**2012312366 Kyu Chul Kim**

**Why use CART (decision Tree) Algorithm?**

What is the purpose of trading strategy? Index such as Sharpe Ratio shows that people’s major interest on investing is “Earning a lot of money at minimum risk”. Therefore, I thought that the goal is to predict the return of investment, and the decision should be made on data-oriented way. However, considering that investors do not give credit to such model considered “Black Box” – investor cannot inspect the process of making decisions – I will use Machine Learning Algorithm called Decision Tree. Decision Tree, or CART, basically by iterating binary split on each independent variable, forms a tree – like model that classifies the observations. In our case, the daily return can be classified, or predicted by binary splitting technical indicators. CART (Classification and Regression Trees) has following advantages. First, since the decision tree grows by splitting observations according to criterion of variables, the model can be literally explained by variables. Second, with the process of “pruning”, the model sets simplest rule possible. These two advantages make the model a “white box”, thus it can be convincing strategy for investors, and at the same time gives flexible predict without complex assumptions such as linear regression. I will not cover details of the algorithm in this paper.

**Two criteria for reducing universe: Volume and Absolute Sharpe Ratio**

In addition to CART algorithm, I will filter the target securities with following two criteria. First, mean volume size of the security. The special lecturer of Robo-Advisor company Qraft stated that when they let their algorithm to work without any limits, it tends to recommend small-cap securities and makes great performance. Under this heuristic, I assumed that there should be some characteristics decided by the volume of security. With this assumption, I will order SP500 securities by its mean volume, and divide it into 5 intervals, then sample 51 securities from each interval. The number 5 and 51 are just arbitrarily chosen. Second, absolute Sharpe Ratio. Since excess return is usually not an extreme number (in the millions or greater), therefore the ratio mainly depends on the standard deviation (volatility).[[1]](#endnote-1) Thus, I used absolute Sharpe Ratio to avoid securities with high volatility.

**Technical Indicators**

So with the filtering process, I sampled 255 securities out of 503 according to its mean volume, and then chose top 50 securities with high absolute Sharpe Ratio. Now with 50 securities, I generated technical indicators for each of them. I utilized 3 types indicators so that my model can catch different types of signal, or rule of deciding the class of observation. And used variety of indicators for each type so that each security can use the fittest indicator. 9 Moving averages (SMA(20days), LMA(50days), EMA, DEMA, EVWMA, EVWMA, ZLEMA, ALMA, HMA) are used to indicate the smoothed moving direction, 2 Strength indicators (RSI, CMO) are used to indicate magnitude of variation and 2 volume indicators (OBV, chaikinAD) are used to indicate money flow into the security.

**Modeling Process**

The modeling process is done for each security. Therefore, the process iterates 50 times.

What we want to know form the data, is whether the return will be positive or negative depending on indicators’ value. The “Class” of data will be “Up” if the daily return is positive, else, the class will be down. Since the indicators are bind to the data with 1 period lagging, it makes sense that the classification based on indicators is a matter of “prediction”.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Time Stamp | OHLCVA | Indicators | Class Predict | Class true |
|  |  |  |  |  |

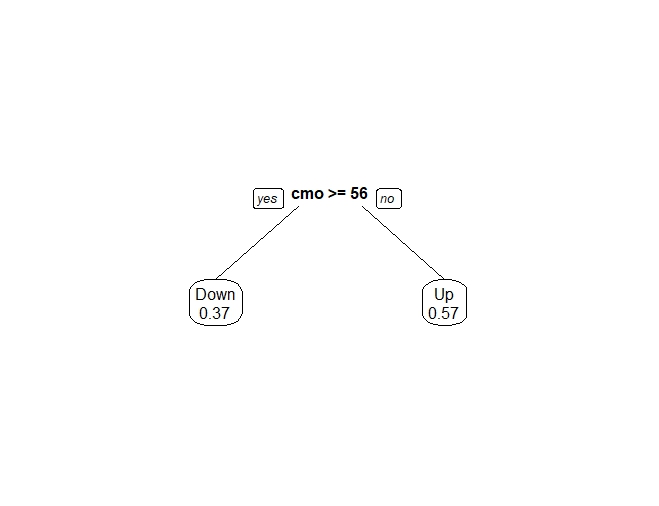
I though 3 is the appropriate number of indicators for the simple model. And the indicators should be uncorrelated. If the variables are correlated, the Linearly Dependent variable is redundant and gives only confusion for the modeling and interpreting. Therefore, I made correlation coefficient matrix for 14 indicators, and chose all possible non-correlated sets of 3 indicators. Let the number of sets is *n*. With Stratified Random partitioning, I’ve split data randomly into train set and test set. Purpose of doing this is to prevent the predict model is overfitted only to the given data. With train data, I made *n* models of decision Tree. And with test set, I calculated the accuracy of each model. The accuracy is yielded as . Among *n* models generated, choose the model with maximum accuracy. Since our class is binomial, I thought the model should be better than just randomly choosing the class. Therefore, I have set the threshold as 0.57 which is slightly higher than 0.5 Now we have 15 securities and models.

