Analysis on $JPM : Machine Learning vs Manual Trading approach.

Kushal Atul Ramaiya
kramaiya3@gatech.edu

1 INTRODUCTION

This report is a comparative analysis of two different trading strategies based on technical indicators for the ticker $JPM. The two strategies are a machine learning-based approach and a manual approach. Both are derived from the same indicators and tested against their in-sample period, starting from January 1, 2008, and lasting until December 31, 2009, and an out-of-sample period from January 1, 2010, to December 31, 2011. The technical indicators used are the MACD, Bollinger Band Percentage, and Momentum. The constraint for these learners was that at any given point, the maximum number of JPM shares held long or short was 1000.

The manual strategy had explicit trading rules based on indicator signals, while the machine learning strategy used a reinforcement model known as Q-Learner. These were compared with a benchmark strategy of buying and holding 1000 shares for both periods.

This project also examines how three different values of impact affect the machine learning algorithm in terms of the number of trades taken, and the Sharpe Ratio is compared accordingly.

The initial hypothesis for this project was that the machine learning strategy would outperform the manual strategy in both in-sample and out-of-sample scenarios, as it is able to generalize and learn patterns well. However, it was expected that both strategies would perform better in the in-sample period than in the out-of-sample period, as the returns were optimized for the in-sample period.

2 INDICATORS

The indicators selected for both strategies were Moving Average Convergence/Divergence Histogram (MACD), Bollinger Band Percentage (%B), and Momentum. Each indicator generated a signal value of -1, 0, or 1, representing 'SELL', 'HOLD', and 'BUY', respectively. The MACD Histogram represents the difference between the MACD Line (calculated as the difference between the 12-day and 26-day exponential moving averages) and the Signal Line (the 9-day exponential moving average of the MACD Line). The histogram values typically range between -1 and 1, with occasional extremes slightly beyond this range. For the manual strategy, the histogram generated a 'BUY' signal when the Signal Line exceeded the MACD Line and a 'SELL' signal when the MACD Line exceeded the Signal Line. A 'HOLD' signal was generated when both lines were equal.

The Bollinger Band Percentage derives its values from Bollinger Bands and presents them as a percentage, providing a singular value from which signals can be generated. Typically, values range between 0 and 1, with occasional peaks and troughs extending beyond these limits, indicating extreme market conditions. For the manual strategy, these extreme values trigger signals: a 'SELL' signal when the value exceeds 1, indicating an overbought zone, and a 'BUY' signal when the value falls below 0, indicating an oversold zone and a potentially advantageous time to purchase. This indicator offers parameterization options for both the window size used to calculate the rolling standard deviation (determining the lower and upper Bollinger Bands) and the simple moving average window for calculating the average price over the preceding N days. The Momentum Indicator is a straightforward tool that measures the rate of price change over a specified period, which can be parameterized in this function. Using a 10-day window typically produces values centered around 0, as returns within such a period rarely reach 100% (which would correspond to a Momentum value of 1). The indicator has a lower bound of -1, representing a maximum possible negative return of -100%, as further declines are not feasible. Based on these parameters, the manual strategy generates a 'BUY' signal when the Momentum value exceeds 0 and a 'SELL' signal when it falls below 0. For the machine learning algorithm, all indicator values use the same parameters as in the Momentum and Bollinger Band Percentage calculations. These values are normalized to range between 0 and 1, then binned into 10 different values from 0 to 9, and finally converted into a three-digit value between 000-999. In this representation, the first digit derives from the MACD Histogram, the second from the Bollinger Band Percentage, and the third from the Momentum indicator. All indicator values were shifted forward by one day to avoid using future data, as these indicators are calculated based on the closing prices of each trading day.

3 MANUAL STRATEGY

The Manual Strategy utilized the combined sum of all indicator signals, with each indicator providing a signal value of -1, 0, or 1. This combined sum would range from -3 to 3. For values greater than 0 in the combined sum, if there was no existing position, the strategy would buy 1000 shares of $JPM (JP Morgan Chase). Similarly, for values lower than 0, the strategy would take a short position of 1000 shares.

In scenarios where a position (long or short) already existed, a combined sum of 0 would trigger an exit, executing the opposite trade with the same number of shares. If a long position existed and the combined sum turned negative, the strategy would implement a 'sell and reverse' approach, selling double the number of shares to both exit the current position and initiate a new short position. Similarly, for an existing short position, if the combined sum became positive, the strategy would trigger a buy position of double the existing position. All other scenarios would result in a trade value of 0, indicating a 'HOLD' position where no action would be taken.

