

Introduction

This report mainly discusses how to format trading problem as a Q learning problem such that a learning trading agent is created to beat the benchmark performance.

Dataset

The data used in this report is the stock price data of symbol JPM, and only stock adjust close price when SPY traded is used here. In sample data is from 01/01/2018 to 12/31/2019 and out sample data is from 01/01/2010 and 12/31/2011.

Format trading problem as Q Learning problem

Q learning problem includes state, action, reward and Q matrix. For trading problem, the action is buy, sell and hold, and stock position is long, short and cash. State is a single value combining the discretization values of three technical indicators. They are momentum, bollinger value and price/sma. Reward is daily reward. The addEvidence method in StrategyLearner file is used to create the q learner for trading problem as follow:

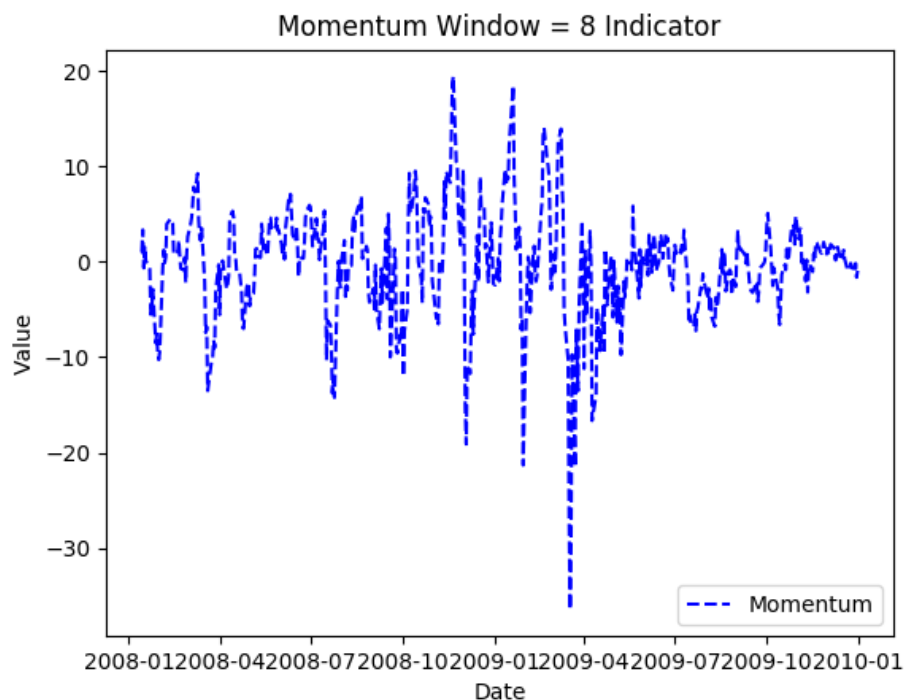
1. get price for the stock
2. calculate and combine all indicators
3. calculate the thresholds for all indicators
4. During each epoch:
 - 4.1 set initial default position to cash
 - 4.2 During each trading date:
 - 4.2.2 get the discretization values of all indicators and transfer them into a single current state
 - 4.2.3 change to opposite position if the date is the last one
 - 4.2.4 just query q learner to get the action and update the position if the date is the first one
 - 4.2.5 for other dates, calculate the daily reward for last action, query and update q learner with current state and reward to get the action, update the position.
 - 4.2.6 save the daily position
 - 4.3 calculate the stock value based on the daily position and get the cumulative return.
 - 4.4 check and repeat from 4.1 until cumulative return is converged.

Technical Indicators

The technical indicators used are momentum, bollinger value and price/sma.

Momentum

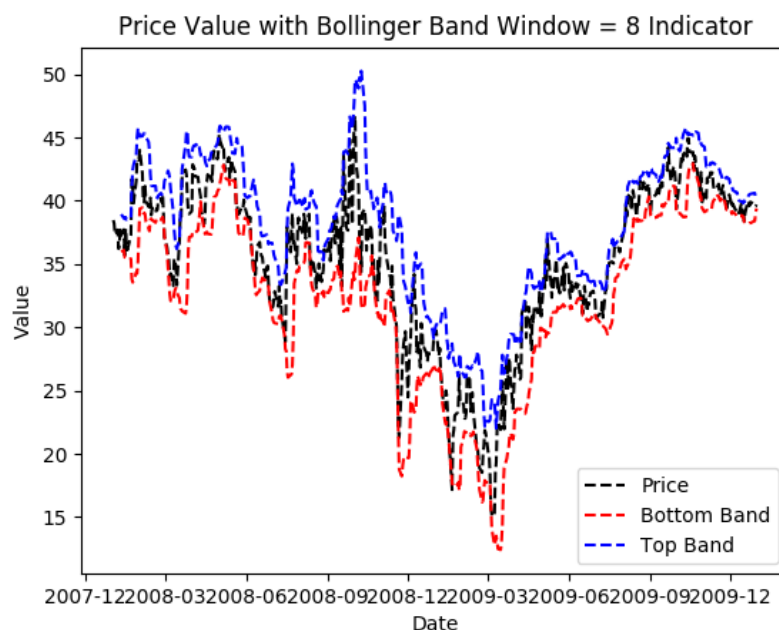
zzhang76 (Zhi Zhang)



