J. COMBIN. MATH. COMBIN. COMPUT. 127b (2025) 9515--9535

Journal of Combinatorial Mathematics and Combinatorial Computing
www.combinatorialpress.com/jcmcc

Combinatorial Press

Modeling and Visualization of Public Opinion Sentiment Evolution after Sichuan Luding MS6.8 Earthquake Based on LSTM Neural Networks

Guoyong Pan[1,2], Ye Ren[1,2], Haiying Yu[1,2], Xiuqing Song[1,2,✉]

[1] *Shanghai Earthquake Agency, Shanghai, 200062, China*

[2] *Shanghai Sheshan National Geophysical Observatory, Shanghai, 200062, China*

ABSTRACT

The article uses web crawling to obtain public opinion data after the Sichuan Luding MS6.8 earthquake and preprocesses this data. Aiming at the limitations of the traditional LDA topic model, an improved topic model based on LDA, TT-LDA, is proposed. the BERT model is used to encode the public opinion data, and on the basis of the BERT embedding, the BiLSTM model is used for contextualized word representation for deep feature extraction to complete the modeling of public opinion sentiment evolution. Combining the crawled data and the model, we analyze the public opinion after the Sichuan Luding MS6.8 earthquake. Three days after the earthquake, positive sentiment, neutral sentiment, and negative sentiment increase to 488498, 466832, and 516560, respectively, a total of 1471890 sentiment data, and after time evolution, the sentiment polarity intensity increases from -0.178 to -0.886, indicating that when the official announcement of the number of casualties of the accident is made, the netizens' negative sentiment fully erupts to show the post-earthquake public opinion sentiment evolution process.

Keywords: BiLSTM model, TT-LDA model, public opinion sentiment evolution, Luding MS6.8 earthquake

1. Introduction

Earthquake is one of the most serious natural disasters, which not only threatens people's life and property safety, but also causes extensive social concern and public opinion. The timely control of earthquake online public opinion has an important impact on the rescue of the disaster area, post-disaster reconstruction and social stability [1-3]. After an earthquake occurs, both online and offline communication channels will have an instantaneous outbreak of information in a short time, including

✉ Corresponding author.
E-mail address: dzzysdyj@126.com (X. Song).
Received 12 January 2024; Revised 05 March 2024; Accepted 25 December 2024; Published Online 16 April 2025.
DOI: 10.61091/jcmcc127b-521

personal disaster situation, rescue progress, and early warnings from professional organizations, etc. This information spreads at a very fast speed on the Internet and social media and other channels, coupled with active participation of a large number of users and retweeting, the information spreads even faster, and it is mostly unverified information or rumor-mongering information that causes panic [4-5]. As the main position of earthquake public opinion dissemination, online media need to review and disseminate information in a comprehensive, timely and accurate way [6]. However, after an earthquake, due to its threat to people's lives and properties, the public tends to be highly agitated and prone to various emotional reactions, such as panic, alarm, anxiety, etc., which may especially trigger some irrational remarks and behaviors. Therefore, there is a need to effectively control public opinion after an earthquake.

According to the authenticity and importance of the information, it is reasonable and timely to guide netizens to express their views rationally by means of official release of information and public opinion guidance to avoid the spread of rumors and social panic, as well as to guide the public to carry out self-rescue and mutual rescue and post-disaster reconstruction [7-9]. Among them, understanding the emotional change state of public opinion after an earthquake is a key point to control public opinion. The evolution of public opinion sentiment after an earthquake concerns the hearts and minds of people and social stability, and the work of the relevant emergency online public opinion supervisory departments after a disaster is particularly important. Through the analysis of various network data, the direction of public sentiment can be obtained, and public opinion can be appropriately guided to alleviate the public's panic about unknown and untrue information; therefore, there are also some studies done on the assessment and control of public opinion risk [10-13]. And Long Short-Term Memory (LSTM) network is a kind of recurrent neural network used for processing and predicting sequential data, which is widely used in the fields of natural language processing, speech recognition, and video analytics [14-16]. In addition, LSTM has been reported studies on sentiment analysis and also applied on earthquake related studies [17-18].

After an earthquake, initial public opinion control is a critical period to prevent public panic. Some studies have shown that the government's public opinion control in the early stage of a disaster or public event is effective and low-cost, and the cooperation between the media and the government can facilitate public opinion management [19]. Social media platforms are the main venues for the fermentation of public opinion after earthquakes, and a number of researchers have analyzed the post-disaster public opinion user sentiment on online social media. Literature [20] utilized online computational methods to extract post-disaster information from social media and public news, detected emotional vocabulary among them, and utilized clustering methods to detect the public's emotional responses during the disaster. Literature [21] utilized the reported intensities in the Modified Mercalli Intensity Scale for post-earthquake sentiment and thematic analysis, and the structure of the sentiment analysis showed that the lowest reported intensities related to user statements showed positivity, while higher intensities showed negativity. Literature [22] analyzed the public mood after earthquake based on the triad of Negative Bias Theory, Integrated Crisis Mapping Model, and Arousal Theory, and the results showed that the public statements of anger mood related words increased after the earthquake and the number of retweets increased, whereas the retweets of anxiety related statements of the influential users decreased but the retweets of sadness related statements increased. Literature [23] designed a public opinion feature extraction algorithm for intelligent analysis of disaster public opinion, which first performs sentiment perception analysis of public opinion samples, and then extracts features for recognition to collect emotional information and get the public's demands. Among them, there is a kind of sentiment analysis applied quite a lot, which has supervised, unsupervised and hybrid methods, all of which can be used as sentiment vocabulary correlation analysis [24]. Literature [25] analyzed the trend of public opinion after the typhoon by using social network analysis and sentiment analysis, and the results showed that the trend of public opinion evolution is “positive-negative-positive”. In [26], the evolutionary mechanism of public emergencies in social networks was investigated using social network analysis and sentiment mining analysis. In addition, literature [27] utilized a pre-trained sentiment analysis model to identify text data in social media, but the accuracy was only 63%. In contrast, literature [28] constructed an opinion sentiment trend analysis model combining XLNet, BiGRU, and attention mechanism, with an accuracy of up to 90%. Literature [29] used deep learning algorithms to extract public sentiment data in social media and analyzed it in combination with geographic information data to obtain sentiment dynamic patterns and optimize disaster reduction strategies. Literature [30] utilized the E-Trans sentiment analysis model to analyze the sentiment tendency of the content posted by microblogging users during the three earthquakes in Sichuan, China in 2022, and the results showed that there are various factors influencing the sentiment evolution of public opinion after earthquakes, which are related to the magnitude of earthquakes, the number of aftershocks, the propaganda of earthquake-related departments, and people's concern issues. In addition, the emergence of earthquake browser provides earthquake visualization visual display for the people to query the magnitude, number, location, duration and other information in the earthquake event [31]. The visualization of earthquake information can eliminate the concerns and anxiety in the minds of the public and correctly treat the emotions of earthquake events. All these studies provide references for public opinion guidance and post-disaster emergency response strategies for earthquake disaster prevention and mitigation.

In this paper, web crawler technology or local direct import and other ways are selected to obtain the data text, transform the cluttered unstructured data into structured data, and carry out the operations of de-weighting, word splitting, de-deactivating words and other operations to realize the main work of data pre-processing. In order to cope with the possible existence of null values in the screened text, the LDA model constructed by the old dictionary is used as an auxiliary model, and the source data corresponding to the null-valued data is used for topic generation, which leads to the construction of a kind of improved topic model TT-LDA based on LDA.The feature representations based on BiLSTM and sentence graph are used as inputs for the convolutional operation, which finally completes the task of postearthquake public opinion sentiment evolution modeling, and the model is used as an auxiliary model. The loss function of the model is also determined. Combined with the research data, the post-earthquake public opinion after the Sichuan Luding MS6.8 earthquake is exemplarily analyzed.

2. Study on public opinion after the MS6.8 earthquake in Sichuan Luding

2.1. Data Acquisition and Processing

Data acquisition generally uses web crawler technology or local direct import and other ways to obtain data text, data preprocessing is mainly to transform the cluttered unstructured data into structured data, and carry out operations such as de-emphasis, lexicalization, and de-discontinuation of words.

2.1.1. Data acquisition. This experiment mainly uses web crawler technology [32]. Based on the framework of requests, we combine the three elements of earthquake (time, location, magnitude) and the keywords of earthquake sensation, such as “sensation”, “shaking” and “shake”, as well as the starting and ending times. “shaking” and other factors such as the starting and ending time, we obtain post-earthquake public opinion data from mainstream social media websites such as Sina Weibo, Sina News, Netease News, Sohu News, Shake, Shutterbugs, etc., over multiple time periods, and obtain the data of 11 destructive earthquakes from 2022 to the present with a total of 419,093 textual comments, including the user name, the content of the post, posting time, posting location, etc.

