

# The RCNN-LDA Network Public Opinion Analysis Model based on the Weibo UGC of Medical Staff

Huilin Gu

School of Management, Shanghai University, Shanghai 200444, China

guhuilin09@163.com

## Abstract

In January 2020, COVID-19 began to spread in China. At present, the national epidemic situation continues to improve, and medical personnel working in the front line in Wuhan have played a key role; the interactive propagation mode of Weibo has accelerated the spread of COVID-19 information speed, User Generated Content (UGC) based on Weibo of anti-epidemic first-line medical staff that recorded epidemic knowledge, real-time treatment information, etc., are of great significance to the national anti-epidemic and understand the true status of the epidemic; Wuhan medical staff fighting against the COVID-19 with all strength are under the great pressure of physiology and psychology, their negative emotions, opinions and other public opinions have an adverse impact on the epidemic work. Therefore, the research based on the Weibo UGC public opinion is of great importance. This paper built the RCNN-LDA(sentiment classification and topic analysis) Network Public Opinion Analysis Model based on the microblog UGC of Wuhan anti-epidemic frontline medical staff, and used the Recurrent Convolutional Neural Network (RCNN) to establish a sentiment praise and depreciation classification model to distinguish UGCs' positive or negative sentiment tendencies; using the Latent Dirichlet Allocation (LDA) document topic generation model to analyze the UGCs of negative tendencies, thus the results of public opinion analysis based on sentiment classification and topic analysis were obtained. The experimental results show that the Network Public Opinion Analysis Model in this paper can effectively express the results of public opinion analysis on the Weibo UGC data of the medical staff of Wuhan's anti-epidemic frontline. Finally, based on the results of public opinion analysis, a discussion and study was carried out, reviewing the efforts made by all sectors of society to solve the problems faced by anti-epidemic medical personnel during the period of COVID-19, analyzing the root cause of the "Medical Trouble" from the course of China medical and health system reform and concluding the historical stage of China current Medical Reform, then some suggestions were proposed.

## Keywords

Covid-19; Emotion Classification; Deep Learning; Topic Analysis; RCNN-LDA Network Public Opinion Analysis Model; New Medical Reform.

## 1. Introduction

Since the outbreak of the new pneumonia, it had attracted great attention from people all over the world. On February 28, 2020, the World Health Organization raised the assessment of novel coronavirus transmission risk and impact risk to "very high". Elisabeth Mahase[1] found that although the mortality rate of novel coronavirus cases were low, the number of deaths caused by covid-19 exceeded the sum of SARS and MERS. According to the National Health Commission and Dr. Ding Xiang Yuan official website statistics, as of 24:00 on April 12, the COVID-19 epidemic had spread to at least 211 countries and regions, and at least 67 countries in the world

had declared a state of emergency. The cumulative number of confirmed cases of COVID-19 in China had reached 83482, and the cumulative number of confirmed cases abroad had achieved 1795163. In December 2019, COVID-19 was first detected in China. In late January, it rapidly spread throughout the country. In early February, the number of confirmed cases in all provinces except Hubei Province dropped continuously, and the COVID-19 epidemic gradually entered the stage of calm. As the main battlefield in the fight against novel coronavirus, a total of 255 medical teams and 32572 medical and nursing personnel had been sent to support Wuhan. The frontline medical personnel in the fight against the epidemic have played a key role in the national epidemic prevention and control.

With the development of new media such as Weibo, they have become the main information platforms for the public to receive information in fragmented time. In the information dissemination of the COVID-19, many medical staff passed the real name authentication of Weibo, and a large number of user generated content (UGC) emerged, which effectively disseminated novel coronavirus related knowledge, medical protection knowledge, treatment information, etc., making the UGC of the frontline medical staff in the epidemic prevention based on Weibo an important source of epidemic information. This was beneficial for the masses to know the latest information of the epidemic situation and the real treatment of medical staff in real time.

Song Quanxi[2] et al. based on the research results of many scholars on the definition of public opinion, proposed that public opinion was a collection of activities reflecting the public's expression and exchange of tendentious views on social emergencies through certain media. The concept of Internet public opinion is based on public opinion. Wang Lianxi[3] believed that Internet public opinion was a collection of activities such as network expression, communication and interaction that took the Internet as the carrier and specific social events trigger Internet users to produce ideologies with emotions, viewpoints, psychology and other tendencies. According to the definition of network public opinion, the emotion classification and opinion analysis of public opinion are important research contents of public opinion analysis in the current network environment. Research on Wuhan anti epidemic frontline medical staff expressing UGCS with emotions and views related to the epidemic through microblogs belongs to the field of online public opinion research. As the main force in the fight against the epidemic, the majority of medical personnel have heavy work tasks and high infection risks. They have born great physical and psychological pressure. Negative emotions have a negative impact on the fight against the epidemic. Analyzing the UGC views of medical personnel on Weibo is conducive to understanding the real treatment situation and needs. Therefore, the public opinion analysis of Wuhan anti epidemic front-line medical staff has important research significance.

This paper analyzes public opinion on the microblog UGC of front-line anti epidemic medical workers in Wuhan. The main research contents are as follows:

(1) RCNN-LDA(Recurrent Convolutional Neural Networks, RCNN; Latent Dirichlet Allocation, LDA; RCNN-LDA) network public opinion analysis model based on UGC of medical staff microblogs is proposed.

Emotion classification and opinion analysis are important research contents of Internet public opinion. In this paper, recurrent convolutional neural networks (RCNN) is used to establish an affective classification model to distinguish the positive and negative affective tendencies of microblog UGCS, and it is compared with the affective classification model based on Convolutional Neural Networks (CNN) [4]and Bidirectional Long Short Term memory Bi-LSTM [5] based on the attention mechanism, Experiments show that the RCNN emotion classification model is better; The hidden Dirichlet allocation (LDA) model is used to divide the negative trend UGC into themes, extract public opinion views, and trace back to the original negative micro blog UGC to determine the theme classification, and output the public opinion analysis results.

The network public opinion analysis model based on deep learning emotion classification and topic analysis ((RCNN-LDA) can effectively express the public opinion analysis results of the UGC data of the microblog of the front-line medical workers in Wuhan.

(2) Based on the results of public opinion analysis, we discussed and studied, focusing on the course of China's medical and health system reform and the root cause of the "medical trouble", summarized the historical stage of China's medical reform, and finally put forward relevant opinions.

This paper discusses and studies the analysis results of the network public opinion analysis model based on the emotion classification and theme analysis of the micro blog UGC of the first-line medical staff in Wuhan, and analyzes the help provided by all walks of life when the medical staff in Wuhan are facing difficulties in their anti epidemic work; In the year of building a well-off society in an all-round way, I reviewed the reform process of China's medical and health system from 1978 to 2020, and drew the root cause of the serious "medical disturbances" during the COVID-19 epidemic from the reform process. I concluded that China's medical and health system reform is in the stage of deepening reform to further reduce the price of medical services and comprehensively improve the quality of medical services, Finally, some suggestions are put forward.

## 2. Literature Review

### 2.1. COVID-19

Since the spread of the novel coronavirus epidemic, researchers at home and abroad have carried out a lot of research work on the policy opinions on regional prevention and control during the epidemic of the novel coronavirus, the characteristics of government information release, the impact of the COVID-19 epidemic on transportation operations and economy, the pathological characteristics of the new coronavirus patients and the changes of chest CT during recovery. Zhou Chenghu[6]and others put forward relevant policy suggestions on the big data analysis of the COVID-19 and the prevention and control in Hubei Province and all regions of the country. Feng Xujie[7]and others studied how the COVID-19 epidemic affected the operation of urban rail transit. Li Yuelin [8] and others discussed the characteristics of government information release during the COVID-19 epidemic. Chen Xikang [9] and others analyzed the impact of China's economy under the COVID-19 epidemic. Dong Hongzhe[10] and others investigated and studied the generation of public opinion secondary to COVID-19 public crisis events, and proposed relevant governance paths according to the public opinion governance model [10]. Wang Yi [11] and others studied the extent to which the public received media information and their mental health status in public health emergencies (COVID-19), providing a basis for scientific dissemination of public health knowledge. Zhe Xu [12] et al. Studied the pathological characteristics of death caused by novel coronavirus infection. Huijun Chen[13] et al. Evaluated the clinical characteristics of covid-19 in pregnant women and the possibility of vertical transmission of covid-19 from mother to child. Ying Liu[14] and other researchers found that compared with SARS coronavirus, covid-19 has a higher reproduction number. Zunyou Wu[15] and others summarized the characteristics and important experience of the outbreak of coronavirus disease 2019 (covid-19) in China in response to the 72314 case reports issued by the Chinese Center for Disease Control and prevention. Feng pan[16] et al. Studied the lung change time course of chest CT during the patient's recovery from novel coronavirus (covid-19) pneumonia in 2019. Ensheng Dong[17] and others studied a web-based interactive dashboard that can track covid-19 in real time, aiming to facilitate public health authorities and the public to track the progress of the epidemic. To sum up, researchers at home and abroad have conducted extensive research in the field of COVID-19 epidemic situation, but there is a lack of research on anti epidemic medical staff. As the main force in the fight against

the epidemic, medical staff have heavy work tasks and high infection risks. They bear both physical and psychological pressures. Negative emotions will have a negative impact on the anti epidemic work. The public opinion analysis of medical staff in the front line of anti epidemic has important research significance, which provides ideas for the selection of research content in this paper.

