Three indicators:

1.Bollinger Band

We use this formula:

$$bb\_value[t] = (price[t] - SMA[t])/(2 * stdev[t])$$

Standard deviation is a measure of volatility. When the stock becomes more volatile, the bands widen. During less volatile period, the bands contract.

This indicator shows the relative position between bands.

The closer the prices move to upper band, the more overbought the stock, and the closer the prices move to lower band, the more oversold the stock.

I normalize that feature so that it will typically provide values between -1.0 and 1.0

2. Moving average crossovers

When shorter-term 12 period moving average (MA) crosses over the longer-term 26-period MA, a buy signal is generated.

Formula:

$$mac=(MA\_12-MA-26)/MA\_26$$

This is like MACD model.

This is simplest and effective momentum indicator.

I normalize that feature so that it will typically provide values between -1.0 and 1.0

3.RSI

This is a well-known momentum oscillator that measures the speed and change of price movements.

Formula:

$$RSI=(100-100/(1+RS))/100$$

$$RS=Average\ Gain/\ Average\ Loss$$

RSI measures the velocity and magnitude of directional price moves and represents the data graphically by oscillating between 0 and 100.

I normalize that feature so that it will typically provide values between -1.0 and 1.0

These three indicators could well reflect momentum and volatility of the stock price.
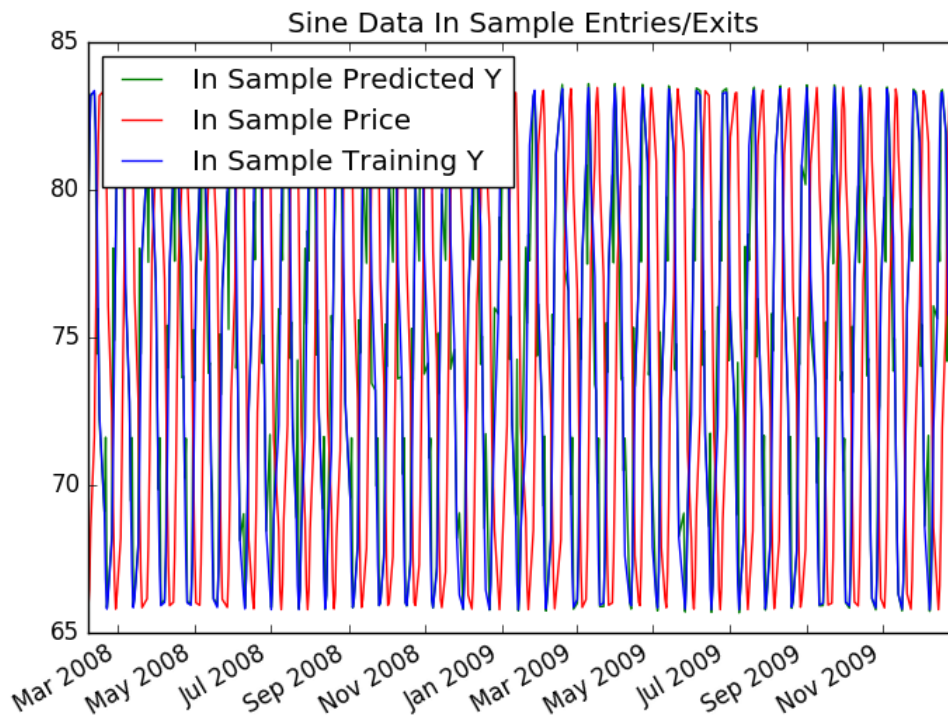
KNN (k=6) regression learner:
My policy:

I choose k=6. When k is small, result is not very good. Overfitting may occur. Based on predictions from KNN learner, we will buy when predicted Y>3%, and we will sell when predicted Y<3% after 5 days.

