# ***Machine Learning for Trading***
# ***Report on Manual Strategy***

# ***Havish Chennamraj***
# ***903201642***

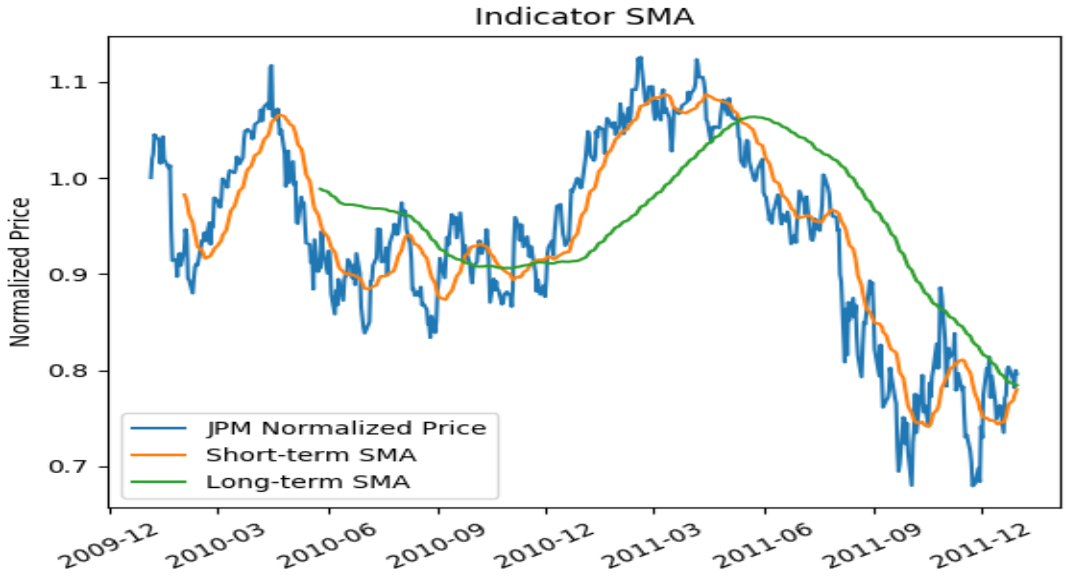

## ***Technical Indicators***

### *1.) Simple Moving Average (window = 20)*

Simple moving average is computed by taking the unweighted average of adjusted close price of a stock for past 'n' days, where 'n' is usually referred to as the 'Window' of the Simple Moving Average. This means that stock value at every time period has equal importance and is weighted equally. As each period ends, the oldest stock value is dropped and the lasted one is added to the computation.

$$SMA(t,n) = \frac{1}{n}\sum_{i=0}^{n-1} Price\,(t-i)$$

I've used a Trading Strategy called "Golden Cross" which uses two moving averages: A long and a short one. When the short-term Moving Average crosses above the Long-term Moving Average, it is a Buy Signal as it indicates that the trend is moving up and vice versa for a Sell Signal.




The graph above depicts how Short-term SMA & Long-Term SMA behave with respect to each other, as well as with JPM's normalized price for a given time duration. I've chosen a window size of 20 for Short-term SMA & 100 for Long-term SMA.

## *2.) **Bollinger Bands***

Bollinger Bands are lines plotted in and around the price structure of a stock to form an envelope. It is usually action of prices near the edge of this envelope that we are interested in.
Bollinger Bands are typically charted over 20 "periods". The middle band of a simple moving average over the before mentioned 20 "periods". The upper and lower bands are set above and below the 20-period SMA by a certain number of standard deviations of the price over the same period thus incorporating volatility.

Bollinger Bands are one of the most powerful concepts available for technical analysis. It is important to note that Bollinger Bands by themselves do not give absolute buy or sell signals, however answers questions of whether prices are high or low on a relative basis. Further, the width of the band can be an indicator of its volatility; narrower bands indicate less volatility while wider ones indicate higher volatility.

