A Quantitative System Conception of the Impact of News on Stock Prices ——An Improvement Based on the Transformer's Self-Attention.

Xin Liu[1,a*], Kuang Du[2,b]

{lx001201@163.com[a], reductionism@foxmail.com[b]}

Institute of Economics, Southwest Minzu University, No. 168, Wenxing Section, Dajian Road, Shuangliu Airport Development Zone, Shuangliu District, Chengdu ,China[1]
Institute of Computer Information Management, Inner Mongolia Finance and Economics College, No. 185 (West), North Second Ring Road, Huimin District, Hohhot, Inner Mongolia ,China [2]

Abstract. News is one of the important factors affecting the volatility of stock prices. However, investors in the Internet era are limited by restricted attention span and information processing heterogeneity under the flood of information, which gives a place to news quantification. This paper focuses on the theoretical issues underlying the different research perspectives in the process of studying the mechanism of news impact on stock prices. Then, based on this, an improvement of Quantitative System Conception of the Impact of News on Stock Price is made using transformer's self-attention.

Keywords: News Quantification, Stock Price Volatility, Transformer

1.Introduction

This paper will begin with a question: "What determines the price of a stock?" The efficient market hypothesis suggest that stock price can accurately reflect changes in information available to investors in a timely manner. As a factor of production, the independence and importance of information become more and more evident. However, at present, China's quantitative trading system mainly focuses on technical quantification, which rarely involves information quantification. In a less mature and policy dependent financial market like China, how to quantify and predict the impact of news on stock prices is a key blockage in the development of the current quantitative investment system in China. From there, the paper continues to focus on how this problem can be solved with the help of a computerized natural language processing.

2.Discussion of theoretical foundation

As shown in **Figure 1**, stock prices can be discounted by future dividends or extrapolated by historical trends, with the former being used mainly for asset pricing in financial theory and the latter for technical analysis. In any case, the news appears as a shock to the current state of the stock price, and the mechanism of this shock has to be done by investors. Whether or not

ICEMME 2023, November 17-19, Beijing, People's Republic of China

DOI 10.4108/eai.17-11-2023.2342762

investors gain more information during the expansion of news dissemination channels, whether or not there is heterogeneity in investors' processing of information, whether or not they take action and how long the duration of the action is. These are all questions that must be discussed before a quantitative system can be established.

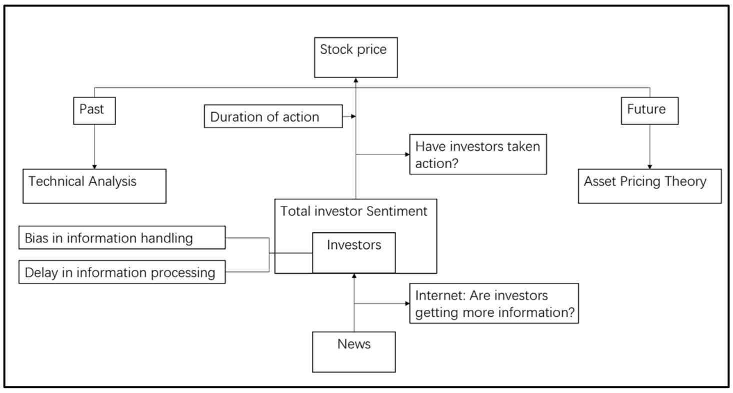
Stock price
Past
Duration of action
Future
Have investors taken action?
Technical Analysis
Asset Pricing Theory
Total investor Sentiment
Bias in information handling
Investors
Delay in information processing
Internet: Are investors getting more information?
News

