

FINM 33160 , FINAL PROJECT

From the nine optimal features files:

ada_opt_feat_4.pkl
ada_opt_feat_5.pkl
ada_opt_feat_6.pkl
gb_opt_feat_1.pkl
gb_opt_feat_2.pkl
gb_opt_feat_3.pkl
opt_feat_4.pkl
opt_feat_5.pkl
opt_feat_6.pkl

Each of these files contain a list of lists of optimal features for various classifiers.

1 Compute the frequency for each feature which occurs as an optimal feature in the totality of these files.

2 Pick a set of 10 features with the highest frequencies and run several back tests with different sets of hyper parameters, using this set of 10 features through the whole period (so we do not change features each quarter) for each of the three classifiers below. Find the hyper parameters for each classifier from these experiments that give the best Sharpe Ratio

- A Random Forest Classifier
- An AdaBoosting Classifier
- A Gradient Boosting Classifier

3 Next use the best hyper parameters you found for each classifier and the 10 features to make a Voting Classifier (see the documentation for VotingClassifier in `sklearn`) and run a back test for this classifier, computing SR, IR, α and Max. Drawdown

4 { Pick a quarter from the period 2000-2018 and for each trading day in the quarter period (i.e. the day after an earnings release) compute the prediction from the Voting Classifier for each stock (with market cap > 1B) in the data set which can be traded on this date. Here we need the symbols and the earnings dates, which are in the data set (the ticker and date columns are dropped when we make the training DataFrames)

Compute the list of all trades predicted by the model.

5 { Compute the **total P/L** for the strategy where for each quarter we allocate our capital equally between the stocks we consider in the quarter, the stocks that are available will vary from quarter to quarter. We go **long each stock** with **prediction +1**, and **short each stock** with **prediction -1** (and **do nothing** if the prediction is **0**) and **hold until the day after the next earnings release**. We start with \$1.

6 { Consider the strategy where we **buy the stock** and **short the SPY** if the **prediction is +1** and **do the opposite** if the **prediction is -1**, with the same allocation to each trade as above and same holding period.

Compute **total P/L** for this strategy.

Compute **SR , IR , α , $MaxDrawdown$** (on an **annual basis**) for **both strategies**