Research on the Fluctuation Characteristics of Social Media Message Sentiment with Time Before and During the COVID-19 Epidemic

Guo, Chaohui; Lin, Shaofu; Huang, Zhisheng; Shi, Chengyu; Yao, Yahong

published in
Health Information Science
2021

DOI (link to publisher)
10.1007/978-3-030-90885-0_1

document version
Publisher's PDF, also known as Version of record

document license
Article 25fa Dutch Copyright Act

Research on the Fluctuation Characteristics of Social Media Message Sentiment with Time Before and During the COVID-19 Epidemic

Chaohui Guo[1], Shaofu Lin[1,2](✉), Zhisheng Huang[3], Chengyu Shi[1], and Yahong Yao[1]

[1] Faculty of Information Technology, Beijing University of Technology, Beijing 100124, China
linshaofu@bjut.edu.cn

[2] Beijing Institute of Smart City, Faculty of Information Technology, Beijing University of Technology, Beijing, China

[3] Department of Computer Science, Vrije University Amsterdam, Amsterdam, The Netherlands

Abstract. "Tree hole" refers to a social media formed after the death of a social media user, in which other users continue to leave messages due to emotional resonance. This paper focuses on exploring the fluctuation of emotions with time in a "tree hole" of social media such as Microblog, and provides ideas and support for suicide warning, rescue, and user portraits of patients with depression in the "tree hole". In this paper, the dataset of 2,356,066 messages captured from the "tree hole" Microblog with the "tree hole" agent (i.e., an AI program) and pre-processed. Subsequently, the effective dataset was labeled by a text sentiment analysis model based on BERT and BiLSTM, and accordingly the sentiment was scored. Then the scored data was visualized and analyzed in the time dimension. Finally, it was found that the sentiment of the "tree hole" messages reached a trough at 4:00 am and a peak around 8:00 am. In addition, the overall trend of "tree hole" sentiment has fluctuated downwards from Monday to Sunday. We have concluded that the sentiment of patients with depression fluctuates regularly at some special time points, and special events such as the outbreak of COVID-19 and so on, have a great impact on the emotions of patients with depression. Therefore, it is necessary to strengthen warning and intervention for those who has expressed thoughts of suicide at special points to prevent the spread and fermentation of suicidal emotions in the "tree hole" in time. In addition, the rescue volunteers for patients with depression as Tree Hole Rescue Team should make corresponding adjustments to the rescue strategy when special events occur. This research is of great significance for the emergency response of "tree hole" depressed users in major events such as COVID-19 epidemic.

Keywords: Microblog tree hole · Message sentiment · COVID-19 epidemic · Scrapy-Redis · BERT-BiLSTM

This research is partially funded by the high-level foreign talent fund program of Beijing (No. C202004).

S. Siuly et al. (Eds.): HIS 2021, LNCS 13079, pp. 3–14, 2021.
https://doi.org/10.1007/978-3-030-90885-0_1

1 Introduction

Depression has become a common disorder. More than 300 million people are living with depression all around the world. It can happen to anybody and begin at any time [1]. According to the 2019 China Mental Health Survey, major depressive disorder was the most prevalent sentiment disorder [2]. Depression brings huge burdens and sufferings to individuals and their friends, leading to an extremely high suicide rate among depressed patients [3]. Close to 800,000 people die due to suicide every year [4]. According to the United Nations report, suicide has become the second leading cause of unnatural death among young people [5]. In addition, more and more young people are addicted to the Internet. As of December 2020, the number of Chinese Internet users has reached 989 million, and the Internet penetration rate has reached 70.4% [6]. With the increasing number of users in the virtual world, online social media has gradually become the main way of socializing for people [7]. The freedom of expression and the instant delivery feature of Microblog [8] have allowed many young patients with depression to express their suicidal feelings and wishes through this platform.

