Neural Computing and Applications (2022) 34:12505–12516
https://doi.org/10.1007/s00521-021-06567-6

Research on the effect of government media and users' emotional experience based on LSTM deep neural network

Nan Wang[1] · Xinlong Lv[1] · Shanwu Sun[1] · Qingjun Wang[2,3]

Received: 28 June 2021 / Accepted: 20 September 2021 / Published online: 6 October 2021

Abstract

Different government media have different communication effects and users' emotional experience. It carries on a comparative research on government media selecting three different types of government media which include China's Police Online, Central Committee of the Communist Youth League, and China's Fire Control in the context of public health emergencies. Based on the deep learning technique, the emotion classification model of long-term memory network is constructed to analyze the emotion of the users' comments of different government media; taking the number of contents, the number of retweets, the number of praises, and the number of comments as evaluating indicators to do comparative analysis to cross platform government medias. Through the comparative results, it is found that different types and platforms of government media have great differences in users' emotional experience; the emotion performance of users' comments is strongly related to the information communication power and effectiveness of government media.

Keywords Government media · Emotion analysis · LSTM · Deep learning · Social media platforms

1 Introduction

With the rapid development of Web2.0 technology, various kinds of social media emerge in people's lives, people are keen to express their views on various social platforms. According to the 44th CNNIC China Internet Report released by China Internet Information Center, as of June 2019, the number of Internet users in China was 854 million, an increase of 25.98 million compared with 2018, and the Internet penetration rate reached 61.2%, 1.6 percentage points higher than the end of 2018 [1].

The activity of Internet users has also led to the explosive growth of government media. Since 2011, the government media in China has developed rapidly, especially the social media represented by micro-blog. Various government departments have started to register official media on micro-blog, so 2011 is also known as "the first year of government microblog." As of June 2019, China's online government service users reached 509 million, accounting for 59.6% of the total Internet users. There are 139,000 government media certified by Sina Weibo platform. Due to the timeliness of micro-blog, local governments at all levels release event notifications and some practical reports through micro-blog, and government micro-blog has become a new interactive platform between government and people. In addition to micro-blog, in recent years, with the blowout of short video platform, more netizens began to enter the short video platform. According to the 44th CNNIC China Internet report, the number of short video

✉ Qingjun Wang
wangqingjun@sau.edu.cn

Nan Wang
ctuwangnan@126.com

Xinlong Lv
xslvxinlong@163.com

Shanwu Sun
ctusunshanwu@126.com

[1] Jilin University of Finance and Economics, Changchun 130117, China

[2] Shenyang Aerospace University, Shenyang 110136, China

[3] Nanjing University of Aeronautics and Astronautics, Nanjing 210016, China

users reached 648 million, accounting for 75.8% of the total netizens. Various government media have also registered official media accounts on various short video platforms. In June 2019, the Government Short Video Development Report pointed out that as of September 14, 2018, 170 network police units registered on DouYin short video platform and established a joint working matrix [2].

After years of rapid growth, the government media in China has formed the mainstream situation of WeChat (a social software), micro-blog and DouYin (a short video platform). In the research of government media, most domestic and foreign scholars focus on quantitative research, the research topic is mainly about the information communication, influencing factors and evaluation system of social media [3], the research content is also focused on the same social media platform, lack of cross platform and cross type government media research.

Based on deep learning emotion analysis technology and taking micro-blog and DouYin as the research platforms, the comparative study is conducted from the perspective of "the same platform, different government media" and "cross platform, the same government media." The overall research framework is shown on Fig. 1.

2 Literature review

2.1 Government media

At present, there are two kinds of research subjects in the field of government media in China. The first one is mostly from the perspective of communication, which studies the government media ontology, including the influencing factors of government media, the construction of influence index system, evaluation system, etc. Yang Changchun et al. [4] put forward a method based on H index to construct the influence index of government micro-blog, which includes four first level indexes (creativity, communication, service and interaction) and 10 s level indexes, and used AHP to give weight to each index. Through the Web crawler to obtain the data of 17 micro-blog government platforms for empirical research, the research shows that the evaluation system has a good application effect. Based on the indicator systems proposed by Yang Changchun, Zhao Amin et al. [5] introduced factor analysis and cluster analysis, and divided the influence of government media into five different types, namely strength leading type, comprehensive medium type, comprehensive backward type, comprehensive leading type and balanced development type. Rong Yihong et al. [6] improved the evaluation index system and established a more comprehensive system including 6 first level indexes, 20 s level indexes and 65 third level indexes.

The second category is mainly the research on the communication effect of government media in network public opinion governance, which is mainly based on the empirical research under some emergencies [7–9]. There is very little research on government media under different platforms. Based on emergencies, taking the government micro-blog and DouYin, the three aspects of quantities, content analysis and "burst point" are compared and analyzed in [10]. This paper focused on the similarities and differences between government DouYin and micro-blog and made a preliminary comparative study of them. However, only one government media of China is selected in this paper, which lacks diversity of samples.

