Capstone Project 2:
Stock Market Analysis

Scott Penn

Project Goals

- Understand the indicators of stock success and failure.
- Create a reusable data pipeline.
- Improve model for use in trading platform.

Trading Platform: Alpaca API

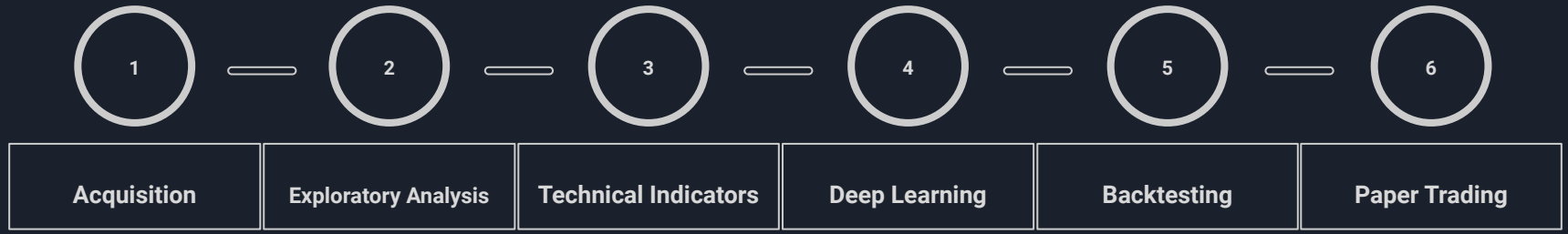
With a free account:

- Use the Alpaca API with REST/Python bindings to download historical stock data.
- Buy and sell stocks using a Paper Account with any amount of Paper Money.

With a verified account (only available to US Citizens currently):

- Backtest your algorithm without needing a custom backtesting suite.
- Access to the Polygon API for more historical stock data acquisition options.

Data Pipeline

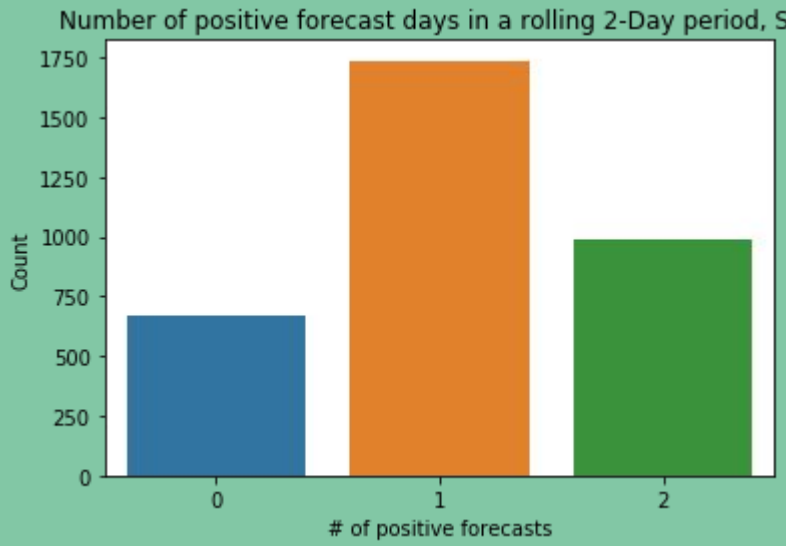


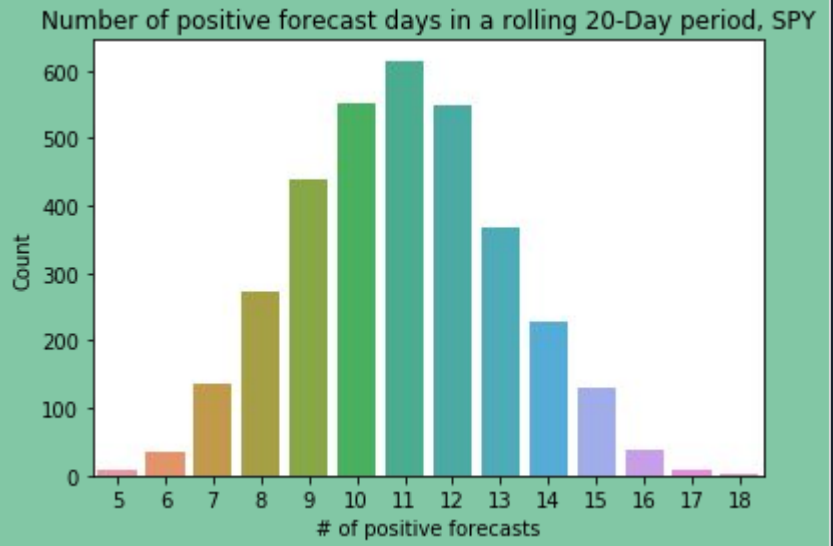
All of these steps can be run separately.
This allows for easy modification of any step.