The term "tree hole" comes from a fairy tale. The barber knew that the emperor had a pair of donkey ears, but was afraid to tell others the secret. So, he told the secret to a hole in a big tree on the mountain [9]. With the popularity of the Internet, more and more people are inclined to tell their secrets through online social media platforms. Many patients with depression often express their real thoughts through social media platforms (such as Microblog) when they have suicidal thoughts, and thus the social media platforms turn to be "tree holes" in the virtual world. On March 17, 2012, the depressed patient with a nickname "Zou Fan" committed suicide after posting the last message on her Microblog account. Since then the message area of the Microblog account of "Zou Fan" has become a "tree hole" for many patients with depression in the virtual world. From 2012 to 2020, "Zou Fan Tree Hole" has more than 2.3 million messages, and it has become the largest "tree hole" in Microblog. Professor Zhisheng Huang has developed a "tree hole" agent (i.e., an AI program) that patrols large "tree holes" in social media and automatically screens people with obvious suicidal tendencies [10].

In different time periods, the sentiments of Microblog "tree hole" users are various, especially when a special event occurs during a specific time period, the overall sentiment of the Microblog "tree hole" will fluctuate, such as the outbreak of COVID-19 epidemic. Most previous analysis of "tree hole" message data was only aimed at the time characteristics of the number of messages [11, 12], whereas the time characteristics of the sentiment of the "tree hole" messages were less analyzed. In this research, we mainly scored the emotional characteristics of the messages from the largest "tree hole" on Microblog, the "Zou Fan Tree Hole", and analyzed the fluctuation of the emotion of the messages in the "tree hole" in time dimension. Furthermore, we analyzed the reasons for "tree hole" emotion fluctuations with time, and provided ideas and data support for suicide warning and intervention.

2 Methods

2.1 Models

Model Selection

There are two main models of sentiment classification. One is based on deep learning, and the other is based on dictionary [13]. The characteristics of the "tree hole" message include short text, large number of messages, and complex thoughts expressed.

The sentiment classification model based on dictionary needs to expand a large number of labelled vocabularies, meanwhile the analysis effect of short text is poor. In addition, the sentiment classification model based on dictionary cannot analyze the relationship between vocabularies and context in the message text. The sentiment classification model based on deep learning have strong generalization ability and outstanding effect on short text analysis. And it also takes into account the order and semantic characteristics of words. Therefore, the sentiment classification model based on deep learning is more suitable for analyzing the emotion of "tree hole" messages.

BERT (Bidirectional Encoder Representations from Transformers)

Since the Microblog comment texts are short texts, the text vectorization layer of this sentiment polarity analysis model does not use the commonly used word2vec [14] text vector representation. And word2vec takes words as processing unit, which needs to go through text preprocessing, feature extraction, feature vector representation, vector stitching, and finally to generate a vector representation of the text. However, BERT offers an advantage over the traditional models such as word2vec. BERT model uses the characters as the processing unit and maps each word into the form of an n-dimensional word vector.

BiLSTM (Bi-directional Long Short-Term Memory)

BiLSTM layer [14] is composed of two parts: forward LSTM and backward LSTM. If only the LSTM model is used to determine the polarity of message sentiment, there will be an inability to encode the message from backward to forward. The structure of LSTM and BiLSTM are shown in Figs. 1 and 2:

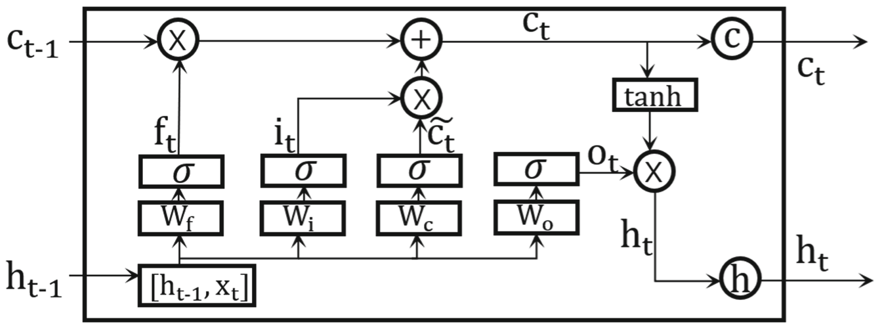

Fig. 1. Structure of LSTM

Fig. 2. Structure of BiLSTM

Compared with other neural networks, BiLSTM can extract text sequence information more effectively, so it has a much better performance in solving the problem of text sentiment classification. The gates in Fig. 1 are expressed in the equations as the following.