momentum compares the current security price with the price n periods ago, it equals $[\text{close price}/\text{close price } n \text{ periods ago}] - 1$. It measures the speed at which prices are changing. A upward surge in momentum reflects a sharp price advance, which might be a signal to buy the stock, while a downward plunge indicates a deep price decline, which might be a signal to sell the stock. The momentum with window 8 from 01/01/2008 to 12/31/2009 is shown above.

bollinger value

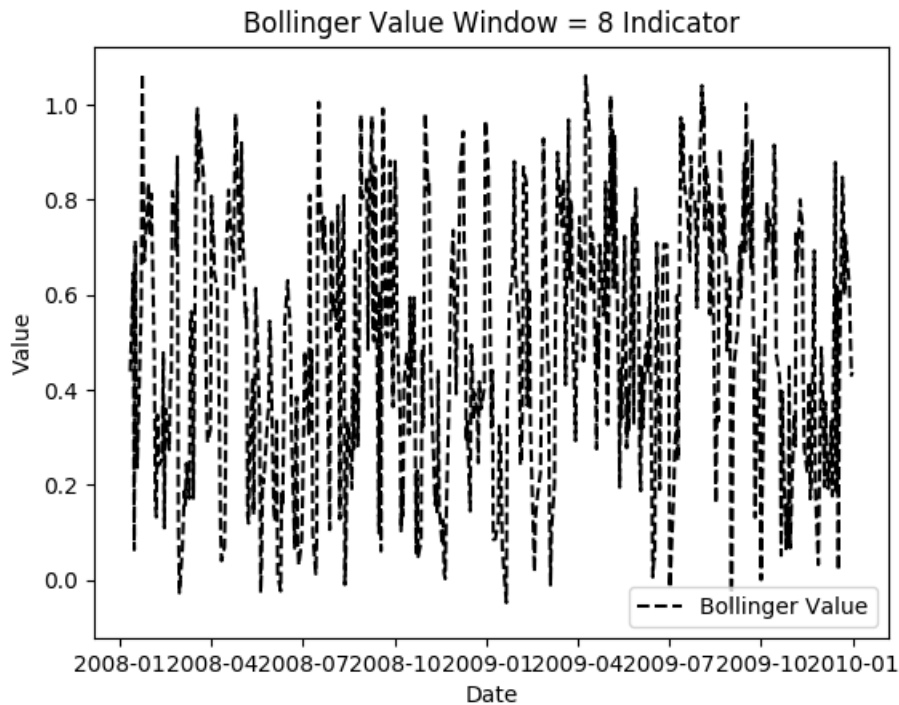
bollinger value or %B quantifies a security's price relative to upper and lower bollinger band. It equals $(\text{price} - \text{lowerband})/(\text{upperband} - \text{lowerband})$. The lowerband equals $(\text{rolling_mean} - 2 \times \text{rolling_std})$, and upperband equals $(\text{rolling_mean} + 2 \times \text{rolling_std})$. A surges towards the upper band show strength, and can be interpreted as overbought sometimes, which might be a signal to sell the stock.



zzhang76 (Zhi Zhang)

A plunge to the lower band show weakness, and can be interpreted as oversold, which might be a signal to buy the stock. The price value with Bollinger Band of window 8 from 01/01/2018 to 12/31/2019 is shown above.

