

ECE 5984 SP22 – Prof. Jones – Group Project II

Due Tuesday, May 3, 2022 – 11:59 PM via Canvas

In this project, your team will develop a model to predict the closing price of a specific stock as traded on the NASDAQ exchange. You are to write Python code to:

- Import a five-year history of the open and close price, trading volume and any other attributes of your team's designated stock (see below);
- Add other features to the dataset as you find them useful (for example, the overall NASDAQ and NYSE index values, daily prices of important commodities such as oil and gasoline, the inflation rate, etc.) – note that some of this data is available on a daily basis while others may be weekly or monthly;
- Preprocess the dataset in the usual manner – check for and correct data problems and missing values, normalize the features, split into training and test sets, etc.;
- Create columns to contain several days' history on one row of the dataset (as you did with the precipitation on Project I);
- Generate your target variable – for this assignment I want you to measure and compare performance in predicting the stock price twenty-eight days in advance (twenty working days);
- Try several models to determine what gives the best performance in predicting the target variables. You MUST try, at a minimum, a linear regression model, a logistic regression model, a neural-network model and an ensemble model of some sort;
- Measure the performance in a manner appropriate for the task;
- Select the best-performing model;
- Generate a graph showing the actual and predicted stock prices plotted by day, for the past five years AND for the past three months;
- Calculate the total income that would be returned, if you had used this model and employed a sensible strategy for when to invest and when not to invest¹. For each day that you decide to invest, assume that you have \$1000 available (this amount does not roll over). On day n , if you choose to, you buy as many shares as \$1000 will buy (partial shares are allowed). Then, on day $n+28$, you sell the shares for the actual price on that day. The income returned after the 28-day period is:

¹ Note: DO NOT DO THIS! Stock prediction is very complex and risky; there are professionals out there who do this. I do NOT want you to lose any money using this approach to actually invest!

$$Income = \frac{\$1000}{ActualClose_n} * ActualClose_{n+28} - \$1000$$

Data:

Here are some data sources that you should look into:

- <https://nasdaq.com> (Look up the symbol for your stock and find the link to download 5-year historical data)
- <https://fred.stlouisfed.org/> Good source of data on general economic indices
- <https://www.indexmundi.com/commodities/>

Stocks:

Team	Stock to use
Team A	Amazon
Team B	AMD
Team C	Apple
Team D	Cisco
Team E	GM
Team F	Google
Team G	IBM
Team H	Microsoft
Team J	Northrop Grumman
Team K	Qualcomm
Team L	Tesla
Team M	Zynga

Report:

Your report shall contain:

- A title
- Team names
- An abstract
- A table showing all datasets that you used – what data is contained, what URL you obtained it from and the date of download
- Data quality reports on the raw data from the file and on the final preprocessed data just prior to partition for modeling
- Performance results for the models that you explored, as described above
- Your choice of the best model – its architecture, parameters and performance
- The graph of actual and predicted stock price for the full five-year period and for the last three months

- Your result for total income
- Discussion – address at least the following questions:
 - Why did you choose the model that you did?
 - Is your model good? Why or why not?
 - Is the computed income a good deal? Why or why not?
 - What would you do to improve this model?
- Conclusions
- Appendix – your Python code pasted in as plain text

Use the IEEE Transactions report template, as in Project I.

Submit your project report and all code, as SEPARATE FILES.