A Strategy Evaluation Journey

Shichao Zhou
szhou401@gatech.edu

1 INTRODUCTION

In the project 8, we will develop and compare two trading strategies: a manual strategy crafted with our insights and an AI-driven strategy learner. By comparing these strategies, we gain valuable insights into their performance and behavior, laying a foundation for continued exploration in trading and machine learning.

2 INDICATOR OVERVIEW

For Manual Strategy and Strategy Learner, I utilized three key technical indicators out of the five implemented in Project 6.

2.1 Relative Strength Index (RSI)

RSI measures the speed and change of price movements. It oscillates between 0 and 100, with values above 70 indicating overbought conditions and values below 30 indicating oversold conditions. I calculated RSI based on a 14-day look-back period using daily price data to find the best balance between sensitivity and reliability.

2.2 Bollinger Bands Percentage (BBP)

BBP assesses a stock's volatility and relative price position within its Bollinger Bands. It quantifies the distance between the stock price and the bands The calculation of BBP involves normalizing this deviation, yielding values within the 0 to 1 range. The optimization process primarily focused on the selection of a 20-day look-back period for computing the Bollinger Bands. Additionally, a threshold value for generating buy/sell signals may be optimized (e.g., BBP > 0.8 for overbought conditions).

2.3 Moving Average Convergence Divergence (MACD)

MACD measures the difference between short-term and long-term exponential moving averages. It includes the MACD line and the MACD histogram. I calculated MACD by subtracting the 26-day exponential moving average (EMA) from the 12-day EMA. The MACD histogram is derived from the difference between the MACD line and the 9-day EMA of the MACD line. The primary parameter optimized is the length of the short-term and long-term EMAs (e.g., 12 and 26 days). Additionally, the threshold value for the MACD histogram (e.g., 0.2 or 0.0) may be optimized for generating buy/sell signals.

3 Manual Strategy

In my manual strategy, I integrated signals from the Relative Strength Index (RSI), Bollinger Bands Percentage (BBP), and Moving Average Convergence Divergence (MACD) to generate a comprehensive trading signal. A buy or sell action is executed only when all three conditions align, thereby reducing investment risk. Regarding the criteria for initiating a long position (buy) or exiting a long position (sell), please refer to the following explanation.

3.1 Buy condition:

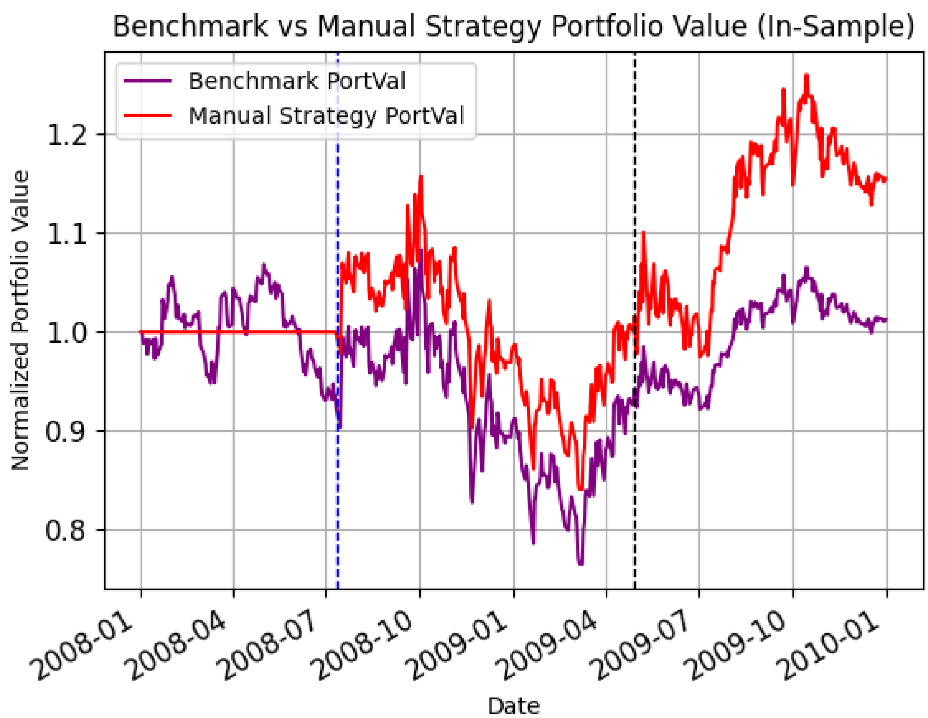
- RSI values below 35: Suggesting potential entry points during oversold or neutral conditions.
- BBP values below 0.2: Indicating a potential price dip within Bollinger Bands, signifying a buying opportunity.
- MACD histogram values above 0.0: Indicating potential upward momentum in the market.

