Description of Indicators:

1. Volatility: Calculated the volatility by taking the rolling standard deviation of the (normalized) daily returns.
2. Momentum: Calculated the (normalized) momentum, using a window of ten. Accomplished this by dividing each day's price by the price of ten days previous, then subtracting one.
3. Bollinger Values: Calculated the Bollinger values, using a standard moving average with a window of ten and a rolling standard deviation with a window of ten.

Description of Strategy and Results:

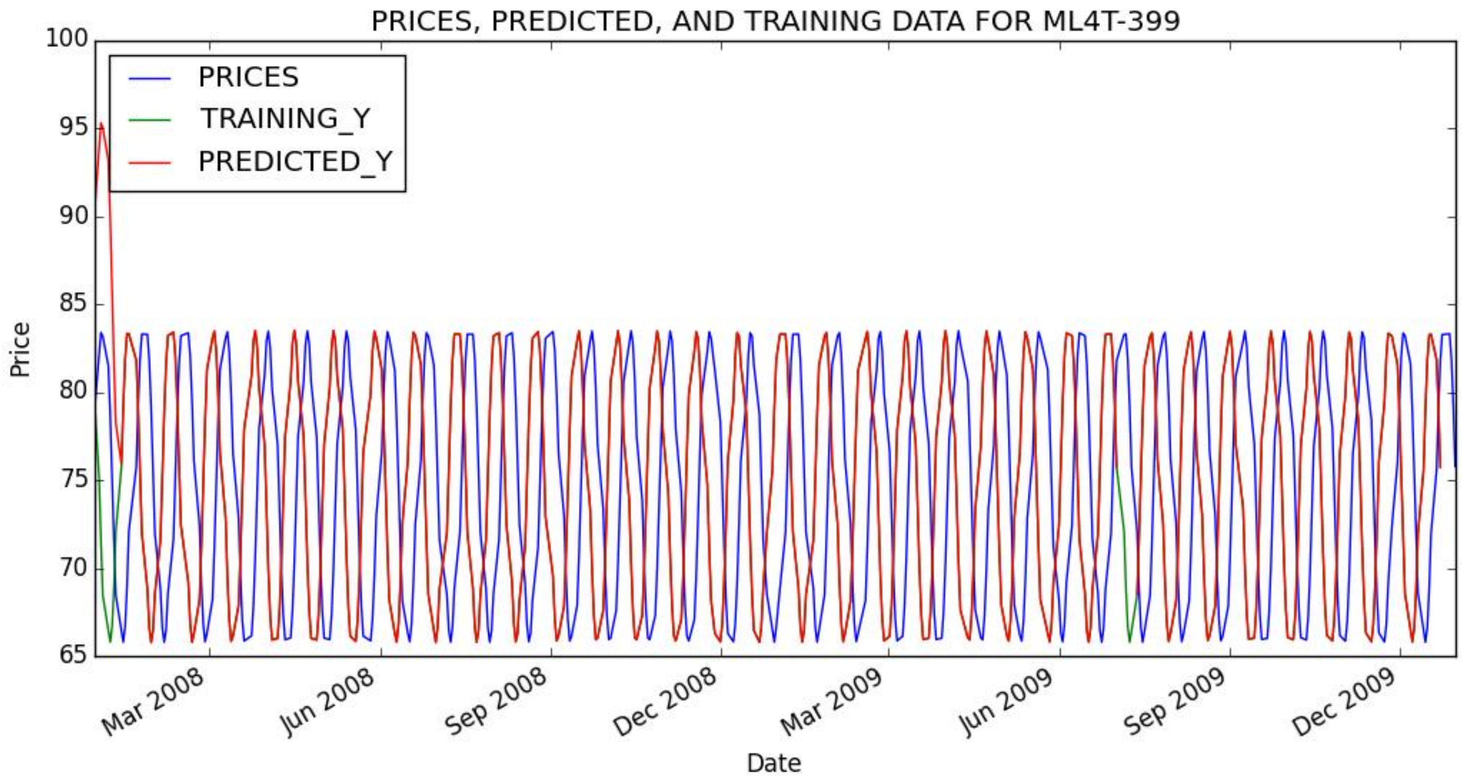
I began with a very basic strategy: Long entry if the predicted return was greater than one percent, hold for five days and then exit. Short entry if the predicted return was less than negative one percent, hold for five days, and then exit. For each entry/exit, I was trading one hundred stocks with a starting portfolio value of one million. When I tested this strategy on the in sample (for sine and IBM), it performed decently, but the returns were not tremendous, at all. The same was true when I tested the strategy on the out of sample sine data. I experimented a bit with changing the numbers around; i.e. entering if the predicted return was greater than two percent, holding for fewer days, etc. I did not notice enough significant variance. So then, I hit upon a simple solution to greatly increase profits. As long as the strategy was yielding consistent profits, and not suffering any losses – then I could simply trade more shares! Voila. I could get up to three hundred percent returns on the in samples, and do extremely well on the out of sample sine, if I traded 5000 shares per entry/exit. There. I had my strategy.