2.1.2. Data pre-processing. The data we collect through web crawlers and other technologies contain some unstructured, useless and repetitive information, such as blank information and missing information, which may have a certain impact on the training of the model, so we need to transform these data into structured data, which additionally contains the steps of word splitting, deactivation and labeling. Segmentation is the basis of data preprocessing, mainly used to cut a continuous text sequence into meaningful words or phrases, the main goal of the segmentation technique is to solve the ambiguity problem in the text, to ensure that each word or phrase can be recognized correctly. English participles are relatively easy, because they can use space as a separator, but Chinese participles are relatively difficult, and the common participles are Padding participles, LAC participles, THULAC participles, LTP participles, Stanford NLP participles, jieba participles, etc. At present, jieba participle is the most widely used, so this paper chooses jieba participle tool for participle operation.

Removing stop words is an important step in data preprocessing, which is mainly used to remove some common words that contribute little or no to the meaning of the text from the text, usually including virtual words, prepositions, particles, modal words, etc., such as "of", "has", "bar", "what", etc. The main purpose is to reduce the dimension of text vectors, reduce the sparsity of data, make text features more prominent, reduce the computational cost of processing data, and make it easier for the model to learn and understand the internal patterns of data, which helps to improve the efficiency and accuracy of subsequent text processing tasks. We generally use a stop word list to stop words, compare each word in the corpus with a stop word list, keep the same if it is different, and remove the same word.

Annotation is also one of the important steps in data preprocessing, which is mainly used to assign a label or tag to each element in the raw text data for subsequent analysis and model training. The main purpose of annotation is to help the machine understand the structure and meaning of the text, providing important information for subsequent text analysis and processing. Annotation can be lexical annotation and can also include named entity recognition (NER), which is used in this paper for extracting geolocation information, and the commonly used annotation methods in this field are sequence annotation methods BIO, BIOES, and BMES. where B (Begin) denotes the location of the beginning of the entity, I (Inside) denotes the location of the inside of the entity, M (Middle) denotes the position of the middle of the entity, O (Outside) denotes that it does not belong to any type of entity, and E (End) denotes the position of the end of the entity, which is generally annotated by using entity labels plus types, and in this paper, we adopt the annotation method of BIO.

2.2. Theme Mining and Visualization of Internet Public Opinion Information

2.2.1. Theme mining tasks. The article is narrated with the theme as the clue, and the theme contains the central idea of the article. As a special kind of article, the word limit of microblog text makes the theme information contained in microblog text more obvious, which can better reflect the real intention of netizens. Traditional theme mining focuses on manual detection, which requires human resources to read through and analyze the article to get the theme of the article. This kind of theme mining requires too much labor cost, and everyone has a different understanding of the article, resulting in a lack of objectivity in the final theme, which is not convincing. In the field of text mining, the theme model can statistically analyze the words in the article, explore the frequency of specific words in the article, and reasonably cluster the words to extract the corresponding theme of the article.

2.2.2. Improved LDA-based topic model TT-LDA. As the most commonly used topic model, LDA is favored by researchers, but LDA is not perfect, and it has some design flaws. In the process of LDA topic model construction, the model will first encode all the words in the corpus to form a dictionary [33-34]. Then, according to the dictionary, all the data are converted into a sparse vector set for topic generation. However, such a way of constructing a lexicon increases the size of the lexicon significantly if the number of corpus words is large, which affects the effectiveness of the model topic generation. To address this problem, an improved topic model TT-LDA based on LDA is proposed, and the specific flow of topic generation based on TT-LDA is shown in Fig. 1. This paper combines TF-IDF and TextRank algorithms to improve the lexicon of the LDA model and increase the model's sensitivity to high-importance words, with a view to improving the effect of model topic generation. First, the TF-IDF and

TextRank algorithms are utilized to rank the importance of all words in the corpus respectively. Then, the size of the lexicon is set to 17957 with reference to the idea of word list design of ERNIE model.Next, the words in the top 17957 of the importance ranking of the two algorithms are screened for the same words by the iterative matching algorithm, respectively. Finally, the set of words remaining from the screening is used as a new lexicon of the LDA model for topic generation. In order to cope with the possible existence of null values in the filtered text, the LDA model constructed from the old dictionary is used as an auxiliary model for topic generation using the source data corresponding to the null value data.

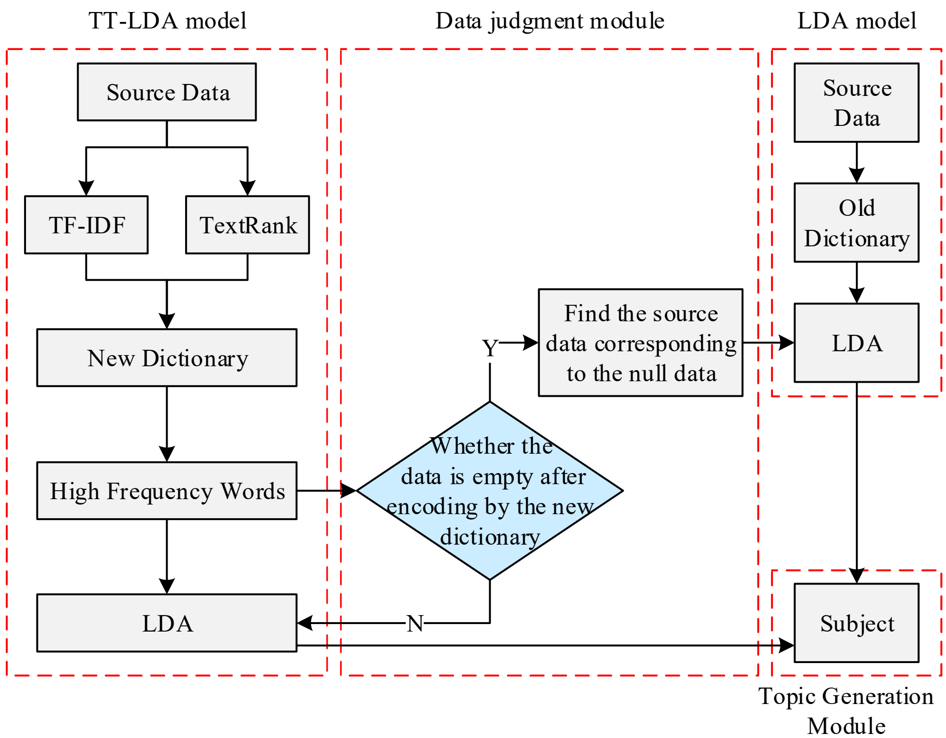

Fig. 1. The specific process based on TT-LDA topic generation

In order to validate the performance of the TT-LDA topic model, the public dataset of Sichuan Luding microblog comments is used as the corpus, and the performance of the topic model is judged by comparing the perplexity and consistency of the four models, namely LDA, TF-IDF-LDA, TextRank-LDA and TT-LDA.

There are two main metrics to measure the performance of topic models: perplexity and consistency. The lower the perplexity and the higher the consistency, the better the topic model is.

1) Confusion degree is often used to measure the degree of a probability distribution or probability model to predict the merits of the sample, can be used to regulate the number of topics, the smaller the degree of confusion, on behalf of the topic model to generate the topic of the ability of the better. Its calculation formula is:

$$Perplexity(D) = \exp \frac{\sum_{d=1}^{M} \log P\left(w_{d_1}\right)}{\sum_{d=1}^{M} N_{d_i}} \tag{1}$$

Where D represents the document, d_i represents the word in document d, $P(W)$ represents the probability distribution of the word in the document, M represents the number of documents, and N represents the number of times d_i appears in the document.

2) Consistency represents the coherence of the words in each topic generated by the topic model, the higher the consistency score, the more coherent the words corresponding to the generated topic. The formula for topic consistency is shown below:

$$Coherence\left(t;V^{(1)}\right)=\sum_{m=2}^{M}\sum_{m=1}^{m-1}\log\frac{D\left(v_m^{(1)},v_1^{(1)}\right)+1}{D\left(v_1^{(1)}\right)} \tag{2}$$

t represents the current topic, $V^{(t)}$ represents the M words most likely to belong to topic t, $D\left(v_m^{(t)},v_l^{(t)}\right)$ is the frequency of co-occurrence of two words v_m and v_l, and $D\left(v_l^{(t)}\right)$ is the frequency of occurrence of word v_l.