Equation (1) is Forget gate:

$$f_t = \sigma(U_f \cdot c_{t-1} + W_f x_t + b_f) \tag{1}$$

Equation (2) is Input gate:

$$i_t = \sigma(U_f \cdot h_{t-1} + W_f x_t + b_i) \tag{2}$$

Equation (3) is Output gate:

$$O_t = \sigma(U_o \cdot h_{t-1} + W_o x_t + b_o) \tag{3}$$

Equation (4) is Memory cell gate:

$$\tilde{c}_t = tanh(U_c \cdot h_{t-1} + W_c x_t + b_c) \tag{4}$$

Equation (5) is Memory cell gate:

$$c_t = f_t \circ c_{t-1} + i_t \circ \tilde{c}_o \tag{5}$$

Equation (6) is Final gate:

$$h_t = O_t \circ tanh(c_t) \tag{6}$$

Model Evaluation

This research uses the BERT-BiLSTM to predict the sentiment of the Treehole message. The accuracy of the BiLSTM model used in text analysis tasks is higher than that of the LSTM and the TextCNN [15]. In addition, because BERT utilizes a powerful Transformer mechanism, BERT-BiLSTM shows better results in the binary classification analysis of text sentiment, as shown in Table 1.

Table 1. Labeled dataset

Models	AccuracyMetric
BiLSTM	0.7548
BERT-BiLSTM	0.8941

2.2 Emotion Scoring Algorithm

This paper uses the mathematical expectation of calculating message sentiment to score each message. The higher the score, the more positive of the message, as shown in Eq. (7).

$$S = P \times \widehat{P} + N \times \widehat{N} \tag{7}$$

In this paper, for the convenience of calculation, the weight of P (Positive) is set to 1, and the weight of N (Negative) is set to 0. $\widehat{P}$ and $\widehat{N}$ refer to the positive and negative probabilities of the message sentiment calculated by the sentiment analysis model, respectively. S represents the total score of the message sentiment, the value range is [0,1]. And the closer S is to 1, the more positive the sentiment is.

Emotion Scoring Algorithm Based on Every Hour in a Day

$$y = \frac{\sum_{i=1}^{max} S(Hour = x)_i}{max} x \in [0, 23] \tag{8}$$

Where i, max and $Hour$ is the serial number, the largest serial number, and the Hour value of each data item in the dataset, respectively. And S is the Eq. (7) under the condition of $Hour = x$. Therefore, the value of y is the average of the sentiment scores of all messages in the dataset, where $Hour$ is equal to x. The value range of y is [0,1], and the sentiment at $y = 0.5$ is neutral.

Emotion Scoring Algorithm Based on Every Day in a Week

$$y = \frac{\sum_{i=1}^{max} S(Week = x)_i}{max} x \in [1, 7] \tag{9}$$

Where i, max and $Week$ is the serial number, the largest serial number, and the Week value of each data item in the dataset, respectively. And S is the Eq. (7) under the condition of $Week = x$. Therefore, the value of y is the average of the sentiment scores of all messages in the dataset, where $Week$ is equal to x. The value range of y is [0,1]. And the larger the value of y, the more positive the sentiment, and the sentiment at $y = 0.5$ is neutral.

3 Data Pre-processing and Sentiment Prediction

3.1 Data Source

The data source of this paper is the Microblog message of “Zou Fan”. The “tree hole” agent is used to capture messages sent by Microblog users through the “tree hole”. And the Scrapy-Redis distributed crawler [17] was written to complete and correct the missing and incorrect fields of the message data. The 2,356,066 messages from March 18, 2012 to August 31, 2020 finally constituted “Raw dataset 1.0”, as shown in Table 2.

Table 2. Raw dataset 1.0

Date	Time	Message-ID	User-ID	User-Name	Message	Other
2012-03-18	10:54	Null	123	Alan	Come back soon	Null
…	…	…	…	…	…	…
2020/08/31	21:29	123456	456	Bob	I will live	Chen XX

3.2 Data Preprocessing

In this experiment, we chose the Scrapy crawler framework to fix the errors and missing fields in the “Raw dataset 1.0”. Since the Scrapy crawler framework does not have distributed characteristics, this experiment used the Redis in-memory database with distributed characteristics to improve the Scrapy crawler framework. Finally, a distributed crawler framework with Scrapy-Redis is formed.

It can be observed that “Raw dataset 1.0” has many problems, such as inconsistent data format and invalid information in the message. In order to solve these problem, we wrote a program to process the “Raw dataset 1.0” as follows: (1) Normalize the time information of each data item in the data set. (2) Calculate the day of the week by year, month, and day. (3) Sort the message data in chronological order to form a “Preprocessed data set”, as shown in Table 3.

