

Stock Trading and Market Analysis Project

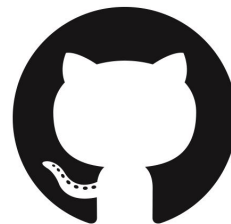
Angelo Vacca, Kasper Seglem, William Ruiz

Introduction

- Dataset and Tools
- Trello
- What We've Done
- Stock Correlation
- Correlations by Year
- LSTM Prediction
- LSTM with Correlation
- Random Prediction
- Analysis and Findings

Dataset and Tools

- S&P 500 companies
 - 2013 to 2018 (5 years)
 - date, open, high, low, close, volume, Name
 - 2013-02-08, 67.7142, 68.4014, 66.8928, 67.8542, 158168416, AAPL
 - 20 year data
- Python
 - matplotlib
 - numpy
 - pandas
 - Seaborn
 - Tensorflow
- GitHub



Trello Board

The screenshot shows a Trello board interface for a project titled "Data Mining Project". The board is organized into five columns: "Milestones", "Todo", "In Progress", "Bugs", and "Complete". Each column has a header with a plus icon and the text "Add a card". The background of the board is a scenic image of a forested mountain with a lake.

Board Header: Trello Workspaces Recent Starred Templates Create Search

Project Name: Data Mining Project

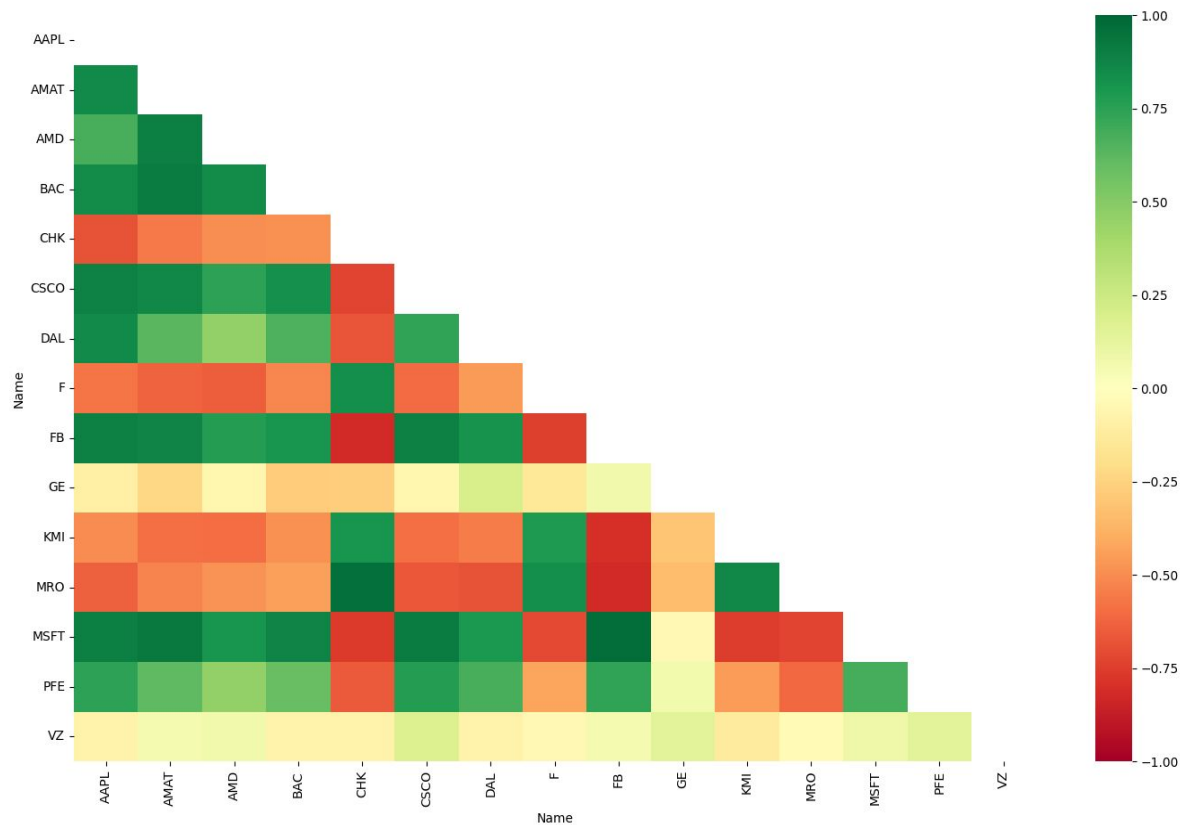
Columns:

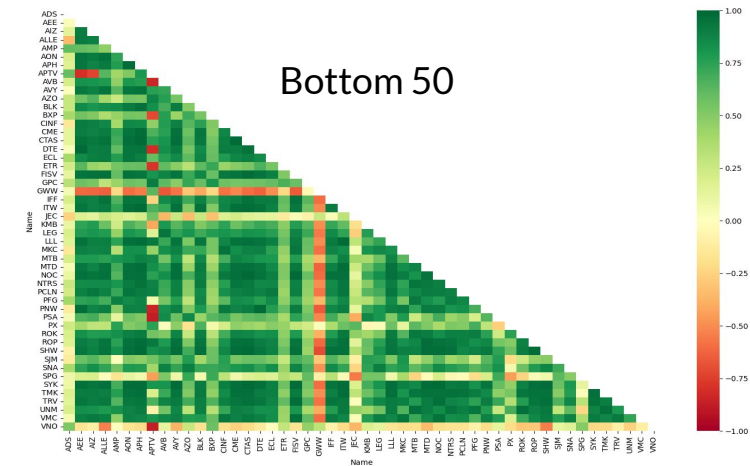
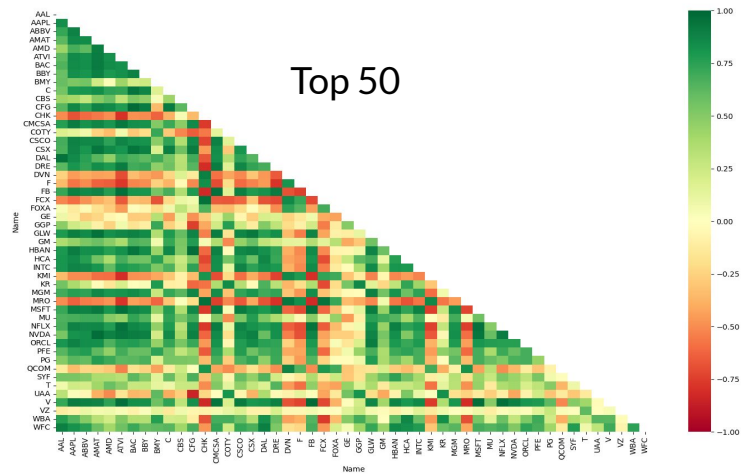
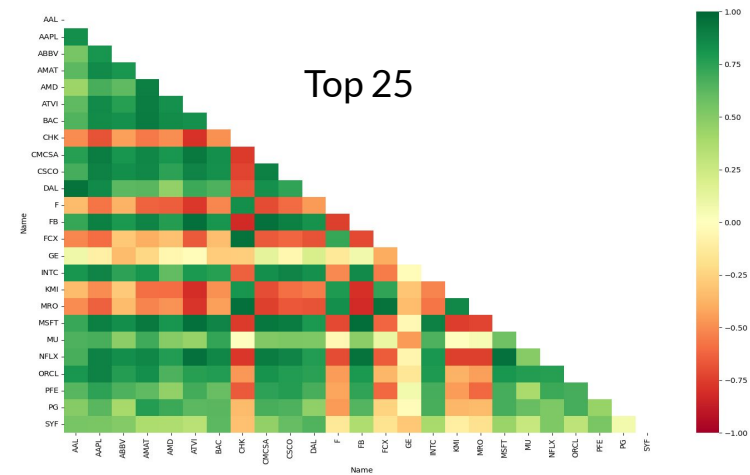
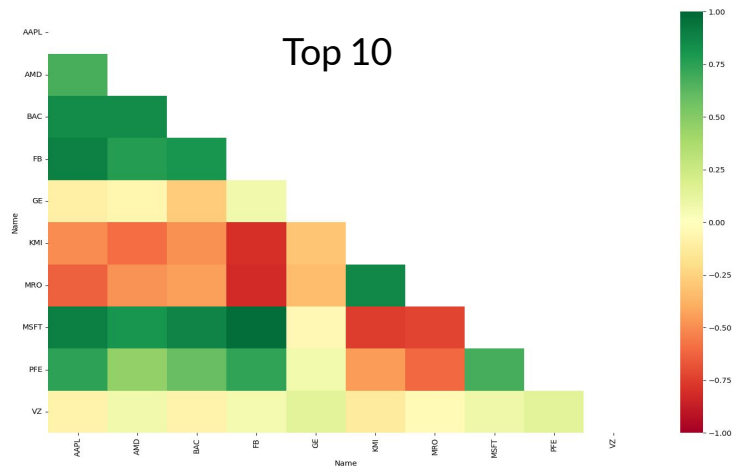
- Milestones:**
 - Collect past and current stock data from the S&P 500. This includes finding an up to date dataset with the history of each stock in the S&P. (Week 7)
 - Study developed trading algorithms and models used to help predict stock movements. This includes learning more about the Hudson River Trading algorithm and the J.A.R.V.I.S. algorithm. (Week 10)
 - Project checkpoint report. Put together a report to document current progress. (Week 12)
 - Study and mine data in order to gain a better understanding of historical stock data by finding trends, patterns and outliers in movements. Use the S&P 500 data and begin to mine patterns and correlations. (Week 13)
 - Use data from the previous milestone to create a rudimentary predictive model/algorithm based
- Todo:**
 - + Add a card
- In Progress:**
 - + Add a card
- Bugs:**
 - + Add a card
- Complete:**
 - Project Checkpoint Slides
 - Download data
 - Small Demo
 - Project Checkpoint Report
 - Analyze data and correlation between a cluster of similar stocks and the market
 - Analyze data and correlation between market and one stock
 - Check LSTM models with correlated stocks
 - Create function to take in a dataset, parameter(volume, open, close, high), top/bottom, num of stocks to show
 - Prediction model for individual stocks