Exploratory Data Analysis

Daily stock data was collected from 2007-2020. Stock symbols from the S&P 500 were used, a collection of frequently traded stocks on the US markets. To track the movement of the market as a whole, the SPY ETF (Exchange-Traded Fund) was used.

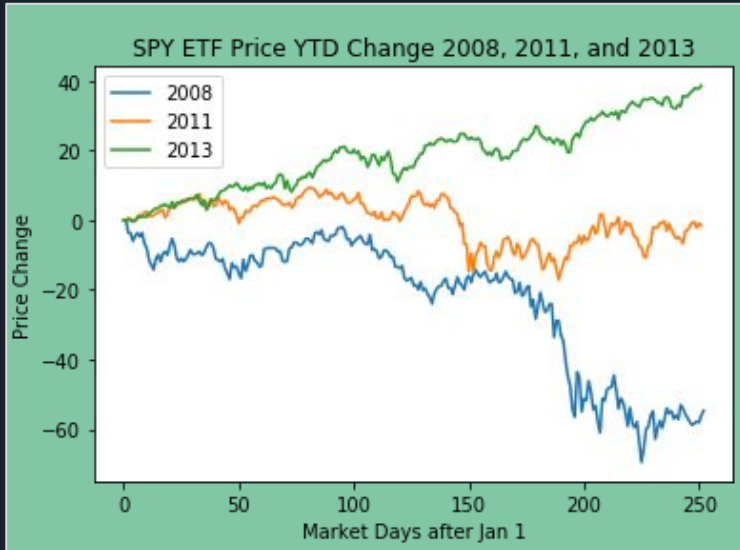
of Positive Market Days in a Rolling Period

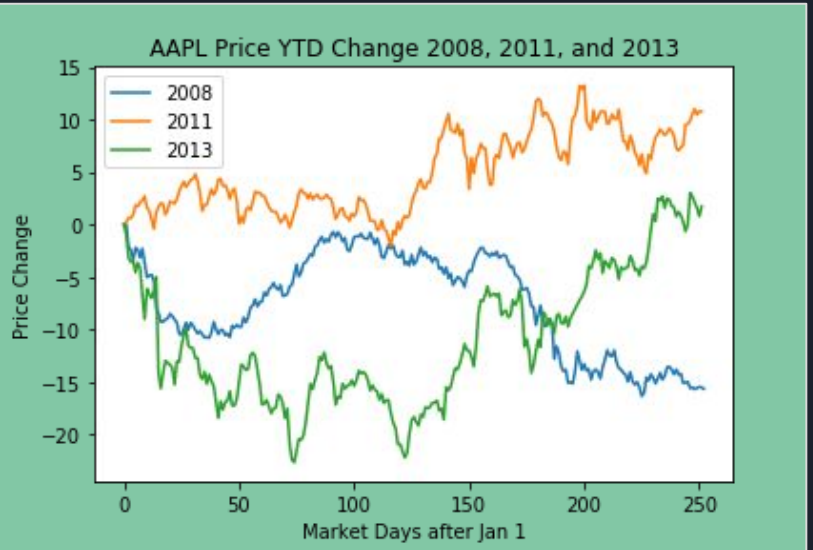




Market Forecast vs Individual Stock Forecast

On average, individual stocks go the same direction as the market 63% of days. This varies by stock and industry, with financial institutions moving with the market 75% of days.