2.2.3. Thematic visualization. Theme mining of text data in the online public opinion credibility feature system based on TT-LDA theme model. In order to get the theme information of all text data, the publisher text content, commenter 1 comment text, commenter 2 comment text, and commenter 3 comment text are integrated together, and the integrated text data are fed into the model for training. Among them, measuring the training effect of the model is mainly determined by the confusion degree and consistency index together, the lower the confusion degree and the higher the consistency, the better the model training effect. Theme visualization is based on data visualization technology to visualize and display the theme information in text-based feature data, making the theme information more clear, concise and easy to observe, based on matplotlib visualization technology for theme visualization.

2.3. Opinion Sentiment Model

2.3.1. Constructing word vectors based on BERT. BERT is a Transformer-based bi-directional encoder representation that pre-trains deep bi-directional representations by jointly regulating the left and right contexts in all layers [35-36]. It is a multi-task model that contains two self-supervised tasks, Masked Language Modeling (MLM) and Next Sentence Prediction (NSP). In the MLM task, the model randomly masks some of the input words and then predicts the masked words. In the NSP task, the model determines whether sentence B is the following of sentence A and outputs 'IsNext' if it is and 'NotNext' otherwise.

The word embedding layer is to map each word to a high-dimensional vector space. In this paper, the text of online public opinion events is trained with the help of BERT model to generate word vectors with semantic information. The comment text of the “Sichuan Luding MS6.8 Earthquake” event on Weibo is used as the dataset, and it is preprocessed. Text t consists of m sentences, i.e., $t=\{s1,s2,\cdots,sm\}$, each sentence consists of n words, and the input sequence s is $s=\{w1,w2,\cdots,wn\}$. For a given sentence $s=\{w1,w2,\cdots,wn\}$, the input is formulated as a sequence $s=\{[CLS],w1,\cdots,wn,[SEP]\}$, with the $[CLS]$ flag placed at the top of the sentence, and the $[SEP]$ flag used to separate the two input sentences. The vector representation of the words is obtained after pre-training by BERT $X=\{x1,x2,\cdots,xn\}$, $xi\in R^{demb\times 1}$, d_{cmb} are the dimensions of the word vectors.

2.3.2. Creating Contextual Word Representations Based on BiLSTM. LSTM is a variant of RNN for the vanishing gradient problem.BiLSTM consists of forward and reverse LSTMs that can learn contextual information and potential meanings of words by reading input sentences in both directions. One LSTM processes sequences from left to right and the other processes sequences from right to left. Experiments have demonstrated that this neural network structural model outperforms a single LSTM structural model in terms of efficiency and performance for text feature extraction. Contextualized word representation converts each word in a sentence into a vector such that each vector is aggregated from the entire vector of the input sentence.

In this paper, BiLSTM model is used for contextualized word representation on BERT embedding for deep feature extraction. Based on the BERT model, the sequence $X=\{x1,x2,\cdots,xn\}$ is used as the input layer of BiLSTM, and the forward vector $\vec{h}_i(i=[1,n])$ and the reverse vector $\overleftarrow{h}_i(i=[n,1])$ can be obtained, and the computational rules are described as follows:

$$\overrightarrow{h_t}=\overrightarrow{LSTM}\left(x_i,\overrightarrow{h_{t-1}}\right) \tag{3}$$

$$\overleftarrow{h_l}=\overleftarrow{LSTM}\left(x_i,\overleftarrow{h_{t+1}}\right) \tag{4}$$

$$h_i=\left[\overrightarrow{h_l},\overleftarrow{h_l}\right]\in R^{2d_{hid}\times 1} \tag{5}$$

Where, x_i is the input of the current unit, $\overrightarrow{h_{t-1}}$ represents the output of the previous unit, $\overleftarrow{h_{t+1}}$ represents the output of the next unit, h_i represents the forward and backward connections of the current unit and d_{hid} is the number of hidden units. Here, BiLSTM is utilized for feature extraction of sentences and $H=\left(h_1,h_2,\cdots,h_n\right)\in R^{2d_{hid}\times n}$ is calculated.

2.3.3. Dependent syntactic information. Sentence-level syntactic and semantic analysis is an important topic in Natural Language Processing (NLP), aiming at revealing the internal structural relationships in sentences, with the basic task of expressing how words are combined to form sentences and the rules that control sentence formation. Currently, dependent syntactic analysis has been used in many ways, e.g., sentiment analysis, machine translation, opinion mining, and information extraction.

2.3.4. Constructing Graph Convolutional Networks. GCN is a commonly used method for processing data containing rich relationships and interdependencies. It efficiently captures graph dependencies through message passing between graph nodes.The input to GCN consists of feature vectors of nodes and the structure of the graph. For each node, GCN takes the relevant information about its neighborhood as a new feature representation vector.

The graph convolution layer is responsible for performing convolution operations using BiLSTM and sentence graph based feature representations as input. The vector H obtained in conjunction with section 2.3.2 is the node feature, which is input into layer GCN and subsequently the output is computed by the following equation. i.e:

$$H^{(l)}=\sigma\left(\tilde{D}^{-\frac{1}{2}}A\tilde{D}^{-\frac{1}{2}}H^{(l-1)}W^{(l-1)}\right) \tag{6}$$

where σ is the activation function, in this case Relu is used, $\tilde{D}$ is the degree matrix of A, $H^{(l-1)}$ is the node features of layer $l-1$, and $W^{(l-1)}$ is the weight matrix learned through network training. To improve the efficiency of the neural network model, the output of the GCN is next average pooled. The feature representation $H^{(l)}$ after graph convolution is average pooled to obtain H^*, which effectively incorporates the semantic as well as syntactic structure of the comment text as an informative feature.

2.3.5. Sentiment Classification and Loss Functions. The sentiment classification layer is mainly based on the output of the graph convolution layer to judge the text sentiment tendency. Through the graph convolution layer obtained node output H^* to the fully connected layer, using the activation function for softmax, combined with the prediction formula (7), to obtain the sentiment label prediction results $\hat{y}$. i.e:

$$\hat{y}=\arg\max\left(soft\max\left(W_E H^*+b\right)\right) \tag{7}$$

where W_E is the weight matrix of the layer and b is the bias term of the layer.

Finally, the cross entropy is chosen as the loss function as follows in equation (8). Meanwhile, regularization coefficients are added in training to achieve regularization constraints thus avoiding overfitting. Namely:

$$L=-\sum_{i=1}^{l} y_i \log\left(\hat{y}_i\right)+\lambda\|\theta\|^2 \tag{8}$$

where l is the number of sentiment classifications in the training set, y_i is the true sentiment distribution for class i, $\hat{y}_i$ is the predicted sentiment distribution for class i, λ denotes the regularization coefficient, and θ is the regularization parameter.

3. Analysis of public opinion after the MS6.8 earthquake in Sichuan Luding

3.1. Visualization and Analysis of Opinion Evolution

3.1.1. Overview of the earthquake situation. According to the official determination of the China Earthquake Network, a MS6.8 earthquake occurred in Luding County, Ganzi Prefecture, Sichuan Province (29.59°N, 102.08°E) at 12:52 on 2022-09-05, with a depth of 16km.After the earthquake, the China Earthquake Administration (CEA) quickly organized emergency response in accordance with the plan to continuously carry out seismic monitoring and tracking as well as trend analysis and judgment, and quickly dispatched an on-site team of more than 160 people from relevant provincial earthquake departments and operational units to the quake area to carry out emergency response work. The China Earthquake Administration (CEA) quickly organized and carried out emergency response work in accordance with the plan, continuously carried out monitoring and tracking of the earthquake situation, analyzed and assessed the trend, and quickly dispatched an on-site task force consisting of relevant provincial seismological departments and operational units, totaling more than 160 people, to the quake area to carry out emergency response. According to the on-site investigation, the earthquake triggered many and extremely serious geologic disasters, resulting in a large number of casualties, damage to housing buildings and infrastructure, and many disruptions to roads, communications, water supply and power supply and other lifelines. According to the seismic intensity map released by the Ministry of Emergency Management on 2022-09-11, the highest intensity of the Sichuan Luding earthquake was IX degree (9 degrees), and the long axis of the isoseismic line was oriented to the north and west, with the long axis of 195km and the short axis of 112km, and the area of the VI degree (6 degrees) zone and above was 19,089km2, which involves a total of 12 counties (municipalities and districts) in 3 cities and states and 82 towns and townships (streets) of Sichuan Province. As of 2022-09-11T17:00, the quake has killed 93 people, including 55 in Ganzi Prefecture and 38 in Ya'an City. MS6.8 earthquake is a strong quake, and generally quakes of magnitude 5 or higher will cause some damage. Because of the large magnitude of this earthquake, it was felt in a wide range of areas and by a large number of people. According to statistics, the earthquake was felt mainly in Sichuan Province, with tremors also reported in Chongqing, Yunnan, Shaanxi and Guizhou Provinces.