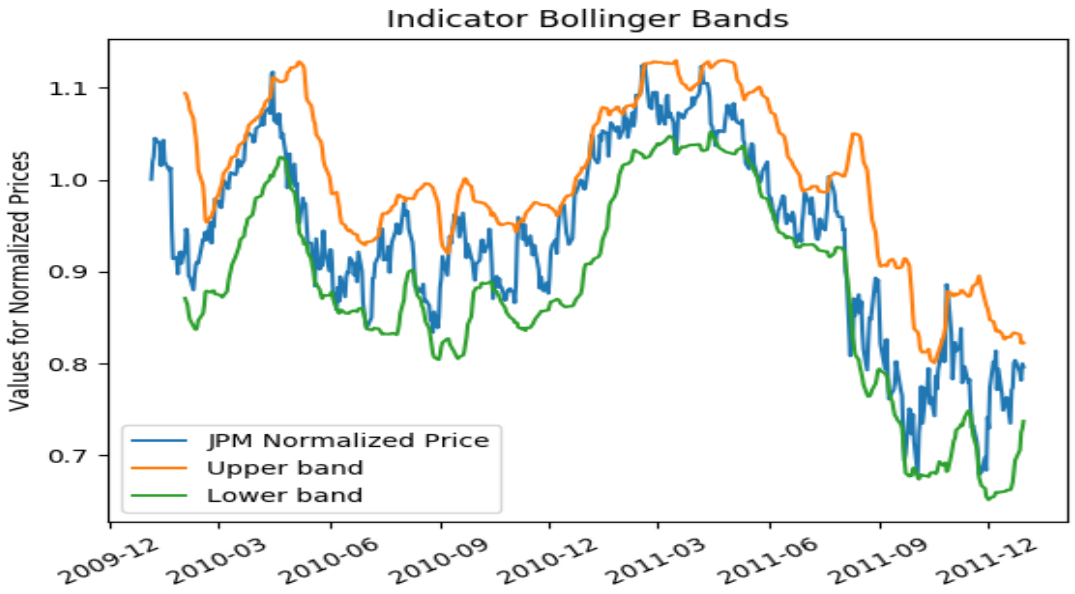
**Bollinger Bands Calculation:**

*MiddleBand = 20-period SMA of Prices => MiddleBand(t) =* $\frac{1}{20}\sum_{i=0}^{19} \boldsymbol{Price}\,(\boldsymbol{t}-\boldsymbol{i})$

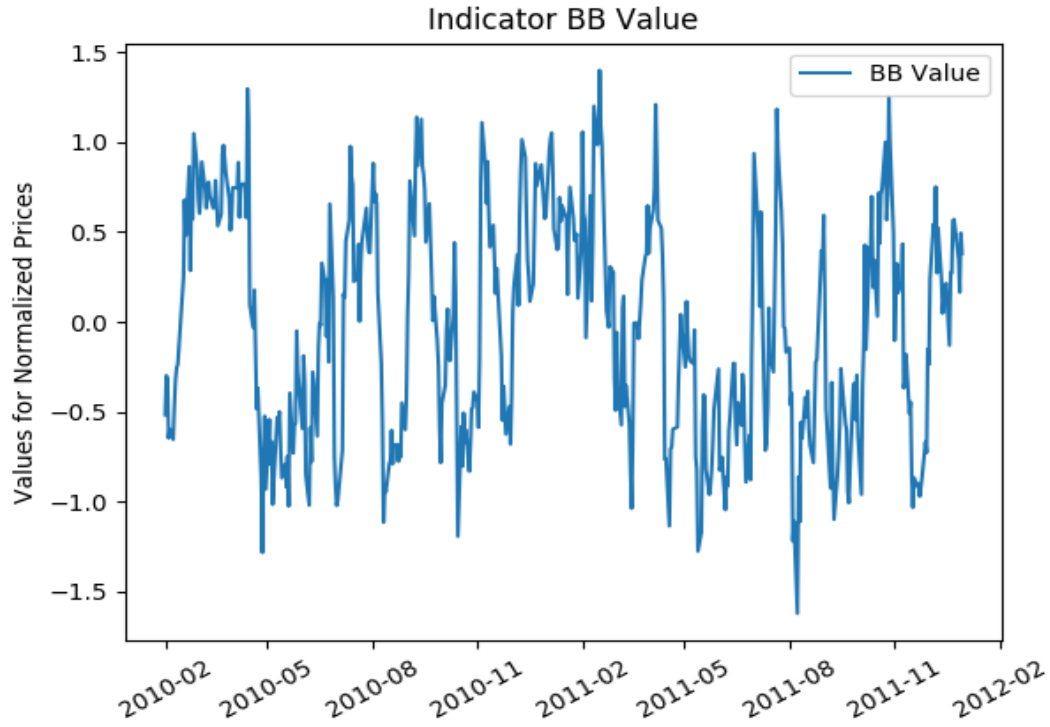
*UpperBand = MiddleBand + 2 * 20-period StandardDev => UpperBand(t) = MiddleBand(t) +* $2 * \sigma(Prices)_{window=20}$

*LowerBand = MiddleBand - 2 * 20-period StandardDev => LowerBand(t) = MiddleBand(t) –* $2 * \sigma(Prices)_{window=20}$

*BB_Value = (Price – MiddleBand)/(2* 20-period StandDev)*



The above graph depicts the Bollinger Bands (Upper and Lower) for JPM and how they change in comparison to the Normalized Price of 'JPM'.