For this project, the focus was on maximizing returns during the in-sample period from January 1, 2008, to December 31, 2009. To achieve this, a grid search was conducted using different parameter values ranging from 1 to 50 for both the Bollinger Band window and the Momentum window to identify the combination that yielded maximum returns. As in Figure1, a subset of the top 100 runs of the

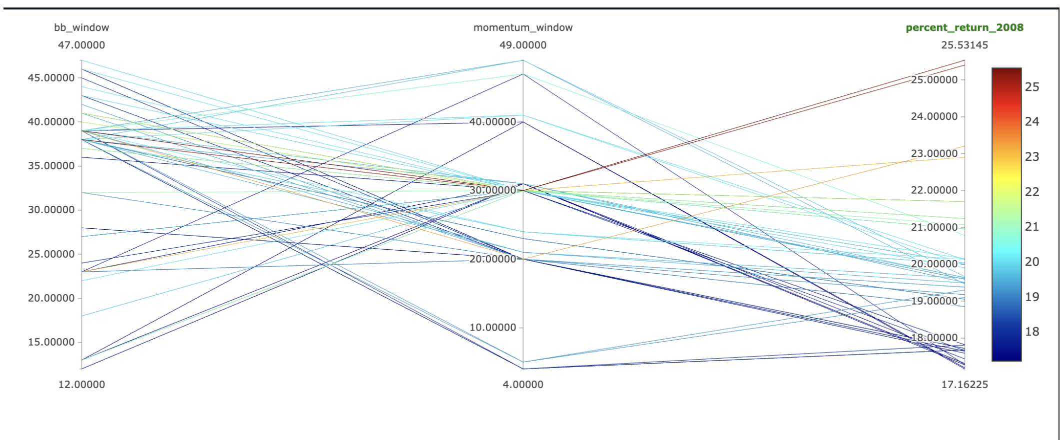

Figure 1—Top 100 runs based on percentage return and their respective bollinger band window and momentum window values

entire grid search is shown here, based on the percentage return on in-sample data. The data shown indicates that only a couple of runs had a value greater than 25, with the maximum distribution between 18 to 23, and the highest percent return of 25.53% in the in-sample period. The parameters for the highest return had a Bollinger Band window of 39 and a Momentum window value of 30. These values were set for indicator values for both the manual strategy and the strategy learner's parameters. Figure2 shows the performance comparison between the

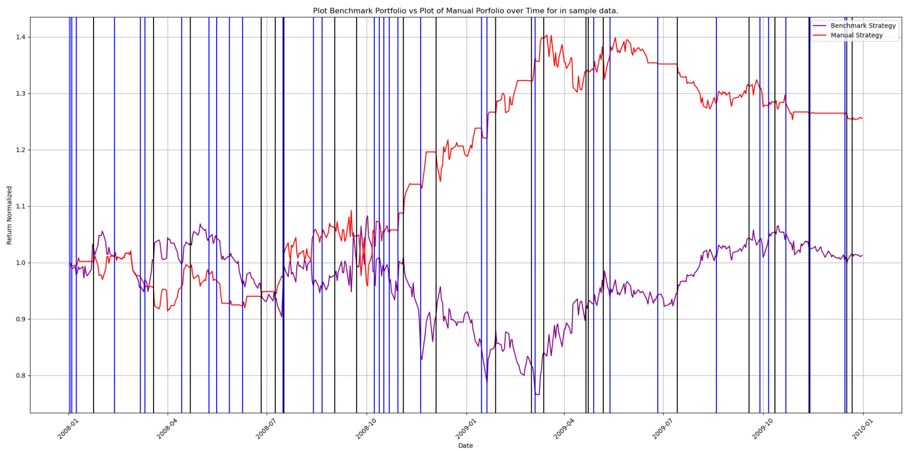

Figure 2—Benchmark Portfolio vs Manual Strategy for in-sample learner.

Manual Strategy and the Benchmark Strategy for the in-sample period. The blue and black lines indicate the entry signals for the Manual Strategy. From the strategy, it is evident that until October 2008, the strategy was not performing well with the indicators. However, during the volatility caused by the housing market crash, it started to gain a significant advantage, though some of the gains capitalized during the crash were later lost.