3.1.2. Analysis of the Number and Trend of Microblog Public Opinion Releases after the Earthquake. Figure 2 shows the microblog posting trend of the MS6.8 earthquake event in Luding County, Ganzi Prefecture, Sichuan within 3d after the earthquake, within 3d after the earthquake, 1259811 microblogging information related to this destructive earthquake event was collected through Sina Weibo, the number of retweets was 933214, the number of comments was 676109, the number of likes was 14131261837, and the exposure was 5326948413161 articles. As a sudden public security incident, this MS6.8 earthquake event in Luding County, Ganzi Prefecture, Sichuan Province, attracted widespread attention from the society.

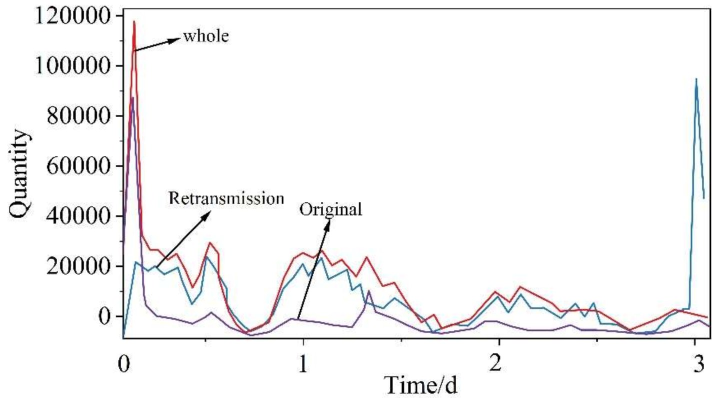

Fig. 2. Weibo released the trend chart

The changes in the number of information released within 72h after the earthquake are shown in Table 1, and the average dissemination rate of microblog opinion information for the whole earthquake event is 4138.239 items/h. The dissemination peaked at 45808 items at 2022-09-05-13:00. At 2022-09-05-14:00, about 3h after the earthquake, the number of original information released on microblog opinion showed a rapid decline, but still maintained a high level of social concern in terms of the overall number of releases and trends.

Table 1. Changes in the number of information releases within 72h after the earthquake

Time/h	Quantity	Time/ h	Quantity	Time/ h	Quantity
2022/9/5-12:00	20325	2022/9/6-12:00	3269	2022/9/7-12:00	2563
2022/9/5-13:00	45808	2022/9/6-13:00	3227	2022/9/7-13:00	2570
2022/9/5-14:00	19599	2022/9/6-14:00	3140	2022/9/7-14:00	2586
2022/9/5-15:00	5555	2022/9/6-15:00	3052	2022/9/7-15:00	2603
2022/9/5-16:00	4766	2022/9/6-16:00	2965	2022/9/7-16:00	2619
2022/9/5-17:00	4293	2022/9/6-17:00	2900	2022/9/7-17:00	2635
2022/9/5-18:00	4150	2022/9/6-18:00	2900	2022/9/7-18:00	2652
2022/9/5-19:00	4032	2022/9/6-19:00	2900	2022/9/7-19:00	2668
2022/9/5-20:00	4148	2022/9/6-20:00	2900	2022/9/7-20:00	2685
2022/9/5-21:00	4264	2022/9/6-21:00	2900	2022/9/7-21:00	2701
2022/9/5-22:00	5040	2022/9/6-22:00	2838	2022/9/7-22:00	2718
2022/9/5-23:00	4469	2022/9/6-23:00	2718	2022/9/7-23:00	2735
2022/9/6-00:00	3817	2022/9/7-00:00	2598	2022/9/8-00:00	2751
2022/9/6-01:00	3475	2022/9/7-01:00	4464	2022/9/8-01:00	2768
2022/9/6-02:00	3133	2022/9/7-02:00	3788	2022/9/8-02:00	2778
2022/9/6-03:00	2807	2022/9/7-03:00	2787	2022/9/7-03:00	2778
2022/9/6-04:00	2483	2022/9/7-04:00	2669	2022/9/8-04:00	2778
2022/9/6-05:00	2848	2022/9/7-05:00	2656	2022/9/8-05:00	2778
2022/9/6-06:00	3327	2022/9/7-06:00	2643	2022/9/8-06:00	2778
2022/9/6-07:00	3542	2022/9/7-07:00	2629	2022/9/8-07:00	2778
2022/9/6-08:00	3457	2022/9/7-08:00	2616	2022/9/8-08:00	2778
2022/9/6-09:00	3371	2022/9/7-09:00	2603	2022/9/8-09:00	2778
2022/9/6-10:00	3334	2022/9/7-10:00	2590	2022/9/8-10:00	2778
2022/9/6-11:00	3301	2022/9/7-11:00	2577	2022/9/8-11:00	2778

3.1.3. Statistical Analysis of Hot Word Frequency of Microblog Public Opinion after the Earthquake. The collected microblog opinion information is text-cleaned to filter out retweeted microblog data without any information. Information related to reflecting the earthquake situation was extracted by artificial intelligence natural language processing technology, and hot word clouds were generated by using a word division tool. The top 100 hot words in frequency of occurrence are selected to visualize the disaster information in the public opinion on the MS6.8 earthquake in Luding County, Ganzi Prefecture, Sichuan in terms of space and size, which can intuitively reflect the first impression of the public opinion information on this earthquake event (see Figure 3).

Fig. 3. Earthquake situation extraction hot word cloud

3.1.4. Thematic evolution analysis. Figure 4 shows the analysis of the "Earthquake Area", Figure 5 shows the analysis results of the "Emergency Rescue Situation", and the topics of "Earthquake Area" and "Emergency Rescue". At the time of the Luding earthquake in Sichuan, the official media updated the situation of the earthquake area in real time, focusing on the words "Luding", "Ganzi Prefecture", "epicenter", "east longitude" and "measurement", and released information such as the earthquake time, the scope of the earthquake, the magnitude, and the latitude and longitude of the earthquake area. Netizens focused on the issue of emergency rescue, and the people in the earthquake area or nearby areas expressed that "the earthquake was felt", "strong" or "obvious", or expressed condolences to the firefighting forces that went to the earthquake area to rescue, which was reflected in words such as "rescue", "firefighting", and "hard work". Topics 1 and 2 have a high degree of aggregation, and the theme of "Earthquake Area Situation" cooperates with "Emergency Rescue" to form a relatively complete information chain of earthquake area. Due to the high degree of damage and the greater impact of earthquakes, netizens are more likely to focus on the treatment of people affected by earthquakes out of concern for life safety.

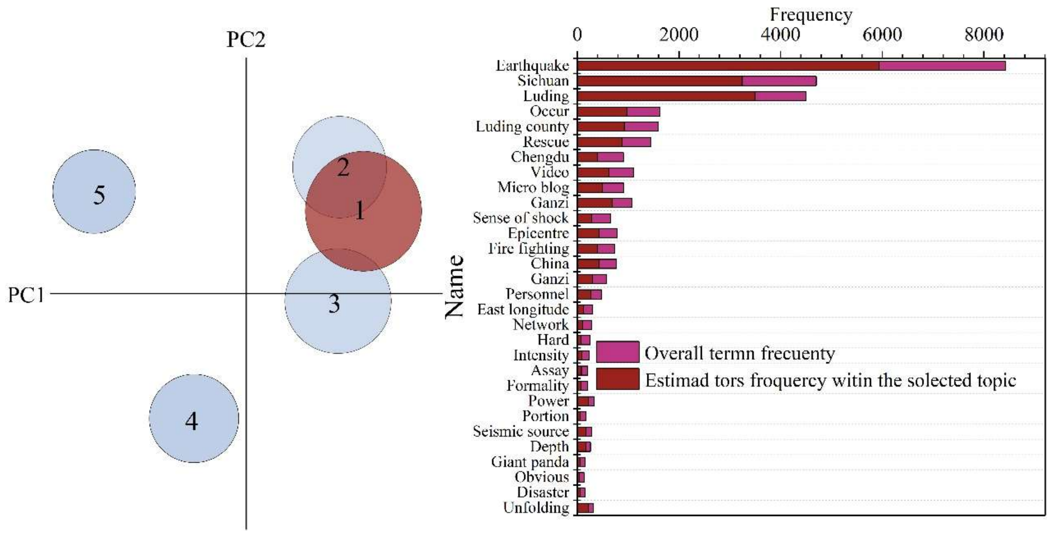

Fig. 4. Topic analysis of "Demand area situation"

