

Chrome Extension for Top Stocks Analysis

Team - JADE

Jhinuk Barman, jbarman2

Anumeha Agrawal anumeha2

Devang Aggarwal devanga3

Team Captain: Anumeha Agrawal

1.Introduction

We have created a Google Chrome extension which shows useful information for the top 10 trending stocks day. Everyday the extension gets refreshed, before the market opens, with the latest trending for the day, including the name of the stock and the price change. We will then use this information that we have scraped to show useful data and visualizations like - latest news on that company and the top trends. We also get the latest news for the day that influences the stock market and do sentiment analysis to show how it impacts the market. We use the CNN finance website to also get data for a specific stock when the user visits the page of that stock. This includes the price change and the user sentiment where positive indicates buy and negative indicates sell. This chrome extension can be used to display the top trending stocks with some visualizations, the latest news headlines and how they affect the stock market and also information about specific stocks when a user visits the CNN finance page for that stock.

2.Software Implementation

This section describes the details of the software implementation and how this chrome extension was developed and can be used.

There are two main components in this software

1. Chrome Extension
 - a. We first set up a flask server that performs data extraction and analysis for the extension and renders data to the chrome extension.
 - b. The base extension involves several cohesive components such as background scripts (in this case the background script fetches data from the flask server and renders to the extension popup), content scripts, UI elements and various logic files. The extension was created with the following web development technologies: HTML, CSS, and JavaScript.
 - c. In order to enhance the UI of the extension we also use pretty table and matplotlib.

2. Web scraping and sentiment analysis

- a. First trending stock data was scraped from <https://money.cnn.com/> using beautiful soup by analyzing the html components of the website. We also extract the top trending news headlines on the website.
- b. We then perform sentiment analysis by extracting data from these scraped news headlines URLs using nltk sentiment analyzer. This helps us analyze how the news affects the stock market and whether it motivates one to buy a certain stock or sell it.
- c. We generate visualizations using matplotlib on the top stocks data which helps us understand the trend better.
- d. We also scrape data and perform sentiment analysis when a user visits a particular stock's website on cnn.

a.Pre-requisites

- 1) Install python3 and conda
- 2) Using conda install flask, bs4, matplotlib, prettytable, nltk, mpld3, pandas
- 3) Once nltk is installed, run the following from a python console
`nltk.download('vader_lexicon')`
- 4) Install git and clone the repository -
<https://github.com/anumehaagrawal/CourseProject>

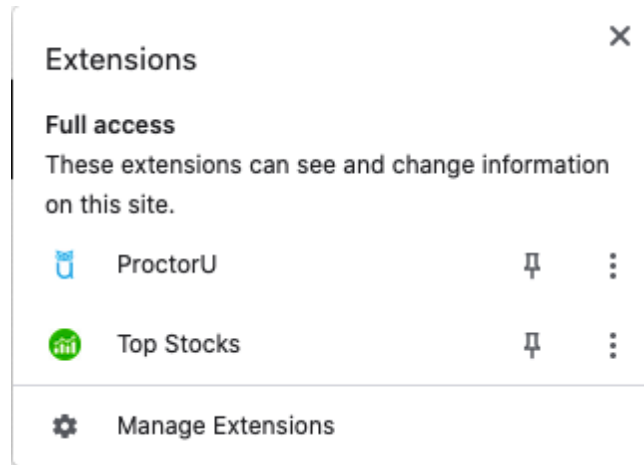
b.Installation Details

- 1) After the repository is cloned, cd into the project directory > flask app and start the server using -> python3 server.py. The flask server is now up and running.
- 2) On the chrome web browser, go to the extensions tab and click on Manage Extensions and turn on developer mode.
3. On the right hand side there is an option called load unpacked and this will allow us to upload the project directory.

The extension is now up and running!

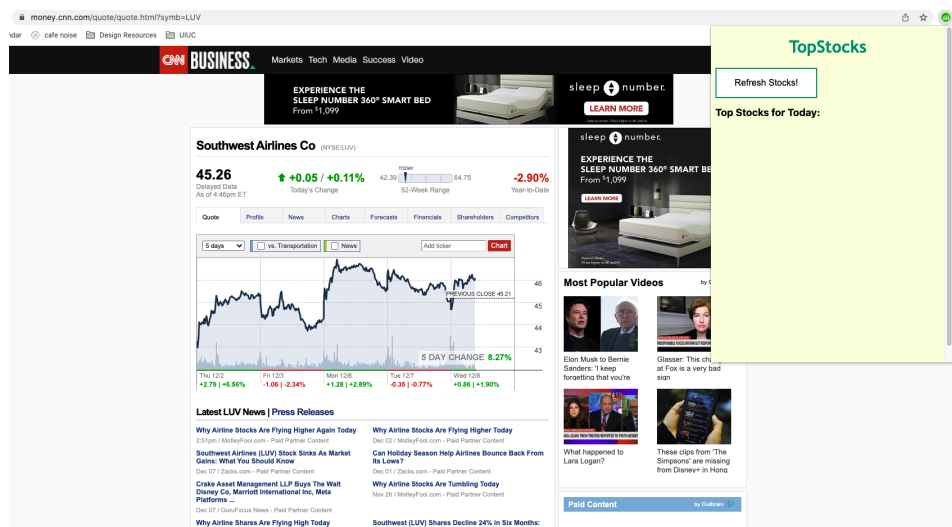
c.Usage

The chrome extension can be accessed via the extension button at the top left and then clicking on "Top Stocks." There will be a green stock market graph icon to the left of it as well. The extension can be accessed from any web page.