2.2 Emotion analysis

Emotion analysis, also known as opinion mining [11], refers to the analysis of a natural language text to determine whether its emotional tendency belongs to positive or negative. At present, Emotion analysis is divided into two categories: based on emotion lexicons and based on

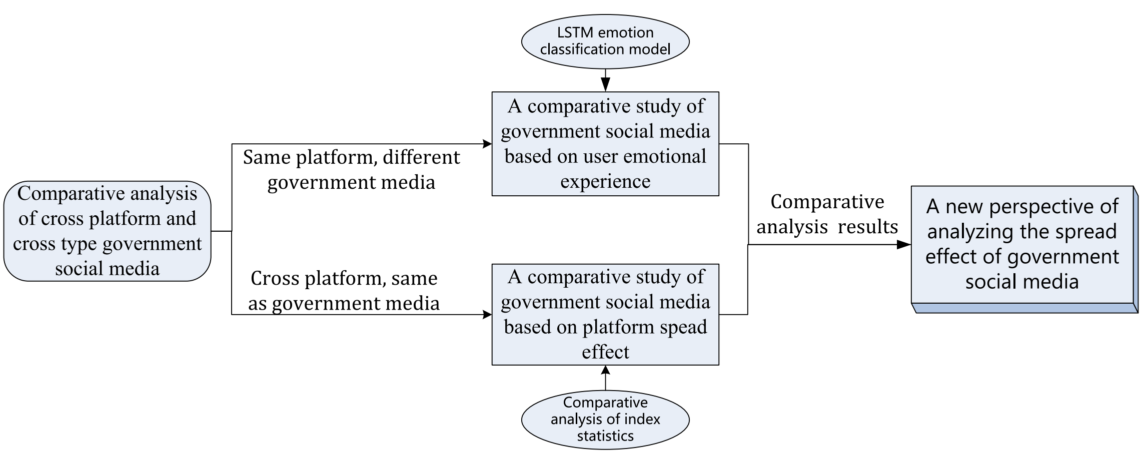


Fig. 1 Research framework

machine learning. Emotion analysis based on emotion lexicons uses the existing knowledge to construct specific emotion lexicons in this field by marking emotion words and calculating the weight of them. After the establishment of emotion lexicons, the text to be predicted is put into the emotion lexicons for matching, so as to get the emotion value [12, 13]. However, the traditional emotion lexicons need a lot of manpower and time costs when it is constructed, as well as the linguistic knowledge of researchers themselves, especially when it involves some professional fields, and it also needs experts in relevant fields to give professional phrase suggestions. With the development of technology, the method of emotion analysis based on machine learning is more and more popular.

The emotion analysis based on machine learning is to learn the characteristics of text data through a small number of tagging words and using the constructed model, so as to achieve the purpose of emotion analysis of massive unknown text, which can save a lot of manpower and time costs, and improve the classification effect significantly. Pang et al. [14] first applied machine learning to text emotion classification in 2002. They applied naive Bayes, maximum entropy and support vector machine algorithm to text emotion analysis and made a comparative study. Experimental results showed that support vector machine algorithm had the best effect in text emotion classification. On this basis, many scholars have made a lot of improvements on these three algorithms [15, 16]. With the increasing popularity of emotion analysis based on machine learning and the emergence of various deep learning models, Bengio et al. [17] constructed language models based on neural networks for the first time. Due to the long training time of the model, in 2013, Mikolov et al. [18] of Google company modified the model to build word2vec model, which has become the main vector representation model in natural language field. The emergence of word2vec model makes the deep neural network model perform better in the classification effect, such as Recurrent Neural Network (RNN), Long Short-Term Memory (LSTM) [19, 20].

Through the literature review in the field of government media, we can see that at present, government media is mainly from the perspective of communication science, most of the research methods are quantitative research and qualitative description, lack of in-depth analysis of government media by using machine learning technology, and most of the research is based on the same platform, less research on cross platform. In the field of emotion analysis, in terms of technology, it is mainly aimed at the innovation and improvement of algorithm, while in the application of emotion analysis and government media, some scholars also studied the evolution of public opinion events based on emotion changes [21]. Generally speaking, there is no research on cross-platform and cross-type government media, and there is no comparative study on government media based on emotion analysis. Therefore, based on machine learning technology, there is a lot of research space for cross-platform government media communication effect and operation management.

3 Construction of emotion classification model based on LSTM

3.1 Research object and data selection

3.1.1 Description of event object

This paper is about different types of government media under the two platforms, micro-blog and DouYin. Based on the classification and ranking given in the 2019 Government Micro-blog Index Report [22], we select three different types of government media: Central Committee of the Communist Youth League (Ranked second in micro-blogs of China government-type institutions), China's Police Online(Ranked first in the top ten public security micro-blogs in China), China's Fire Control(Ranked first in the top ten emergency management micro-blogs in China).

In order to make the research more concrete and comparative, this paper selects the relevant data of the three major government media about the same event in the same period. The COVID-19 event was selected and described as follows:

In late December 2019, patients with pulmonary infection of unknown etiology were first found in Wuhan. The novel coronavirus was discovered by experts after two weeks of research. With the spread of the disease, different levels of infected people began to appear all over the country. On January 23, 2019, Wuhan announced locking down. Then the first level response of health and safety was launched in succession all over the country, and the major government media continued to follow-up and report. The people of the whole country worked together to launch an anti-epidemic war.

3.1.2 Data acquisition and preprocessing

Data are selected from 12 days of relevant data between January 29, 2019, and February 9, 2020.

Training data: by searching the keyword "Novel Coronavirus" or "2019-nCoV" in the micro-blog platform, we can get the comment data under different micro-blogs by using the technology of web crawler. In order to ensure the quality of training data to the maximum extent and train a high-quality classification model, we obtained 9671 pieces of effective data through data filtering, removing invalid characters and other data cleaning operations, and divided the data into two categories: 4132 pieces of comments with positive emotion (positive samples) and 5539 pieces of comments with negative emotion(negative samples).

The classification principle of data labels is as follows: those with words such as blessing and salute and the meaning expressed by the statement of positive energy are regarded as positive samples; those with words such as rumor and apology and the meaning expressed by the statement biased toward dissatisfaction and accusation are regarded as negative samples.

Test data: using web crawler technology to obtain the comment data of Novel Coronavirus reports from January 29, 2019, to February 9, 2020, in three government media, namely Central Committee of the Communist Youth League, China's Police Online and China's Fire Control, a total of 57,491 pieces of data. The details of data distribution are shown in Table 1.