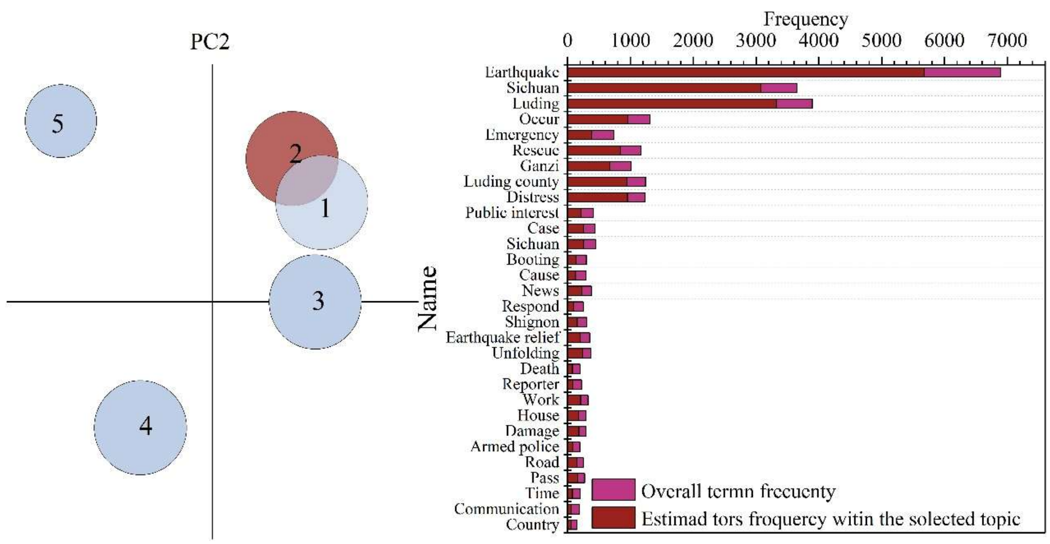

Fig. 5. Emergency rescue "theme analysis results

The results of the analysis of the theme "Pray for Peace" are shown in Figure 6. Theme 3 "Praying for Peace" and Theme 1 "Earthquake Zone Situation" partially overlap. Most of the netizens under this theme are people outside the earthquake area, and after learning about the situation in the earthquake area, they pray for the people affected by the disaster and the firefighting forces who go to the rescue, so the word "peace" is a keyword with high frequency under the whole theme. Words such as "killed", "deceased" and "rest in peace" express the feelings of pity for the victims, words such as "pay attention to safety" and "tide over difficulties together" reflect the people's ardent concern and spiritual encouragement for the people in the disaster areas, and words such as "strong" and "hope" convey the public's respect and belief in life and encourage the people affected by the disaster to persevere.

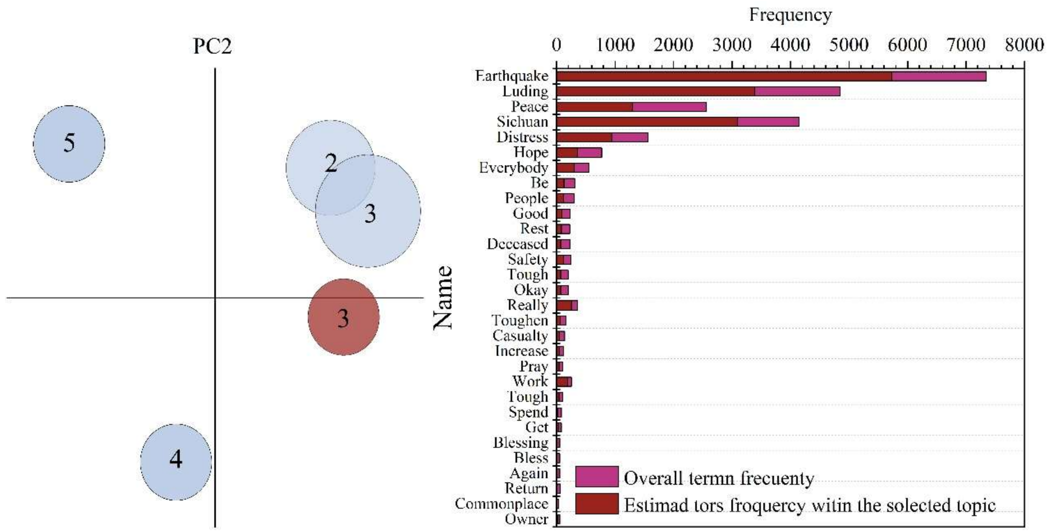

Fig. 6. Pray for peace" theme analysis results

As shown in Figure 7, the number of casualties caused by natural disasters is staggering, and although human power is still small in the face of natural disasters, the continuous development of early warning technology has continuously minimized the damage caused by disasters. Under theme 4, words such as "peace and security" and "heroes" convey the sincere prayers of netizens for the safety of the social forces that go to the disaster area for rescue, and praise their noble character of selfless dedication. Words such as "salute", "cherish" and "live well" express people's feelings about life after the disaster, and they must know how to cherish the experience of life in the present after experiencing disasters. Under this theme, netizens expressed their understanding of life more emotionally, full of positive optimism, and brought spiritual comfort to more people by forwarding the good news of successful rescue and survivorship.

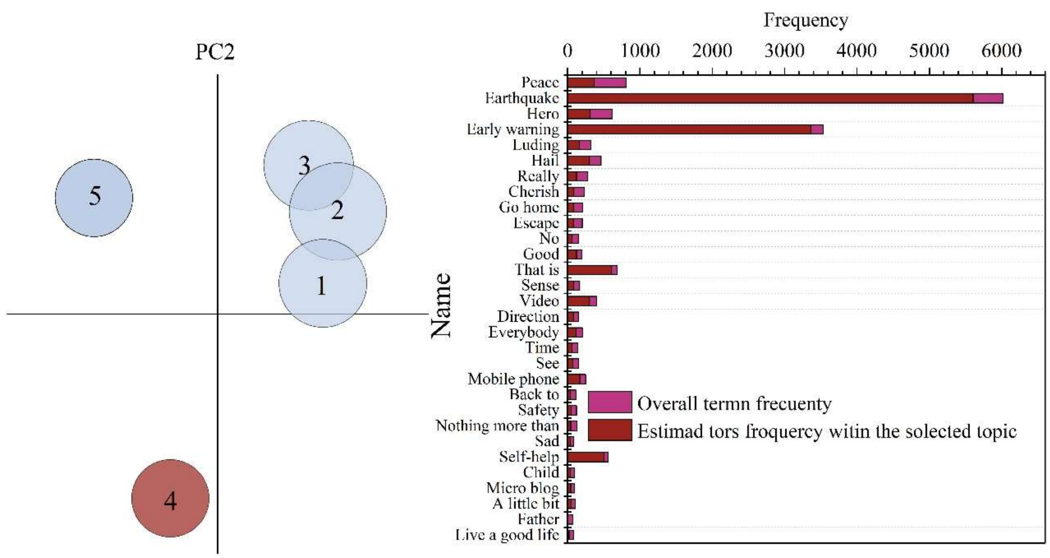

Fig. 7. Tribute to life theme analysis

As shown in Figure 8, prosocial behavior is the behavior that the actor consciously and voluntarily brings benefits to the behavior recipient, which is in line with the social hope and has no obvious benefit to the actor himself, and comforting the victim and donating money are all prosocial behaviors. Under this theme, "refueling" is the main keyword, and netizens summarized the difficulties of the Sichuan people after experiencing "high temperatures" and "power rationing" since the summer, and then encountering earthquakes, cheering for the people of Sichuan. After the earthquake, many public groups and celebrities donated money or materials to Luding, and words such as "moved" and "thank you" expressed the netizens' emotion for the unity of the Chinese people. In addition, with the help and encouragement of rescue forces and other social forces, the affected people gradually recovered, and many netizens also expressed their blessings to the victims and their satisfaction for the rest of their lives.