**[Attachment1]** shows a simple tree of a security. We can interpret this strategy as if CMO is equal to or greater than 56, the observation will go down at 37% chance. Otherwise, CMO has 57% chance of going up. With this model, we can make signal such as to buy when CMO is under 56, and sell otherwise.

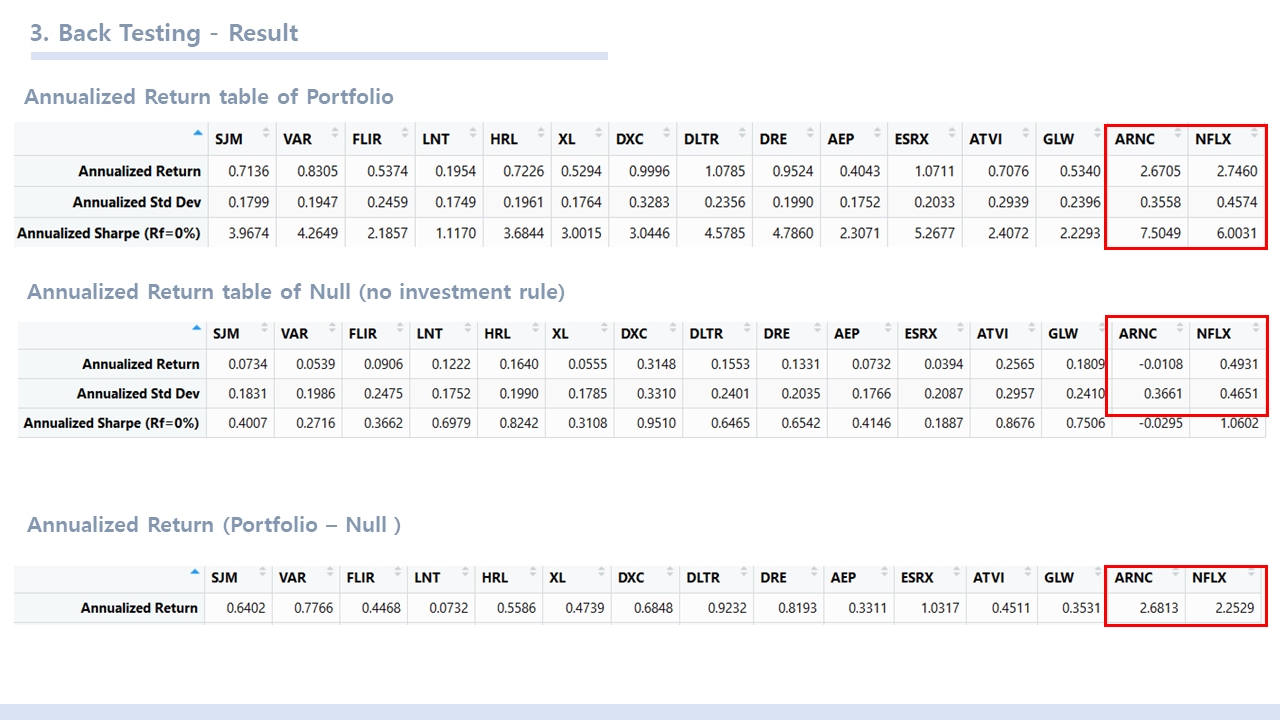
**[Attachment 2]** compares the 3 years annualized return of my 15 models to null status. Null status means that buying the security at starting point and sell at ending period. We can observe that ARNC and NTFLX show remarkable annualized return compared to null status. And both also has very high Sharpe Ratio at 7.5049 and 6.0031. With **[Attachment 3]** showing the drawdowns, cumulative return and hit ratio of ARNC, we can observe that cumulative return is much higher at 139.23 and has consistent upward slope. In addition, in contrast to Null’s steep (max -0.6 )and frequent drawdowns, portfolio show less frequent and shallow drawdowns ( max -0.2).

