

# CS 7646 001- MC3 P2

glD : nghosh9

## ML4T-399

Indicators Used – Learner used was KNN (N=3)

- Bollinger Bands** – I have used Bollinger band indicators to find a value when the price goes above the 20d SMA and then normalized it so that it takes in values between (-1,1)
- Momentum** - I have used 5d momentum values and scaled the same so that the value is between (-1,1)
- Volatility** – I have used volatility (rolling std deviation) over a 5d period and scaled the values as above.

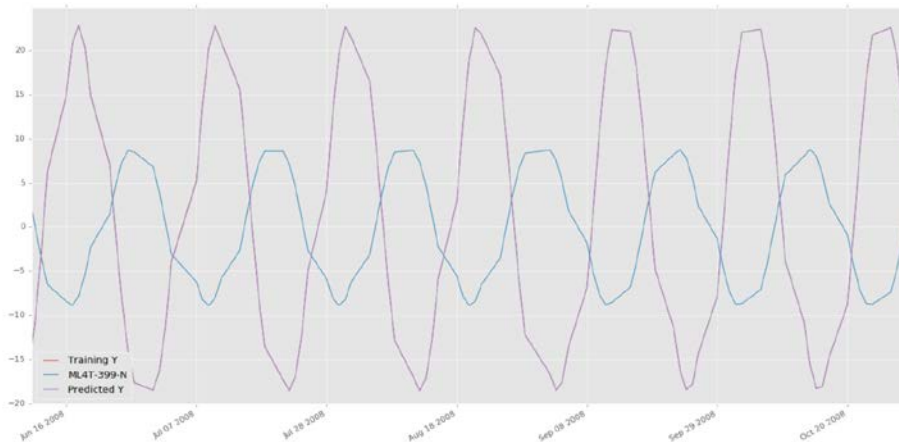
### Trading Policy

**Long Entries/Exits** : I have used a trading policy where I enter long when the predicted 5d return is positive and exit the moment I see that the 5d return is negative.

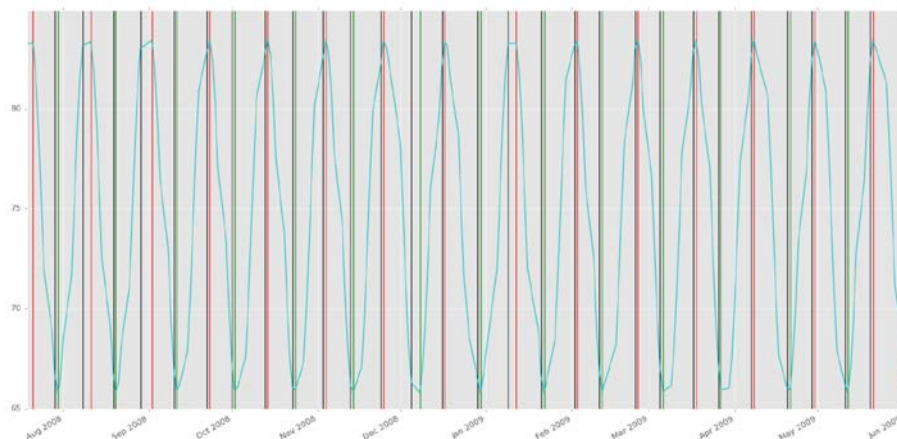
**Long Entries/Exits** : I have used a trading policy where I enter short when the predicted 5d return is negative and exit the moment I see that the 5d return is positive.

### Plot 1 : Training Y/ Price / Predicted Y

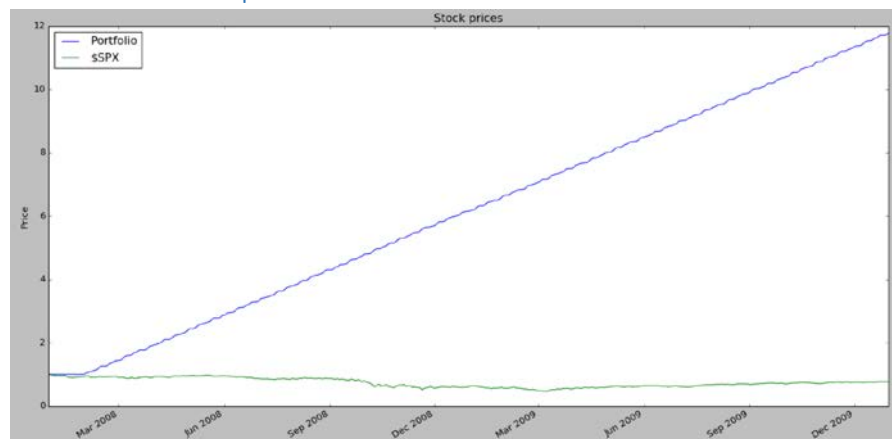
I have scaled the data such that they all hover around the y=0 axis