Fig 1 The path of news affecting stock price

2.1 News and Stock Price

Although there are many innovative and exciting studies in the field of news quantification and many methods and models have been developed to better solve the problem of how news affects stock prices, this paper argues that most of the studies have not clearly sorted out the economic logic behind the problem or the logic of investor behavior before building the models. As an example, the "Efficient Market Hypothesis" is often mentioned in many papers constructing investor sentiment indices. According to Efficient Markets Hypothesis, it can divide stock market into strong effective market, partly strong effective and weak effective marketplace. With weak efficient markets, it is not effective for us to use history to predict the future. In a partly strong efficient market, only future information can be used to predict stock prices. However, in the published studies, we have seen little verification of the extent of the effectiveness of the market. You can show support or doubt for the theory instead of a lack of discussion. This is why academics often have criticized this type of research. From this perspective, the purpose of this paper is also to try to provide a more solid theoretical foundation and a more rigorous logic for this type of research.

2.2 What determines the price of a stock?

The human cognition of stock market fluctuation law is a very challenging world-class problem. To date, no theory or methodology has been convincing and has stood the test of time. In 2000, the famous economist Robert Shiller pointed out in his book "Irrational Exuberance" that "we should keep in mind that stock pricing is not yet a full science". In 2013, the Royal Swedish Academy of Sciences, in awarding the Nobel Prize in economics for that year to Eugene Fama, a leading expert on the efficient market hypothesis, and Robert Shiller, a professor of behavioral finance, noted that there are few ways to accurately predict the direction of the stock market bond market in the coming days or weeks[1]. Currently, the representative theories about the pricing of financial assets and the logic of stock market volatility are as follows: Random Walk

Theory (RWT), Modern Portfolio Theory (MPT), Efficient Market Hypothesis (EMH), Behavioral Finance (BF), etc.

Regardless of which theory determines the stock price, the formation of the stock price is an equilibrium state reached at a certain moment, such as the reasonable value of the company based on financial valuation, the price of the stock based on dividend discounting, etc. Any movement away from the equilibrium state cannot be achieved without the intervention of additional forces. In stochastic processes it is noise perturbation, in efficient markets and in line finance it is event driven. And no matter which of the three is the case, in the current society the dissemination of additional information is achieved through the channel of news. And, we can also get from many empirical studies that stock price movements are indeed related to news events.

2.3 Development of News Media

Recently, with the development of the era, the Internet has become the most mainstream way and means for human beings to obtain information resources with unparalleled advantages. The massive number of resources and information on the Internet, and the speedy and convenient network information service all determine the Internet to stand at the top of all kinds of communication media with its transcendent communication advantages. No matter who the users are, they can find the information they need on the Internet. The Internet has certainly expanded the access to information and increased its availability. Even global news can be accessed more quickly by investors. But in such an age of information explosion, on the one hand, the validity of information is rather to be tested, and on the other hand, investors are unable to handle the flood of information with their limited physical capacity. This provides a place for news quantification.

If you want to consider news quantification, it is not enough to consider these alone. Are investors actually getting more information? Or are they trapped in their own information cocoons? The latency of the internet is low, but how does one go about determining if there is a delay in investors' access to information? It comes down to the subjectivity of the investor, i.e., whether they have the ability and willingness to get information.

2.4 What investors do with information

This process is the most critical part of determining the entire quantitative system. Unfortunately, it is rarely covered in many quantitative papers and is only discussed in the field of behavioral finance. It is the key to determining that news has a quantifiable relationship with stock prices. Both efficient markets and modern portfolio theory make ideal, homogeneous assumptions about the process by which investors process information. But behavioral finance tells us that investors process information with a bias - limited attention.

Today, financial reporting is readily available, which seems consistent with efficient market theory. However, this is often not the case. Of concern is that firms are adept at exploiting investors' limited attention span by choosing to release bad news during periods of low attention and then finding ways to capture investors' attention when good news is available to drive stock prices higher. The "drift" after earnings announcements means that investors who do not follow the news closely may take a few days to catch up with investors and analysts who follow the news. When there are many "distracted" investors, the stock's immediate reaction to earnings news will be weakened or slowed, sometimes both[2].