## 2.2. Text Emotion Classification

At present, there are two kinds of text emotion classification methods. One is the classification method based on emotion dictionary without training data; The second is the training set to be labeled, a classification method based on machine learning [18].

(1) Text emotion classification analysis based on emotion dictionary.

The emotion classification method based on the emotion dictionary mainly queries the weight of words in the emotion dictionary after establishing the emotion dictionary, and calculates the emotion value of the text. The research on emotion dictionaries in foreign countries started early, and there are mature emotion dictionaries such as sentiwordnet, general inquirer, opinion lexicon, etc; However, there are relatively few standardized Chinese affective dictionaries in China. The commonly used affective dictionaries include hownet affective dictionary and ntusd Chinese affective polarity Dictionary of National Taiwan University. V hatzivassiloglou[19] and others identified and verified the constraints of positive and negative semantic tendencies of adjectives from a large number of corpora, and built a log linear regression model considering these constraints to predict the positive and negative semantic tendencies of adjectives. Y Shen [20] and others built attitude dictionary, negative dictionary, degree dictionary and exclamation dictionary, and used the constructed dictionary to identify the emotion expressed by Weibo data. Hong Qing [21] and others conducted emotional analysis and user classification on bullet screen users by constructing an online bullet screen emotion dictionary. At present, the text emotion classification method based on the emotion dictionary mainly has the following three problems: 1) when some emotional words in the text sentence appear repeatedly or the positive and negative narrative dialectics appear, it will produce a great error to predict the positive and negative semantic tendencies of the sentence only by the method of weight calculation. 2) Affective dictionaries are usually composed of words with prominent affective features, which can not contain all affective words, resulting in ambiguity in affective dictionaries. 3) In the process of building an emotion dictionary, there are many difficulties such as the adaptation of the emotion dictionary field and the disambiguation of emotion words[22,23]. The above problems lead to the low recognition rate of positive and negative emotions expressed by Weibo data when emotion dictionary is used to classify Weibo data. Therefore, this paper adopts the text emotion classification analysis method based on machine learning.

(2) Text emotion classification and analysis based on machine learning.

The machine learning emotion classification method mainly uses the machine learning model to learn the features of the training set, save the relevant parameters and construct the model, so as to classify and judge the test set. Alexander Pak[24] and others proposed a method of automatically collecting twitter corpus, and used a polynomial Naïve Bayesian classifier to classify twitter data emotionally. Liu Zhiming[25] et al. Used three traditional machine learning algorithms, namely n-meta language model, Naïve Bayes and support vector machine (SVM), to conduct case studies on emotion classification of microblog data, and compared and analyzed the advantages and disadvantages of the three algorithm models in different feature weight calculation methods and feature selection methods. Cai Huiping[26] and others constructed a convolutional neural network (CNN) emotion classification model and compared it with traditional machine learning methods. Traditional machine learning text classifiers rely on the functions of dictionaries and knowledge bases that can be designed by others. Siwei Lai[27] and

others introduced the cyclic convolutional neural network (RCNN) into the text classification task. This RCNN text emotion classification model can effectively capture context information when learning features from the training set, and will introduce less noise than the traditional window based neural network; The model also adopts a maximum pooling layer, which can automatically judge which words play a key role in text classification to capture the key components in the text. This paper also compares and analyzes RCNN classification models. Experiments show that RCNN emotion classification model is obviously superior to other classification models such as convolutional neural network (CNN). In view of this, this paper uses deep learning RCNN emotion classification model to classify the emotion of Weibo UGC data.

### 2.3. Subject Analysis

In recent years, topic analysis has received extensive attention from scholars, and LDA topic classification model is the most widely used analysis model. Hidden Dirichlet distribution (LDA) [28] is a generation probability model used to collect discrete data. This model assumes that each document hides multiple topics, and finally outputs them in the form of probability distribution. The advantages of LDA include: 1) LDA is an unsupervised learning algorithm, which only needs to input the training document set and the number  $k$  of specified topics, and does not need to manually mark the training set. 2) For each topic, LDA can output some descriptive words. R Krestel[29] and others studied a label recommendation method based on potential Dirichlet assignment (LDA) to improve the search. The analysis shows that this method has higher accuracy and recall. Using these recommended labels to expand resources can significantly improve the search for new resources. Shi Jing[30] and others used LDA method to model and extracted the central theme of the full text. Yue Guo[31] and others used LDA method to mine data from online ratings and reviews of hotel customer service and extract key dimensions of visitor satisfaction. By studying 266544 online reviews from 25670 hotels in 16 countries / regions, they found 19 dimensions that are crucial to hotel management and visitor interaction. Hamed Jelodar[32] and others summarized the challenges faced by current topic Modeling Research on the basis of studying academic articles related to topic modeling based on LDA methods from 2003 to 2016, and introduced famous tools and data sets in topic modeling based on LDA. It can be seen from the above that the hidden Dirichlet distribution document topic generation model is widely used in the field of text topic recognition. In view of this, this paper uses this method to extract the theme of the negative original Weibo UGC after emotion classification.

### 2.4. Network Public Opinion Analysis Model

With the rapid development of the Internet industry, a large number of scholars have carried out a lot of research work in the field of Internet public opinion research. Juliane urban[33] and others proposed a new method to measure public opinion on the Internet, and quantitatively analyzed the construction process of network public opinion on specific events with the help of social networks. Wang Lancheng[34] and others carried out in-depth research on the literature review of the evolution law, early warning mechanism, coping theory and evaluation of online public opinion at home and abroad. The research shows that the relevant research of online public opinion in China is mainly theoretical research, lacking the research combining qualitative and quantitative methods. In recent years, with the development of machine learning, natural language processing and other technologies, researchers have carried out extensive research on online public opinion analysis models. LAN Yuexin[35] and others built a differential equation model to analyze the formation process of network public opinion in emergencies and its influencing factors. Zheng Changxing[36] and others proposed an analysis model of online public opinion based on the perspective of stakeholders, which provides theoretical support for the response and handling of online public opinion in emergencies. An



Lu[37] et al. Proposed an analysis model for the evolution of public opinion on microblogs in emergencies. This model integrates theme and emotion features. Based on word2vec technology, the microblog themes at various stages of the life cycle of the event are extracted respectively, and the evaluation emotions under different themes are finely divided using the emotion analysis method based on dictionaries. Wang Xiufang[38] and others proposed a microblog public opinion analysis model, which quickly clusters microblog topics, quantifies their emotional intensity using an emotional dictionary, and tracks and predicts the emotional changes of hot topics through time series regression. Ye Jian[18] and others proposed a public opinion analysis model of bullet screen data, specifically, they established an emotion classification model based on convolutional neural network (CNN), and used the model to divide the positive and negative bullet screen data, and obtained the public opinion analysis results on this basis. To sum up, due to the late start of the domestic Internet, the current domestic research on online public opinion is still at the initial stage, lacking of qualitative and quantitative research; For the research of network public opinion analysis model, there are mainly methods to build public opinion analysis model based on differential equation model, relevant feature indicators, fusion theme and emotional features. According to the definition of online public opinion in Chapter 1.1.2, the emotional classification and opinion analysis of public opinion are important research contents of public opinion analysis in the current network environment. At present, the research on emotion classification and opinion analysis of network public opinion mainly uses the technical methods based on emotion dictionary and topic division, which has problems such as low accuracy of emotion classification of public opinion and unstable results of topic division. In view of this, this paper constructs an RNN-LDA network public opinion analysis model that integrates RCNN deep learning emotion classification and LDA theme analysis, and analyzes the microblog UGC of front-line medical workers in Wuhan.