3.2 Model constructing

3.2.1 LSTM

LSTM, also known as long short-term memory cyclic network, is an improved model developed on the basis of classical cyclic neural network. The characteristic of LSTM network is to use memory module instead of common hidden nodes, so as to ensure that the gradient does not disappear or explode after passing over many time steps, and overcome some difficulties encountered in traditional cyclic neural network training. Long- and short-term memory networks are very suitable for learning classification from experience, as well as for processing and predicting time series with unknown time delay between important events.

Figure 2 is a network structure diagram of a neural unit. Based on the traditional cyclic neural network, LSTM adds a memory block, which consists of three gates: forget gate, input gate and output gate.

Forget gate: filter some unnecessary or useless sentences through the knowledge learned in the last moment.

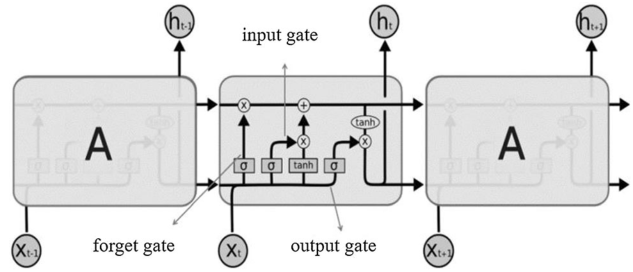


Fig. 2 Internal structure of LSTM neurons

Input gate: the sentences retained in the forget gate are updated through the input gate.

Output gate: control unit status, which determines how many values are output to the current output value of LSTM.

3.2.2 Emotion classification model based on LSTM

1) Vectorization of text

   In the field of natural language processing, the first problem is how to transform text information into a form that can be recognized by computer, that is, to model natural language. Natural language modeling method has experienced a transformation from rule-based method to statistical method. In the context of the research on statistical language models, Google released a software tool for training word vectors, Word2vec, in 2013. Word2vec can express a word into a vector form quickly and effectively through the optimized training model according to the given corpus, that is, to follow certain coding rules to convert the word into a computer-readable vector form. Su Jianlin [23] compared three coding modes, one-hot, one-embedding and word-embedding, and found that the accuracy of classification was almost the same. In this paper, we choose word-embedding as coding mode, namely word vector mode, which is a matrix-based distributed representation. In this representation, a row in the matrix becomes the representation of corresponding words. Firstly, a sentence is divided into the combinations of word

Table 1 Number of comments by date

Date Government media	1.29	1.30	1.31	2.1	2.2	2.3	2.4	2.5	2.6	2.7	2.8	2.9
Central Committee of the Communist Youth League	2921	2696	2810	3235	2322	2537	4266	1835	4131	1823	2736	1426
China's Police Online	1943	1587	901	1057	1399	957	1000	1336	1251	1540	965	1395
China's Fire Control	1104	668	531	1003	842	1893	479	480	544	458	884	536

and word by using the word segmentation tool of Jieba. For example, 希望你们平安归来 (Hope you can come back safely). After word segmentation, it becomes 希望/你们/平安/归来(Hope/you/can/come/back/safely). Suppose that after word-embedding coding, it becomes a matrix in the form of [12, 13, 65, 47], then a group of numbers represents a word, and a matrix represents a sentence.

2) Input of the model

The sentences after word segmentation are input into the model after vector transformation. However, because the Keras framework used in this paper requires that the input data of the model have the same data length, we set the maximum length of the data to 50 according to the length of the comment information. The excess part is processed by truncation, and the insufficient part is supplemented by 0.

For example, 希望你们平安归来 (Hope you can come back safely). The representation matrix of this comment is [12, 13, 65, 47], after the insufficient part is supplemented, it is expressed as [12,13,65,47,0,0,0,0,0,0,0,0,0,0,0,0,...].

3) Parameter optimization

In terms of model parameter setting, through multi round test, the parameter Batchsize is set to 48, and the accuracy of the model basically does not change after 25 rounds of training, so the parameter Epoch is set to 25. The Sigmoid function is selected as the activation function. The Sigmoid function is one of the most commonly used activation functions. Its output is a value in the interval (0, 1), which is very suitable for the task of binary classification. The Sigmoid function is expressed as:

$$f(z) = \frac{1}{1 + \exp(-z)} \tag{1}$$

Adam algorithm was first proposed in 2014 [24], which considers both one-stage momentum and two-stage momentum updating. Because of its outstanding performance, it is widely used in deep neural network and is the most recognized adaptive learning optimizer at present. Therefore, Adam algorithm is chosen as the optimizer to optimize the objective function and update the parameters in the model.

3.3 Model training and evaluation

We use 20% of the training data as the verification set of the model, i.e., 7403 pieces of data as the training set and 1850 pieces as the verification set. The training results show that the accuracy of the model in the training data set is over 90%, and the accuracy in the verification set is stable at about 98%. This model can be used as the emotion classification model in this study. The accuracy and loss rate of model training history are shown in Fig. 3.

Through observation of a large number of prediction results, it is found that the default threshold value of the system, that is, the default output of the probability value greater than 0.5 is 1, and the default output of the probability value less than 0.5 is 0. Such decision conditions are not accurate in this situation. The setting of this threshold value has stronger recognition ability for positive samples and weaker recognition ability for negative samples. The reason is that some sentences in the context of public opinion tend to be ambiguous and difficult to define.

In order to solve this problem, instead of determining emotional tendency by using values 0 and 1, we output specific emotion probability values, divide different emotion fields according to the probability values, observe a large number of prediction results, and divide the differentiation criteria. Table 2 is the emotion domain division standard designed in this paper.