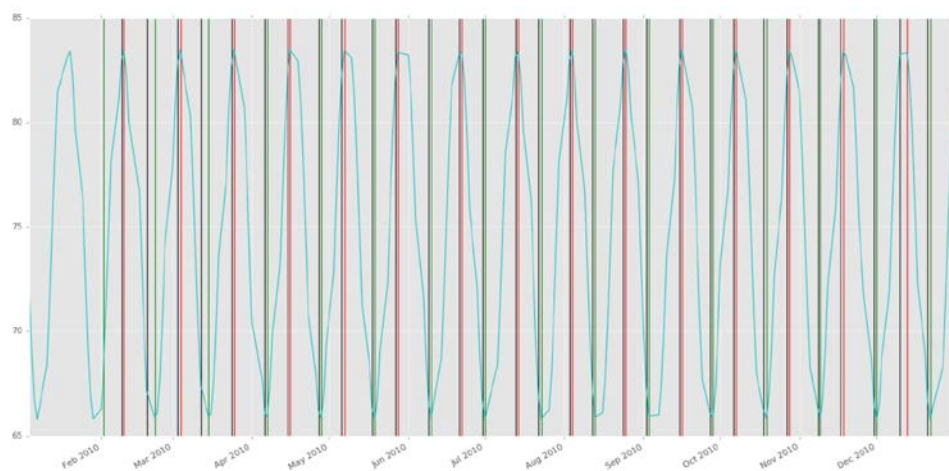
### Plot 2 : Sine Data In Sample Entries/Exits – 2008-09



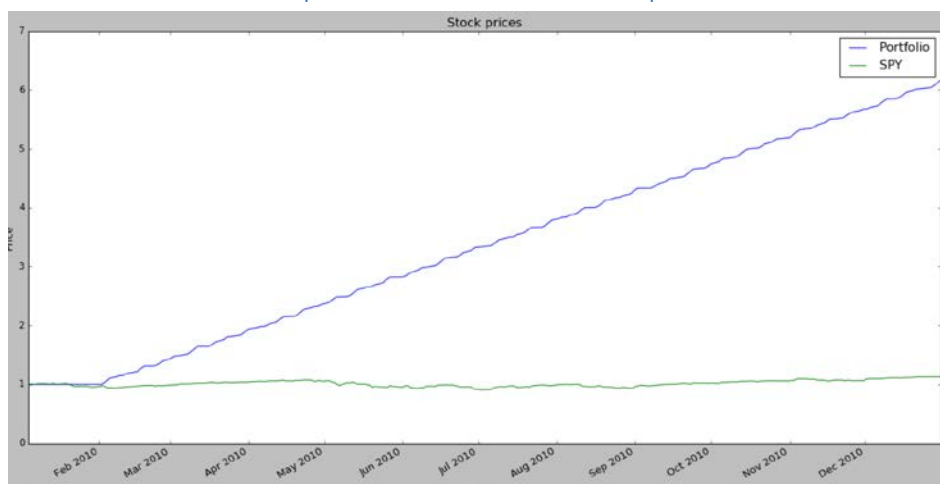
Plot 3 : Sine Data In Sample Backtest – 2008-09



