Challenges Confronting the Sustainability of Anti-Epidemic Policies Based on the BERT-Pair-ABSA Model

Yanchun Zhu[1], Luyao Li[2], Chenguang Li[3*], Wei Zhang[2]

[1] Business School, Beijing Normal University, 100875 Beijing, China
[2] School of Information, Central University of Finance and Economics, 100081 Beijing, China
[3] School of Insurance, Central University of Finance and Economics, 100081 Beijing, China

[*] Correspondence: Chenguang Li (lichenguang@cufe.edu.cn)

Abstract: As a major public health emergency, the COVID-19 pandemic has some uncertainties. Coupling of the uncertainties and anti-epidemic policies easily leads to the spread of negative emotions. It is challenging to maintain the sustainability of anti-epidemic measures. Therefore, this paper aimed to analyze the challenges that the sustainability of anti-epidemic measures is faced with. A topic clustering extension method was proposed, which integrated Latent Dirichlet Allocation (LDA) topic information with Bidirectional Encoder Representations from Transformers (BERT) contextual information through aspect-based sentiment analysis. In addition, this paper constructed a thesaurus of aspect words from the two stages of dynamic zero-COVID and orderly relaxation of epidemic control. This paper established the BERT-pair-ABSA model for semantic expansion of auxiliary sentences and calculated sentiment polarity to gain insight into the changes in netizens' concerns, emotional states and evolution trends at different stages. The research results showed: (1) Compared with the benchmark model, the proposed sentiment analysis model had better classification accuracy and was applicable to the sentiment classification of short texts in the epidemic situation; (2) During the dynamic zero-COVID stage, netizens paid attention to grassroots epidemic management and the scope of lockdown and epidemic control, which were closely related to both specific lockdown and control management, and the implementation of regional epidemic management; and (3) in the orderly relaxation stage of epidemic control, netizens were concerned about drug guarantee, medical care guarantee, personal health protection and health protection of special population groups, and negative emotions always dominated in drug guarantee, medical care guarantee and health protection of special groups. The negative sentiment of drug guarantee, medical care guarantee and health protection of special groups always dominated. The results provided an empirical basis for the optimization and adjustment of the anti-epidemic policies.

Keywords: Aspect-based sentiment analysis; Netizen concerns; Anti-epidemic policies; Online emotions

1. Introduction

As a major public health emergency, COVID-19 is characterized by its suddenness, danger, variability and relevance to all people (Dillard & Meier, 2023). This epidemic triggered great concern in society and bred negative emotions, such as panic, anxiety and anger (Liang et al., 2023; Zhang et al., 2022c). With the intertwining and collision of various negative emotions, under the constraints of limited time and resources and limited by knowledge, experience and cognitive ability, people easily lost their ability to perceive risks and think and judge rationally under normal circumstances (De Landsheer & Walburg, 2022). In addition, irrational behaviors, such as rumor spread, panic buying, and regional stigmatization occurred from time to time, which dissipated social cohesion and disturbed the social order (Inam et al., 2022).

Based on the characteristics of virus mutation and spread at different stages, and starting from crisis resolution and satisfaction of netizens' needs, governments made dynamic adjustments to their anti-epidemic policies. The adjustments and promulgation of these policies stimulated corresponding emotional feedback from netizens

(Campbell, 2012), which directly involved the acceptance of the policies by netizens and then affected the implementation and effects of the policies (Zhang et al., 2020).

The combination of uncertainties of the COVID-19, continuous adjustments of the anti-epidemic policies, and the spread of negative emotions among netizens increased the complexity and difficulty of crisis management. Therefore, it is challenging to maintain the sustainability of the anti-epidemic policies.

Most of the existing studies have focused on a certain epidemic, and analyzed the emotional states and distribution characteristics of netizens under the policy framework by considering the anti-epidemic policies as a whole. However, few studies have done longitudinal research on the impact of different anti-epidemic policies on the concerns and emotional changes of netizens (Sukhwal & Kankanhalli, 2022; Wen & Zheng, 2022). Due to the rapid spread and widespread infections, the anti-epidemic policies had a wide range of contents. Different contents led to different concerns and emotional feedback. However, the existing studies have treated the anti-epidemic policies in a holistic manner, which made it difficult to discern which part of the policies finally caused the sentiment changes of netizens, thus limiting subsequent adjustments to improve the policies.

Therefore, this paper proposed to adopt an aspect-based sentiment analysis method to deconstruct the anti-epidemic policies from the concerns of netizens, and analyzed in depth the changes in public concerns and the evolution of emotions in the two stages of dynamic zero-COVID and orderly relaxation of epidemic control, thus providing more specific and practical direction guidance and decision-making basis for the continuous adjustment and optimization of policy measures.

2. Literature Review

2.1 Online Emotions and Impact

Online emotions is a collection of similar emotional experiences that a certain number of internet users share about an event or a social phenomenon within a certain time frame. Current studies on online emotions focus on two aspects: (1) the classification of different dimensions of online emotions; and (2) the formation factors of online emotions and their effects (Zhang et al., 2022a). For the research on online emotions in public health emergencyies, most scholars have focused on a specific epidemic, identified the overall emotional states and characteristics of netizens using social media data, and explored and identified the role of the epidemic in influencing the emotions of netizens, such as the aggravation of stress, anxiety, depression, panic and other emotions, as well as the decrease in well-being of life (Ahmad & Murad, 2020; Lin et al., 2020). However, few studies have focused on the effects of anti-epidemic policies and adjustments on netizens' emotions, such as delay of school opening, lockdown, etc. (Wen & Zheng, 2022; Zhang et al., 2022a). Sukhwal & Kankanhalli (2022) calculated the daily sentiment values of Singaporean netizens regarding the COVID-19 as a whole using related Facebook posts and used the Natural Language Processing (NLP) approach and constructed a multiple regression model. The empirical analysis found that some policy measures had a significant negative effect on netizens' sentiment, such as control of international flights and cancellation of public events, and some had a significant positive effect, such as ordering of masks and income support. The lockdown and shutdown measures from implementation to gradual relaxation first increased and then decreased the sentiment values of netizens (Wen & Zheng, 2022). These studies have only explored the sentiment changes of netizens under specific anti-epidemic policies, such as delay of school opening and lockdown. However, the anti-epidemic policies had extensive contents and different contents led to different concerns and emotional feedback. Therefore, this paper aimed to identify the reasons for the ultimate changes in the overall sentiment of netizens caused by anti-epidemic policies, and make a longitudinal research on the impact of different anti-epidemic policies on the concerns and sentiment changes of netizens.

2.2 Aspect-Based Sentiment Analysis

As for the problem of sentiment misclassification due to the lack of consideration of sentiment targets in traditional sentiment analysis studies, aspect-based sentiment analysis (ABSA) has been proposed in recent years, where the concerned target on which the sentiment is expressed shifts from an entire sentence or document to an entity or a certain aspect of an entity (Nazir et al., 2020).

ABSA consists of two core tasks: aspectual lexicon construction and polarity classification (Zhang et al., 2022b). Two main methods have been used to construct the aspectual lexicon and one is to obtain an exclusive lexicon that matches the research target by deduplication and merging based on the existing lexicon. As for the other method, the bag-of-words-based LDA topic model or target document embedding is used first. Then clustering algorithms are used to identify similar structures in the vector space together with other methods in order to achieve topic extraction. Finally manual evaluation and screening are supplemented to form the lexicon.

In terms of polarity classification, the strength of the ABSA approach lies in its emphasis on aspect-specific sentiment judgments using contexts associated with aspects. To improve the classification performance of this

type of models, scholars have tried single or multilayer attention mechanisms to fuse the information of aspect words and context more effectively, thus prompting the models to automatically focus on the important parts of the sentences (Yu et al., 2022). Du & Liu (2020) proposed a BHAN (Bert-based Helical Attention Network) model based on Helical Attention Networks, which, calculated the weights of aspect words with the new representation after obtaining a weighted contextual representation based on aspect words, and recalculated the weights of context with this new representation of aspect words. The performance of the aspect-level sentiment analysis model was further improved by fully realizing the interaction between the aspect words and the context, thus spirally improving the representation of both the context and the aspect words.

Han et al. (2020) obtained the spatiotemporal evolution distribution of seven themes and thirteen subdivision aspects related to COVID-19 and the number of their opinions from microblogging text, based on the LDA model and random forest algorithm. But the study only stopped at hotspot analysis without measuring the sentiment status of netizens in different aspects (Han et al., 2020). As the COVID-19 continued and the information demand for crisis management work increased, Jang et al. (2021) first identified 20 key aspect words and 60 opinion sentiment words through LDA topic extraction and manual review by domain experts in Twitter's epidemic opinion analysis (Jang et al., 2021). Based on the measurement of seven fine-grained sentiment types of the epidemic as a whole, Wen & Zheng (2022) further explored the spatial distribution characteristics of different sentiment types, and made refinements to the analysis of epidemic public opinion from sentiment granularity and spatiotemporal perspectives. Although these studies have made efforts in disentangling the overall sentiment measurement of the epidemic towards the topic level, the assumption that a single text expressed a single sentiment for a single topic still existed, which did not match most of the cases in practice, resulting in the loss of sentiment information and the bias of sentiment calculation.