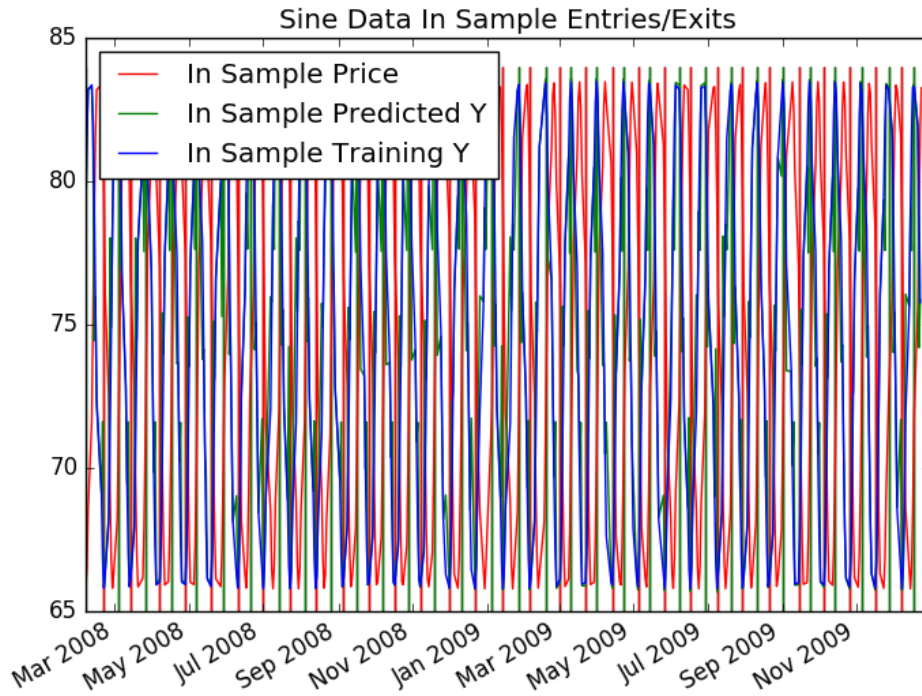If we choose >2% or 1%, lots of trades will occur. When volatility is high, result is not great.
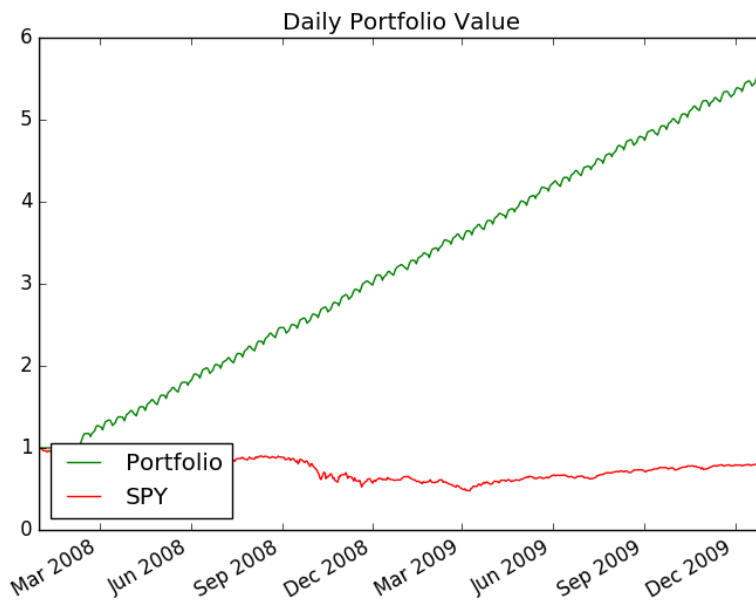
Sine Data:

1.Training Y/Price/Predicted Y:

2. Sine Data In Sample Entries/Exits
Red line means 'Sell'. Green line means 'Buy'. Black line means 'Exit'.

Sine Data In Sample Entries/Exits

3.Sine Data In Sample Backtest

Daily Portfolio Value

Data Range: 2007-12-31 to 2009-12-31
Start-val=10000
Sharpe Ratio of Fund: 5.50430969731
Sharpe Ratio of SPY: -0.149575888341
Cumulative Return of Fund: 4.45118455
Cumulative Return of SPY: -0.201395139514
Standard Deviation of Fund: 0.00983976647495

