

Capstone_3 Final Report

Stocks Analysis and Trading Strategies Final Report

Introduction

In 2022, 58% of adults in the United States had some kind of exposure to the stock market.[1] Investors and traders use different methods to decide entry and exit points for trading stocks. Although there are several trading strategies, they do not work in all scenarios and due to external factors like macroeconomic conditions, financial results and other factors it becomes difficult to develop profitable trading strategies. Data Science can be leveraged to develop and back test different strategies to see which strategies are more profitable than others in different market scenarios.

Problem Statement

How can Data Science be leveraged to evaluate and develop trading strategies, which can outperform the market returns for year 2022 for selected US Stocks and ETFs in 3 months

Scope

Daily closing price data for stocks and etfs available at Yahoo Finance API will be used for developing and back testing different models

Constraints within solution space

- Stock prices are affected by economic factors like inflation, bond prices and federal interest rates and may adversely impact the trading models.
- Fluctuation of stock price within day could not be considered

Data

Daily closing price for exchange traded funds like QQQ, SPY, DIA and VYM for the period 2010 to 2023 is considered for analysis.

Criteria for Success

Successfully develop models based on Bollinger bands, MACD, Relative Strength index, neural network in 3 months

Bollinger Band Strategy

Bollinger Bands®[2] are composed of three lines. One of the more common calculations uses a 20-day simple moving average (SMA) for the middle band. The upper band is calculated by taking the middle band and adding twice the daily standard deviation to that amount. The lower band is calculated by taking the middle band minus two times the daily standard deviation

Bollinger band strategy is mean reverting strategy and is based on the assumption that the market can be overbought and oversold. At the most basic level when a stock is oversold traders make an entry long position (buy the stock) and exit the long position (sell the stock)

The strategy was tested for 4 etfs for uptrending, downtrending and non-trending scenarios. Bollinger band strategy performs poorly in an uptrending market. However, It performs better compared to buy and

hold strategy in non-trending and down-trending markets.

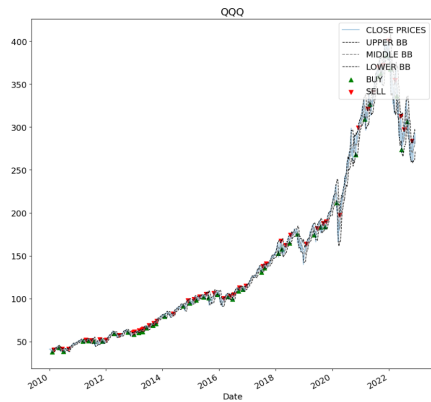


Fig.1 QQQ Uptrending

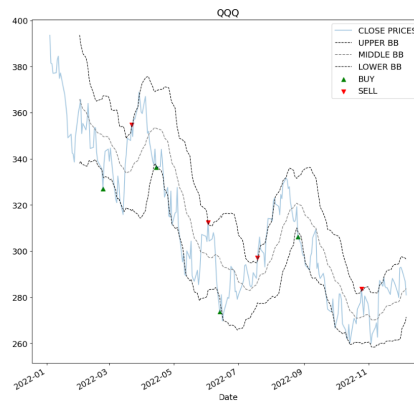


Fig.2 QQQ Downtrending

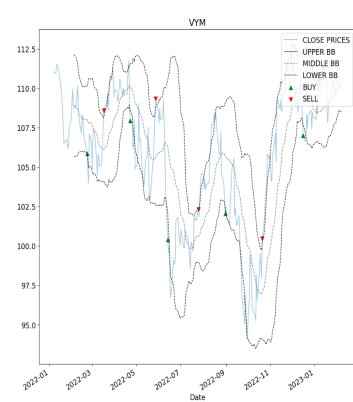


Fig.3 VYM No trend

Ticker	Period	One Time Investment Profit%	Bollinger Band Strategy Profit%
QQQ(Uptrend)	2010-01-04 TO 2022-12-07	584.23%	2.23%
QQQ(Downtrend)	2022-01-04 TO 2022-12-07	-28.6%	0.37%
VYM(No trend)	2022-01-07 TO 2023-02-08	-0.33%	1.58%

