Journal of Applied Finance & Banking, Vol. 10, No. 5, 2020, 201-213
ISSN: 1792-6580 (print version), 1792-6599(online)
Scientific Press International Limited

The Impact of the Public Opinion on Stock Market: Evidence from Weibo in China

Dongqi Cui[1] and Yuhan Cheng[2]

Abstract

We examine the effect of public opinion measured by Weibo tweets' tone about the COVID-19 virus pneumonia event on firm's stock return. During January and February 2020, tweets about new virus pneumonia experience an explosive increase with the event continually discussed, and the proportion of positive mood has a gradually growth. Our baseline results show that higher positive tweets ratio will lead to higher stock return, compared to both total tweets' number or negative tweets' number. The result will be suggestive to countries and governments.

JEL classification numbers: G30
Keywords: COVID-19, public opinion, Weibo, tweets, stock market.

---

[1] PBC School of Finance, Tsinghua University.
[2] PBC School of Finance, Tsinghua University.

Article Info: *Received:* March 10, 2020. *Revised:* March 25, 2020.
Published online: May 1, 2020.

1. Introduction

Both MERS and SARS are known to be caused by coronavirus. Coronavirus is a large family of viruses that can infect humans, mice, pigs, cats, dogs, cattle, mink, bats, camels, hedgehogs and other mammals as well as a variety of birds. It was first discovered in the 1960s. The virus was formally named a coronavirus in 1975 because of its spike shell, which looks like a crown.

The coronavirus is about 80 to 120 nanometers in diameter and is the largest known RNA virus. Coronaviruses are divided into four genera, among which the coronaviruses of SARS and MERS, which cause severe diseases and are able to spread widely and cause epidemics, belong to the phylum coronaviruses.

In December 2019, pneumonia of unknown cause appears in several medical institutions in Wuhan, Hubei. As of December 31, 2019, twenty-seven cases of viral pneumonia were found, all of which were diagnosed as viral pneumonia/pulmonary infection.

On January 7, 2020, the laboratory isolated the virus-like strain from the sample of a positive patient, and observed the virus particle morphology under the electron microscope. On January 10, laboratory staff obtained the full genome sequence of the virus. By sequencing the virus's genome, the scientists found that the pathogen that caused the Wuhan pneumonia was the closest relative to a SARS-like coronavirus found in bats, followed by the closest. But it has formed a new branch in the evolutionary relationship, so it is considered a new type of coronavirus. The virus was named COVID-19 by the World Health Organization on January 12, 2020.

The virus and pneumonia were first discovered in China in late 2019. Because of China’s high population density and the unique time node (Lunar Year and Spring Festival travel rush subsequently), it is really hard to prevent and control the increase and spread of the new virus. As cases of the new virus are increasing, more and more people voice their opinion in public platform and social media. Weibo, the China’s largest personal social platform, reflects people’s mood toward this big event. The discussion amounts will have explosive increase with the virus event going on, and considering its wide spread and great influence in mainland China, we believe that Weibo’s sentiment index calculated by users’ text and tone behind it, will affect China’s stock market. And we truly find evidence about that.

The rest of the paper is organized as follows. Section 2 discusses the economic background and the related literature. Section 3 discusses event development and data selection. Section 4 presents the empirical results. Section 5 discusses and concludes.

2. The Economic Background and Literature Review

The time of the virus’ outbreak is really special for China. It is about twenty days before the most important festival, Spring Festival, for Chinese people. Almost everyone will go home to reunite with family at that time, and because many people work/study far from home province, they have to take train or plane back home and return to workplace after the festival, which is named Spring Festival travel rush.

Chinese people also have many recreational activities during the festival holiday, such as watching movies and dinning together with friends. However, central government discourage both the travel rush and recreational activities this year, which must have influence on the economic development during the period. On the other hand, government and the central bank actively respond to this event, releasing liquidity to enhance public confidence. Thus, the impact of this virus event on Chinese economy may not to be too bad.

