News-Based Trading Strategies: An Exploration

Vivek Patel

Professor Dhar

Trading Strategies and Systems

Spring 2011

**INTRODUCTION**

News is a vehicle for investment, providing investors with the information they want and need to help them make investments. It informs them of movements, analyst estimates, earnings reports, and breaking news about markets which would otherwise escape their sight. It also seems to encompass a mass sentiment about certain companies, markets, or movements. We can gauge sentiment when reading articles – it is something humans are inherently skilled at doing. This mass sentiment potentially can move markets, determine prices and movements. The question follows: does sentiment correlate with future returns? If it does, we can build a strategy on that relationship.

Some have already built strategies on news-based methods, trading off “exposure” to certain micro factors. If a company is “exposed” to a certain element that analysts see as misaligned with the market, this strategy would call to trade until the market recognizes the mismatch, or after a fixed amount of time. In theory, this database of micro factors grows and becomes more robust, developing more complex relationships and connections. This strategy also allows for revisions as more recent news comes out.

While this strategy seems to have performed well in the past, the process is labor-intensive. Humans spend much time consuming news because we have an exceptional ability to understand language. However, we can also argue we are also subject to bias and ignorance when it comes time to make an investment decisions. We also have limited resources and can consume so much.

Computers fill in these deficiencies. They easily scan thousands of articles a minute, and are unbiased in their investment judgment. They struggle with understanding language, however. Over the last semester, we sought to delve into an automated approach to news-based trading.

The process was exploratory. We created a program to parse business news articles supplied by Reuters, determine average daily sentiment, and then invest based on a series of different strategies built on a common sentiment analysis algorithm.

This research stood on three assumptions:

1. News sufficiently measures sentiment.

2. Markets react to news.

3. A lag exists between news and investment action.

**METHODS**

**1. Prototype**

Before building the scaled program, a prototype was built for two purposes, 1) to get familiarized with the Python programming language, and 2) to see if scaling such a program would be feasible. This prototype would only generate sentiments on a pre-populated file of query results. Connecting to the Reuters database, we queried the database with the following query:

SELECT STORY\_DATE\_TIME, CAST(TAKE\_TEXT AS TEXT) FROM Reuters.dbo.news WHERE RELATED\_RICS LIKE ‘%WMT%’ AND LANGUAGE = ‘en’ AND EVENT\_TYPE = ‘STORY\_TAKE\_OVERWRITE’ AND STORY\_DATE\_TIME >= 2003-01-01 AND STORY\_DATE\_TIME <= 2003-03-01

This would effectively fetch all articles related to Wal-Mart Stores, Inc. (WMT) in the Reuters database from January 1st, 2003, to March 1st, 2003, a period of 2 months. For an initial test, two months seemed reasonable, considering the approximately 150 articles that Wal-Mart has mentions in. We then exported the results as one text file.

To conduct sentiment analysis, we parsed the text by new-lines, and then looked at positive and negative word matches (Harvard Psychology provided us with a lexicon of words with positive or negative associations). The program did not stem any words at this stage. The daily average sentiment (DAS) was calculated as follows,

Here, pa is the positive word count, na is the negative word count, a is the start article index and b is the end article index for a particular day. Before it would be possible to test for correlations, however, WMT prices would have to be obtained.

We exported historic stock price data from Yahoo Finance and cleaned the data so only days where both an article and stock price existed. Over two months of data, this was whittled to a relatively small number of 22 DAS data points. To determine whether correlation today tended to correlate with returns tomorrow, we conducted analysis in Excel.

The results seemed promising, but we stayed wary of the little amount of data we could build assumptions on. The results showed relatively strong correlations, but they wavered and had indescribable patterns.

Regardless, we decided to build a program more scalable and automated: one that could fetch a year’s worth of articles, generate sentiment, and run a strategy automatically.

**2. Build Infrastructure**

The scaled program would work as follows.

Query Reuters database for company articles

Query Yahoo database for company stock prices

Generate sentiment

Merge and obtain EMA

Run strategy!

