**World Quant University**

**Professor: Greg Ciresi**

**Alpha Design 3**

Erich Ratzat - [eratzat@hotmail.com](mailto:eratzat@hotmail.com)

Nikolas Lippmann Pareschi - [nikolaslippmann@gmail.com](mailto:nikolaslippmann@gmail.com)

**Final Project**

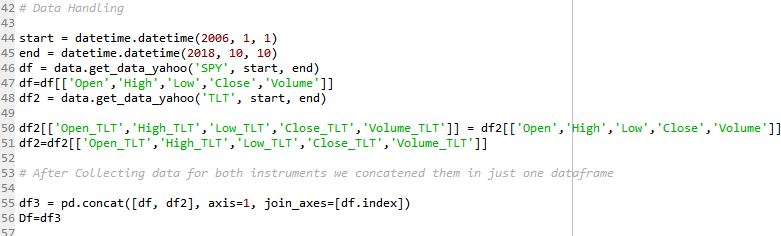
### Overview

In this project, in part 1 we will use a decision tree classifier and intermarket analysis to predict future returns on SPY and TLT. Afterwards we combine the predictions into one portfolio. The idea is to improve the sharp ratio, using the Treasuries (TLT ETF) to hedge the SP500 (SPY) due to the fact that these 2 instruments present negative correlation during sell offs.

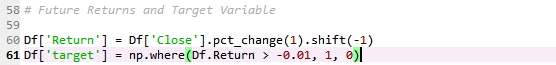
In part 2 we use Linear Regression to construct a band. If the price is above the predicted up price of the band we sell gold, the commodity chosen for this analysis. If the price is below the predict down price of the band we buy gold. The forecasting techniques (the predictors) used were RSI and Open/Close prices in Gold and RSI and Open/Close prices in US dollar. The US dollar is known as a big driver in the commodities world and also in gold.

PART 1 – Decision Tree Classifier – SPY, TLT and SPY Hedged:

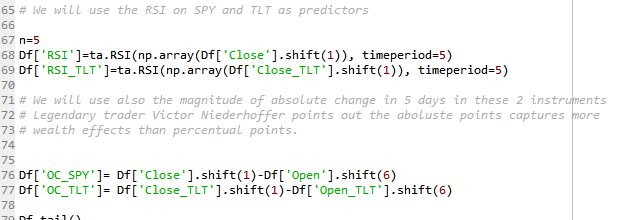
First, we download and handle the data:



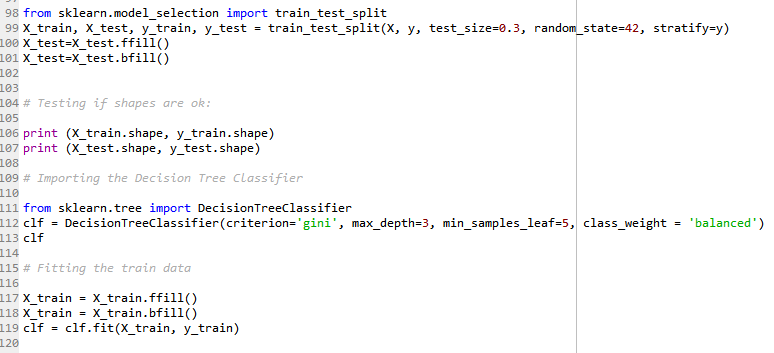
We compute the future returns. They are what we want to predict with the machine learning algorithm:



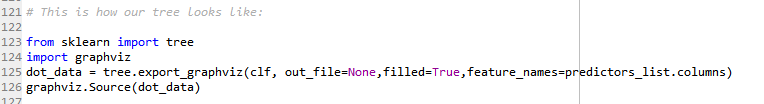
Below we have the predictors (RSI, Open and Close) that we will use:

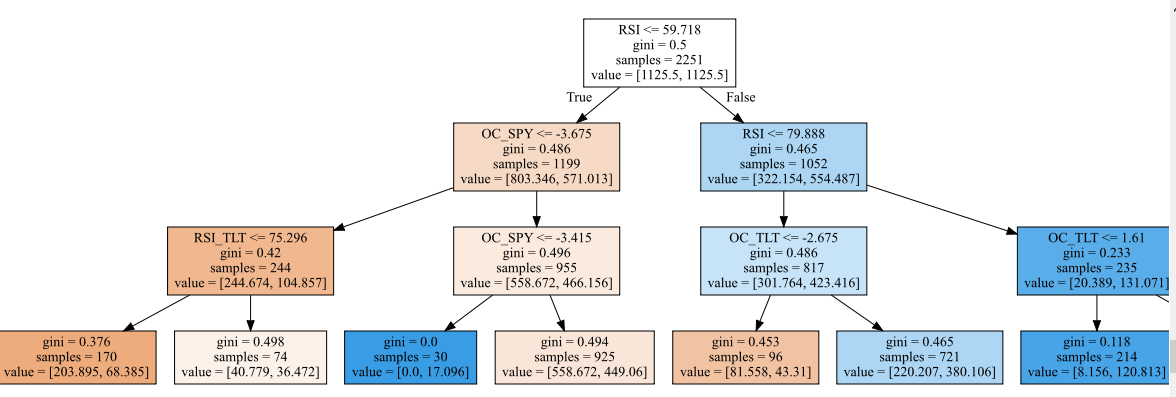
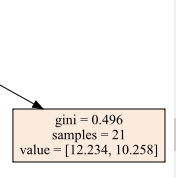


Next step is to train the classifier algorithm using 30% of the data:

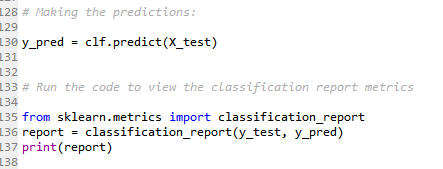


Bellow we printed the tree:

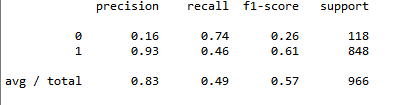


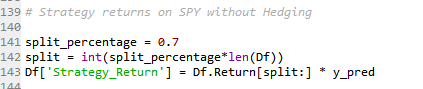
Now we make our prediction and we analyze its statistics:



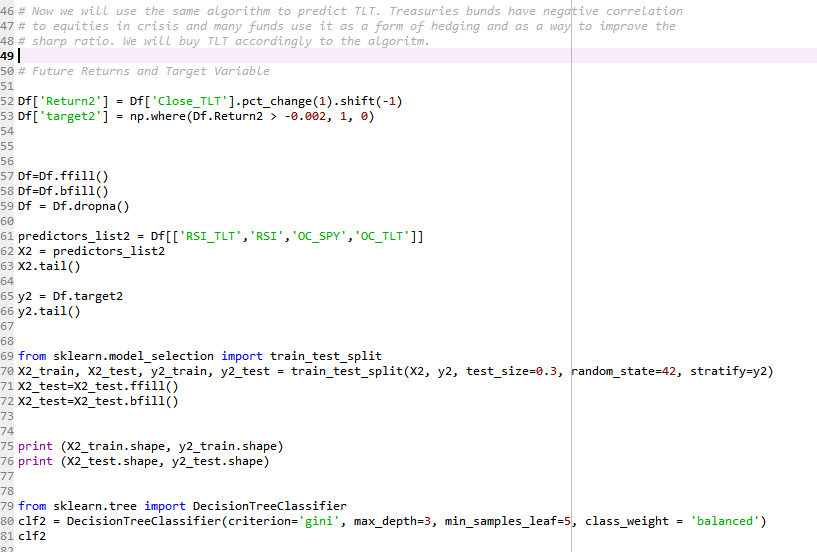
The tree is showing a f1-score higher than 0.5 which is very good and a 0.61 for predicting the days above the threshold 0f -0.005:



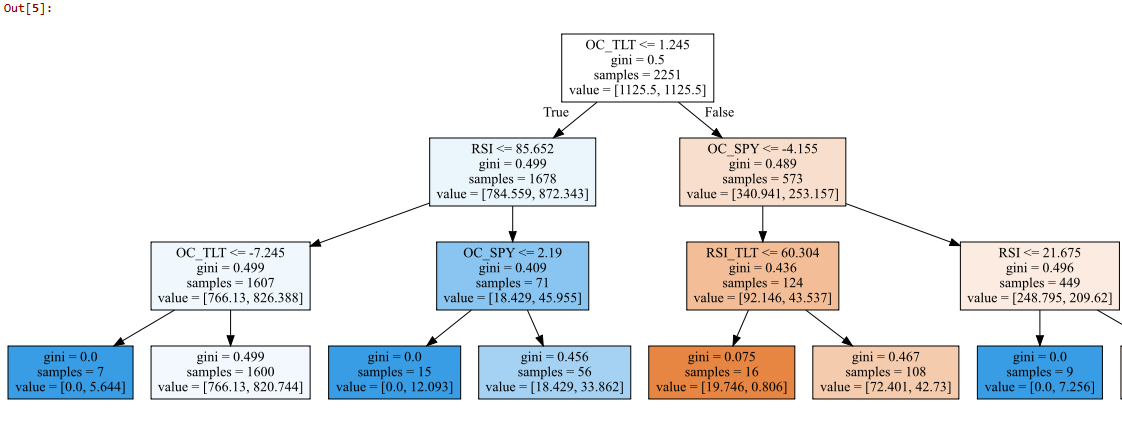
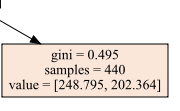
We compute the strategy returns:



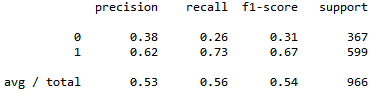
Now we do the same procedure, but now we want to predict the TLT. This is the code for the TLT strategy. Afterwards we combine both strategies:



This is the tree for the TLT strategy:

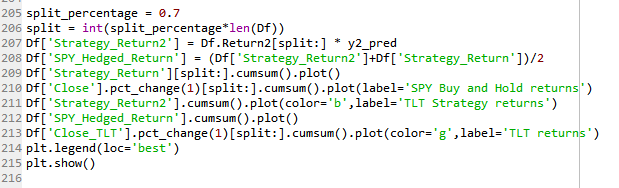
 

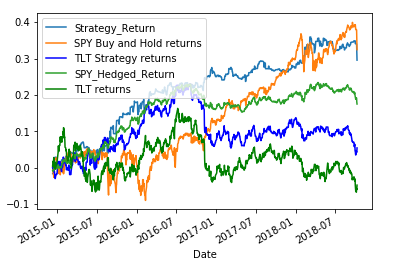
The F-Scores for the TLT strategy:



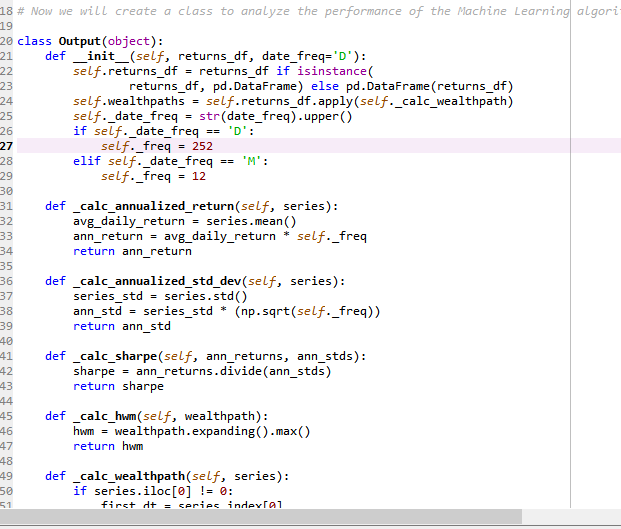
We got also a f1-score higher than 0.50 and we got a 0.67 score to predict the up days in TLT.