It is well known that public opinion has an impact on company performance and the economy in China. Xu, Yang and Huang (2020) used public opinions on tourism in China in 2010–2017 as exogenous shocks, and found that firms located in provinces where public opinions broke out experience significant decreases in cumulated abnormal returns. Ni, Su, Wang and Ying (2019) showed that both general sentiment of individual investors and the evaluation of consulting institutions surely influence the change of stock prices, and the change of stock prices can be expressed by the public opinion well. Scholars also discovered that network public opinions have a potential impact on vegetable prices volatility and can be treated as potential factor to predict vegetable prices, such as Li, Zhou and Lin (2019).

When focusing on personal social platform's effect, Zu, Diao and Meng (2019) found that launching Weibo accounts does indeed enhance the performance of firms, and greater Weibo input intensity leads to a greater improvement in performance. Zhang, Yuan and Wu (2019) showed that the variation tendency of Sina Weibo Index is highly correlated with stock market volatility, especially during the boom and crisis periods. Sun, Liu, Chen, Hao and Zhang (2019) also verified that news media attention plays a positive moderating effect in the relationship between investor attention and the stock return and social interaction could positively moderate the effect of news media's and investor's sentiments on stock return.

Our paper makes a number of contributions to the corporate financial theory: First, our data is newly hand-collected from Sina Weibo's website, and it is not just focusing on numbers and attentions, but in terms of tone and attitude of all pneumonia-related tweets. Second, the virus COVID-19 has spread many countries in year 2020 and great attention has been attached to it, our study providing efficient evidence that country's public opinion about it will indeed influence economic development.

3. Event Development and Data Selection

3.1 Event Development

Since first appearance in December 2019, the development of pneumonia has undergone a rapid and complex process. Table 1 presents the timeline of events related to the pneumonia.

Table 1: Timeline of pneumonia events

Day	Related Events
2020/12/8	The first confirmed case of the new coronavirus pneumonia got symptomatic.
2019/12/29	The epidemiological investigation began in Wuhan.
2019/12/31	The expert group of the National Health Commission arrived in Wuhan to carry out relevant inspection and verification work.
2020/1/3	A total of 44 cases of "viral pneumonia with unknown causes" were found, but no "obvious human-to-human transmission" or "medical care infection" were found, according to the Wuhan Health Commission.
2020/1/8	The expert group of National Health Commission preliminarily confirmed the virus COVID-19 as the pathogen of this epidemic.
2020/1/19	National Health Commission has announced the start of decentralizing testing kits.
2020/1/20	Xi has given important instructions on the pneumonia epidemic caused by the virus COVID-19.
	Zhong confirmed the new virus has human-to-human transmission in press conference of National Health Commission.
2020/1/22	Hubei provincial government has decided to launch a second-level emergency response to a public health emergency.
2020/1/23	Wuhan was closed at 10 in the morning.
2020/1/26	The State Treasury has allocated 11.21 billion yuan in subsidies for epidemic prevention and control.
2020/2/2	Vulcan mountain hospital was formally delivered.
2020/2/15	State Administration for Market Regulation introduced ten measures to resume production and production.
2020/2/20	The State Council released three measures to help enterprises tide over the difficulties.

From Table 1 we can see that the peak of the pneumonia epidemic is in late January, and has gradually eased since February in mainland China. In the same vein, the discussion on Weibo about the outbreak of pneumonia continues to simmer.

3.2 Data Selection

3.2.1 Weibo Data Selection

The Weibo data we used was from January 1 to February 23, covering most of the development of pneumonia. We use keywords related to the event as well as the virus and most attacked regions, such as “冠状” (Corona), “疫情” (Outbreak), “武汉” (Wuhan) and “肺炎” (Pneumonia). Total keywords are listed in Table 2.