3.2 Sell condition:

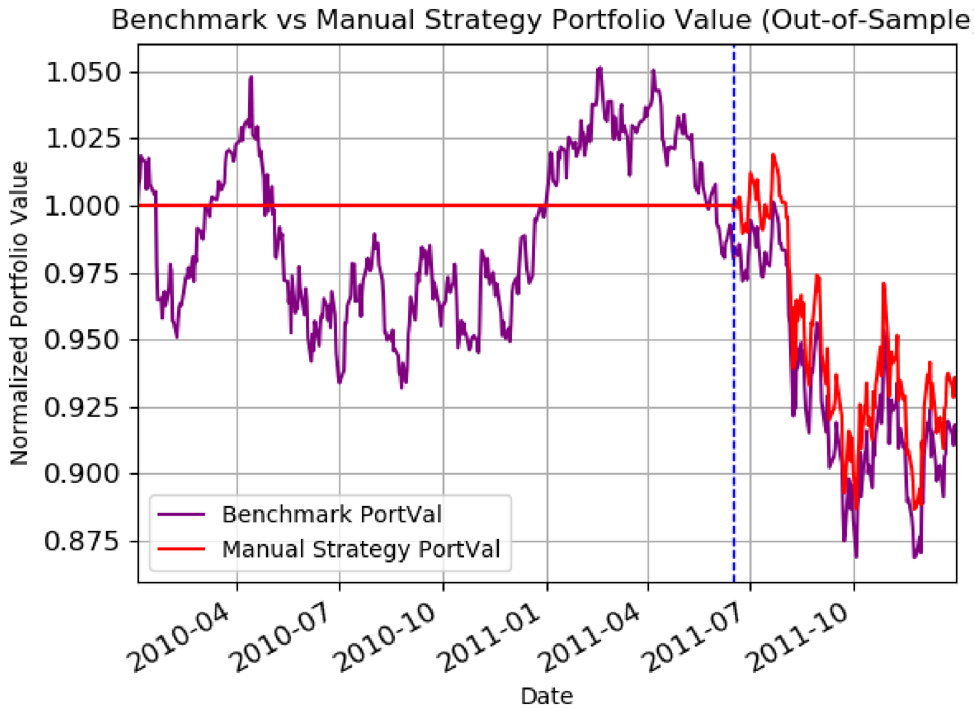
- RSI values above 59: Suggesting potential exit points during overbought conditions. Under normal circumstances, an RSI value above 70 indicates an overbought condition, signaling an opportunity to sell stocks. However, during periods of weak market sentiment, low liquidity, or unfavorable conditions like the 2008 financial crisis, it is safer to lower the RSI threshold to the range of 58-60. In my strategy, I have set the RSI threshold to 59.

- BBP values above 0.8: Signifying potential overextended price levels.
- MACD histogram values below 0.0: Indicating a potential downturn in market momentum.

The strategy's effectiveness can be assessed through a performance comparison with a benchmark. I employed a benchmark that represents a baseline trading strategy. This benchmark involves starting with $100,000 in cash, investing in 1000



shares of the symbol in use on the first trading day and holding that position. I then compared the strategy's performance against this Manual Strategy. In the in-sample period, the Manual Strategy demonstrated superior performance compared to the benchmark. The manual strategy achieved a cumulative return of 0.1546, surpassing the benchmark returns of 0.0123.



