

Enio Peza & Özyürek Mustafa https://github.com/Mustafa6614/Project\_10

# FENTUS STOCK MARKET

IR Project – 2024/2025

# **Final Report**

# **Assignment**

The goal of this project is to implement a working prototype of an information retrieval system. The system needs to gather a large collection of samples that are related to the assigned topic and be able to search over this collection. The system that has to be implemented must provide an interface for searching, browsing and presentation of the data to the user

### Websites

Our assigned topic is Stocks and we were provided with a website by our professor to crawl data from: "<a href="https://www.borsaitaliana.it/borsa/azioni/listino-a-z.html?initial=A">https://www.borsaitaliana.it/borsa/azioni/listino-a-z.html?initial=A"</a>. After this site we had to identify at least other two sites similar to the first site from which we should retrieve data to expand our dataset. After a long search, we selected two other websites.

The second website is <a href="https://www.tradingview.com/markets/stocks-usa/market-moversgainers/">https://www.tradingsiew.com/markets/stocks-usa/market-moversgainers/</a>, which is based on the best gainers from the U.S market. The third site is: "<a href="https://www.tradingsat.com/actions-de-a-z/">https://www.tradingsat.com/actions-de-a-z/</a>, which is a website based on French stocks mainly and not only. So, we are gathering data from these three sites with one containing information about the Italian stock market, one for best gainers in the U.S market and the last one for the French stock market.

#### **Features**

The system that we should build must also have at least two additional features (one for each category), which are divided in *simple* and complex *features*.

We chose three features for our websites. Two are simple features and the third one is complex;

- Simple Features:
   Result Presentation (results displayed in tabular format)
  - Filtering: Name (Alphabetical order, ascending or descending)
    - Price (Numerical order, ascending or descending)
    - Volume (Numerical order, ascending or descending)
- Complex Feature:
   Automatic Recommendation (suggest similar stocks)

## **Implementation**

#### Data

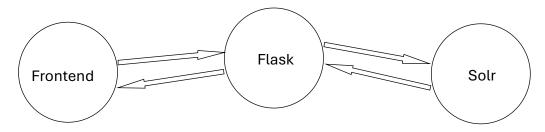
As it was mentioned the data for the stocks was crawled from three different websites. The data that was crawled than was saved in 3 different JSON documents. However, to ensure that data was clean and suitable for indexing in SOLR they were normalized which involved converting, cleaning and standardizing the data fields. To ensure that all records followed a consistent structure, the missing fields were converted to "null" identified as N/A in our IR System. Each stock record contained the fields such as: name, last price, high, low, volume and market gap. So, there are present 3 different JSON documents with three JSON normalized documents in python.

#### Apache SOLR



Apache SOLR is a free, open-source search engine based on the Apache Lucene library. In this project SOLR was the backbone of the search engine, providing efficient indexing, querying and retrieval of data related to stocks. SOLR *indexed* the stock data to allow the fast and efficient searches but as we said before the data could be normalized by SOLR it first had to be normalized. SOLR\_Client class encapsulated the logic for communicating with SOLR by sending HTTP requests to SOLR'S API endpoint. It handled various types of search queries as: full text queries, range queries and filtering (one of the features that had to be implemented). Also, the complex feature *recommendation* was implemented thanks to SOLR's search capabilities to provide stocks with similar attributes.

#### Flask



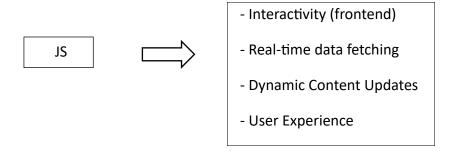
Flask is a lightweight web framework used to create the backend API for processing search queries and communicating with SOLR. It acted as an intermediary between the frontend and SOLR. It received search requests from the frontend, constructed appropriate SOLR queries and fetched the results from SOLR.

In our project is presented in *app.py* to handle the backend.

#### HTML and CSS

HTML and CSS were used to design and structure the *frontend* of the application. The user interface includes search bar, filter options, display section for search results and recommendations.

#### **JavaScript**



JavaScript has also a really important role on the creation of this application. JavaScript handles the interactivity of the frontend, such as sending search requests, updating the UI dynamically and implementing the pagination and filtering for our search engine. It handles the User Interactions for example: search button click, enter key etc. JS uses the fetch API to make HTTP requests to the Flask backend. JS dynamically generates HTML content to display search results in a tabular format (simple feature) and inserted them into the DOM (Document object model) ensuring this way that the results were displayed instantly to the user without a page reload.