The existing studies have provided some insight into the relationship between the revision of anti-epidemic policies and the emotional adjustment of netizens. However, the following shortcomings exist: (1) the studies on the impact of public policies on the emotional feedback of netizens mainly focus on the situation and characteristics of the emotional feedback of netizens in a specific policy situation, lacking in longitudinal studies on the impact of different epidemic policies on the concerns and emotional changes of netizens. The studies usually consider the anti-epidemic policies as a whole, and analyzes the impact of policy promulgation and implementation on the sentiment of netizens. However, major anti-epidemic policies involve a wide range of fields and often give rise to a series of chain reactions in different aspects, and the changes in the sentiment presented by netizens under epidemic policies may be the effect of simultaneous changes in different aspects of the policies or the result of mutual offsetting. Therefore, macroscopic calculations of the changes in the sentiment of netizens are needed only. But it is limited and fails to provide specific directional guidance for subsequent policy optimization and adjustment. Therefore, ABSA of netizens is necessary in the epidemic management process. (2) ABSA research has made great progress in algorithm refinement and optimization, but its application in COVID-19 is not deep, failing to make full use of the existing technical achievements to fully explore the emotional information in the text of netizens'comments.

To this end, this paper constructed a model based on microblogging data using the ABSA method to explore the emotional representations and evolutionary distributions of netizens for various aspects of the anti-epidemic policies, starting from the stages of dynamic zero-COVID and orderly relaxation of epidemic control in China.

3. Construction of Aspect Word Lexicon

3.1 Topic Clustering Extension Model Based on LDA+BERT+K-Means

LDA model tends to perform poorly in short texts, data with sparse content, or implicit opinion expressions because it ignores contextual information by disregarding order and syntax (Li et al., 2023). At this point, additional more comprehensive contextual information is needed to assist in identifying the central theme of the text. Therefore, this paper introduced a BERT word vector model containing deep bidirectional contextual information, which combined the advantages of both models well, and identified a scheme for building an aspectual word lexicon, that incorporated LDA topic information and BERT semantic expansion into K-Means clustering. Specifically, the following key steps were included (Figure 1).

(1) The original corpus was refined by the LDA model to obtain the topic probability distribution vector v_lda.

$$v\_lda_i = [t_{i1}, t_{i2}, \ldots, t_{ik}] \tag{1}$$

where, t_{ij} denotes the probability of the i-th text in the j-th topic, and k is the number of topics.

(2) The original corpus was pretrained by BERT to obtain a sentence vector representation v_bert containing contextual information.

$$v\_bert_i = [w_{i1}, w_{i2}, \ldots, w_{im}] \tag{2}$$

where, v_bert_i is the pooler_output of BERT's pretrained model, which is usually used to represent the sentence vector of the whole sentence, and m is the number of hidden layer layers.

(3) $v\_ldabert$ was obtained through weighted splicing of the topic distribution vector of LDA and the BERT sentence vector, which fused the contextual semantic information in BERT on the basis of LDA topic information, thus achieving semantic expansion.

$$v\_ldabert_i = \lambda \cdot v\_lda_i \oplus v\_bert_i \tag{3}$$

where, λ is used to adjust the weight of thematic and contextual information.

(4) The splicing vector was input into the self-encoder to achieve dimensionality reduction and obtain a low-dimensional potential space representation z, then $\chi = v\_ldabert$.

$$z = f\left(W^{(1)}x + b^{(1)}\right) \tag{4}$$

$$\hat{x} = f\left(W^{(2)}z + b^{(2)}\right) \tag{5}$$

$$MinimizeLoss = \text{dist}(x, \hat{x}) \tag{6}$$

where, $W^{(1)}$, $W^{(2)}$, $b^{(1)}$, and $b^{(2)}$ are the learning parameters of the self-encoder, x is the high-dimensional input vector, z is the low-dimensional hidden variable, and $\hat{x}$ is the reduced output variable. The model kept learning by minimizing the loss function, which enabled the reduced output variable $\hat{x}$ reconstructed by the low-dimensional hidden variable z to be infinitely close to the initial input variable x. Then, the hidden variable z was considered to have learned the most informative features of x, and the high-dimensional input vector was replaced by the low-dimensional hidden vector to achieve data dimensionality reduction.

(5) The low-dimensional hidden variable z was input into the K-Means algorithm and cluster to obtain the topic words and the keywords under the topic.

(6) Finally, the final aspect words and corresponding keyword lists were obtained by manual screening, merging or stratification according to the idea of progressive refinement.
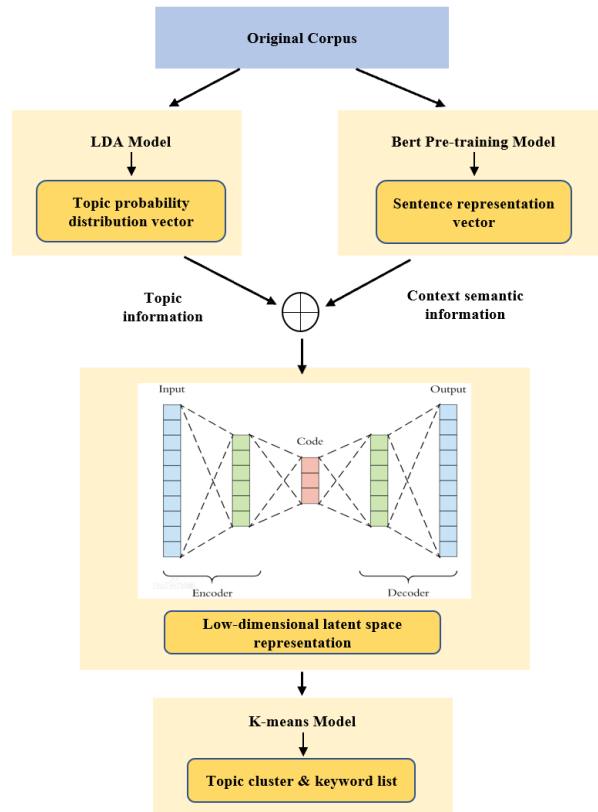

Figure 1. Flow chart of the topic clustering extension model of LDA+BERT+K-Means

3.2 Experimental Analysis and Validation

3.2.1 Data acquisition and cleaning

Taking into account the representativeness of local anti-epidemic policy adjustments and the availability of official microblogs, this paper selected "Shijiazhuang Release", "Health Guangzhou", and "Capital Health" from the Sina Weibo Health Commission's public affairs hotlist. The official microblogs of three local health care commissions were looked through, including the dynamic zero-COVID stage (June-December 2022) and the orderly relaxation stage of epidemic control (November 2022-January 2023). The data of user comments related to the anti-epidemic policies exceeded 10,000 items each.

3.2.2 Experimental group design and comparison of results

The LDA+Bert+Cluster model was used as the experimental group, and the LDA model and K-Means model as the control group. Among them, the K-Means algorithm was implemented by the K-Means model and mini_batch-Kmeans model, and the batch_size in the mini_batch-Kmeans model was set to 1000. The experimental scheme was designed as below. According to the model evaluation results in Table 1, this paper incorporated LDA topic information with BERT semantic information. The topic clustering model incorporating LDA topic information and BERT semantic information significantly improved the topic extraction task of short text of microblog comments.

(1) TF-IDF + K-Means: TF-IDF was first used to obtain the weight of each word in the comment text, and the highest frequency threshold was set to 0.5 to obtain a text-word matrix with dimension (10636, 12628), which was then input to the K-Means algorithm as features, and the optimal clustering result was output after averaging for ten times.

(2) hash+K-Means: hash was used to reduce the dimension of high-dimensional sparse text features and the hash dimension was set to 5000. The text representation with dimension (10636, 5000) was obtained and then input as features into K-Means algorithm. The optimal clustering result was output after averaging for ten times.

(3) TF-IDF+LSA+K-Means: The text representation obtained by TF-IDF was dimensionally reduced by the LSA model using Singular Value Decomposition (SVD), and the implicit semantic information was captured at the same time to obtain a low-dimensional matrix with dimensions (10636, 100). Then it was fed into the K-Means algorithm as features to output the optimal clustering result after averaging for ten times.

(4) hash+LSA+K-Means: The text representation obtained by hash was dimensionally reduced by the LSA model using SVD, and the implicit semantic information was captured at the same time to obtain a low-dimensional matrix with dimension (10636, 100). Then it was input to the K-Means algorithm as features to output the optimal clustering result after averaging for ten times.

(5) LDAMallet: The Mallet version of the LDA model usually provided higher quality themes.

(6) Bert+K-Means: The hfl/Chinese-roberta-wwm pretrained model was used to obtain word vectors, which were then fed into the K-Means algorithm as features.

(7) LDA+Bert+K-Means: the LDA-based topic distribution vector and the BERT-based word embedding vector drawn were stitched together, and a weight value of 15 was set for the former to emphasize the topic information. Adam optimizer was used for the autoencoder, with epochs set to 200 and batch_size set to 128. The resulting low-dimensional representation was used as the feature input to the K-Means algorithm, and the optimal clustering result was output after averaging for ten times.

Table 1. Comparison of theme extraction model results

Models	Orderly relaxation stage of epidemic control		Dynamic zero-COVID stage	
	Text representation	Contour factor	Text representation	Contour factor
Mini-Kmeans	X_tfidf	0.026	X_tfidf	0.031
	X_hash	0.059	X_hash	0.046
	X_tfidf_lsa	0.088	X_tfidf_lsa	0.071
	X_hash_lsa	0.120	X_hash_lsa	0.105
K-Means	X_tfidf	0.028	X_tfidf	0.035
	X_hash	0.059	X_hash	0.046
	X_tfidf_lsa	0.104	X_tfidf_lsa	0.093
	X_hash_lsa	0.138	X_hash_lsa	0.108
Models	Consistency score	Contour factor	Consistency score	Contour factor
LdaMallet	0.388	/	0.389	/
Bert+K-Means	0.287	0.133	0.281	0.128
LDA+Bert+K-Means	0.397	0.279	0.425	0.255

3.2.3 Aspect word lexicon display

In the dynamic zero-COVID stage, seven topics, twelve aspects and their keywords were finally obtained by combining the clustering results and manual analysis and evaluation, and the results are shown in Table 2.