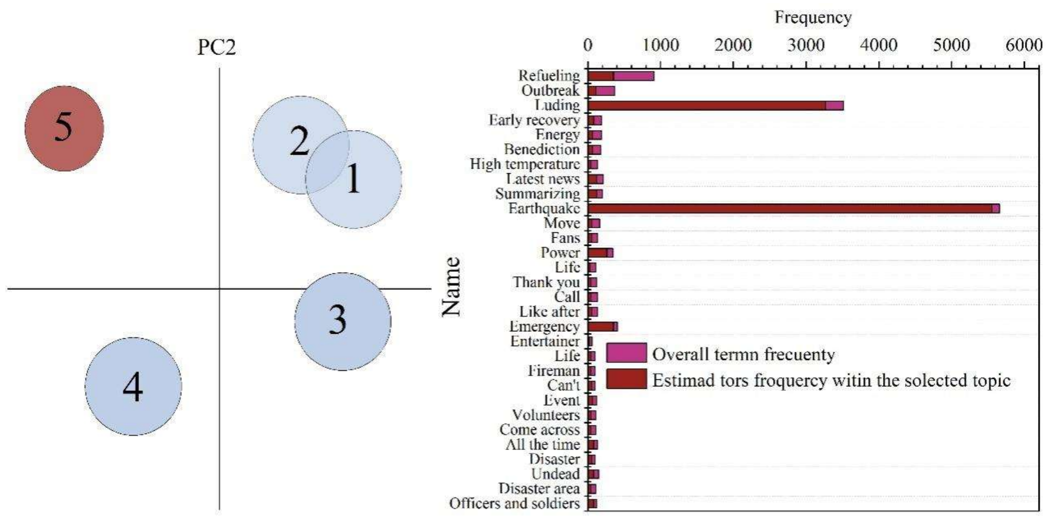

Fig. 8. Thematic analysis of "prosocial behavior"

Figure 9 shows the word cloud map of Sichuan Luding earthquake at various stages, where (a) ~ (d) are the eruption period, diffusion period, treatment period, and aftermath period, respectively, and the Sichuan Luding earthquake is a more serious seismic hazard event. During the outbreak period netizens outside the earthquake area were concerned about the location, magnitude and duration of the earthquake by browsing and reading the official media releases on the situation of the earthquake area. Internet users inside the earthquake zone, on the other hand, express their self-perception after the earthquake. With enough information, netizens tend to pay attention to the rescue situation of the earthquake, and in the diffusion period, there are a few comparative tendencies, i.e., comparing the domestic handling of the earthquake with that of earthquake-prone countries, and expressing their opinions and suggestions on the rescue situation, while the majority of netizens continue to pray for the people in the disaster area. During the processing period, the media updated the situation of all-out earthquake relief and casualties, and occasionally rumors about the disaster area came in, most netizens chose to adhere to the attitude of not believing rumors and not spreading them, and netizens focused on the rescue speech with peace and hope. Handling the results and aftermath of the period is similar to the post-disaster summary, in the event of the role models have become a source of dissemination of positive energy, whether it is the compatriots to help the public figures to the disaster area of the love of donations, or the rescue forces of the safety of the home, they are centered on the positive energy to set off a wave of moving, for the other end of the netizens are worthy of the news of relief, feel happy. Overall, in the Sichuan Luding earthquake incident, the netizens' topic has been centered on rescue, and the concern pulled by the disaster situation has been intensified with the latest notification of the situation in the quake zone over and over again, and the figures of the number of people affected by the disaster and the economic losses have pulled the public's emotions until the spread of the spirit of the role model, and the public's topic has gradually shown relief and gratification. Moreover, compared with the previous reports on the disaster area, netizens are more eager to participate in the discussion of positive topics, reflecting that most of the netizens are used to remain optimistic in the face of unavoidable natural disasters.

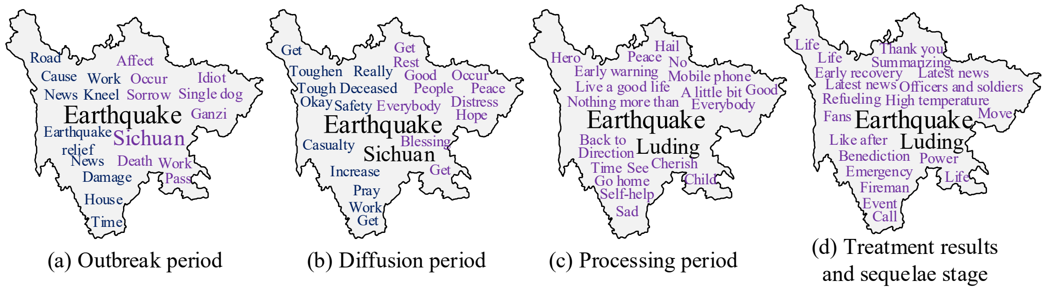

Fig. 9. Cloud map of various stages of Luding earthquake in Sichuan

3.1.5. Thematic Model Validation Analysis. In order to verify the performance of the TT-LDA topic model, the public dataset of Sichuan Luding earthquake microblog comments is used as the corpus, and the performance of the topic model is judged by comparing the perplexity and consistency of the four models, namely, LDA, TF-IDF-LDA, TextRank-LDA, and TT-LDA, and the results of the topic model validation analysis are shown in Table 2. It can be seen that the perplexity and consistency of the TT-LDA model are 446.02 and 0.3942, respectively, and compared with LDA, TF-IDF-LDA, TextRank-LDA, the theme model constructed in this paper is particularly effective in extracting the themes in terms of the Sichuan Luding earthquake.

Table 2. Topic model validation analysis results

Model	Perplexity	Coherencet
LDA	535.14	0.2153
TF-IDF-LDA	466.48	0.3433
TextRank-LDA	448.06	0.3017
TT-LDA	446.02	0.3942

3.2. Analysis of Opinion Sentiment Evolution

3.2.1. Description of the experiment. In this paper, the sentiment recognition and classification of the microblog comment corpus of the “Sichuan Luding Earthquake” emergent social network public opinion event is carried out. Since this paper adopts a supervised deep neural network model, the BiLSTM-GCN classifier model needs to be pre-trained. In this paper, because manual labeling is labor-intensive and time-consuming, the dataset used is the data provided by DataFountain on the competition of sudden public opinion events as the training set of this paper, which classifies the netizen's sentiment categories into three major categories, namely, positive/negative/neutral. The experimental design steps of this paper are shown below:

1) In this paper, we first load the training dataset and perform a series of preprocessing on the original dataset in python environment.
2) Adopt BERT pre-training model (BERT pre-training model uses pytorch pretrained and bert-base-chinese Chinese pre-trained model) to extract text features.
3) Construct data iterator to build Iterater according to batch_size=128 for each batch of data.
4) Load the BERT pre-trained model and BERT confing (original hyperparameters), update the parameters of each layer of BERT to TURE, and adjust the parameters appropriately according to the training results during training.
5) Then add two layers of BiLSTM-GCN network in the last layer of the loaded BERT, Dropout layer to prevent overfitting, and fully connected layer as the category output.
6) The features and sentiment labels in the training set are input into the BiLSTM-GCN network to train the classification, compare with the known labels, analyze and evaluate its performance, and take the average value of the text for several times of training as the final result, and finally input the

experimental data into the text sentiment classification model, and perform sentiment annotation on the microblog comment data related to the "Sichuan Luding Earthquake" event. Finally, we input the experimental data into the text sentiment classification model to label the sentiment of the microblog comment data related to the "Sichuan Luding earthquake" event.

3.2.2. Parameter setting. Different hyperparameters may have different effects on the experimental results, this paper consults several experimental literature for comparison to set the initial parameter size, and then for the consideration of the overall effect of the experiment on the pre-training model, the hyperparameters of the BiLSTM-GCN network are finally determined, with the word vector dimensions of 600, Epochs of 50, the LSTM hidden layer dimensions of 500, the number of GCN layers of 3, the learning rate of 0.001, loss rate is 0.15, and L2 regularization parameter is 0.0005.

3.2.3. Assessment of indicators. In order to quantitatively evaluate the performance of BiLSTM-GCN network model for emotion recognition, three metrics, namely, accuracy, recall, and F-value, are used above the selection of evaluation metrics to validate the classification results of the training set. The specific calculation is shown in the following formula:

$$precision = \frac{TP}{TP + FP} \tag{9}$$

$$\mathrm{Re}\,call = \frac{TP}{TP + FN} \tag{10}$$

$$F1 - Score = 2 * \frac{\mathrm{Pr}\,ecision * \mathrm{Re}\,call}{\mathrm{Pr}\,ecision + \mathrm{Re}\,call} \tag{11}$$

Wherein, TP denotes as the number of positive samples predicted as positive samples, FP denotes the number of negative samples predicted as positive samples, FN denotes the number of positive samples predicted as negative samples, and TN denotes the number of negative samples predicted as negative samples.