- Based on the graphical representations, it is evident that the out-of-sample cumulative return of the manual strategy (-0.0657) exhibits a smaller negative value compared to the benchmark's cumulative return (-0.0835), indicating a performance advantage of 0.0178. Under the manual strategy, the confluence of these three indicators led to buy signals during the out-of-sample period but did not trigger any sell signals. To potentially enhance the strategy's effectiveness in capturing sell opportunities at appropriate high JPM stock prices, one could consider adjusting the MACD histogram's threshold range from values less than 0 to values less than 0.25. However, it is important to note that using a threshold of 0.25 may introduce subjectivity and the risk of lookahead bias. When analyzing the performance disparities between the in-sample and out-of-sample periods, it becomes evident that the manual strategy did not fare favorably during the latter period, despite surpassing the benchmark in cumulative return within the former. Several factors can be attributed to this discrepancy:

- Market Dynamics: Market conditions and trends can exhibit significant variations between the two periods, thereby impacting the strategy's effectiveness.
- Parameter Sensitivity: Parameters optimized for the in-sample period may not remain as effective when applied to the out-of-sample period, primarily due to evolving market dynamics.

4 STRATEGY LEARNER

To frame the trading problem as a learning problem for the Strategy Learner, the following main steps were taken:

1. Indicator: According to project 6, I had five indicators as options, and based on the results of cumulative return, I ultimately chose three technical indicators, namely the Relative Strength Index (RSI), Bollinger Bands % (bbp), and Moving Average Convergence Divergence (MACD), which I also utilized in the manual strategy. From the table below, we can observe that the combination of golden cross, RSI, and MACD yields the highest cumulative return when the symbol is JPM. However, this combination did not pass all the auto-grading test cases. That is why I ultimately selected RSI, BBP, and MACD

Indicator	JPM Strategy CR
Golden Cross、BBP、RSI	-0.024008
Golden Cross、BBP、MACD	-0.476455
Golden Cross、BBP、Momentum	0.0392185
Golden Cross、RSI、MACD	0.287956
Golden Cross、RSI、Momentum	-0.398141
Golden Cross、MACD、Momentum	-0.220886
BBP、RSI、MACD	0.1879945
BBP、RSI、Momentum	-0.1153735
BBP、MACD、Momentum	-0.003022
RSI、MACD、Momentum	-0.0781355

2. Discretization: To enable the Q-Learner to work with the continuous indicator values, discretization was performed. The RSI, BBP, and MACD values were divided into discrete bins to categorize them. The QCUT method from NumPy was employed to perform the discretization of each indicator's values. This method segmented the indicator values into 10 bins, resulting in a new dataframe for each indicator. Each date index in these dataframes was paired with a discrete value ranging from 0 to 9.

3. State Creation: The Q Learning agent had a total of 1,000 possible states, ranging from 000 to 999, to exist in during its learning process. In the part, these discretized indicator values served as the specific state of the Q Learning agent. To create a unique state representation for the agent, the discrete values from the three indicators were concatenated into a three-digit number. For instance, if the bins for the three indicators(rsi,bbp,macd) on a specific date were [1, 0, 1], the resultant state would be 101.

4. Q-Learner Initialization: I initialized the Q-Learner with hyperparameters, the number of states for the Q-Learner was set to 1000. The number of actions for the Q-Learner was set to 3, corresponding to BUY, SELL, and NOTHING). The random action rate for the Q-Learner was set to 0. The impact parameter, which can be adjusted as needed. The Q-Learner would learn to make trading decisions based on these parameters.

5. Training: I iteratively trained the Q-Learner by determining the daily returns and updating the state. The Q-Learner provided actions based on the current state and executed these actions while considering transaction costs (impact and commission). This process continued until convergence or a maximum number of iterations (MAX_ITERATIONS) was reached. Within the specified 25-second time window, increasing ITERATIONS can enhance the cumulative return, so I set max_interations as 600.