Table 3. Preprocessed dataset

Year	Month	Day	Hour	Minute	Week	User-ID	User-Name	Message
2012	03	18	10	54	7	123	Alan	Come back soon
…	…	…	…	…	…	…	…	…
2020	08	31	23	59	1	456	Bob	I will live

3.3 Sentiment Prediction

Training Dataset

This paper used the deep learning method to determine the sentiment polarity of the “tree hole” message, so a tremendous amount of labelled data is needed as the training set of the deep learning model. In this experiment, we randomly selected 100,000 pieces of data from the “Preprocessed dataset” as the “Labelled dataset”. The sentiment polarity of the messages in the “Labelled dataset” was automatically determined by a publicly available generic sentiment analysis tool [18], where the label “0” represents negative sentiment and the label “1” represents positive sentiment. However, public sentiment analysis tools are not effective in specific scenarios. This experiment used a combination of public sentiment analysis tools and manual screening. After labelling the “Labelled dataset” by sentiment analysis tool, we wrote rules to further filter the "Labelled dataset". The main rules are as follows: (1) Delete messages where the emotions represented by the emoji and the emotion represented by the labels are different; (2) Delete non-Chinese messages; (3) Delete messages composed of symbols; (4) Delete redundant messages, and only keep emotion labels and message texts. Finally, the "Labelled dataset" consisting of 40,000 messages with positive emotions and 40,000 messages with negative emotions was retained after screening, as shown in Table 4.

Table 4. Labelled dataset

Target	Message
1	Thanks for the hug. I'm much better. Good night
…	…
0	Suddenly want to die

Model Training and Output

The experiment chose a deep learning model based on BERT-BiLSTM for the sentiment polarity analysis. The vector dimension output after the BiLSTM layer is the specified units, which is different from the vector dimension of the label. Therefore, this experiment adds a fully connected layer after the BiLSTM model. Finally, the model outputs the results to the sentiment prediction output layer [19].

The sentiment prediction output layer of this model is the SoftMax function, which maps the input to real numbers between 0 and 1 and ensures that the sum of the probabilities of classification is exactly 1. In this experiment, the output of the SoftMax function is $(\hat{x}, \hat{y})$, where $\hat{x}$ is the probability of the message sentiment being negative, and $\hat{y}$ is the probability of the message sentiment being positive, and the sum of both is exactly 1.

Dataset Sentiment Prediction

In the experiment, the "Labelled dataset" was divided into training set, validation set and test set according to the ratio of 9:0.5:0.5.

This paper selected the BERT-BiLSTM as the training model to judge the sentiment polarity of all messages in the "Preprocessed dataset". And the "Sentiment prediction dataset" is obtained, as shown in Table 5. The value of Emotion represents the sentiment polarity of the message, which is divided into positive emotion (1) and negative emotion (0), $\widehat{P}$ and $\widehat{N}$ represent the probability of positive emotion and negative emotion of the message, respectively. The prediction results of the BERT-BiLSTM model are shown in Table 5.

Table 5. Sentiment prediction dataset

Year	Month	Day	Hour	Minute	Week	Message	Emotion	$\widehat{P}$	$\widehat{N}$
2012	03	18	10	54	7	Come back soon	1	0.272	0.728
…	…	…	…	…	…	…	…	…	…
2020	08	31	23	59	1	I will live	1	0.273	0.727

4 Experimental Results and Analysis

4.1 Analysis of "Tree Hole" Message Sentiment in a Day

In order to study the impact of the COVID-19 epidemic on the sentiment of the "tree hole", the "Sentiment prediction dataset" was divided into two parts, that is, the data from 2012 to 2019 and the data formed after the outbreak of the COVID-19 in 2020.

This paper uses Eq. (8) to score the information sentiment of every one hour period of a day and plot the results. The result is shown in Fig. 3.