Financial technology may have found a way to compensate for human shortcomings. Robo-analysts using advanced tools such as natural language processing and machine learning can absorb and update news faster than their human counterparts. "Evidence of all kinds suggests that robot-analysts are less affected by limited attention than traditional financial analysts and can help sellers achieve high-quality output."

2.5 Total investor’s sentiment

All investors are subject to emotions when they engage in investment behavior, and this emotional dominance may be difficult to detect. In academic terms, these can be called rational and irrational investors. Rational investors usually make investment decisions based on fundamental information and engage in arbitrage when they find that the stock price deviates from its intrinsic value, while irrational investors make "wrong" investment decisions based on their personal judgment, such as overconfidence or pessimism. There are studies that prove that those professional institutional investors are more influenced by "emotions" and thus make irrational trading behavior[3].

Neither complete rationality nor irrationality can be verified, and the above is only a description and definition of these two behaviors. The irrational behavior of rational investors is called the bandwagon effect. Investor sentiment is a variable that measures this "irrationality".

Although the concept of investor sentiment has been introduced early, the study of "sentiment" has always been in an awkward position. The most direct reason for this is that "sentiment" is difficult to measure. Many scholars have doubts about how this indicator is made. These include the earliest metric used as a proxy for sentiment which was the clos2-end fund discount, the most widely used BWI and the HJTZI and the calculation of positive word frequency. Later, it was found that some words may appear to have a negative message on the surface, yet they may indicate a positive message. Nowadays, most of the studies have started to favor the use of specific lexicons to classify textual information[4].

Studying the impact of overall investor sentiment on stock prices is absolutely the right direction to go. It is the overall buying and selling game of investors that determines the price direction of a stock. However, if we want to quantify it, the object of study must be selected with accuracy and representativeness. Commentary, as the mainstream research variable chosen, is by far the closest variable to investor sentiment. We do not have direct access to the sentiment in investors' minds, but can only measure it from its expression. Reviews, however, have the most fatal flaw, which is the inconsistency between Cognition and action. For example, an investor who has analyzed the stock price to conclude that it may still go up, but he or she will indicate in the commentary that he or she is not bullish on this uptrend or that the uptrend has come to an end and the reversal will start next. In this way, he tries to confuse other investors, hoping to get them to sell the stock so that he can continue to make profits. This phenomenon is quite common in China. The opposite opinion is expressed as a way to create a panic. So, the choice of a single channel to obtain comments will aggravate this problem.

2.6. How long can news affect?

News has a short or long window of influence on stock prices. The event study method in the financial world requires an artificially determined time window for each calculation. The field of artificial intelligence uses RNN (short term) and LSTM (long term) models to predict stock price movements within the window, respectively. Financial report releases look like short-term effects, macro policy changes may even have permanent effects. Therefore, a reasonable approach is to dynamically adjust the impact time frame of different news to improve the prediction accuracy.

3. Quantitative system design for the impact of news on stock prices

3.1 Previous System Design Review

In my early paper[5], an analysis system is proposed to quantify the impact of news events into two components: "duration" and " volatility of stock price". Firstly, the news is divided into three types: company, industry and macro class, and the corresponding historical news event database would be established. The symbol of the company history news database is a clear company name or stock code in the title or content of the news. Similarly, the symbol of industry class historical news database is a clear industry name or sector concept, and the rest of the news events without obvious characteristics are classified as macro class historical news database. One news item corresponds to the calculation of a stock price fluctuation, and the industry and macro correspond to the index fluctuation.

The most critical link of the whole system is to get the similarity by text matching. However, this traditional matching method only solves the problem of similarity at the lexical level, and cannot identify the problems that the meanings may be different in different contexts, and different sequences may have the same meaning, etc.

3.2 Transformer and text similarity

This section is a general overview of the relevant technical approaches, as shown in **Table 1**. Broadly, it is divided into three parts, the first is the main introduction of traditional semantic similarity algorithms, followed by machine learning-based semantic similarity algorithms, and then deep learning-based semantic similarity algorithms.