# View of the IR System

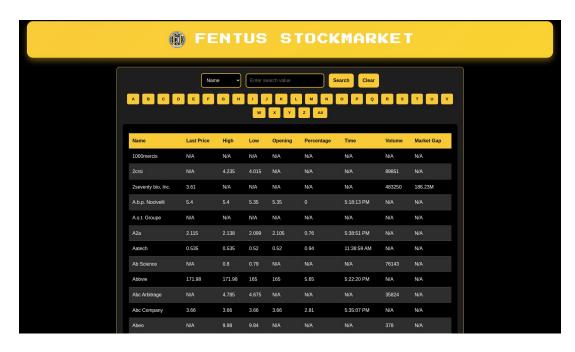


Figure 1 - Main Page

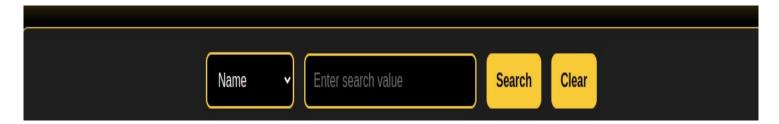


Figure 2 - Search Panel

Users can search their stocks by writing their name directly in the second tab where is written "Enter the search value" and then execute the search by pressing the search button on the right or by simple pressing Enter on your keyboard. As it will be explained under you can also filter the results and search about stock that you want not only by name but depending on other attributes you are searching your stock to have.



Figure 3 - Filter Panel

As said above users can search about their stocks not only by their name but selecting different attributes. Users can select one of the attributes as filter to find their stock. We have different attributes such as name, last price, high, low, volume, market gap.

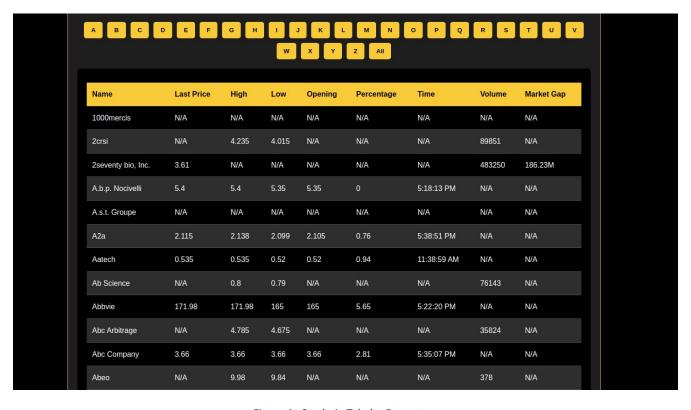


Figure 4 - Stocks in Tabular Format

All the data is arranged in tabular format, and they are ordered in alphabetic order with the stock starting with number which are on top. There is the option that you can press any of the letter button on yellow to go to the stocks which start with that letter of the alphabet and find them in order.



Figure 5 - Previous and Next Button

Related to the stocks tabular format you can go to the previous or the next page of the stocks by pressing one of these two buttons.

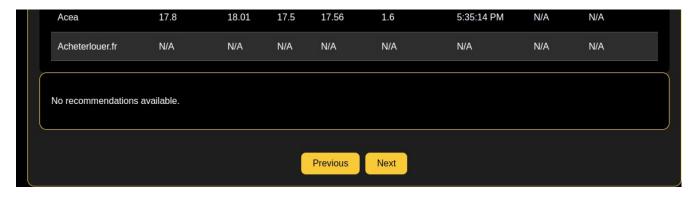


Figure 6 - Recommendation Panel (in the Main Page)

