ML4T Project 8

Chase Brooks
DBrooks43@gatech.edu

1 INTRODUCTION

One of the quintessential problems many traders try to solve is how to reliably beat the market. Over the years, various trading strategies have been introduced with the hopes of generating outsized returns. This paper aims to compare the performance of three different investing strategies: buy and hold, using technical indicators, and a bag learner of random forests trained on historical data. I will use the following parameters to standardize the results of each strategy: only trade JPM, starting cash balance of $100,000, in sample period 1/1/2008-12/31/2009, out of sample period 1/1/2010-12/31/2011. Based on countless musings of investing experts that it is very difficult to reliably beat the market, I expect the buy and hold approach to have the highest returns with the lowest standard deviation, the technical indicators strategy to have the second highest returns, and the learner approach to have the lowest returns.

2 INDICATOR OVERVIEW

For this project, I used three indicators to create my trading strategy.

2.1 Momentum

Momentum allows us to analyze both the magnitude and direction of a stock's price change over a given number of days. Tunable parameters include the lookback period/window and the threshold for a buy/sell opportunity. I tried many different parameter combinations, but settled on a 10 day lookback period.

2.2 Bollinger Band Percentiles (BBP)

Bollinger Bands allow traders to analyze current price in relation to the historical pricing action by incorporating views of a 20-day moving average as well as upper and lower bounds of two standard deviations. By comparing current price to the moving average, we can quickly see if a stock is over/under priced. Additionally, we can look at the current price to estimate the significance of the deviation in current price from average price. Tunable parameters include the number of days used in the moving average, how many standard deviations to use for the bollinger bands, and the threshold used to indicate a buying/selling opportunity. To simplify result interpretation, I transformed the bollinger bands into the bollinger band percentiles with the following formula: (price - lower band) / (upper band - lower band).

2.3 Relative Strength Index (RSI)

RSI is a momentum oscillator that gives traders an overbought and oversold indication. RSI incorporates the average gain/loss over a given window. To calculate RSI, we can use the following formula:

100 - (100/ ((1/moving average up) / (1/moving average down))), where moving average up is the average percentage gained for the stock on a day where the price increased and moving average down is the average percentage lost for the stock on a day where the price decreased.

3 MANUAL STRATEGY

3.1 Using indicators to create a signal

The overall premise I was striving for when creating a buy/sell signal relied on the assumption that mean reversion generally holds true. Even though stocks can have massive run ups or crashes, they generally revert back to their average price over a longer period of time. Assuming mean reversion to be a winning strategy, I first used RSI to determine when a stock was overbought/oversold. The assumption here is that when a stock is overbought, a sell off is likely to occur in the future and vice versa. Just because a stock is overbought/sold does not mean that it’s trajectory will change immediately. As such, I also used momentum to confirm when a stock’s trend began reverting to the mean. Finally, I used Bollinger Bands to confirm that a stock’s price had deviated significantly enough from the trailing average to warrant a reversal.

3.2 Choosing indicator thresholds to create a signal

To create my signals, I generally followed an educated guess and check approach to maximize returns. I first started with threshold recommendations from websites such as investopedia for each indicator. Through online research I found that the standard practice is to use rsi <= 0.3 as an oversold indicator (buying opportunity) and rsi >= 0.7 as an overbought indicator (selling opportunity). Using these thresholds in combination with my other indicators yielded no trading activity, so I changed the threshold to rsi > 0.5 is a selling opportunity and rsi < 0.5 is a buying opportunity. These thresholds in conjunction with my other indicators yielded satisfactory results. For BBP I used the 20 day moving average, 2 standard deviations, and a threshold of <= 0.2 to buy and >= 0.8 to sell. For momentum I used a threshold of > 0 for a buying opportunity and < 0 for a selling opportunity. Thus, my initial signals were as follows: **Buy**: rsi <= 30, momentum > 0, BBP <= 0. **Sell**: rsi >= 70, momentum > 0, BBP >= 1.0.

Upon running these strategies on JPM for the in-sample date range, I realized that no trades were being generated by the approach so I would need to relax my thresholds. After many rounds of trial and error, I settled on the following thresholds that balanced trading activity with returns. **Buy**: rsi <= 50, momentum > 0, BBP <= 0.2. **Sell**: rsi > 50, momentum < 0, BBP >= 0.8.

