

Market Efficiency of News Events

Adrian Beer

Contents

1	Introduction	1
1.1	Categories	1
1.1.1	Labeling	1
2	Related Work	2
2.1	Topic Modeling	3
3	Methodology	3
3.1	Holding Period and Exit-Strategies	4
3.2	Labeling the Data Set	4
3.2.1	Adjusting Returns	5
3.3	Backtesting	6
3.4	Historical Data	6
3.5	Price Data Set	6
3.5.1	News Data Set	6
3.5.2	Ticker Consolidation	6

1 Introduction

One possible explanation is that "low-capitalisation, young, unprofitable, low-dividend-paying, high-volatility and highgrowth companies are difficult to arbitrage or value according to traditional financial theory and are therefore very sensitive to investor sentiment".

1.1 Categories

Peng and Jiang (2015): Bag of Keywords, Polarity Score, Category Tagging... Using initial seed words for each category. Categories: new-product, acquisition, pricerise, price-drop, law-suit, fiscal-report, investment, bankrupt, government, analyst-highlights.

1.1.1 Labeling

- Next Days Close
- Next Close/Open
- Trailing Stop-Loss

2 Related Work

”Following the work of Luss and d’Aspremont (2015), Ding et al. (2015), Xu and Cohen (2018), Ke et al. (2019), we formulate the stock movement prediction as a binary classification task”

? investigates the impact of financial news on the daily returns of S&P 500 constituents using a self-fine-tuned BERT model. Their algorithm produce the CLS Token from FinNewsBert and a Topic vector from Text2Topic and feed these into a standard feed forward NN, which then produces probabilities for the 3 classes. They find that the news is priced in within one day, but not instantaneously. Their FinNewsBert model is retrained every two years in order to obtain a large time-frame for their out-of-sample study. They consider the problem of look-ahead-bias of the base BERT model, due to the data that it is trained on. They use a dataset of financial news published by Refinitiv from 1996 to 2020 and from Refinitive Datastream daily price data of 1330 companies, which were at some point in time S&P consituents. As part of the preprocessing they convert news articles to lower case, remove all numbers, punctuation marks and brackets, so that only letters remain. Multiple news articles about one company published in quick succession are merged into a single document. (Q: How is this handled coupled with the token limitations of BERT?) An algorithms is used to discern fresh ans stale news. The labels for the training of FinNewsBert are constructed based on the idiosyncratic price movement following the news, namely on the following formula:

$$IR_{i,t} = R_{i,t} - R_{f,t} - \beta_{i,t} * (R_{S\&P500,t} - R_{f,t})$$

Afterwards z-scores are calculated for each stock, in order to avoid over-weighting price movements in smaller, more volatile stocks.

$$z_{i,t} = \frac{IR_{i,t} - \mu_{i,t}}{\sigma_{i,t}}$$

They find improved classification performance, with the topic vector being used as input. With the Text2Vec algorithm they differentiate between analyst forecasts, earnings reports, monetary policy and business/strategic.

Ke et al. (2020) propose a text-mining methodology, named SESTM, to extract sentiment information from text. They used a bag-of-words representation of news articles. To distinguish fresh from stale news a measure of novelty was constructed based on the cosine similarity of an articles with all other articles about the same firm five trading days prior.

$$\text{Novelty}_{i,t} = 1 - \max_{j \in \chi_{i,t}} \left(\frac{d_{i,t} \cdot d_j}{\|d_{i,t}\| \|d_j\|} \right)$$

They find that price responses to news are larger for fresh news. They also found larger price responses for smaller and also for more volatile stocks. Price movements for more volatile stocks last up to three days on average, while news about low volatility stocks are incorporated after just one day of trading.

Liu and Stentoft (2023) analyze the predictability of stock returns for different time horizons, ranging from 1-min to 30-min, and market sectors between 2005 and 2016. They find high profitability even after transaction costs. Intraday predictability decreased with higher time horizons

These results are consistent with the observations of Baker (2007b), who observes that low-capitalisation, young, unprofitable, low-dividend-paying, high-volatility and highgrowth companies are difficult to arbitrage or value according to traditional financial theory and are therefore very sensitive to investor sentiment.