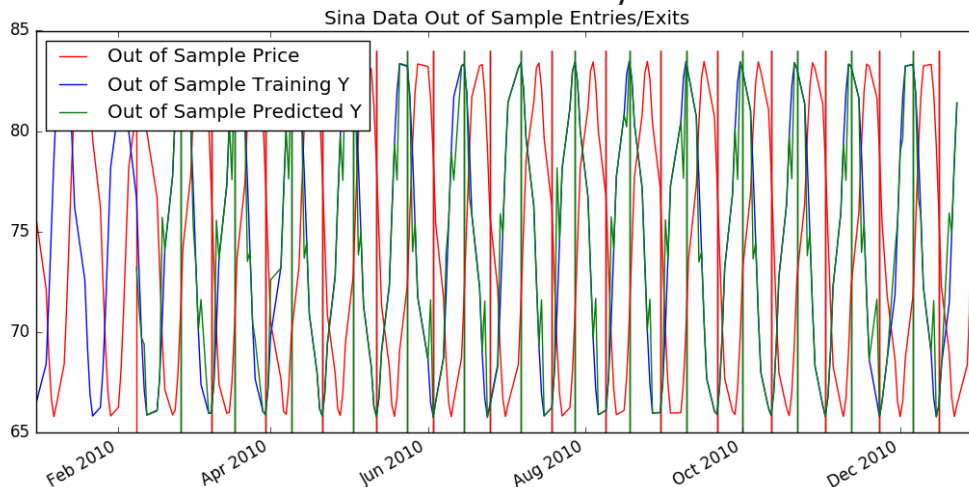Standard Deviation of SPY: 0.0219136847778
Average Daily Return of Fund: 0.00341182999078
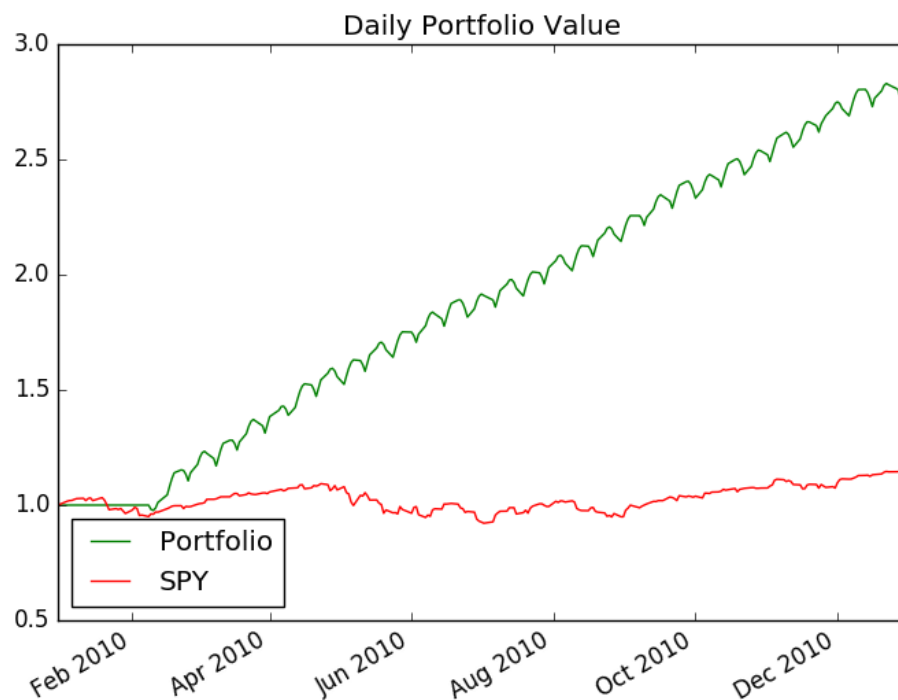Average Daily Return of SPY: -0.000206479400499
Final Portfolio Value: 54511.8455
4.Sine Data Out of Sample Entries/Exits
Red line means 'Sell'. Green line means 'Buy'. Black line means 'Exit'.

5. Sine Data Out of Sample Backtest

Data Range: 2009-12-31 to 2010-12-31
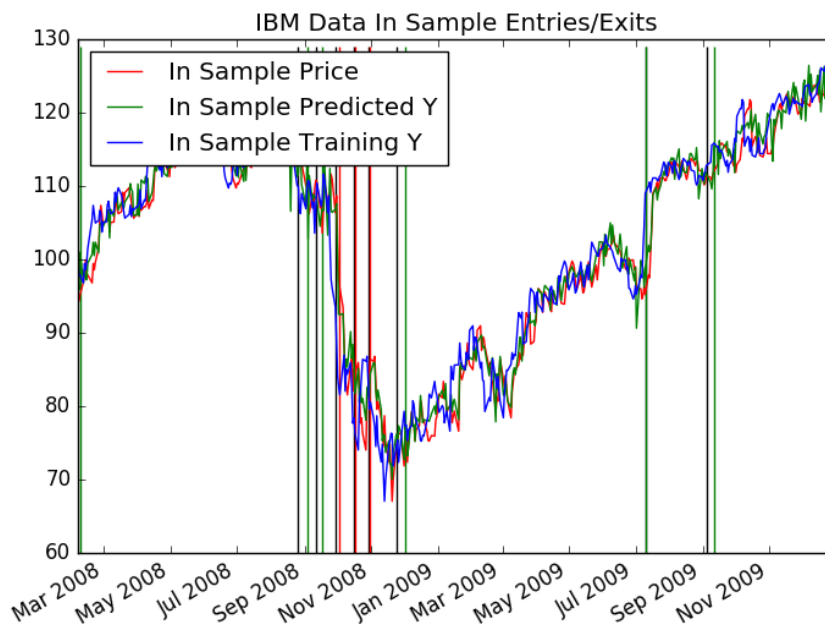Start-val=10000
Sharpe Ratio of Fund: 4.72643629421

