Use of NLP-Powered Sentiment Analysis in Trading Strategy

Zanyang (Neil) Chen(✉)

UWC Changshu China, 88 W. Kunchenghu Road, Changshu, Jiangsu, China
neilchenhenri@gmail.com

Abstract. The research paper focused on sentiment analysis powered by natural language processing and its use in trading strategy. Previous sentiment analysis based trading solutions are usually proprietary and produce market reports on a daily or weekly basis. We, therefore, aimed to address the lack of transparency and execution efficiency in these solutions. By taking advantage of the algorithmic trading platform QuantConnect and referring to its documentation, we built our very own algorithm directly executable for live trading. The algorithm adopted a rule-based dictionary approach to analyze the market sentiment towards different stocks and the insight weight method to construct the portfolio based on the sentiment analysis results. The algorithm was backtested using the historical data and optimized for higher performance. The final testing results showed that the sentiment analysis powered trading strategy is able to achieve a similar performance to the benchmark, which is the S&P500 index, both in periods of bull markets and bear markets, with the potential of outperforming under favorable market conditions.

Keywords: sentiment analysis · trading strategy · execution · rule-based dictionary · insight weight portfolio construction method · backtest

1 Introduction

From the stock market to ordinary investment in banks, trading is ubiquitous among us. With the increase in disposable income and the democratization of financial data around the globe, the barrier to trade has been significantly lowered. As a result, retail investors with various degrees of expertise have flooded into the market, which further undermined the assumption of an effective market in traditional finance theories, and their trading behavior can be significantly impacted by the news or rumors that they receive. Their investing time horizon is also shorter than professional asset management practice and can change drastically over a short period of time. Therefore, to stand out from the crowds and respond swiftly to changes in the market, the use of sentiment analysis in trading strategy is essential and worth researching.

Currently, in the market, there exist several sentiment analysis solutions. For example, the SentimenTrader, is a sentiment-based trading service provided by Sundial Capital Research, an independent investment research firm dedicated to the application of mass

J. Yen et al. (Eds.): ICBIS 2023, AHCS 14, pp. 109–115, 2023.
https://doi.org/10.2991/978-94-6463-198-2_13

psychology to the financial markets. In short, SentimenTrader has developed a proprietary model for market sentiment analysis. The model is used to generate daily and weekly sentiment reports to help customers with their decision-making [1].

However, despite the ability to instantly research and update changes in sentiments, problems persist in the existing business solutions. The models are not open-sourced, resulting in low confidence in using them. The execution may be delayed if the users choose to make decisions on their own. If customers decide to fully automate the process, accountability risk may rise as they possess limited knowledge of the strategy. In addition, the user does not have the access to fully examine the performance of the model over a different time period, nor does he have the flexibility to customize the model.

Therefore, this research brought up a new approach to public attention, by addressing the aforementioned problems. In this research, we will implement sentiment analysis on QuantConnect, an online algorithmic trading platform. The algorithm is fully transparent, and the strategy could be deployed in live trading.

2 Methodology

2.1 Brief Introduction to the Strategy

To use sentiment analysis to make trading decisions, a strategy needs to be decided beforehand. The most fundamental principle is that we will buy more shares when the market sentiment towards a specific stock becomes positive and sell more when it turns negative.

One important factor is how our portfolio should be constructed, specifically, the number of holdings and the position of each ticker [2]. When deciding to buy, we also need to decide how much to buy or sell each time. In other words, we should know how many stocks to hold. It is also important for us to allocate the cash, because assuming the cash stays constant, buying more of one means selling of another. The question will become complicated when multiple stocks show positive sentiments simultaneously. When decided to sell after a negative sentiment on the market, we need to decide whether to do short or not, depending on our own long-short strategy.

Above all, rebalance or trade frequency is another important factor of portfolio management. We may want to rebalance our share of different stocks every quarter, month, week, or day, in order to reduce the risk of suffering from a slump made by a once-positive stock.

That said, risk management is another important aspect of the strategy [3]. Multiple risk metrics or indicators, such as absolute risk exposure, need to be applied to better inform us about the risks of every decision made. We should decide at which level of risk we would make a change according to the sentiment analysis results.

2.2 Sentiment Analysis Logic

At the core of the strategy is how we determine the sentiment towards each stock. As a result, we need to set up our own criteria when doing the sentiment analysis. Since most sentiments are derived from news, we will be mainly analyzing sentiment in news articles. We could simply classify all news into the categories of "good news" or "bad news". Although such classification could tell us whether to buy or sell, it cannot help us to decide how much to buy or sell. Therefore, we also need to evaluate the strength of optimism or pessimism in news articles. For example, while all imply an optimistic sentiment, "great", "good", "better" and "best" have different strengths, with "best" being likely the strongest. We made the above conclusion because "best" is a superlative and it implies a much higher degree of optimism than the rest as far as we have learned in our language classes. For machines, the way of teaching is to quantify sentiments by assigning scores to each word. We could either subjectively evaluate meanings and assign the word with a score, or adopt a frequency-driven approach as stronger emotions are likely to appear less often.