Fig.4 Table showing profit and loss with Bollinger Band Strategy in different scenarios

Relative Strength Index (RSI) Strategy

Relative Strength Index[3] measures the momentum and is an oscillator which moves between 0 and 100. Generally, when the RSI drops below 30, it indicates the market is oversold and when it goes above 70 it indicates overbought level. When a stock is oversold, traders make an entry long position (buy the stock) and exit the long position (sell the stock). RSI calculation is done in two steps. In Step1, average gain or average loss is calculated. For a predetermined period (Generally 14 days), the days with gain only are used for calculation for average gain and days with losses only are used for calculating average loss.

Step 1

$$RSI_{\text{step one}} = 100 - \left[\frac{100}{1 + \frac{\text{Average gain}}{\text{Average loss}}} \right]$$

After the initial 14 days RSI is calculated as in step 1, later periods RSI is calculated in Step 2

Step 2

$$RSI_{\text{step two}} = 100 - \left[\frac{100}{1 + \frac{(\text{Previous Average Gain} \times 13) + \text{Current Gain}}{((\text{Previous Average Loss} \times 13) + \text{Current Loss})}} \right]$$

The strategy was tested as before, and similar to bollinger band strategy did not perform well on uptrending scenarios but outperformed in downtrend and no trend scenario. The threshold limits for RSI and calculation period can be modified as needed.

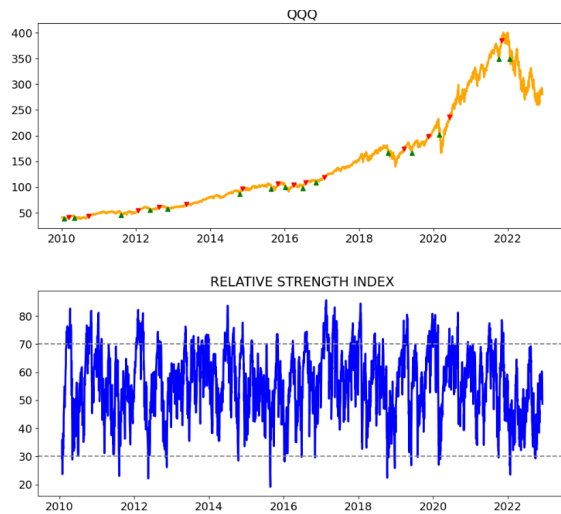


Fig.5 QQQ Uptrending (RSI)

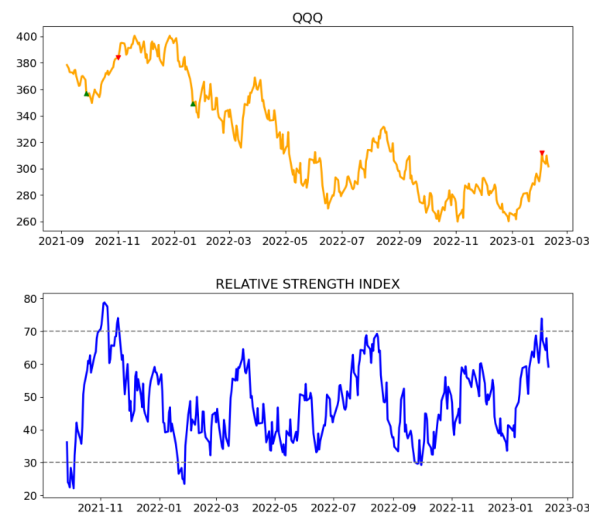


Fig.6 QQQ Downtrending (RSI)

Ticker	Period	One Time Investment Profit%	RSI Strategy Profit%
QQQ(Uptrend)	2010-01-04 TO 2022-12-07	584.23%	6.0%
QQQ(Downtrend	2021-09-04 TO 2023-02-10	-15.44%	1.43%
VYM(No trend)	2021-09-04 TO 2023-02-10	4.69%	8.9 %