Table 2. Thesaurus of aspect words in dynamic zero-COVID stage (partial)

Topic	Aspect	Keywords
0: Anti-epidemic policies	Anti-epidemic policies	Anti-epidemic; Anti-epidemic policy; Stabilization; Sealing; Prevention and control; Control criteria; Silence; Lockdown; Risk zone; Accurate countermeasures; Dynamic clearing; Notification; Society; People; Twenty measures; Adjustment and optimization...
	Government work	Government leadership; Press conference; Epidemiological survey; Screening; Law enforcement management; Regulation Treatment; Disposal; Penalty; Charge card; Escape; Disorderly walk; Clothes delivery; Livelihood; Hotline 12345; Coordination...
	Grassroots management	Community; grassroots; business unit; Sub-district; Property management division; village; one-size-fits-all; Incremental; Over-proofing; lift lockdown; Stay-at-home order; Residents; Lift lockdown; Disinfect...
1: Nucleic acid policies	Lockdown scope	Area; End lockdown; Extension; Cross-zone; Risk zone division; Low risk; Medium-high risk; Stay-at home; Lockdown Zones; Controlled Zones; Recovery; Large-scale; New case; Adding; Precise; Semi-confined; Outdoor; Public place...
	Necessity of nucleic acid testing	Swab; Cancel; Gather; One-day test; Three-day test; Mass nucleic acid test; Invalid; Charge; Workplace; Go to work; Negative nucleic acid results; Positive; Green code; Normal; Yellow code...
	Convenience of nucleic acid testing	Nucleic acid sites; Testing sites; Queuing; Time limit; Convenience; Test results; Fees Charges; Close; Reserve; Restore…
2: Health code policies	Health code policies	Pop-ups; Capital Region; Delayed; Complaints; Stranded; Outside of Beijing Health code; Sweeping Code; Assignment Code; Transfer Code; Place Code; Low Risk…
3: Isolation policies	Isolation policies	Quarantine; Residence; Close contact; Square cabin; Community; Positive; Hotel; Negative; Entry; Stay-at home; Policy; Transfer; Red code; Isolation place; Isolation environment…
4: Information disclosure	Information content	Social side; Announcement; Cases; Notification; Specific tracks chain; Details Explanation causes; Risk sites; Crowd places; Disclosure; Science Propaganda; New case; Screening; Zero-covid; Isolation observation; Cured discharged asymptomatic infected; Imported from abroad; Local…
	Information quality	Update time; Delay; Send early; Underreporting; Epidemic data; Accurate; Fake; Statistics; Publish; Reference; Notification; Number of cases; Rumor; …
5: Livelihood guarantee	Livelihood guarantee	Delivery service; Recovery; Neighborhood; Return; Life support; Dispatch; Delivery; Go to work; Vegetables; Cross-district; Public transport; Subway; New Year…
6: Vaccination	Vaccination	Vaccine; Vaccination; Fourth infection inhalation; Virus; Urgent; Elderly; Seriously ill; Appointment…

Table 3. Thesaurus for the orderly relaxation stage of epidemic control (partial)

Topic	Aspect	Keywords
0. Health code policies	Health code policies	Pop-ups; Cancel; Go to Beijing; Optimize; Abolish; Train Tickets; Ten measures; Violation; Return to Beijing; Complaint; Health code; Place code…
1. Anti-epidemic policies	Anti-epidemic policies	Releases; Open policy; Epidemic prevention and control measures; Science; Freedom; Economy; Local implementation; Pilot; Intervention; Public opinion; National; Dynamic Zero-COVID; One size fits all; Masks order…
	Government work	Government; Leadership; Exemplary; Take charge; Public concerns; Comment ;12345 hotline; Solve; traceability…
2. Nucleic acid policies	Necessity of nucleic acid testing	Nucleic acid test; Public places; Hospitals; Work places; Full staff; Certification; Checking; Negative...
	Convenience of nucleic acid testing	Nucleic acid sites；Testing sites; Queuing; Single tube; Mixed tube; Mixed collection; Mixed test; Cost...
3. Medical resource guarantee	Drug guarantee	Antigens; Self-test; Anti-fever medicine; Ordinary people; Testing; Fever; Confirmation; Reagents...
	Medical care guarantee	Hospital; Fever; Outpatient; Beds; Patients; Square Cabin; Critical Care…

4. Personal health protection	Personal health protection	Health; First responder; Protection; Proactive consultation; Symptoms; Infection; Aftereffects; Appointment...
	Health protection of special population groups	Widows; Orphans; Seniors; Children; Infants; Pregnant; Maternal; Critical Illness; Medical staff...
5. Information disclosure	Information quality	Official words; Set words; Truth; Data; Meaning; Social aspects; Reference; Accuracy; Report...
	Information content	Tracking; Key sites; involved in the epidemic sites; Deaths; New case; Confirmed case; critical illnesses; High-risk areas...
6. Vaccination	Vaccination	Vaccine; Vaccination; Voluntary; Effectiveness; Inactivation; Urgent; Import; Inhalation...
7. Livelihood guarantee	Livelihood guarantee	Express delivery; logistics; takeaway; delivery; back to school; back to work; commuting; Celebrate the New Year on the spot...

In the orderly relaxation stage of epidemic control, eight topics, thirteen aspects and their keywords were finally obtained, and the results are shown in Table 3.

The thesaurus results showed that the two stages had commonalities and differences in terms of the concerns of netizens. In terms of commonality, netizens' concerns were consistently expressed during the epidemic management process, such as feedback on anti-epidemic policies, government work, livelihood guarantee, necessity and convenience of nucleic acid testing, health code policies, information content and quality, and vaccination. In terms of differences, aspects related to specific control measures occurred in the concerns of netizens in the control phase, such as grassroots management, lockdown scope, and isolation policies. However, these aspects were no longer widely considered and discussed in the liberalization stage. Instead, aspects more relevant to the new anti-epidemic situation occurred, such as drug guarantee, medical care guarantee, personal health protection, and health protection of special population groups.

4. Aspect-Level Sentiment Analysis of Microblog Comments

4.1 The BERT-Pair-ABSA Model for Semantic Expansion of Auxiliary Sentences

Since the pretraining of the existing BERT model mainly relied on the corpus data of Wikipedia, which deviated from the context and experience concerning the anti-epidemic policies of this paper, this seriously affected the understanding of the BERT model to recognize the specific task of this paper. Given the availability of comment data related to anti-epidemic policies, the amount of task-related corpus available for this paper did meet the magnitude of the post-training of the BERT model. Therefore, on the basis of transforming the ABSA task into a sentence pair classification task by constructing auxiliary sentences, the form of auxiliary sentences was further improved in the aspect-level sentiment analysis of this paper. Instead of directly using simple aspect words with high generality, this paper utilized the aspect words corresponding to keyword sequences to extend it.

FOR example, the simple aspect word "government work" in the dynamic zero-COVID stage was replaced with the corresponding keyword sequence of "government leadership, release, screening, implementation, enforcement, disposal, punishment, punishment, severe punishment, punch card, spill, escape, clothes delivery, livelihood, hotline 12345 hat implementation coordination release notification districts citizens". In the specific context of epidemic control, the regulation and treatment of "clothes delivery", "punch card" and "epidemic escape" were also part of the government work and had strong semantic relevance to it. However, if such associations were not included in the pretraining and no additional information was added in the task-specific fine-tuning, the BERT model easily ignored the context-specific associations or the implicit expressions in the language, which eventually led to biased aspect-level sentiment analysis.

Therefore, on the one hand, the Bert-pair-ABSA model with extended auxiliary sentence semantics proposed by this paper made use of the inherent advantage of "next sentence prediction" in BERT by constructing auxiliary sentences, thus enabling the BERT pretraining model to better match and fit with the task nature of ABSA. On the other hand, more context-specific a priori information and knowledge were introduced into the Bert model, by using keyword sequences to expand the semantics of auxiliary sentences, thus helping the BERT pre-training model to focus on the textual information related to specific aspects more comprehensively and precisely, which was conducive to capturing context-specific associations and implicit expressions in the text more effectively. The specific flowchart is shown in Figure 2.

