Long-short trading strategy with StockTwits sentiment data

Joyce Yu Cahoon, North Carolina State University

Objectives

We were tasked with constructing a cross-sectional, long-short US equity strategy with the constraints:

- Trade liquid stocks
- No more than 5% of capital invested in any one asset
- No more than 10% net dollar exposure,
- Mean daily turnover between 5% and 65% over a 63-trading-day rolling window
- Gross leverage between 0.8x and 1.1x
- Low correlation to the market
- Less than 20% exposed to each of the 11 sectors as defined
- Positive returns

Trading Strategy

Universe Definition Liquid assets from `QTradeableStocksUS` that pass the following filters:

- Not trading within 2 days of any earnings announcements
- Not announced as an acquisition target
- Can calculate a 5 day moving average of the bull-minus-bear signal from the `StockTwits` API

Alpha Discovery

Ranked assets by its bull-to-bear intensity and found a set of portfolio weights that maximizes the sum of each asset's weight times this alpha value.

Trading

Each week, we calculate the portfolio that maximizes the alpha-weighted sum of our position sizes, subject to the constraints of max 1.0x gross leverage, limit 5% exposure to any asset, and max daily turnover of 80%.

References

Fan, J., and Yao, Q. (2015), *The elements of financial econometrics*, Science Press.; Hong, L., Harrison, and Stein, J. (2002), *Bad news travels slowly: Size, analyst coverage and profitability of momentum strategies*, The Journal of Finance.; Jegadeesh, N., and Titman, S. (2002), *Profitability of momentum strategies: An evaluation of alternative explanations*, The Journal of Finance.; Mackenzie, D. (2018), *Difference between specific and total returns.*; Smith, C. (2017), *StockTwits sentiment analysis.*; Quantopian.

