

... → COMPUTER SCIENCE, DECISION-MAKING, AND DATA → FURTHER EDUCATION

Algorithmic and advanced Programming in Python

Final project instructions

Reminder of the objective of this course

- People often learn about data structures out of context
- But in this course you will learn foundational concepts by building a real application with python and Flask
- To learn the ins and outs of the essential data structure, experiencing in practice has proved to be a much more powerful way to learn data structures
- The mid term evaluation is precisely to start playing with the data structure

Exam day

- We will take the same groups as previously formed and keep the same draw order
- <https://docs.google.com/spreadsheets/d/16AY8vnm69kuHfFu3RM4u3ZJI6dCm13UI2UUCMrCkgzM/edit?usp=sharing>

Instructions for preparing the final project

- Download the data.zip that contains 7 market price information (up to date)
 1. BTC: Price of the Bitcoin in USD updated (BTC_USD Bitfinex Historical Data.csv)
 2. EUR_USD: Value of EUR/USD to reflect appreciation or depreciation of the USD (EUR_USD Historical Data.csv)
 3. Nasdaq: Value of US technological stocks reflected by Nasdaq 100 Futures prices (Nasdaq 100 Futures Historical Data.csv)
 4. VIX: Level of stress in the market reflected by the S&P 500 VIX Futures (S&P 500 VIX Futures Historical Data.csv)
 5. TSLA: Value of TESLA to see any relationship between Bitcoin and Tesla stock (TSLA Historical Data.csv)
 6. Gold: Value of Gold (Gold Futures Historical Data.csv)
 7. SP500: Value of Large cap stocks reflected by S&P 500 futures index (S&P 500 Futures Historical Data.csv)

Tasks to do

1. Add a drop down list to choose which market to load. Markets are BTC, EUR_USD, Nasdaq, VIX, TSLA, Gold, SP500.
2. Use mid term project to select 2 dates and n to display last nth element for the selected data
3. Be able to display the nth best
4. Compute percentage change over 5, 10, 20, 30 and 90 days for all markets (BTC, EUR_USD, Nasdaq, VIX, TSLA, Gold, SP500). Percentage change is computed by $\text{Price}[t]/\text{Price}[t-k] - 1$ where $k = 5, 10, 20, 30$ and 90 days . As we have 7 markets and 5 percentage changes, you should compute in total 35 features.

Tasks to do

5. Use XGBoost as in lab of session 8 to predict the next 5 days percentage change for BTC using the features 35 features computed at time t with a train set up to end 2019 and a test set after (beginning 2020 to today).
6. Compare prediction versus real value by displaying corresponding cumulative returns $(1+\text{RealReturn}).\text{cumprod}()$ versus $(1+\text{EstReturn}).\text{cumprod}()$
7. Comment your results
8. Bonus question: add drop down list for hyper parameters of XGBoost to measure their impact
 - learning_rate
 - max_depth
 - subsample
 - colsample_bytree
 - n_estimators

Instructions for the presentation

- Take 3 minutes to make a demo of the website on your laptop
 - Explain your data structure decision in 2 mn **without any slide!**
 - Leave 4 minutes for Question and answers
 - Your instructor will warn you after 5 and 8 minutes
 - Presentation stop after 9 minutes
-
- **Send us after the presentation within a day your final code.**

Some tips and advices

- Do not start at the last minute!
- If you have technical problems, liaise with the rest of the class and let us know who managed to help you!
- Work as a group and not individually!
- Test before the presentation that everything runs well on your computer to avoid blank presentation in the due day!