraries

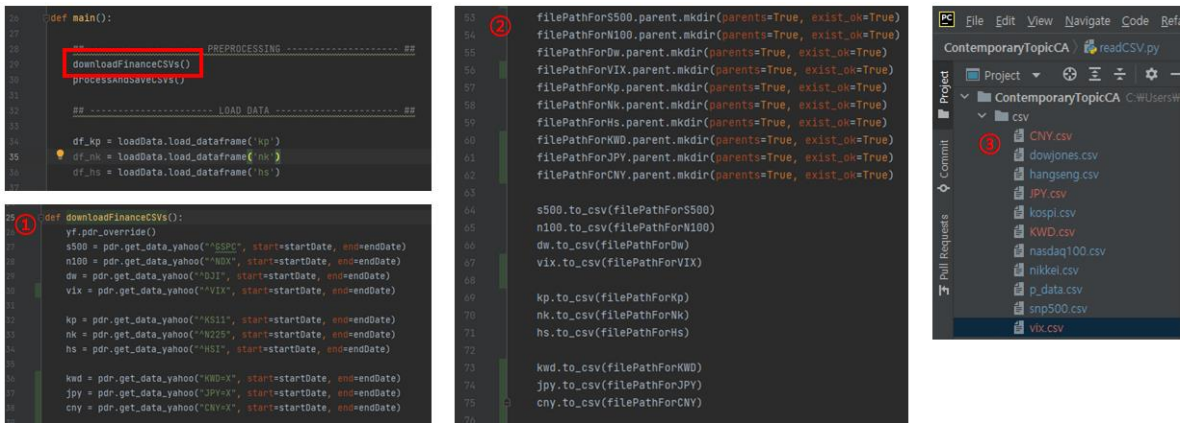
In this project, data were analyzed using 17 libraries. Financial data provided by Yahoo Finance was extracted using the yfinance library. In addition, data processing libraries and machine learning algorithms required for various data analysis were utilized as follows. If the library below is not present, the program may not work.

Libraries	Purpose
`import tensorflow as tf`	Related Libraries for Neural Network
`from keras.callbacks import EarlyStopping`	
`from keras.layers import Dense`	
`from keras.models import Sequential`	
`from tensorflow import metrics`	
`from tensorflow.python.keras.callbacks import EarlyStopping`	
`from tensorflow.python.keras.layers import LSTM`	long short-term memory (LSTM)
`from tensorflow.python.keras.layers import RNN, SimpleRNN`	Recurrent Neural Network
`from sklearn.linear_model import LinearRegression`	Linear Regression
`from sklearn.ensemble import RandomForestRegressor`	Random Forest
`from sklearn import svm`	Suport Vector Machine
`import yfinance as yf`	Yahoo Finance Data Extraction
`from pandas_datareader import data as pdr`	Data Processing
`import pandas as pd`	
`import numpy as np`	
`import matplotlib.pyplot as plt`	
`from sklearn.preprocessing import MinMaxScaler, StandardScaler`	

3 Program Operation Process

The program operation phase is divided into five stages: data, loading data pre-processing, data separation by market, algorithm execution, and data analysis. The details of operation in each detailed step are dealt with in the following detailed section.

3.1 Data Download from Yahoo Finance

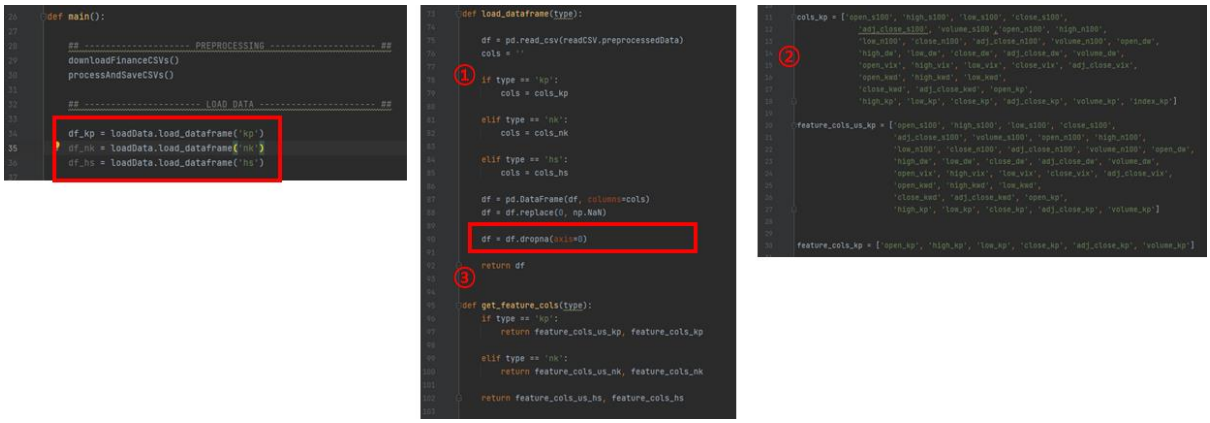


The first step is to extract data from Yahoo Finance. Data from January 1, 2001 to May 31, 2022 were extracted using the Yahoo Finance API. This content is introduced in ①. Next, each extracted data was saved as a CSV file. These steps are shown in ② and ③.

