STOCK FUNDAMENTALS AND TECHNICAL INDICATOR ANALYSER

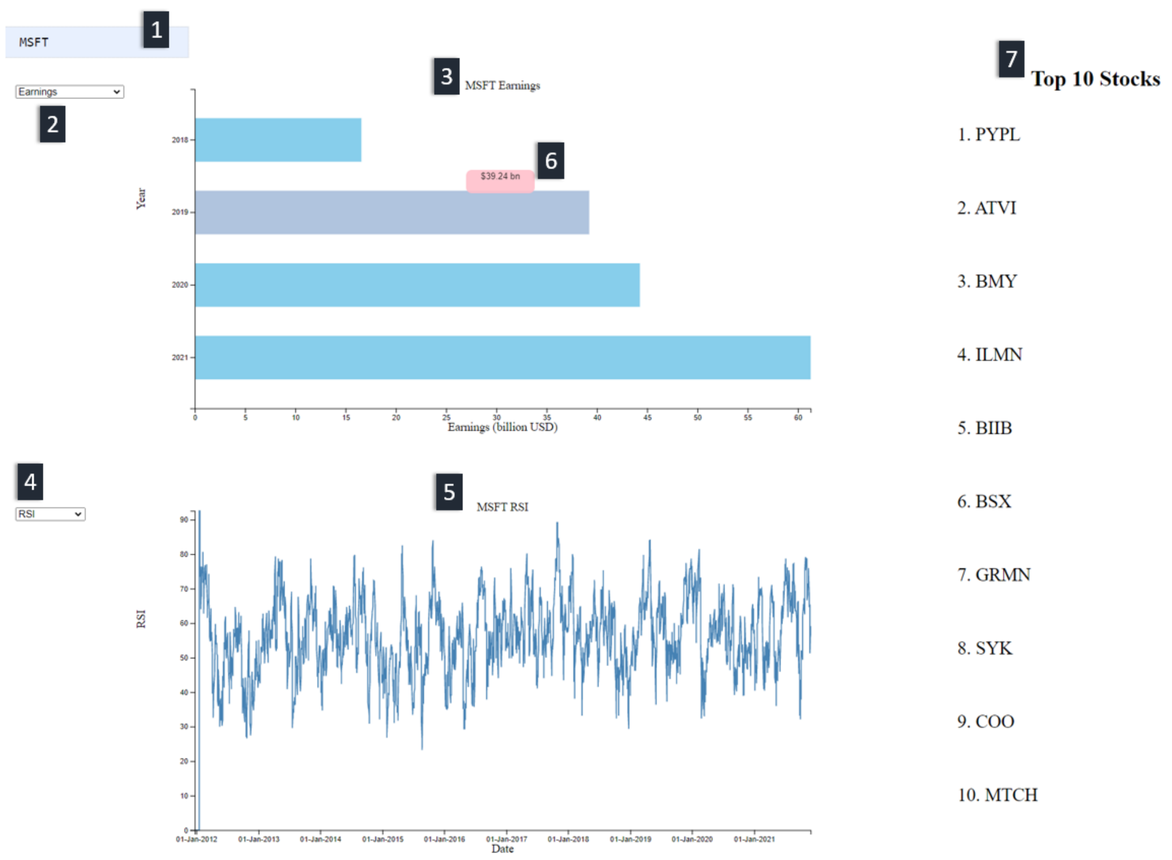
Harshad Hazariwala | Daniel Muehring | Michael Nock | Soham Pandya | Richard Reube | Somreeta Roy

Motivation

There is currently no mainstream website that provides visualizations of both fundamental data and technical indicators on the same page, along with stock recommendations.
This project provides a convenient way to visualize both technical and fundamental indicators to guide investors towards the most lucrative financial choices.

Approach

The visualization is interactive in which the user can select any stock from S&P 500 and the visualization will populate both the fundamental and technical data for the selected stock. Dropdown menus for the charts enable the user to view the data of their choice. Based on a logistic regression algorithm that runs in the backend, the visualization will suggest the top 10 stocks to buy.



[1] Search bar to type stock TIKR of choice. [2] Drop down to provide options for fundamental data. [3] Bar chart for fundamental data. [4] Drop down menu to provide options for technical indicators. [5] Line chart for technical indicators. [6] Tool tip showing the exact value the bar represents. [7] Section showing the top 10 stocks to buy according to the logistic regression model

Our financial data, sourced from **Yahoo Finance** APIs is stored in an **SQLite database**. We use it to calculate technical indicators in **python** utilizing pandas data frames, which will also be stored with the financial data in the same database. The aforementioned logistic regression algorithm combines the technical (RSI, MACD, EMA) and fundamental signals (Debt-to-Equity Ratio, Price-to-Earnings Ratio and Price-to-Revenue Ratio) into a score that can be used to make a buy recommendation. The data from the database is served on request in JSON format (a REST API) via a small **Flask app** running on the backend. The JSON data we need can be directly requested and used by an **html script using D3** to provide a user-friendly interactive **visualization of fundamental and technical data**.

The product provides a **simple, intuitive and novel visualization** for investors that currently is not available.

Logistic Regression has never been implemented on this combination of technical and fundamental indicators.