What We've Accomplished

- Studied trading algorithms
- Found and cleaned up dataset
- Found correlation between stocks
- Compared correlations by year
- LSTM prediction model
- LSTM model with correlated stocks
- Random prediction model

Correlated Stocks



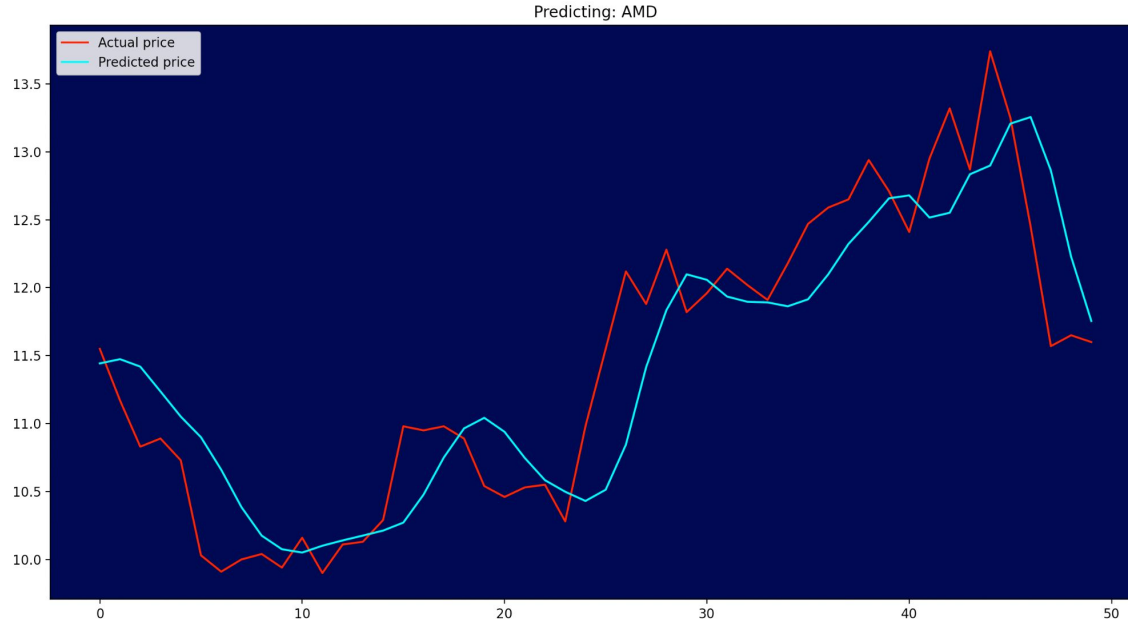


About Correlation

- Can vary greatly depending on time frame
- Correlation does not imply causation
- Helpful tool for diversifying portfolio
 - Negatively correlated stocks for limited losses
- Delayed correlation may work
 - Industry trends

Using machine learning to predict prices

- Using Long Short-Term Memory machine learning model
 - We were able to utilize the LSTM model in order to roughly predict stock prices
 - Utilized the “close” attribute of the stock data to train the model to predict prices