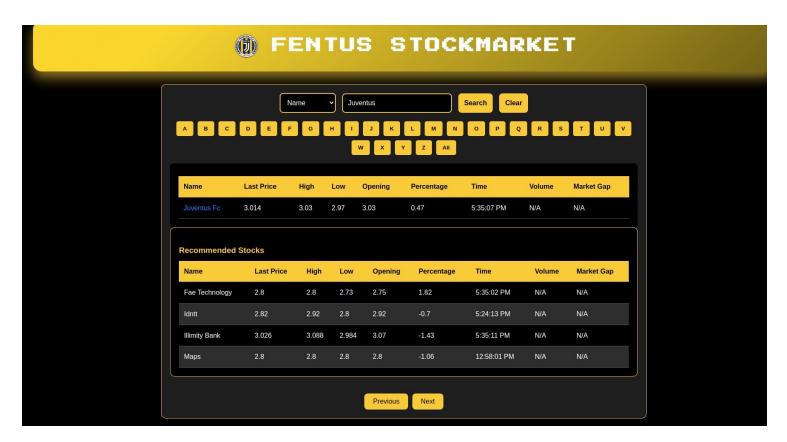


Figure 7 - Searching "Juventus" and the recommended stocks

In this case the user has searched for the name stock "Juventus" which results in a stock and finds it how we can see above. The main difference that you will see in the case of the searching is the recommendation panel which now actually recommends real stock. In the main page before the search, you will never see recommended stocks because you are not interacting just watching the stocks not searching anything. But the moment that you put a query, you will get recommended stock that are similar in attributes with the stock that you initially searched.



Figure 8 - Searching "Roma"

As we see in this case no stock "Roma" is found so we get the message "no results for your query" and we don't get any stock recommended since the stock is not there so we can't have stocks that in a certain way are similar to the search query "Roma".

# **User Evaluation**

To understand the usability and effectiveness of our IR System, we asked 6 of our classmates to perform 5 tasks. All the participants are male and range of age 21-27. These tasks were designed to evaluate the system's search functionality, ease of navigation and overall user experience. They are also asked to grade from 1-10 the difficulty of the task (10 - extremely difficult, 1 - extremely easy).

The tasks that the participants had to complete were:

- 1) Search Juventus / Ferrari then click on it.
- 2) Search a stock that has volume 50.
- 3) Clear the search of number 2) and then go to the second page of the stocks that begin with the letter B.
- 4) Search about a stock that you are interested in and observe the recommended stocks if they are interesting since they should be similar to the stock that you just searched (if the stock you searched isn't found then you will get "no result found for your query" so then you just search about another stock.

5) Just navigate around the IR system and try to get opinions that might be useful to us if the system is easy / hard to navigate, satisfies your research, is done well, you like the structure, you like the design etc.

# **Results**

#### • Task 1

All users were able to successfully complete these tasks, and they didn't have any issues.

6/6 graded this task: 1

#### • Task 2

All six users were able to search for stock with volume 50. They had positive feedback on how the filter panel was implemented and how it was easy to use.

6/6 graded this task: 1

#### Task 3

Only four users were able to successfully complete this task. The other two users had the same problem. They instinctively pressed the letter B directly which would get them nowhere since they first had to clear the search of the question number two, then go to letter B and press "Next" at the end of the page only once (this is how the task would have been completely successful). After we explained to them how to solve it right, they responded that it was not so obvious that they had to clear before searching for another stock. Even though they admitted that the task was actually clearly explained, and they reacted directly without thinking on pressing directly B before pressing "Clear" button.

2/6 graded this task: 3 1/6 graded this task: 4 1/6 graded this task: 5 2/6 graded this task: 9

#### Task 4

Only three users completed the task successfully. They were able to find the query they were looking for on the first try. Out of these three users that completed the task two of them found the recommended stocks interesting and useful. The other three users didn't find the stock they were looking for after two queries.

3/6 graded this task: 1 2/6 graded this task: 3 1/6 graded this task: 4

They didn't find this task difficult to execute they just couldn't find.

#### Task 5

In general, the users had a positive impression of the IR system. They really like the system's interface and design. They said that the combination of dark background with bright highlights made the page visually appealing and easy to read. They liked the tabular format and the attributes that were chosen for the different stocks. The filter panel was easy to use and very intuitive. Same as the letters of the alphabet that are used to search stocks that begin with that letter. They suggested us to retrieve even more data, so we have even more stocks than the ones that we already have. This suggestion comes especially from the students that were not able to successfully complete task 4 since they couldn't find the stocks that they were looking for. They also liked the recommended panel which only recommends stocks only when they search for one specific stock and the stocks that were recommend were actually of their interest. So, all together, we can say that they liked the IR system and found it to be well-designed, functional and user-friendly.

6/6 graded this task: 1

# **Individual Evaluation**

Both of us give each other a 10. We worked very well together, effectively dividing tasks based on our strengths and interests. Setting clear deadlines for specific tasks helped us stay organized and on track during this entire time with the project. We consistently shared ideas and suggestions on doubts or problems for the different tasks that we had to do. We maintained an open communication, which allowed us to solve problems quickly. We are both really proud of how the project turned out and feel that our efforts resulted in a functional and well-designed IR System.