The above graph plots the BB value which shows if JPM's price is relatively high or relatively low. My strategy is to use 'BB_Value' in combination with other indicators to enter/exit positions. If 'BB_Value' is tending towards 1, I use it as a Sell signal and if 'BB_Value' is tending towards -1 I use it as a Buy signal.

### 3.) Moving Average Convergence Divergence (MACD)

MACD is a very simple and effective momentum indicator, which uses two exponential moving averages to measure the momentum in the price. The idea behind calculating this momentum indicator is to measure short-term momentum compared to a longer term momentum to help signal the current directon of momentum.

**MACD Calculation:**

*MACD = (12-period_EMA – 26-period_EMA)*
*Where EMA stands for Exponential Moving Average.*

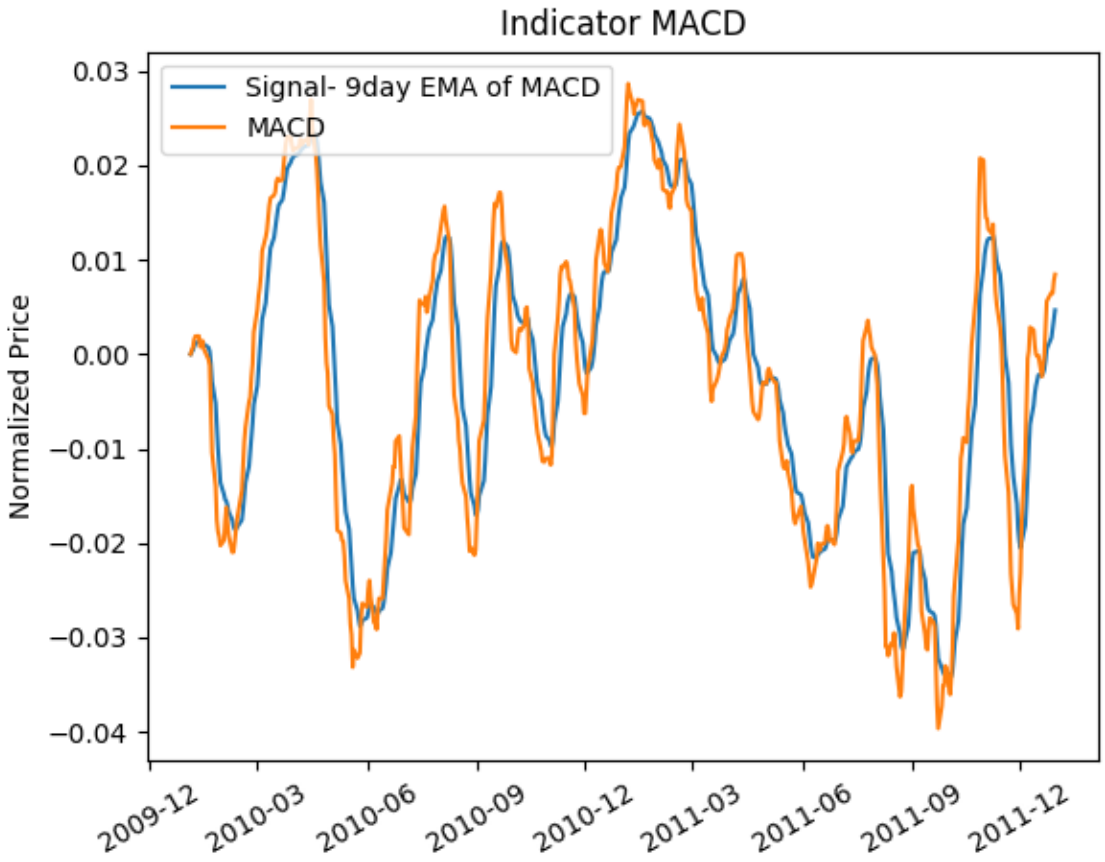
Exponential Moving Average is very similar to Simple Moving Average. It only differs in the fact that unlike Simple Moving Average which is un-weighted average of Stock Prices for the past 'X' periods, EMA is a weighted average that places a greater weight and significane on the most recent

stock prices. It is also noteworthy to point out that because of this, EMA reacts more significantly to recent price changes compared to SMA.