Table 1 overview of the relevant technical approaches

summary	Technical method
Traditional Contextual Similarity	TF-IDF
	BM25
	Sim Hash
Contextual similarity based on machine learning	VSM
	LSA
	WMD
Contextual similarity based on deep learning	Simase
	RCNN
	BIMPM
	DSSM

The Transformer model, proposed by Google in 2017, is mainly used in the field of natural language processing and has a complete Encoder-Decoder framework, which mainly consists of the attention mechanism[6][7][8]. Each encoder is composed of two main sub-layers, the self-attention mechanism and the feed-forward neural network. Transformer is designed to process sequential data (e.g., natural language), but does not require sequential processing of sequential data. Compared to recurrent neural networks (RNNs)[9], Transformer allows more parallelization, thus reducing training time and allowing training on larger datasets. Since its introduction, Transformer has become the model of choice for solving many problems in NLP, replacing the older recurrent neural network models[10].

Most of the existing text matching methods only consider the internal information of the text itself, ignoring the interaction information between two texts, or only perform one interaction after extracting text features, which can only obtain single-level interaction information, but lose multi-level interaction information[11]. With the concept of transformer self-attention, we can improve and upgrade the key parts of the original system.

3.3Underlying assumptions

The previous system was based on the assumption that the market would be efficient if it were. In order to fit the new system, we directly adopt the assumptions of technical analysis. These are: market behavior covers all information, security prices move along a trend, and history repeats itself.

1. Market behavior covers all information

2. Security prices move along the trend

3. History will repeat itself

3.4 New Analysis System

Now, let's see how the new system looks like, as shown in **Figure 2**. After we get a news story, we first perform named entity identification to see whether this belongs to a company news story, an industry news story, or a macro news story. Next, the news is returned to the corresponding past news database and a similarity score to the past news is calculated. Next, this score is used to weight the stock price volatility and duration corresponding to each past news article. This allows us to characterize how long and how much a real-time news story can actually affect.

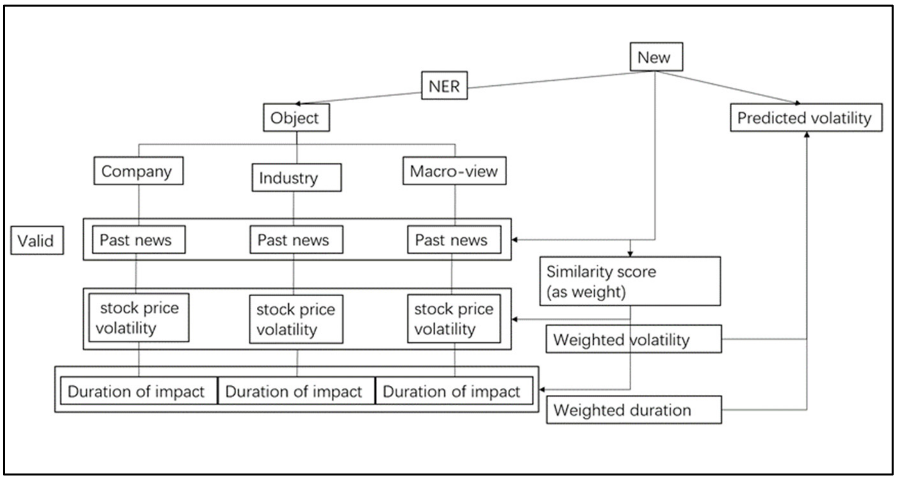

Fig 2 The process of the designed system

3.4.1 News Correlation

If you are holding a news story, we definitely have to think about finding the news story that is most likely to cause the same impact as this one among all the news stories. Self-attention is precisely a mechanism that considers the entire sequence of all information. We take a piece of immediate news item after encoding as a^0 and the rest of the past news items as a^1 to a^n. From a^0 to b^0, it can be divided into two steps. The first step is to find all the news items related to a^0 between a^1 and a^n. So how does self-attention determine the correlation of two vectors? There are many ways, such as Dot-product, Additive etc. we use the first method. The second step is from a^0 to b^0.

