

MC3-Project-3

GT Account Name: xtao41

Part 1: Technical Indicators

1) MACD

MACD (Moving Average Convergence/Divergence Oscillator) is a momentum oscillator based on difference between two EMAs - a longer moving average (usually 26 days) and a shorter moving average (usually 12 days). EMA (Exponential Moving Average) is calculated by applying weighting multiplier to the most recent prices (depending on window size), which reduces the lag of moving average calculation.

STEPS of calculating EMA of n day period

- a. Calculate SMA= $n \text{ day sum} / n$
- b. Calculate multiplier= $2 / (n + 1)$
- c. Calculate EMA= {Closing Price – EMA (previous day)} * multiplier + EMA (previous day)

MACD consists of three major elements: MACD line (difference of two EMAs), signal line (9-day EMA of MACD line to identify turns) and MACD Histograms (difference between MACD and signal line)

STEPS of calculating MACD (12, 26, 9)

12, 26, 9 are the typical settings used in MACD for fast EMA, short EMA and signal line period

- a. Calculate MACD Line= EMA_Fast (12-day EMA)- EMA_Slow (26 day EMA)
- b. Calculate Signal Line= 9 -day EMA of MACD Line
- c. Calculate MACD Histogram= MACD Line – Signal Line

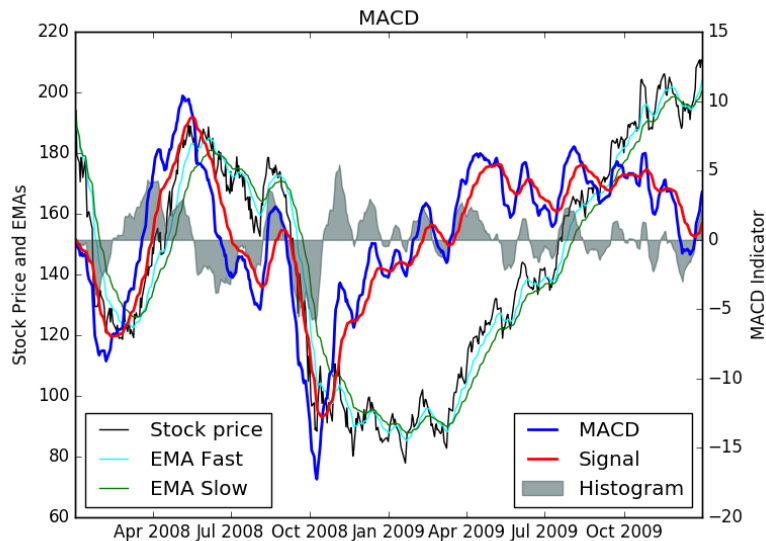
MACD provides both trend following and momentum. When the MACD trend towards Zero line (center line), the moving averages converge, while as the MACD moves away from Zero line, the moving averages diverge. Positive (negative) MACD indicates fast EMA is above (below) slow EMA. Positive/negative values increase as fast EMA diverges further from slow EMA- upside (downside) momentum is increasing. The trading opportunities can be implied by signal line crossovers, centerline crossovers and divergences.

Interpretation

- a. MACD Line crosses centerline: A bullish centerline crossover occurs when the MACD Line moves above the zero line to turn positive. A bearish centerline crossover occurs when the MACD moves below zero line and turn negative.
- b. MACD Line crosses Signal line: A bullish crossover occurs when the MACD turns up and crosses above the signal line. A bearish crossover occurs when the MACD turns down and crosses below the signal line.
- c. Divergences: A bullish divergence forms when a security (stock or bond price) records a lower low and the MACD forms a higher low. A bearish divergence forms when a security records a higher high and the MACD line forms a lower high.

Chart:

showing stock price and 12-day and 26 day EMAs on the left axis, and MACD line, Signal Line (9-day) and MACD Histogram



2) RSI

RSI (Relative Strength Index) is a momentum indicator and measures the how fast a stock is moving in its current direction. RSI is calculated from RS (relative strength), which is defined by average gain/average loss in a given lookback window. Both gain and loss are expressed as positive values.

