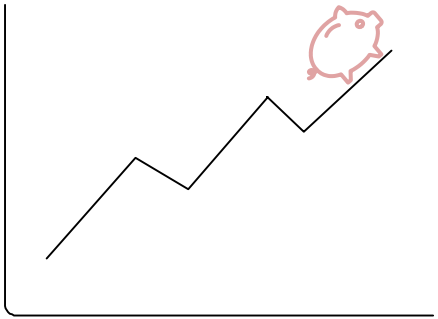
CrowdValue

A prediction model for “crowdsourced fundamental analysis”

Jason Rosenfeld



View full source code here (WIP):
https://github.com/jrosenfeld13/VIC-Fundamental-Analysis

The sum of a group of investors is greater than its parts

Can professional investors outperform the market?

- Yes!* (*not accounting for risk)
- Using investor forums, we can test this hypothesis
- We will use ValueInvestorsClub.com

Can we compile the theses of professional investors to perform even better?

- Crowdsourcing casts a wider net of investment write-ups
- A machine learning algorithm identifies the best ones

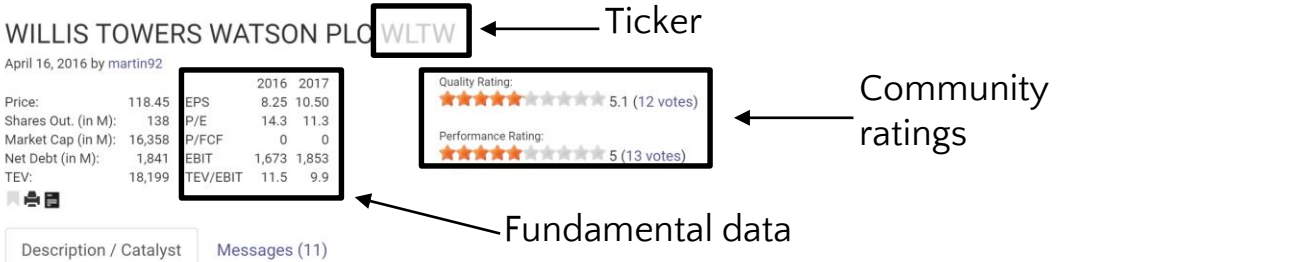
Previous “literature”: Do Hedge Fund Managers Have Stock-Picking Skills? http://www.retailinvestor.org/pdf/HedgeFund.pdf

ValueInvestorsClub.com (VIC)

What is ValueInvestorsClub?

- Private access, members only investment forum
- ~800 users
- Members must post at least 2 investment ideas per year
- Members vote on quality of other investment ideas and comment on write-ups

Typical investment write-up



Description

Thesis →

Willis Towers Watson (WLTW) is an attractive business - stable and defensive with opportunities for growth. It's run by a management team with a strong track record and trades at a reasonable valuation. Krusty75 posted a great write-up in December and I would refer you to that report (as well the thread), especially for a background on the business segments and merger developments. Since that time, the price has moved a bit lower and there have been a few new events. As such, I'm re-posting this idea for the VIC community because I think it's a compelling one.

WLTW is a global advisory, broking and solutions provider. The current business was created as the result of a merger of equals between Willis Group and Towers Watson that was announced in June of 2015 and closed at the beginning of January. The deal has striking similarities to the successful Aon/Hewitt marriage from 2010 and presents a number of opportunities going forward.

The new WLTW looks identical to its two peers, AON and MMC - just smaller. WLTW's sales and EBITDA of $8b and $2b, make it the #3 player in the insurance brokerage and benefits consulting business. AON and MMC have $13-$14b in sales and ~$3b in EBITDA. All of three players have a roughly 50/50 split between "risk" and "benefits" and all derive about 50% of sales from OUS. The main difference is that WLTW has much greater exposure to private healthcare exchanges, small today but a high growth future opportunity. Industry reports have pegged the WLTW opportunity at $2b (~25% of current sales) vs. $1b for AON and MMC (<10% of their current sales). This is why Towers traded a growthy multiple pre-merger.

Despite better opportunities for growth, WLTW trades at a discount to AON and MMC and an attractive absolute valuation. This is because insurance analysts had come to view stand-alone Willis as under-performer due to a subpar acquisition in 2008, and ongoing restructuring efforts that entailed big margin promises but have been to slow to show any results. Burned in the past, analysts are yet to fully appreciate the benefits of the merger going forward, and management is too conservative to fully voice them.