Figure3 shows the out-of-sample performance for both strategies. The benchmark as well as the manual strategy both have negative returns, but the manual strategy barely manages to beat the cumulative return for the out-of-sample data. This shows that while the manual strategy is able to beat the benchmark during the in-sample period, it barely manages to outperform and in fact underperforms for most of the out-of-sample period.

Thus, the strategy is not effective as it does not generalize well to the out-of-sample period and may even result in negative returns. Another reason for the difference in returns could be the volatility in $JPM's price movement during

Table 1—Comparison of Metrics with Benchmark and Manual Strategy

Metric	**Benchmark Strategy**	**Manual Strategy**
Cumulative Return in-sample	0.012325	0.255315
Daily Return Std in-sample	0.017041	0.01189
Daily Return Mean in-sample	0.000169	0.000522
Cumulative Return out-of-sample	-0.083579	-0.083323
Daily Return Std out-of-sample	0.0085	0.007808
Daily Return Mean out-of-sample	-0.000137	-0.000143

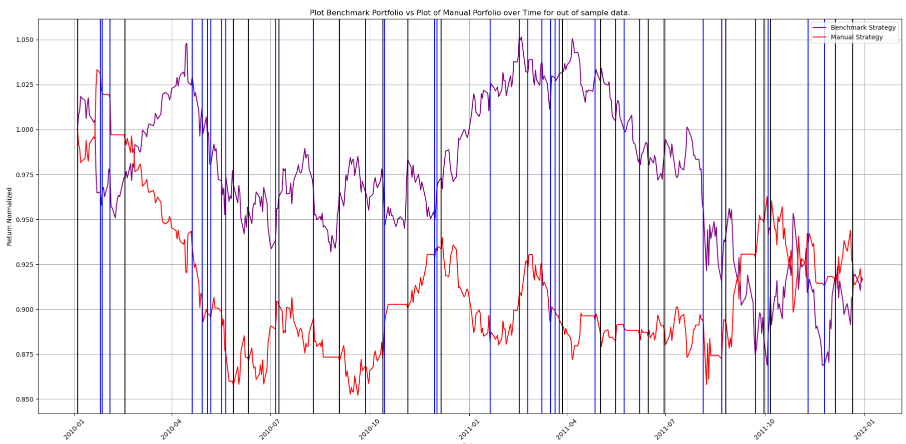

Figure 3—Benchmark Portfolio vs Manual Strategy for out-of-sample learner.

those periods. The in-sample period had extreme volatility, and the manual strategy earned most of its profits during the market downturn. In contrast, during the out-of-sample period, the market was in recovery, and the manual learner had negative returns during the upward movement of $JPM prices. Since the parameters for the indicators were optimized for the in-sample period, they likely overfit to that period for maximum returns during a drawdown in the market, and hence were not able to capitalize during the recovery phase.

4 STRATEGY LEARNER

The strategy learner was a machine learning algorithm known as Q-Learner, which is a model-free reinforcement learning algorithm. This means that the learner can learn the optimal policy by directly interacting with the environment without requiring a model of the dynamics. For this project, the Q-Learner was chosen because the learner's current actions affect future states, which aligns with how actions in the stock market such as 'BUY', 'SELL', or 'HOLD' affect the portfolio.

The trading problem had to be specified as a learning problem with a state, action, and reward. The indicators used the same window values as those tuned in the Manual Strategy, and the state for this model would be those indicator values. However, since these indicator values were continuous, the state space would be infinite, as the values would form a set of real numbers. Hence, to discretize and manage the state space, the indicator values were normalized to a range between 0 and 1. Then, the data was binned into 10 different bins, with values ranging from 0 to less than 0.1 as 0, 0.1 to less than 0.2 as 1, and so on. This allowed all the indicators to take discrete values from 0 to 9.

The resulting state space was 10^3, ranging from 000 to 999, with each digit representing one of the indicator values. For example, if the indicators on a particular day, after preprocessing, had values [5, 3, 8], the state would be 538. This allowed us to limit the state space to only 1000 possible states and avoid excessive sparsity.