Top Stocks:

After clicking on the extension a “Refresh Stocks” button will appear. The user can click on this button to get the information on the top 10 daily stocks in the format of tables and graphs. The CNN finance page will be web scraped to collect this information. Additionally there will be tables and graphs for sentiment analysis on the top stock headlines from the CNN finance page.



If the user is on a specific stock page on CNN (for example: the Southwest Airlines stock shown in the image above is at the link <https://money.cnn.com/quote/quote.html?symb=LUV>) then a summary of the stock will appear in a table. If the user is on any other webpage, everything except for this specific stock summary table will be displayed.

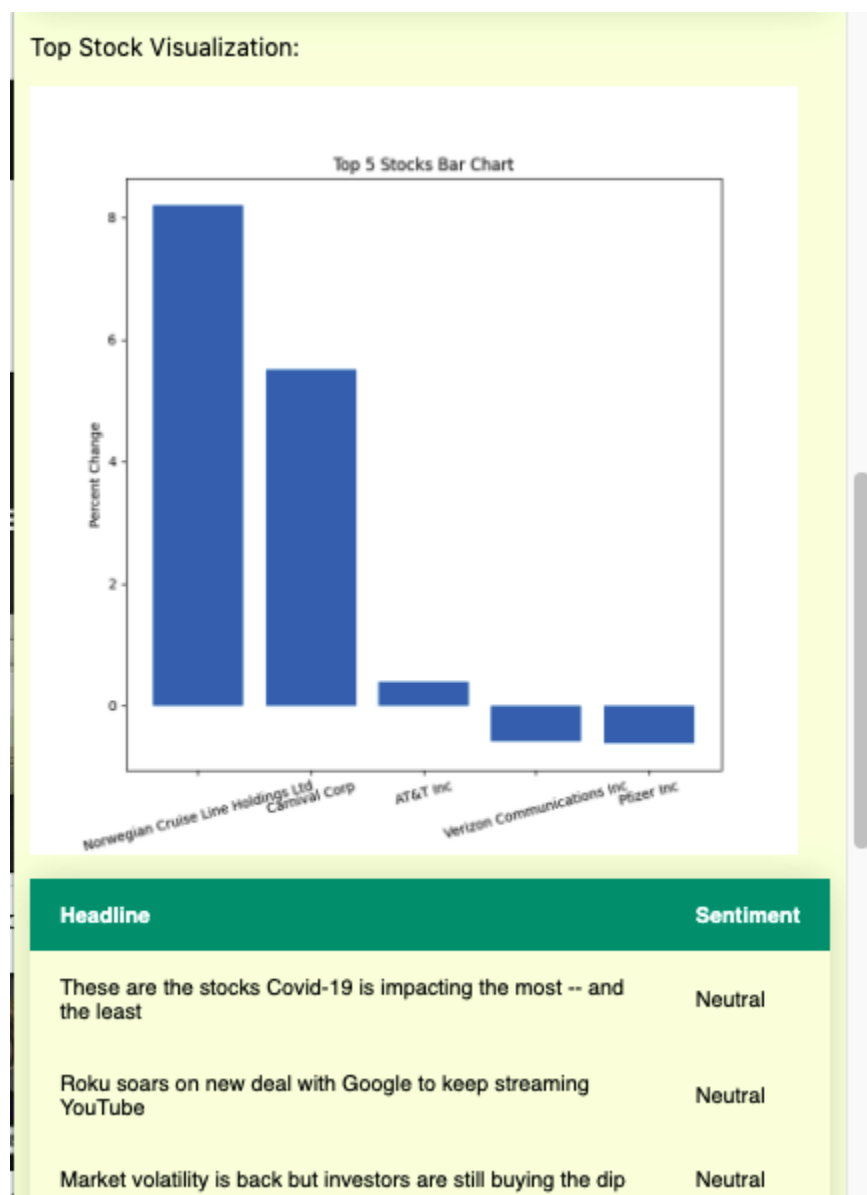
The summary table includes the stock name, daily stock price, percent change and sentiment. The sentiment analysis is generated by web scraping and collecting the top headlines related to this specific stock. After conducting sentiment analysis on these headlines, a polarity score is assigned to each headline. The polarity score determines if the sentiment is positive, neutral or negative. The sentiment is positive if the polarity score is greater than 0, neutral if the score is equal to 0 and negative if the score is less than 0. The majority of the sentiments determined the final sentiment for the stock.