As shown in Figure 2, the BERT-pair-ABSA model for semantic extension of auxiliary sentences in this paper contains three major modules: sentence-to-task transformation, downstream fine-tuning of the BERT model, and prediction using the fine-tuning model.
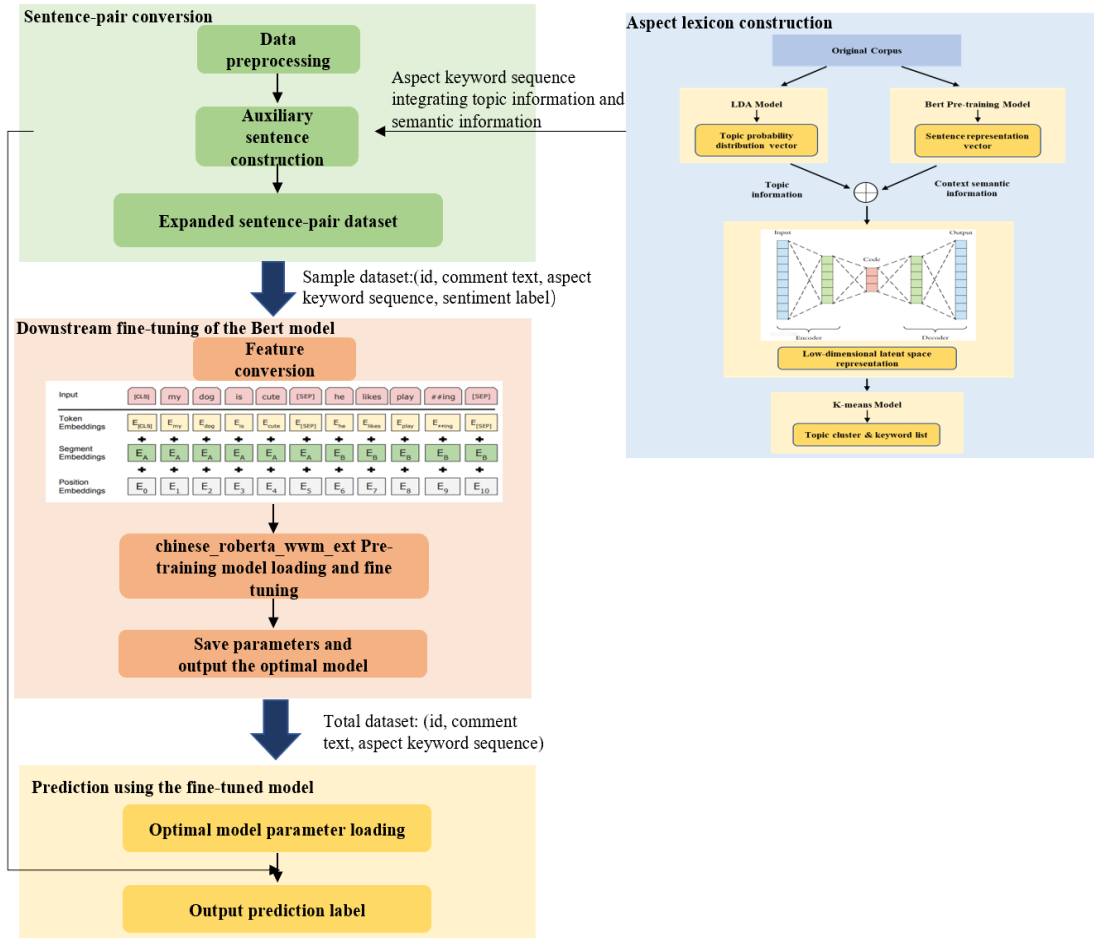
Figure 2. Flowchart of the BERT-pair-ABSA model for semantic expansion of auxiliary sentences

This paper took comment data as an example to show the processing process of the model on the data. Due to the limited space, the auxiliary sentences in the data shown in this paper were represented by aspect words. While the auxiliary sentences were expanded to a series of keywords corresponding to the aspect words in the actual experimental process.

(1) The labeled data was organized into the format of Table 4.

Table 4. Example of data labeling format

Id	Text	Opinions
433	The control cannot control, the release of the dare not put! The most trouble is to draw the circle large enough, a cut, is the most secure way to keep the hat on the head. What is the use of this to you?	[{"aspect": "Epidemic policy", "sentiment": "negative"}, {"aspect": "Sealing and control scope", "sentiment": "negative"}, {"aspect": "Government work", "sentiment": "negative"}]

(2) According to the aspect words and their keyword sequences obtained from the aspect word lexicon construction module in section 3.1, the highly generalized aspect was replaced by the keyword sequence. Due to the limited space, the auxiliary sentences in the example were still represented by aspect words, but the auxiliary sentences were expanded into the corresponding keyword sequences in the actual experimental process.

(3) By Json parsing, the aspect in Opinions corresponded to sentence 2 and sentiment corresponded to label, respectively (Table 5).

(4) The unmentioned aspects in the Opinions were added, and the corresponding emotional polarity was marked as None. Sentence 1 and sentence 2 were divided into words. Thus far, the sentence pair conversion module was completed, and the data form of (id, sentence 1, sentence 2, label) was obtained.

(5) The BERT model further converted the data of (id, sentence1, sentence2, label) into the features of (input_ids, segment_ids, input_mask, label_id), which was the final input model.

Table 5. Example of Jason parsing results

Id	Sentence1	Sentence2	Label
433	The control cannot control, the release of the dare not put! The most trouble is to draw the circle large enough, a cut, is the most secure way to keep the hat on the head. What is the use of this to you?	Outbreak Policy	Negative
433	The control cannot control, the release of the dare not put! The most trouble is to draw the circle large enough, a cut, is the most secure way to keep the hat on the head. What is the use of this to you?	Seal control range	Negative
433	The control cannot control, the release of the dare not put! The most trouble is to draw the circle large enough, a cut, is the most secure way to keep the hat on the head. What is the use of this to you?	Government wotk	Negative

4.2 Experimental Validation and Analysis

4.2.1 Sample dataset

A total of 1,400 data points, randomly selected from each of the 10,000+ data points in both stages, were manually labeled as the sample dataset. Of these data points, 1,000 were used as the training set, 200 as the validation set, and 200 as the test set. Through the construction of auxiliary sentences, each piece of data in the dataset was eventually expanded to the same number of entries as the number of aspects, achieving exponential expansion of the sample dataset. Therefore, after preprocessing, the size of the training set in the dynamic zero-COVID stage was 12,000, and the size of both the validation and test sets was 2,400; the size of the training set in the orderly relaxation stage of epidemic control was 13,000, and the size of both the validation and test sets was 2,600.

4.2.2 Model superparameter setting and evaluation results
• Parameter setting
Fine tuning was performed using the pretrained English_roberta_wwm_ext model (Table 6).

Table 6. Parameter setting

Hyperparameters	Set value
Number of hidden layers	768
Self-attended head count	12
Dropout	0.1
Learning Rate	2e-5
Epochs	8
Batch Size	24

• Solution of category distribution imbalance
Processing of constructing auxiliary sentences exponentially expanded the dataset. But it also caused serious imbalance in the distribution of categories, which led to the model classification favoring categories with large sample data and seriously affected the model performance. Therefore, this paper addressed the imbalance of sample distribution in the task at the loss function level by using Focal_loss instead of the cross-entropy loss function. focal_loss started from the perspective of easy sample classification, which reduced the weights of easily classified samples and increased the weights of hard-to-classify samples. Since intuitively sample imbalance made it more difficult to classify the categories with small sample sizes, the easy classification of samples led to the loss focusing on difficult samples, thus solving the problem of low accuracy of classification of categories with small samples.

• Evaluation metrics
As a common measure of classifier performance, accuracy (Acc) was calculated in a binary classification task as follows:

$$Acc = (TP + TN) / (TP + TN + FP + FN) \tag{7}$$

In a multiclassification task, accuracy is the ratio of the total number of correctly classified samples to the total number of samples and measures the proportion of correct predictions made by the model on the dataset.

- Experimental results

The sentence pair model without context-specific semantic expansion was used as the baseline model to verify the improvement of the model proposed in this paper. The results are shown in Table 7.

Table 7. Results of the aspect-level sentiment classification model

Models	Stage	Dataset	Loss	Acc
Bert_pair_base	Dynamic Zero-COVID	Validation set	0.030	0.866
		Test set	0.020	0.869
	Orderly Release	Validation set	0.013	0.881
		Test set	0.021	0.908
Bert_pair_auxiliary_clause_extension	Dynamic Zero-COVID	Validation set	0.021	0.908
		Test set	0.017	0.905
	Orderly Release	Validation set	0.016	0.908
		Test set	0.020	0.921

4.2.3 Presentation and analysis of model results
- Dynamic zero-COVID stage

Using the official announcement date of the twenty measures as the key division point, the first substage ranged "from the announcement of the ninth edition of the program to the announcement of the twenty measures", and the period "after the announcement of the twenty measures" was the second substage. This paper explored the impact of the promulgation of the two anti-epidemic policies on the emotional feedback of netizens.

According to the obtained statistical results, the distribution of netizens' attention hotspots is organized in Figure 3 and Figure 4.
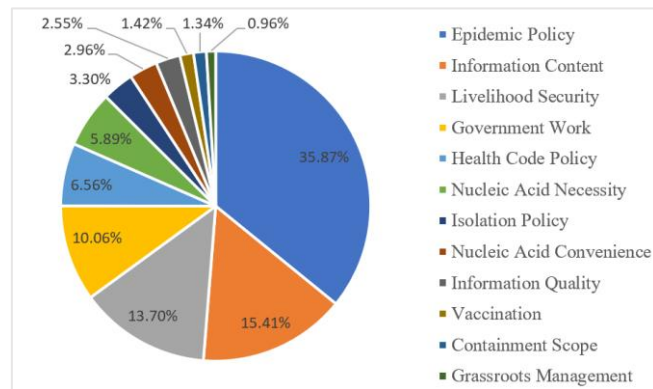
Figure 3. Distribution of netizens' concern in the first substage
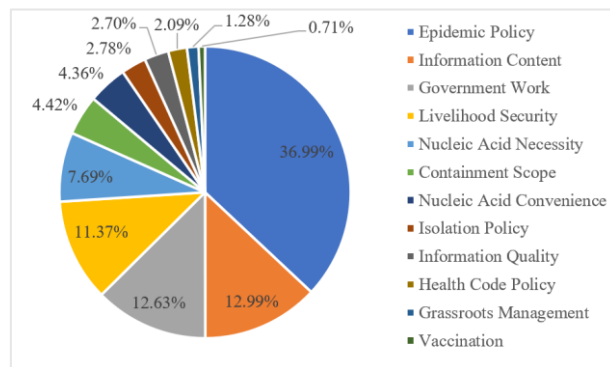
Figure 4. Distribution of hotspots of netizens' concern in the second substage

First, in terms of commonality, the above-mentioned results showed that four major aspects, namely, "anti-epidemic policies", "information content", "livelihood guarantee" and "government work", accounted for approximately 75% of the feedback content in the first and second substages, which were the main concern of netizens during the dynamic zero-COVID stage of the COVID-19.