2.3 Strategy Implementation

To implement the strategy, we will first determine the time period of historical data to be used, the amount of starting cash, and the universe, which is the pool of tickers that we are interested in. We will pull in the universe-related data feed from the Tiingo dataset [4], which is to be used in this algorithm as the news source, after every hour or another outlined time period.

When the news data come in, the algorithm will scan through all of them and find those containing the keywords. The keywords are to be matched to the rule dictionary to assign a corresponding score to each keyword, and the total score of each stock ticker is thus aggregated. The algorithm will rank the tickers in descending order of their total scores and choose the top several tickers into the final list. A portfolio is constructed based on the ranking, either in equal weight or being score weighted.

The above process repeats on an agreed frequency and the portfolio is regularly adjusted according to the latest market sentiment.

3 Algorithmic Training and Evaluation

In machine learning, we typically train and optimize our algorithm using a labeled training set. In the field of quantitative trading, the historical record on the market is the training set. The process of looking back over the history to test for the profit and return of an algorithm is known as backtest.

For each round of backtest, the sentiment score dictionary, the portfolio construction method, and the execution frequency of the algorithm are variables to be determined. Sentiment scores are to be changed to better reflect the strength of emotions each of the words conveys and the dictionary is to be enriched to cover different types of news. The portfolio construction method, on the other hand, is more generic. After ranking to look for the top ten tickers, the portfolio construction method decides whether the number of shares purchased will be equally distributed or distributed based on ranking. If it is to be distributed based on ranking, the specific distribution could also be varied. Finally, the model could be run daily, meaning changing the composition of stocks in the portfolio once a day. It could also be run weekly or monthly and is up to our discretion.

Fig. 1. Backtest Results of the Original Solution

To better visualize the process of backtest, the following tests are conducted on QuantConnect, an online platform supporting the programming of algorithms using Python.

3.1 Original Solution

The original solution uses a limited dictionary, in which each word was assigned a score solely dependent on the polarity of the word, either positive or negative 0.5. The portfolio is constructed based on the insight weighting model. The stocks ranking in the top ten are to be purchased and the percentage of cash allocated to each stock is sentiment score dependent. By default, the model runs daily.

As shown in Fig. 1, the return of the algorithm is 8.23%, with the Sharpe Ratio [5] (return per unit risk) being 0.568 and alpha [6] (extra return other than the change in the general stock market) being 0.021.

3.2 Enriched Dictionary Solution

It is obvious that the dictionary in the original solution is limited, both in terms of the size of the dictionary and the accuracy of the sentiment score. Therefore, a new dictionary, the VADER lexicon open-sourced on GitHub [7], is used to improve the performance of the sentiment analysis solution.

The first method was to replace the original sentiment scores of the existing dictionary with the new sentiment scores provided by the VADER lexicon. Words that are missing in the VADER lexicon were removed.

The method showed an insignificant boost in performance of around 1%. To further improve, an enriched dictionary was adapted from the VADER lexicon. The new dictionary came with over 1600 words, increasing the return of the algorithm from around 9% to around 50%. Sharpe Ratio increased to 0.815 and alpha increased to 0.166. More details could be found below, in Fig. 2.

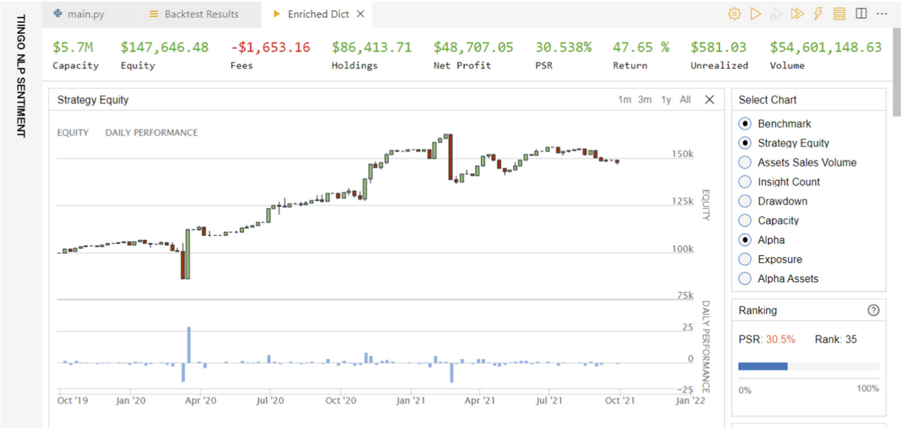

Fig. 2. Backtest Results of the Enriched Dictionary Solution

3.3 Equal Weighting Portfolio Construction Model Solution

Different from the insight weighting portfolio construction model implemented in the original solution, the equal weighting model evenly purchases the ten top-ranking tickers. In the test, the equal weighting model outperformed the insight weighting model by an insignificant 0.24%. Both the Sharpe Ratio and alpha declined. As a result, the final solution will still adopt the insight weighting portfolio construction model.