We then combine the strategies and plot the returns of: SPY buy and hold, SPY strategy, TLT buy and hold, TLT strategy, SPY hedged with TLT strategy:

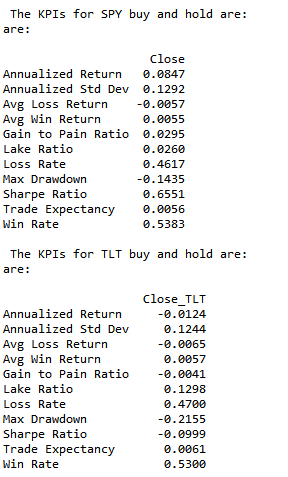
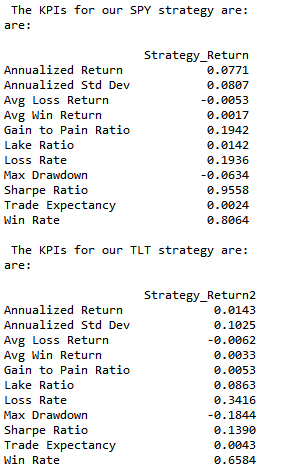


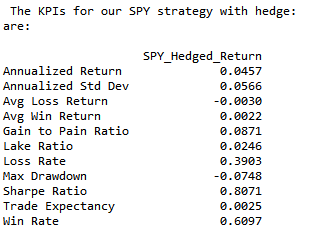


We then compute several metrics as Drawdowns, Sharp Ratio and others. This Class was provided by professor Steven Stelk in the course Risk Management:



And these are the results:

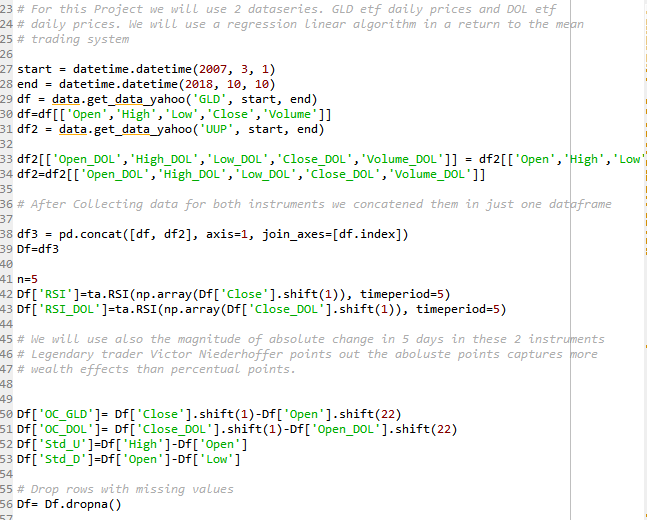


The SPY strategy and the TLT strategy have presented higher sharp ratios than the buy and hold in the out sample data. The portfolio hedged has cut the SPY drawdown in 50%, but was unable to deliver a superior sharp ratio.

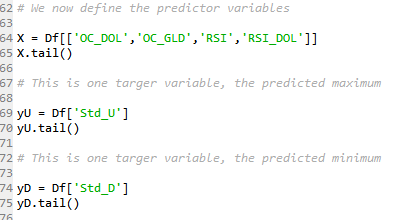
PART 2:

We used in part 2 the Machine Learning Linear Regression from sklearn to construct a band. The logic was: if the price is above the predicted up price of the band we sell gold, the commodity chosen for this analysis. If the price is below the the predicted down price of the band we buy gold. The predictors used were RSI and Open/Close prices in Gold and RSI and Open/Close prices in US dollar. The US dollar is known as a big driver in the commodities world and also in gold.