Using machine learning to predict prices

- Using the Random Tree Classifier model
 - This model allowed us to utilize all of the attributes of the stock data
 - We could also add more information to the table by calculating trends
 - More accurate model and could be utilized for actual trading vs basic prediction
- Initial findings (AMD: 09/15/2017 - 02/07/2018)

Initial money: 1000

Total money: 1064.0800000000004

Total shares: 0

Total trades: 48

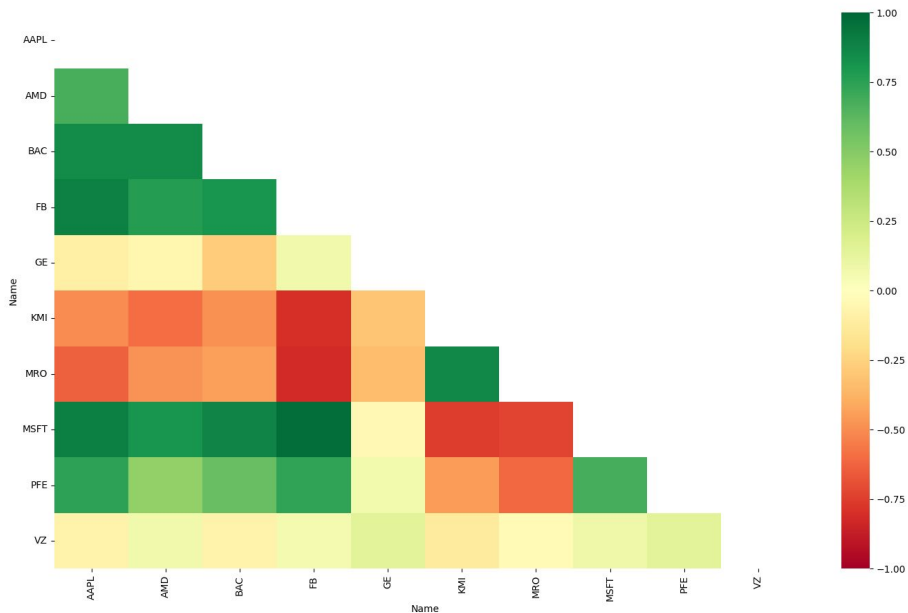
Percent gain or loss: 6.4080000000000036%

Using machine learning to predict prices

- Using the Random Tree Classifier model
- More intensive run:

Initial money: 1000
Total money: 1090.7799999999997
Total shares: 0
Total trades: 55
Percent gain or loss: 9.0779999999999964%

Total money: 1000
Total money: 1015.08
Total money: 988.859999999
Total money: 988.859999999
Total money: 513.389999999
Total money: 1001.579999999
Total money: 1001.579999999
Total money: 961.679999999
Total money: 43.629999999
Total money: 481.289999999
Total money: 947.869999999
Total money: 933.589999999
Total money: 470.339999999
Total money: 66.729999999
Total money: 1016.55
Total money: 260.289999999
Total money: 1081.33
Total money: 1088.199999999
Total money: 1090.779999999
Total money: 1090.779999999



Findings

- Use correlation data to build predictive models
 - We were able to utilize our correlated stocks in order to train predictive models that fit well for more than one company
 - Highly correlated stocks tend to have a more accurate predictive price versus non-correlated stocks
- LSTM
 - Using a Long Short Term Memory model we were able to predict stock prices for various companies
 - Trained multiple models to predict many different stocks
- Random Tree Classifier
 - Using the Random Tree Classifier we were able to utilize more parameters to train a model
 - Using the predictions we can calculate the amount of money we could make
 - Found that on average we would make money, around 7% profit over 100 days
- Testing
 - As we hoped, we were able to create models with positive returns when tested on market data

Evaluation

- Work has be evaluated by observing if our algorithm successfully predicts the actual change in the market.
- Metrics for evaluation will be the accuracy or error of our predictions, especially when compared to that of other prediction algorithms.
- Our models have been tested by letting it invest and make its own decisions
- We have compared our work with the standard trading trends in the market
- We have compared our work with random guesswork

Milestones

- Study other trading algorithms and models (Week 7) COMPLETED
- Collect past and current stock data (Week 10) COMPLETED
- Find correlation between the top S&P 500 stocks COMPLETED
- Project Checkpoint Report (Week 12) COMPLETED
- Study and mine data for patterns, trends, outliers, etc. (Week 13) COMPLETED
- Create a predictive model/algorithm based off patterns and trends from mined data (Week 14)COMPLETED
- Create/use a site to host the application (Week 15) Possible
- Final Project Report (Week 16)

We are happy with our results!

Questions?