As shown in **Figure 3**. The similarity between $a_0$and historical news we use alpha (α) to portray, $\alpha_{0,1}$indicates the similarity between target news and past news, this α is actually the attention score. We let a^0multiply by W^q (query vector) to get q^0 vector and let { a^1, a^2…a^n} multiply by W^k (key vector) to get {k^1, k^2.... k^n}. The inner product of q^0 and k^1 is the $\alpha_{0,1}$ which represents the similarity between the two vectors, that is, the similarity between the target news and the past_news_1. This alpha is actually the attention score. In fact, generally in practice, we also calculate the correlation between q0 and itself, that is, $\alpha_{0,0}$. This step is important although it may seem to be of little significance.

After calculating the attention score, we find which past news is highly correlated with the target news. SoftMax () function normalizes all input vectors and maps them to a probability distribution $\{\alpha_{0,0}, \alpha_{0,1} \cdots \alpha_{0,n}\}$.

At this point each alpha can be used as a weight to weight the stock price volatility and duration of impact corresponding to each past news event.

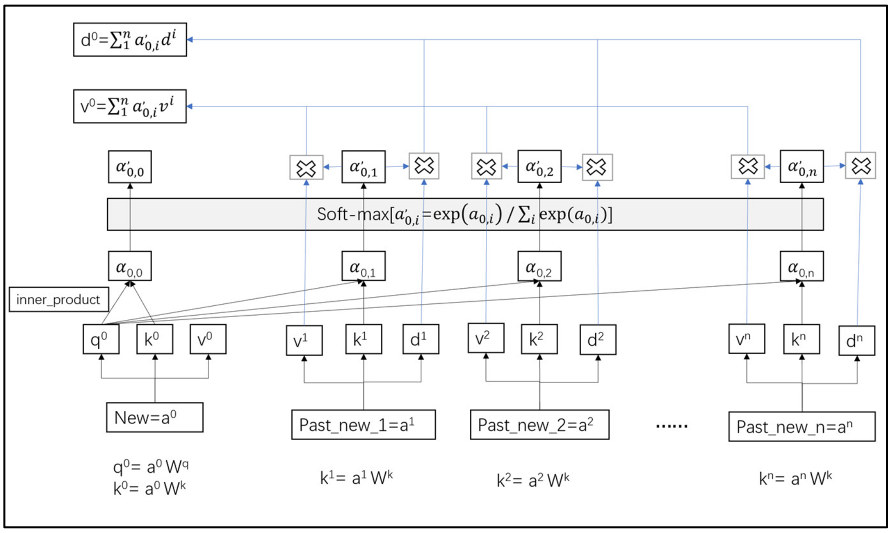

Fig 3 Calculation of similarity score

3.4.2 Stock price volatility

After obtaining the similarity, we still need to calculate the stock price volatility on its basis to get the final prediction. The average of the daily Opening Price P_1 and closing price P_2 is used as the mean value of the volatility，let the value be $Y = (P_1 + P_2)/2$; The distance between the maximum price P_{max} and the mean of volatility as the volatility of the maximum price，let the value be $M_1 = P_{max} - Y$, M_1/Y is the relative volatility of the maximum price；Likewise, Let the distance between the minimum price P_{min} and the volatility mean be the volatility distance of the minimum price, let the value be $M_2 = P_{min} - Y$, M_2/Y is the relative volatility to the minimum price.

Suppose the effect time limit of an event is n days, the opening price of day i is P_{1i}，the closing price is P_{2i}; if the highest price is P_{maxi} and the lowest price is P_{mini},then the highest price fluctuation is $M_{1i} = P_{maxi} - Y_i/Y_i$, the fluctuation of the lowest price is $M_{2i} = P_{mini} - Y_i/Yi$, The range of volatility on day i is $M_i = P_{maxi} - P_{mini}/Yi$.