Then, I tested the out of sample IBM… and saw that I was *losing* money! So the more shares I traded, the more I was losing. Apparently, this strategy worked on the predictable, patterned sine data. It also worked on the in sample IBM data, since it had been trained on it. But for the out of sample IBM data, which does not follow a predictable pattern like the out of sample sine data, it was relying too heavily on the predicted five day returns. That is to say: the fact that a predicted five day return is above one percent, most definitely indicates that the price is on the rise (and vice versa for below one percent). However, to assume that the price would indeed rise to that predicted point precisely over the course of five days, was trusting the learner too much. (The reason the out of sample sine did so well is because the sine is so patterned and predictable. In fact, it is interesting to note that out of sample sine does even better than in sample IBM. Obviously this is a function of the sine pattern, and not just the learner.) So I decided to try holding for a shorter amount of time. In other words: I was trusting my forecaster that the price would increase. It is reasonable to assume so; after all, the forecaster is basing its query results on momentum, volatility, and Bollinger values, all of which are sound technical indicators. It has been trained well enough so that its results are reasonably reliable. But I am not completely trusting that my learner's five day prediction is perfectly accurate, and therefore I will not wait five days to sell. Instead, I now hold the stock for three days (and sell at the end of the third). This resulted in terrific performances for in sample sine, in sample IBM, and out of sample sine. The out of sample IBM also did pretty nicely now. I did try holding for different amount of days. Holding for one day less did not perform as well, and neither did holding for one day more (i.e. sell on fourth day). So apparently, the learner's predictions should best be trusted to hold for three days. Here are the cumulative returns for all the data sets:

Starting Portfolio Value: 1,000,000
Shares Traded per Entry/Exit: 5000
Long Entry: predicted five day return greater than one percent
Long Exit: after holding position for three days
Short Entry: predicted five day return less than negative one percent
Short Exit: after holding position for three days

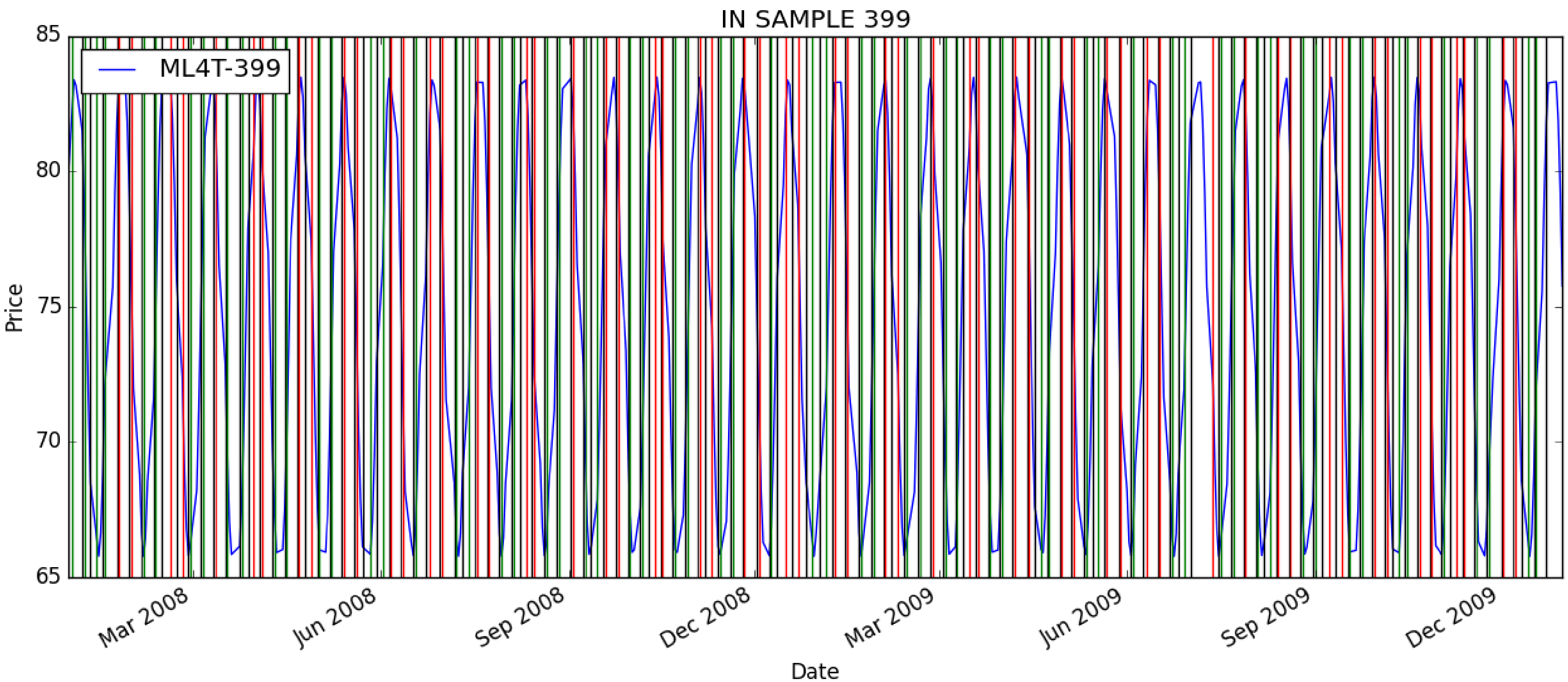
Cumulative Returns:
In Sample Sine: 3.68614941
In Sample IBM: 0.745
Out of Sample Sine: 1.674681565
Out of Sample IBM: 0.12335