Table 2: Keywords used in drawing tweets

Keywords	
SARS	钟南山 (Zhong Nanshan)
冠状 (Corona)	武汉 (Wuhan)
小汤山 (Xiaotangshan)	肺炎 (Pneumonia)
雷神山 (Leishenshan)	疫情 (Outbreak)
火神山 (Huoshenshan)	肺炎一号 (Pneumonia No.1)
李兰娟 (Li Lanjuan)	瑞德西韦 (Remdesivir)
李文亮 (Li Wenliang)	重大突发公共卫生事件一级响应 (First-level response to major public health emergencies)

Figure 1 presents the amount change as time going on of total tweets, original tweets and comments. The left axis shows total and original tweets’ rapid growth while the right one reflects the total comments for all Weibo tweets. We can see all the three indexes had explosive increase at around January 20, very closed to the date of outbreak of pneumonia.

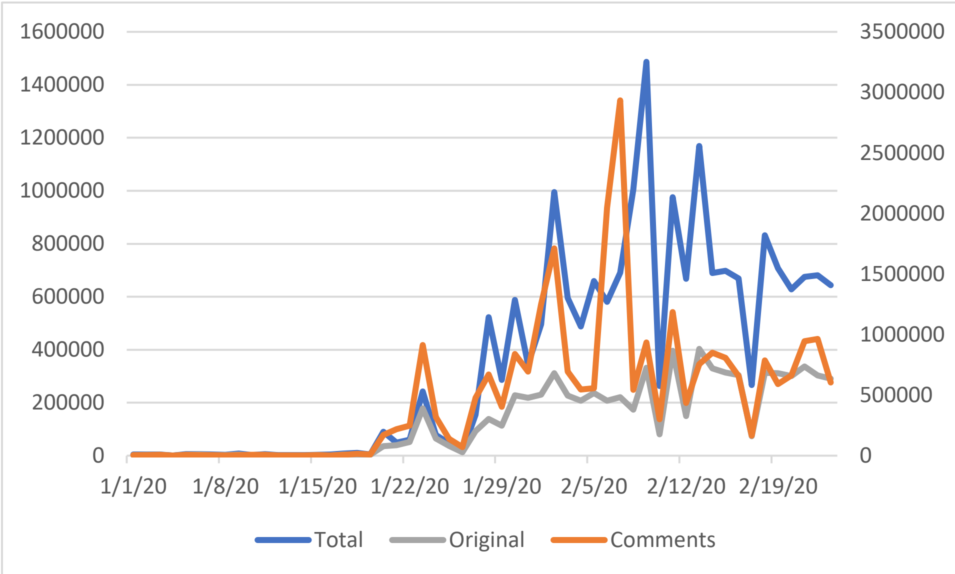

Figure 1: Total amounts for tweets and comments

We also count the number of tweets in different tone, that is, positive, negative and neutral. It is shown that neutral tweets' number is highly below both positive tweets and negative tweets, indicating people have own attitude toward this event. More important, we find that positive tweets' number is stably above negative tweets since February 6, stating that people's confidence is largely encouraged after government's active movements.