Sharpe Ratio of SPY: 0.848235559397
Cumulative Return of Fund: 1.67369987
Cumulative Return of SPY: 0.145862684324
Standard Deviation of Fund: 0.0134339042542
Standard Deviation of SPY: 0.011306957999
Average Daily Return of Fund: 0.00399977707495
Average Daily Return of SPY: 0.000604173932453
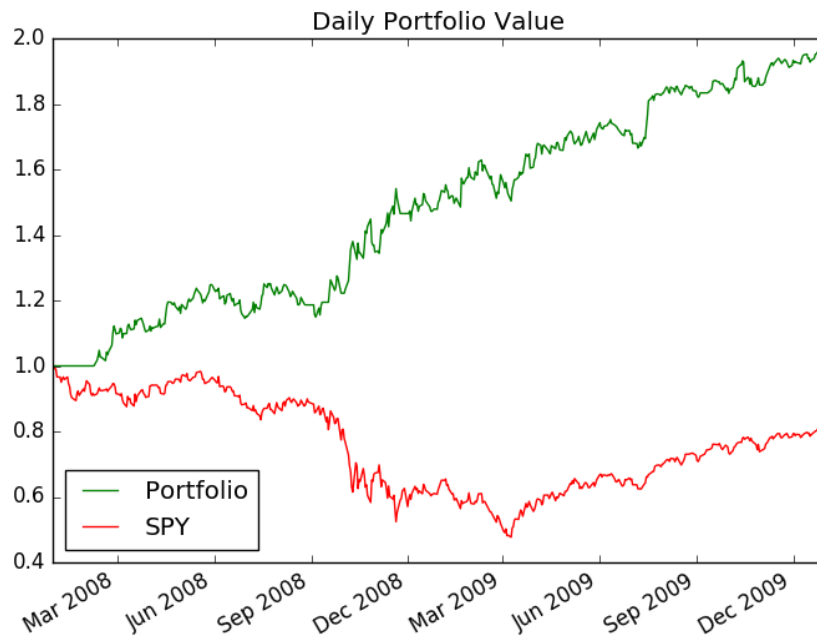Final Portfolio Value: 26736.9987

IBM Data

6. IBM Data In Sample Entries/Exits:

Red line means 'Sell'. Green line means 'Buy'. Black line means 'Exit'.