In order to verify the scientific nature of emotion domain division and the reliability of model classification, 200 samples are randomly selected from the predicted samples, among which the positive samples are absolutely positive, i.e., the emotion probability value is (0.6–1], the negative samples are absolutely negative, i.e., the emotion probability value is [0–0.3), the emotion tendency is manually marked, and the confusion matrix of model prediction result is calculated The recall rate and specificity were used to evaluate the model. The definition of the four indexes of the confusion matrix is shown in Table 3, and Fig. 4 is the confusion matrix of the sample.

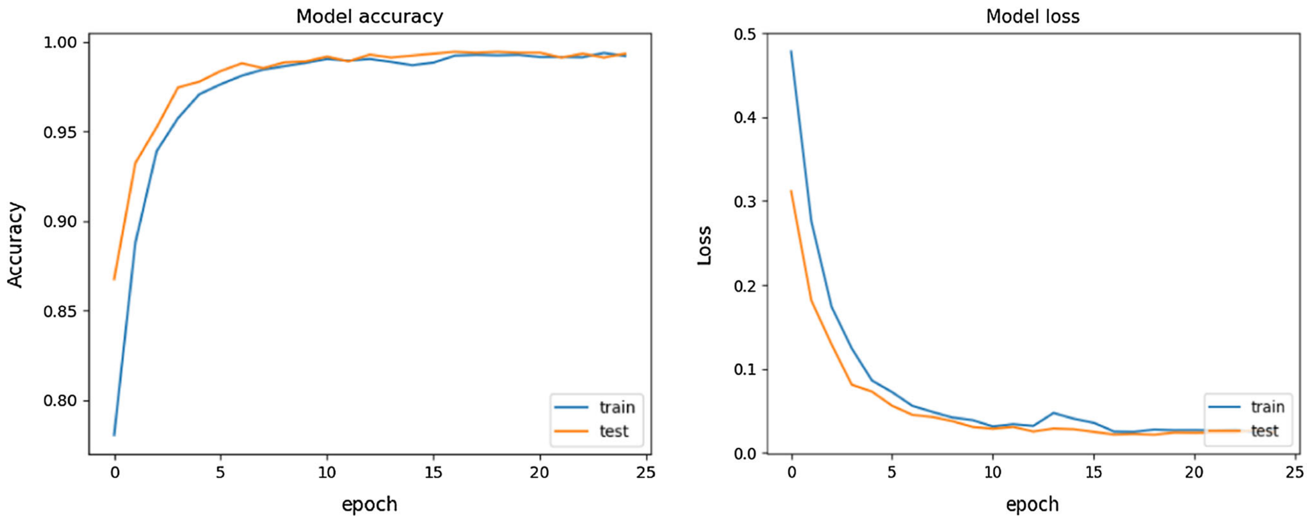
Recall rate: the proportion of samples correctly predicted to be positive to all actually positive samples. The higher the recall rate is, the better the prediction ability of positive samples is. That is to say, as long as it is positive emotion comment, it can be identified.

$$Recall = \frac{TP}{TP + FN} \tag{2}$$

Specificity: the proportion of samples correctly predicted to be negative to all actually negative samples. The higher the specificity, the better the prediction ability of negative samples, that is, as long as it is negative emotion comment, it can be identified.

$$Specificity = \frac{TN}{TN + FP} \tag{3}$$

It can be seen from Fig. 4 that the recall rate of 200 random samples is 96% and the specificity is 87%. The recognition ability of both positive and negative samples is at a high level. Therefore, this model can be used as a model for follow-up research, and the classification criteria is reasonable, that is, sentences with a probability of (0.6–



(a) accuracy rate

(b) loss rate

Fig. 3 Accuracy and loss rate of model training history

Table 2 Criteria for emotion domain division

Emotion domain	Absolutely positive	Weakly positive	Weakly negative	Absolutely negative
Emotion probability value	(0.6–1.0]	(0.5–0.6]	[0.3–0.5]	[0–0.3)

Table 3 Confusion matrix

TN	The prediction is negative sample (N), and the real value is also the number of negative sample (N), that is, the model prediction is correct (True)
FP	The prediction is positive sample (N), and the real value is the number of negative sample (N), that is, the model prediction is wrong (False)
FN	The prediction is negative sample (N), and the real value is the number of positive sample (N), that is, the model prediction is wrong (False)
TP	The prediction is positive sample (N), and the real value is also the number of positive sample (N), that is, the model prediction is correct (True)

1] are classified into absolutely positive emotion, which are regarded as positive samples, and sentences with a probability of [0–0.3) are classified into absolutely negative emotion, which are regarded as negative samples.

4 Experimental results and analysis

We carry out experimental analysis from two perspectives. First, we choose the government media of different government departments under the same social media platform for comparative analysis. In this paper, we choose Central Committee of the Communist Youth League, China's Police Online, China's Fire Control on the micro-blog platform. The comparison method is based on the emotion classification model of LSTM. Then, we choose the government media of the same government department under different social media platforms, the two representative platforms DouYin and micro-blog are chosen here.

4.1 Comparison of the same platform and different government media

This part is based on the users' perspective, that is, from the users' emotional experience to compare and analyze various emotion indicators of different types of

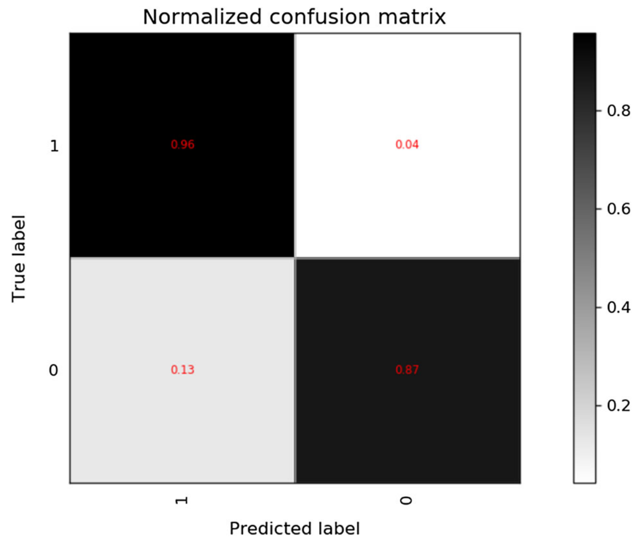


Fig. 4 Confusion matrix

government media. Use emotion classification model of LSTM to analyze the emotion of three types of government media, and compare the two aspects of emotion fluctuation and emotion distribution. The analysis results are shown in Fig. 5.