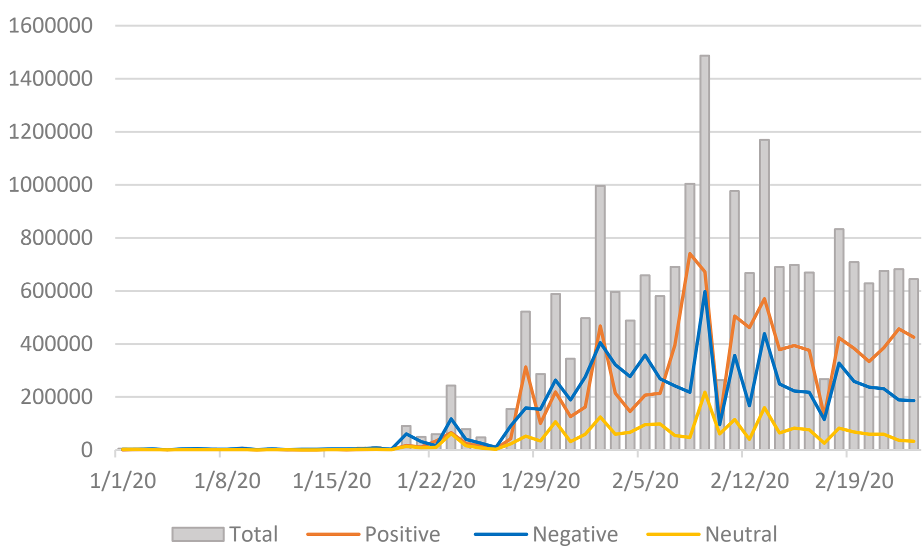

Figure 2: Tweets tone classification

3.2.2 Company and Market Data Selection

The sample examined in this paper includes all Chinese listed firms in China Stock Market Accounting Research (CSMAR) database during the period of January 1 to February 23, 2020. The address information for all companies is from Wind database, and the stock return and all other firm's financial variables are obtained from CSMAR database or CNRDS database, which two are the biggest databases providing listed firms' information in China.

The core independent variable is the public opinion measured by Weibo data, and the core dependent variable is the stock return meanwhile. There are also a set of control variables which may affect stock return as well, and we add them all in the regressions. The control variables in the baseline regressions include firm size, measured by the natural logarithm of firm total assets; firm age, measured by the natural logarithm of one plus the number of years the firm is listed; firm's book-to-market ratio; ownership of property, equals 1 if the firm is state-owned and 0 otherwise; firm's liability risks, measured by total debt-to-total assets; profitability, measured by return on assets. We also control province fixed effect and industry fixed effect, according to CSRC 2012 industry classification. We provide detailed variable definitions in Table 3.

Table 3: Definition of variables

Variable	Definition
Measure of dependent variables	
stock_return	Firm *i*'s stock return (scaled by 100) in day *t*.
lstock_return	Firm *i*'s stock return (scaled by 100) in day *t+1*.
Measure of independent variables	
posi_nega_ratio	The ratio of positive tweets to negative tweets in day *t*.
posi_total_ratio	The ratio of positive tweets to total tweets.
Measure of control variables	
lnasset	Natural logarithm of the firm's total assets in year 2019.
lage	Natural logarithm of one plus the firm's listed years in year 2019.
leverage	Firm's total debt-to-total assets ratio in year 2019.
SOE	A binary variable equal one if the firm is a state-owned enterprise in year 2019.
ROA	Firm's return on assets ratio in year 2019.
BM	Firm's book-to-market ratio in year 2019.
province	Firm's business address in year 2019.
industry	Firm belongs to which industry according to CSRC 2012 industry classification.

4. Empirical Results

We first confirm that there's no significant multicollinearity in the regressions. Table 4 presents the correlation matrix for each of our variables. Panel A includes the variables used in the baseline OLS model with $posi\_nega\_ratio$ while panel B includes the variables used with $posi\_total\_ratio$. Because in this analysis there are significant inter correlations among some variables, we computed variance inflation factors (VIFs) to investigate whether there was a potential multicollinearity problem. The mean VIF in the two models are both 1.43, hardly below the rule-of-thumb cutoff of 10 (Ryan 1997). Therefore, we do not believe that multicollinearity significantly affects our results.

Table 4: Correlations for each variable

Panel A: Correlation Matrix for variables with posi_nega_ratio							
	posi_nega_ratio	**lnasset**	**lage**	**leverage**	**SOE**	**ROA**	**BM**
posi_nega_ratio	1.000						
lnasset	-0.001	1.000					
lage	0.000	0.328	1.000				
leverage	-0.000	0.395	0.213	1.000			
SOE	0.000	0.343	0.426	0.161	1.000		
ROA	-0.000	0.020	-0.138	-0.416	-0.046	1.000	
BM	-0.001	0.591	0.211	0.270	0.241	-0.108	1.000
Panel B: Correlation Matrix for variables with posi_total_ratio							
	posi_nega_ratio	**lnasset**	**lage**	**leverage**	**SOE**	**ROA**	**BM**
posi_total_ratio	1.000						
lnasset	-0.001	1.000					
lage	0.001	0.328	1.000				
leverage	-0.000	0.395	0.213	1.000			
SOE	0.000	0.343	0.426	0.161	1.000		
ROA	-0.000	0.020	-0.138	-0.416	-0.046	1.000	
BM	-0.001	0.591	0.211	0.270	0.241	-0.108	1.000
This table reports the correlations for variables used in baseline regressions constructed based on the sample of Chinese public firms during January and February 2020. Definitions of the variables are listed in Table 3.							