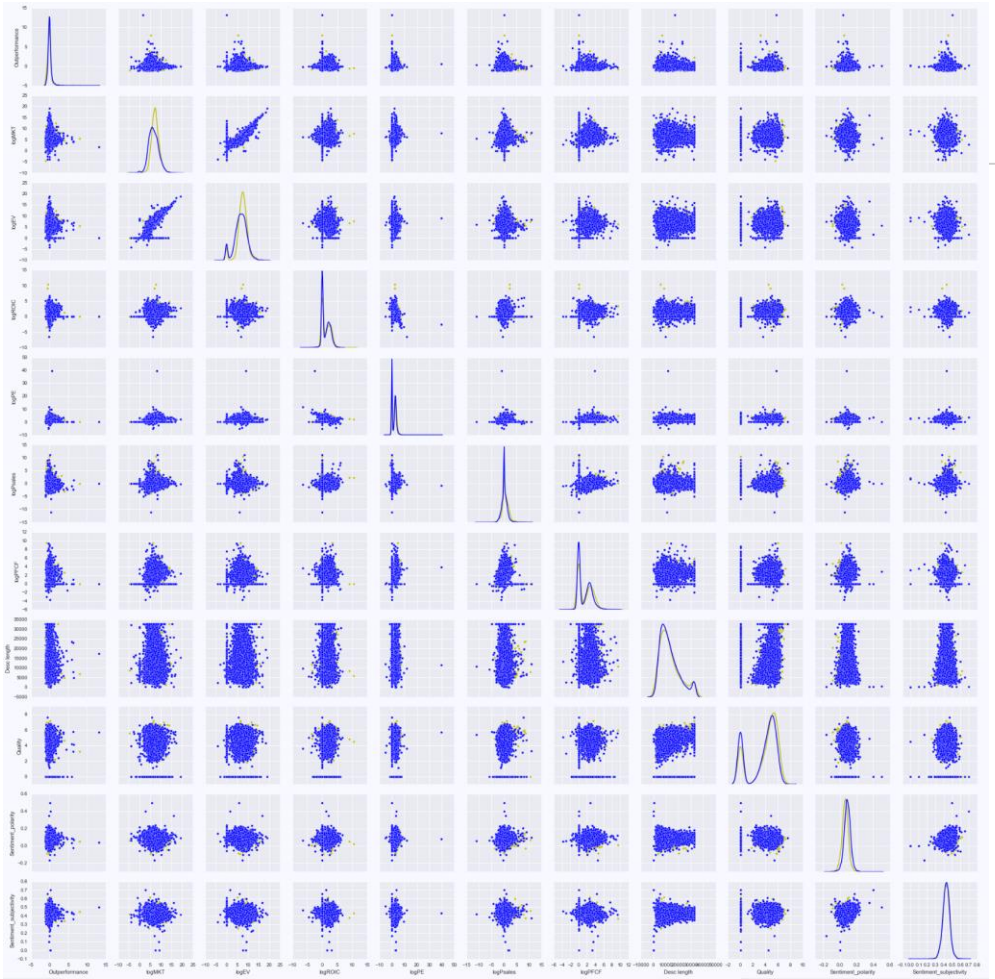
WLTW today trades at 14x cash EPS on 2016 (or 10x EBITDA and 15x NOPAT). In comparison, AON and MMC trade at 13x and 12x EBITDA. While I would argue that WLTW should trade

A treasure trove of data

~5,000 investment write-ups

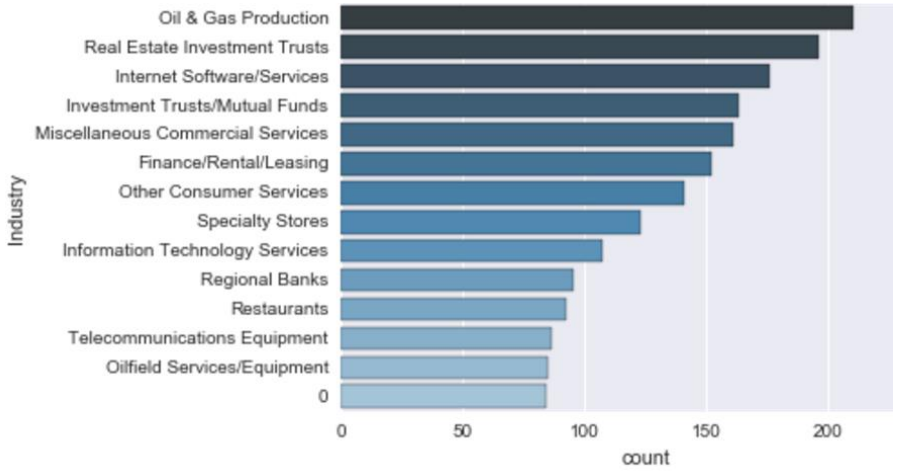
- Data includes:
  - Market cap, Enterprise value
  - Write-up text (i.e. investment thesis)
  - Valuation metrics (e.g. P/E, ROIC)
  - One year performance vs. S&P500
- Bi-modal normal distributions
- Often random correlation