From the analysis of emotion fluctuation graph, the emotion fluctuation trend of different emotion domains of three government media is synchronous, that is, the same increase and the same decrease, which is related to the number of all comments on a certain day. When the total number of comments increases, the comments of different sentiment domains naturally change. From January 29 to January 31, the emotion of China's Police Online fluctuated obviously, but after January 31, it was more relaxed. From February 2 to February 4, the emotion of China's Fire Control fluctuated greatly, but after February 4, it was more relaxed. Compared with the former two, the number of comments from the Central Committee of the Communist Youth League was higher, and the total number of comments in a day was maintained at about 2000, highest up to 4131 comments a day, and its emotion fluctuation was also relatively frequent, with a total of five major emotion fluctuations.

After observation, we found that the emotion fluctuation of netizens largely depends on the theme of micro-blogging. Take Central Committee of the Communist Youth League as an example: the blog in January 31st "Novel coronavirus can be inhibited by Chinese medicine Shuanghuanglian oral liquid, according to the joint discovery between Shanghai Institute of Medicine and Wuhan Institute of Virology." had up to 18,006 comments; the blog released on February 6 "State Council: local governments shall not retain or transfer medical materials in any name." had 3274 comments; on February 7, the blog post "Farewell! Xu Hui, an anti-epidemic doctor in Nanjing, died" had 2324 comments. Among them, because the "Shuanghuanglian" incident was issued by an authoritative organization, the citizens rushed to buy; in the "intercepted medical materials" incident, because the intercepted materials were used to support Wuhan, the people were angry. The themes of these blog posts can tightly mobilize the emotions of netizens and cause different emotional resonance. Overall, the emotion trend of Central Committee of the Communist Youth League is the most unstable, followed by China's Police Online and China's Fire Control.

Therefore, we can draw a conclusion: the more obvious the theme emotion fluctuation of the blog posts sent by the government media, the more different emotion resonance can be caused, thus enhancing the communication power and influence of the government media.

According to the analysis of the emotion distribution graph, China's Police Online has the highest proportion of negative comments, which reached 60% from February 5 to 7, followed by Central Committee of the Communist Youth League, and China's Fire Control has the most peaceful emotion distribution. By analyzing the reasons, we can find that different types of government media will lead to different themes of micro-blogs. Because China's Police Online is public security government media, most of the blog posts are about some cases of illegal crimes, and the frequency of releasing blog posts is very high, which makes its negative comments a relatively high proportion, while Central Committee of the Communist Youth League and China's Fire Control is more life style, and the positive comments are higher than the negative ones. Therefore, there is a certain correlation between the netizens' emotional experience and the type of government media.

In order to further study the areas where the emotional distribution of government media is concentrated, we substitute the classified absolute positive emotional comments and absolute negative emotional comments into the word cloud model, as shown in Fig. 6. From the negative word cloud, we can see that the distribution of negative comments is scattered, mainly in the areas of shortage of materials(物资), rumors(谣言), people can not buy masks (口罩), Hubei Red Cross corruption(红十字会), and can not go to work(在家). These themes are the midpoint of netizens making complaints about public opinion events, so netizens' emotions are biased toward negative ones. The distribution of positive comments is relatively concentrated, and most of them are words such as encouragement (加油), saluting(致敬), safe(平安) and moving(感动). This is mainly the blessing of netizens to the medical staff struggling in the front line of anti-epidemic. It is the positive emotion of netizens, which also verifies the effectiveness of model classification from the other side.