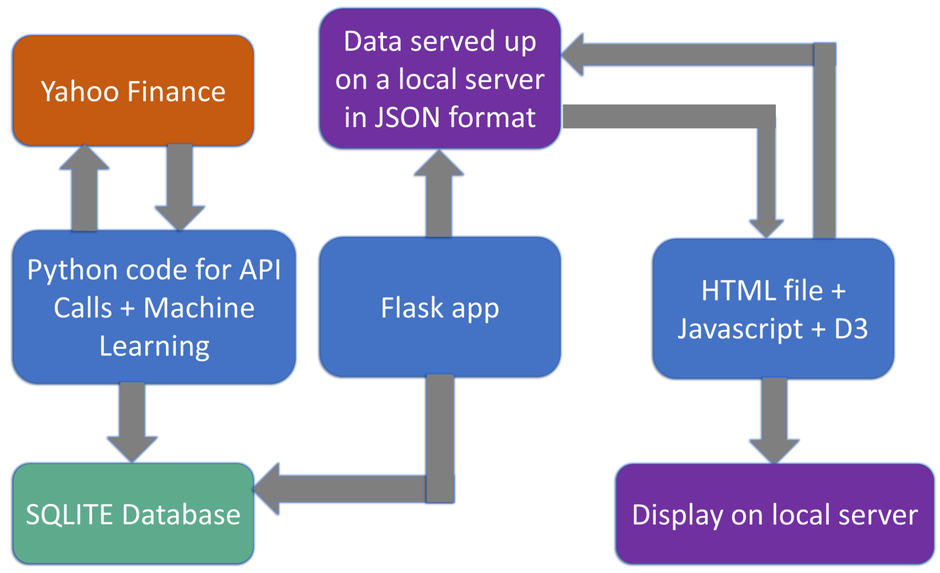


Illustration to show how our approach works and how the different modules interact

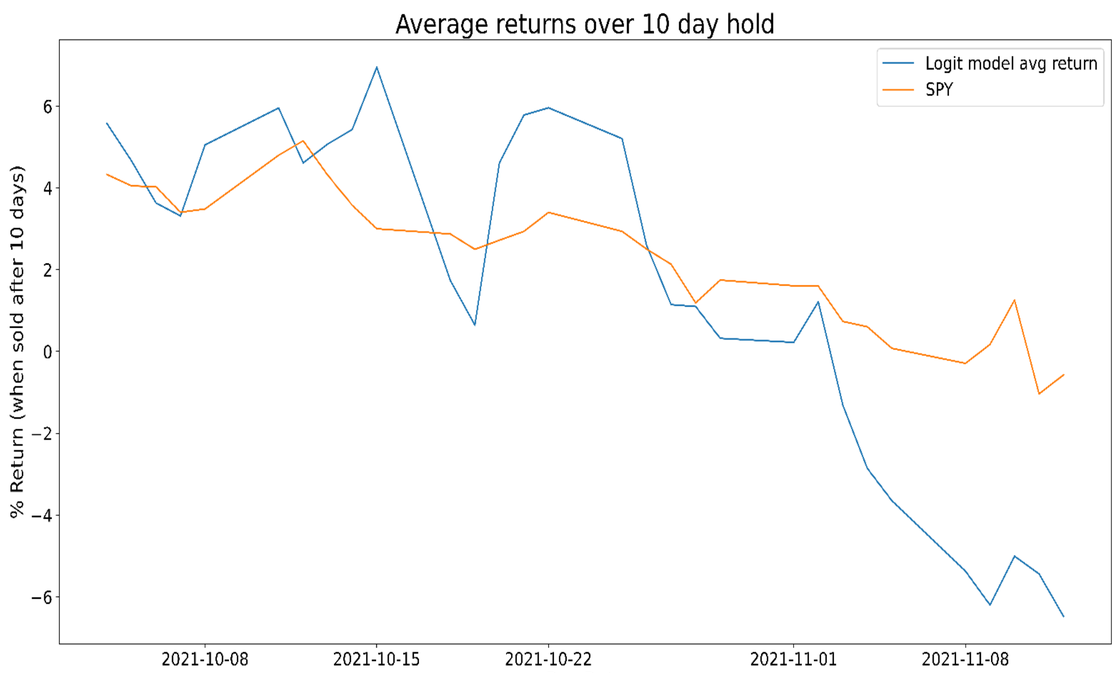
We compared logistic regression against other approaches (linear regression, KNN, SVM etc.) and found that this approach will provide a better indication (higher returns) of when one should buy to realize a profit within a 10-day trading window.

Data

Raw financial data for the past five years for 500 companies which make up the S&P 500 were obtained through Yahoo Finance APIs, using Python and stored in a SQLITE database. The size on disk for the temporal data obtained is 260 megabytes, with the number of records at 1.25 million.

Experiments and Results

The prediction accuracy for the average returns proposed by the logit model was back-tested on a sample of 500 stocks with the buy dates over a 30-day time period. The benchmark is that the model needs to perform better than S&P 500.



Results of evaluation of the logistic regression model

Our model outperforms S&P500 for the first 15 days, however when the market turns lower, our model underperforms. One possible explanation is that RSI is not a good indicator when the market is trending downwards. The RSI would hit a bottom very quickly which would falsely indicate to the model that it's a good time to buy. One strategy to improve our model would be to detect when the market is in a downtrend. During such periods, the model can either suggest the user to avoid trading or extend the hold period until an uptrend has been detected. Simple K Nearest Neighbors and Linear SVM models were also run to see if any other classifiers might perform better. Neither of these options were substantially better.