7. IBM Data In Sample Backtest

Daily Portfolio Value

Data Range: 2007-12-31 to 2009-12-31
Start-val=10000
Sharpe Ratio of Fund: 1.85002133674
Sharpe Ratio of SPY: -0.149575888341
Cumulative Return of Fund: 0.9611
Cumulative Return of SPY: -0.201395139514
Standard Deviation of Fund: 0.0120736876543
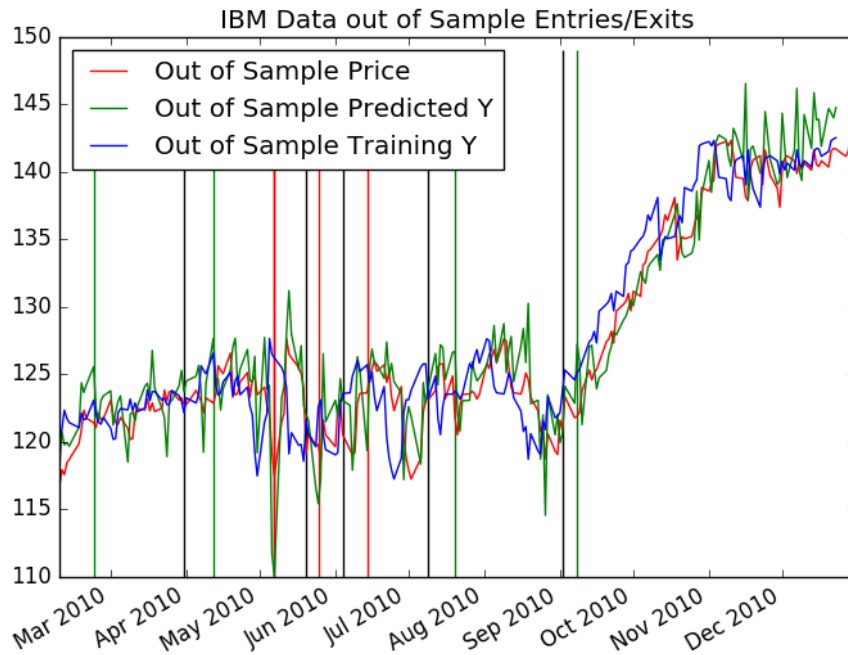Standard Deviation of SPY: 0.0219136847778
Average Daily Return of Fund: 0.00140707226716
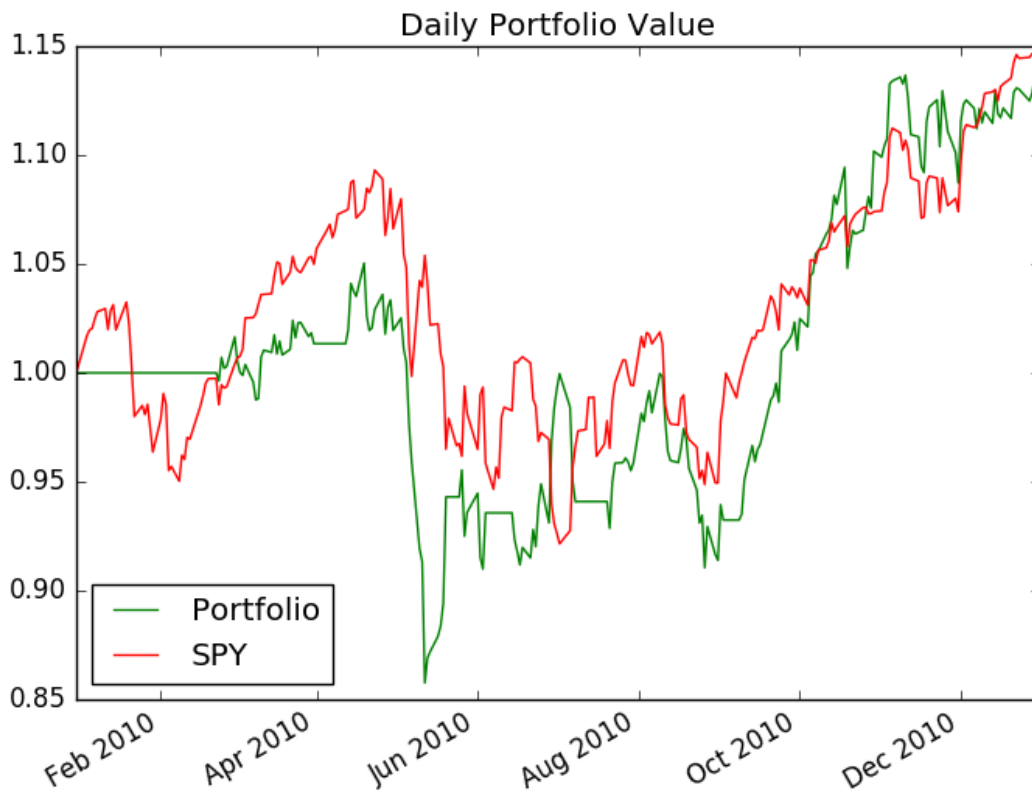Average Daily Return of SPY: -0.000206479400499
Final Portfolio Value: 19611.0
8.IBM Data Out of Sample Entries/Exits
Red line means 'Sell'. Green line means 'Buy'. Black line means 'Exit'.

IBM Data out of Sample Entries/Exits

9. IBM Data Out of Sample Backtest

Daily Portfolio Value

Data Range: 2009-12-31 to 2010-12-31
Start-val=10000

Sharpe Ratio of Fund: 0.765012090438
Sharpe Ratio of SPY: 0.848235559397
Cumulative Return of Fund: 0.1387
Cumulative Return of SPY: 0.145862684324
Standard Deviation of Fund: 0.0122555099052
Standard Deviation of SPY: 0.011306957999
Average Daily Return of Fund: 0.000590608120319
Average Daily Return of SPY: 0.000604173932453
Final Portfolio Value: 11387.0

Result:

This policy does well with ML4T-399 and IBM.

According to results, portfolio value is higher than SPY.

Generally, KNN can well predict future stock price. But it is hard to decide when to buy and when to sell. Because ML4T-399 is sine data whose data show very high return, sell/buy is easy to decide. When its price is low, a buy is generated. When its price is high, a sell is generated.

For normal stock price, news and market can impact it. 1% and 2% increasing/decreasing are very common. Risk would be high. We hope we can catch an uptrend or downtrend, not a small fluctuation. For a big uptrend/downtrend, variation is relatively large. So we choose 3%.

During 2010, we found IBM fluctuated back and forth for a long time. It is very hard to decide the buy/sell signal because of high risk. In later design, I will try to avoid trades during this kind of fluctuation. KNN is hard to predict dramatic changes due to earnings report, or other news. For unpredictable results, I will stop trade. Or based on fundamental analysis, I do corresponding operations.