Fazlija and Harder (2022) also use BERT for directional prediction, however not of individual stocks, but of the S&P 500 Index itself. Binary Classification task using Random Forest. Using the Financial Phrase Bank for fine-tuning the BERT model. Using a dataset of news articles from Bloomberg and Reuters between 2007 and 2016. They analyze the usefulness of title- and content sentiment separately. No transactions costs considered. No buy or sell threshold, positive sentiment results in buy, negative sentiment in sell. Using the individual sentiment scores of 58 news articles per day and some technical indicators as features for the Random Forest, they achieve a marginal outperformance, even though it doesn't seem to have been sustainable and statistically significant.

2.1 Topic Modeling

Salbrechter use Text2Topic, which is based on Word2Vec and cosine similarity.

3 Methodology

Because pre-news price movements have shown to correlate with the sentiment of the news, we incorporate this information into our feature set Ke et al. (2020).

3.1 Holding Period and Exit-Strategies

Our neural network tells us when we should enter, but not necessarily, for how long we should hold the position, i.e. the exit strategy. An important variable of any signal-based trading strategy is the exit-strategy. Depending on how it is chosen, the characteristics of a trading strategy can vary widely, e.g. regarding the return distribution of trades or how much capital can/should be allocated on average.

Theoretically, different models could be trained, where the target output of the model changes based on the exit-strategy, however this would involve fine-tuning multiple models and could strongly increase the variance of its predictions, e.g. for longer holding-periods, and distract the model from its intended task which is quantifying the sentiment of a news article.

Hence we will choose one more-or-less arbitrary holding period which is hopefully long enough, so that the price change reflects the change in market sentiment due to the message, but not so long that other events start to contaminate this price response to the news. Considering the analysis of Ke et al. (2020) we choose a holding period of one to two days. Specifically, for intra-day news the holding period will last until the next day's close or in other words the second closing price which will be observed after receiving the news. For overnight news the holding period lasts until the next observed close.

The closing price is a natural choice for the end of the holding period, because it is a single price which reflects the value of the stock. This is compared to the bid and ask prices that are available during the day and can show significant spreads. One disadvantage of this holding period is that the holding periods differ and so intra-day news which are released early in the day will lead to an observed price reaction which encompasses a longer time frame than for news which were released only later that day. However when choosing a fixed holding period of e.g. 30 minutes we would have to differentiate between intra-day news and overnight news, since there are no observed prices directly after news publishment for a large number of overnight news.

So this decision introduces some variance into the recorded price reactions, but it lets us treat intra-day and overnight news the same and reduces complexity w.r.t. the algorithm and analysis.

3.2 Labeling the Data Set

When choosing what we want to choose as the output of the Neural Network we are confronted with various options. Most stock prediction mod-

els using news are labeled in a binary or ternary way, i.e. up/down or up/down/neutral. Aside from this specific approach to predicting stock prices, other approaches usually employ point forecasts, which are then evaluated using the Mean Squared Error (MSE) or Mean Absolute Error (MAE). Other approaches exist, but are rare, e.g. Martinez et al. (2009) forecast the high and low of a stock on a daily basis in order to construct a profitable trading system.

3.2.1 Adjusting Returns

Adjusting returns to the market is important.

Compare Salbrechter (2021):

$$IR_{i,t} = R_{i,t} - R_{f,t} - \beta_{i,t} * (R_{S\&P500,t} - R_{f,t})$$

$$z_{i,t} = \frac{\overline{IR}_{i,t} - \mu_{i,t}}{\sigma_{i,t}}$$

Chen (2021):

$$r_{s,t} = \frac{P_{s,t+\Delta t}}{P_{s,t}} - \frac{P_{m,t+\Delta t}}{P_{m,t}}$$

Comparing the approaches between Chen and Salbrechter, we see that Chen basically assumes a constant beta of 1 for all stocks. Beta is notoriously unreliable and changes, depending on the market environment. E.g. all stock betas tend to 1 as the large market returns are observed during a crash. This follows from the CAPM and has been observed empirically.