The choice of using the insight weighting method is also cross-validated with the results from the enriched dictionary. In this case, the insight weighting method outperformed the equal weighting method by roughly 10%. The result also matched our basic and common knowledge of investment.

3.4 Solution of Selecting Different Numbers of Securities to Purchase

In the original solution, the top ten tickers are to be purchased. The algorithm has been tested with another five options: purchasing only the most positive ticker, purchasing the top 3, 5, 8, or 15. The results turned out that purchasing the top 1, 3, and 5 results in lower returns whereas purchasing 8 and 15 results in the same return as purchasing 10. To balance between return and risk, we will continue to use the original solution of purchasing the top ten tickers.

3.5 Final Solution

The final algorithm will consist of the enriched dictionary and the insight weighting portfolio construction model. The algorithm is still to be executed on a daily basis.

4 Testing and Results

To avoid the problem of overfitting, the aforementioned training only used data from October 1, 2019, to October 1, 2021. The data from there on until June 1, 2022, are used together as the testing set.

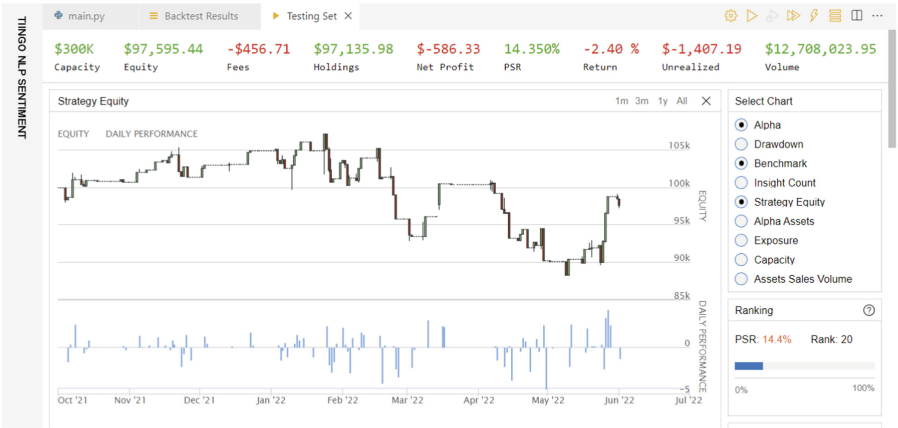

Fig. 3. Backtest Results using the Final Solution and Testing Set

Applying the final solution and backtesting it using the testing set, the return from October 1, 2021, to June 1, 2022, turned out to be -2.40%. The change in the benchmark during the same period is $\frac{409.4913-424.4793}{424.4793} = \frac{-14.988}{424.4793} = -3.53\%$. Comparing the result of the algorithm with the change in the benchmark, the algorithm has demonstrated a fair performance that slightly outperformed the market. Considering the unusual volatility, unprecedented significant events, and drastic change in the market backdrop in the past 6 months, we believe the model performed very well.

Figure 3 below shows the model performance in detail.

5 Conclusion and Future Works

As shown by the training process and testing results, the final algorithm on QuantConnect successfully applied the approach of sentiment analysis in trading strategy and achieved satisfactory outcomes.

In the future, the program may be further improved to push the results forward. Firstly, as part of the limitations of the dictionary approach of sentiment analysis, the program may not be able to handle sarcasm in news articles. The trend of the increased use of sarcasm will compromise the effectiveness of the algorithm. Secondly, from the results obtained from the testing set, we have speculated that the algorithm failed to gain positive returns because the program blindly purchased the top ten tickers. The stock market crashed in early 2022, leaving the extreme circumstances of having all stocks showing negative sentiment. In this case, the algorithm is to be optimized to activate a protection mechanism to reduce the loss.

References

1. "Our Core Mission." *SentimenTrader*, https://sentimentrader.com/about.
2. "Documentation - Algorithm Framework - Portfolio Construction." *QuantConnect.com*, https://www.quantconnect.com/docs/algorithm-framework/portfolio-construction.

3. "Documentation - Algorithm Framework - Risk Management." *QuantConnect.com*, https://www.quantconnect.com/docs/algorithm-framework/risk-management.
4. "Documentation - Alternative Data - Tiingo News." *QuantConnect.com*, https://www.quantconnect.com/docs/alternative-data/tiingo-news.
5. Fernando, Jason. "Sharpe Ratio Definition." *Investopedia*, Investopedia, 6 June 2022, https://www.investopedia.com/terms/s/sharperatio.asp.
6. Chen, James. "Alpha Definition." *Investopedia*, Investopedia, 14 June 2022, https://www.investopedia.com/terms/a/alpha.asp.
7. Hutto, C.J.. "VADER Sentiment Analysis." *GitHub*, https://github.com/cjhutto/vaderSentiment.