3.3 Experimental Results

Figure 1 shows the returns for manual strategy outlined in section 2 when trading JPM from 1/1/2008-12/31/2009. The benchmark strategy of buying and holding 1000 shares of JPM generated a return 0.02% over the period compared to a 36% cumulative return for the manual strategy.

Figure 2 shows the returns for manual strategy outlined in section 2 when trading JPM from 1/1/2010-12/31/2011. The benchmark strategy of buying and holding 1000 shares of JPM generated a return -9% over the period compared to a 0.5% cumulative return for the manual strategy.

Tables 1 and 2 show the cumulative return, average daily return, and standard deviation of returns for both in-sample and out of sample trading periods. The manual strategy outperformed the benchmark buy and hold strategy in both periods while maintaining a similar standard deviation. The majority of the difference in gains from each respective strategy in both time periods comes from the manual strategy shorting JPM. Since the benchmark is unable to profit when JPM is losing value, this can cause a divergence in returns if the manual strategy enters a short position before JPM's stock price decreases. This phenomenon leads to a high average daily return as well. Somewhat counterintuitively both strategies have a similar standard deviation of returns which is likely due to the fact that even though the manual strategy has more trades, it each position it enters is usually a prolonged period of weeks to months, so the majority of the volatility in returns it experiences is also experienced by the long term buy and hold strategy.

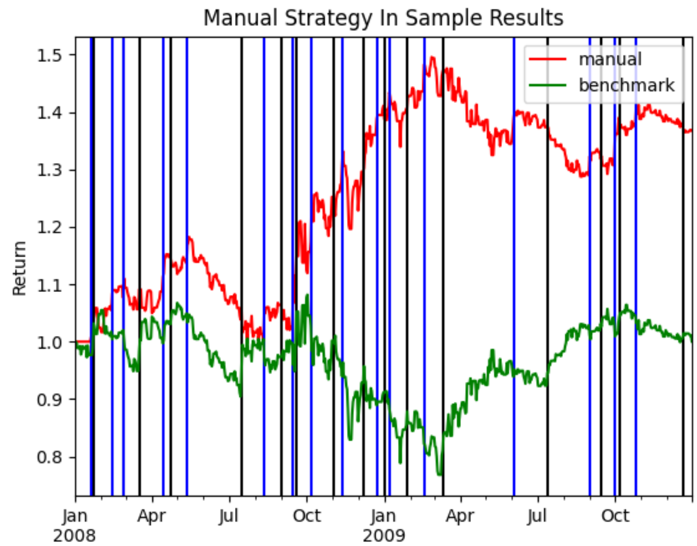

Figure 1— Manual Strategy vs. Benchmark returns trading JPM from 1/1/2008 through 12/31/2009 (in-sample).

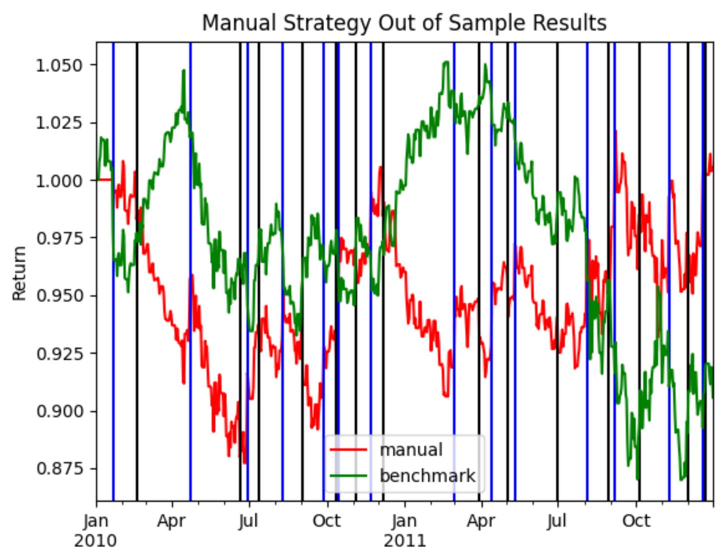

Figure 2— Manual Strategy vs. Benchmark returns trading JPM from 1/1/2010 through 12/31/2011 (out of sample).

Table 1— Manual Strategy vs. Benchmark returns trading JPM from 1/1/2008 through 12/31/2009 (in-sample).