**Strategy: Zero-Line Crossovers & Signal line crossovers**

Moves across the zero in the MACD line indicate the crossing between 12-period EMA and 26-period EMA. Thus when the MACD line is crossing zero with a positive slope, it's a signal for a new uptrend and the line crossing zero with a negative slope signals a downtrend. The strategy is to hold long trades until the MACD line cross back below the zero line and hold short trades until the MACD line crosses back above zero. During choppy conditions, this strategy results in loss.

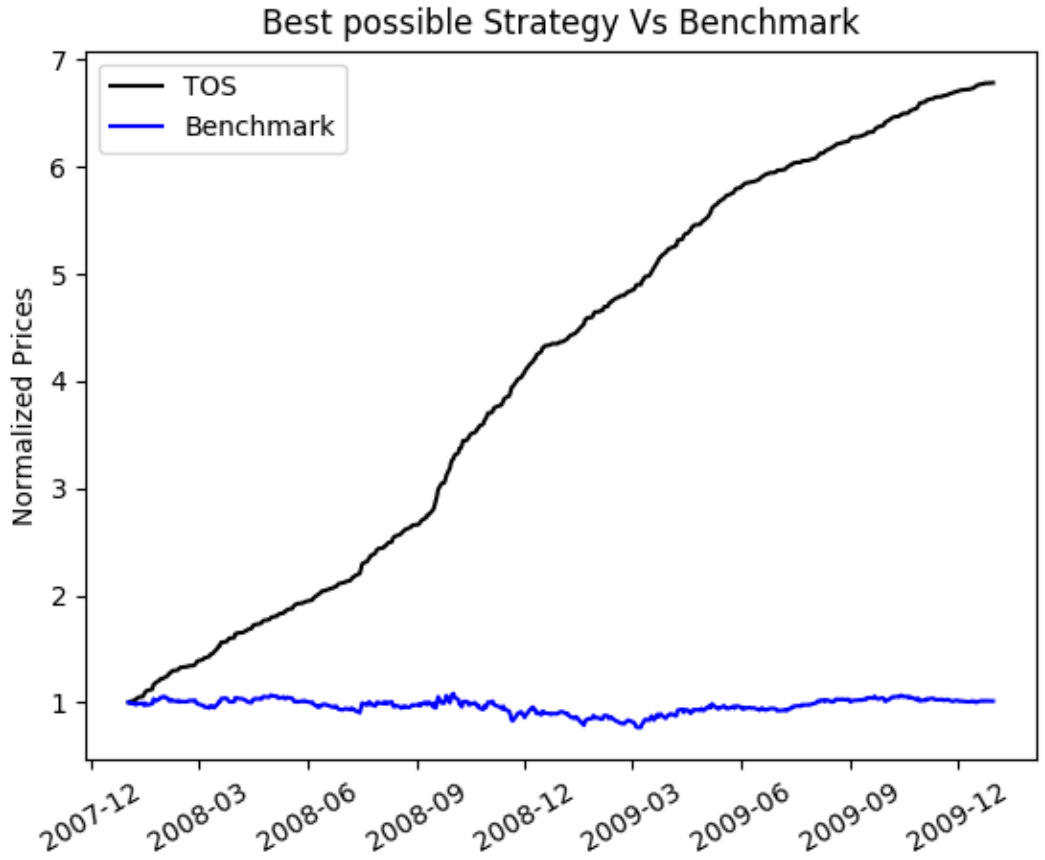
Thus, I chose to opt for the **Signal Line Crossove**r where in addition to the MACD line, a Signal line which is equal to the 9-day EMA of MACD line is used. In this strategy, a sell signal occurs when MACD line crosses below the Signal line and a Buy signal occurs when MACD line crosses above the Signal line. Below graph shows both the Signal line and MACD line using which I implement as my 3rd indicator employing the 'Signal Line Crossover' strategy.