Sharpe Ratio:
In Sample Sine: 11.1110494691
In Sample IBM: 3.41857994481
Out of Sample Sine: 11.0632991099
Out of Sample IBM: 1.36629685473

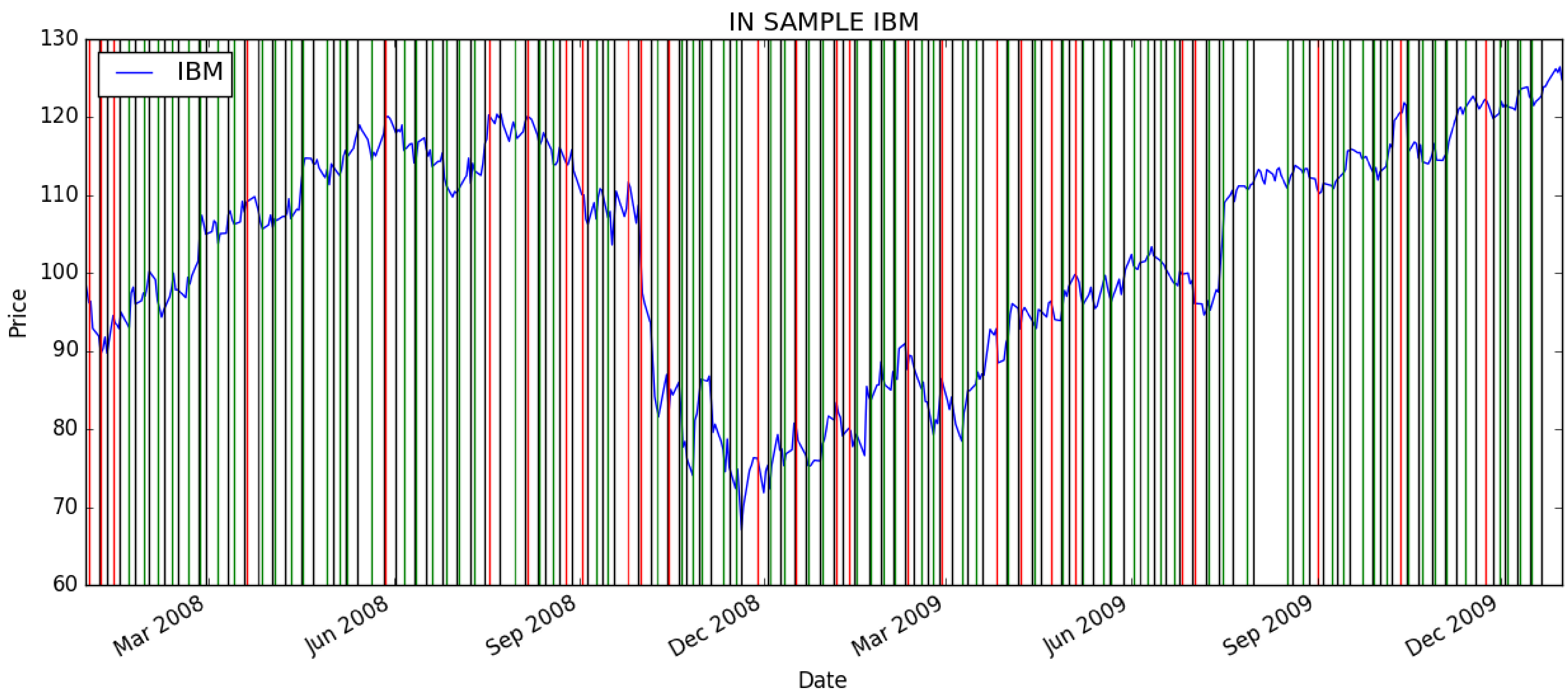


Note that the above chart has a peak at the start of the predicted Ys. This is because at the start of the date range, there is no data