### 3. RCNN-LDA Network Public Opinion Analysis Model

#### 3.1. RCNN-LDA Network Public Opinion Analysis Model Framework

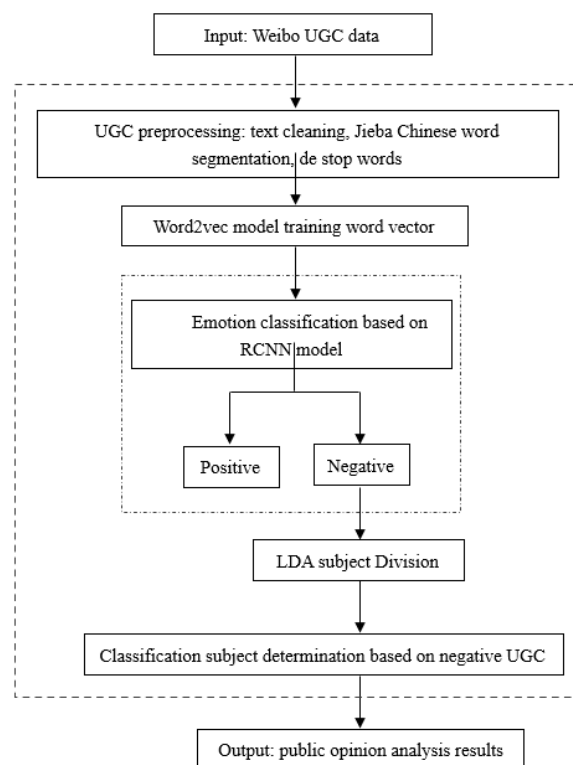


Figure 1. RCNN-LDA Network public opinion analysis model diagram

The framework of the network public opinion analysis model of Weibo UGC in this paper is shown in Figure 1. The model is divided into three processes: input, processing and output, and includes seven key steps. In the first stage, the input Weibo UGC data will be input into the model from the collected Weibo UGC data of the front-line medical and health care workers in Wuhan. In the second stage, the collected micro blog UGC is preprocessed, including cleaning the text, removing low-frequency words after sentence segmentation, and finally using the stop word list to stop words. In the third stage, the word2vec model is used to train the preprocessed Weibo UGC data and obtain the word vector corresponding to each word. In the fourth stage, the classification model based on RCNN is used to classify emotions, and positive UGC and negative UGC are obtained. In the fifth stage, LDA theme division model is used to divide the negative Weibo UGC data into themes. In the sixth stage, according to the obtained theme classification results, the original negative UGC data is traced back to further determine the classification theme and mark the UGC. In the seventh stage, the public opinion analysis results are output according to the classification theme and the marked negative UGC.

### 3.2. RCNN-LDA Network Public Opinion Analysis Model Training

According to the framework of network public opinion analysis model constructed in Chapter 3.1, the training of rcnn-lda network public opinion analysis model is mainly divided into the following seven steps.

Step 1: input Weibo UGC data.

Based on the research results of Shi Yaolin[39], this paper divides the development of COVID-19 epidemic into four stages. Specifically, in the first phase (from December 1, 2019 to January 23, 2020), before the closure of Wuhan, the nucleic acid detection standard was high, the epidemic prevention work was not strict, and the epidemic spread; In the second phase (from January 24 to February 1), after the closure of Wuhan City, traffic control was implemented to limit the flow of population and prevent the continuous and large-scale spread of novel coronavirus; In the third phase (from February 2 to February 8), "Huoshen mountain" and "Leishen mountain" hospitals and many "shelter hospitals" were built and put into use one after another, and medical rescue teams from all over the country rushed to help. At the same time, Wuhan communities implemented a management model of "comprehensive screening without missing one household and full prevention and control without missing one person", which further improved Wuhan's medical level and prevention and control, and accelerated the digestion of Wuhan's suspected patient stock; In the fourth stage (after February 9), Wuhan adopted a strict policy of community isolation and comprehensive screening of suspected infected people, and basically realized "all due and all due" for covid-19 patients. Except for Hubei, the confirmed cases in other provinces and cities continued to decline, and entered the phase of epidemic regression.

Therefore, the micro blog corpus data in this paper includes 6085 micro blog (www.weibbo.com) UGC data of Wuhan anti epidemic front-line medical personnel from December 1, 2019 to February 23, 2020, and 120000 micro blog corpus data from 2017 to 2018.

Step 2: UGC preprocessing.

First of all, the original UGC data of 6085 Wuhan anti epidemic front-line medical personnel collected were text cleaned, and special symbols such as numbers, Chinese and English were removed; After that, add user-defined dictionaries (such as the proper nouns "Huoshen mountain" and "Leishen mountain"), and use the Jieba Chinese word segmentation module in Python to segment the original UGC data and remove words with a word frequency less than 5; Finally, the halting vocabulary list of Harbin Institute of technology is adopted to Delete stop words.

Step 3: train word vectors.

The preprocessed words belong to natural language, and the natural language needs to be represented by data before it can be recognized by the computer. This process is an important step of natural language processing (NLP). The expression methods of words are divided into traditional semantic dictionary method, one hot representation method and current mainstream distributed representation method. The distributed expression methods of words are divided into three categories: distributed expression based on neural network, matrix and cluster. The distributed expression method based on neural network is widely used, also known as word vector, word embedding or distributed representation, Compared with other representation methods, word vectors can represent more complex and rich contextual semantic information[40]. Tomas mikolov [41] and others proposed word2vec algorithm, including skip gram model and cbow (continuous bag of words) model. Both models can be used to train word vectors. Skip gram model is worse than cbow model in grammatical tasks and takes longer to train, but it is much better than cbow model in semantic tasks. Therefore, this paper will apply skip gram model to train word vectors, which is implemented through Python's gensim Library Word2vec algorithm in Realized. Specifically, the word2vec model is used to train the 120000 pieces of Weibo data corpus built in this paper, and obtain the word vector corresponding to each word. The training results are saved as word vectors for use in subsequent stages.

Step 4: train RCNN classification model and output positive and negative Weibo UGC.

In this paper, RCNN's supervised learning model is used to classify the preprocessed Weibo UGC data. The experiment is developed using Python 3.6, and the deep learning framework used is tensorflow 1.8.0. The data set includes a training set for training the model and a test set for testing the model. On the training set, due to the outbreak of COVID-19, and social media being used as a platform for epidemic propagation for the first time, this paper adds some manually labeled epidemic data to the marked Weibo data from 2017 to 2018 and uses them as a training set, and uses the labeled epidemic Weibo data as a test set to establish a training model for supervised learning. Specifically, the data set used for training the model includes 25000 pieces of manually labeled Weibo data from 2017 to 2018 and 1000 pieces of manually labeled epidemic data of medical personnel on the front-line of anti epidemic in Wuhan from January 2019 to February 23, 2020, of which 12600 are positive and negative. The data set used for the test is 300 pieces of COVID-19 data manually marked, including 150 pieces of positive and negative data. Finally, the RCNN model prediction results are output.

Step 5: LDA theme division.

This paper uses the hidden Dirichlet allocation (LDA) model described in Chapter 2.3 to divide the theme of the negative micro blog UGC. First, the Jieba part of speech tagging module is used for the preprocessed negative Weibo UGC data output in step 4 Tag to get the attributes of words, filter out nominal keywords and carry out word frequency statistics. Then, in order to reduce the impact of low-frequency words on topic classification, this paper selects nominal keywords with word frequency greater than or equal to 20 to form a nominal keyword dataset marked with word frequency. Next, using Python's gensim Library LDA model in digs out potential topics from the data set.

Step 6: determine the classification subject based on negative UGC.

In order to reduce the uncertainty of LDA theme analysis results, this paper traces back the original negative micro blog UGC of Wuhan anti epidemic front-line medical staff to further determine the theme classification. Specifically, according to the theme classification results output by LDA model, we compared and analyzed them with the original negative Weibo UGC, verified the effectiveness of the LDA theme analysis results, improved the classification theme and labeled the negative original UGC.

Step 7: output the results of public opinion analysis.



Carry out multi-dimensional analysis according to the tagging results of the original negative micro blog UGC, such as analyzing the changes of various classified themes over time and the dynamic changes of various themes at different stages of development of the COVID-19 epidemic.

## 4. Results

### 4.1. RCNN Emotion Classification Results and Analysis

Through experiments, the classification model based on RCNN is compared with the classification models based on CNN and Bi-LSTM based on the attention mechanism. For the training set, the Weibo corpus from 2017 to 2018 is mainly used. As can be seen from Table 1, the classification effect of RCNN model is the best, and there is no significant difference with CNN model and Bi LSTM + Attention model.

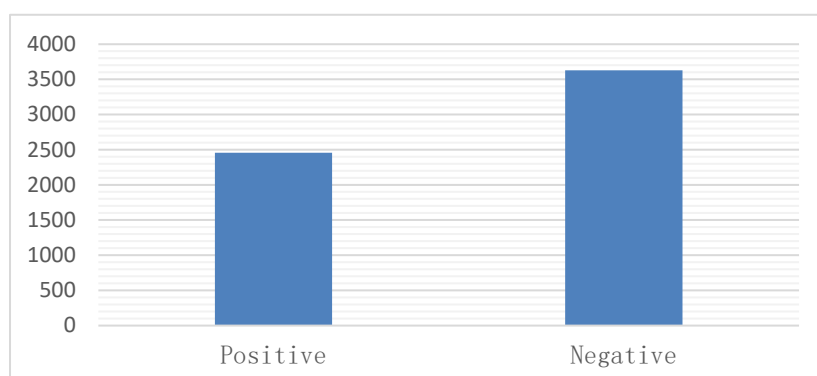
**Table 1.** evaluation of training set emotion classification model

Model	Accuracy	Recall	Precision	F1-Score
RCNN	0.98	1.0	0.97	0.98
CNN	0.97	0.98	0.98	0.96
Bi-LSTM+Attention	0.96	0.95	0.97	0.96

In order to verify the generalization ability of this model in the emotional classification of Weibo epidemic data, 300 pieces of epidemic data of Wuhan anti epidemic frontline medical workers have been marked to test and evaluate the model. Table 2 shows the test results of this paper. It can be seen from table 3 that since Weibo is the platform for large-scale epidemic information dissemination for the first time, the previous Weibo corpus does not contain COVID-19 related data, so the classification effect of each model relative to the training set has decreased, but the classification effect based on RCNN model is significantly better than the other two classification models. Therefore, the classification model based on RCNN can more effectively express the emotional classification results of the Weibo UGC data of Wuhan anti epidemic front-line medical staff.