# Best Possible Strategy

In the best possible strategy, it is assumed that we can see into the future. However, we are constrained by the portfolio size (1000 shares) and order size (can only trade in 3 states: 0, +1000 and -1000). The strategy is to look one day forward to decide on the trade based on the price of stock on the current day and its very next day. The strategy can be summarized as follows:

1.) If the price of Stock on day 'i' is greater than the price of the stock on day 'i+1', short the stock. It is important to note that if I'm already a short position on 1000 shares (-1000), I will not perform any trade on that day.
2.) If the price of Stock on day 'i' is lesser than the price of the stock on the day 'i+1', buy the stock (enter a long position). It is important to note that if I've already bought 1000 shares, I would not buy any more shares. In other words, the constraints are honored.
3.) If the price of Stock on day 'i' is equal to the price of the stock on the day 'i+1', do nothing.



The above graph plots the portfolio value (normalized) for JPM Stock when both Benchmark and Best Possible Strategy is employed. Below are few of the metrics which provide better insight into how much better the 'Best Possible Strategy' is than the 'Benchmark – Buy & Hold Strategy'.

| | Best Possible Strategy | Benchmark |
|---|---|---|
| Cumulative Return | 5.7861000000000002 | 0. 012299999999999978 |
| Mean of Daily Returns | 0.003809228158479885 | 0.0001677541326895741 1 |
| Stddev of Daily Returns | 0.0045464828147862889 | 0.016987490130867061 |

# ***Manual Strategy***

The manual strategy is a rule based one, with the rules being defined using the 3 indicators mentioned in the **'Technical Indicators'** section. Below is a pseudo code explaining how the indicators are combined to create an overall BUY/SELL Signal.

```
If( buySmaSignal(day) or buyBollingerSignal(day) or buyMACDSignal(day) )
        // Signal to buy
else if( sellSmaSignal(day) or sellBollingerSignal(day) or sellMACDSignal(day) )
        // Signal to sell
else if( SMA[day] > 1 and SMA[day-1] <= 1 and 'Current Holdnigs' == 1000)
        // Close the current Long position and enter Short position
else if( SMA[day] < 1 and SMA[day-1] >= 1 and 'Current Holdings' == -1000)
        // Close the current Short position and enter Long Position
else
        // Do nothing
```

Following is how each indicator is used to decide to enter/exit positions

### ***Indicator – SMA***

The Subroutine to compute Simple Moving Average in 'Indicators.py' returns two Data frames 'smallSMA' & 'largeSMA'. 'smallSMA' is computed for a 20-day time period, while the 'largeSMA is computed for a 100-day time period. Both the data frames are indexed over the give date range.

As mentioned in 'Simple Moving Average' under 'Technical Indicators' section, I've used the 'Golden Cross' strategy to indicate BUY/SELL Signals. Below is the pseudo code and the explanation of the same.

For a buy Signal, the 'Golden Cross' strategy dictates that when Short-term Moving Average (smallSMA) crosses above the Long-term Moving Average (largeSMA), it is a BUY Signal and when it cross below the Long-term Moving Average it is a SELL Signal.

```
buySmaSignal(day)
{
        buySignal = smallSMA[day] > largeSMA[day] and smallSMA[day-1] <=
largeSMA[day-1];
        return buySignal;
}

smallSmaSignal(day)
{
        sellSignal = smallSMA[day] < largeSMA[day] and smallSMA[day-1] >=
largeSMA[day-1];
        return buySignal;
}
```

***Indicator – Bollinger Value***

The Subroutine to compute Bollinger Bands and Bollinger Value in 'Indicators.py' returns three dataframes 'lowerBB', 'upperBB' & 'bollingerIndex', all of which are computed for the same 20-day time period. I've computed 'bollingerIndex' (which stores the BB_values for the given date range) using the following formula:

```
Middle_Band = 20-period SMA
BB_Value = (Price – Middle Band)/(2*20-period Stand_Dev)
```

BB_Value is a reflection how relatively high/low the current price. For instance, if BB_Value is tending towards 1, the current price is relatively high and indicates a SELL Signal and if BB_Value is tending towards -1, the current price is relatively low and indicates a BUY Signal. Below is the pseudo code reflecting the same strategy.

```
buyBollingerSignal(day){
        buySignal = bollingerIndex[day] > 0.9
        return buySignal
}

sellBollingerSignal(day){
        sellSignal = bollinderIndex[day] < 0.65
        return sellSignal
}
```

The benchmarks of '0.9' for BUY & '-0.65' for Sell were set for the '**BollingerIndex**' after trying out various combinations using trial and error and finally zeroing on the above mentioned

values because it yields the highest Cumulative return for the Training Set (JPM's stocks from January 1, 2008 to December 31 2009)

**Indicator – Moving Average Convergence Divergence (MACD)**

The subroutine to compute MACD in 'Indicators.py' returns two dataframes 'MACD' & 'ema9MACD', both of which are computed for 20-day period using the formula mentioned in '**Moving Average Convergence Divergence (MACD)**' under the 'Technical Indicators' section. I've used the 'Signal Line Crossover' strategy which uses the 'ema9MACD' which is basically the 9-day exponential moving average of 'MACD' as a signal Line to figure out BUY/SELL Signals. If the 'MACD' line crosses below the Signal Line, it is a SELL Signal and if it crosses above, it is a BUY Signal. Below is the pseudo code reflecting the same.