The actions were encoded as 0, 1, or 2, representing 'SELL', 'HOLD', or 'BUY' respectively. The reward for this model was defined as the daily return (if the model held any position) minus the impact cost. With this setup, the learner could be trained using the state and reward. The learner was optimized to achieve higher returns by passing through the entire dataset multiple times, allowing it to efficiently take trades that would outperform both the benchmark and the manual strategy during in-sample learning.

5 EXPERIMENT 1

For Experiment 1, the project focused on evaluating the performance of the Manual Strategy, Strategy Learner, and Benchmark. To conduct this experiment, the prices of JP Morgan (Ticker: $JPM) were used, with the in-sample period spanning from January 1, 2008, to December 31, 2009. The in-sample period was used to train the Strategy Learner, which discretized the data as mentioned earlier and trained on it. Similarly, the same data was passed to the Manual Strategy, and the Benchmark return was collected. An impact cost of 0.005 and a commission cost of $9.95 were applied to all strategies. The impact cost represents 0.5% of the trading volume, and the commission simulates real-life costs associated with paying brokers and accounting for market impact.

The hypothesis was that the Strategy Learner would slightly outperform both in-sample and out-of-sample, as the learner is able to iterate through the data in each epoch and estimate returns, while the Manual Strategy was optimized using fixed parameters for the indicators.

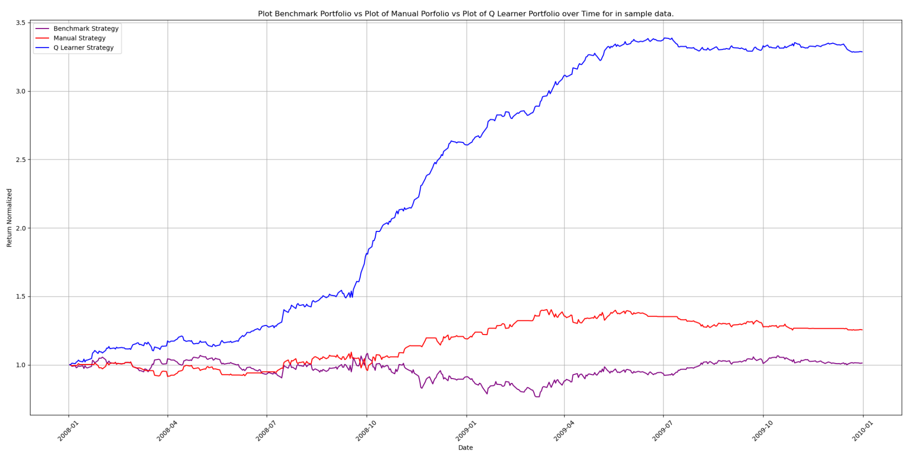

Figure 4—Benchmark Portfolio vs Manual vs Strategy Learner Portfolio for in-sample.

Figure4 shows that the returns from the Q-Learner Strategy portfolio outperformed both the Manual Strategy and the Benchmark, with a cumulative return of 2.25% relative to the initial value, compared to 0.25% for the Manual Strategy and 0.01% for the Benchmark. This confirms the hypothesis that the learner would outperform on the data it was trained on.

For the second part of the experiment, the learners were trained on the in-sample period but evaluated against the out-of-sample period. Figure5 shows that even with strong performance in the in-sample period, the model fails to outperform the Benchmark and the Manual Strategy during the out-of-sample period. This indicates that the model is overfitting to the in-sample data and is not generalizing well, whereas the Manual Strategy, while more generalized, does not outperform the market by a margin that instills confidence in the system.

This demonstrates that the Q-Learner's performance during the in-sample period consistently generates higher returns, as it learns patterns in the data and