It would start by prompting the user for a stock ticker, a start date and an end date. Querying the Reuters database for articles that mention the stock ticker between the start and end dates, the program would then fetch those articles and combine them and ready them for sentiment analysis. Sentiment analysis was conducted first by generating daily sentiment:

Here, pa is the positive word count, na is the negative word count, a is the start article index and b is the end article index for a particular day. ti is the total word count and ui is the uncertain word count. Note the dampening effect of uncertainty on DAS. Investors typically have more reservations in times of uncertainty and tend to trade less, making markets slower.

Since news tends to have a cumulative effect, we then calculated an exponential moving average for each day:

Today’s sentiment has a weight of 1, yesterday’s sentiment has a weight of ½, two day prior has ¼, etc. The division by two at the end comes from the sum of those weights, the sum of a geometric series. , where ½ is the r, the ratio.

Once generating sentiment, we used the Yahoo Finance API to fetch the closing prices between the date of the first and last articles in the range. We merged the close, DAS and EMA by date. Now, we could run the strategy.

**3. Build Strategy**

Four parts composed the strategy: 1) preliminary analytics to choose various metrics to choose position on, 2) generating positions, 3) running the strategy and collecting metrics, 4) generating the output and plotting the results.

During preliminary analytics, we calculated the lagged correlations and calculated returns.

Generating positions was then calculated by looking at daily change in EMA. We decided the threshold should change – when sentiment is extreme in a certain direction, it is assumed changes in sentiment won’t have much of an effect on prices. If sentiment is relatively neutral, or is at one extreme, then goes to the other extreme, this intuitively should have a larger effect. The threshold was calculated dynamically, using a modified, windowed coefficient of variation:

Note that if the mean approaches 0, the threshold approaches infinity. To avoid this, we add 0.5 to the denominator. We employed a counter-trend strategy: if a drop in sentiment happens, price may take a hit, but it will likely go back up; conversely for a jump in sentiment. Part of this reasoning also comes from the fact that we considered only closing prices. Most business news is released intra-day, so action in a certain direction is likely to have happened at that point.

The pseudo-code for the strategy is as follows:

If daily change in EMA > threshold:

Go short

Else if daily change in EMA < –(threshold):

Go long

Else

Take no position

One positions were generated, we ran the strategy. If the signal said to trade at the end of one day, we would realize the returns the next day.

Metrics were collected from the results thereafter. We calculated the Sharpe ratios of both our strategy and of the companies, the number of trades, the percentage in, the number of winning and losing days, and the average win and loss.

The program then would output the metrics and plot the data. We plotted three pieces of data: the equity value, the stock price, and the EMA.

We decided to run the strategy on five stock tickers: INTC, IBM, AAPL, GOOG, and HPQ. The main limitation in this study was speed – the Reuter's servers took several hours to serve a year's worth of articles. The stocks were chosen from the technology sector for comparison purposes.

**RESULTS**

A backtest was conducted on 2005-01-01 to 2007-01-01. The results are as follows:

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **BACKTEST** | **2 years** | 1/1/2005 | to 1/1/2007 |  |  |  |
|  |  |  |  |  |  |  |
|  |  | **GOOG** | **AAPL** | **YHOO** | **IBM** | **INTC** |
| **EMA STRATEGY** | Sharpe | -1.19 | 1.22 | -0.84 | 0.06 | 0.79 |
|  | Number of trades | 199 | 220 | 223 | 222 | 118 |
|  | Percentage in | 43.93% | 46.61% | 47.05% | 47.33% | 24.43% |
|  | Wins, losses | 97, 102 | 112, 108 | 98, 125 | 112, 110 | 60, 58 |
|  | Average win, loss | 0.016, -0.019 | 0.019, -0.016 | 0.014, -0.013 | 0.008, -0.008 | 0.014, -0.012 |

The backtest results were varied, with only two stocks performing positively using this strategy. GOOG and YHOO performed relatively badly, and IBM stagnated. We decided the backtest would not necessarily give a good indication of future performance and moved to running the same strategy on the next two years, 2005 to 2007:

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **TEST** | **2 years** | 1/1/2007 | to 1/1/2009 |  |  |  |
|  |  |  |  |  |  |  |
|  |  | **GOOG** | **AAPL** | **YHOO** | **IBM** | **INTC** |
| **EMA STRATEGY** | Sharpe | 1.45 | -0.76 | -0.12 | -0.76 | 0.67 |
|  | Number of trades | 223 | 200 | 216 | 205 | 193 |
|  | Percentage in | 45.23% | 40.73% | 44.54% | 42.27% | 38.83% |
|  | Wins, losses | 124, 99 | 98, 102 | 100, 116 | 110, 95 | 99, 94 |
|  | Average win, loss | 0.023, -0.021 | 0.02, -0.023 | 0.017, -0.015 | 0.01, -0.014 | 0.019, -0.017 |

The only consistent positive performer in both the backtest and test was INTC. Otherwise, performance generally flipped in a rather extreme opposite direction. Most two-year strategies did not work.

**DISCUSSION**

Why did the results have such variation? Among companies, sentiment patterns varied. Though this should be expected, certain companies seemed to consistently carry negative sentiment or positive sentiment. However, we anticipated this and attempted to control for stagnation in sentiment by varying the threshold.

First, sentiment analysis is very difficult for computers. Humans can recognize positive or negative sentiments in a piece of text easily, but training a computer to generate sentiment requires a vast network of relationships and context. This strategy also employed simple word count means – this leaves the strategy victim to instances of negating bigrams, or the nuances in larger blocks of text.

Considering the extreme flips in the performance, it seems another factor also plays into market response to sentiment. Perhaps the volume of articles also influences market response to sentiment. A high volume might reflect a closer public following of a company. A spike in volume might also indicate a tighter following.

Also, we looked at only closing prices. Much news occurs intra-day and is acted upon quickly. We realized this, which explains the counter-trend strategy employed: good news moves the market in a theoretically exaggerated positive direction, and is thought to correct itself the day after. Likewise for negative sentiments.

The fourth reason we think of comes from the issue of classification. Much of the news might encompass reactions to past changes in stock price, and not contain any actionable content. In fact, when the strategy was run with no threshold, trading positively on positive sentiment, negatively on negative sentiment, the equity graph followed the stock price movements almost exactly, except lagged by one or two days.

However, sentiment on occasion seems to have some predictive power. Running lagged correlations demonstrates this.

**CONCLUSION**

**Future Research**

The nature of this project, having to do with natural language processing, means it is highly multidimensional. Language can be interpreted in many ways and quantified in so many more. The best we can do is 1) create a robust and intuitive framework, and 2) ensure that the extra parts in the strategy tend to move independently.

Next, we plan on integrating financial terms into sentiment analysis. The Harvard lexicon may support general sentiment, but fails to count for IPOs, Obamanomics, arbitrage, etc. The financial lexicon of terms likely has a large influence on market reaction, so seems most pressing to implement.

Also, I think other indicators should be considered. Volume suggests how closely and how actively the public responds to news. Perhaps the counter-sentiment strategy works during high volume, and a regular sentiment strategy works during low volumes. Classification also seems significant – earnings announcements versus bankruptcy will likely have the same calculated sentiment, but clearly different actual sentiments.

Lastly, I think the ultimate step in this leg of research would be to create a portfolio of stocks. Individually, a stock may perform badly, but together, due to the benefits of diversification, the portfolio will likely perform better and without the high risk putting full equity into one stock.

**Takeaways**

News-based trading is at the next frontier in systematic trading. The relationship between news and investment action looks intuitive. And as computers become more powerful, and natural language processing techniques increase in robustness and accuracy, this kind of strategy looks continually more viable.

Many opportunities for loss exist, however. As evidenced by the results of this project, sentiment does not always correlate with future returns. Or perhaps, the results call for more accurate sentiment analysis. Or a different method of calculating the moving average. Or a consideration of volume. The higher dimensionality of natural language throws all these options at us. It also presents an opportunity to learn more about relationships between sentiment and markets. The true value is contained in these relationships, because then, the possibilities extend past trading strategies. It expands into research, market analysis, crisis management. Getting lost, making blind assumptions in this higher dimensionality can be easy, but with careful, stepped progress, we have much to learn and gain from.