In-Sample	Cumulative Return	Average Daily Return	Standard Deviation
Benchmark	0.02%	0.0098%	0.014
Manual Strategy	36%	0.049%	0.011

Table 2— Manual Strategy vs. Benchmark returns trading JPM from 1/1/2010 through 12/31/2011 (out of sample).

Out of Sample	Cumulative Return	Average Daily Return	Standard Deviation
Benchmark	-9.4%	-0.011%	0.007
Manual Strategy	0.5%	0.0039%	0.008

4 STRATEGY LEARNER

My strategy learner is a baglearner consisting of 25 bags of random forests.

4.1 Framing the problem

The first step in creating a random forest trader is to process the data such that we can use it to train a decision tree. I used the same indicators as the manual strategy (momentum, RSI, and BBP) for the splitters, but I needed to manually create the target variable that would be used in the leaf nodes. To accomplish this, I looked forward N days and measured the return during the training period. I artificially set a return threshold whereby if the future return was greater than my threshold I would enter a long position and if the future return was less than the negative threshold, I would enter a short position. The final training data set consisted of an X_Train of discretized indicators (see section 4.3) and a corresponding Y_Train array with an indicator of +1 for long, -1 for short, and 0 for do nothing.

4.2 Hyperparameters

My Strategy learner utilized three hyperparameters: number of bags, lookback window, and a return threshold. My final implementation used 25 bags, a lookback window of 7 days, and a return threshold of 5%. Additionally, each random forest learner has a leaf size of 10. To arrive at these values, I created a brute force hyperparameter testing script that searched through a list of possible parameters and selected the combination that generated the highest return amongst stocks used in the test cases (JPM, AAPL, SINE_FAST_NOISE, ML4T-220, and UNH).

4.3 Processing Data

To process the data, I discretized the outputs of each of the three indicators. I used the cut method from pandas to classify each indicator's output into 10 discrete bins. Completing this action simplified my decision tree and allowed for quick iteration to tune my hyperparameters and maximize the learner's returns. Since I set the random seed for this experiment, successive runs with the same hyperparameters return the same trades and thus returns.

5 EXPERIMENT 1

My strategy learner outperforms the benchmark and performs similarly to my manual strategy for JPM during the in-sample time period. I could certainly change the hyperparameters to increase performance of the strategy learner for JPM in-sample, but this would likely cause overfitting and poor performance on other tickers and time periods. In selecting my current hyperparameters, I also looked at other time periods and tickers and the performance was satisfactory. I initially hypothesized that the strategy learner would underperform both the benchmark and manual strategy, so I did not anticipate my strategy learner performing on par with the manual strategy and better than the benchmark buy and hold strategy.

Assumptions: 9.95 commission, 0.05 impact, only trade JPM, date range: 1/1/2008-12/31/2009, only legal holding positions are +1000, 0, -1000, and trades of +2000 or -2000 are legal if the holding position constraint is met.

Hyperparameters: 25 bags, 7 day lookback period, 5% return threshold to enter into a long/short position.

Table 3— Manual Strategy vs. Benchmark vs. Strategy Learner returns trading JPM from 1/1/2010 through 12/31/2011 (out of sample).

In-Sample	Cumulative Return	Average Daily Return	Standard Deviation
Benchmark	0.021%	0.0098%	0.014
Manual Strategy	36%	0.049%	0.011
Learner	38%	0.051%	0.011

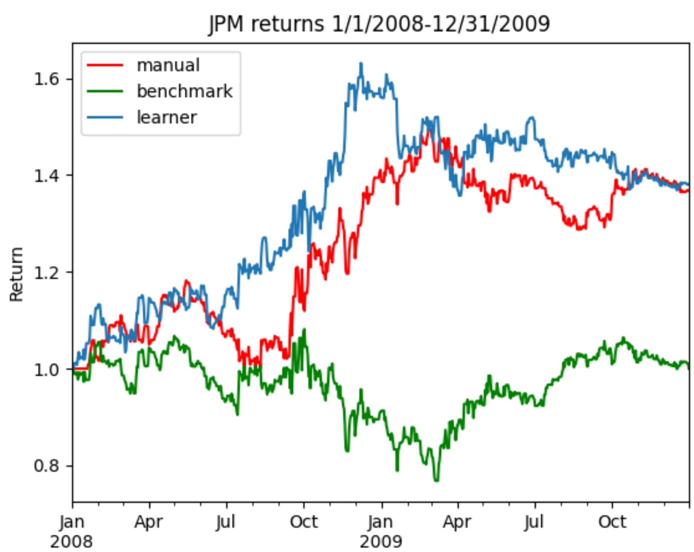

Figure 3— Manual Strategy vs. Benchmark vs. Strategy Learner returns trading JPM from 1/1/2010 through 12/31/2011 (out of sample).