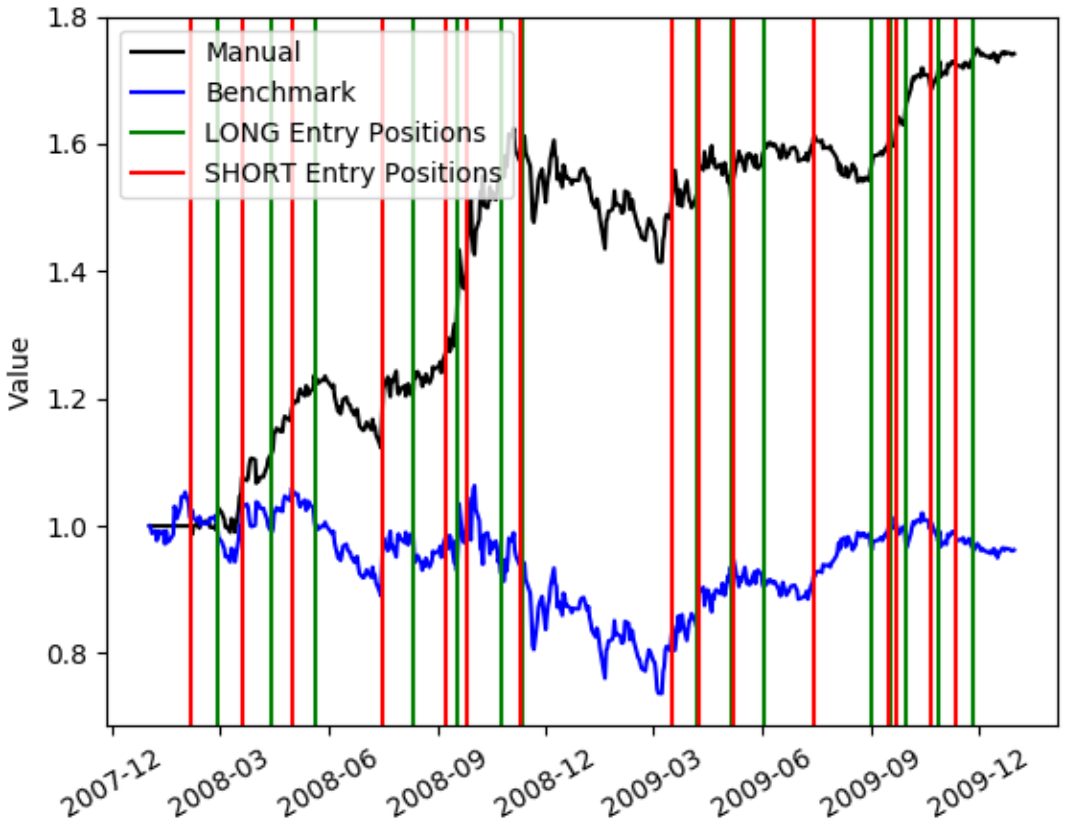
```
buyMACDSignal(day){
        buySignal = MACD[day] > ema9MACD[day] and MACD[day-1] <= ema9MACD[day-1]
        return buySignal
}

sellMACDSignal(day){
        sellSignal = MACD[day] < ema9MACD[day] and MACD[day-1] >= ema9MACD[day-1]
        return sellSignal
}
```

## *Why does the Manual Strategy work and its Shortcomings?*

I would basically enter a LONG position if any of my indicators emit a BUY Signal and enter a SHORT position if any of my indicators emit a SELL Signal. Because of this, I've noticed that I enter into SHORT/LONG positions more frequently and end up making greater profit (higher Cumulative Return) on the In-Sample period as opposed to combining the indicators with 'AND' (although I strategy I've employed might be slightly risker). The strategy would work because whenever I enter a BUY/SELL position, I do only after consulting with all of my indicators and at least one of them says 'YES'.