…let's take a closer look

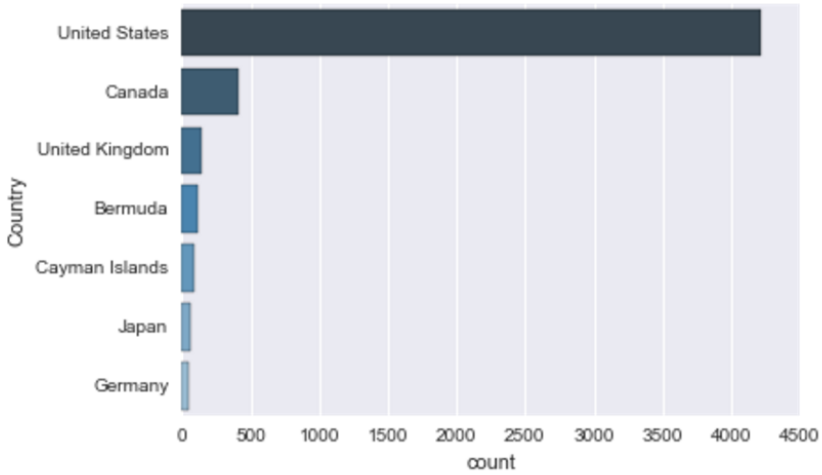


Quick facts about the write-ups

Industry



Country



VIC members outperform the market, on average

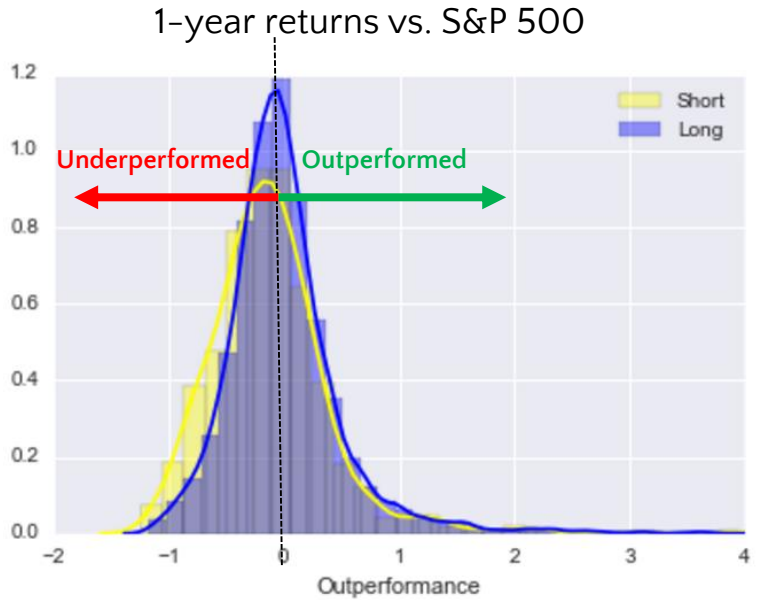
Buy and hold for one year

Long Write-ups (blue)

- **Mean outperformance: 0.88%**
- Percent outperformed: 42.9%
- Write-up sample size: 4,285

Short Write-ups (yellow)

- **Mean outperformance: (11.5)%**
- Percent outperformed: 34.12%
- Write-up sample size: 800



Identifying VIC's best ideas

What makes one investment idea better than another?

- We will only look at longs
- We already have the data
- Let's use machine learning to do the dirty work

Our model can identify succesful write-ups

Using machine learning, we were able to identify likely outperformers as soon as they were posted, improving average outperformance by almost 10%.

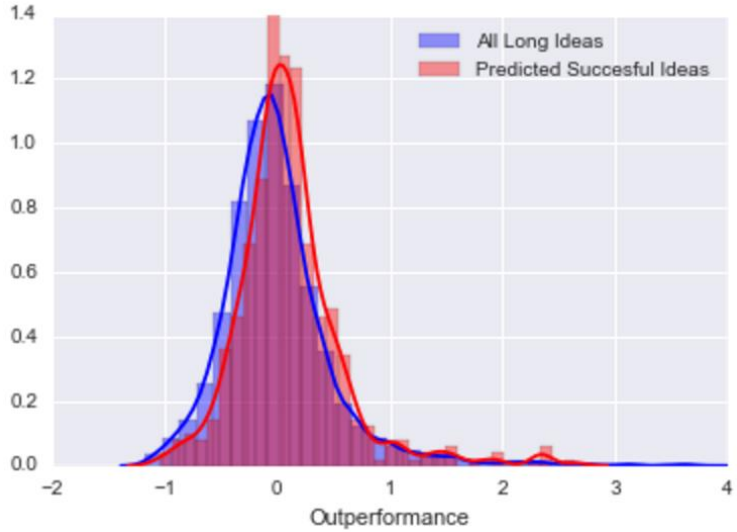
All long write-ups (blue)