Fig. 5 Table showing profit and loss with RSI strategy

Moving Average Convergence & Divergence (MACD) Strategy

MACD[4] is a momentum-based trend indicator and is based on the convergence and divergence of 12 and 26 period exponential moving average(EMA) and consists of following components:

- Fast Line : This line represents short term exponential moving average, which is 12 period EMA
- Slow Line : This line represents long term exponential moving average, which is 26 period EMA
- MACD Line : 12 period EMA- 26 period EMA
- Signal Line: This line is the 9 period EMA of MACD
- MACD histogram : MACD Line -Signal Line
- Zero Line : It's the base line with value of zero, and on this line MACD histogram is plotted. Generally, positive values of MACD bars are shown as Green and negative values are shown as Red

In general, when the MACD line turns positive coming from below zero, it represents a bullish signal also

called golden cross and when MACD Line turn negative coming from above zero, it represents a bearish signal also called death cross.

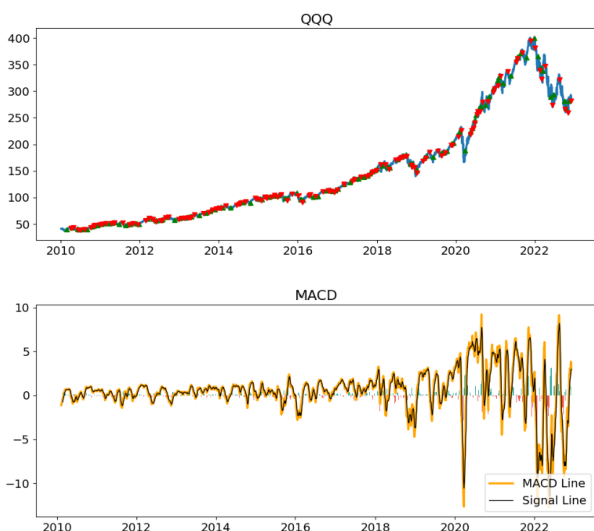


Fig. 7 QQQ Uptrending(MACD)

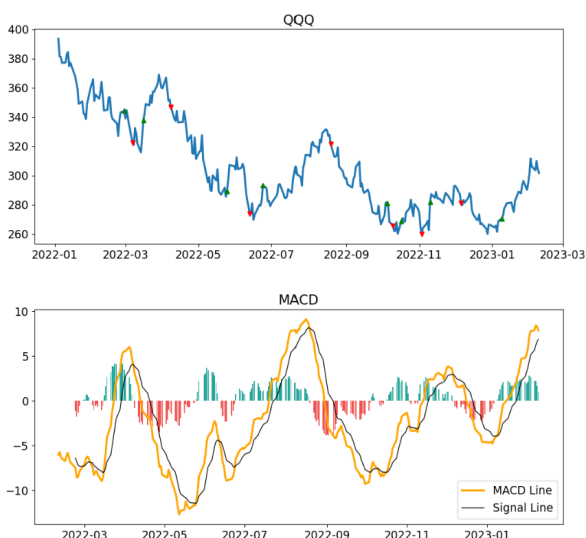


Fig. 8 QQQ Downtrending(MACD)

Ticker	Period	One Time Investment Profit%	RSI Strategy Profit%
QQQ(Uptrend)	2010-01-04 TO 2022-12-07	584.23%	0.73%
QQQ(Downtrend	2022-01-04 TO 2023-02-10	-23.35%	-1.22%
VYM(No trend)	2022-01-04 TO 2023-02-10	-1.53%	0.01 %

Fig.9 Table showing profit and loss with MACD Strategy

Long Short Term Memory (LSTM) Neural Network

LSTM is a type of artificial recurrent neural network, which is used in the field of deep learning. It has feed back connections and can process sequences of data points. LSTM neural network can be used for short term movement of stocks over a long period of time. In this project, "Adam" optimizer and "mean squared error" are used as loss functions. ETF "QQQ" is tested for the period 2010-07-01 to 2022-11-30. Model was trained on 3060 days of data for training and 60 days of data for predictions of stock prices. Loss function reduced from 0.0203 in the first epoch to 0.000705 after the last epoch.

```
model.summary()
```