To assess how Weibo tweets' tone affects firm's stock return, we estimate the following model:

$$stock\_return_{i,t} = \alpha + \beta \cdot posi\_nega\_ratio_t + \gamma \cdot Z_i + industry_i + province_i + \varepsilon_{i,t} \quad (1)$$

where *i* indexes firm and *t* indexes time. The dependent variable captures firm's market performance. The core independent variable, $posi\_nega\_ratio_t$, is the ratio of positive tweets to negative tweets in day t. Z is a vector of firm that may affect a firm's stock return as we discussed in Section 3.2.2. Industry and province capture industry fixed effects and region fixed effects, respectively.

In order to make the results more readable in magnitudes, we make the stock return multiplied by 100 and it will not generate other effects. Table 5 reports the OLS regression results with different fixed effects. The coefficient estimates of $posi\_nega\_ratio_t$ are positive and significant in all four columns, suggesting that positive tweets' tone indeed has positive effect on the whole firms' stock returns. Numerically analyzing, when the ratio of positive tweets to negative tweets increase by 0.1, every firm's stock return will increase by above 0.04%.

Table 5: Regressions of stock return on Weibo tweets' tone

	(1)	**(2)**	**(3)**	**(4)**
Dep. Var	**stock_return**	**stock_return**	**stock_return**	**stock_return**
posi_nega_ratio	0.436***	0.436***	0.436***	0.436***
	(0.015)	(0.015)	(0.015)	(0.015)
lnasset	0.024**	0.021**	0.032***	0.033***
	(0.009)	(0.010)	(0.011)	(0.011)
BM	-0.460***	-0.458***	-0.204***	-0.214***
	(0.049)	(0.050)	(0.056)	(0.056)
leverage	-0.110**	-0.105**	0.017	0.028
	(0.049)	(0.049)	(0.051)	(0.051)
ROA	0.397**	0.411**	0.832***	0.831***
	(0.159)	(0.161)	(0.164)	(0.166)
lage	-0.058***	-0.046***	-0.048***	-0.042***
	(0.015)	(0.015)	(0.016)	(0.016)
SOE	-0.142***	-0.128***	-0.091***	-0.083***
	(0.025)	(0.026)	(0.027)	(0.028)
Constant	-0.152	-0.149	-0.433	-0.435
	(0.183)	(0.192)	(0.266)	(0.278)
Industry fixed effects	No	No	Yes	Yes
Province fixed effects	No	Yes	No	Yes
Observations	110,400	110,400	110,400	110,400
R-squared	0.010	0.011	0.014	0.014

This table reports regressions of the stock return on Weibo tweets' tone and other control variables. Definitions of control variables are listed in Table 3. Each regression includes a separate intercept. Robust standard errors are displayed in parentheses. ***, **, and * indicate significance at the 1, 5, and 10 percent levels, respectively.

We change the independent variable's measure to assess our result is robust:

$$stock\_return_{i,t} = \alpha + \beta \cdot posi\_total\_ratio_t + \gamma \cdot Z_i + industry_i + province_i + \varepsilon_{i,t} \quad (2)$$

where the core independent variable, $posi\_total\_ratio_t$, is the ratio of positive tweets to total tweets in day t. All other settings are the same as equation (1) and the results are reported in Table 6. Similar to the above, coefficient estimates of $posi\_total\_ratio_t$ are positive and significant in all four columns, and the numerical value is larger than $posi\_nega\_ratio_t$ in equation (1), which indicates that compared to the relation between positive mood and negative mood, market and investors pay more attention to the ratio of positive information in the whole sample.