STEPS of calculating RSI with n day lookback (default=14)

- Average Gain = $[(\text{previous Average Gain}) \times (n-1) + \text{current Gain}] / n$.
- Average Loss = $[(\text{previous Average Loss}) \times (n-1) + \text{current Loss}] / n$.
- $RS = \text{Average Gain} / \text{Average Loss}$
- $RSI = 100 - 100 / (1 + RS)$

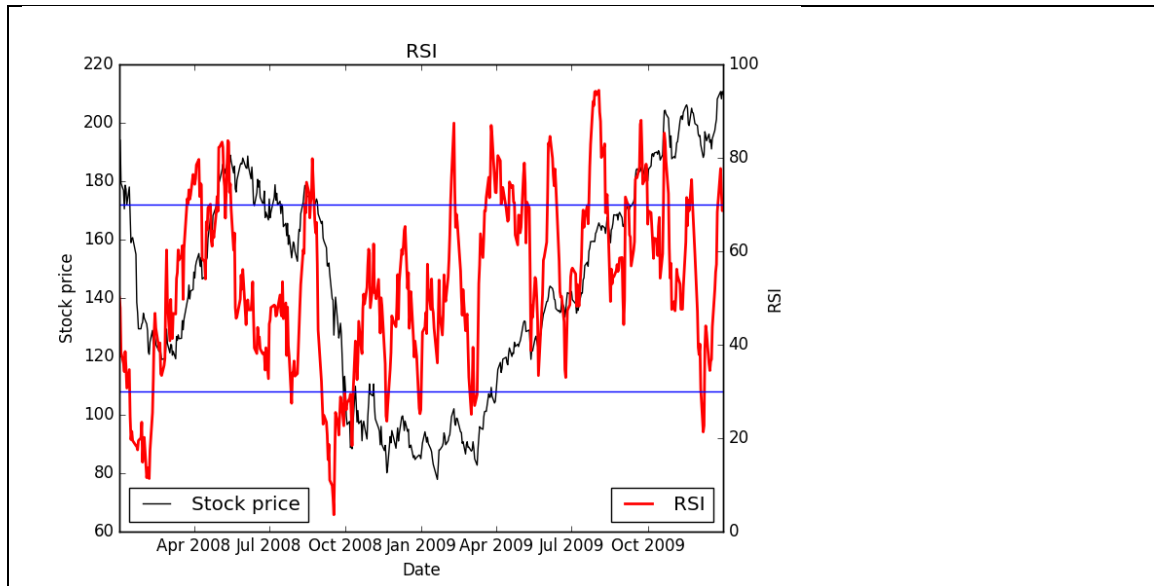
RSI oscillates between 0 and 100, and is traditionally considered overbought when above 70 and oversold when below 30. A low number reading of RSI (typical <30) suggests a stock is oversold and ready to bounce. On the other hand a large reading (typical >70) suggest an overbought.

Interpretation

- Buy signal: when RSI goes from below 30 to above 30. (oversold and is now heading upward)
- Sell signal: when RSI goes from above 70 to below 70. (overbought and is now selling)
- Divergence: A bullish divergence occurs when the underlying security makes a lower low and RSI forms a higher low. A bearish divergence forms when the security records a higher high and RSI forms a lower high.

Chart:

Showing stock price and RSI. Oversold limit=30, and overbought limit=70.



