



# Emotion Analysis of Microblog Epidemic Coexistence Based on BERT

Yijie Su<sup>(✉)</sup>

Jinshan College, Fujian Agricultural and Forestry University, Fuzhou, China  
manzouyj24@163.com

**Abstract.** Since the outbreak of the novel coronavirus, a large number of scholars have put forward scientific and effective analysis on epidemic prevention and control through obtaining microblog data. The spatial-temporal analysis of the evolution of public opinion on COVID-19 based on microblog data is conducive to better fight against the epidemic. In this paper, a large number of public comments on Sina Weibo in different periods of time were counted, and BERT technology was used to make emotional analysis, so as to judge whether the public attitude towards the epidemic is coexistence or zero elimination.

**Keywords:** COVID-19 · BERT · Emotion Analysis · Weibo

## 1 Introduction

Since the outbreak of the novel coronavirus, citizens' daily life has changed from offline to online. Internet users obtain relevant information such as epidemic situation, policy release and social news through social media platforms, and express their opinions or emotions on news content. Because of the asymmetric information access of citizens in the early stage of the epidemic, panic and anxiety are more common. The negative network atmosphere leads to the complex and changeable situation of network public opinion and numerous hidden dangers. Early detection of hidden dangers of public opinion and positive response, correct guidance of public opinion through interpretation, disclosure, punishment and supervision can timely control the situation of public opinion and avoid adverse effects in a wider range [1].

A large number of scholars have put forward scientific and effective analysis on epidemic prevention and control through obtaining microblog data. The analysis on the evolution of public opinion on COVID-19 based on microblog data is conducive to better fight against the epidemic. Online work has played a pivotal role in controlling the public opinion situation of the epidemic. Timely disclosure of relevant epidemic prevention policies and trends will enable citizens to learn the latest policies and epidemic situation, so as to have a clear mind and reduce unnecessary panic. We will quickly respond to public opinion, maintain the government's credibility, and make citizens feel reliable and credible. By statistical analysis of Sina Weibo posts and a large number of comments from the public in different periods of time, and by using BERT technology, public opinion analysis on COVID-19-related topics can more objectively show the evolution

process of online public opinion in this epidemic event. In the current context, it can also judge the public's attitude towards the virus, so as to provide reference for future epidemic prevention and control and policy making, and avoid social events.

In order to accurately monitor the development process of public opinion and control the development direction of public opinion, it is increasingly important to rely on a variety of monitoring indicators for comprehensive research and quantitative analysis. Traditional analysis methods mainly rely on various monitoring indexes, such as forwarding volume, search volume, comment volume, etc. these indexes describe the popularity and change trend of events from the side. These indicators reflect changes in the popularity of public opinion, while more detailed public opinion themes and changes and users' emotional views and changes are difficult to reflect through these indicators, requiring manual observation and extremely labor-intensive. With the development of deep neural network and natural language processing technology, more text processing methods are widely used in public opinion monitoring and governance, helping to extract more direct and subtle indicators. It can be found from the researches on relevant experiments conducted by domestic and foreign scholars that the use of big data monitoring system to analyze and study public concern has become the current research trend at home and abroad.

Therefore, this paper chooses the more practical BERT model. With the continuous development of deep learning, devlin and others first proposed the Bert model (bidirectional encoder representations from transformers, Bert) in 2018, which has become a very representative pre training model in the field of natural language processing and is widely used in various downstream tasks [2]. Bidirectional Encoder Representation from Transformers (BERT) is a pre-trained language Representation model. It emphasizes the use of new Masked Language Model (MLM) instead of the traditional one-way language model or the method of shallow splicing of two one-way language models for pre-training, so that in-depth bidirectional language representation can be generated [3].

## 2 Research Basis

In this kind of research, scholars crawled the news and comments related to COVID-19 from Weibo, and obtained the news corpus and comment corpus through data preprocessing [4]. Various mathematical models were used to extract epidemic events and their components around various topics such as prevention and control, clinical practice and closure. Then, on the basis of introducing the emotion dictionary and the modifier dictionary, we calculate the emotion intensity. Then, models such as BiLSTM were used to obtain the emotional distribution of netizens, and then the impact of the epidemic on netizens' emotions was obtained. They try to scientifically and accurately grasp the public opinion of the epidemic situation for governments at all levels and provide strong theoretical support for effective prevention and control publicity and public opinion guidance [5].

The marginal contribution of this paper is to analyze domestic people's attitudes towards COVID-19 prevention and control by combining the latest trends of domestic epidemic prevention and control and the popular BERT model used for text sentiment analysis, so as to provide suggestions and references for the formulation of national epidemic prevention and control policies.