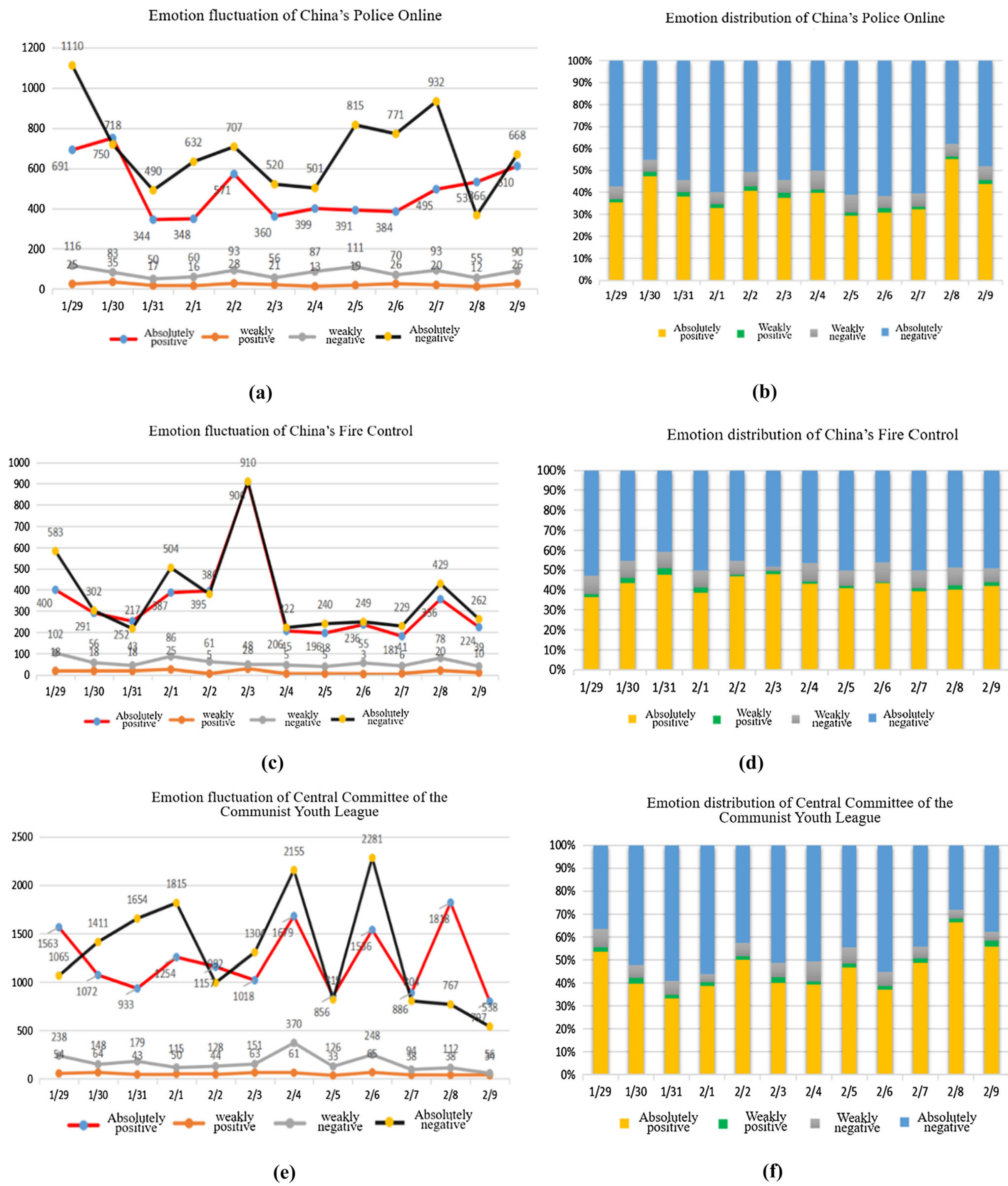
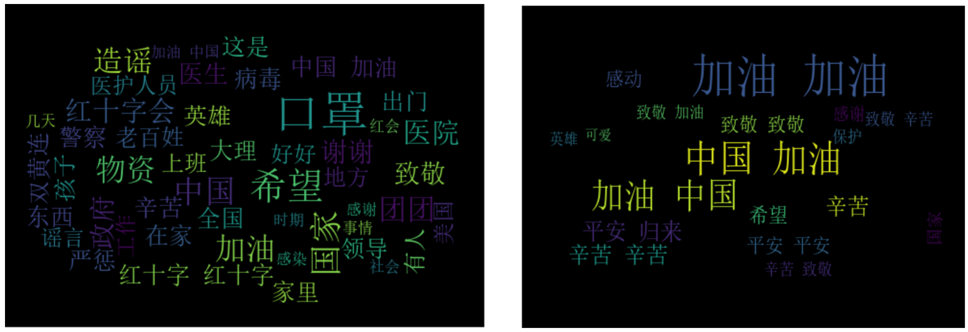


Fig. 5 Emotion comparison



(a) absolute negative emotional comments **(b)** absolute positive emotional comments

Fig. 6 Word cloud model

4.2 Comparison of the same type and cross platform government media

Here, we compare the communication effect and strength of each platform by comparing the four indicators, number of contents, number of retweets, number of praise and number of comments of the same government media in two representative social media platforms, micro-blog and DouYin.

The number of released contents of all three government medias on DouYin is much larger than micro-blog. The total number of contents on micro-blog platform remains stable, the average number of online articles sent by China's Police Online is 40, China's Fire Control is 12 and Central Committee of the Communist Youth League is 16, and the reports on incidents are frequent; while the releasing frequency of DouYin was very low, basically 2 to 3 days in a day, and there would be "zero releasing" phenomenon, such as China's Police Online in February 9 and Central Committee of the Communist Youth League zero releasing for three times. It can be seen that at present, the government media use micro-blog as the main operation platform. Micro-blog focuses on event notification and daily interaction with netizens, and users have high stickiness. DouYin is a short video platform, which is different from micro-blog. It focuses on releasing some videos that encourage morale, so as to publicize the government's image.

The total number of comments and retweets of micro-blog is higher than DouYin. The public's recognition of micro-blog is even higher. On the one hand, more contents are released on a single day. On the other hand, the released content is more popular with netizens. The contents of DouYin are relatively simple, which are hard to arouse the discussion of netizens.

The number of praises on DouYin is much higher than micro-blog, and there often appears "phenomenon" blog, that is, the number of praises of single video is as high as several hundred thousand. China's Fire Control released only one video about the novel coronavirus pneumonia on February 5, but the number of praises reached 1.277 million. It can be seen that in short video platform, video has a stronger impact on people. When released content can mobilize the emotions of netizens, people are more keen to praise video content (Table 4).

In order to make a more scientific comparison, we divide the number of comments, retweets and praises by the number of contents on the same day, calculate the rate of comments, retweets and praises, and average the 12-day indexes. Table 5 shows the details of the three indicators, and Fig. 7 shows the comparison of the three indicators under the two platforms.

The forward rate and comment rate are different from each platform. China's Police Online and China's Fire Control have higher forward rate and comment rate on DouYin, but Central Committee of the Communist Youth League is higher on micro-blog platform. On the indicator of praise rate, the three government medias on DouYin is far higher than on micro-blog. Overall, China's Police Online and China's Fire Control have better communication effect on the platform of DouYin, while Central Committee of the Communist Youth League has better communication effect on micro-blog platform.