3.2 Data Pre-processing

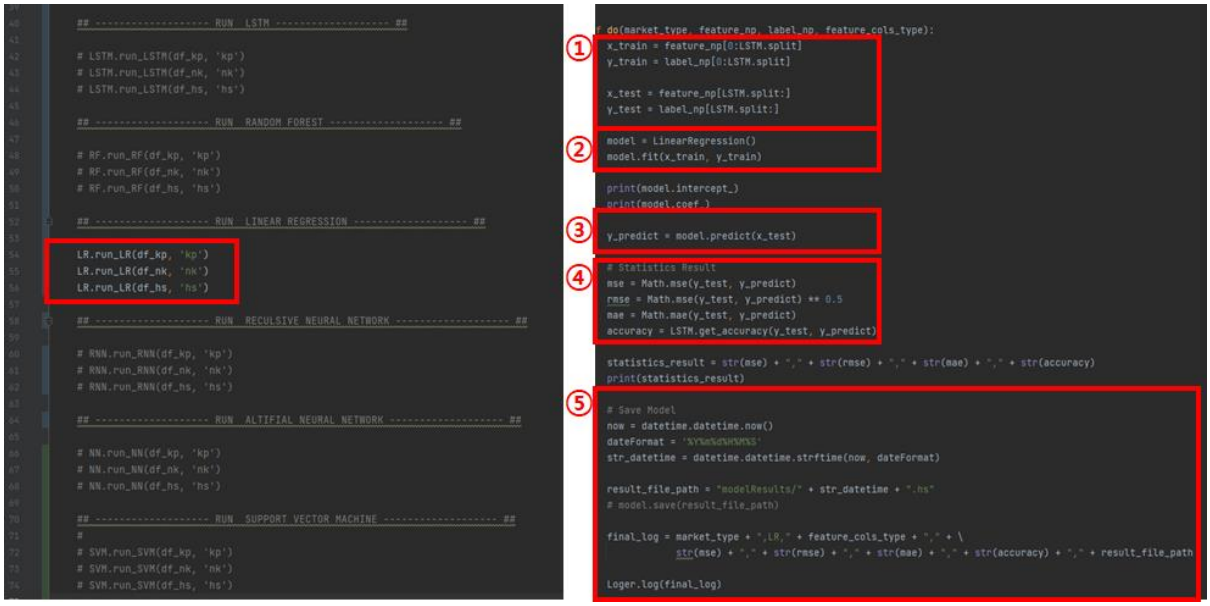
The next step is to pre-process the extracted data. ① As the first step, load the previously downloaded CSV file. ② Using the loaded data, each U.S. Financial market data, exchange rate, VIX index, and Asian stock indexes are all combined by date. ③ Finally, the final data is completed by combining the market price data of the Asian stock index on the next day to be predicted. (For reference, null or blank data processing is performed in the next step.)

3.3 Data Separation Based on Markets



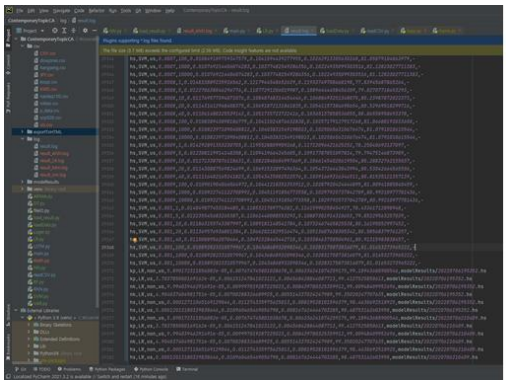
Next, it is a classification of the aggregated data by each market. ① is the process of extracting the entire data for each market, and at the end of the process, data with no value is deleted. The data of each market is separated according to the column shown in ②. ③ In addition, the separation into data with and without US data is made when each algorithm calls data.

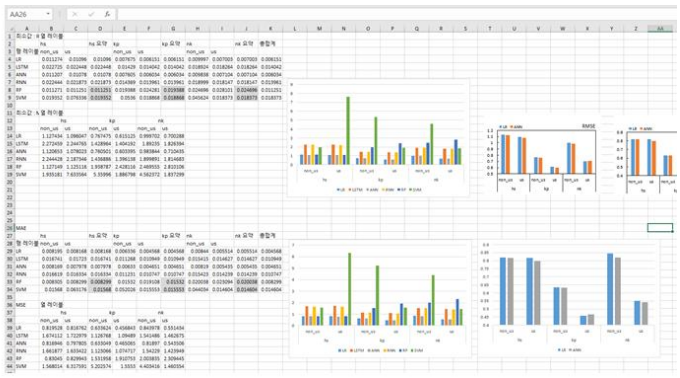
3.4 Run Each Algorithm and Save the Results



Each algorithm is then executed using the prepared data. Each algorithm can be run by uncommenting it as shown in the figure on the left. All algorithms work in the same way as shown below. ① The first step is to separate the training data and the verification data. All data except for the last 100 are used as training data. ② Afterwards, go through the process of learning. ③ Next, prediction is made using 100 verification data. ④ Finally, the error value is calculated by comparing the predicted value with the actual value, and ⑤ the final result is saved. The final result is saved in ‘log/result.log’.

3.5 Analyse the data result





As mentioned above, the results obtained by executing each algorithm are saved in ‘log/result.log’. These saved data are transferred to Excel and the error values of each algorithm are compared. After that, errors by data type, algorithm, and market can be compared, so that the result data necessary for the study can be finally obtained.