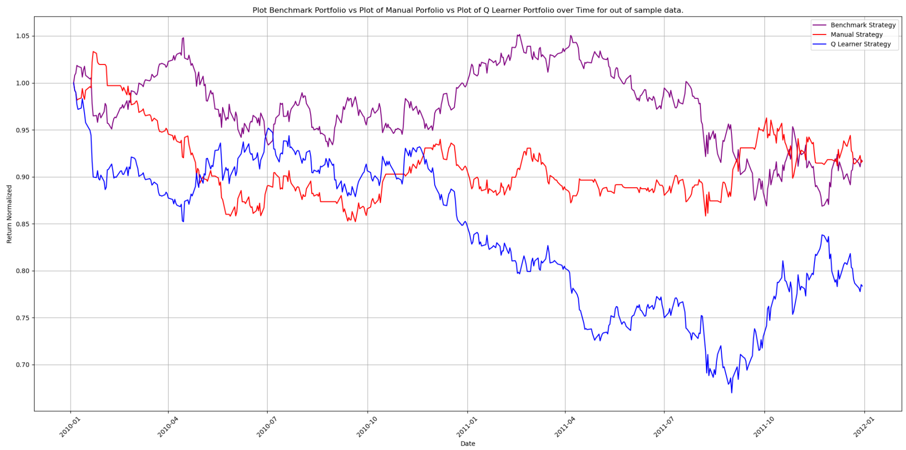

Figure 5—Benchmark Portfolio vs Manual vs Strategy Learner Portfolio for out-of-sample.

attempts to maximize returns. However, its performance in the out-of-sample period does not replicate these results. In contrast, there is not enough data to conclude whether the Manual Strategy will always outperform or match the Benchmark in the out-of-sample period, although there are intervals where it does outperform the Benchmark over short durations.

6 EXPERIMENT 2

The impact parameter represents the cost of taking a trading position. In the case of large orders, the price of the stock typically moves against the trader, and this phenomenon is captured by the impact parameter. This experiment investigates how different impact values affect the returns of the Strategy Learner. It is hypothesized that increasing the impact parameter will influence the number of trades taken by the Strategy Learner and affect its overall performance.

Specifically, the number of trades should decrease as the impact increases, since the chances of achieving a positive return are reduced. For instance, a trade that may have resulted in a small profit under a low impact setting could result in a negative return if the impact is high enough to offset that small gain. Additionally, cumulative return is expected to be inversely proportional to the impact parameter: if higher returns are observed under high impact values, then a learner using lower impact values should be able to take the same trades and achieve even higher returns.

To conduct this experiment, three Q-Learners were initialized with impact values of 0.005, 0.010, and 0.025, with a commission cost of 0. Each learner was trained on the same in-sample period, from January 1, 2008, to December 31, 2009. The trades generated by each learner were used to compute the portfolio value for each day, and the results were compared by evaluating the daily portfolio values.

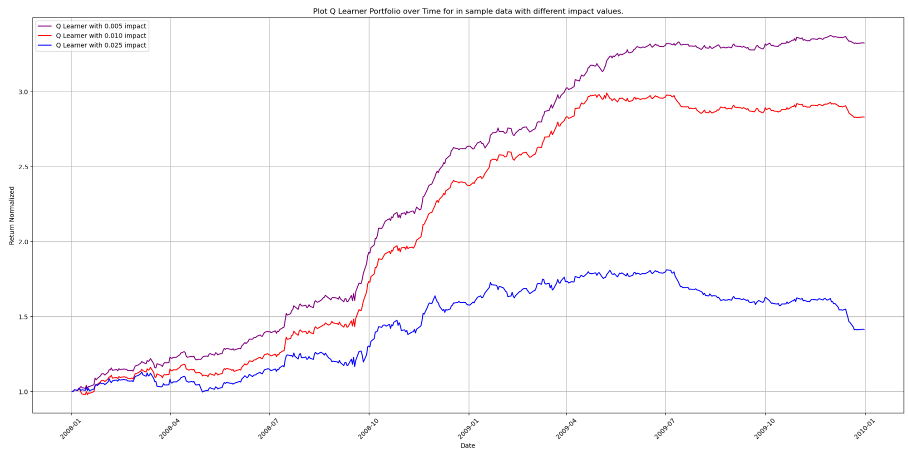

Figure 6—Portfolio values of different Q-leaners with differing impact values

As shown in Figure6, our hypothesis is correct: each learner was trained on the same data, and the learner with the lower impact value was able to generate higher returns. The performance metrics for Q-Learners with different impact values are presented in Table2.

Table 2—Comparison of Metrics with different Q learners

Metric	**Impact 0.005**	**Impact 0.010**	**Impact 0.025**
Number of Trades	149	129	65
Cumulative Returns	2.427168	1.833752	0.464639
Daily Returns	0.002478	0.002104	0.00083
Sharpe Ratio	4.961899	3.949625	1.09112

These parameters show how the impact value affects the learner, and even a small change in impact value can affect the performance significantly. This would be similar to trading stocks where liquidity is low, as buying larger quantities can cause the stock price to move significantly.