3) Bollinger Bands (BB)

BBs volatility indicators which shows the upper and lower limits of price movement based on standard deviation of the prices. The outer bands are set 2 stdev above and below the middle band. The calculations are as follows

STEPS of calculating BB with n day lookback

- Middle band=SMA (simple moving average) n day sum/n
- Upper band= SMA+2*stdev
- Lower band=SMA-2*stdev

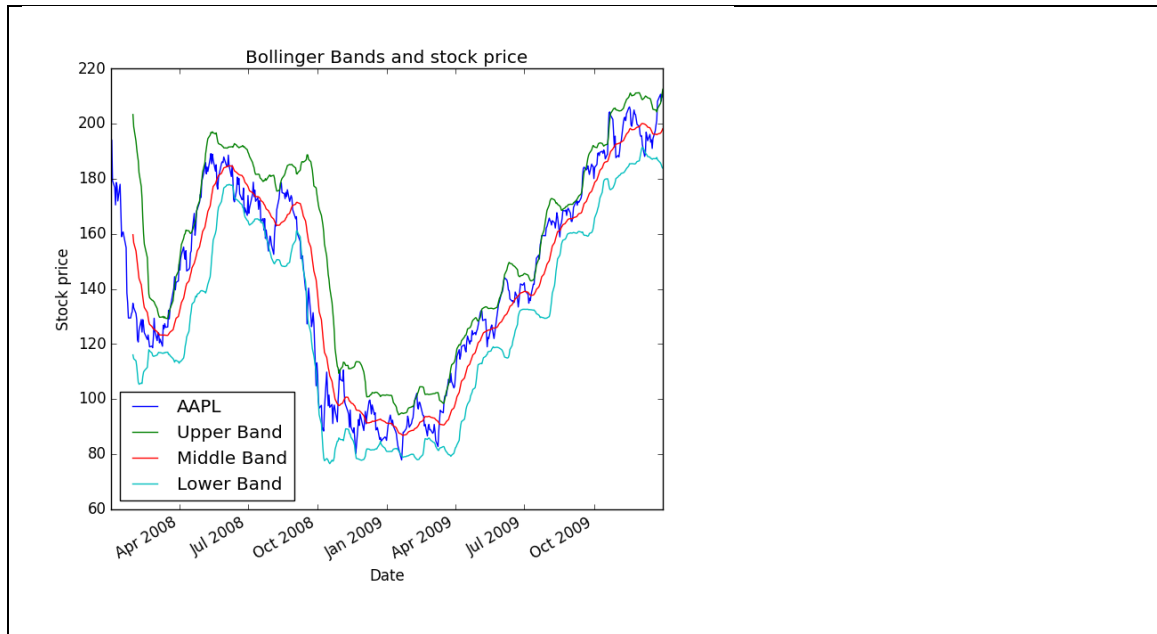
Since BB reflect direction of SMA and volatility, it can be used to determine if prices are relatively high or low. Technically, prices are relatively high when above the upper band and relatively low when below the lower band. However, these are not necessarily indicating a buy/sell signal.

Interpretations

- Price Crossovers: moving from above upper band toward middle band might indicate a Sell signal; moving from below lower band toward middle band might indicate a Buy signal.
- W-Bottoms (Buy signal): forms in a downtrend involve two reaction lows (second low is lower than the first, but holds above the lower band)
- M-Tops (Sell signal): opposite to W-Bottoms

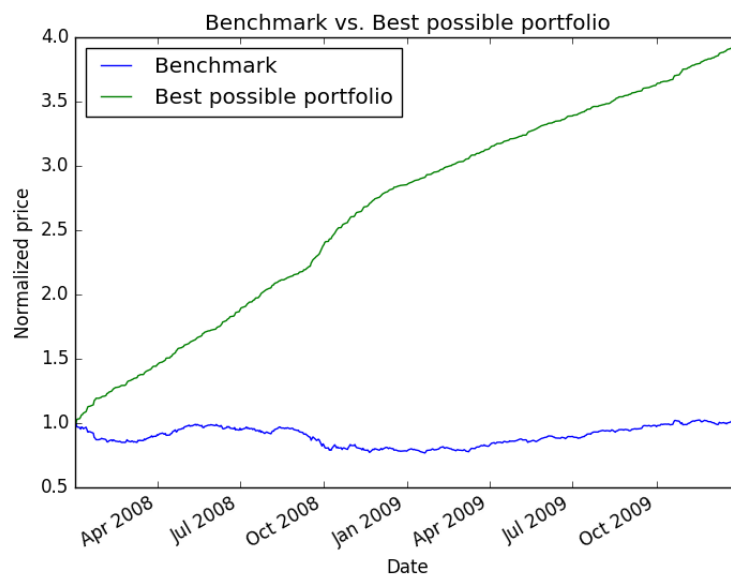
Chart:

Showing stock prices and bollinger bands



Part 2: Best Possible strategy

Since the allowable positions are 200 long, 200 short or 0 shares, the best possible strategy I think is that whenever next day stock price is higher than current day, we long 200 shares; whenever next day price is lower than current day, we short 200 shares; If stays the same, then hold 0 shares. Daily returns would be always ≥ 0 and is $200 \times \text{price change}$. Figure below shows the Benchmark vs. Best possible portfolio.



The statistics are calculated the results are as below.

	Cumulative return	Stdev of daily returns	Mean of daily returns
Benchmark	0.03164	0.0087	0.0001
Best possible portfolio	2.94948	0.0032	0.00273

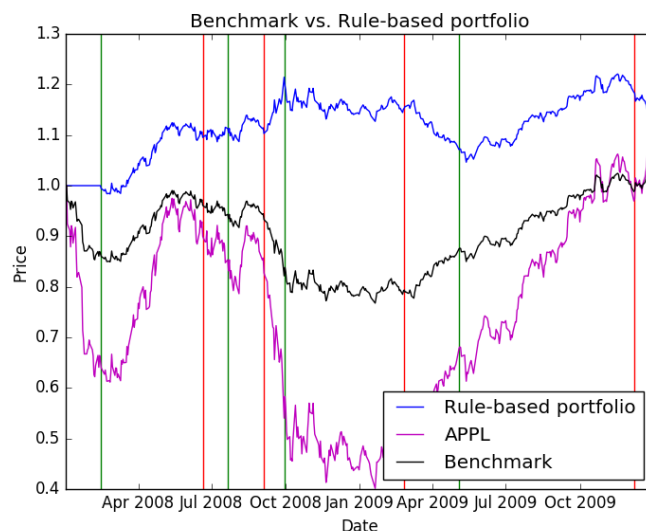
Part 3: Manual Rule-Based Trader

Trading Rules

- 1) Short:
 - a. whenever there is a short signal, short the max possible amount, i.e. 200 + current holdings (can be positive (in long position) or negative (in short position)), with the restriction that there must not exist trading within past 21 trading days, and there is still room for increase of holdings.
 - b. Short signals: i) when the MACD line moves below zero line and turn negative OR ii) the MACD turns down and crosses below the signal line OR iii) stock prices moves from above upper band toward middle band AND iv) RSI above 70
- 2) Long
 - a. whenever there is a long signal, long the max possible amount, i.e. 200 - current holdings (can be positive (in long position) or negative (in short position)), with the restriction that there must not exist trading within past 21 trading days, and there is still room for increase of holdings.
 - b. Long signals: i) when the MACD line moves above zero line and turn positive OR ii) the MACD turns up and crosses above the signal line OR iii) stock price moves from below lower band toward middle band

The Rule design is a combination of three indicators. Actually, it is found that the MACD indicator is the most useful in this particular case.

Chart showing the results



Part 4. ML Trader

In this part, I choose the three indicators described in the above paragraphs to train and test ML trader. Specifically, the following indicators were calculated from MACD, Bollinger Band, and RSI.

1. MACD Line value: calculated by $\text{EMA_Fast (12-day EMA)} - \text{EMA_Slow (26 day EMA)}$
2. MACD signal line value: 9 -day EMA of MACD Line
3. SAM: 21 days of SAM
4. $\text{BB_value} = (\text{price}[t] - \text{SMA}[t]) / (\text{stddev}[t])$
5. RSI value.