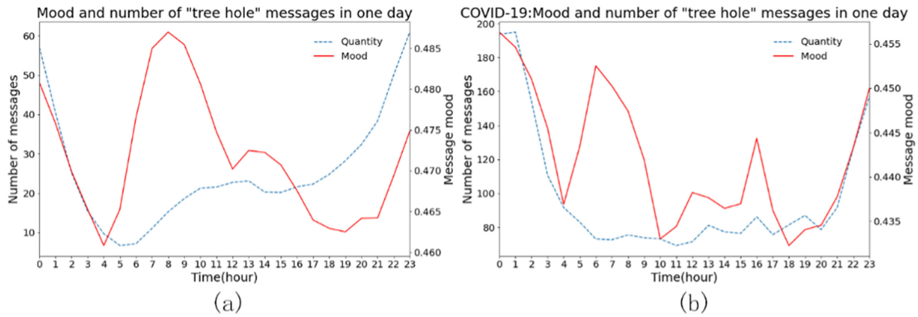

Fig. 3. The number and the fluctuation of sentiment of "tree hole" messages during 24 h a day before the COVID-19 epidemic (a) and during the epidemic (b)

It can be seen that there is obvious difference between the sentiment of the "tree hole" on Microblog before the COVID-19 epidemic and during the epidemic.

1. In daily life before the COVID-19 epidemic, the lowest point of "tree hole" sentiment in a day appeared around 4 am. The highest point of "tree hole" sentiment appeared in the early morning from 7 am to 9 am. After that the sentiment in the "tree hole" gradually declined until 12 at noon when there was a rebound, and reached the peak around 1 pm. After 1pm, the sentiment in the "tree hole" gradually declined, and reached the second low point of sentiment around 6 pm. Then the sentiment in the "tree hole" started to rise rapidly and reached to the peak at about 1 am.
2. From the above curves, it can be found that negative emotions dominate in the "tree hole". And the emotions in the "tree hole" during the epidemic were much lower than those in the "tree hole" before the epidemic. The peak point of "tree hole" emotion

in a day during the epidemic became around 0:00 a.m., while the lowest point was around 10 am.

3. The number of messages was correlated with the sentiment in the "tree hole". The sentiment of the messages was also relatively positive during the time period when the number of messages was high, and conversely the sentiment of the messages was relatively low. However, there was an anomaly in the early morning from 7am to 8 am. The sentiment turned to be relatively positive in that time period, but the number of messages was low.

4.2 Analysis of "Tree Hole" Message Sentiment in a Week

After analyzing the sentiment fluctuations of the "tree hole" within a day, we continued to draw and analyze the weekly sentiment fluctuations of the "tree hole" before and during COVID-19 epidemic.

This paper uses Eq. (9) to score the information sentiment of each time period of the day and plot the results. The result is shown in Fig. 4. The x-axis represents Monday to Sunday, and the y-axis represents the "tree hole" sentiment score.

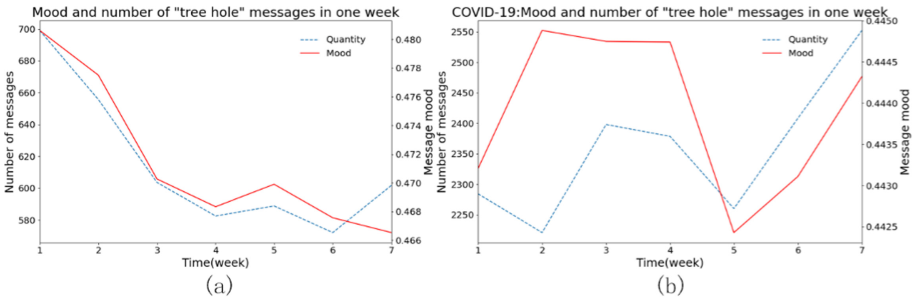

Fig. 4. The number and the fluctuation of sentiment of "tree hole" messages during a week before the COVID-19 epidemic (a) and during the epidemic (b)

As can be seen from Fig. 4, the situation of messages and sentiment swings within a week during the epidemic differed significantly from that during the epidemic.

1. In general, the sentiment in the "tree hole" of Microblog fluctuates and decreases throughout the week. The highest point of the "tree hole" sentiment usually appeared on Monday, and this sentiment gradually declined day by day until a brief rebound on Friday and then continued to go down on weekend.
2. The highest point of "tree hole" sentiment during the epidemic was also much lower than the lowest point of "tree hole" sentiment before the epidemic. The sentiments of "tree hole" users during the epidemic reached its lowest point on Friday, the opposite of the daily situation.
3. Under normal circumstances, the number of "tree hole" messages in a week showed a downward trend in daily situations. During the epidemic, the number of "tree hole" messages showed big fluctuations in a week. The number of "tree hole" messages during the epidemic was less on Tuesday and Friday, and more on weekends.