Table 6: Regressions of stock return on Weibo tweets' tone-another measure

	(1)	**(2)**	**(3)**	**(4)**
Dep. Var	**stock_return**	**stock_return**	**stock_return**	**stock_return**
posi_total_ratio	1.060***	1.060***	1.060***	1.060***
	(0.053)	(0.053)	(0.053)	(0.053)
lnasset	0.024**	0.021**	0.032***	0.033***
	(0.009)	(0.010)	(0.011)	(0.011)
BM	-0.460***	-0.458***	-0.205***	-0.214***
	(0.050)	(0.050)	(0.056)	(0.056)
leverage	-0.110**	-0.105**	0.017	0.028
	(0.049)	(0.049)	(0.051)	(0.051)
ROA	0.397**	0.411**	0.832***	0.831***
	(0.159)	(0.162)	(0.164)	(0.166)
lage	-0.058***	-0.046***	-0.048***	-0.042**
	(0.015)	(0.015)	(0.016)	(0.016)
SOE	-0.142***	-0.128***	-0.091***	-0.083***
	(0.025)	(0.026)	(0.027)	(0.028)
Constant	-0.140	-0.137	-0.420	-0.423
	(0.184)	(0.193)	(0.267)	(0.279)
Industry fixed effects	No	No	Yes	Yes
Province fixed effects	No	Yes	No	Yes
Observations	110,400	110,400	110,400	110,400
R-squared	0.006	0.006	0.010	0.010

This table reports regressions of the stock return on Weibo tweets' tone and other control variables. Definitions of control variables are listed in Table 3. Each regression includes a separate intercept. Robust standard errors are displayed in parentheses. ***, **, and * indicate significance at the 1, 5, and 10 percent levels, respectively.

Another concern is that there may be some synchronism problems if we use Weibo tweets tone to explain stock return in the same day. We lag the return for one market day to solve the potential problem:

$$stock\_return_{i,t+1} = \alpha + \beta \cdot posi\_nega\_ratio_t + \gamma \cdot Z_i + industry_i + province_i + \varepsilon_{i,t} \quad (3)$$

where the dependent variable, $stock\_return_{i,t+1}$, is the market return at the day after Weibo tweets' collection. Table 7 reports the OLS regression results with different fixed effects. We can see that the ratio of positive tweets to negative tweets still has positive effect on stock return of the day after, though the magnitude is somehow smaller. That's to say, this significant positive influence on stock return is not (only) due to synchronism.

Table 7: Regressions of lagged stock return on Weibo tweets' tone

	(1)	**(2)**	**(3)**	**(4)**
Dep. Var	**lstock_return**	**lstock_return**	**lstock_return**	**lstock_return**
posi_nega_ratio	0.307***	0.307***	0.307***	0.307***
	(0.016)	(0.016)	(0.016)	(0.016)
lnasset	0.032***	0.032***	0.048***	0.050***
	(0.011)	(0.011)	(0.013)	(0.014)
BM	-0.484***	-0.493***	-0.239***	-0.251***
	(0.058)	(0.059)	(0.066)	(0.067)
leverage	-0.113**	-0.109*	0.034	0.046
	(0.058)	(0.058)	(0.060)	(0.061)
ROA	0.378**	0.374**	0.761***	0.759***
	(0.188)	(0.191)	(0.194)	(0.196)
lage	-0.064***	-0.052***	-0.056***	-0.049**
	(0.018)	(0.018)	(0.019)	(0.019)
SOE	-0.184***	-0.167***	-0.128***	-0.117***
	(0.030)	(0.031)	(0.032)	(0.033)
Constant	-0.101	-0.146	-0.509	-0.521
	(0.217)	(0.228)	(0.315)	(0.329)
Industry fixed effects	No	No	Yes	Yes
Province fixed effects	No	Yes	No	Yes
Observations	85,448	85,448	85,448	85,448
R-squared	0.007	0.007	0.011	0.011

This table reports regressions of the lagged stock return on Weibo tweets' tone and other control variables. Definitions of control variables are listed in Table 3. Each regression includes a separate intercept. Robust standard errors are displayed in parentheses. ***, **, and * indicate significance at the 1, 5, and 10 percent levels, respectively.