Model: "sequential"

Layer (type)	Output Shape	Param #
lstm (LSTM)	(None, 5, 50)	10400
dropout (Dropout)	(None, 5, 50)	0
lstm_1 (LSTM)	(None, 5, 50)	20200
dropout_1 (Dropout)	(None, 5, 50)	0
lstm_2 (LSTM)	(None, 50)	20200
dropout_2 (Dropout)	(None, 50)	0
dense (Dense)	(None, 1)	51

=====
Total params: 50,851
Trainable params: 50,851
Non-trainable params: 0

Fig 10 Model Summary

Train on 3060 samples

Epoch 1/50
3060/3060 [=====] - 2s 605us/sample - loss: 0.0221
Epoch 2/50
3060/3060 [=====] - 1s 351us/sample - loss: 0.0020
Epoch 3/50
3060/3060 [=====] - 1s 390us/sample - loss: 0.0018
Epoch 4/50
3060/3060 [=====] - 1s 263us/sample - loss: 0.0016
Epoch 5/50
3060/3060 [=====] - 1s 224us/sample - loss: 0.0015
Epoch 6/50

Fig11 Ist 6 epoch performance

Epoch 45/50
3060/3060 [=====] - 1s 418us/sample - loss: 7.2356e-04
Epoch 46/50
3060/3060 [=====] - 1s 483us/sample - loss: 8.4449e-04
Epoch 47/50
3060/3060 [=====] - 2s 522us/sample - loss: 6.9153e-04
Epoch 48/50
3060/3060 [=====] - 2s 507us/sample - loss: 6.8070e-04
Epoch 49/50
3060/3060 [=====] - 2s 498us/sample - loss: 7.4703e-04
Epoch 50/50
3060/3060 [=====] - 2s 500us/sample - loss: 7.0570e-04

Fig 12 Last 6 Epochs

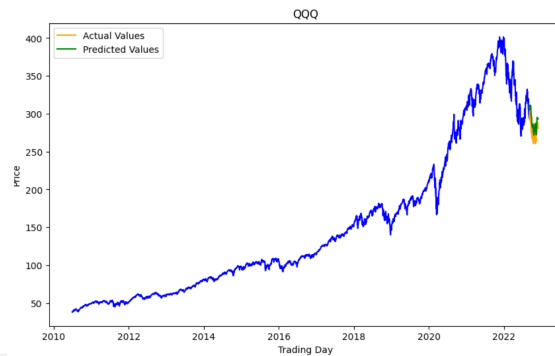


Fig 13 Plot Comparison with actual and forecast

Conclusion

- Bollinger Band Strategy, Relative Strength Strategy and MACD did not perform well on the up trending market, compared to buy and hold strategy.
- Bollinger Band Strategy, Relative Strength Strategy and MACD outperformed the buy and hold strategy for both down trending and non-trending stocks
- Relative Strength Index Strategy performed far better than Bollinger Band and MACD strategy for all three price trend scenarios
- LSTM neural network did show good predictability for exchange traded fund "QQQ"

Future Work

In the future reinforcement learning can be explored for developing trading strategies. Additionally, existing strategies can be modified to improve the performance of the trading strategy as below:

- Bollinger Band Strategy can be modified using different periods for calculating moving average, along with modifying the upper and lower limit of bollinger bands by altering the no. of standard deviations from the mean. In the future, variables can be modified to see if they perform better.
- RSI strategy can be modified using different periods for calculating RSI, along with modifying the upper and lower limit of relative strength threshold. decisions.
- MACD strategy can be optimized using different periods for calculating slow, fast and signal line periods.

References

1. <https://www.statista.com/statistics/270034/percentage-of-us-adults-to-have-money-invested-in-the-stock-market/>
2. <https://www.investopedia.com/trading/using-bollinger-bands-to-gauge-trends/>
3. <https://www.fidelity.com/learning-center/trading-investing/technical-analysis/technical-indicator-guide/RSI>
4. <https://www.investopedia.com/terms/m/macd.asp>