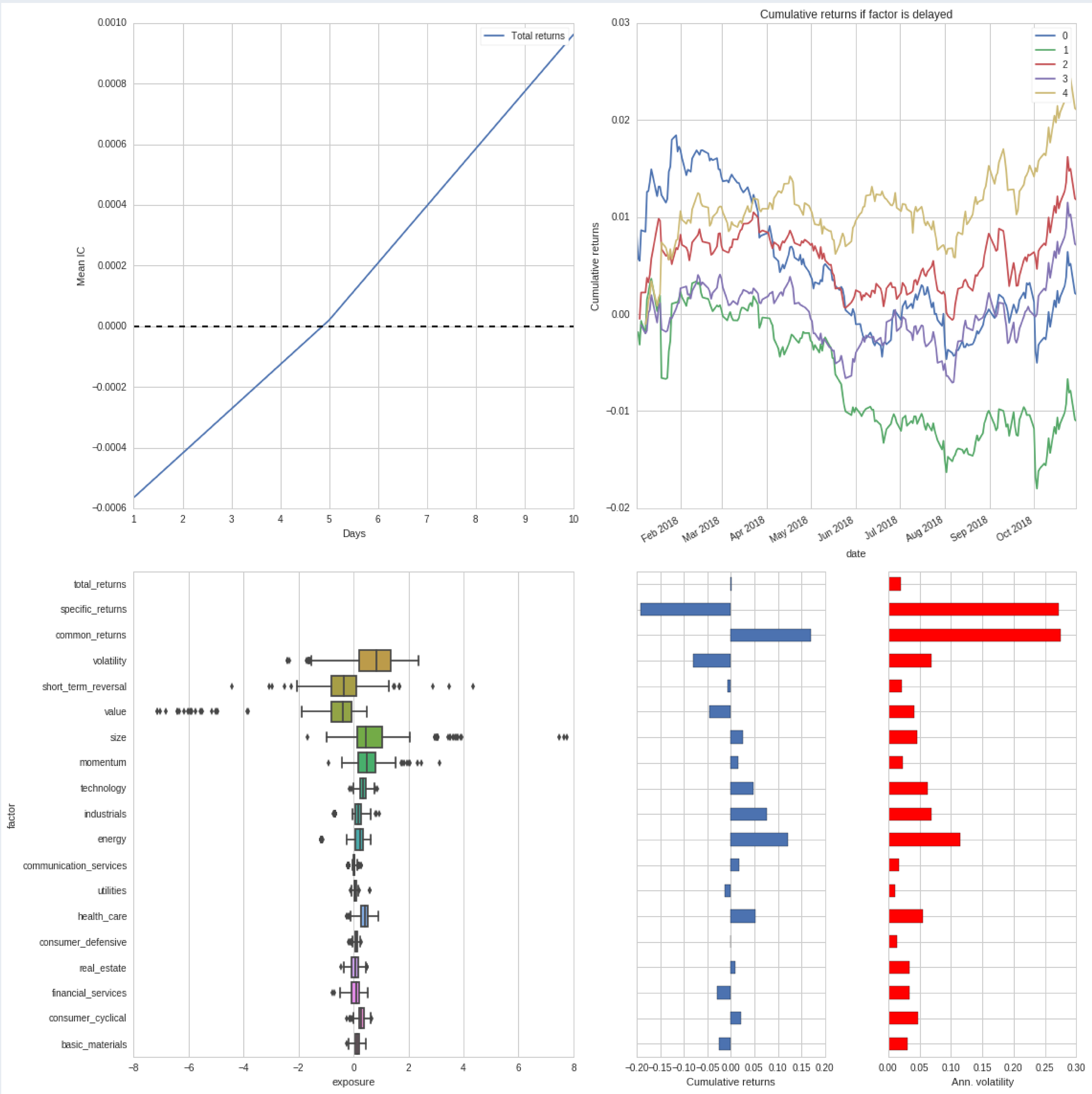
Sentiment Analysis

In addition to the ease of implementation, we chose to rely on the sentiment factor as it seems to have some good characteristics:

1. Predictive alpha: Calculated the mean information coefficient using the built-in function in `alphalens` and found our sentiment signal matches the direction of actual asset returns.
2. Low exposure risk: Quantified exposures via the `perf_attrib` function in `pyfolio`. Sentiment factor appears to have relatively low exposures throughout all categories.

Viability of our sentiment signal

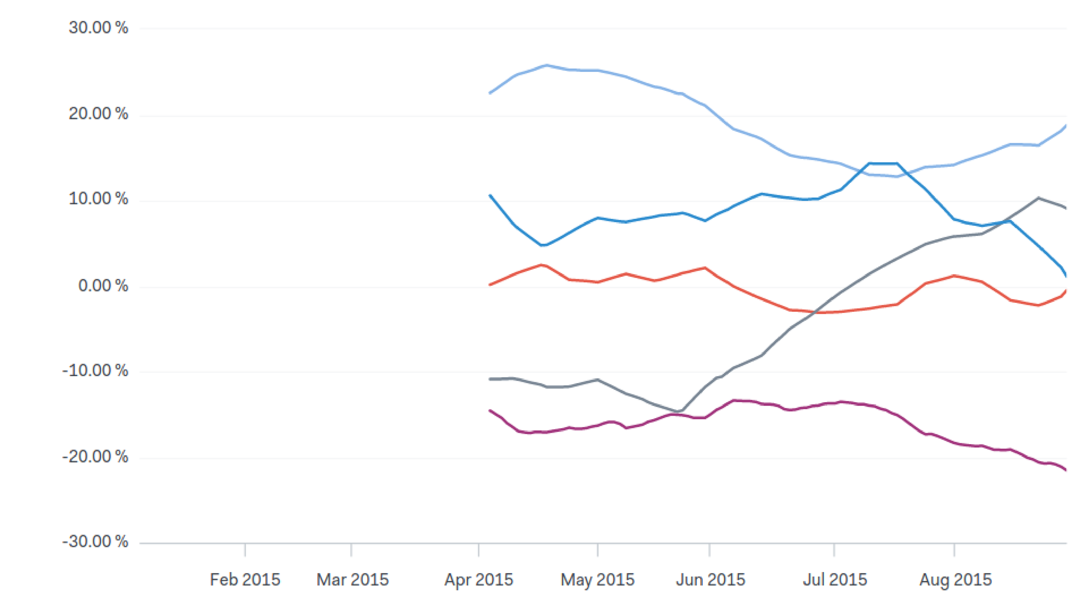
We do not know the natural language processing engine that calculates the bullish intensity and bearish intensity of a stock, but we do know how the `bull_minus_bear` signal arises; traders attached either a bull emoji or a bear emoji to any message they release on the `StockTwits` platform as well as a cashtag that identifies the asset under discussion.



Examination of (top-right) information coefficient averaged over all possible asset returns, (top-left) cumulative returns when factor is delayed, (bottom-left) risk exposures, and (bottom-right) cumulative returns and volatility.

Backtesting

Despite contradictory results from our data exploration, given the extensive literature that exists on sentiment data in predicting stock price, we forged ahead and backtested our trading strategy. Most of our returns are not only from specific returns, but:



Exposure to 5 different investing styles: momentum (light blue), volatility (dark blue), size (orange), value (gray) and short term reversal (purple).

Rolling beta to SPY calculated on a 6 month rolling basis. We sought for low beta, but this negative beta benefits us during our set trading window of November since the market has trended down.

Given the constraints we fed in our `optimize` API call, we are able to achieve (i) leverage between 0.95x and 1.05x; (ii) net dollar exposure between -2.00% and 2.00%; (iii) turnover rate between 17% and 20%.

Final Rank: #66