Members Contributions

	Contributions	%
Jean-Gabriel	Web Scraper Charts News rating (LLM) Local Web page Research	50
Nicholas	IBKR API implementation Strategies IBKR TWS setup Backtesting Research	50

Meeting minutes : around 2h30

Summary of issues:

- Web Scraper
    - Yahoo finance stopped working
        - At first, we used yahoo finance to scrape news in their news tab for each stock. However, at some point the web scraper stopped updating the news shown. We tried a few different things to fix it but were not able to. We think it might have been because of a call limit to the site. To resolve this, we modified the code a bit to scrape news from Finviz, which is a financial platform with a lot of useful data, including news.
    - Most articles need a click on read more to see the full text
        - A lot of finance related articles only show a portion of the text and have a button to expand the page to show the rest of the content. Since each sites have different HTML tags for the content, we used Selenium to click through each site if there is a “Read more” or “Continue reading”. Using this method, we can scrape more sites. However, this method is not perfect, and some sites do not return any content which is why we will continue to work on the web scraper to improve it.
- Charts
    - First library didn’t work
        - For the charts on our web page, we wanted to used a python library based on TradingView charts which are some of the best in the financial field. This library was working but did not work well on the web page. After trying multiple different approaches and doing more research, we decided to use the library directly from TradingView, which use mostly Javascript and comes with some pre-built elements. It was also a lot easier to use since we could put the javascript code directly into the HTML.
- LLM
    - Finding a free option

        - One of the difficult part to implementing a LLM in our project was to find which one to use. We obviously started our research by looking if we could use ChatGPT, Copilot or Bard since they are the most well-known. However, all these were paid options which we didn't want to use in the early stages of our project. We did more research and considered a few options such as roBERTa, VADER, BERT, etc. We decided to use BERT since after watching a few tutorials for each options it seem that BERT was one of the easiest to implement into our project and since it was made by Google, the pre-trained model should be quite good. Our project currently use BERT with a publicly available pre-trained model which does respond to our need as of now for POC, but we are considering to moving to a more advanced option later to get more accurate ratings.
    - Reading the data from the spreadsheet
        - Other issues we encountered during the semester was feeding the scraped data to the LLM. A lot of them were caused because the data was saved incorrectly in the .csv file, which caused the wrong data or no data to be accessible to the LLM. After modifying the code to how the scraped data was saved, it was a lot easier to work with the .csv file and after a bit of research we were able to find the correct commands to read specific columns and rows in the csv file.
- Web page
    - Displaying the analysed data in an organized manner
        - Displaying the data to be easy to read on the web page was a bit difficult because the content of the article was usually a lot longer than the headline. This problem was apparent in the first iteration when we tried to use table/grid. We decided to try the bootstrap accordion which was a lot better since we could only show the headlines and expend the content when they were clicked. A bit of work was needed to display to correct content in each box but it made the news ratings a lot easier to read.

- IBKR API
    - Which Library to choose
        - There are various libraries that can interact with the IBKR API, we are currently operating with the ib_insync library. This library is free, and fulfills most of our basic needs and concerns. However, it remains to be seen if we can utilize it to fulfill all our requirements.
            - An alternate library might be used in the future if it is impossible to continue with our requirements. The logic will remain the same; currently, strategies are treated separately from the API call file.
            - If funding is acquired, a pivot to IBKR's Web API / FIX API / or TWS API might be used instead.
    - Securing Connection
        - We are using ib_insync to connect to our TWS API, the code does not divulge sensitive information since we are connecting directly to TWS API which is on the user's computer through a port connection. If the user does not have IBKR TWS, it will not work.
- Research
    - Understanding of financial models and macroeconomic concepts
        - During the meeting with our client, we discussed various financial and economic concepts which could be applied to our project. Some of these concepts were quite advanced for our understanding since they were beyond the scope of the Telfer classes we took the date. Because of this, a lot of time this semester was spent researching and learning financial notions since our client wanted to take a macroeconomic approach to the algorithmic trading. Moreover, we also needed to determine what data/information was useful for our strategies.

Estimated Time Spent

Issues	Time spent (h)	Contributors
Web Scraping	8	J-G
Charts	6	J-G
Web page	12	J-G
LLM	12	J-G
Backtesting	8	Nic
IBRK (API & TWS)	10	Nic
Strategies	20	75% - Nic, 25% - J-G
Research – Macro concepts	30	J-G & Nic - 50/50
Research – Financial models	30	J-G & Nic – 50/50
Research – Fundamental Data	25	J-G & Nic – 50/50
Planning	10	J-G & Nic – 50/50

Have you achieved what you planned for the first semester?

- Yes, we achieved what we wanted this semester. We knew the first semester would be mostly research and setting up the groundwork of our project. As of now, we have a good base that we can improve as we learn more about the markets and the technologies we use.

Plan the next semester:

- Summer
    - We also plan to work on our capstone project with our client during the summer even if we did not take the official class. Our plan for the summer is to start backtesting the different strategies we have, in order to find best performing ones. We also plan to try live paper trading in the second half of the summer if backtesting goes well.
- Fall
    - For the fall we plan to continue live paper trade or set it up if we were not able to during the summer. Based on the results we obtain we will try to refine some aspect of the strategy and maybe even create strategies for specific stocks of sectors. Eventually, if everything goes well, we might try to do live trading with real money using our algorithm.