**Table 2.** evaluation of emotion classification model of test set

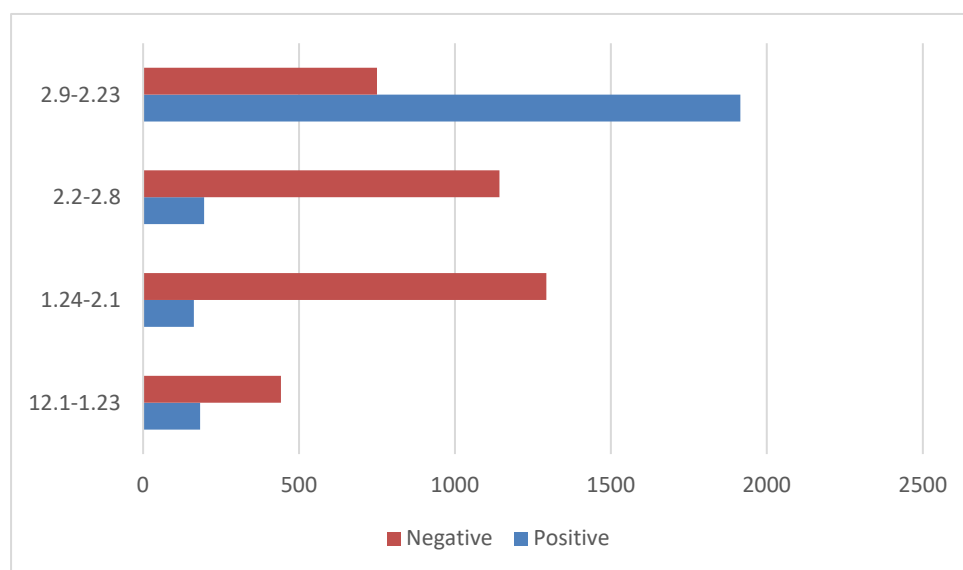
Model	Accuracy	Recall	Precision	F1-Score
RCNN	0.73	0.86	0.67	0.76
CNN	0.64	0.70	0.62	0.66
Bi-LSTM+Attention	0.63	0.41	0.74	0.53



**Figure 2.** Distribution of positive and negative micro blog UGCs

The distribution of the positive and negative micro blog UGC epidemic data obtained from the classification of 6085 micro blog UGC data of medical personnel in Wuhan from December 1, 2019 to February 23, 2020 collected in this paper is shown in Figure 3. Among them, the positive emotional tendency UGC accounts for about 40% (2457), and the negative emotional tendency UGC accounts for about 60% (3628), as shown in Figure 2.

Analyze Weibo UGCs with positive and negative tendencies in stages, as shown in Figure 3. The analysis results show that the negative emotion expressed by UGC on Weibo rises continuously in the first stage, the second stage and the third stage, and falls in the fourth stage; Positive emotions first decreased and then increased in the first stage, the second stage and the third stage, and tended to be stable in general, but suddenly increased in the fourth stage; The gap between positive and negative emotional UGC is small in the first stage and the largest in the fourth stage. In general, with the development of the epidemic, the positive and negative emotions of the first-line anti epidemic medical personnel in Wuhan have a dynamic trend, that is, in the epidemic diffusion period, the epidemic outbreak period and the epidemic stabilization period (there is still a huge stock of infection cases in Wuhan), the negative emotional UGC of the first-line anti epidemic medical personnel in Wuhan is significantly higher than the positive emotional UGC; But in the regression period of the epidemic, the positive emotional UGC increased significantly and the negative emotional UGC decreased.

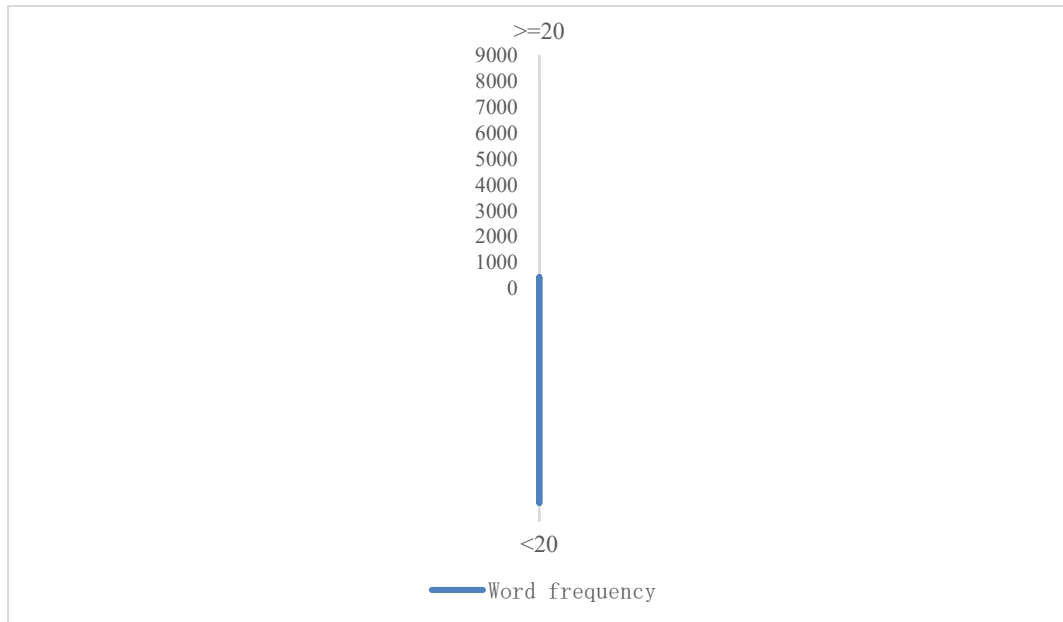


**Figure 3.** Positive and negative Weibo UGC stage distribution

#### 4.2. LDA Subject Division Results and Analysis

According to the description of LDA theme analysis model steps in the previous chapter 3.1, 8730 non repetitive noun keywords were collected from the negative Weibo UGC after emotion classification, including 422 high-frequency nouns with word frequency greater than or equal to 20. The word frequency distribution of negative Weibo UGC noun keywords is shown in Figure 4. Among them, the opinions of the first-line anti epidemic medical staff on Weibo UGC mainly focus on COVID-19, rescue materials, treatment work, doctors and patients.

According to the above analysis results of high-frequency words, this paper assumes that there are four themes in the negative Weibo UGC data set, so the number of themes K of the LDA theme analysis method is set equal to 4. After LDA theme analysis, the Weibo UGC dataset of Wuhan anti epidemic front-line medical personnel was divided into four themes, and 10 most likely words and their corresponding probabilities were generated under each theme.



**Figure 4.** Word frequency distribution of nominal keywords in negative Weibo UGC

After analysis and collation, table 3 shows the main contents of LDA model after theme division. Among them, the effective keywords screened in theme 1 are mainly related to rescue materials. The effective keywords screened in topic 2 are mainly related to rescue work. The effective keywords screened in theme 3 are mainly related to social events. The effective keywords screened in topic 4 are mainly related to doctor-patient relationship.

**Table 3.** LDA subject classification results

Topic 1	Topic 2	Topic 3	Topic 4
nucleic acid	Wuhan Union Medical College Hospital	Red Cross Society	Doctor patient
Shuang huang lian	Hospital director	society	family members
disease	lung	Old men	Central South
guide	citizen	war	illness
many	result	pneumonia	stage

### 4.3. Classification Subject Determination based on Negative UGC

According to the LDA theme analysis results, this paper compares and analyzes 3628 Weibo UGCs with negative emotion tendency after RCNN emotion classification. Specifically, theme 1 related to rescue materials, Theme 2 related to rescue work, Theme 3 related to social time, and theme 4 related to doctor-patient relationship are marked as 1, 2, 3, and 4 respectively, and the original negative micro blog UGC is marked.