Second, in terms of differences, the two substages were again different in the trend of discussion heat and the sentiment evolution in various aspects. In general, the twenty measures elicited significantly more discussion heat

than the ninth version of the program. As seen from the display results, the first substage caused 2,395 opinion discussions, while the second substage caused 9,490. The ninth edition of the program made three revisions to the quarantine time limit and mode, the criteria for dividing risk zones and epidemic surveillance. Based on the new changes in the mutation and transmission pattern of the COVID-19, the twenty measures were revised, optimized and adjusted in several areas, such as reducing the anti-epidemic cost, keeping the exchanges and expanding the openness, strengthening prediction and stratification, and emphasizing people's livelihood, thus meeting the new requirements. Although both measures realized the "taking small steps and not stopping" policy, the twenty measures were more characterized by small and quick steps, and responded to the outstanding problems existing in the past epidemic prevention and control process and people's concerns, thus causing more intense social discussions and emotional expressions.

In addition, the trends of public opinion in various aspects showed the following characteristics: (1) The trends of the evolution of netizens' sentiment were highly consistent in "anti-epidemic policies" and "government work". (2) The peaks of feedback and sentiment of "anti-epidemic policies", "government work", "livelihood guarantee" and "grassroots management" were the same as that of "government work". The peaks were closely related to the regional lockdown and control measures. As shown in Figures 5-8, the two obvious peaks occurred at the end of August 2022 (implementation of regional control measures in Shijiazhuang) and after mid-November 2022 (implementation of regional control measures in Guangzhou). Both "anti-epidemic policies" and "government work" promptly elicited positive sentiment after the promulgation of the twenty measures, but stronger negative sentiment peaks also emerged subsequently. The percentages of opinions on "lockdown scope", "necessity of nucleic acid testing", and "convenience of nucleic acid testing" all increased significantly after the promulgation of the twenty measures. The proportion of opinions on "information content" remained high before and after the promulgation of the two measures, and the polarity of social sentiment was mainly neutral (Figure 13).

Based on the above characteristics, we can see two things. First, as decision makers, implementers and supervisors, government departments had a closed loop workflow, ranging from scientific formulation of decisions, efficient implementation, thoroughness of implementation, smooth reflection of public voices, reguidance to optimization to further improve the rationality of decisions, which had a real impact on the effective implementation of anti-epidemic work. Therefore, the opinions and emotions of netizens on "anti-epidemic policies" basically focused on their satisfaction with or opinions and suggestions for "government work", thus making the two show an in the process of implementing the regional preventive and control measures. Netizens had heated discussions in these two aspects. In addition, as the last link of policy implementation, the "grassroots management" was lagging, simplistic and one-size-fits-all, with excessive measures taken, which seriously affected the normal life of the public and easily aroused the opposition of netizens. At the same time, the regional lockdown measures influenced people's normal production and life. As the lockdown duration became longer, livelihood guarantee issues, such as food, clothing, housing and transportation, also triggered increasingly more negative emotions, such as anxiety, worries and complaints from netizens (Figures 5-8).

Second, at the beginning of the announcement of the twenty measures, it can be seen that the positive sentiment of netizens toward "anti-epidemic policies" and "government work" increased significantly. This indicated that the twenty measures responded to the concerns of public opinion and achieved remarkable results in following public opinion and stabilizing people's minds. However, it also meant that netizens had greater expectations and demands for the actual implementation of subsequent policies, which was evidenced by the sentiment evolution trend in the Guangzhou epidemic lockdown incident that immediately followed. The twenty measures emphasized optimizing the scope of lockdown and nucleic acid testing to minimize the number of personnel in lockdown and standardize nucleic acid testing practices, which led to significantly increased social attention in the "lockdown scope" and "necessity of nucleic acid testing" in the Guangzhou lockdown incident two days after the official announcement of the measures. When the epidemic prevention and control measures failed to meet netizens' expectations, which led to their disappointment with the implementation and execution of the policies, the netizens had stronger emotional changes in the anti-epidemic policies, government credibility, and the scope of lockdown and nucleic acid testing emphasized in the twenty measures (Figure 9-11).
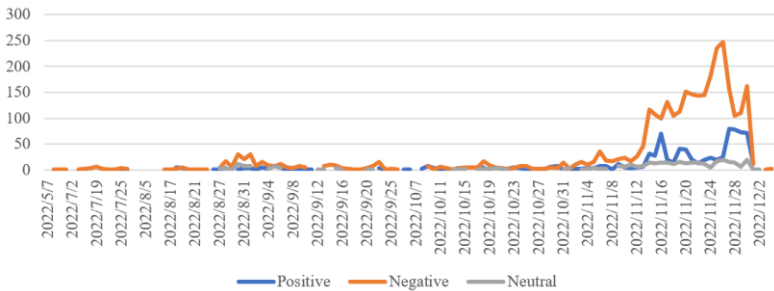

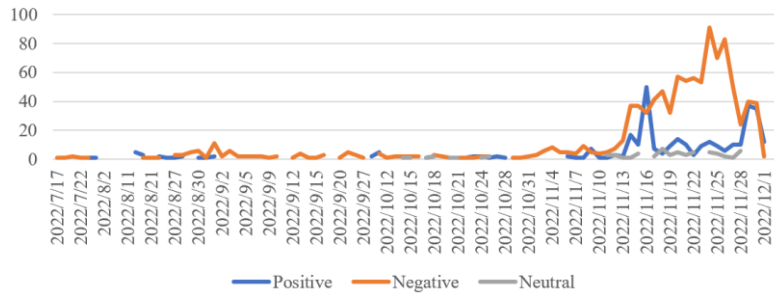**Figure 5.** Trend of sentiment evolution of anti-epidemic policies during the dynamic zero-COVID stage

Figure 6. Trend of sentiment evolution of government work during the dynamic zero-COVID stage
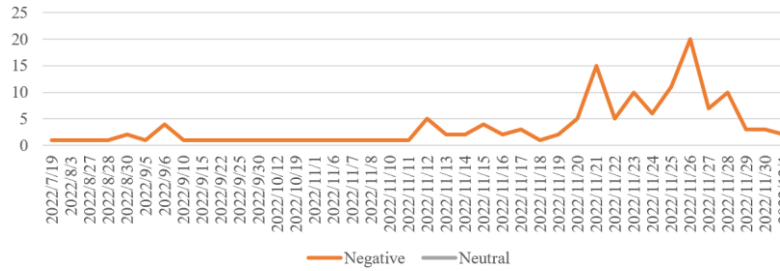

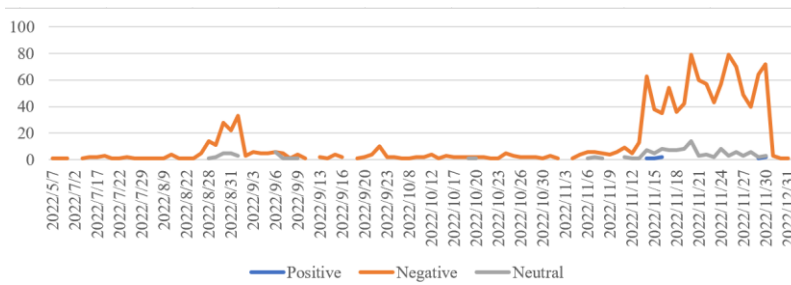**Figure 7.** Trend of sentiment evolution of grassroots management during the dynamic zero-COVID stage

Figure 8. Trend of sentiment evolution of livelihood guarantee during the dynamic zero-COVID stage
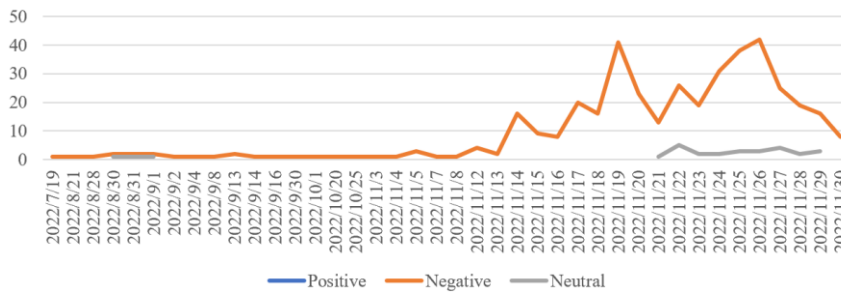

