**Capstone Project** Theo Lekkas

Machine Learning Engineer Nanodegree June 2016

**Definition**

**Project Overview**

In this project I look to determine whether or not the Relative Strength Index (RSI) is an effective tool for buying & selling stocks. First I begin with a definition of the RSI:

“The relative strength index (RSI) is a technical momentum indicator that compares the magnitude of recent gains to recent losses in an attempt to determine overbought and oversold conditions of an asset. It is calculated using the following formula.”[[1]](#footnote-1)

The RSI is supposed to be able to determine both good buy & sell points for stocks (it is also used in trading futures, but I will restrict my analysis to stock for this project.) The purpose of this project is to determine whether or not buying a stock when it’s oversold & then selling the stock when it’s oversold is a profitable strategy. Below, I will more clearly state the parameters of this problem & further clarify various definitions.

**Problem Statement**

This project intends to determine whether trading strategy of going long (buying) a stock when it is considered oversold using the RSI is a viable (profitable) strategy. Profitability, however, is not enough for a viable strategy. We will also want a comparison against some benchmark; the reason for this is to determine whether or not it was better to simply invest in the broader market rather than deploying a specific trading/investing strategy. For this project I will use the returns to the S&P 500 over the time period used for this strategy. There are other more sophisticated metrics that look at the volatility of the strategy (such as Sharpe ratios), but for two reasons I will not be looking at these: 1) for the purposes of this project I am only interested whether or not RSI is really predictive and nominally better than the broader market 2) I don’t much care about volatility, to quote Warren Buffet “I would much rather earn a lumpy 15% over time than a smooth 12%.”[[2]](#footnote-2)

The buy signal for a stock will be when the RSI indicates oversold & the sell signal will be when the RSI indicates overbought. I will be using daily prices. The buy will be made on the closing price & the sell will be made on the closing price as well. In the initial attempt I will not consider commissions, slippage, or other market frictions; these can be added later if the strategy actually seems viable. Furthermore, I will also look at various moving averages, which the RSI is dependent upon. I will also look at different levels of oversold & overbought (these will become clearer below when the RSI formula is broken down.)

I will use two different machine learning models, decision trees & logistic regression, to see if there is predictive value to this strategy.

Below I will more clearly define RSI and various elements/concepts that will be used throughout the project.

**Metrics**

The RSI is a momentum indicator that attempts to determine short-term bullish (positive) & bearish (negative) positions for stocks. The RSI looks at the price action of a stock and does not consider other elements (such as sentiment, news, accounting metrics, etc.) The RSI is a pure “technical” indicator, its only concerned with the price movement of the instrument being analyzed.

The formula for RSI is as follows:



Created with <http://www.sciweavers.org/free-online-latex-equation-editor>

As can be interpreted from the formula above, the RSI is an index between 0 & 100.

The RS is the relative strength metric, it is defined as average of up days over n periods divided by the average of down days over n periods. There are a couple of ways to calculate the averages; smoothed moving average[[3]](#footnote-3) or exponential moving average[[4]](#footnote-4). Below is formula for the RS using a smoothed average: [decide which to use first 🡪 both being tested could be beneficial]



Created with <http://www.sciweavers.org/free-online-latex-equation-editor>

The U & D in the above formula represent ‘Up’ & ‘Down’ which are calculated by subtracting the current close with the previous days close[[5]](#footnote-5).

Finally, we need to define the meaning of Oversold & Overbought. These are somewhat arbitrary, but for the initial runs I will use the traditional definitions. Oversold on the RSI is < 30 and Overbought > 70. These two definitions are definitely ripe for fine tuning, but we will use the traditional definitions at first.

**Analysis**

**Data Exploration**

**Exploratory Visualization**

**Algorithms and Techniques**

**Benchmark**

**Methodology**

**Data Preprocessing**

**Implementation**

**Refinement**

**Results**

**Model Evaluation and Validation**

**Justification**

**Conclusion**

**Free-Form Visualization**

**Reflection**

**Improvement**

1. Relative Strength Index - RSI http://www.investopedia.com/terms/r/rsi.asp [↑](#footnote-ref-1)
2. 1996 Chairman's Letter - http://www.berkshirehathaway.com/letters/1996.html [↑](#footnote-ref-2)
3. https://en.wikipedia.org/wiki/Moving\_average#Modified\_moving\_average [↑](#footnote-ref-3)
4. https://en.wikipedia.org/wiki/Moving\_average#Exponential\_moving\_average [↑](#footnote-ref-4)
5. https://en.wikipedia.org/wiki/Relative\_strength\_index#Calculation [↑](#footnote-ref-5)