Thus, we find overall that Weibo tweets' mood indeed affect stock market, and positive tweets' tone will lead to a higher stock return.

5. Discussion and Conclusion

This study focuses on whether and how public opinion in Weibo tweets about COVID-19 virus pneumonia will affect Chinese stock market and companies' stock returns. Our results point to a positive channel for the relation.

Weibo tweets about pneumonia explosively increase at late January, 2020, of which most were in negative mood. However, because Chinese government's active movements, the proportion of positive tweets gradually rise after early February. Besides amount growth, we also find that the ratio of positive tweets to negative/total tweets has positive effect on stock market, and it is not caused by synchronism. The reason behind it may be that positive public opinion reflects investors' confidence, and will encourage the stock return to increase.

Overall, our study indicates that public opinion reflected in social media platform indeed has significant effect on country's stock market and economy development, if there appears national event spread to the whole country. The result will be suggestive to countries and governments.

References

[1] Chemmanur, T. J., & Yan, A., Advertising, attention, and stock returns. Quarterly Journal of Finance, 9(03), (2019), 1950009.

[2] Chen, J., Zhou, H., Hu, H., Song, Y., Gifu, D., Li, Y., & Huang, Y., Research on agricultural monitoring system based on convolutional neural network. Future Generation Computer Systems, 88, (2018), 271-278.

[3] Deng, S., Huang, Z. J., Sinha, A. P., & Zhao, H., The interaction between microblog sentiment and stock return: An empirical examination. MIS quarterly, 42(3), (2018), 895-918.

[4] Hu, X., Tang, L., Tang, J., & Liu, H., Exploiting social relations for sentiment analysis in microblogging. In Proceedings of the sixth ACM international conference on Web search and data mining, (2013), (pp. 537-546).

[5] Li, Y., Zhou, H., Lin, Z., Wang, Y., Chen, S., Liu, C., ... & Xia, J., Investigation in the influences of public opinion indicators on vegetable prices by corpora construction and WeChat article analysis. Future Generation Computer Systems, 102, (2020), 876-888.

[6] Ni, Y., Su, Z., Wang, W., & Ying, Y., A novel stock evaluation index based on public opinion analysis. Procedia computer science, 147, (2019), 581-587.

[7] Runpeng, H., Wenming, Z., & Lingyan, B. · Predicting the Stock Market Based on Microblog Mood. Journal of Industrial Engineering and Engineering Management, 29(1), (2015), 47-52.

[8] Schlagwein, D., & Prasarnphanich, P., Organizational Social Media around the GLOBE. In 2011 IEEE Ninth International Conference on Dependable, Autonomic and Secure Computing, (2011), (pp. 924-931). IEEE.

[9] Sun, Y., Liu, X., Chen, G., Hao, Y., & Zhang, Z. J., How mood affects the stock market: Empirical evidence from microblogs. Information & Management, (2019), 103181.

[10] Tang, L. Y. N., Zhang, Y. M., Dai, F., Yoon, Y. J., & Song, Y. Q., Sentiment Analysis for the Construction Industry: A Case Study of Weibo in China. In Computing in Civil Engineering 2017, (2017), (pp. 270-281).

[11] Zhang, T., Yuan, Y., & Wu, X., Is microblogging data reflected in stock market volatility? Evidence from Sina Weibo. Finance Research Letters, 32, (2020), 101173.

[12] Zhou, Z., Zhao, J., & Xu, K., Can online emotions predict the stock market in China?. In International conference on web information systems engineering, (2016), (pp. 328-342). Springer, Cham.

[13] Zu, X., Diao, X., & Meng, Z., The impact of social media input intensity on firm performance: Evidence from Sina Weibo. Physica A: Statistical Mechanics and its Applications, 536, (2019), 122556.