

## Home Work due Sunday, Feb. 9, before 12:00am

1. Insert a column in the data set where the entries are **1** if the stock outperforms SPY in the earnings period and **-1** if it underperforms or has the same return.

Create a **BaggingClassifier** with a **DecisionTreeClassifier** with **max\_level = 6** and **n\_estimators = 40** new column as labels.

Use the period **2007-2010** to train the model on the data set constructed above and construct a set of features that optimizes the profit of the long-short strategy on the quarter **2009-07-30 – 2009-10-30**

2010 ?

Backtest the performance of the model with the optimal features over the period **2010 – 2018**.

Compute **Sharpe Ratio**, **Information Ratio** and **alpha** for the strategy and for the buy-and-hold strategy for SPY

Change hyper parameters for the **BaggingClassifier** and the **DecisionTree** and try at least 3 combinations of max\_depth and n\_estimators. Find optimal features and backtest with each of the combinations of hyperparameters.

Take the **best performing hyperparameter combination** and run the code that changes optimal features each quarter and compute the performance characteristics of the strategy over the time period above

2. Insert a column in the data set with entries: **2** if the stock return is more than 5% higher than the SPY return, **1** if it is more than 1% but less than 5% higher, **0** if it is between -1% and 1%, **-2** if the stock underperforms the SPY by more than -5% and **-1** if the performance is between -1% and -5%

Same questions for this dataset as in problem 1 but where the **strategy** is to **invest**

? **\$x** where x is the predicted label