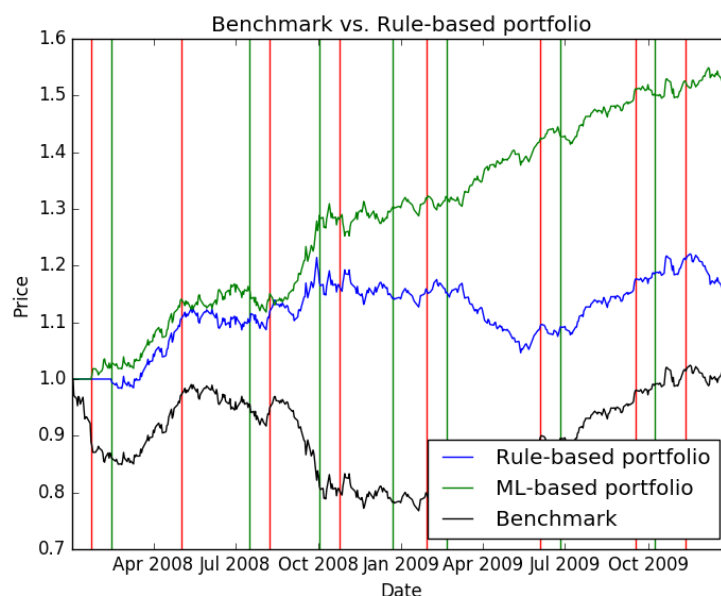
The training parameter are tweaked so that they give the optimal return, as below:

- $\text{YSELL} = -0.05$
- $\text{YBUY} = 0.05$
- $\text{Leaf_size} = 5$
- $\text{Bagging} = 10 \text{ bags}$

The X values of the training set are the indicators described above, but standardized to help all indicators are considered with equal importance for the learner. The formula below describes the standardization process.

$$\text{Standardized value} = (\text{values} - \text{mean}) / \text{values.std}()$$

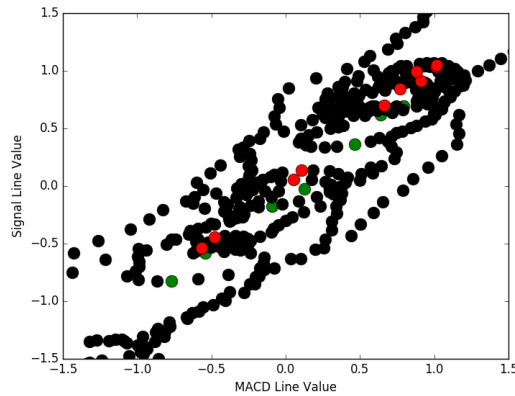
The classification learner uses 1 for long, 0 do nothing and -1 for short to train and test, and then build the orders book. I use 100% of the in-sample data for training and 100% of it for learning. The figure below shows the result of trading strategy developed by ML-based learner and the portfolio profile. ML-based portfolio substantially out-performances both of benchmark and Rule-based portfolio.



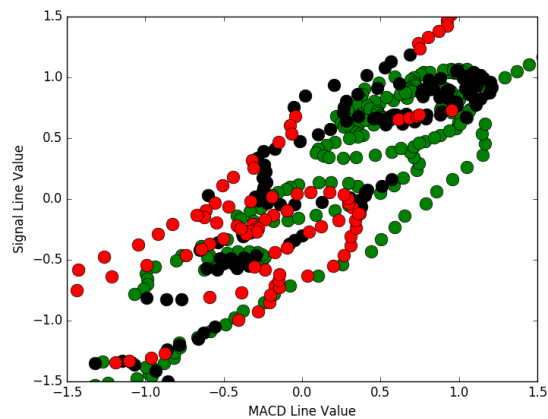
Part 5. ML Trader

I choose two indicators, one is MACD Line value, the other is Signal Line value, with the definition described above. The charts below shows scatter plots of these two indicators and green circles are “long”, Red circles are “short”, and Black circles are neither of these two conditions. The 21-day restriction is not applied here.