5 Discussion and Suggestions

5.1 Possible Causes of Sentiment Swings of Microblog "Tree Hole" Messages in General

As shown in Figs. 3a and 4a, we can infer that the possible reasons for forming the sentiment fluctuation pattern of the "tree hole" messages in daily situation are as follows.

1. The sentiment in the "tree hole" fluctuates regularly over time. But the negative sentiment always dominates in the "tree hole" messages.
2. In the tree hole, users are most negative at 4 am, and most active at around 8 am. This shows that "tree hole" users who stay up late and have insomnia will have a lot of negative sentiment. On the contrary, after a moderate amount of sleep, the sentiment of "tree hole" users will be relieved, leading to a rise in emotions of the "tree hole". This shows that sleep condition has a direct impact on the sentiment of the "tree hole" users.
3. Around 6 pm is also the time when there are more negative emotions in the "tree hole" message, because people need to vent their negative emotions during a day of study and work. This proves that students and office workers who are confronted with great pressures are the dominant players of negative emotions in the "Zou Fan Tree Hole".
4. The sentiment of the "tree hole" messages tend to fluctuate downwards during the week. However, due to the approach of the weekend, the emotional score in the "tree hole" messages will rise on Friday. This further supports the fact that students and office staff under pressure are the dominant negative emotions in the "tree hole".

5.2 Possible Causes of Sentiment Swings in Microblog "Tree Hole" Messages During the COVID-19 Epidemic

As shown in Figs. 3b and 4b, we can infer the sentiment fluctuation pattern and possible causes of the "tree hole" messages during the COVID-19 epidemic are as follows.

1. After the outbreak of the COVID-19 epidemic, the sentiments of the "tree hole" has changed significantly, showing a trend different from previous fluctuations.
2. During the epidemic, the daily sentiment swing in the "tree hole" became more dramatic. This led to four troughs in the sentiment of the messages every day. These four troughs appeared at 4 a.m., 10 a.m., 3 p.m., and 6 p.m.
3. During the epidemic, "tree hole" sentiment was lower than normal throughout the week.

5.3 Suggestions

Based on the above analysis, the following suggestions are made for suicide warning and rescue of "tree hole" users.

1. It is suggested that the Tree Hole Rescue Team should enhance suicide prevention and control mechanisms based on monitoring and warning suicide information. Mainly aimed at the "double low" period when mood is extremely low and the number of people is small, and passive early warning is transformed into active intervention.
2. The sensitivity of suicide warning in "tree hole" is adjusted dynamically according to the changing law of message mood. It is recommended to allocate more rescue resources at the lowest emotional point of the "tree hole" within a certain period.
3. Increase the attention on the "tree-hole" users who send messages early in the morning, maybe staying up late and suffering from insomnia. Actively discover "tree hole" crowds with irregular work and rest, and accordingly increase the alert sensitivity to such crowds.
4. The Tree Hole Rescue Team should stay vigilant persistently in the event of the COVID-19 epidemic or other special event. The intensity of suicide warning and suicide intervention should be adjusted according to the changing law of the sentiments of "tree hole" users. For example, during the COVID-19 epidemic, the Tree Hole Rescue Team should improve the intensity and sensitivity of suicide warning, especially on Fridays at 4 am, 10am and 6 pm. If necessary, we should take the initiative to post positive messages in the "tree hole" to stop the spread of negative emotions. This can be used to intervene in advance of suicide. By intervening in suicidal tendencies in advance, our research results can be helpful to the Tree Hole Rescue Team to better carry out rescue work.

6 Conclusion and Outlook

In this paper, we analyze the sentiment and quantity of messages in the "tree hole" of Microblog in time dimension, and we can conclude that the overall sentiment of the messages in the "tree hole" fluctuates regularly with time. And the fluctuation pattern will change when certain big events occur. In addition, few messages and low sentiment of the "double bottom" time period, should be appropriate to increase the sensitivity of suicide warning. If necessary, sending positive messages to the "tree hole" through the rescue team during the "double bottom" period turns passive rescue into active early intervention. This can assist the rescue team to carry out better work.

