

cpts515 Final Exam (Take Home)

0. Many people think that the dynamics of the stock market is not even predictable. However, after closely looking at a chart, you would agree that the trend of a stock should be predictable; at least, the price movement isn't as random as we thought (see an example chart on page 3). Nowadays, extremely short term price predication is a mature technology, pioneered by mathematician Jim Simons who successfully developed an algorithm in 1980s in predicting stock prices. Nowadays, 70% of stock trades in the United States are done by algorithms and hence, as a computer science graduate student, you shall understand how algorithms are, profitably, applied in the real world.

1. There is a theory called Efficient Market Hypothesis which says that all market information available is already factored into the price of a stock. In other words, stock charts themselves are already enough in predicating market's movements, nothing else needed. Therefore, your predicting algorithms shall **only** be based on stock charts.

2. Surprisingly, a stock chart  $\alpha$  has a very simple data structure: an array

$$\alpha[1...k],$$

for some large  $k$ . Each element  $\alpha[i]$  is the market data for the stock of day  $i$ , which is represented by a tuple of five numbers:

- open price of the day;
- close price of the day;
- highest price of the day;
- lowest price of the day; and
- number of shares traded in the day (called volume).

Notice that when  $i = k$ , day  $i$  refers to today. For major stocks traded in the United States, the charts are publically available for download; most of them are with a history of more than 2000 trading days (10 years); i.e., with the  $k \geq 2000$ .

3. In this exam, you are going to write a mini paper to address the following two problems:

- A. (30%) Given a stock chart  $\alpha$  with  $k \geq 2000$ , design one or more algorithms to predict the price for the next day. I will grade on efficiency, depth and correctness of your ideas.

- B. (30%) Given two stock charts  $\alpha$  and  $\beta$ , design one or more algorithms to measure the similarity between the two. I will grade on efficiency, depth and correctness of your ideas.

The remaining 40% contributes to the quality and clarity of your writing. I shall emphasize again that your algorithms shall **only** take input as the stock chart(s) given.

4. If you only use high-school mathematics (like computing 20-day moving average, drawing some trend lines, etc.), you will receive zero and retake the course. There is a reason why Goldman Sachs is the largest employer for PhDs in physics etc. and the fact that Goldman Sachs does not hire a lot of high school students. I would suggest you to look at Markov chain, graph and finite automata, signal processing, linear algebra and matrix, etc. Also, machine learning would also **NOT** help since for problem A, the input is a single chart  $\alpha$ , and for problem B, the input is two charts  $\alpha$  and  $\beta$  and you won't have enough training data available for the stock. So, for this exam, do not try to use machine learning approaches. Also, do not use ideas like market psychology or social media mood etc since it is not part of the input.

**In zoom review sessions next week, I will provide my own approach using de Bruijn graph to solve the problem. However, if you follow my approach, you will at most receive 90% (i.e., 10% discount)!**

5. **IMPORTANT.** Your mini paper shall be prepared in IEEE proceedings double column format, using the templates at

<https://www.ieee.org/conferences/publishing/templates.html>

with **at least 3 pages**. You will upload your minipaper in pdf. Zero will be given if you do not follow these instructions.