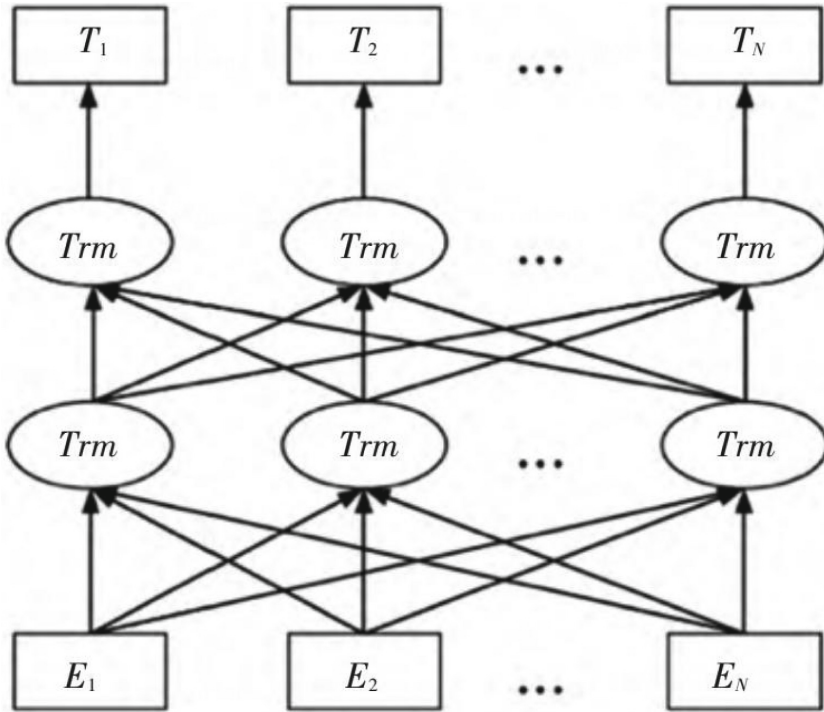
At present, the analysis of public attention at home and abroad is mainly divided into two types: supply-based and demand-based. Supply-based data mainly refers to the data that netizens actively post on Weibo, Facebook, blogs and other social media, while demand-based data refers to the data of people's search behavior in the network [6].

Public opinion analysis on COVID-19, a major public health emergency, is conducive to the formulation of epidemic prevention and control measures in key regions [7]. Public opinion analysis method based on this build the framework, the data acquisition, index evaluation, information extraction, time and space analysis and visual expression are decomposed, the introduction of natural language processing technology and the community network discovery algorithm parse complex text corpus and different area of the mining index information, can be a major public event of public opinion analysis and evaluation to provide reference and reference [8]. In the process of writing, this paper refers to the experience of previous scholars to a large extent, and improves some limitations existing in it.

However, due to the limitations of technology and objective environment, this paper also has some limitations as follows: In terms of the selection of text corpus, this kind of experiment only takes Weibo comments as the object [9]. Most of the comments are young people who are active in social media, and network users of all ages are not covered. In addition, BERT's sentiment analysis can accurately identify the evaluations with clear public orientation, but it is difficult to obtain correct scores for relatively vague or neutral public emotions. In this kind of experimental research, Chinese and foreign scholars not only conducted data analysis of epidemic situation from different dimensions (space and time, etc.) through various research statistical analysis methods, but also put forward specific suggestions on the governance of online rumor spreading from administrative and legal perspectives [10]. Because of different types of emergencies content and rumor propagation characteristics of each are not identical, especially a public health emergency professional stronger, the rumor spread more "concealment" and "professional", should be aimed at different types of emergencies rumors difference analysis and discussion with spam, can put forward in accordance with event development regularity of outbreaks.

### 3 Research Content

The research process of this paper is mainly divided into three parts, including data acquisition and processing, empirical analysis and research conclusion analysis. The article started with comments on popular Weibo posts about the epidemic on People's Daily, CCTV News and other official Weibo accounts with more than 100,000 followers. For the comments on Weibo, a simple descriptive analysis was conducted through the high-frequency words and word cloud map to understand netizens' concerns and discussion hot spots about the epidemic. Secondly, Chinese emotion word ontology was used to conduct corpus emotion analysis through BERT model, and comment emotion words were extracted to classify netizens' emotions and obtain netizens' emotional tendency towards the epidemic. The chart shows the changes of netizens' emotional values at key time points, and based on the above analysis results, it explores people's emotional changes on epidemic control policies under the situation of normalized epidemic [11].



**Fig. 1.** The BERT model pretrains the bidirectional Transformers through MLM (by owner-draw)

The BERT model pretrains the bidirectional Transformers through MLM to generate bidirectional language representation, as shown in the figure. Through word vector, text vector and position vector, the sum of the three parts is taken as the input layer, then the representation vector with semantic information can be obtained (Fig. 1).