6. Testing Policy: In the test policy method, the training and convergence steps were not relevant. Instead, the method performed a single pass through the date loop to determine the best actions to take based on the

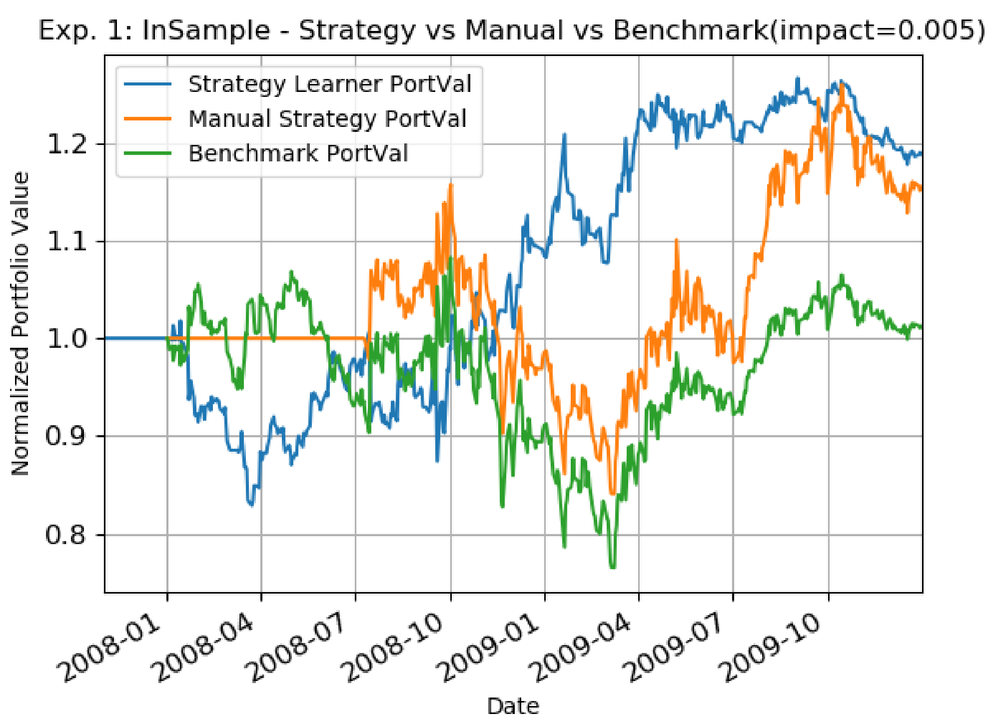
model created during training. The agent's actions were selected based on the learned policy.

