MScFE 622 STOCHASTIC MODELING

Group Work Project # 3

See grading rubric here.

Scenario

The asset management team has benefited tremendously from the work this group has done. They are now asking for direct methodologies that can help facilitate stock selection. This group will follow the steps of applying reinforcement learning to portfolio selection.

Tasks

Step 1

Individually, each team member selectively reads the following paper:

> Huo "Risk-aware multi-armed bandit problem with application to portfolio selection" 2017 https://royalsocietypublishing.org/doi/full/10.1098/rsos.171377 paper.

Selective reading means that students focus on some areas and ignore others:

- Students do **not** need to read or understand the section on graph theory. Therefore, they can skip reading Section 2.2. The only key information from that section will be the tickers from the 30 chosen stocks.
- Students do **not** need to read or understand Theorems 2.1, 2.5, and 2.6, and Definitions 2.2, 2.3, 2.4. These are interesting properties of conditional Value-at-Risk, to be sure, but not the focus of the assignments.
- Students **must** closely read other sections, especially Algorithm 1.

Step 2

Team members collaborate and describe the portfolio selection problem as a multi-armed bandit problem. The group will write **pseudocode** that elaborates the Steps in Model 1: Sequential Portfolio Selection Problem (from the Huo paper).

Step 3

Individually, students will collect the sample data used in the paper. The data comes from daily returns during Sep 2008 and Oct 2008.

a. Member A collects the data for 15 financial institutions (JPM,WFC, BAC, C, GS, USB, MS, KEY, PNC, COF, AXP, PRU, SCHW, BBT, STI)
b. Member B collects the data for 15 non-financial institutions (KR, PFE, XOM, WMT, DAL, CSCO, HCP, EQIX, DUK, NFLX, GE, APA, F, REGN, CMS).
c. Member C combines the data into a suitable Python time series data structure. Member C will also compute the daily returns of all 30 series.

Note: In groups of 2, Member A collects all the data, and Member B performs Member C's responsibility.

Step 4

As a group, using the Jupyter notebook, students will then compute a 30 by 30 correlation matrix. They will also produce a heatmap of the correlation matrix. Discussions should focus on how to sort the securities so that similar correlations are closer to each other. The group should write 1 - 2 paragraphs as to how they chose the criteria to sort the 30 stocks.

Step 5

Using the multi-armed bandit methodology studied in Module 6, students will get together and discuss the performance of the algorithm using the Upper-Confidence Bound (UBC) algorithm as mentioned in the module's lesson notes and not the Huo paper. This discussion will help the team understand the algorithm in order to replicate it in the next step. As a reminder, the assigned reading from Sutton/Barto Chapter 2 will help for this and the next few questions. Note: there are no deliverables required for this step.

Step 6

Individually:

a. Member A writes pseudocode that describes the UBC algorithm.
b. Using A's pseudocode, Member B implements those steps in Python. Note that member B is welcome to use Python packages.
c. Member C provides detailed comments of B's code. Member C also applies B's code to the data set.

Note: In groups of 2**,** *Member A writes pseudocode; Member B implements those steps in Python; Member A applies B's code to the data set. (No comments needed).*

Step 7

Using the "epsilon-greedy" policy studied in Module 6, students will get together and discuss the performance of the algorithm using the epsilon-greedy algorithm. This discussion will help the team understand the algorithm in order to replicate it in the next step. As a reminder, the assigned reading from Sutton/Barto Chapter 2 will help for this and the next few questions. Note: there are no deliverables required for this step.

Step 8

Individually:

a. Member B writes pseudocode that describes the epsilon-greedy algorithm.
b. Using B's pseudocode, Member C implements those steps in Python. Note that member C is welcome to use Python packages.
c. Member A provides detailed comments of B's code. Member A applies B's code to the data set.

Note: In groups of 2, Member B writes pseudocode; Member A implements those steps in Python; Member B applies A's code to the data set. (No comments needed).

Step 9

Collaboratively, students will compare the results (i.e. UCB and epsilon-greed). Similarly, they will also compare their results with that of the paper. Collaboration should:

- Describe the results the group produced (1 - 2 pages)
- Compare the group's results with the results of the Huo paper (1 page)
- Present the results in a manner that emphasizes the key differences (1 page, including graphs)

Step 10

Team members will be tasked with updating 30 data series. Collaboratively, members agree on the specific time periods. (e.g. March and April 2020).

a. Member C will import and structure15 financial companies (preferably the same 15 as before)
b. Member A will import and structure 15 non-financial companies (preferably the same 15 as before).

c. Member B will merge the series into a single data structure and compute the returns.

Note: In groups of 2, Member B imports and structures all 30 companies; Member A will merge the series into a single data structure and compute the returns.

Step 11

In groups, team members rerun the algorithms (both UCB and epsilon-greedy) on the new data. Instead of merely rerunning the algorithms with the same parameters, members are **encouraged** to vary other parameters, such as holding period, or use other policies that illustrate their understanding of learning strategies from the curriculum in modules 5 and 6.

Then, students write a 1-page technical report addressing the following question:

What are the results when more recent data is used?

Submission Requirements and Format

One team member submits the following on behalf of the entire group:

1. **1 zipped folder** including:
    a. An executable Jupyter notebook* that includes the code, its output, and the answer to each question along with the solution
    b. A duplicate version of the Jupyter notebook above in PDF or HTML format. In order to include the output of the code, you must RUN the code before downloading the PDF.
2. **1 PDF document** containing ONLY the answers to the questions, EXCLUDING code
    a. Use the available Report Template and fill out the required information in the first page.

****Use Google Colab or GitHub to collaborate** in completing the executable Python program.*

NOTE: *The PDF must be uploaded **separately** from the zipped folder that includes any other types of files. This allows Turnitin to generate a similarity report.*

Rubric

Your instructor will evaluate your group submission for GWP1 using the following rubric:

Quantitative Analysis (Open-Ended Questions) 40 Points	Technical and Non-Technical Reports 30 Points	Writing and Formatting 20 Points
The group is able to apply results, formulas, and their knowledge of theory to real-life finance scenarios by doing the following: ● Providing all the necessary information to support their arguments. ● Presenting arguments that reflect group discussion and research. ● Using authoritative references to support a position and provide updated information. ● Concluding with practical takeaways for more insightful financial decision-making.	Technical reports contain 3 parts: 1) **code** for each question (be sure to explicitly state the question number), 2) the corresponding output of that **code**, and 3) **interpretations** and/or **recommended courses of action** that reasonably follow from those results. **Note**: Technical reports will include the technicalities of models, such as names, methods of estimation, parameter values, etc., and exclude generalities about the work done. It should NOT include names of Python code that were used.	A submission that looks professional should: ● Include the axes, labels, and scales in graphs. ● Be free of significant grammatical errors or typos. ● Be an organized, well-structured, and easy-to-read document. ● Include proper citations and a bibliography in MLA format.
	Non-technical reports contain 3 parts: 1) clear explanation of results; 2) the recommended course of action that follows; and 3) the identification of factors that impact each portfolio. **Note**: AVOID all references to model names, algorithms, and unnecessary details. Instead, focus on the investment decision.	