5 Summary

This paper takes 2019 novel coronavirus pneumonia as the research event, selects three different types of government media: public security government media "China's Police Online," public security management government media "China's Fire Control," central government media "Central Committee of the Communist Youth League" to make a comparative study on the communication effectiveness of government media from two different perspectives, namely "the same platform, different government media" and "different platforms, the same government media." Using deep learning technology to build LSTM emotion

Table 4 Summary of statistical data of four indicators

Media	Indicator (number of)	Platform	Date											
			1.29	**1.30**	**1.31**	**2.1**	**2.2**	**2.3**	**2.4**	**2.5**	**2.6**	**2.7**	**2.8**	**2.9**
China Police Online	Contents	micro-blog	40	41	34	37	41	43	37	37	41	42	38	42
		DouYin	1	2	5	4	3	2	7	1	0	0	0	1
	Retweets	micro-blog	3846	6896	2531	2273	4199	2811	2605	3091	2989	3036	4454	3453
		DouYin	322	235	1284	640	584	122	913	136	0	0	0	101
	Comments	micro-blog	3529	5691	1880	1981	2403	1601	1469	2291	1909	2802	1994	2117
		DouYin	86	70	278	182	305	115	314	136	0	0	0	48
	Praise	micro-blog	58,677	103,841	12,644	13,453	17,521	10,260	13,711	15,202	14,593	14,719	12,533	25,102
		DouYin	17,000	34,000	81,000	50,786	64,000	40,000	144,000	22,000	0	0	0	29,000
China Fire Control	Contents	micro-blog	13	13	8	15	19	15	9	20	16	15	15	16
		DouYin	1	1	4	1	2	2	1	1	1	1	1	0
	Retweets	micro-blog	1449	6882	1774	2648	3565	2477	1531	14,691	35,925	2363	2381	2296
		DouYin	138	84	359	555	377	921	2658	6531	50	364	95	0
	Comments	micro-blog	1400	2565	678	2148	2678	2152	1211	4840	9753	1952	1640	1981
		DouYin	37	95	294	27	339	10,052	77	412	47	28	29	0
	Praise	micro-blog	5400	19,396	3425	7671	13,125	7928	4041	64,606	16,285	8966	12,788	6585
		DouYin	7380	23,000	146,760	6282	69,000	659,000	63,000	1,276,000	7254	3948	4170	0
Central Committee of the Communist Youth League	Contents	micro-blog	14	17	16	21	16	24	16	18	14	13	15	15
		DouYin	1	4	1	0	1	1	6	3	2	0	1	0
	Retweets	micro-blog	7500	11,114	8798	62,074	5750	9703	15,033	1,611,935	7691	18,911	37,803	7325
		DouYin	104	6593	37	0	1706	482	1044	498	197	0	175	0
	Comments	micro-blog	3554	8526	11,290	16,787	4670	9927	9775	23,327	10,537	23,912	12,659	5489
		DouYin	40	252	25	0	145	190	449	296	53	0	63	0
	Praise	micro-blog	299,861	151,425	171,818	90,509	193,990	106,882	124,522	45,130	204,375	100,961	86,314	43,474
		DouYin	11,000	379,469	12,000	0	63,000	25,000	218,983	47,000	35,859	0	3826	0

Table 5 Details of the three indicators

Media	Platform	retweet rate	comment rate	praise rate
China's Police Online	Micro-blog	88	62	651
	DouYin	164	67	19,533
China's Fire Control	Micro-blog	416	180	885
	DouYin	1019	547	162,884
Central Committee of the Communist Youth League	Micro-blog	8413	722	8604
	DouYin	510	80	31,087

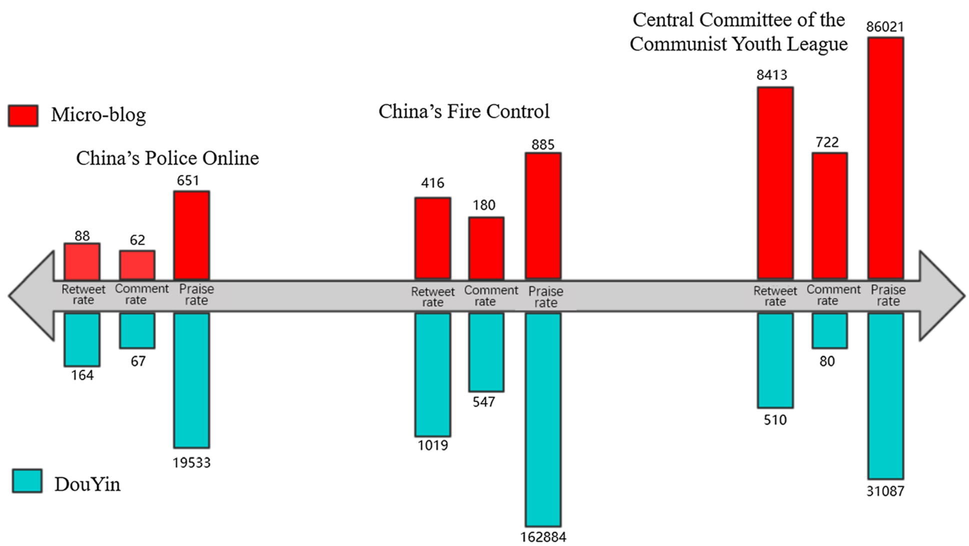


Fig. 7 Comparison of comment rate, retweet rate and praise rate of each government media on the two platforms

classification model, based on the user perspective to analyze the user comments in micro-blog government media, from the two aspects of emotion fluctuation and emotion distribution, it is concluded that: the users' emotion trend of Central Committee of the Communist Youth League is the most unstable, China's Police Online has the highest proportion of negative comments, and China's Fire Control has the most smooth emotion distribution. Four indicators of the same time period: the number of contents, comments, praises and retweets of the above three government media on micro-blog and DouYin are statistically analyzed. Results show that China's Police Online has more outstanding communication effect on DouYin than on micro-blog, while Central Committee of the Communist Youth League has better communication effect on micro-blog platform.