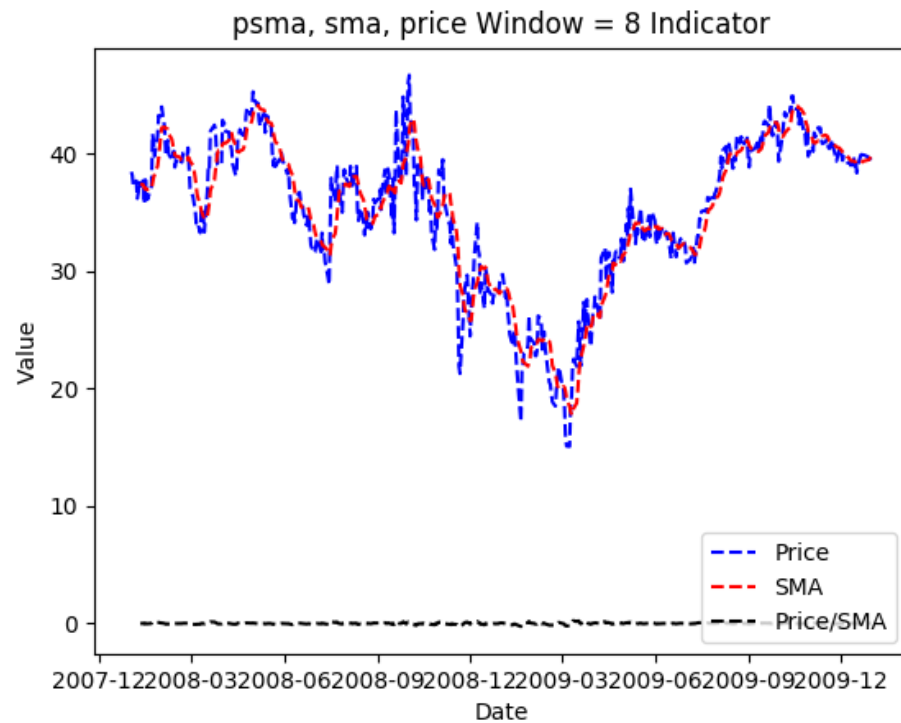
The real Bollinger value with window 8 from 01/01/2018 to 12/31/2019 is shown below:



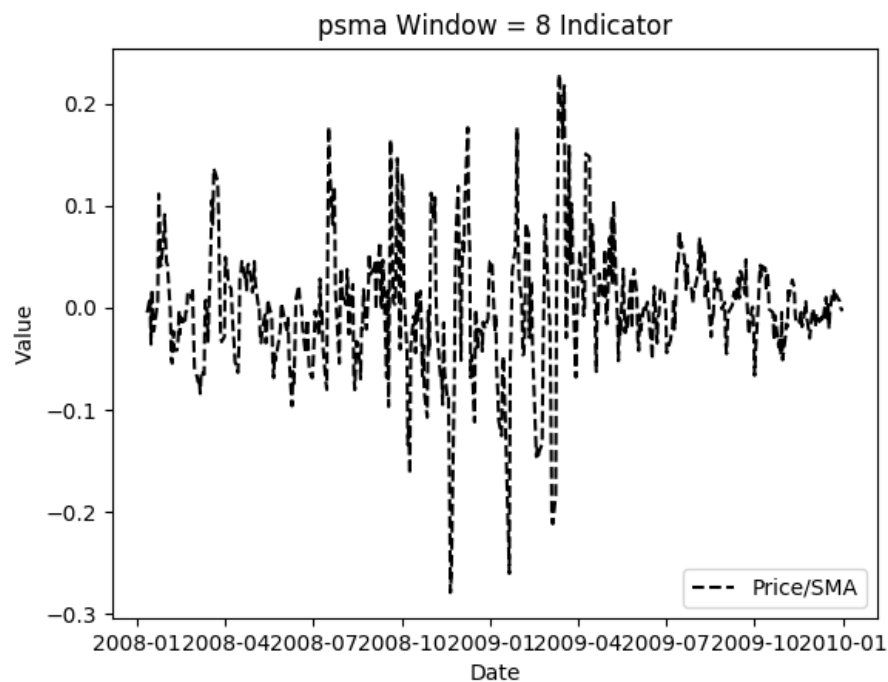
price/sma

A simple moving average is the average price of a security over a specific number of periods. A rising moving average shows that price are generally increasing, which might be a signal to buy the stock. A falling moving average indicate that price are generally decreasing, which might be a signal to sell the stock. Price/SMA is a normalized version of SMA. The price, sma and price/sma value with window 8 from 01/01/2018 to 12/31/2019 is shown below:

zzhang76 (Zhi Zhang)