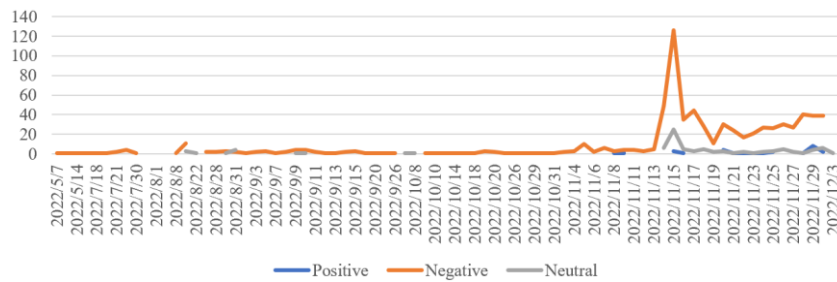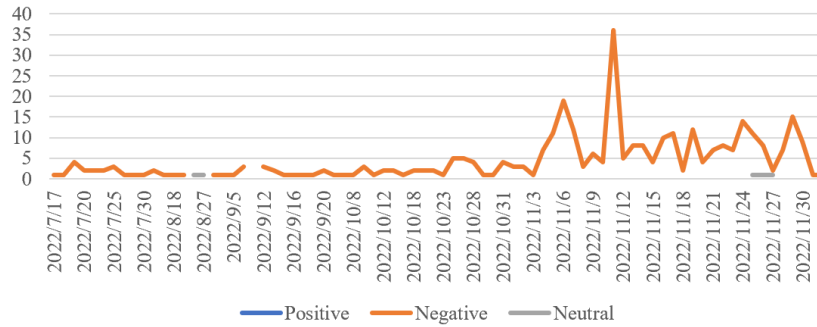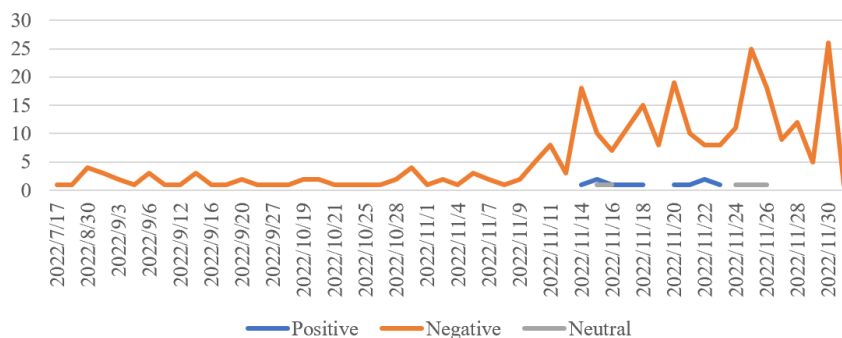**Figure 9.** Trend of sentiment evolution of lockdown scope during the dynamic zero-COVID stage

Figure 10. Trend of sentiment evolution of the necessity of nucleic acid testing in the dynamic zero-COVID stage
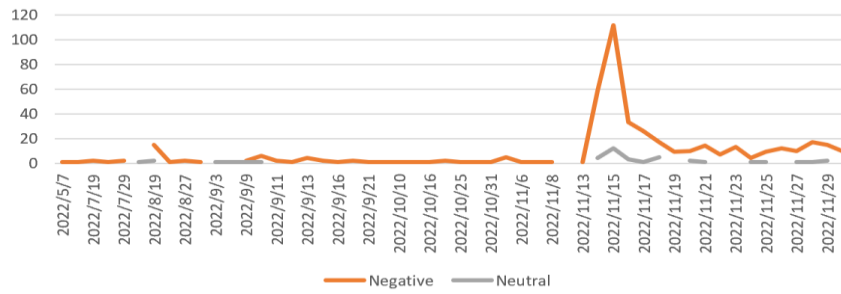
Figure 11. Trend of sentiment evolution in convenience of nucleic acid testing during the dynamic zero-COVID stage

Figure 12. Trend of sentiment evolution of health code policies during the dynamic zero-COVID stage
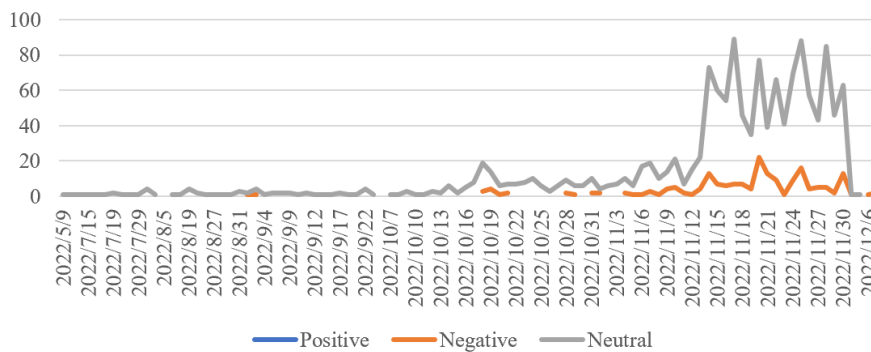
Figure 13. Trend of sentiment evolution of information content in the dynamic zero-COVID stage

Figure 14. Trend of sentiment evolution of information quality during the dynamic zero-COVID stage

In addition, after the twenty measures were announced, netizens also expressed strong negative emotions about the "health code policies", which was not mentioned in the measures, mainly because the measures did not notice and solve several problems, such as abnormal pop-ups, restrictions of pop-ups on normal travel needs, difficulties in the appeal process, inefficiency or ineffectiveness, etc., which resulted in the intensification of negative emotions. Therefore, managers should pay attention to and respond to this type of views in a timely manner and

further add and improve the perspective of policy optimization (Figure 12).

The content of epidemic information disclosure was always a hot topic of concern for the public, and the emotional feedback from netizens was dominated by inquiries about specific epidemic information in neutral emotions. When the official information disclosure was not timely and effective enough, which lagged behind or seriously mismatched the information needs of netizens, the information disclosure work also lost its meaning of establishing epidemic risk expectations and relieving people's anxiety, which also led to an increase in negative social sentiment (Figure 13 and Figure 14).

Based on the above discussion, this paper summarized two management insights. On the one hand, managers need to not only pay attention to the improvement of the epidemic information release mechanism to meet the basic information needs of netizens, but also focus on improving their own ability to communicate with the public when major policies are announced by describing policies in a clear and logical manner, without ambiguity or obscurity and supplemented by illustrations, conferences, etc. Timely and efficient policy interpretation in diverse forms are needed to minimize the deviation of understanding of major policies and the resulting deviation of expectations of netizens, thus reducing the generation of negative emotions at the source. In the face of major policies or events that have aroused high-degree public concern, managers cannot be indifferent or cover up and should be the first to make a clear response and explanation, using science to explain doubts, thus ensuring the bridge between the government and the public. Only in this way can they continue to strengthen the bond of trust between the government and the public and avoid the continuous intensification of negative public opinion. On the other hand, attention should be paid to improving the quality and efficiency of the spatial diffusion of the Party Central Committee's decisions. First, the diffusion path from the central-local-grassroots level should be analyzed on a case-by-case basis and the thoughts and actions of management departments at all levels should be unified, thus ensuring that the policy diffusion does not go out of shape or run out of tune. Second, a scientific and efficient policy learning plan should be developed through scheduling, training, issuance of supporting guidelines and other forms, and the implementation of policies at all levels of should be promoted quickly, thus avoiding confusion in epidemic disposal due to inconsistent management and measures of various departments, as well as paralysis and slackness due to waiting of managers, and ineffective policy implementation.

- Orderly relaxation stage of epidemic control (November 2022-January 2023)

Based on three events, namely, the orderly relaxation of epidemic control in both Shijiazhuang and Beijing, and the announcement of the new ten measures, this paper tried to explore the impact of the implementation of anti-epidemic policies on the emotional feedback of netizens in this stage. The distribution of netizens' concern is shown in Figure 15.

First, the above results showed that opinions on "anti-epidemic policies" and "government work" still dominated the comments on anti-epidemic microblogs like the dynamic zero-COVID stage. However, differences existed as follows: (1) The peaks of "anti-epidemic policies" and "government work" were extremely high, which appeared and disappeared quickly with the announcement of policies related to the hard landing of the epidemic; (2) The emotional polarity of "anti-epidemic policies" and "government work" was mainly positive (Figure 16 and Figure 17); (3) Considering the difference between "being outside" and "being inside", this paper further divided netizens into two types, namely, netizens inside and outside Hebei, and netizens inside and outside Beijing, by using the IP addresses of comments obtained, and explored whether obvious regional differences existed in the positive sentiment toward the liberalization policy and government work in corresponding events. Although netizens outside the pilot regions had significantly stronger positive sentiment toward the "anti-epidemic policies" and "government work", netizens in Hebei and Beijing alone had mixed sentiment, with positive sentiment more than negative sentiment. Thus, in this new stage of anti-epidemic prevention and control, the announcement of the orderly relaxation policies played an important role in responding to public opinion and inspiring people.
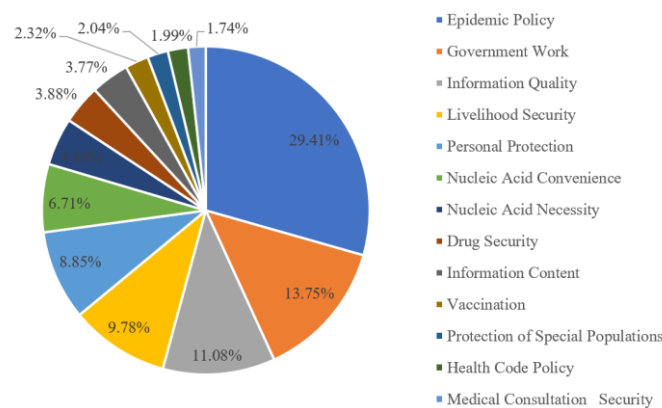