After comparing and marking 3628 original negative micro blog UGCs, the results show that a total of 1793 UGCs are related to the divided themes, including 248 for theme 1, 818 for Theme 2, 649 for Theme 3 and 78 for theme 4, as shown in Figure 5. Then, the negative original micro blog UGC was analyzed, and the four themes that made the front-line medical staff in Wuhan have negative emotions were identified, namely, theme 1 showed that there was a shortage of medical resources in the early stage, insufficient medical staff and insufficient logistics support; Theme 2 indicates that the medical staff in the anti epidemic front line are overloaded, at high risk of infection, and under great physiological and psychological pressure; Theme 3: social negative events and rumor spreading; Theme 4 is medical disputes and injuries.

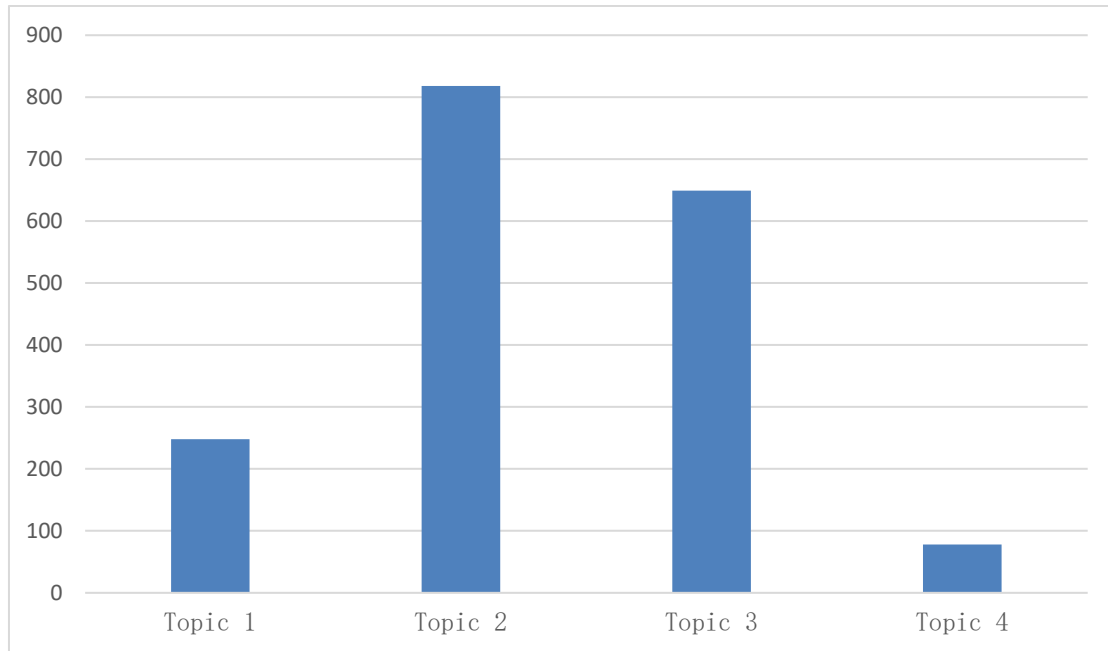


Figure 5. Original negative UGC subject mark

#### 4.4. Public Opinion Analysis Results

Since the outbreak of the COVID-19 epidemic, medical personnel have worked tirelessly in the forefront of the fight against the epidemic regardless of their personal safety. Their overload work and high risk of infection have almost reached the limit of what they can bear physically and psychologically. Negative emotions will have a negative impact on the physical and mental health of medical staff and the anti epidemic work. This paper will conduct multi angle public opinion analysis according to the theme analysis results output in Chapter 4.3. The detailed analysis results are as follows.

(1) There is a shortage of medical resources, medical staff and logistics support in the early stage.

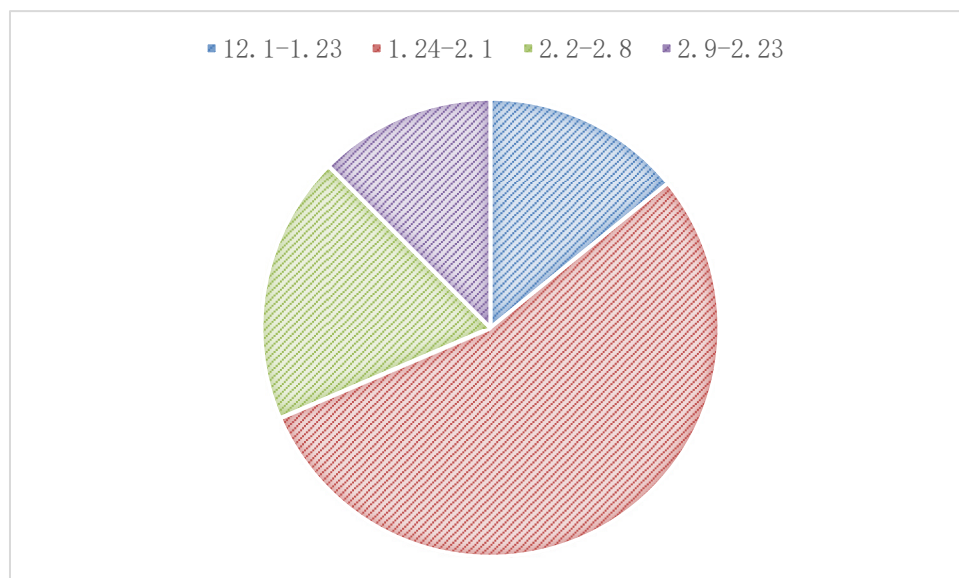


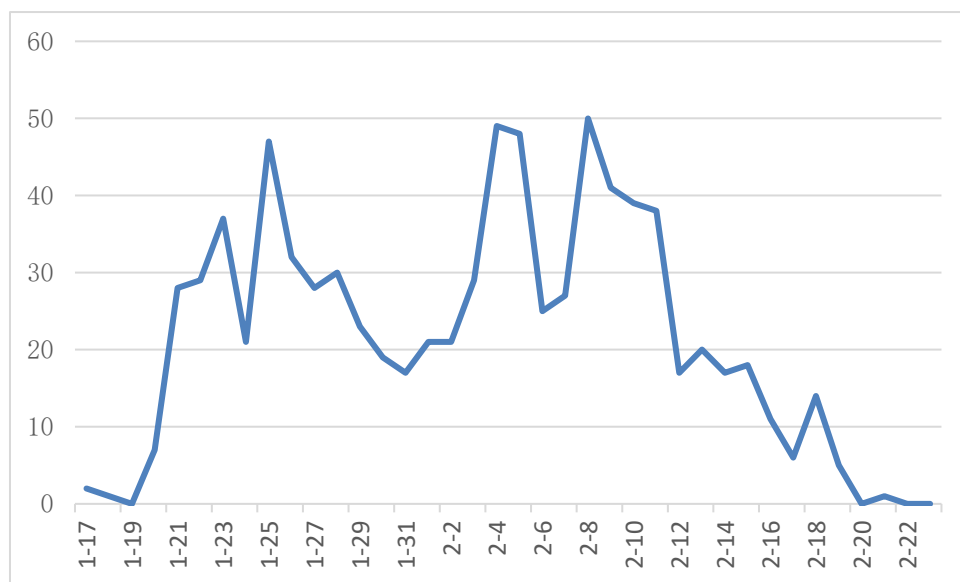
Figure 6. UGC stage distribution diagram of shortage of medical resources, medical staff and logistics support

After analysis, a total of 248 original UGCs showed a shortage of medical resources, insufficient medical staff and insufficient logistics support in the early stage, as shown in Figure 6. After the

closure of Wuhan City, Wuhan citizens went to the hospital in panic, and a large number of people poured into the hospital. Medical resources, especially masks, protective clothing and goggles, were consumed at an accelerated pace. During the Spring Festival holiday, the supply of medical resources could not be sustained, and medical resources were run. Hospitals are overcrowded, resulting in a shortage of medical staff. After the closure of Wuhan, the stoppage of public transport has brought inconvenience to medical staff, and the supply of hospital meals and other logistics support is insufficient. Therefore, during the spread and outbreak of the COVID-19 epidemic, the panic of the masses caused a shortage of medical resources, insufficient medical staff and insufficient logistics support, which led to the negative feelings of anti epidemic medical staff.

(2) The treatment work is overloaded, the infection risk is high, and the physiological and psychological pressure is high.

After analysis, a total of 818 original UGCS showed that medical staff were overloaded with rescue work, high risk of infection, and high physiological and psychological pressure, as shown in Figure 7. In late January, the number of patients infected with COVID-19 in Wuhan hospitals surged, leading to a run on hospital medical resources and a collapse of the medical system. At the same time, medical and nursing personnel also need to risk their lives to carry out treatment. In the early stage of the epidemic alone, more than 3000 medical and nursing personnel in Hubei Province were infected. The high-intensity and even overloaded rescue work and the harsh and dangerous working environment have all caused tremendous pressure on the psychology and physiology of the front-line medical staff. At the beginning of February, with the implementation of the policy of community screening and full collection of receivables, the number of confirmed patients increased dramatically, and the number of COVID-19 infection channels increased. The virus may mutate, further aggravating the physical and mental pressure of the front-line medical personnel in the fight against the epidemic.