On the basis of BERT model, we made statistics of emotional segmentation in the comments section of Weibo in different periods and drew a chart. Through the analysis of the trend, it is easy to see that at the beginning of the epidemic, the public’s attitude towards epidemic prevention and control was relatively uniform, and the vast majority of people supported the zero-out policy to protect their lives and health.

But over time, people’s views diverged. After entering 2021, the virus has evolved into the Delta model. While the transmission capacity is strengthened, the mortality rate has decreased to a certain extent. In addition, during the Spring Festival, the strong containment measures have led to people’s dissatisfaction, so the proportion of people supporting the elimination of the epidemic has increased.

The biggest change in attitude occurred at the beginning of 2022, when the virality of omicron was greatly increased and the death rate was greatly reduced. At the same time, every country except China gradually took policy measures to lift the containment. But many Chinese cities have already continued the previous measures, which have hit people’s basic life security.

## 4 Conclusion

From the above research and analysis, it is not difficult to see the evolution of public opinion. Therefore, in the process of policy implementation, the government should further strengthen the flexibility of epidemic prevention and control measures, and invest more resources in drug research and development and vaccine promotion to avoid long-term extension of containment measures. At the same time, local governments should also implement the core policy of dynamic zero clearance. It should be noted that “dynamic clearance” does not mean “zero infection”. It means that early detection, diagnosis, isolation and treatment should be made as early as possible in the normal prevention and control phase to resolutely prevent the continued spread of the epidemic in communities. Remain on high alert when there are no indigenous cases; When local cases emerge, they will be promptly detected, promptly handled, controlled and effectively treated. “Dynamic clearance” can not only effectively control the outbreak and spread of the epidemic, minimize infection, morbidity and death, and avoid a drain on medical resources, but also minimize the impact of the epidemic and reduce its impact on the economy and society. It can be said that “dynamic zero” strikes a good balance between epidemic prevention and control and economic and social development, enabling us to achieve maximum results at minimum cost and effectively protect people’s lives and health as well as economic and social development. In this regard, we must adhere to the “dynamic zero elimination” unswervingly, and make greater efforts to “scientific precision”, to eradicate the epidemic as soon as it is discovered. It should also be noted that adhering to the “dynamic zero clearance” will inevitably have a certain impact on the daily life of the masses in the short term, and all aspects should take corresponding safeguard measures.

## References

1. Wang Xiyi, Du Mingkun, & Zhang Shan (2020). Research on the popularity of online public opinion based on deep learning Wireless Internet technology, 17 (22), 2
2. Jacob Devlin, Ming-Wei Chang, Kenton Lee, Kristina Toutanova (2018). Bert: Pre-training of deep bidirectional transformers for language understanding. <https://doi.org/10.48550/arXiv.1810.04805>
3. Benitez-Andrades, J.A., Alija-Perez, J.M., Garcia-Rodriguez, I., et al. (2021) BERT Model-Based Approach For Detecting Categories of Tweets in the Field of Eating Disorders (ED). 2021 IEEE 34th International Symposium on Computer-Based Medical Systems (CBMS). IEEE.
4. Xiang, N., Jia, Q.Q., & Wang, Y.D. (2021) Sentiment analysis of Chinese Weibo combining BERT model and Hawkes process.
5. Alami, F., Alaoui, S., & Ennahahi, N. (2021) Contextual semantic embeddings based on fine-tuned arabert model for arabic text multi-class categorization. Journal of King Saud University-Computer and Information Sciences(2).
6. Shi, W.Z., Zeng, F., Zhang, A., et al. (2022) Online public opinion during the first epidemic wave of covid-19 in china based on weibo data. Palgrave Communications, 9.
7. Yang, W., Wu, Z., Mok, N.Y., et al. (2022) How to save lives with microblogs? lessons from the usage of weibo for requests for medical assistance during covid-19.

8. Pei, J., Lu, Z., & Yang, X. (2021) What drives people to repost social media messages during the covid-19 pandemic? evidence from the weibo news microblog. *Growth and Change*.
9. Fan, Y. (2020) The Social Amplification of Risk on Weibo-Take the COVID-19 Epidemic in China as an Example. *CAIH2020: 2020 Conference on Artificial Intelligence and Healthcare*.
10. Nandy, R. (2021) Facebook and the covid-19 crisis: building solidarity through community feeling. *Human Arenas*.
11. Kuchler, T., Russel, D., & Stroebel, J. (2021) Jue insight: the geographic spread of covid-19 correlates with the structure of social networks as measured by facebook. *Journal of Urban Economics*.

**Open Access** This chapter is licensed under the terms of the Creative Commons Attribution-NonCommercial 4.0 International License (<http://creativecommons.org/licenses/by-nc/4.0/>), which permits any noncommercial use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.