The code bellow shows how we downloaded the data and created the predictors used in the algorithm:



Defining the predictors´s list:



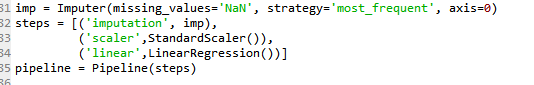
Note that we will make 2 predictions, the predicted maximum for the next day and the predicted minimum for the next day. If the price goes above the predicted up value we sell, if it goes below we buy.

In pipeline we will compute the following sequence:

1) The input function which deals with the NaN values;

2) the scaler that deal with asymmetries and outliers;

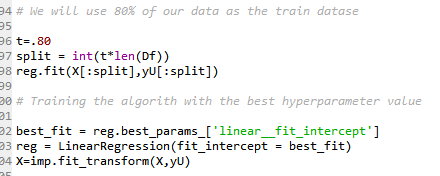
3) The Linear Regression algorithm that will be used to predict yU and yD:



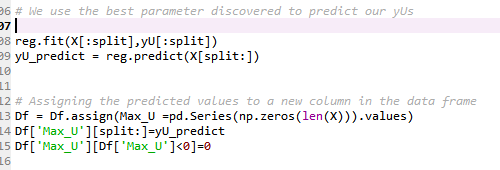
We define also the search space for the hyperparameter Intercept. The cross-validation method will be used in the training data with a number of 5.



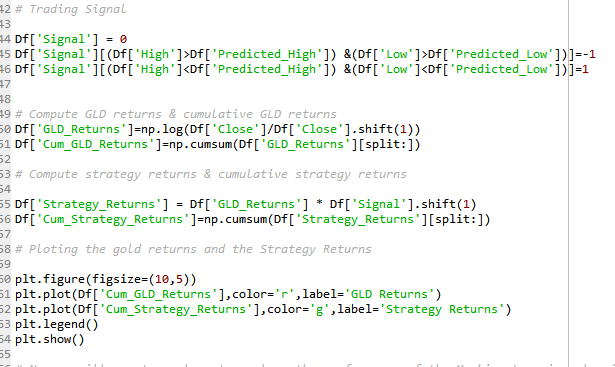
We used 80% of the data to train the algorithm and find the best hyperparameter value (the intercept):

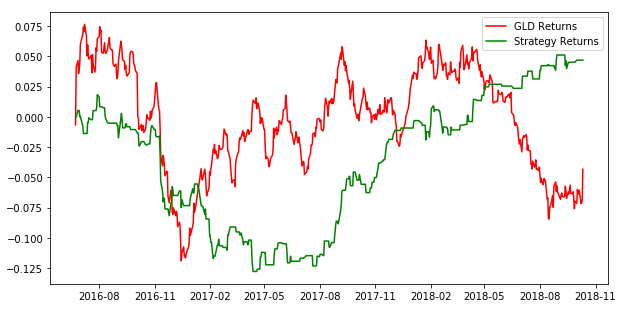


We proceed using the best parameter to predict the next day maximums:

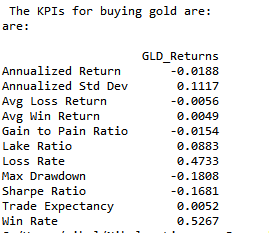
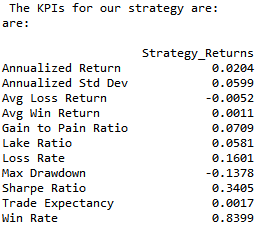


The calculation of the down band is analogous. With the predicted values we create our strategy and plot the graph of the results:





These are the KPIs provided by the code (same used in the Part I algorithm)

The strategy was superior in the out sample when comparing with a buy and hold strategy and important KPI’s such as drawdown and sharp ratio were also better. We were able to deliver some alpha, despite being low. Many managers use gold in their portfolio’s so the algorithm has indeed some value.