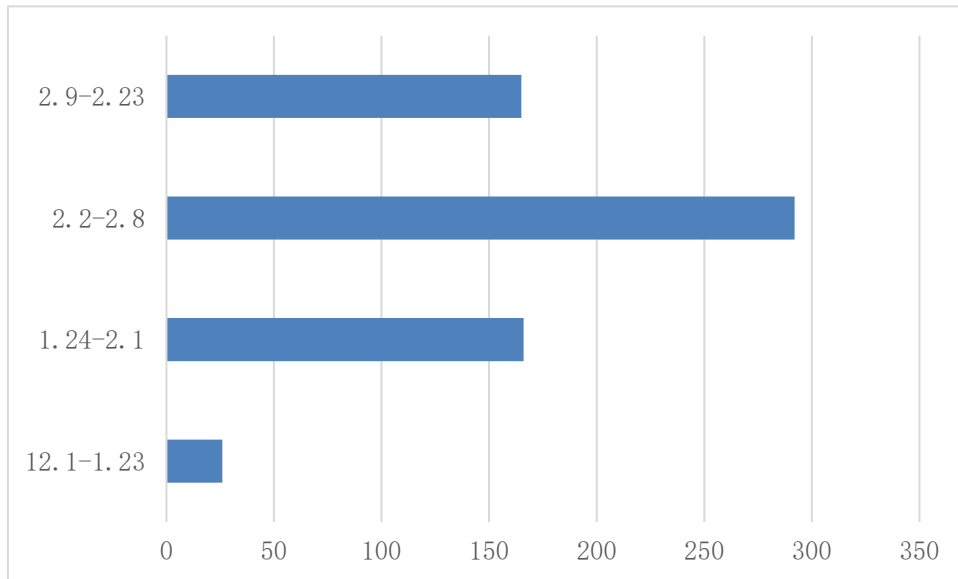
**Figure 7.** UGC time distribution of medical staff with overload, high risk of infection and high physiological and psychological pressure

(3) Negative social events and rumors.

After analysis, a total of 649 original UGCS showed that the spread of social negative events and rumors caused negative emotions among medical personnel, as shown in Figure 8. On January 30, the official website of the Red Cross Society of Hubei Province announced the list of the use of donated materials received for the first time since the outbreak of pneumonia. The list shows

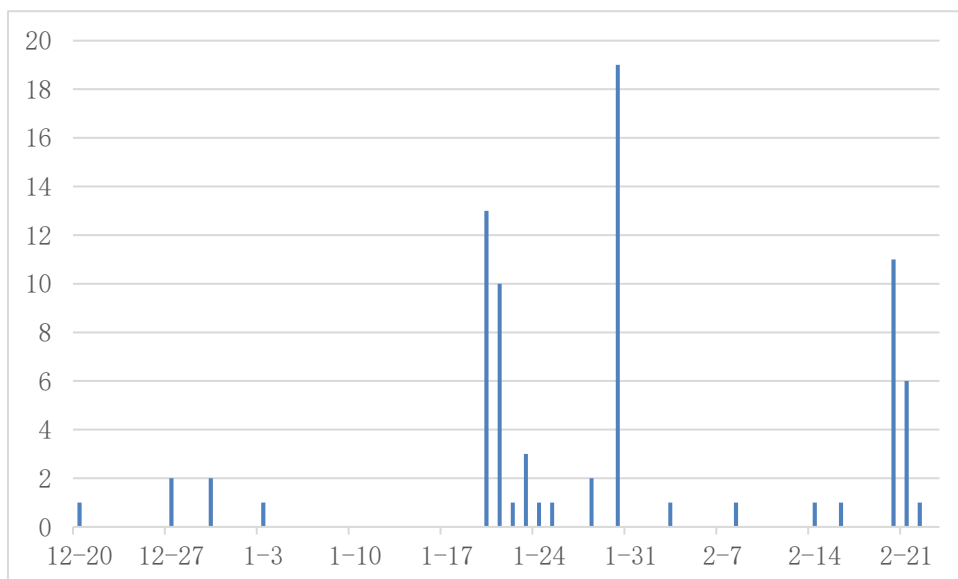


that the Red Cross Society donated 16000 masks to a Putian private hospital, while Wuhan Union Medical College Hospital received only 3000 masks from individuals, triggering a storm in mask distribution. Despite their personal safety, medical staff in public hospitals are at the forefront of the fight against the epidemic, but they can only be allocated very few masks, which makes countless medical staff feel cold and desperate. In the early stage of the epidemic, rumor mongers spread rumors about the death of infected medical staff (such as the "Li Wenliang" incident), which caused panic. In the face of infection and death of peers and patients, it increased the negative emotions and sense of helplessness of medical staff.



**Figure 8.** UGC stage distribution of social negative events and rumors

(4) Medical disputes and injuries.



**Figure 9.** UGC event distribution of medical incidents and injuries

After analysis, a total of 78 original UGCS showed that medical personnel had negative emotions due to medical disturbances and injuries, as shown in Figure 9. On December 24, 2019, Dr. Yang Wen of Beijing Civil Aviation General Hospital was killed; On January 20, 2020, Dr. Tao Yong of

Beijing Chaoyang Hospital was hacked; On January 30, two doctors in Wuhan were injured by the patients' families; On February 20, an official and patient scolded the nurse for no reason. During the COVID-19 epidemic, the majority of medical and health care workers rushed to the forefront of the epidemic, risking their lives to treat patients, but there was a personal risk of being injured by patients. Medical incidents such as medical disturbances and injuries have chilled countless medical personnel and impacted their professional belief and sense of industry honor of "saving the life and healing the wounded".

#### 4.5. Discussion and Suggestions

Since the outbreak of covid-19, the party and the state have attached great importance to it and actively carried out epidemic prevention and control; Large enterprises, the public, overseas Chinese and international friends have donated money and raised materials; In particular, the vast number of medical personnel have been bleeding to fight the epidemic, making China's COVID-19 epidemic gradually enter the stage of calm after February 9. The following is a review of the measures taken by all sectors of society to solve the problems exposed during the anti epidemic period.

(1) In view of the shortage of medical resources, insufficient medical staff and insufficient logistics support. The state uniformly allocates anti epidemic protection resources, and caring enterprises and individuals continue to donate. After the resumption of work and production, the output of protective products has increased greatly, so that medical resources are no longer short; A total of 255 medical teams and 32572 medical and nursing personnel were sent from various provinces and cities to support the establishment and delivery of "thunder god mountain", "Huoshen mountain" and "shelter hospital" in Wuhan, resulting in the phenomenon of "beds waiting for people" in the later stage; Alibaba and other enterprises provide free meals for medical personnel, and volunteer online car Hailing has further improved the logistical support of medical personnel.

(2) In view of the overload of rescue work, high risk of infection and great physiological and psychological pressure. On February 11, the State Council issued the notice on several measures to improve the working conditions of front-line medical personnel and effectively care for their physical and mental health, which gave clear instructions on the treatment, physical and mental health, and work treatment of medical personnel, improved the working environment of front-line medical personnel in the fight against epidemic and reduced their physical and mental pressure; The Hubei provincial government has even proposed a policy of giving 10 points to the children of frontline medical workers in Hubei in this year's mid-term exam, so as to reward medical workers who fight in the frontline.

(3) Aiming at social negative events and rumor spreading. The government promptly came forward to deal with social negative events, punished those who spread rumors and incidents, and actively guided the dissemination of positive "positive energy" news. For example, the investigation team appointed by the state supervision and administration of the people's Republic of China investigated Li Wenliang's problem, promptly carried out investigation and verification on the distribution of anti epidemic donated materials by the Hubei Red Cross Society, and held its main leaders accountable and removed from office.

(4) Aiming at medical trouble and injury. Punishment shall be given to medical personnel who make trouble or injure medical personnel. However, during the epidemic, there were still many medical incidents. After Dr. Yang Wen was killed, the 15th meeting of the Standing Committee of the 13th National People's Congress on December 28, 2019 voted to adopt China's first basic law in the field of health - the law on basic medical care and health promotion. However, on January 20, there was another serious medical disturbance. Although the final perpetrators were punished, it caused a serious psychological shadow to the medical personnel at the forefront of the fight against the epidemic and constantly called for the reform of the medical

and health system. Next, this article will analyze the root cause of the "medical trouble" event from the course of China's medical and health system reform, and clarify the focus of China's medical and health system reform.