3.2.4. Data analysis. These three evaluation metrics have different focuses in different application scenarios, in which the accuracy rate applies to the prediction result evaluation to determine whether the model classifies the positive samples accurately or not. Recall is applicable to the evaluation of training results, to evaluate how many positive samples have been recalled by the training model. From the applicable scenarios and the calculation logic, it can be seen that the accuracy and recall are two contradictory indicators, while the F1 value is a combination of the first two, if the sample category ratio is not balanced using the F1 value for evaluation is relatively accurate. The evaluation results of the training set are shown in Table 3. From the performance evaluation table, it can be seen that the emotion recognition model constructed in this paper has a fairly high accuracy rate and can better recognize the emotions of netizens, so it can be applied to the "Sichuan Luding Earthquake" unexpected online public opinion events to classify the attitudes of netizens.

Table 3. The training set evaluates the results

Affective type	Accuracy rate	Recall rate	F1 value
The right side	0.9521	0.9335	0.9427
Neutrality	0.9686	0.9733	0.9713
negative	0.9915	0.9935	0.9932

Sentiment data statistics are shown in Table 4, in which A1~A3 represent positive, neutral and negative sentiments, respectively. In this paper, after the pre-training of the sentiment classification model, the collected comment data of "Sichuan Luding Earthquake" is inputted into the BiLSTM-GCN model to classify the sentiment types, and the preprocessed 1471890 data are classified by the time period from 9-5 to 9-7, and finally the following data are obtained 488498 positive comment data, 466832 neutral comment data, and 516560 negative comments. It is not difficult to find that negative emotions have been in the dominant position in the whole life cycle of the event, and the public has been looking at the development of the whole event with a negative attitude, and even appeared the phenomenon of "one-sided" polarization of emotions. This shows that the incident has had a great adverse impact on society. Although there are many positive comments during the whole evolution process, they are not enough to influence the public's negative emotions towards the incident. From the table, it is not difficult to find that there are some gaps in the emotional polarity of the event at different stages, and each stage of emotional polarity shows different evolutionary patterns and development trends.

Table 4. Affective data statistics

Time/h	A1	A2	A3	Time/s	A1	A2	A3	Time/h	A1	A2	A3
9/5-12:00	10431	10571	10995	9/6-12:00	3109	3165	3062	9/7-12:00	9426	10444	10755
9/5-13:00	10095	10133	10750	9/6-13:00	4139	3264	3147	9/7-13:00	9383	10390	10473
9/5-14:00	9969	9880	10668	9/6-14:00	4220	3380	3216	9/7-14:00	9245	9956	10339
9/5-15:00	8914	9591	9177	9/6-15:00	4282	3443	3626	9/7-15:00	9028	9638	9724
9/5-16:00	8905	9493	9131	9/6-16:00	4288	3903	4124	9/7-16:00	8220	9442	9541
9/5-17:00	8812	9364	8656	9/6-17:00	4604	4002	4432	9/7-17:00	8174	8017	9480
9/5-18:00	8723	8461	8493	9/6-18:00	5857	4195	4818	9/7-18:00	8108	7632	9450
9/5-19:00	7976	8354	8294	9/6-19:00	6142	4277	5210	9/7-19:00	7874	7179	9217
9/5-20:00	7182	8235	7677	9/6-20:00	6147	4556	5281	9/7-20:00	7519	7162	9067
9/5-21:00	7149	8054	7535	9/6-21:00	6215	4903	5725	9/7-21:00	7185	7058	8441
9/5-22:00	6851	7908	7501	9/6-22:00	6558	5370	5976	9/7-22:00	7023	6934	8396
9/5-23:00	6395	7387	7489	9/6-23:00	6944	5418	6678	9/7-23:00	6642	6120	8310
9/6-00:00	6373	7100	7121	9/7-00:00	7672	5671	6970	9/8-00:00	6498	5863	8127
9/6-01:00	6246	7029	7098	9/7-01:00	8089	5775	8138	9/8-01:00	6211	5398	6988
9/6-02:00	6094	6600	6626	9/7-02:00	8176	5809	8139	9/8-02:00	5525	5233	6893
9/6-03:00	5620	6438	6296	9/7-03:00	8503	6216	8279	9/7-03:00	5514	5032	6772
9/6-04:00	5319	6359	6270	9/7-04:00	8921	6492	8380	9/8-04:00	5139	5029	6456
9/6-05:00	4841	5494	5867	9/7-05:00	8927	6956	8413	9/8-05:00	5086	4926	6285
9/6-06:00	4736	5110	5570	9/7-06:00	9398	6990	8458	9/8-06:00	4873	4868	6245
9/6-07:00	4373	4597	5507	9/7-07:00	9587	7152	9201	9/8-07:00	4687	4711	5681
9/6-08:00	3941	4345	4600	9/7-08:00	9849	9272	9732	9/8-08:00	4413	4195	3983
9/6-09:00	3800	3558	4578	9/7-09:00	10416	9682	10033	9/8-09:00	3524	4002	3664
9/6-10:00	3535	3262	4339	9/7-10:00	10611	9860	10197	9/8-10:00	3339	3878	3383
9/6-11:00	3164	3057	3585	9/7-11:00	10759	9961	10671	9/8-11:00	3005	3633	3161

The statistics of emotional intensity in each evolutionary stage are shown in Table 5. The event in real life once occurred that quickly spread on the microblogging social network, after a series of posting, retweeting and commenting on the event quickly broke out, attracting more netizens to pay attention to the event, the first outbreak of the stage of the network is filled with the heartache of the victims before the regret and the safety of the life of the heart, at this time, the average polarity strength of -0.178 (this paper, in the calculation of the emotional intensity of the polarity) (In this paper, when calculating the strength of emotional polarity, the number of positive comments, the number of neutral comments, and the number of negative comments are given a weight of 2, 0, and -2, respectively, for the weighted equalization calculation). With the rescue related to the Sichuan Luding incident, the exposure of related rescue information, information on the victims, and videos of the earthquake moments, the negative emotions of netizens were enhanced, and the process of the first recession was filled with netizens' compassion for the victims. In the second outbreak stage, when the number of casualties was officially announced, netizens' negative emotions erupted in full force, with an average polarity as high as -0.886. In the second recession stage, after netizens knew the original cause of the whole incident, the polarity of the negative emotions began to level off, but compared to the first outbreak stage, there were still more negative emotions.

Table 5. Statistical results of emotion intensity in each evolutionary stage

Evolution stage	Affective polarity				Mean polar intensity
	Positive comments	Neutral comment count	Negative review count	Total comments	
First burst	135895	146451	145644	427990	-0.178
First recession	86054	73805	83474	243333	0.289
Second burst	164384	147723	166923	479030	0.886
Second recession	102165	98853	120519	321537	0.446
Total	488498	466832	516560	1471890	0.409

4. Conclusion

In this paper, we jointly study the public opinion sentiment after the MS6.8 earthquake in Sichuan Luding with the help of TT-LDA and GCN-BiLSTM models under the perspective of deep learning.

1) Three days after the Luding MS6.8 earthquake, the number of retweets, comments, and likes are 933214, 676109, and 1413261837 respectively, and the speed of microblogging public opinion dissemination of the earthquake event is 4138.239, and the perplexity and consistency of this paper's T-LDA model is 446.02 and 0.3942, which is of higher quality compared with the performance of the other three topic models.

2) Positive ratings, neutral comment data, and negative comment data from 2022-September-5 to 2022-September-7 were 488498, 466832, and 516560, and the first stage was filled with compassion, with an average polarity strength of -0.886, and the entire process was dominated by negative sentiment data.

Funding

Project (2024SSY11) supported by Shanghai Sheshan National Geophysical Observatory.

Project (23DZ1203902, 22DZ1200200, 22DZ1201400) supported by Shanghai Science and Technology Program Project.

References

[1] He, W., Yuan, Q., & Li, N. (2023). Research on the Characteristics of Internet Public Opinion and Public

Sentiment after the Sichuan Earthquake Based on the Perspective of Weibo. Applied Sciences (2076-3417), 13(3).