6 EXPERIMENT 2

Generally the higher the impact, the lower the expected returns. As such, we should expect both the value of returns as well as the number of trades to decrease as impact increases. Figure 4 shows that having 0 impact generates higher returns than having a non zero impact. Figure 5 shows that the number of trades also decreases as impact increases. This is due to the expected return decreasing, which causes less scenarios where return exceeds the minimum return threshold for a buy/sell opportunity in the training data. With less buy/sell signals, there are less scenarios where the strategy learner would recommend a trade, thus a lower trading volume.

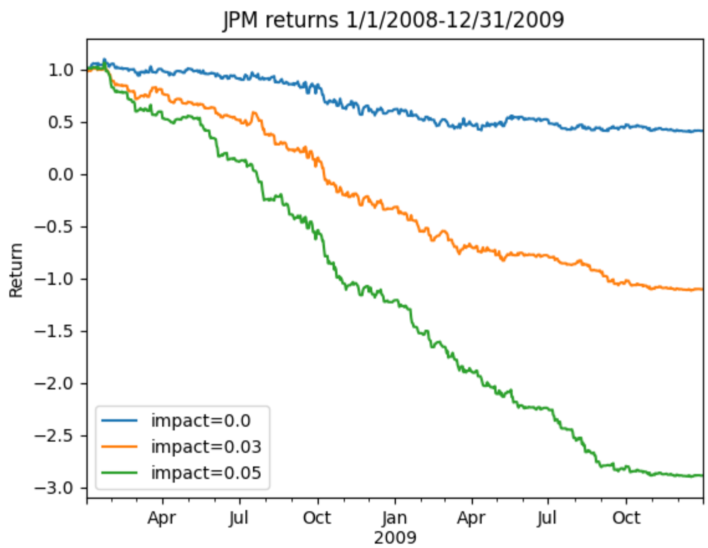

Figure 4— Strategy Learner returns with various impact values trading JPM from 1/1/2008 through 12/31/2009 (in-sample).

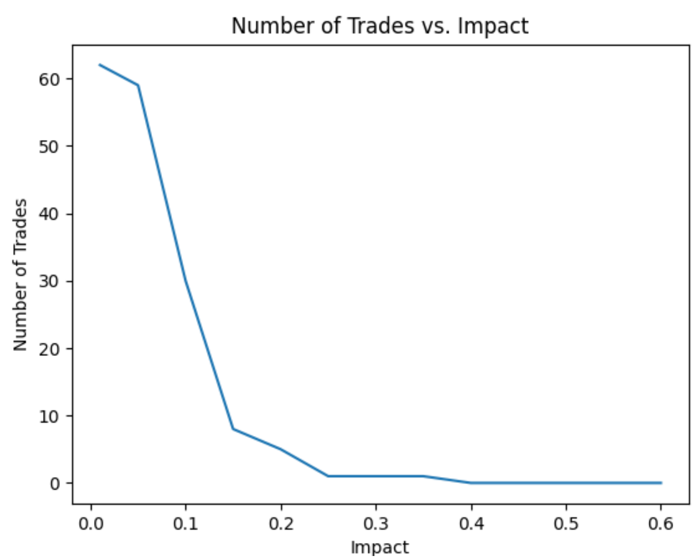

Figure 5— Strategy Learner number of trades vs. impact values trading JPM from 1/1/2008 through 12/31/2009 (in-sample).

7 REFERENCES

1. Fernando, J. (2021, November 20). Relative strength index (RSI). Investopedia. Retrieved November 29, 2021, from https://www.investopedia.com/terms/r/rsi.asp.
2. Barone, A. (2021, November 20). *What is momentum trading?* Investopedia. Retrieved November 29, 2021, from https://www.investopedia.com/trading/introduction-to-momentum-trading/.
3. Hayes, A. (2021, July 28). Bollinger Band®. Investopedia. Retrieved November 29, 2021, from https://www.investopedia.com/terms/b/bollingerbands.asp.