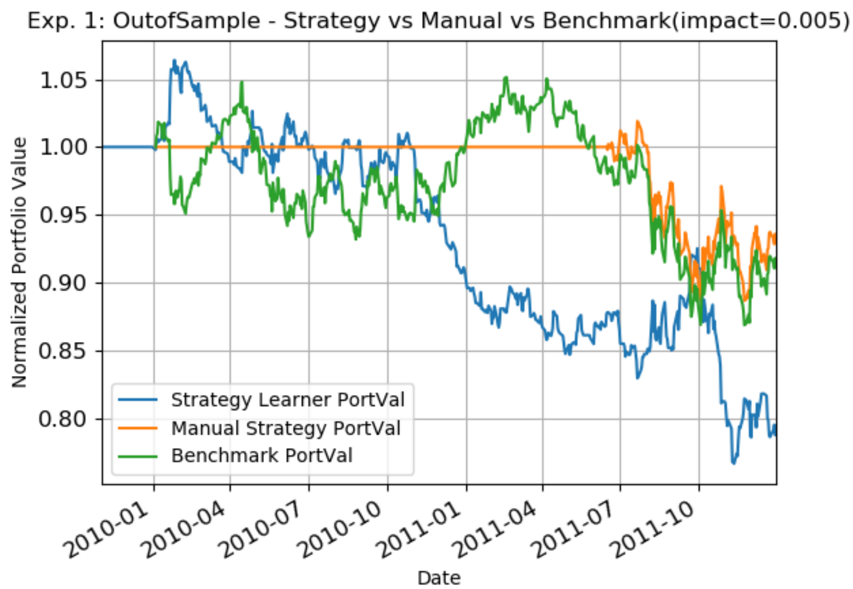
5 EXPERIMENT-1

This experiment aims to compare the performance of the Strategy Learner, Manual Strategy, and a benchmark strategy. The experiment is conducted using two different time periods: an in-sample period (from 2008-01-01 to 2009-12-31) and an out-of-sample period (from 2010-01-01 to 2011-12-31). The impact parameter for trading was set to 0.005 and commission was set to 9.95 for both the strategy learner, manual strategy and a benchmark strategy. As illustrated in the tables and figures below, the Strategy Learner outperformed the Manual Strategy in the in-sample period. This outcome is logical since the Strategy Learner employs a "learning" approach, dynamically adapting its actions based on the indicators, while the Manual Strategy relies on a fixed set of static conditions for the indicators on every date. Consequently, the Strategy Learner strives to identify the most optimal actions and constructs a model for them, whereas the Manual Strategy passively selects actions based on predefined indicator conditions that provide a reasonable estimate of future price fluctuations. Additionally, the Strategy Learner's better performance in in-sample period is expected because it is tested on the same data it was trained on, implying a good model fit indeed, it should achieve a perfect fit as I set the random action rate (rar) of my Q Learner to 0.0 for making the agent rely entirely on its known best strategy. However, the manual strategy performed better than the strategy learner during the out-of-sample period. In the out-of-sample period, RAR set to 0 could prevent the Strategy Learner from adapting to new market conditions, resulting in poor performance. The second reason I assumed is the market in the out-of-sample period may have higher levels of noise and volatility, which could lead to a decrease in the Strategy Learner's performance.

	in-sample		
	benchmark	manual strategy	strategy learner
cumulative return	0.012325	0.154621	0.187995
avg daily return	0.000168	0.000436	0.000401
standard deviation of daily return	0.017041	0.017392	0.013031
sharpe ratio	0.157205	0.397683	0.489076

	out-of-sample		
	benchmark	manual strategy	strategy learner
cumulative return	-0.083579	-0.065744	-0.212539
avg daily return	-0.000137	-0.000117	-0.000409
standard deviation of daily return	0.0085	0.005868	0.007632
sharpe ratio	-0.256656	-0.319157	-0.851192

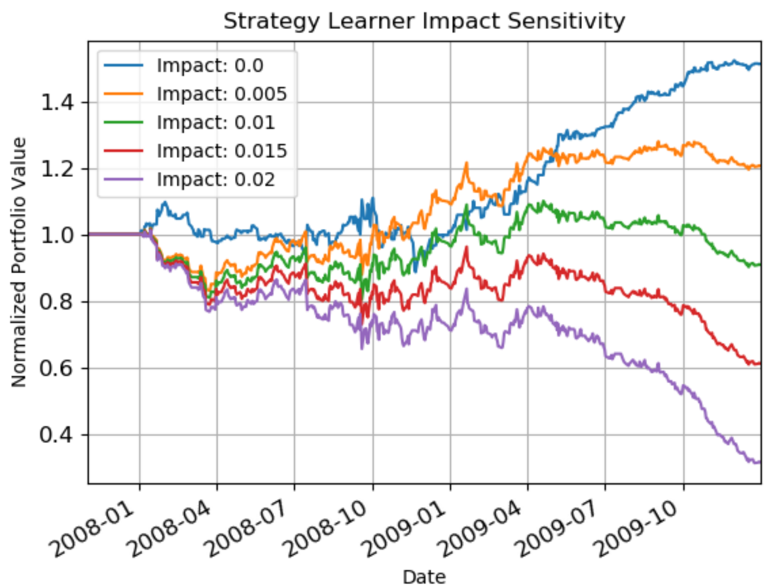




6 EXPERIMENT-2

In this experimental setup, I established two hypotheses. To test these hypotheses, I systematically varied the impact level from 0.000 to 0.02, adjusting it in increments of 0.005. 5.

- From the below first figure, the increase in the impact level is anticipated to result in a decline in cumulative returns for the trading strategies. This expectation is rooted in the fact that the impact, which is integrated into the portfolio value calculation within marketsim, affects the costs associated with buying and the proceeds from selling. Consequently, it leads to reduced returns for each trade, ultimately translating to an overall decrease in portfolio value.



- Based on the second figure, it is hypothesized that as the impact level rises, it may not significantly alter the frequency of trades. A lower reward may reduce the agent's motivation, but it is unlikely to substantially affect the trade frequency.