[2] CONTRERAS, D., SHRESTHA, R., DHUNGEL, R., Shrestha, S., PAUDEL, K., SHARMA, S., ... & Milillo, P. (2023). ASSESSMENT OF THE RECOVERY PROCESS USING SENTIMENT AND TOPIC ANALYSIS: THE CASE STUDY OF THE 2015 GORKHA EARTHQUAKE. SECED 2023.

[3] Bossu, R., Fallou, L., Landès, M., Roussel, F., Julien-Laferrière, S., Roch, J., & Steed, R. (2021). Rapid Public Information and Situational Awareness After the November 26, 2019, Albania Earthquake: Lessons Learned From the LastQuake System. The Power of Citizen Seismology: Science and Social Impacts.

[4] Finazzi, F. (2020). The Earthquake Network project: a platform for earthquake early warning, rapid impact assessment and search and rescue. FRONTIERS IN EARTH SCIENCE, 8(art. n. 243), 1-10.

[5] Amiresmaili, M., Zolala, F., Nekoei-Moghadam, M., Salavatian, S., Chashmyazdan, M., Soltani, A., & Savabi, J. (2021). Role of social media in earthquake: A systematic review. Iran. Red. Crescent Med. J, 23(5).

[6] Mano, R. M., Kirshcenbaum, A., & Rapaport, C. (2019). Earthquake preparedness: A Social Media Fit perspective to accessing and disseminating earthquake information. International Journal of Disaster Risk Management, 1(2), 19-31.

[7] Qiu, Q., He, J., Chen, H., & Qiu, J. (2019, December). Research on the evolution law of emergency network public opinion. In 2019 12th International Symposium on Computational Intelligence and Design (ISCID) (Vol. 2, pp. 157-161). IEEE.

[8] You, Y., Huang, Y., & Zhuang, Y. (2020). Natural disaster and political trust: A natural experiment study of the impact of the Wenchuan earthquake. Chinese journal of sociology, 6(1), 140-165.

[9] Ying, L., Zhiming, L., ZHANG, H., & Haixia, J. I. N. (2023). A Study on Public Opinion Strategy of New Media Platforms of Earthquake in Emergency Responses. North China Earthquake Sciences, 41(2), 32-36.

[10] Huang, X., Jin, H., & Zhang, Y. (2019). Risk assessment of earthquake network public opinion based on global search BP neural network. PLoS ONE, 14(3).

[11] Fersini, E., Messina, V., & Pozzi, F. (2017). Earthquake management: a decision support system based on natural language processing. JOURNAL OF AMBIENT INTELLIGENCE AND HUMANIZED COMPUTING, 8(1), 37-45.

[12] Su, X., Ma, S., Qiu, X., Shi, J., Zhang, X., & Chen, F. (2021). Microblog Topic-Words Detection Model for Earthquake Emergency Responses Based on Information Classification Hierarchy. IJERPH, 18(15), 1-20.

[13] Pirnau, M. (2018, June). Sentiment analysis for the tweets that contain the word "earthquake". In 2018 10th International Conference on Electronics, Computers and Artificial Intelligence (ECAI) (pp. 1-6). IEEE.

[14] Wang, J. H., Liu, T. W., Luo, X., & Wang, L. (2018). An LSTM Approach to Short Text Sentiment Classification with Word Embeddings. ROCLING 2018, 214.

[15] Kumar, S., Nirschl, M., Holtmann-Rice, D., Liao, H., Suresh, A. T., & Yu, F. (2017). LATTICE RESCORING STRATEGIES FOR LONG SHORT TERM MEMORY LANGUAGE MODELS IN SPEECH RECOGNITION. stat, 1050, 15.

[16] Bin, Y., Yang, Y., Shen, F., Xie, N., Shen, H. T., & Li, X. (2019). Describing Video With Attention-Based Bidirectional LSTM. IEEE transactions on cybernetics, 49(7), 2631-2641.

[17] Zhang, R., Chen, Z., Chen, S., Zheng, J., Büyüköztürk, O., & Sun, H. (2019). Deep long short-term memory networks for nonlinear structural seismic response prediction. Computers & Structures, 220, 55-68.

[18] Gandhi, U. D., Malarvizhi Kumar, P., Chandra Babu, G., & Karthick, G. (2021). Sentiment Analysis on Twitter Data by Using Convolutional Neural Network (CNN) and Long Short Term Memory (LSTM). Wireless

Personal Communications.

[19] Li, S., Liu, Z., & Li, Y. (2019). Temporal and spatial evolution of online public sentiment on emergencies. Information Processing & Management, 57(2), 102177-102177.

[20] Li, X., Wang, Z., Gao, C., & Shi, L. (2017). Reasoning human emotional responses from large-scale social and public media. Applied Mathematics and Computation, 310(C), 182-193.

[21] Contreras Mojica, D., Wilkinson, S., Didem Aktas, Y., Fallou, L., Bossu, R., & Landès, M. (2022). Intensity-based sentiment and topic analysis. The case of the 2020 Aegean Earthquake. Frontiers in Built Environment, 8.

[22] LI, L., WANG, Z., ZHANG, Q., & WEN, H. (2020). Effect of anger, anxiety, and sadness on the propagation scale of social media posts after natural disasters. Information Processing and Management.

[23] Li, S., & Sun, X. (2023). Application of public emotion feature extraction algorithm based on social media communication in public opinion analysis of natural disasters. PeerJ Computer Science, 9.

[24] Aoudi, S., & Malik, A. (2018, November). Lexicon based sentiment comparison of iPhone and android tweets during the Iran-Iraq earthquake. In 2018 Fifth HCT Information Technology Trends (ITT) (pp. 232-238). IEEE.

[25] Ma, X., Liu, W., Zhou, X., Qin, C., Chen, Y., Xiang, Y., ... & Zhao, M. (2020). Evolution of online public opinion during meteorological disasters. Environmental Hazards, 19(4), 375-397.

[26] Wei-dong, H., Wang, Q., & Jie, C. (2018). Tracing Public Opinion Propagation and Emotional Evolution Based on Public Emergencies in Social Networks. International Journal of Computers, Communications and Control, 13(1).

[27] Contreras, D., Wilkinson, S., Alterman, E., & Hervas, J. (2022). Accuracy of a pre-trained sentiment analysis (SA) classification model on tweets related to emergency response and early recovery assessment: the case of 2019 Albanian earthquake. Natural Hazards.

[28] Chenxi, L., Jilin, F., Meng, H., & Zhonghao, W. (2022, May). Research on Post Earthquake Public Opinion Analysis Based on XLNet-BiGRU-A Algorithm. In 2022 IEEE 2nd International Conference on Computer Communication and Artificial Intelligence (CCAI) (pp. 81-84). IEEE.

[29] Yang, T., Xie, J., Li, G., Mou, N., Li, Z., Tian, C., & Zhao, J. (2019). Social Media Big Data Mining and Spatio-Temporal Analysis on Public Emotions for Disaster Mitigation. ISPRS International Journal of Geo-Information, 8(1), 29.

[30] Yuan, Q., Wang, S., & Li, N. (2024). Research on emotional tendency of earthquake disaster based on E-Trans model: take the topic of “Sichuan Earthquake” on microblog as an example. Natural Hazards: Journal of the International Society for the Prevention and Mitigation of Natural Hazards, 120(6), 5057-5074.

[31] Sumy, D. F., Welti, R., & Hubenthal, M. (2020). Applications and Evaluation of the IRIS Earthquake Browser: A Web-Based Tool That Enables Multidimensional Earthquake Visualization. Seismological Research Letters, 91(5), 2922-2935.

[32] Xiaoyu Wang,Junping Xu & Younghwan Pan. (2024). Generation Z Satisfaction with Smart Homestays: ASCI and Web Crawler Insights from China. Electronics(20),4003-4003.

[33] Liang Gao,Fang Cui & Chengbo Zhang. (2024). Library Similar Literature Screening System Research Based on LDA Topic Model. Journal of Information & Knowledge Management(05),

[34] Ju Hyun Lee& Michael J. Ostwald. (2024). Latent Dirichlet Allocation (LDA) topic models for Space Syntax studies on spatial experience. City, Territory and Architecture(1),3-3.

[35] Mengdi Zhang,Qiao Shen,Zhiheng Zhao,Shuaian Wang & George Q. Huang. (2025). Optimizing ESG reporting: Innovating with E-BERT models in nature language processing. Expert Systems With Applications125931-125931.

[36] Yanhong Wang,Chen Gong,Xiaodong Ji & Qi Yuan. (2025). Text classification for evaluating digital technology adoption maturity based on BERT: An evidence of Industrial AI from China. Technological Forecasting & Social Change123903-123903.