Research on the Support Degree of Netizen Comments on the Proposals of the National Committee Members of the National People's Congress

Weibo Dong

School of Software, North University of China, Taiyuan, 038507, China

Abstract

With the popularization of the Internet, people try to express their opinions on online social platforms, the government and relevant departments are gradually paying attention to the intentions of the people on public social platforms. This paper collects the popular proposals during the two sessions in 2021 for a systematic classification, and uses the latest BERT model in NLP to perform sentiment analysis on Weibo user comments to understand the people's emotional attitudes towards different types of committee proposals. Research has shown that most netizens are more concerned about Science, Education, Culture, Sports and Social Security, the proposals on Science, Education, Culture, Sports and Labor and Personnel are more controversial. This work not only helps the government understand the social issues that most netizens care about and their demands for the future, but also helps to observe the quality of the proposals made by the members of the two sessions.

Keywords

Sentiment analysis; Sina Weibo; Committee proposal.

1. Introduction

In recent years, with the active social media on the Internet, more and more users have begun to express their views on certain things on various social media. With the maturity of 5G technology and the vigorous development of the short video industry, China Mobile's social user scale continues to grow. In 2020, its user scale will exceed 900 million, an increase of 7.1% from 2019. [1] As one of the social media with the largest user base in China, Sina Weibo has 523 million monthly active users in 2020. Not only that, the proportion of users born in the 80s and 90s is as high as 78%, and the overall trend is younger, [2] So the platform can better reflect the voice of contemporary young people.

With the increase in the number of users, the government and state agencies are paying more and more attention to the voices of the increasingly large number of users on Sina Weibo. In 2020, the number of government affairs agency accounts authenticated by Weibo will exceed 140,000, and the total number of fans will exceed 3 billion. The total number of Weibo posts published has exceeded 450 billion times. Authoritative epidemic information, social hotspot responses, and positive-energy heart-warming stories are what users most want to see when paying attention to government affairs microblogs. After the outbreak of the epidemic, topics such as health care and the construction of central enterprises have received widespread attention from netizens. In addition, as an authoritative voice and responsible for positive energy, the hot topics and positive energy topics hosted by government microblogs also have a high degree of attention. [3] During the National Two Sessions in 2021, various committee proposals frequently appeared on hot searches, and a large number of users expressed their opinions on different proposals, which became a current political hotspot at the time. Therefore,

we studied the opinions of users on social media during the National Two Sessions. There is a certain necessity for the emotional attitude of the CPPCC members' proposals.

The analysis of committee proposals generally belongs to the task of public opinion analysis in the field of NLP. Many scholars at home and abroad have done various degrees of research on public opinion analysis, but the research on the proposals of the two committees of the country has not yet been involved. The analysis of online public opinion on the members' proposals during the two sessions not only helps the country understand the demands of the people, but also helps the committee members to put forward more constructive proposals in a more targeted manner.

2. Research Method

2.1. Data Acquisition

As one of the largest social media in China, Sina Weibo is especially loved by young people. Therefore, this article uses the Weibo platform and Python crawlers to obtain data. The National Two Sessions in 2021 will be held from March 4 to March 11, but various committee proposals have already been searched on the eve of the two sessions, so the selection date for this article is from March 2, 2021 to March 11, 2021. Data for the day. In the selection of the data range, this article uses the following three ways to obtain user comments: 1. Proposals that have been listed on Weibo hot searches during the selection period. 2. Choose the most popular proposal in the official Weibo of People's Daily. 3. Choose the most popular proposal on the official Weibo of China News Network. A total of 39 proposal microblogs and 28574 comments were selected.

2.2. Data Cleaning

After getting the comments, many comments include hyperlinks, pictures, @ other users, etc. that are of no reference value for the research itself. This article uses regular expressions to clean the data. The main tasks are: 1) Remove all HTML tags, Hashtags, hyperlinks, pictures, emoticons, and @users. 2) Use the "Beautiful Soup" Python library to remove punctuation. In the end, a total of 27,349 comments with reference significance were obtained.

3. Data Classification and Statistics

In order to more intuitively show the social issues concerned by Sina Weibo user groups and to further understand the demands of users of the platform, this article systematically classifies and counts 39 highly popular proposals, which are divided into 5 categories in total, as follows as shown in the figure.

Table 1. Classification of proposals

Proposal Type	Total number of proposals (items)	Total number of comments (items)
Science, Education, Culture and Sports	12	9073
Medicine and Health	6	3166
Social Security	9	4307
Democracy and Legal System	6	2855
Labor and Personnel	6	6948

4. Data Processing

This article uses the latest BERT model to fine-tune the sentiment analysis of comments. The main tasks are: 1) Download the Pytorch version of the BERT model from the official website and the Chinese pre-training model provided by Google. 2) Fine-tune the parameters of the BERT model, the specific parameters are as follows:

Table 2. BERT model parameters

Hidden_size	Num_epochs	Batch_size	Pad_size	Learning_rate
12	3	3	3	5e-5

3) Perform sentiment analysis on 27349 comment data, and divide the results into three categories: support/against, and neutral. 4) Count and calculate the emotional attitude of different types of proposals.

4.1. Result

Emotional attitude statistics are carried out for different types of proposals. The statistical results are as follows:

Table 3. Proposal support results

Emotional Attitude / Proposal Type	Support	Opposition	Neutrality
Science, Education, Culture and Sports	0.5236	0.2151	0.2613
Medicine and Health	0.7024	0.1117	0.1859
Social Security	0.8206	0.0656	0.1138
Democracy and Legal System	0.8145	0.0336	0.1519
Labor and Personnel	0.4721	0.3361	0.1918

4.2. Conclusion and Outlook

4.2.1. Research Conclusion

Through sentiment analysis and statistics, this article draws the following conclusions:

(1) In terms of proposal classification and statistics, among the 39 proposals with high popularity, 21 are related to science, education, culture, sports and social security. This shows that most netizens are more concerned about these two types of proposals. There are more social demands.

(2) In terms of support sentiment analysis, the debates on the proposals for science, education, culture and sports and labor and personnel are fierce, indicating that netizens have different responses to the two types of proposals; while the social security and democracy and legal system proposals have more consistent responses. It proves that the two types of proposals are more supported by netizens.

4.2.2. Research Deficiencies and Prospects

This article has conducted a reasonable analysis of the emotional attitudes of the members of the National Committee of the People's Republic of China and the National Committee of the People's Republic of China on social media. However, there are still some shortcomings, which can be further in-depth in the following aspects:

(1) The selected data are all from the comments on different proposals made by netizens on Sina Weibo. Although Sina Weibo is one of the largest social media in China, its users are younger, born in the 1980s and older. There are fewer users and a narrower age distribution. You can also choose more social media as a means of data acquisition, making the results more representative.

(2) In terms of sentiment analysis, this article adopts the latest BERT model and fine-tuning to analyze sentiment tendency, but the pre-training uses the Chinese pre-training model provided by Google. If the existing data can be used for pre-training, the accuracy of sentiment analysis may be higher.

References

[1] iiMedia (2021). 2020-2021 China Mobile Social Industry Research Report. https://www.iimedia.cn/c400/76205.html

[2] Sina Weibo Data Center (2021).Weibo 2020 User Development Report. https://data.weibo.com/report/reportDetail?id=456

[3] Cao Zenghui, Senior Vice President of Weibo (2020). 2020V Impact Summit. https://auto.sina.com.cn/yx/yd/yxdt/2020-10-22/detail-iiznctkc6664724.shtml