- Mean outperformance: 0.88%
- # of write-ups: 4285

Prediction model (red)

- Mean outperformance: 10.53%
- # of write-ups: 446

Model vs. All VIC Ideas



Notice how the model's (red) distribution is "pushed" to the right compared to that of the original (blue)

Machine learning makes the crowd smarter

“Use machine learning to take data and do something that is better than what the humans are doing. Take the human crowdsourced data and you’re going to compute something new.”
-Eric Schmidt

Thanks

Send questions and comments

Contact: JRosen.1392@gmail.com

Appendix: model exploration

- **Features:**
  TF-IDF (n-grams), valuation metrics, sentiment, etc
- **Target:**
  Outperformed (0 or 1)
- 3 or 5 fold cross validation

1) Random Forest Classifier
Number of estimators: 400
ROC AUC score: 0.56 – 0.60

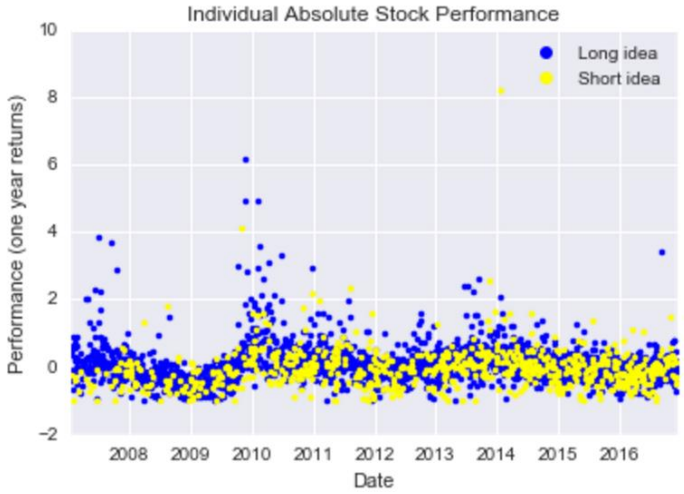
2) AdaBoost Classifier
Number of estimators: 50
ROC AUC score: 0.54 – 0.55

3) Naïve Bayes Classifier
Number of estimators: 50
ROC AUC score: 0.54 – 0.55

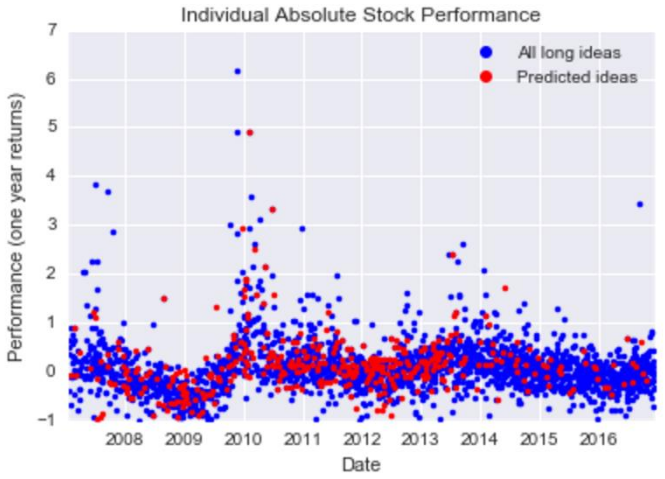
4) Neural Network Classifier
Work in progress

Appendix: misc visualization

All write-ups



Random Forest Classifier predictions



Appendix: exploring the model

Black-box models make it difficult to interpret how features/variables are being used to classify each write-up

Important features include:

- Various text weightings of cash flow references
- Various text weightings of valuation terms
- Misc. valuation metrics

"Important" Features

	Features	Importance Score
536	cash flow	0.004345
976	free cash	0.003240
911	ev ebitda	0.002927
1710	share price	0.002815
429	balance sheet	0.002806

Appendix: future steps

Extra model tuning

- Feature engineering
- Regression vs. Classification
- Continuous model return (+XX%) vs. Discrete outperform (y/n)
- Minimize error metrics

Time Series

- ROI optimization

More data

- Scrape message content
- Include more general financial data

Automated trading algorithm

- Deploy model into simulated trading environment

Appendix: limitations

Financial Data is expensive and/or often lumpy

- Gaps in data make it difficult to use valuation metrics
- Lack of free, scalable API

“Exotic” Investments have not been priced

- Special Situations, Debt, M&A, Warrants are mostly excluded

No message scraping

- Responses to write-ups can be indicators of future outperformance

Write-up descriptions

- 35,000 word-count max
- Sentiment analysis used in model was trained on movie reviews

Sample Size

Sources

- ValueInvestorsClub.com
- FactSet
- Presentation template by SlidesCarnival
- Photographs by Unsplash