Vincent La

Georgia Tech ID (vla6)

Strategy Learner Report

# Introduction

*Describe the steps you took to frame the trading problem as a learning problem for your learner. What are your indicators? Did you adjust the data in any way (discretization, standardization)? Why or why not?*

To frame the trading problem as a learning problem, I first took the data frame of prices for the stock. The overarching idea is that we are given a data frame of prices and we need to return a prediction for whether we need to BUY, SELL, or STAY (where stay is keep the same position as the previous day. To do this we need to do two steps:

1. Feature Engineering: How can we convert the prices data to features with values for each day to make the prediction?
2. Transforming the Data to create a Y value: What is the value we actually want to predict? How do we create that?

**Feature Engineering**

For feature engineering we created technical indicators as we did in the ManualStrategy assignment. The indicators that we use are:

1. Bollinger Bands
2. SMA (Simple Moving Average)
3. Rate of Change (ROC) aka Momentum
4. EWMA (Exponential Weighted Moving Average)

**Creating Y Value**

To create Y value we calculate the return in the training set after 20 days. We use 20 days as there are approximately 20 trading days in a month. This seems to be the standard in the industry.

Once we calculated the return in 20 trading days for each date, we then compared the return to constants, call them YBUY and YSELL. The definitions for these are as follows. YBUY represents a constant upon which if the N day return exceeds YBUY then you would want to take a LONG position. On the other hand YSELL represents a constant upon which if the N day return is below YSELL then you would want to take a SHORT position. In any other cases, we keep the same position as the previous day.

**Learner**

For the actual learner we used the RTLearner that we implemented in a previous assignment. We used a leaf size of 5. Furthermore we set YBUY and YSELL to 0.04 and -0.04, respectively. How we came to those values is that there’s a tradeoff in YBUY and YSELL in setting too low or too high. For example, if you set YBUY too low (eg. 0.00001) what’s happening here is that in some sense you are “overfitting” because in your training data you’re telling the algorithm that you want to take a long position even if the return in next 20 days is just barely positive. Thus, you’re adding noise to the dataset because in our construction the algorithm cannot necessarily tell difference between when return is very highly positive or just barely positive. Similarly, you can say the same thing about setting YSELL. However, you also don’t want to set YBUY too high or else you would never change your position and you have too low variance.

Finally, to summarize and synthesize, we use an RTLearner to predict whether or not we should take a long or short position using the technical indicators described earlier as features and setting the y variable mapping 20 day returns to long or short decisions based on comparisons to YBUY and YSELL.

# Experiment 1

*Using exactly the same indicators that you used in manual\_strategy, compare your manual strategy with your learning strategy* ***in sample****. Plot the performance of both strategies* ***in sample*** *along with the benchmark. Trade only the symbol JPM for this evaluation.*

**Description of Experiment (Assumptions/Parameter Values)**

In this experiment, we test three portfolio values:

1. Benchmark (Buy and Hold JPM starting 2008-01-01)
2. Manual Strategy (This is the same as the manual strategy we used in the previous assignment. We will describe more below, but the manual strategy essentially uses Bollinger bands to inform buy/sell decisions)
3. Strategy Learner (This uses RT Learner as described in Section 1 of this report).

For our manual strategy we will use the Bollinger Bands technical indicator to guide our Sell or

Buy decisions. In particular:

1. Start by buying 1000 shares of JPM on the first day.
2. When the price hits the lower band, we buy.
3. When the price hits the higher band, we sell.

In terms of Assumptions and Parameter Values:

1. In Sample Start Date: 2008-01-01
2. In Sample End Date: 2009-12-31
3. Commission = 0
4. Impact = 0
5. Symbol traded = ‘JPM’

# Experiment 2

If impact is lower, then better results and more trades. If impact is higher, worse returns and lower trades and more closely matched to benchmark.