1. Sharpe Ratio Range of Possible Values. (n.d.). Retrieved December 13, 2017, from <http://www.macroption.com/sharpe-ratio-range/>

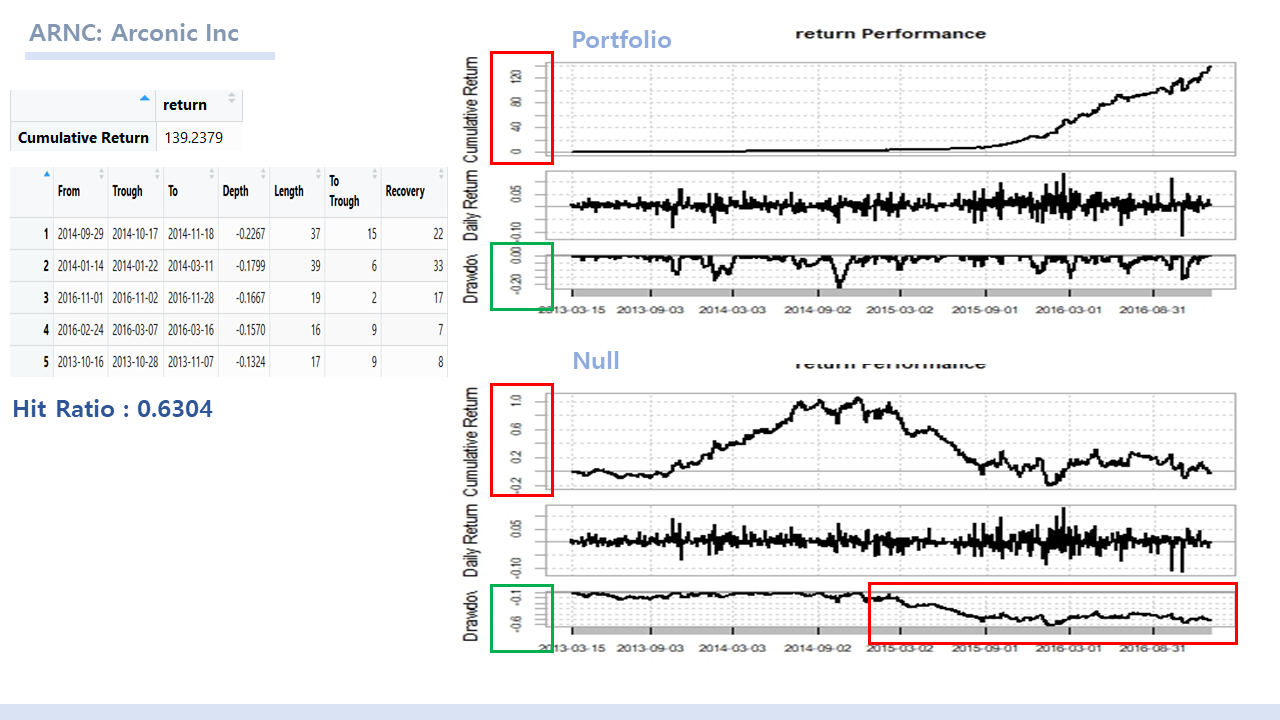
   **[Attachment 1]**

   ****

   **[Attachment 2]**

   ****

   **[Attachment 3]**

    [↑](#endnote-ref-1)