Our main concern are large returns which are caused due to market risk events - not due to idiosyncratic risk events. These would erroneously create labels of large quantity which result in faulty sentiment signals. Hence working with a beta that's too small and is based on an estimation during calm market phases would very quickly result in many false signals during financial crises.

On the other hand, using betas of one across the board will create another bias. Namely, *ceteris paribus* stocks with the same amount of idiosyncratic risk but a lower beta will have exhibit a higher variance in the adjusted returns than stocks with a higher beta. However high $IR_{i,t}(beta = 1)$ will occur more randomly across time. It's hard to make statements about the magnitude...

Adjusting the returns by their (estimated) volatility is also necessary, as otherwise for volatile stocks one would observe many large returns, which are just normal price movements for that stock but the same price movement may represent significant changes in sentiment due to news for another low volatility stock.

3.3 Backtesting

3.4 Historical Data

3.5 Price Data Set

IQFeed provides us with 1-minute intraday data that is unadjusted to neither splits nor dividends and daily data that is only split adjusted. The 1-minute data is edge labeled on the right, i.e. a candle labeled 16:01 captures the time interval between 16:00:00 and 16:00:59. Since we don't have historical bid and ask prices we will only consider the trading period from 9:30am to 4pm and the closing and opening auctions. Otherwise trades are too intransparent and illiquid.

As we will use the intra-day data from IQFeed that is unadjusted we backwards adjust each stock time series ourselves using Split and Dividend data that is available from Yahoo Finance via the python library `yfinance`. Since this requires the stock to be made available by Yahoo Finance, this automatically eliminates all tickers for which Yahoo Finance does not provide data. It is important that we adjust the data, especially for splits, since these can lead to extremely high erroneous one-day returns in the data base, if left unadjusted.

3.5.1 News Data Set

We get our news from the Benzinga News API. Each news is tagged with stock tickers, associated to the news. We consider all news which are tagged with exactly one stock ticker.

3.5.2 Ticker Consolidation

In some cases we will have multiple price time series for the same company. E.g. in case of Alphabet (Google) we have two different tickers and two different stock prices for the same underlying company. Here 'GOOG' and 'GOOGL' describe two different classes of stock for the same company.

I consider it too hard of a problem to differentiate between different classes of stock when automatically assessing news. Hence we will choose one ticker to represent the company. If there is only one price time series available for the company, we simply group together the tickers. We utilize the simple rule which chooses the ticker with the most recent values and the longest price history.

References

- Qinkai Chen. Stock Movement Prediction with Financial News using Contextualized Embedding from BERT, July 2021. URL <http://arxiv.org/abs/2107.08721>. arXiv:2107.08721 [cs, q-fin].
- Bledar Fazlija and Pedro Harder. Using Financial News Sentiment for Stock Price Direction Prediction. *Mathematics*, 10(13):1–20, 2022. URL <https://ideas.repec.org//a/gam/jmathe/v10y2022i13p2156-d843797.html>. Publisher: MDPI.
- Zheng Tracy Ke, Bryan T. Kelly, and Dacheng Xiu. Predicting Returns with Text Data, September 2020. URL <https://papers.ssrn.com/abstract=3389884>.
- Fred Liu and Lars Stentoft. Intraday Stock Predictability Everywhere. *SSRN Electronic Journal*, 2023. ISSN 1556-5068. doi: 10.2139/ssrn.4496917. URL <https://www.ssrn.com/abstract=4496917>.
- Leonardo Martinez, Diego Hora, João Palotti, Wagner Meira Jr, and Gisele Pappa. From an artificial neural network to a stock market day-trading system: A case study on the BM&F BOVESPA. pages 2006–2013, June 2009. doi: 10.1109/IJCNN.2009.5179050.
- Yangtuo Peng and Hui Jiang. Leverage Financial News to Predict Stock Price Movements Using Word Embeddings and Deep Neural Networks, June 2015. URL <http://arxiv.org/abs/1506.07220>. arXiv:1506.07220 [cs] version: 1.
- Stefan Salbrechter. Financial News Sentiment Learned by BERT: A Strict Out-of-Sample Study, November 2021. URL <https://papers.ssrn.com/abstract=3971880>.