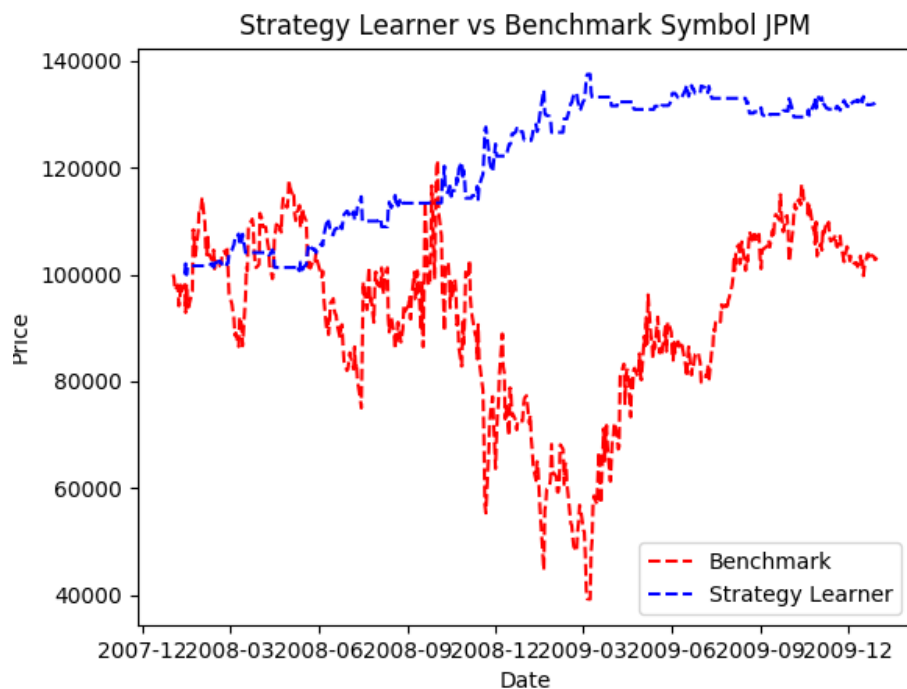


The real price/sma value with window 8 from 01/01/2018 to 12/31/2019 is shown below:



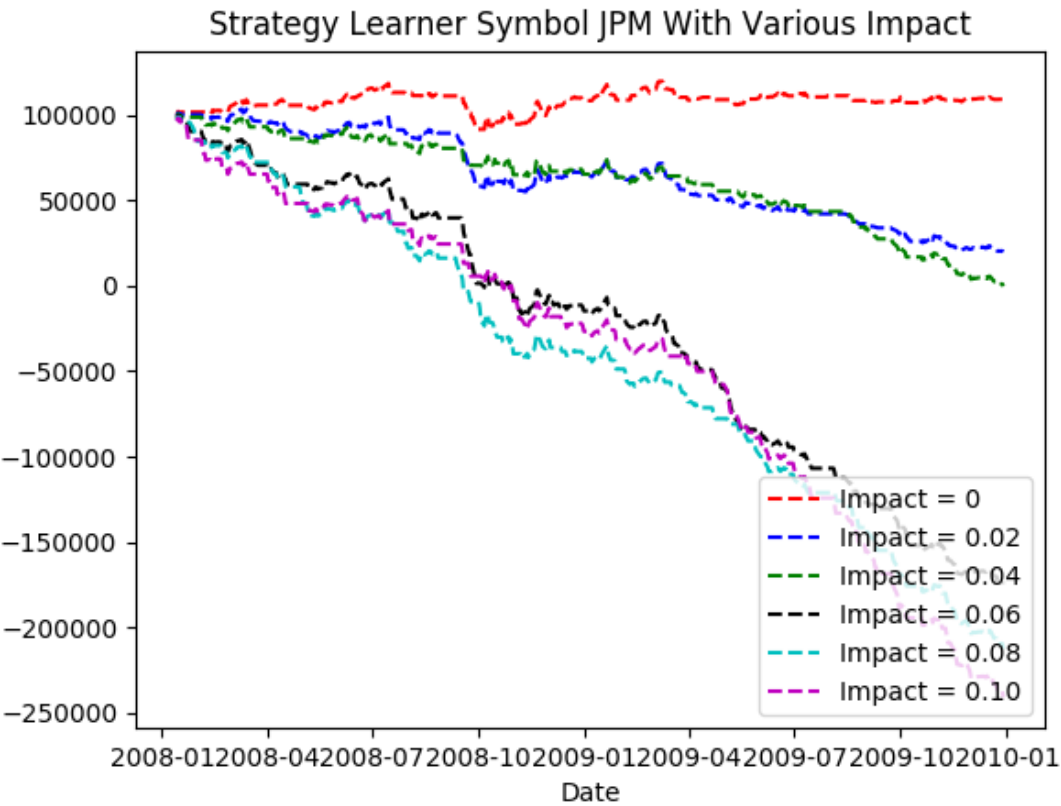
Experiment1

In order to compare the performance of the strategy learner, we choose a basic benchmark. The benchmark is the performance of a portfolio starting with \$100,000 cash, investing in 1000 shares of the symbol JPM and holding that position. The portfolio value of stock JPM under strategy learner is also drawn. The in-sample data from 01/01/2018 and 12/31/2019 is used. The outcome is that strategy learner has better performance than the benchmark and it is about 30% more. The relative result may not remain the same for in-sample data since the strategy learner may predict bad actions and generate bad performance. It only consider the information in all technical indicators. Sometimes, a rare chance event can happen such that even a buy at the first day and sell at the last day as the benchmark shows may have better performance.



Experiment2

The increased value of impact will decrease the performance of strategy learner, and it will also decrease the number of trade. The performance with various impact from in-sample 01/01/2018 to 12/31/2019 is shown below:



However, the number of trade does not decrease as I think and it is shown below:

Impact	0	0.02	0.04	0.06	0.08	0.1
# of trade	106	120	92	128	120	102