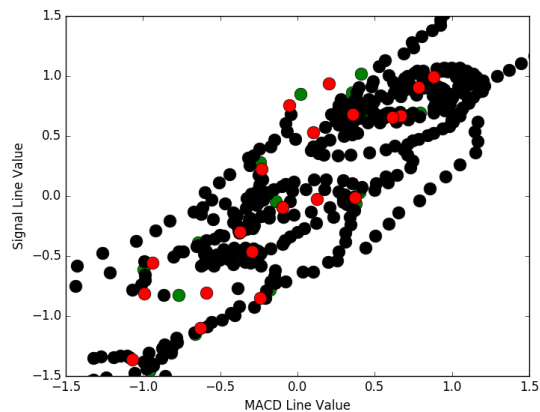
a) Rule-Based Strategy



b) Training data for ML strategy

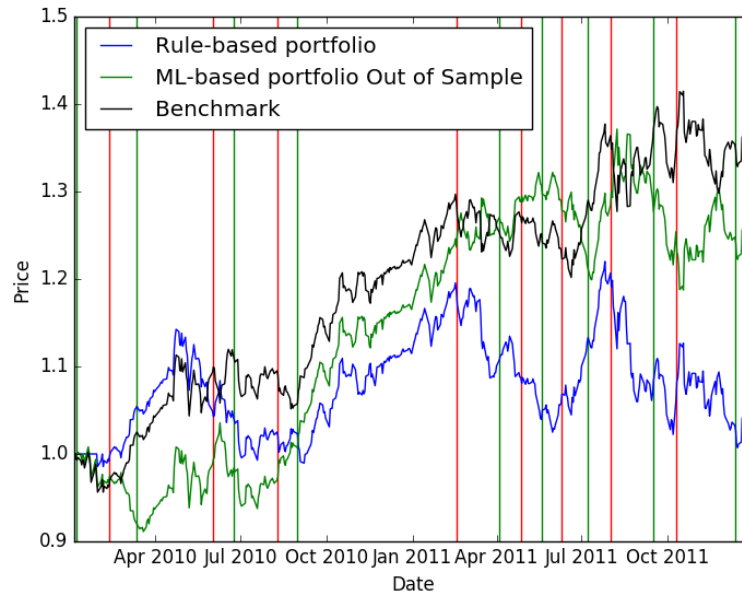


c) Response of learner when queried with same data after training.



Part 6: Comparative Analysis

The plot below shows out of sample test results of benchmark, Rule-based manual strategy and ML strategy performance.



The summary of all performances (regarding cumulative return) of in sample and out-of-sample are summarized as below.

	Benchmark	Manual Strategy	ML Strategy
In-sample	0.035	0.14	0.51
Out-of-sample	0.38	0.06	0.28

Cumulative Return is one of the best factors for comparing these strategies. Another good factor is sharp ratio, which reflects the risk of the trading. From my experience, the MACD values are one of the best indicators of stock price trending and oscillations.

The performance of out-of-sample is worse than in sample because in sample train and test using the same dataset, thus the classification pattern is more likely to be the same using decision tree classification learner. For a new dataset, it is more unpredictable and often not fit well for the trained pattern.

For both in sample and out-of-sample, the ML strategies all out-perform the Manual strategies. I think the reason is the manual method is based on certain rules based on our experience and are not working all the time. Also, the rules are limited. There are often hidden patterns that can only be realized by machine learning method.

Manual strategy is more susceptible to the same underlying flaw, because the rules might inherently have flaws and incompleteness. Under different trending pattern, it might be useless or even give the opposite signals. On the other hand, the ML method does not need

to understand or find the rules. It is completely data-orientated and thus less susceptible to the flaws.