Plot 4 : Sine Data Out of Sample Entries/Exits: -2010



Plot 5 : Sine Data Out of Sample Sine Data Out of Sample Backtest - 2010

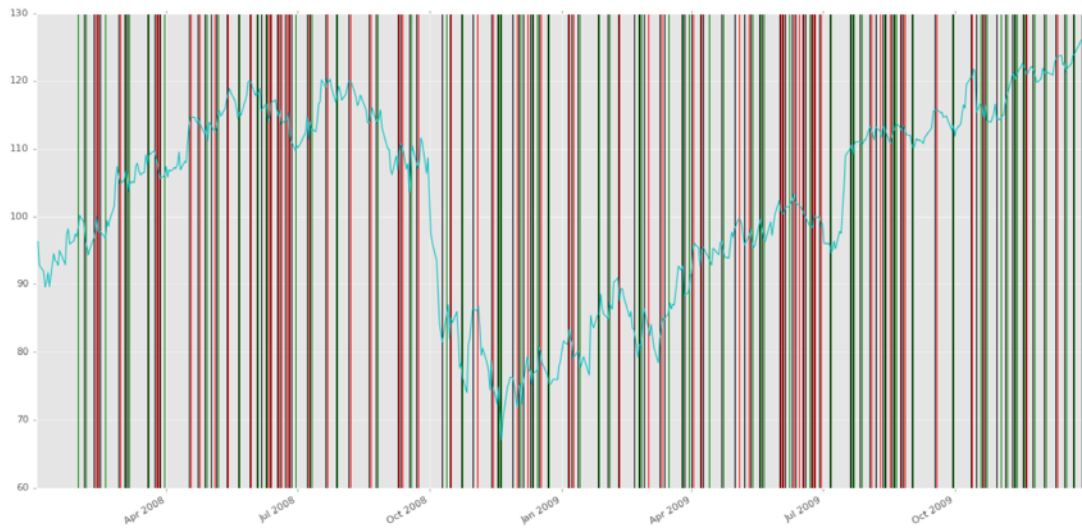


## Results :

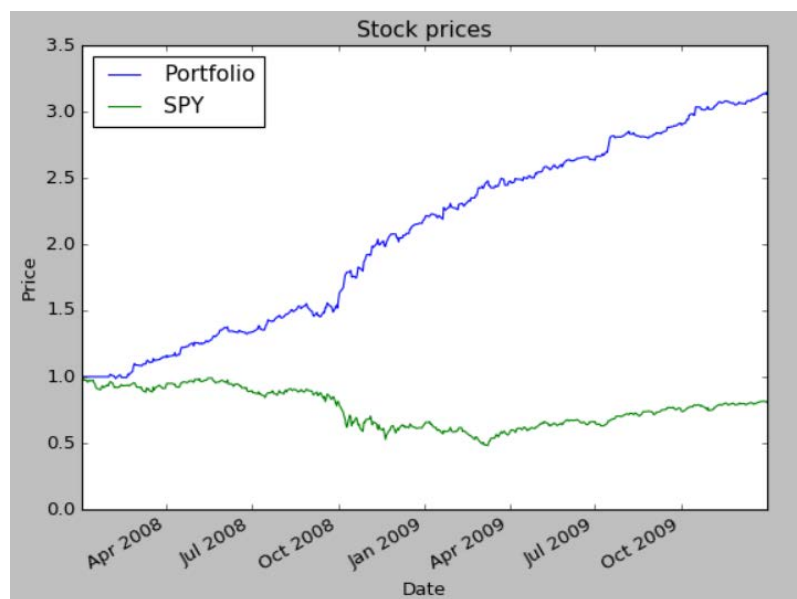
The results come out really well (6 fold increase) and we see continuous increase in the portfolio percentage over time as well as it performs much better than SPY. This is expected as the stock values are taken from a sine curve which is very “predictable” by the KNN regression learner.