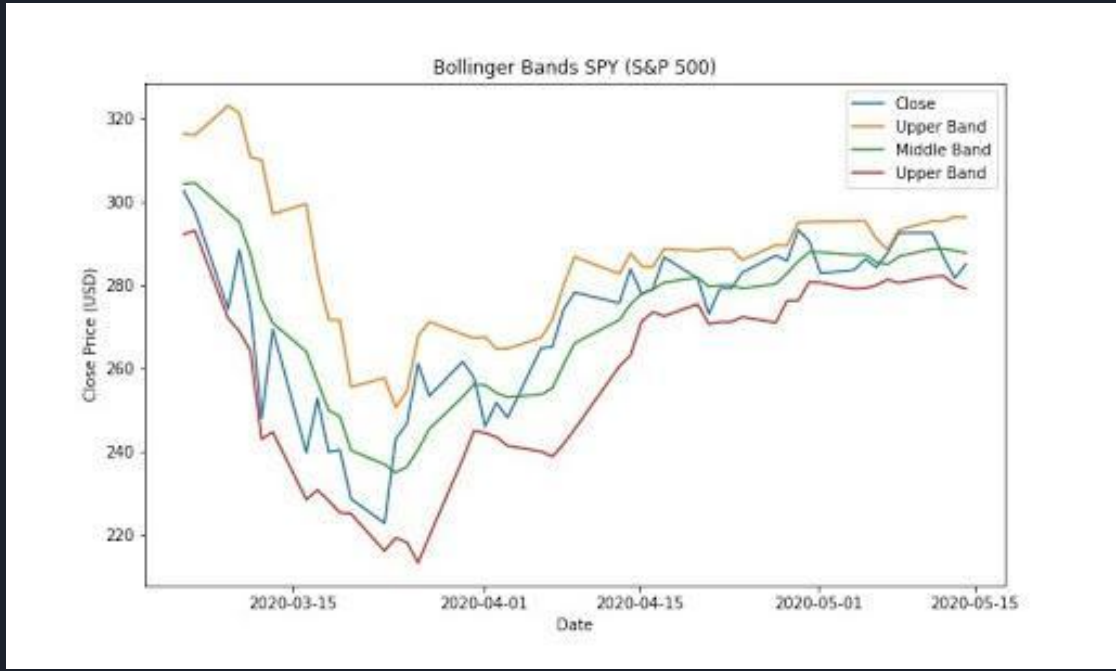


Technical Indicators

A technical indicator is a fancy term for a function that modifies historical stock prices. This modification usually reveals a trend or a signal that can be used to make decisions when buying or selling stocks.

Types of Technical Indicators

- Overlap Studies
- Momentum
- Volume Indicators
- Statistic Functions



Machine Learning

3 Approaches:

1. All Features Included (Neural Net)
2. Market Features Excluded (Neural Net)
3. Market Features Excluded (XGBoost)

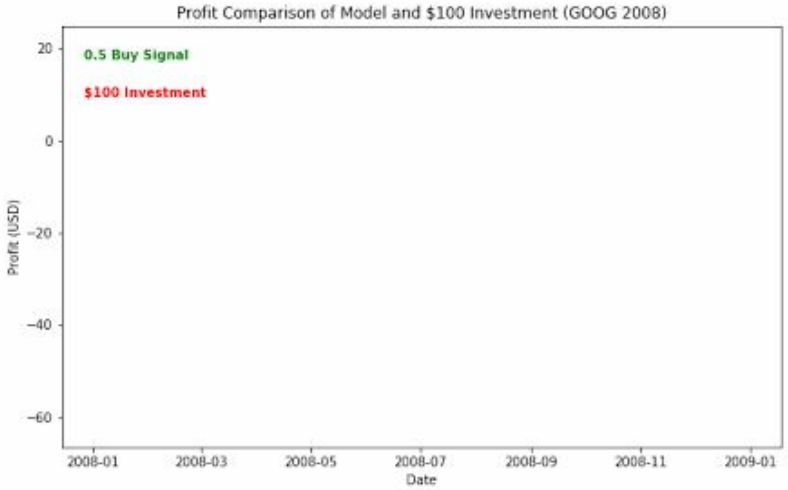
	w/ Market Features	w/o Market Features	XGBoost
Accuracy	~71%	~56%	~53%
F1 Score	0.737	0.621	0.567

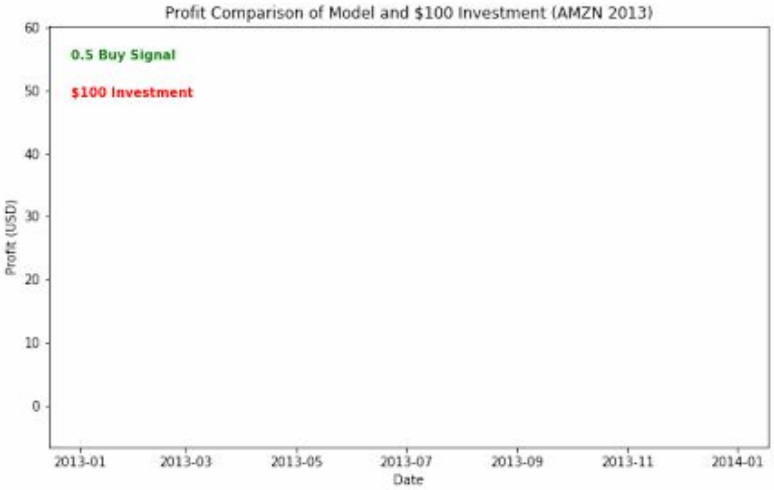
Loss vs. Profit or Why Backtesting is Important

While the Market model performed best by machine learning metrics, it performed the worst in terms of potential profit. By predicting market success, it missed stocks that went against the predicted market movement.

Loss may be an important metric, but it might not be the metric you are looking for when determining model success.

Backtesting Results in 2008 and 2013





Next Steps

- Incorporate New Features
- Trade on the Paper Markets
- Tune Buy Signal
- Automate Trading Script

Conclusions

While my models outperformed the market in specific circumstances, almost every model was beat by the market in the long run. Investing is always a sound strategy, but any attempt at trading stocks should be seen as gambling. Stocks are influenced by so much more than the sum of their past prices. However, the analysis of prices can be used to gain a better understanding of past events and their impact on individual companies and the market as a whole.