It is important to note here that Indicators I've used also come with some shortcomings. For instance, I've used only BB_Value in the **'Bollinger Bands'** indicator to decide on my positions. It might back fire when there is a very dominant uptrend/downtrend because the prices can "walk the band" during such strong up-trend or down-trend. This means that there are repeated instances of a price touching or breaking though the upper or lower band. That is why we might **not** to action when the price breaches either of the bands.

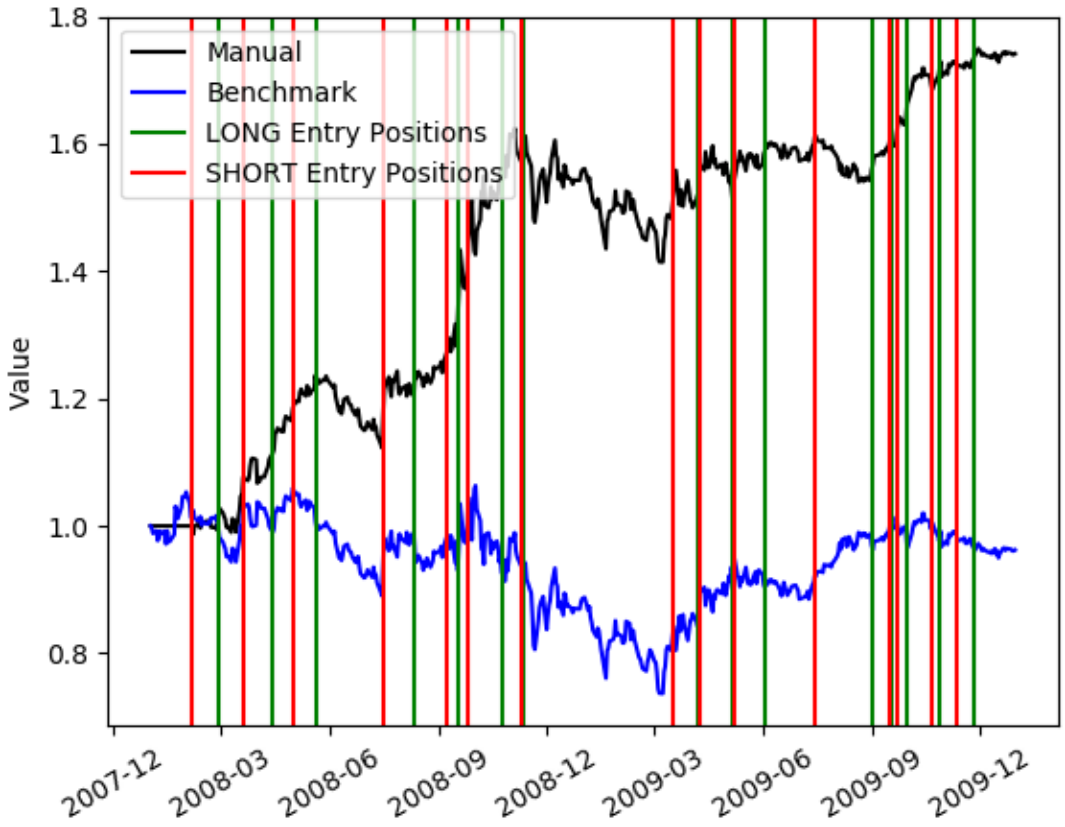
Indicators using Moving Averages also come with disadvantages too. For instance, moving averages are calculated from historical data, and nothing about the calculation is predictive in nature. In other words, if the prices behave radically different from the historical set which we used for our computation, a SMA indicator might do us more harm than good.




The above graph portrays the performance of the Manual Strategy I've explained above with the Benchmark-Buy & Hold Strategy for the 'In-Sample' period. The 'Green' vertical lines indicate the points in time where I've chosen to enter LONG position (BUY) & the 'Red' vertical lines indicate the points in time where I've chosen to enter the SHORT position (SELL). We can clearly see that the manual strategy performs better than the benchmark yielding a profit of roughly 80%.