Figure 15. Distribution of netizens' concern during the orderly relaxation stage of epidemic control
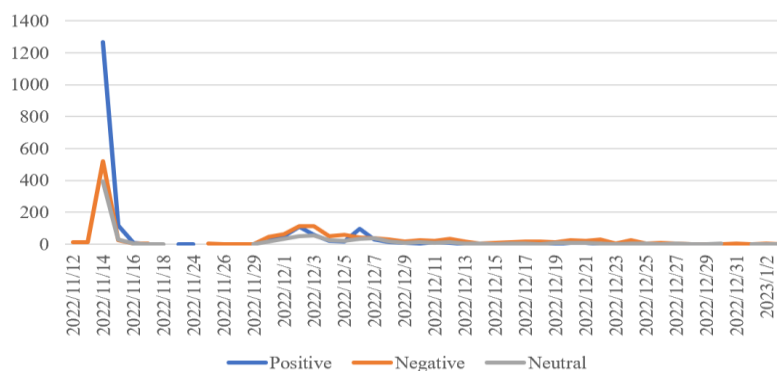
Figure 16. Trend of sentiment evolution of anti-epidemic policies during the orderly relaxation stage of epidemic control
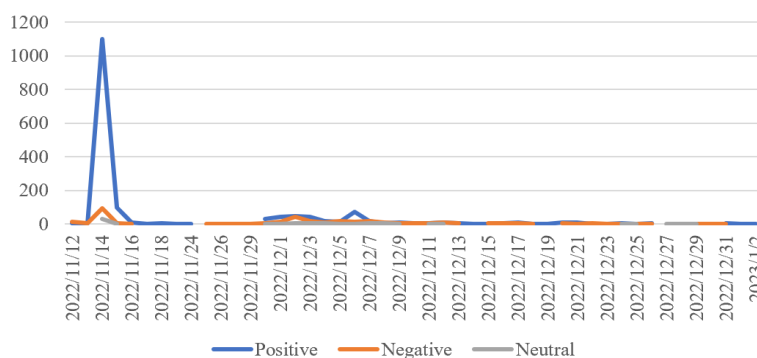

Figure 17. Trend of sentiment evolution of government work during the orderly relaxation stage of epidemic control

Second, the trend of sentiment evolution was also closely related to the announcement of the epidemic liberalization policy in three aspects: "personal protection", "protection of special population groups" and "livelihood guarantee". Despite the fear and resistance to virus infection, based on the deepening understanding of the new characteristics and laws of virus mutation and transmission, as well as the uplifting and encouraging effect of anti-epidemic policies that changed over time and in different situations, netizens mostly showed a strong sense of self-efficacy in personal health protection that "everyone is the first person responsible for his or her own health" after the policy implementation. However, the positive sentiment of netizens fell back along with the extention of the observation period, and their negative sentiment on "personal protection" lasted longer, which indicated that the self-efficacy of netizens on personal protection urgently needed supporting measures and solid resources to truly play a role. In addition, as for special groups, such as elderly individuals, children, pregnant women, and people with basic illnesses, the negative sentiment of netizens on "personal health protection" lasted a few times longer. For the health protection of special groups and health care workers, their vulnerability or special nature was more likely to trigger negative emotions, such as worry and sadness among netizens. Therefore, the policies should not be generalized but required additional more targeted supporting measures and resources from management. In addition, the hard landing of the anti-epidemic policies has aroused the urgent demand for the resumption of a series of normal production and life operations, such as work, production, business, schools, logistics, transportation, commuting, supply, and superstores, which were suspended by the COVID-19, thus making the demand present a complicated and trivial feature. However, the demand was most closely related to the basic concerns of people and needed to be solved most urgently. Therefore, related specific measures significantly affected netizens' perception of the policy value and were important for netizens to measure and judge whether a policy was truly "a governmental response to people's call" or just "official words and empty talk". Before the demand is properly satisfied, the negative emotions arising from it will continue. Government departments need to unify their cooperation and coordination in order to do their best to meet the basic needs of the people more practically and comprehensively while ensuring the planning and orderliness of restoring normal production and life (Figures 18-20).

As mentioned earlier, the call for netizens to be the first person responsible for their personal health required solid resources. Therefore, three aspects of "drug guarantee", "vaccination", and "medical care guarantee" received extensive attention and discussion in the new stage of the orderly relaxation stage of epidemic control, which was characterized by a longer duration of negative emotions than livelihood guarantee until the demand was met and resolved (Figures 21-23).
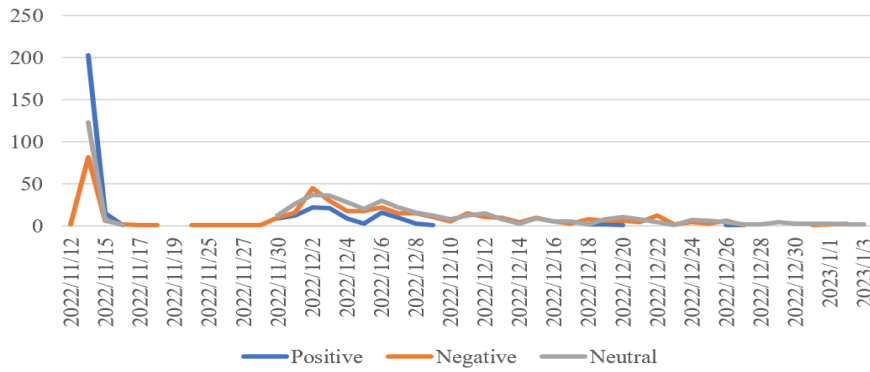
Figure 18. Trend of sentiment evolution of personal health protection in the orderly relaxation stage of epidemic control
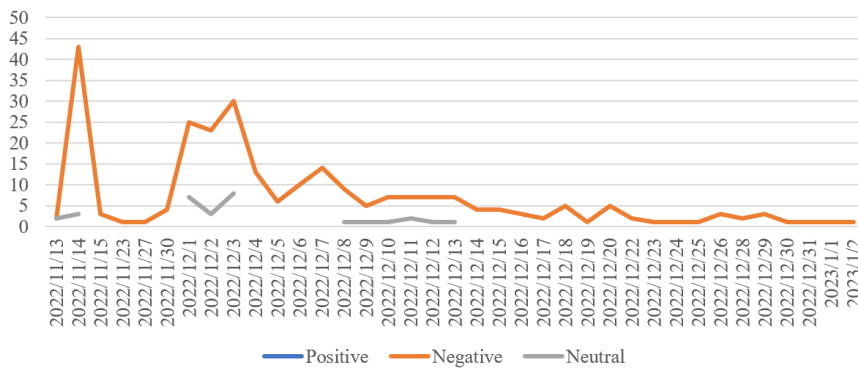

Figure 19. Trend of sentiment evolution of health protection of special population groups in the orderly relaxation stage of epidemic control
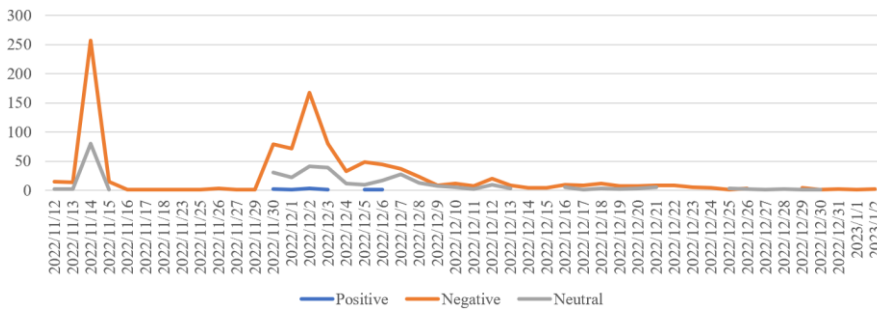

Figure 20. Trend of sentiment evolution of livelihood guarantee during the orderly relaxation stage of epidemic control
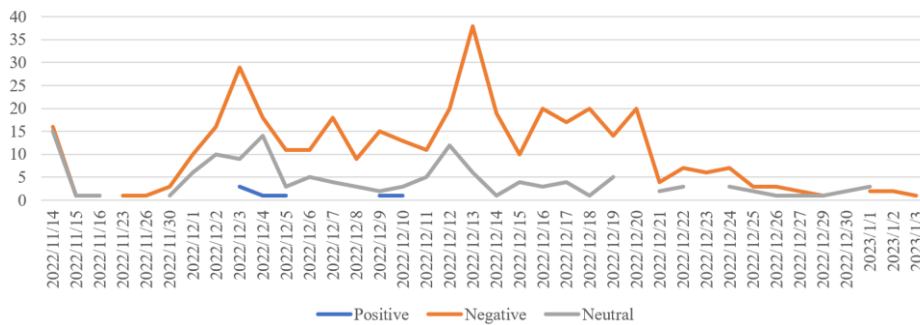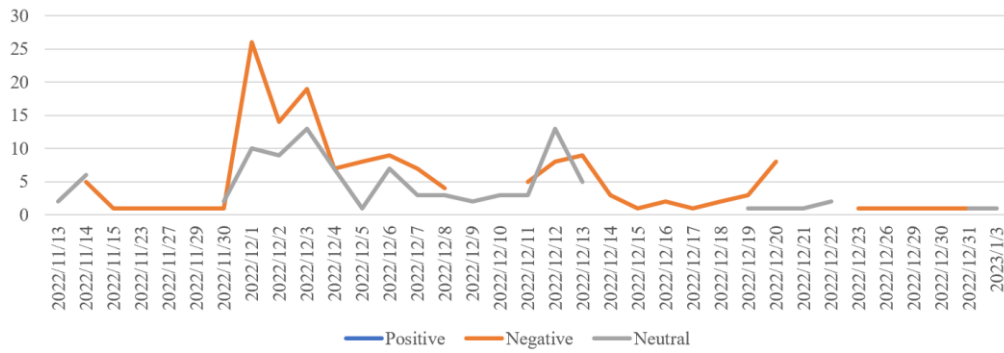

Figure 21. Trend of sentiment evolution of drug guarantee during the orderly relaxation stage of epidemic control

