EE798 FISA

Endterm project

Tushar Sandhan

sandhan@iitk.ac.in

29 June, 2023

Market Index Prediction

In the financial markets, stock indexes are essential because they offer useful information on the general performance of the industries or the entire market. Stock market indices are used by analysts and researchers to analyze the market thoroughly and identify patterns. They employ indices to analyze past performance, monitor market cycles, and assess the effects of economic events on particular industries. There are two types of investment strategies, passive and active. A passive investment approach, commonly referred to as passive management or index investing, tries to replicate the performance of a certain market index or benchmark. Active investment is an investment strategy that involves actively picking and managing individual stocks, bonds, or other assets with the goal of outperforming the overall market or a specific benchmark.

In this assignment you have to write a python program (no need to submit any `.pdf` report) that will predict the closing price of the given index for the next 2 days, given the information of 50 past days. You are provided with a sample_test.py file which contains the sample function used for evaluation. "You should NOT modify the evaluate() function". Write your code in the predict_func() which should take pandas dataframe as input and returns a python list with two values (closing prices). We will invoke evaluate() for evaluating your code. You are provided with a csv file STOCK_INDEX.csv and you have to use this dataset to train and test your model. You can split the give dataset according to your wish for training your model. To evaluate your model we will use our new test data samples. Below are few samples from the dataset provided to you.

	A	B	C	D	E	F	G
	Date	Open	High	Low	Close	Adj Close	Volume
1	Date	Open	High	Low	Close	Adj Close	Volume
2	2010-01-04	5200.899902	5238.450195	5167.100098	5232.200195	5232.200195	0
3	2010-01-05	5277.149902	5288.350098	5242.399902	5277.899902	5277.899902	0
4	2010-01-06	5278.149902	5310.850098	5260.049805	5281.799805	5281.799805	0
5	2010-01-07	5281.799805	5302.549805	5244.75	5263.100098	5263.100098	0
6	2010-01-08	5264.25	5276.75	5234.700195	5244.75	5244.75	0
7	2010-01-11	5263.799805	5287.200195	5227.799805	5249.399902	5249.399902	0
8	2010-01-12	5251.100098	5300.5	5200.950195	5210.399902	5210.399902	0
9	2010-01-13	5212.600098	5239.200195	5169.549805	5233.950195	5233.950195	0
10	2010-01-14	5234.5	5272.850098	5232.5	5259.899902	5259.899902	0

Figure 1: First 10 samples from the dataset

2010-02-02	4907.850098	4951.149902	4814.100098	4830.100098	4830.100098	0
2010-02-03	4831	4949.149902	4831	4931.850098	4931.850098	0
2010-02-04	4931.299805	4931.299805	4832.350098	4845.350098	4845.350098	0
2010-02-05	4819.649902	4827	4692.350098	4718.649902	4718.649902	0
2010-02-06	null	null	null	null	null	null
2010-02-08	4755.350098	4799.049805	4675.399902	4760.399902	4760.399902	0
2010-02-09	4760.549805	4810.399902	4739.350098	4792.649902	4792.649902	0
2010-02-10	4793	4826.850098	4748.100098	4757.200195	4757.200195	0

Figure 2: The figure shows presence of null values in the dataset which should be handled by you

Your assignment/code will be evaluated on 3 below metrics:

1. Mean Square Error (10%)

The mean square error (MSE) for all the test cases combined will be calculated and compared to other student's MSE scores. The student with the highest MSE (lowest rank) will receive 0 marks out of 10, while the student with the lowest MSE (highest rank) will receive 10 marks. The marks for other students will be scaled linearly based on their rank.

2. Directional Accuracy (10%)

The Directional accuracy will be calculated of all the samples by determining the percentage of correct predictions. For instance, if the directional accuracy is 85.3% the corresponding score will be 8.53 out of 10.

3. Coding Efforts (7%)

If your code does not run, it will be evaluated based on your efforts in completing the assignment. Some points are awarded for good programming methods such as proper variable names, useful comments, indentation, and so on.

NOTE: Your code will be automatically evaluated with no human intervention. If your code does not run on our system, you will be given 5 minutes to make your code run on our system otherwise you will be allotted 0% for Metrics 1 and 2 and your assignment will be evaluated out of 7% only.

Important Guidelines

1. You will have to submit a zip file which must contain a Readme file (with the required libraries and other instructions) and a .py file with the name {Rollno}_{Name}.py ex: 210123_Sahil.py
2. You are strongly advised to not use any unfair means. Use of any unfair means or plagiarism could lead to complain to SSAC and F grade in the course.