## IBM

Plot 6 : Sample Entries/Exits – 2008-09

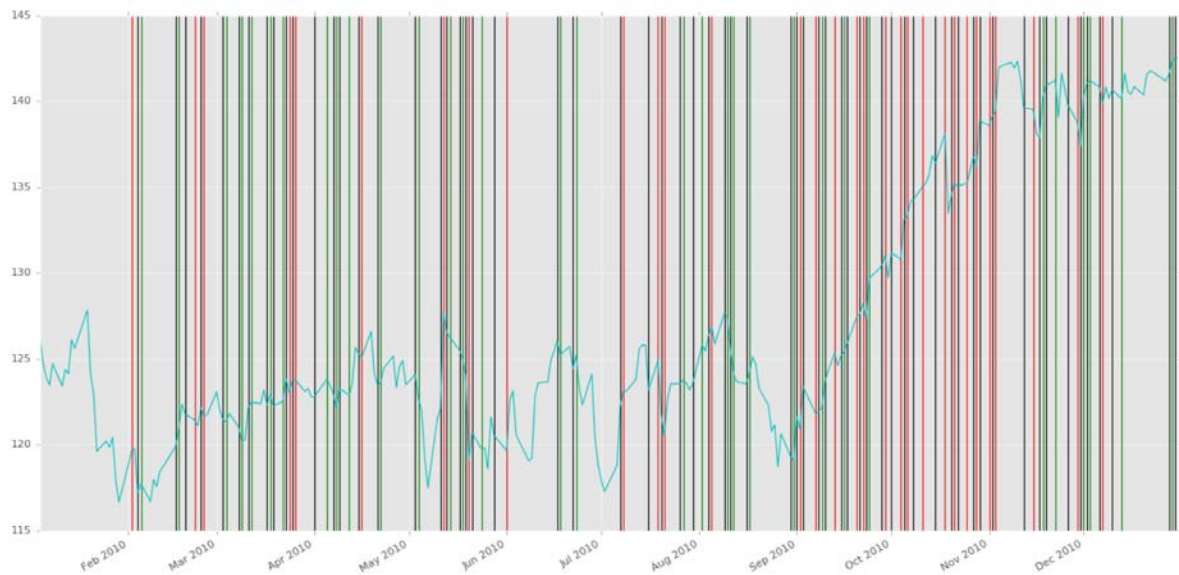


Plot 7 : In Sample Backtest– 2008-09

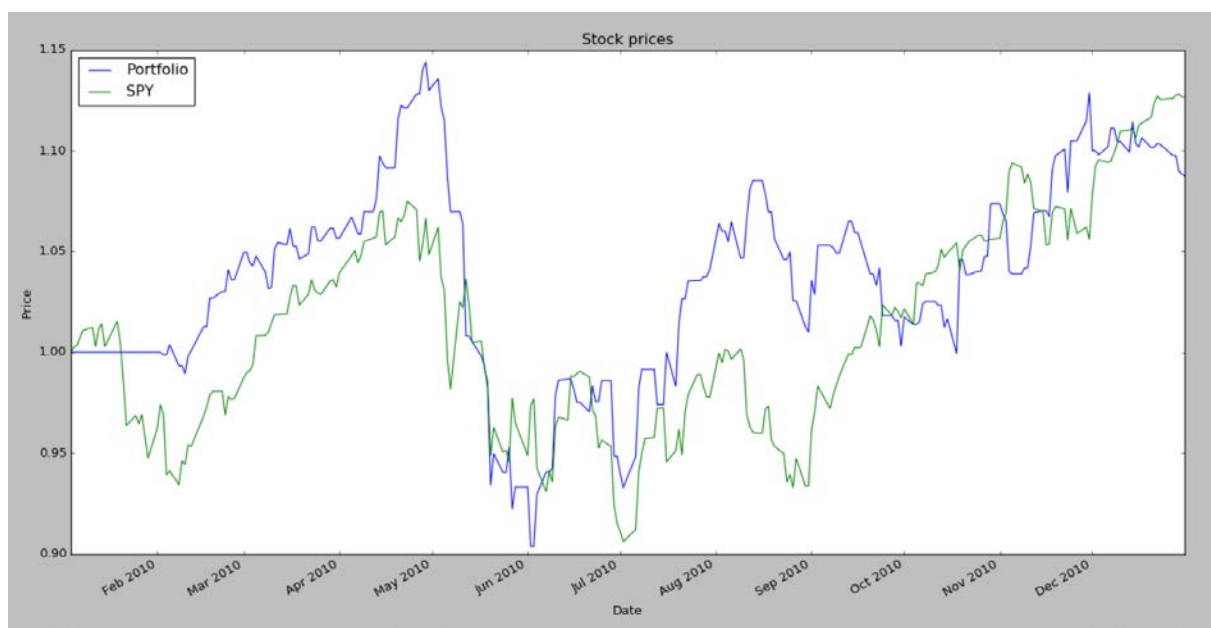


The in sample backtest performed pretty well with **3 fold returns**.

Plot 8 : Out of Sample Entries/Exits – 2010



Plot 9 : Out of Sample Backtest – 2010



### Results :

The results for IBM during the period 2010 did not come out too well and the portfolio value gained just 7% as opposed to the SPY which made around 12%. This could be because of the following reasons.

- a. IBM's stock price started rising sharply to levels (around Aug-Sep 2010) that were not foreseen in the training data. The indicators were not suited to predict such a scenario

As a different approach

- a. Test with more indicators/learners (in different combinations) to see which of them give consistently better results on out of sample data.
- b. Use bagging to average out RMS errors as much as possible.
- c. Personally I believe we should look to use volatility in a more refined way and not just the plain values as used here.