Since the reform and opening up in 1978, the reform of the medical and health system has also begun. In the early stage of reform, due to the limited financial capacity of the government, the investment in medical and health undertakings was ignored, which was manifested in the policy of giving no money in the 1980s; Medical marketization in the 1990s; The property rights reform in 2000 gradually led to the problem of "expensive and difficult medical treatment" for the people; SARS in 2003 exposed the fragility of China's public health system. In 2005, the development research center of the State Council issued a report saying that China's medical and health system reform was "generally unsuccessful"; The year 2006 set the tone of "government led" for the direction of medical reform, and became a turning year of medical reform in China; After 2007, medical reform policies were continuously introduced until the publication of the opinions of the CPC Central Committee and the State Council on deepening the reform of the medical and health system in April 2009, marking the official opening of China's "new medical reform" [42-44].

At present, China's "new medical reform" is divided into three stages.

(1) From 2009 to 2011. In April 2009, the opinions of the Central Committee of the Communist Party of China and the State Council on deepening the reform of the medical and health system were issued, which defined five goals: by 2011, the basic medical security system will fully cover urban and rural residents, the primary medical and health service system in urban and rural areas will be further improved, the basic drug system will be initially established, basic public health services will be popularized, and the pilot reform of public hospitals will make breakthroughs.

(2) From 2012 to 2016. In October 2012, the "12th Five Year Plan" and "12th Five Year Plan" for the development of health services were issued, which emphasized the key reform of public hospitals, improving the accessibility, efficiency and quality of medical services, and reducing personal medical expenses. In October 2016, the outline of the "healthy China 2030" plan was released, emphasizing the health of the masses and promoting the construction of a healthy China.

(3) 2017-2020. In January 2017, the notice of the State Council on printing and distributing the 13th five year plan for deepening the reform of the medical and health system was issued, which set the main objectives of deepening the reform of the medical and health system by 2017 and 2020. The third stage is the key period for building a moderately prosperous society in all respects by 2020. Therefore, the national medical reform has been speeding up. In 2017, the "implementation of the healthy China strategy" was written into the report of the 19th National Congress of the Communist Party of China, and the medical reform began to deepen; In 2018, the Health Commission, the medical insurance bureau, and the pharmaceutical Bureau were established to intensify the linkage reform of the three major institutions of medical insurance and medicine; In 2019, the notice of the general office of the State Council on printing and distributing the key tasks of deepening the reform of the medical and health system in 2019 was issued. Various medical reform measures were constantly refined, and the reform of the medical and health system was further deepened with the centralized procurement and use of drugs as a breakthrough.

From the above analysis, it can be seen that from the 1980s to the early 2000s, the government's low investment in the medical and health industry led to the "marketization of medical care", resulting in the problem of "expensive and difficult medical treatment" for the masses, which is the root of the continuous "medical trouble" events in recent years. After the outbreak of SARS in 2003, the central government realized the fragility of China's public health service system

and began to emphasize the characteristics of "government led" and "public welfare". 2020 is the year to achieve the goal of building a moderately prosperous society in an all-round way, and it is also a key stage to deepen the reform of the medical and health system. It is of great significance to clarify the goal of medical reform. The medical and health system reform should achieve three basic goals [42]: 1) fairness. The coverage of medical and health security is wide and the price is low. 2) Efficiency. The efficiency of medical and health investment is high, and the operation efficiency of all links of medical services is high. 3) Service quality. Meet the medical needs of the masses and improve the quality of medical services. At present, the participation rate of basic medical insurance for residents in China is stable at more than 95%, and the drug price and the burden of the masses have been continuously reduced; Hierarchical diagnosis and treatment and the establishment and improvement of the primary medical care system have basically met the medical needs of the masses; However, China has a large population, a large number of medical needs and limited medical service capacity, which leads to the reduction of doctors' treatment time, the decline of people's satisfaction with medical treatment and the decline of medical service quality.

To sum up, the medical reform over the past 30 years has basically achieved the two reform goals of fairness and efficiency, and basically solved the problem of "expensive and difficult to see a doctor". China's medical correction is at the stage of deepening the reform to further lower the price of medical services and comprehensively improve the quality of medical services. With the continuous deepening of the reform of China's medical and health system, it is expected to achieve the three goals of fairness, efficiency and high quality of service of the medical and health system in the future, and effectively solve the problem of "expensive and difficult" medical treatment for ordinary people, so as to fundamentally put an end to "medical trouble". Next, from the perspective of solving the problem of "medical trouble", this paper puts forward suggestions for the deepening reform of the medical and health system.

(1) Increase the cost of crime and form a legal deterrent.

Some medical troublemakers obtain illegal benefits by hyping up medical disputes, which brings potential personal danger to medical personnel. With the implementation of the policy of "governing the country according to law" in China, the cost of crime should be greatly increased for medical incidents. On the one hand, criminal cases are handed over to the procuratorate for prosecution and severe punishment; On the other hand, criminals are severely punished in civil matters. Only by forming legal deterrence in criminal and civil affairs can most of the "medical trouble" criminal incidents be eliminated

(2) We will continue to promote the Internet-based medical care, improve the quality of medical services, and further lower medical prices.

In recent years, the government has tried to use Internet technology to promote the reform of the medical system. 2020 is the final year of building a moderately prosperous society in all respects, and also the key node of the new medical reform. It is necessary to continue to promote the reform of medical Internet. On the one hand, Internet technology is used to further transform hospitals. Big data and historical medical records can greatly improve doctors' diagnosis and treatment efficiency, reduce people's waiting time for medical treatment, and thus improve the quality of medical services; On the other hand, after the medical Internet, the medical prescriptions will be highly unified, and the medical procurement will be carried out in a market competitive manner in the world, which will further depress the medical prices and enable the masses to pay for medical expenses. In particular, after the medical Internet, patients and their families who have medical troubles will be included in the blacklist of the national public hospital network. Most ordinary people can't afford the high service fees of private hospitals. In order to meet their medical needs and avoid being listed in the "black list", they will spontaneously put an end to medical disturbances.

(3) Stimulate the increment of medical staff and maintain their sense of honor.

Continuous medical killings will lead to a gap on the supply side of the medical industry in the future; During the fight against COVID-19, medical staff were under physical and psychological pressure. The serious medical incidents during the fight against covid-19 made many of them doubt their careers. On the one hand, for the possible supply side gap of medical staff in the future, the Ministry of education and other departments need to continue to encourage young people to engage in the medical industry, increase the supply side supply, and stimulate the increase of medical staff; On the other hand, medical and health personnel are engaged in high-pressure and busy treatment work all the year round, most of which come from their professionalism of "saving the life and healing the wounded" and their sense of industry honor. Therefore, the health commission and other departments need to advocate that the whole society respects medical and health personnel and maintains their sense of honor.

## 5. Conclusion

The main contribution of this paper is to build a rcnn-lda network public opinion analysis model based on UGC of medical staff's microblog. By using the rcnn-lda network public opinion analysis model constructed in this paper, the public opinion of the micro blog UGC of the front-line medical staff in Wuhan was analyzed, the results of the public opinion analysis were discussed and relevant suggestions were put forward. First of all, more than 6000 pieces of UGC data of Wuhan anti epidemic front-line medical personnel were collected on the Weibo platform. After text preprocessing, the emotional classification model based on RCNN was used to classify them, and positive UGC data and negative UGC data were obtained, and compared with the deep learning classification model based on CNN and Bi LSTM based on the attention mechanism. Next, LDA theme analysis was performed on UGCS with negative tendency after emotion classification, and the original negative Weibo UGC data was traced back to further determine the classification theme, and then the public opinion analysis results were output. Finally, the results of the public opinion analysis were discussed. With the efforts of all sectors of society, the problems of shortage of medical resources and shortage of medical staff in Wuhan in the early stage were solved. China has entered the stage of calming down the epidemic; In view of the serious medical disturbances that occurred during the epidemic, this paper analyzes the root causes from the reform process of China's medical and health system and puts forward relevant opinions accordingly. The analysis results show that China's medical and health system reform is in the stage of deepening the reform to further reduce the price of medical services and comprehensively improve the quality of medical services. It is expected to fundamentally eliminate medical disturbances and injuries in the future.

Due to the limitation of my research level and time, I can further study the following aspects in the future:

(1) During the period of COVID-19, the majority of medical personnel effectively disseminated novel coronavirus related knowledge, medical protection knowledge, treatment information, etc. through Weibo real name authentication. Since Weibo is the first platform for the dissemination of large-scale epidemic information, this paper uses the Weibo data from 2017 to 2018 as the training corpus of the emotional classification model. With the continuous improvement of the national epidemic situation, in the future, large-scale epidemic Weibo data can be collected as a training set to train the classifier, so as to improve the effectiveness of the emotional classification model. In addition, we can also build an emotion dictionary to identify the neutral emotions of the Weibo UGC data of Wuhan frontline anti epidemic medical personnel, and then use the emotion classification model built in this paper to classify the positive and negative emotional tendencies, so as to improve the accuracy and precision of emotion classification in the online public opinion analysis model.