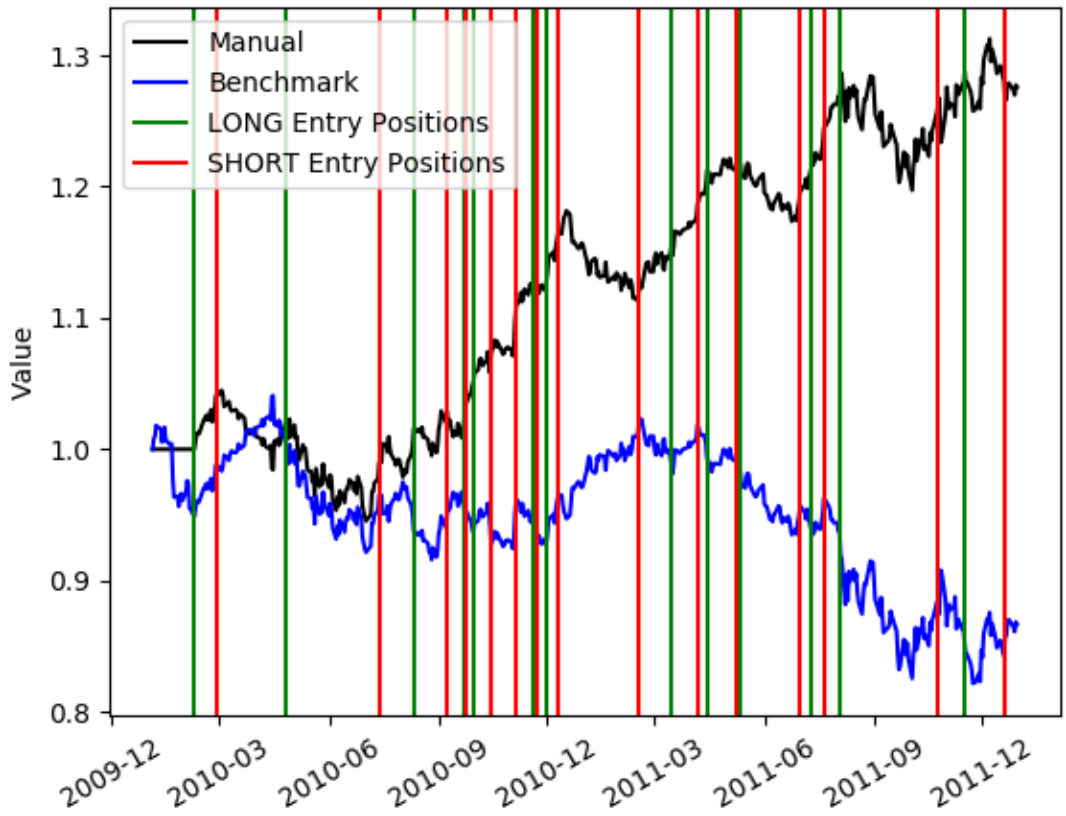
# ***Comparative Analysis***

***Comparing Manual Strategy and Benchmark for 'In sample' periods.***




| | Manual Strategy | Benchmark |
|---|---|---|
| Cumulative Return | 0.74140000000000006 | - 0.037779090938128257 |
| Mean of Daily Returns | 0.0011612014453154238 | 7.3957630440739174e-05 |
| Stddev of Daily Returns | 0.011185065022615777 | 0.017377360179350241 |

***Comparing Manual Strategy and Benchmark for 'Out sample' periods.***




| | Manual Strategy | Benchmark |
|---|---|---|
| Cumulative Return | 0.27570000000000006 | - 0.13318957945753407 |
| Mean of Daily Returns | 0.0005069308254232 4207 | -0.0002454800207874 8758 |
| Stddev of Daily Returns | 0.0068920125833814936 | 0.0087330710716286356 |

**Comparing Performance of Manual Strategy for both In-Sample and Out-Sample Periods**

| | Manual Strategy (In-Sample) | Manual Strategy (Out-Sample) |
|---|---|---|
| Cumulative Return | 0.74140000000000006 | 0.27570000000000006 |
| Mean of Daily Returns | 0.0011612014453154238 | 0.00050693082542324207 |
| Stddev of Daily Returns | 0.011185065022615777 | 0.0068920125833814936 |

***Why do these differences occur?***

One of the reasons I can think of to attribute to the difference in performance of Manual Strategy for In Sample and Out Sample periods is the difference in Volatility. The Market in the In-Sample period is more volatile enabling higher opportunities for trading as compared to out sample period. Like I've explained in my "***Why does the Manual Strategy work***" section, my strategy tends to make profits when higher number of trades are performed. I've noticed with trial and error on the In-Sample period that the more frequently I enter and exit LONG/SHORT positions, the more profitable my strategy is. This can be attributed to the fact that although the profits you make from one single trade might be low, the cumulative profit that results from those high volume of trades overshadows a strategy which picks BUY/SELL positions very carefully (with tighter constraints).