Duration is measured with a Var model. Var models were developed using the daily closing prices in the time interval between the preceding and following news events, and impulse response analysis was performed to determine the timeliness of the preceding news event. In the impulse response analysis process, the degree of impulse response is used as the influence of news events on daily stock price movements. Suppose the influence of day i is F_i; Percentile the influence, the weight of day i is $Wi = F_i/\sum_{i=1}^n F_i$. Then the lower bound of the relative range of stock price volatility in n days is：

$$MIN = \sum_{i=1}^n (w_i * M_{2i}) = \sum_{i=1}^n [\frac{Fi}{\sum_{i=1}^n Fi} * \frac{P_{min} - (\frac{P_{1i}+P_{2i}}{2})}{\frac{P_{1i}+P_{2i}}{2}}] \quad (1)$$

The upper limit is：

$$MAX = \sum_{i=1}^{n}(w_i * M_{1i}) = \sum_{i=1}^{n}[\frac{Fi}{\sum_{i=1}^{n} Fi} * \frac{P_{max}-(\frac{P_{1i}+P_{2i}}{2})}{\frac{P_{1i}+P_{2i}}{2}}] \tag{2}$$

The volatility of the stock price during the effective period is：$\{MIN, MAX\}$。

For both broad market indices and sector indices, the idea of building a library of stock price movements is used, simply by replacing the stock prices with the corresponding indices to obtain the relative volatility of the indices over the validity period of the news event. The calculated volatility over the validity period of each news event is stored in the index database, and a corresponding link is established with the past news database.

4.Conclusion

News is one of the most important factors influencing stock price volatility. However, how to quantify news is a difficult problem for investors in the era of information flooding. The continuous development of NLP gives us more means to deal with different domains. In this paper, we improve the similarity calculation method based on the original system by combining transformer.

References

[1]. Zhong Xuezhong ,Liu Kai ,Wang Chuncheng . Time-varying economic dominance in financial markets: a bistable dynamics approach[J]. Chaos,2018,28(5).

[2]. Shijia Wu (2002). What Limited Attention Does to Efficient Market Theory[D], Handbook of Financial Decision Making.

[3]. Mao Yipeng. Trading Information, Asset Prices and the Asymptotically Efficient Market Hypothesis [D]. Shanghai University of Finance and Economics,2022.

[4]. Brown, G. W., & Cliff, M. T. (2004). Investor sentiment and the near-term stock market. Journal of empirical finance,11(1), 1–27.

[5]. Liu Xin, Sheng Mingcai, Huang Xi, Su Ganya. The Conception of Stock Price Volatility Analysis System Based on News Events[P]. 5th International Conference on Social Sciences and Economic Development (ICSSED 2020),2020.

[6]. Suranga M T S S .Multimodal Emotion Recognition With Transformer-Based Self Supervised Feature Fusion[J].IEEE ACCESS,2020,8176274-176285.

[7]. Yulu Wang. Research on Chinese named entity recognition technology and its application [D]. Beijing University of Posts and Telecommunications, 2022.

[8]. Wang Yannian,Ruan Pei,Lian Jihong et al. Research on semantic segmentation method based on dual attention mechanism[J/OL]. Journal of Xi'an Engineering University:1-7(2023)

[9]. Yulu Xia. A review on the development of recurrent neural networks [J]. Computer Knowledge and Technology, 2019, 15(21): 182-184.

[10]. XiPeng Q ,TianXiang S ,YiGe X , et al.Pre-trained models for natural language processing: A survey[J].Science China(Technological Sciences),2020,63(10):1872-1897.

[11]. Tongyue S, Zhongqin W. An overview of pre-trained language models for natural language processing based on Transformer[J]. Information and Computer (Theoretical Edition),2022,34(10):52-56.