(2) During the period of COVID-19, the government implemented an open and transparent information disclosure mechanism, and the effective publicity of epidemic information reduced the panic of the masses. Wechat, Tiktok, BiliBili and other social media platforms, with their interactive communication mode, have also played an important role in helping the public understand the epidemic information and reduce panic. In the future, the epidemic data of wechat, Tiktok, BiliBili and other platforms can be collected to study the user behavior of different platforms under the COVID-19.

## References

- [1] Mahase E. Coronavirus: covid-19 has killed more people than SARS and MERS combined, despite lower case fatality rate[J]. *British Medical Journal*, 2020, 368: m641.
- [2] Song Quanxi, looking for his front, Ning Yina Research and development of Internet public opinion -- a case study of police related Internet public opinion [J]. *Operations research and management*, 2018, 27(4): 93-104.
- [3] Wang Lianxi Analysis on the distribution and relationship of related concepts in the field of Internet public opinion [J] *Modern intelligence*, 2019, 39(6): 132-141.
- [4] Kim Y. Convolutional neural networks for sentence classification[J]. *Eprint ArXiv*, 2014.
- [5] Zhou P, Shi W, Tian J, et al. Attention-based bidirectional long short-term memory networks for relation classification[C]. *Proceedings of the 54th annual meeting of the association for computational linguistics (volume 2: Short papers)*, 2016: 207-212.
- [6] Zhou Chenghu, Pei Tao, Du Yunyan, et al. COVID-19 big data analysis and regional prevention and control policy recommendations [J] *Journal of the Chinese Academy of Sciences*, 2020, 35(2): 200-203.
- [7] Feng Xujie, Wang Yang, Liu Shuhao, et al. The impact of COVID-19 on the operation of urban rail transit [J] *Transportation Research*, 2020, 6(01): 45-49.
- [8] Li Yuelin, Wang Shanshan Analysis on the characteristics of government information release during the COVID-19 epidemic [J] *Books and information*, 2020, 40(01): 27-34.
- [9] Chen Xikang, Yang Cuihong, Bao Qin, et al. Analysis of the impact of the COVID-19 epidemic on China's economy and countermeasures [J] *Journal of the Chinese Academy of Sciences*, 2020.
- [10] Dong Hongzhe, Zheng Yongfang Fire prevention and extinguishing: generation and management of secondary public opinion in COVID-19 public crisis [J] *Journal of University of Electronic Science and Technology (SOCIAL SCIENCE EDITION)*, 2020, 22(2): 1-7.
- [11] Wang Yi, Gao Junling, Chen Hao, et al. Public media exposure during the COVID-19 epidemic and its relationship with mental health [J] *Journal of Fudan University (Medical Edition)*, 2020, 47(2): 173-178.
- [12] Xu Z, Shi L, Wang Y, et al. Pathological findings of COVID-19 associated with acute respiratory distress syndrome[J]. *The Lancet respiratory medicine*, 2020, 8(4): 420-422.
- [13] Chen H, Guo J, Wang C, et al. Clinical characteristics and intrauterine vertical transmission potential of COVID-19 infection in nine pregnant women: a retrospective review of medical records[J]. *The Lancet*, 2020, 395(10226): 809-815.
- [14] Liu Y, Gayle A A, Wilder-Smith A, et al. The reproductive number of COVID-19 is higher compared to SARS coronavirus[J]. *Journal of Travel Medicine*, 2020, 27(2).
- [15] Wu Z, Mcgoogan J M. Characteristics of and Important Lessons From the Coronavirus Disease 2019 (COVID-19) Outbreak in China: Summary of a Report of 72314 Cases From the Chinese Center for Disease Control and Prevention[J]. *The Journal of the American Medical Association*, 2020, 323(13): 1239-1242.
- [16] Pan F, Ye T, Sun P, et al. Time Course of Lung Changes On Chest CT During Recovery From 2019 Novel Coronavirus (COVID-19) Pneumonia[J]. *Radiology*, 2020, 0(0): 200370.
- [17] Dong E, Du H, Gardner L. An interactive web-based dashboard to track COVID-19 in real time[J]. *The Lancet Infectious Diseases*, 2020, 0(0).

- [18] Ye Jian, Zhao Hui Public opinion analysis model based on large-scale bullet screen data monitoring and emotion classification [J] Journal of East China Normal University (NATURAL SCIENCE EDITION), 2019, 3(2019): 86.
- [19] Hatzivassiloglou V, Mckeown K R. Predicting the semantic orientation of adjectives[C]. Proceedings of the 35th annual meeting of the association for computational linguistics and eighth conference of the european chapter of the association for computational linguistics, 1997: 174-181.
- [20] Shen Y, Li S, Zheng L, et al. Emotion mining research on micro-blog[C]. 2009 1st IEEE symposium on web society, 2009: 71-75.
- [21] Hong Qing, Wang siyao, Zhao qinpei, et al. Video user group classification based on bullet screen emotion analysis and clustering algorithm [J] Computer engineering and Science, 2018, 40(6): 1125-1139.
- [22] Wang Ke, Xia Rui Overview of automatic construction methods of emotion DICTIONARY [J] Journal of automation, 2016, 42(4): 495-511.
- [23] Zhou shengchen, Qu Wenting, Shi yingzi, et al. A review of research on emotion analysis of Chinese microblogs [J] Computer application and software, 2013, 030(003): 161-164,181.
- [24] Pak A, Paroubek P. Twitter as a corpus for sentiment analysis and opinion mining[C]. LREc, 2010: 1320-1326.
- [25] Liu Zhiming, Liu Lu Empirical research on emotion classification of Chinese microblog based on machine learning [J] Computer engineering and Application, 2012, 48(01): 1-4.
- [26] Cai Huiping, Wang Lidan, Duan Shukai Emotion classification model based on word embedding and CNN [J] Computer application research, 2016, 33(10): 2902-2905.
- [27] Lai S, Xu L, Liu K, et al. Recurrent convolutional neural networks for text classification[C]. Twenty-ninth AAAI conference on artificial intelligence, 2015.
- [28] Blei D M, Ng A Y, Jordan M I. Latent dirichlet allocation[J]. Journal of machine Learning research, 2003, 3: 993-1022.
- [29] Krestel R, Fankhauser P, Nejdl W. Latent dirichlet allocation for tag recommendation[C]. Proceedings of the third ACM conference on Recommender systems, 2009: 61-68.
- [30] Shi Jing, fan Meng, Li Wanlong Theme analysis based on LDA model [J] Journal of automation, 2009, 35(12): 1586-1592.
- [31] Guo Y, Barnes S J, Jia Q. Mining meaning from online ratings and reviews: Tourist satisfaction analysis using latent dirichlet allocation[J]. Tourism Management, 2017, 59: 467-483.
- [32] Jelodar H, Wang Y, Yuan C, et al. Latent Dirichlet Allocation (LDA) and Topic modeling: models, applications, a survey[J]. Multimedia Tools and Applications, 2019, 78(11): 15169-15211.
- [33] Urban J, Bulkow K. Tracing Public Opinion Online-An Example of Use for Social Network Analysis in Communication Research[J]. Procedia-Social and Behavioral Sciences, 2013, 100: 108-126.
- [34] Wang Lancheng, Chen Lifu Summary of domestic and foreign research on the evolution, early warning and response theory of Internet public opinion [J] Library magazine, 2018, 37(12): 4-13.
- [35] LAN Yuexin Research on the diffusion law model of microblog public opinion in emergencies [J] Information Science, 2013, 31(03): 31-34.
- [36] Zheng Changxing, Su Xinning, Liu Xiwen Construction of network public opinion analysis model for emergencies -- Based on the perspective of stakeholders [J] Intelligence magazine, 2015, 34(04): 71-75.
- [37] An Lu, Wu Lin Analysis on the evolution of public opinion of emergency microblog integrating theme and emotional characteristics [J] Library and information work, 2017, 61(15): 120-129.
- [38] Wang Xiufang, Sheng Shu, Lu Yan A Weibo public opinion analysis model based on topic clustering and emotion intensity [J] Modern library and information technology, 2018, 002(006): 37-47.
- [39] Shi Yaolin, Cheng Huihong, Huang Luyuan, et al. Study on epidemiological dynamics of COVID-19 with discrete stochastic model [J] Journal of the Chinese Academy of Sciences, 2020, 37(2): 145-154.

- [40] Yu Zheng Research and application of text Vectorization Based on deep learning [D] Shanghai: East China Normal University, 2016.
- [41] Mikolov T, Chen K, Corrado G, et al. Efficient estimation of word representations in vector space[C]. ICLR (Workshop Poster), 2013.
- [42] Li Ling, Jiang Yu, Chen Qiulin 30 years of China's medical reform under the background of reform and opening up [J] China's health economy, 2008, 27(2): 5-9.
- [43] Li Ling, Jiang Yu 2006: the turning point of China's Medical Reform [J] China's health economy, 2007, 26(4): 5-9.
- [44] Li Ling Review on the progress of new medical reform [J] China's health economy, 2012, 31(01): 5-9.