on which the prediction can be based, and so that data will inevitably be inaccurate. (As per conversation during office hours with Depriviya and Ukshat, this feature of the graph would not result in a grading deduction if noted and explained.)

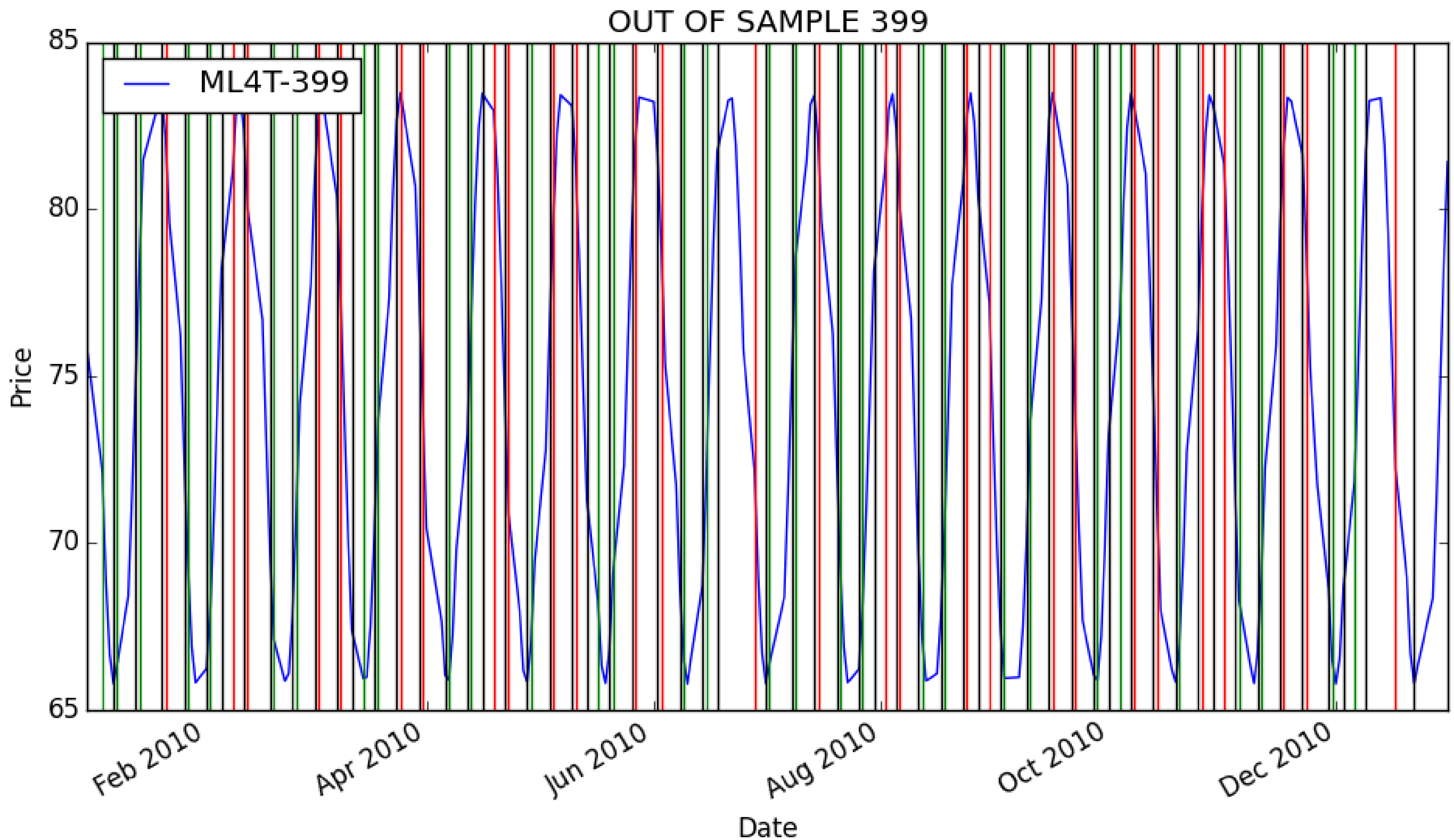
IN SAMPLE 399 STRATEGY CHART:



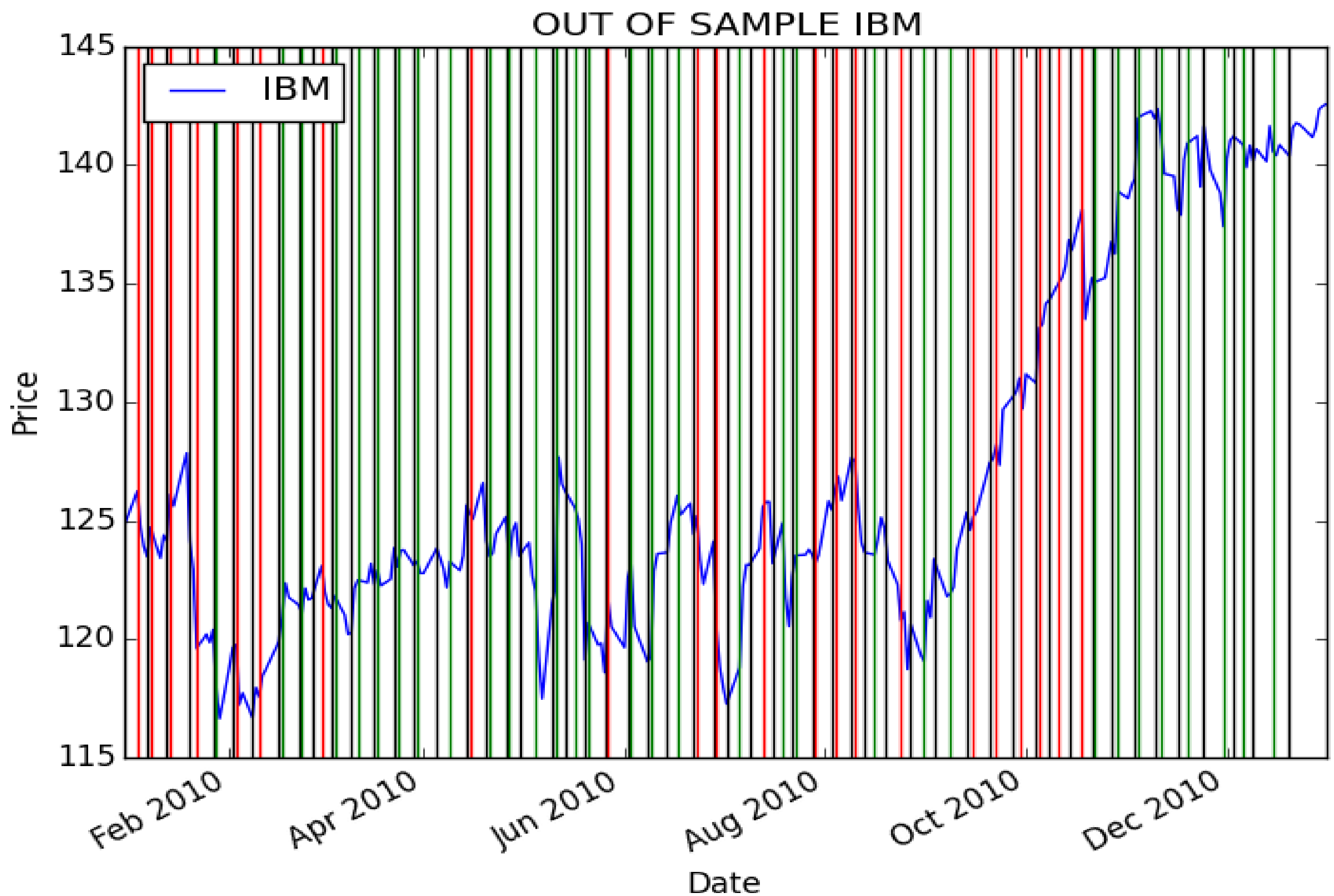
IN SAMPLE IBM CHART:



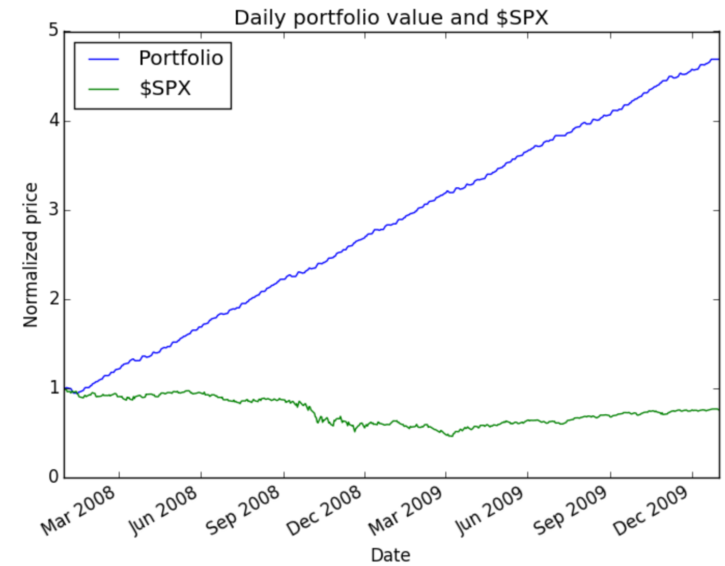
OUT OF SAMPLE 399 CHART:



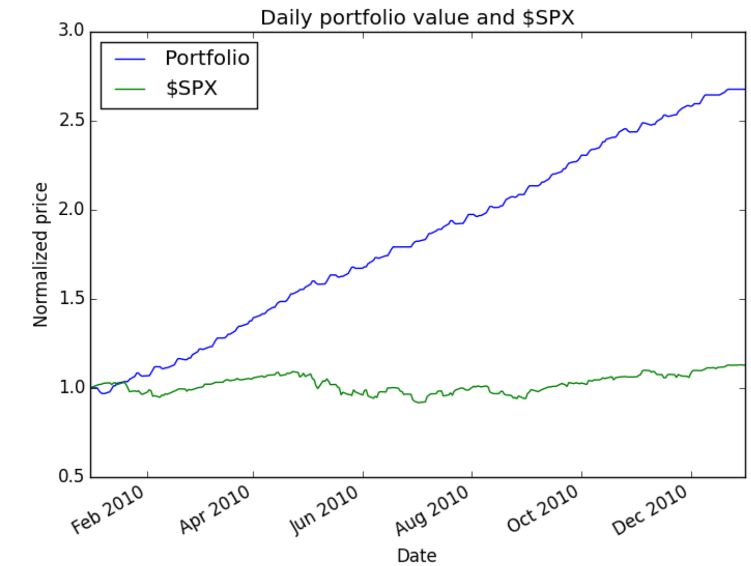
OUT OF SAMPLE IBM CHART:



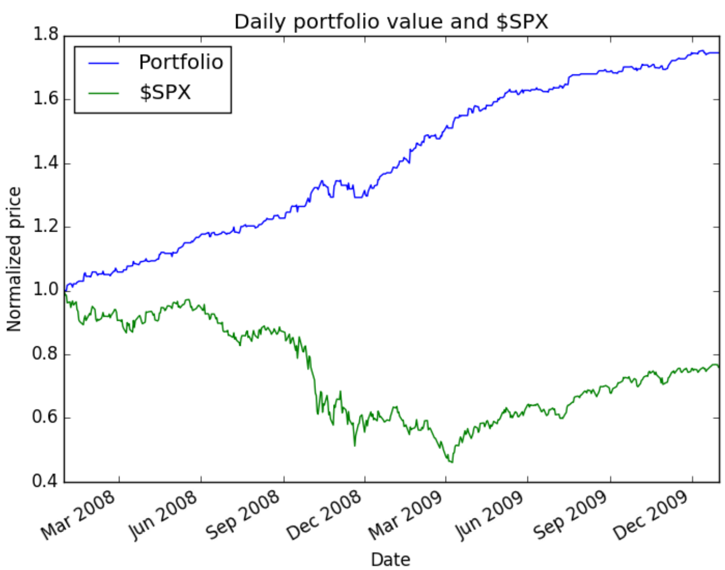
IN SAMPLE 399 BACKTEST



OUT OF SAMPLE 399 BACKTEST



IN SAMPLE IBM BACKTEST



OUT OF SAMPLE IBM BACKTEST