Introduction to Machine Learning and AI

Group Project: Thinking about Purchasing Stock Case Study

OBJECTIVES:

To complete Data Collection, Data Cleanse, and Analysis to Make Good Decision. Student teams should select their real world stock data set, refine the data set, and conduct exploratory data analysis using machine learning algorithms to predict good stock prices.

BACKGROUND:

Thinking about Purchasing Stock?

Questions:

1- What is the stock price trend for?
2- Is it possible to estimate the future stock price value?
3- Select any two stocks to compare and to recommend:
   - Amazon
   - Google
   - Netflix
   - Tesla
   - Boeing
   - Walmart
   - Facebook
   - CVS
   - Microsoft

For example, to understand how well Tesla stock price (TSLA), or other stock you select from: Amazon, Google, Netflix, etc., is doing, nineteen years (2000 – 2019) time series, more or less, need to be analyzed to understand and predict the stock price trend. *Do not use 2020 data at all due to COVID-19 impacts*.

Sample of the data. A subset of the data is shown in as below:

https://finance.yahoo.com/quote/TSLA/history?p=TSLA

Download stock data: https://finance.yahoo.com/quote

Date	Open	High	Low	Close	Adj Close	Volume
6/29/2010	19.00	25.00	17.54	23.89	23.89	18766300
6/30/2010	25.79	30.42	23.30	23.83	23.83	17187100
7/1/2010	25.00	25.92	20.27	21.96	21.96	8218800
7/2/2010	23.00	23.10	18.71	19.20	19.20	5139800
7/6/2010	20.00	20.00	15.83	16.11	16.11	6866900
7/7/2010	16.40	16.63	14.98	15.80	15.80	6921700
7/8/2010	16.14	17.52	15.57	17.46	17.46	7711400
7/9/2010	17.58	17.90	16.55	17.40	17.40	4050600
7/12/2010	17.95	18.07	17.00	17.05	17.05	2202500

There were six main variables for the data set named open, high, low, close, adjusted close, and volume. The first four capture how the stock was received in the marked when the market opens, closes and the high and low pick the market. The volume provided a means to understand how many stocks were sold and how much was earned as a result of the transaction. You may observer, there was a visible fluctuation on the volume variable against time but there was positive trend indicated on the variables even if some significant decline was observed for certain period.

Our goal is to use machine learning in order to developing the predicting model in order to show more reasonable predicting values.

For example, selecting your Tesla stock prices will be compared with another one stock, let us say Amazon (AMZN), for example, to recommend which one is to purchase per your models.

TASKS

Feel free to use any regression model or algorithms, showing decision trees, clustering, classification, and visualization, you learned on this class, to find best stock between the two you selected.

TURN IN/SUBMISSION

- Turn in a screen shots of your successful output of code models, in addition to your final code.
- Click on the Final Project folder button to submit your Final project files.
-
    - Stock1.csv & Stock2.csv (rename stock1 and stock2 to real name of your stocks)
    - Save your python code file "finalproject_yourteam..py".
    - Provide snapshot of your output on PDF File called it FinalProject_yourteam.pdf".

- All files will be submitted to the Canvas as a final submission. **One** submission only is allowed.
- **All work will be accepted to the Canvas course web site, no email will be accepted.**

Thank you.

Dr. Majed Al-Ghandour

RUBRIC:

1. Each stock model worths 50% points.
2. Running code is needed to get full mark.
3. Nice and clear python code comments.
4. Clear screen shots on PDF file or HTML if using jupyter.