There are also some shortcomings in this study. On the one hand, there are not enough samples of training statements to cover the vocabulary and specific expressions involved in the whole public opinion event, which has an impact on the generalization ability of the model, so the prediction accuracy is deficient; on the other hand, only three different types of government media are selected, so the representativeness of the study is not enough. In future research, we consider increasing the number of training samples to further improve the classification accuracy of the model and choose more types of government media to enhance the representativeness of the research.

Author Contributions Design of topic selection and research framework, WN; Data collection and analysis and paper writing and revision, LX and SS. All authors have read and agreed to the published version of the manuscript.

Funding This work was supported by Key Project of Jilin Province Education Science During the 13th Five Year Plan in 2020: Research on new teaching mode in big data cloud education environment (ZD20024).

Declarations

Conflicts of interest The authors declare no conflict of interest.

References

1. China Internet Information Center, 44th CNNIC China Internet Report [EB/OL]. [2019–08–30]. http://www.cac.gov.cn/2019-08/30/c_1124938750.htm.
2. The 7th China Internet audio visual conference “Government communication in 5G era- Audio visual New Media Development Forum”, Research report on the development of government short video[EB/OL].[2019–05–29].http://media.people.com.cn/n1/2019/0529/c14677-31109068-2.html.
3. Xiaobo T, Shixuan Li, Mingliang T et al (2020) Review and Prospects of Study on Social Media in Government[J]. J Modern Inf 40(1):159–166
4. Rui YCW (2018) Applying H-Index to the Research on the Influence of Government Microblogging[J]. J Modern Information 38(3):110–123
5. Guiquan ZAC (2014) Positive Study on Evaluation and Comparison of Government Affairs Micro - Blog Influence: Based on Factor Analysis and Cluster Analysis[J]. J Intelligence 33 (3):108–112
6. Yihong R, Le Liu Xu, Eryu. (2016) “Internet+” Oriented Microblog Rransformation Strategy: Based on the Evaluation of the Effectiveness of the Government Micro-blog of Beijing, Shanghai, Guangzhou and Shen Zhen[J]. E-Gov 8:53–63
7. Guohua W, Tengfei WCY et al (2015) A Study on the Response of Governmental Weibo to Internet Public Opinion Crisis in Emergency —A Case Study of @ Shanghai Announcement in Shanghai Stampede[J]. J Intelligence 34(4):66–70
8. Zhao J, Sun N (2020) Government subsidies-based profits distribution pattern analysis in closed-loop supply chain using game theory. Neural Comput & Applic 32:1715–1724
9. Liu C, Yu T (2020) The regulatory warning model of regional product quality based on the back-propagation artificial neural network. Neural Comput & Applic 32:1639–1648
10. Jing Jiang, Wentao Wang (2020) Research on Government Douyin for Public Opinion of Public Emergencies: Comparison with Government Microblog. J Intelligence 39(1):100–106
11. Yang Kaimo Wu, Mingfen Chen Tao (2019) Generalized Text Sentiment Analysis Review. J Comput Appl 39:6–14
12. Xiaoping Y, Zhongxia Z, Liang W et al (2017) Automatic Construction and Optimization of Sentiment Lexicon Based on Word2vec[J]. Comput Sci 44(1):43–74
13. Li Shixuan, Shi Wenxuan, Wang Jiancheng, Zhou Heshen (2021) A Deep Learning-Based Approach to Constructing a Domain Sentiment Lexicon: a Case Study in Financial Distress Prediction. Inf Process Manag 58(5):102673
14. Pang B, Lee L, Vaithyanathan S et al (2002) Sentiment Classification Using Machine Learning Techniques[J]. Empir Methods Nat Lang Process 9:79–86
15. Zeng Yu, Peiyu L, Wenfeng L et al (2017) Naive Bayesian Algorithm for Text Sentiment Classification Based Feature Weighting Integration[J]. J Northwest Normal Univ (Nat Sci) 53 (4):56–73
16. Chang Jun, Wenting Tu, Changrui Yu, Qin Chuan (2021) Assessing dynamic qualities of investor sentiments for stock recommendation. Inf Process Manag 58(2):102452
17. Bengio Y, Ducharme R, Vincent P et al (2003) A Neural Probabilistic Language Mode[J]. J f Mach Learning Res 3:1137–1155
18. Mikolov T, Sutskever I, Chen Kai, et al. Distributed Representations of Words and Phrases and Their Compositionality[C]. Proc of the 27th International Conference on Neural Information Processing Systems.USA: Curran Associates Inc., 2013: 3111–3119.
19. Zhanyong J, Yapeng T, Mang B (2019) Sentiment Analysis on Network Public Opinion of Sudden Disasters Based on Long-short Term Memory Network[J]. Inf Sci 37(5):142–154
20. Ronglei Hu, Rui lu, Qi xiao, et al (2019) Text Sentiment Analysis Based on Recurrent Neural Networks and Attention Model. Appl Res Comput 36(11):3282–3285
21. Tian Qianjin Yu, Guanghui JL (2018) Evolution of Online Public Opinion of Zhongtai Waste Incineration Power Plant Incidents[J]. Inf Res 8:33–36
22. List of government affairs. 2019 government microblog influence Report [EB/OL]. [2020–01–17]. http://yuqing.people.com.cn/n1/2020/0117/c209043-31553643.html.
23. Su J, Text Sentiment Classification (3): Segmentation or Non-segmentation [EB/OL]. [2016–06–29]. https://kexue.fm/archives/3863.
24. Kingma D , Ba J. Adam (2014) A Method for Stochastic Optimization. Comput Sci