After this summary, a table of the top 10 daily stocks will be displayed including the stock name, price and percent change. The stocks are sorted by decreasing percent change.

TopStocks			
Refresh Stocks!			
Stock Name	Stock Price	Percent Change	Sentiment
Southwest Airlines Co (NYSE:LUV)	45.26	+0.05	Neutral

Top Stocks for Today:			
	Company	Price	Percent Change
1	Norwegian Cruise Line Holdings Ltd	21.77	8.2
2	Carnival Corp	19.52	5.51
3	AT&T Inc	23.17	0.39
4	Verizon Communications Inc	50.49	-0.59
5	Pfizer Inc	51.4	-0.62
6	Citigroup Inc	62.46	-0.7
7	Ford Motor Co	19.81	-0.75

Below the top stocks table is a bar graph visualization of the top 5 daily stocks. The x-axis is the name of the stock in the order of the top stocks table. The y-axis is the percent change of each stock. This will show how the percent change of the top stocks vary.



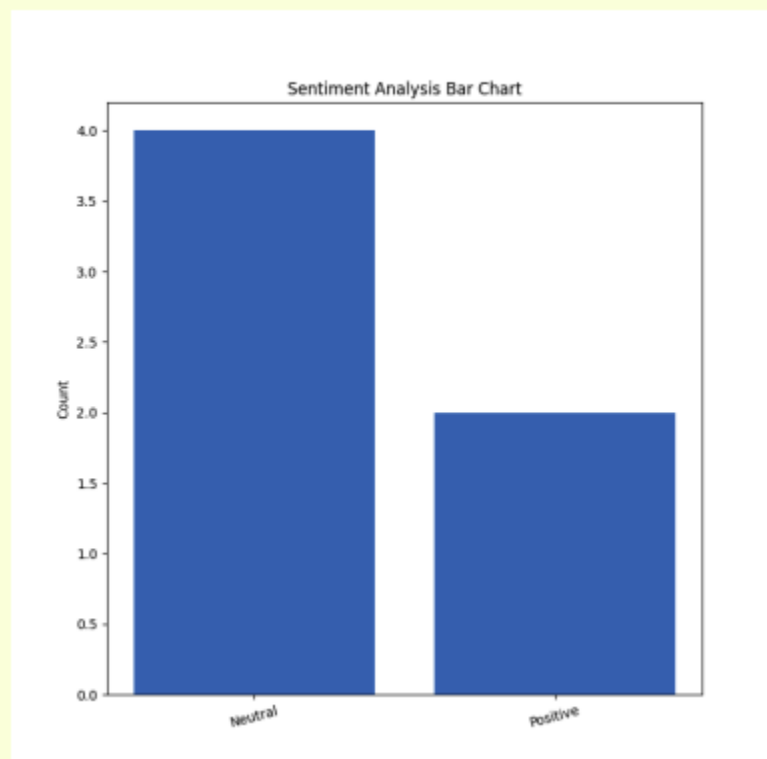
Sentiment Analysis:

The next section of the extension will show a sentiment analysis table. The columns are the Headline and Sentiment for the top stock headlines. The headlines are web scraped from the CNN finance page. The sentiment analysis is conducted on each headline and given a polarity score. The process for determining positive, neutral or negative sentiment is the same as described earlier for specific stocks (polarity score of 0 is neutral, less than 0 is negative and greater than 0 is positive).

Headline	Sentiment
These are the stocks Covid-19 is impacting the most -- and the least	Neutral
Roku soars on new deal with Google to keep streaming YouTube	Neutral
Market volatility is back but investors are still buying the dip	Neutral
Trying to make sense of the Omicron market? Look at bank stocks	Neutral
Wall Street extends rally	Positive
There's little reason for Robinhood investors to be merry	Positive

The last part of the extension will show a visualization of the sentiment analysis on the daily top stock headlines. This bar chart shows the count of each sentiment. For example in the bar graph below, the count of neutral sentiments is 4, positive sentiments is 2 and negative sentiments is 0.

Sentiment Visualization:



Lastly, the refresh button can be used to refresh the stock data and receive real-time information in the form of tables and visualizations.

3.Contributions

The entire project work was divided equally amongst the three team members and each one of us worked on a crucial component of the project. Anumeha worked on developing the flask server and the end to end chrome extension. Devang worked on web scraping and getting the data for top trending stocks and the top news headlines. Jhinuk worked on the sentiment analysis of the news data along with enhancing the UI of the extensions and adding visualizations.

4.References

<https://developer.chrome.com/docs/extensions/mv3/getstarted/>

<https://www.nltk.org/>

<https://www.digitalocean.com/community/tutorials/how-to-make-a-web-application-using-flask-in-python-3>

<https://beautiful-soup-4.readthedocs.io/en/latest/>

<https://matplotlib.org/>