This paper has completed the analysis of the temporal characteristics of the sentiment of the Microblog "tree hole" messages in terms of hours in a day and days in a week, and compared the changes of sentiment fluctuations before and during the COVID-19 epidemic. However, we haven't analyzed the temporal characteristics of the sentiment of "tree hole" messages in terms of months and years. Since the data is only available until August 2020, we haven't analyzed the temporal characteristics of the sentiment of the "tree hole" messages in the context of regular epidemic prevention and control. Therefore, we will continue to complete the dataset and expand to more social media. And we plan to use more sentiment analysis methods to analyze the sentiment of the "tree hole" messages in a month, a year and other temporal dimensions.

References

1. Losekam, S., Konrad, C.: Depression. In: Kircher, T. (ed.) Kompendium der Psychotherapie, pp. 101–143. Springer, Heidelberg (2019). https://doi.org/10.1007/978-3-662-57287-0_7

2. Huang, Y., et al.: Prevalence of mental disorders in China: a cross-sectional epidemiological study. Lancet Psychiatry **6**(3), 211–224 (2019)
3. Zhu, W.: The formation, harm and response of depression. J. Polit. Sci. Eng. **2021**(03), 70–71 (2021)
4. WHO TEAM: Suicide in the world. World Health Organization, pp. 1–32 (2019)
5. Rozanov, V.A., Rakhimkulova, A.S.: Suicidal ideation in adolescents—a transcultural analysis. In: Kumar, U. (ed.) Handbook of Suicidal Behaviour. Springer, Singapore (2017). https://doi.org/10.1007/978-981-10-4816-6_15
6. CNNIC: The 47th China Statistical Report on Internet Development. The 47th China Statistical Report on Internet Development, pp. 1–118 (2021)
7. Li, C.: Research on group loneliness in online social media environment. Huazhong Normal University (2019)
8. Yang, F., Huang, Z., Yang, B., Ruan, J., Nie, W., Fang, S.: Analysis of suicidal ideation of Microblog "tree hole" users based on artificial intelligence technology. J. Nurs. **34**(24), 42–45 (2019)
9. Lv, S.: On the anonymity effect of tree hole communication. Literary Educ. **2021**(12), 33–34 (2021)
10. Feng, X.-H., Huang, Z.: Using AI to brighten the lives of depressed patients. China Sci. Technol. Ind. **2019**(10), 65–66 (2019)
11. Gong, J., Lin, S., Huang, Z.: Study on the spatial characteristics of depression patients' data on Microblog "tree holes." China Digit. Med. **15**(04), 70–74 (2020)
12. Jing, X., Lin, S., Huang, Z.: Research on the behavior pattern of microblog "tree hole" users with their temporal characteristics. In: Huang, Z., Siuly, S., Wang, H., Zhou, R., Zhang, Y. (eds.) Health Information Science. HIS 2020. LNCS, vol. 12435, pp. 25-34. Springer, Cham (2020). https://doi.org/10.1007/978-3-030-61951-0_3
13. Zhang, J.: A combination of Lexicon-based and classified-based methods for sentiment classification based on Bert. J. Phys.: Conf. Ser. 1802(3), 032113 (2021)
14. Mikolov, T., Chen, K., Corrado, G.S., et al.: Efficient estimation of word representations in vector space. In: International Conference on Learning Representations (2013)
15. Liu, W., Li, Y., Luo, J., Li, W., Fu, S.: Chinese short text sentiment analysis based on BERT and BiLSTM. J. Taiyuan Normal Univ. (Nat. Sci. Edn.) **19**(04), 52–58 (2020)
16. Fu, Y., Liao, J., Li, Y., Wang, S., Li, D., Li, X.: Multiple perspective attention based on double BiLSTM for aspect and sentiment pair extraction. Neurocomputing **438**, 302–311 (2021)
17. Yin, F., He, X., Liu, Z.: Research on scrapy-based distributed crawler system for crawling semi-structure information at high speed. In: 2018 IEEE 4th International Conference on Computer and Communications (ICCC), pp. 1–6 (2018)
18. Baidu.Sentiment_Classify[EB/OL] (2019). https://ai.baidu.com/tech/nlp/sentiment/_classify
19. Yan, Y.: Research on text sentiment analysis of microblog hot topics based on deep learning. Nanchang University (2020)