Figure 22. Trend of sentiment evolution of medical care guarantee during the orderly relaxation stage of epidemic control
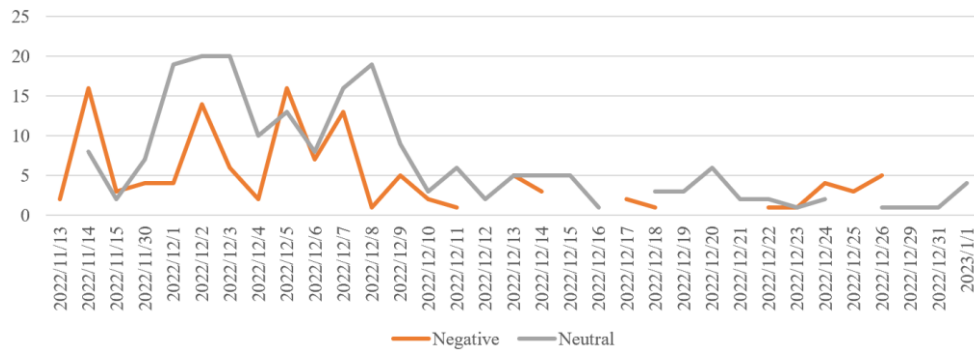

Figure 23. Trend of sentiment evolution of vaccination during the orderly relaxation stage of epidemic control
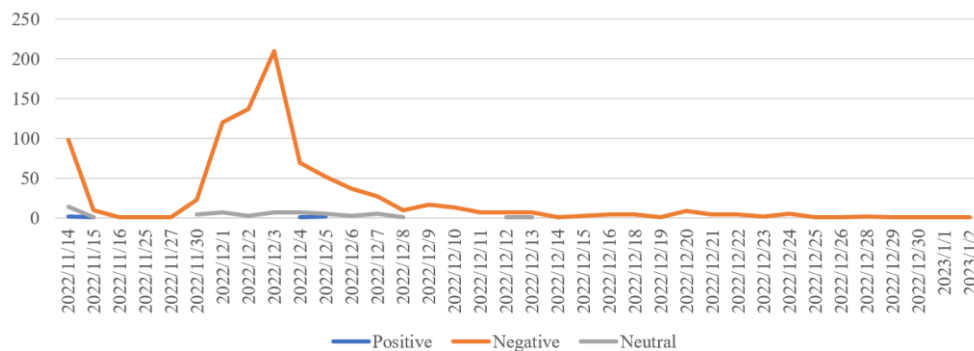

Figure 24. Trend of sentiment evolution of convenience of nucleic acid testing during the orderly relaxation stage of epidemic control
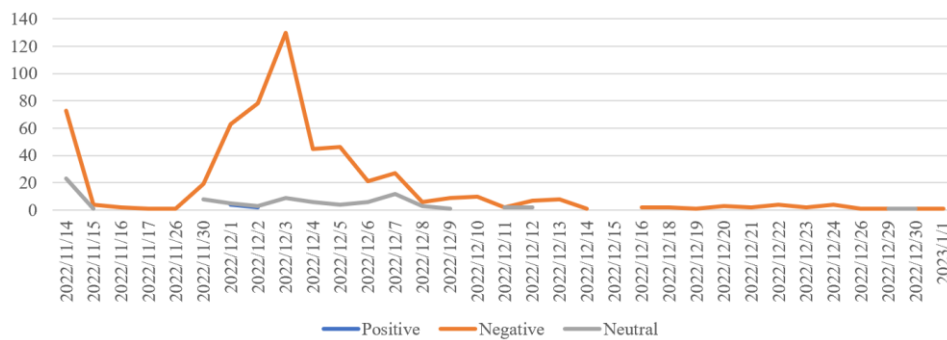

Figure 25. Trend of sentiment evolution of the necessity of nucleic acid testing during the orderly relaxation stage of epidemic control
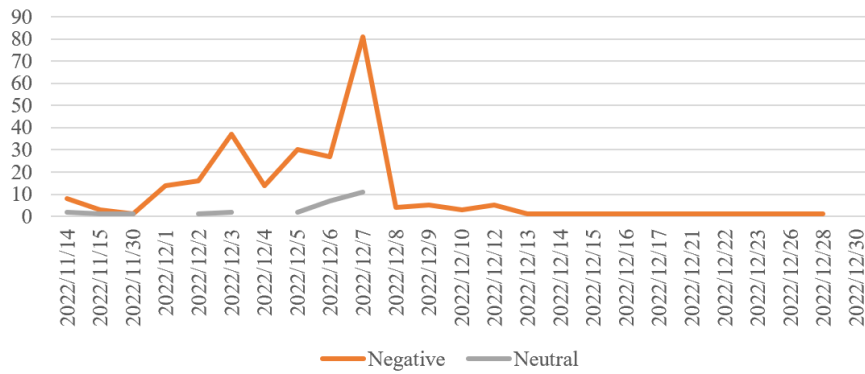
Figure 26. Trend of sentiment evolution of health code policies during the orderly relaxation stage of epidemic control
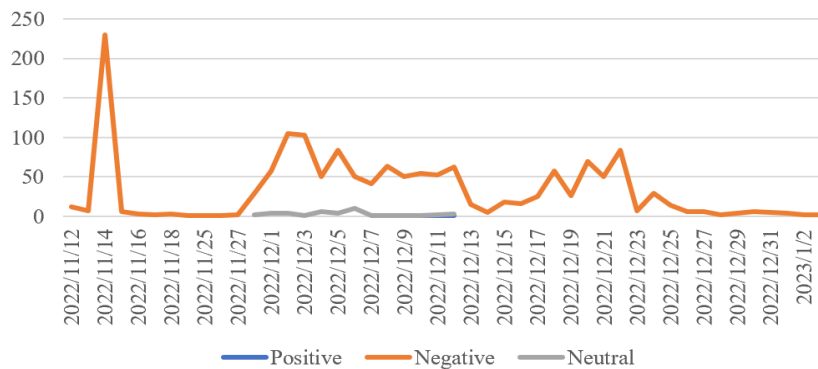

Figure 27. Trends of sentiment evolution of information quality in the orderly relaxation stage of epidemic control
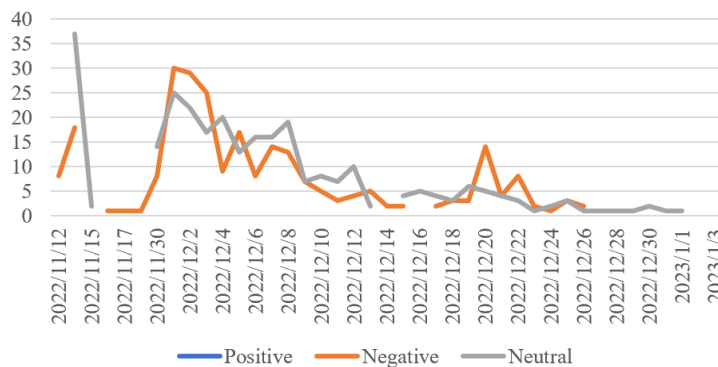

Figure 28. Trend of sentiment evolution of information content in the orderly relaxation stage of epidemic control

In addition, convenient sampling points were established during the normalization period of epidemic prevention and control to stimulate new momentum of the public for nucleic acid testing, which opened up the "last mile" of nucleic acid testing and built up a 15-minute sampling service circle for the public. Since the scope of nucleic acid testing was regulated and optimized by the twenty measures, the social sampling points began to be removed one after another, and the discussion of "convenience of nucleic acid testing" gradually increased. The mass removal of social nucleic acid testing points was even more important till the anti-epidemic policies were relaxed. The discussion of "convenience of nucleic acid testing" was pushed to a climax when the anti-epidemic policies were relaxed and the large-scale nucleic acid testing sites in the community were removed. The removal and closure of testing sites was not the crux of the issue. As shown in Figures 24-25, the trends of "convenience of nucleic acid testing" and "necessity of nucleic acid testing" were highly consistent, because many scenes or places in daily life still required nucleic acid testing results despite the rapid withdrawal of nucleic acid testing sites. The serious disconnect between the two policies caused much inconvenience to the public and aroused strong

discontent among netizens. On December 7, 2022, as a positive response to the outstanding problems strongly reflected by the masses in the ninth edition of the program and the twenty measures that were not in place and inaccurate, the new ten measures were officially released, which explicitly mentioned the optimization of the nucleic acid testing and health code inspection sites. It can be seen in Figure 26 that the landing of this supplementary policy has extremely effectively alleviated the negative emotions of these aspects. This also revealed that policy makers needed to pay more attention to the synchronization and coordination between policies with "complementary relationship" when adjusting policies, thus avoiding the disconnect and conflict between policies to intensify the negative emotions of society.

Finally, in terms of information disclosure, due to its special nature of frequent updating, the evolution of netizens' sentiment related to it was no longer limited to the period before and after the announcement of the anti-epidemic policies. In addition, compared with the dynamic zero-COVID stage, the trends of sentiment evolution of netizens showed distinct differences regarding "information quality" and "information content" in this stage. First, netizens' concern and negative sentiment in terms of "information quality" increased significantly. With the withdrawal of social nucleic acid testing sites and the rise of antigen testing, the statistical data related to the epidemic was bound to be no longer true and accurate. While the public was always informed of wrong data without reference value at this time, which made relevant management departments form a negative image of rigid work and waste time on meaningless matters in the minds of netizens, thus causing strong resentment among them. Second, the negative emotion of "information content" also increased significantly. Based on the theory of emotion cognition, information is an important trigger of emotion generation, and negative emotion generates when information on the event cannot meet the necessary information needs held by the organism based on past memories and future expectations. Under the new situation of epidemic control relaxation, netizens had a series of new information needs, such as the location of existing nucleic acid testing points, critical care, available medical resources, etc. At this time, relevant departments did not make timely changes and flexible adjustments, or they still rigidly disclosed the information that netizens were concerned about in the dynamic zero-COVID stage (Figure 27 and Figure 28).

5. Conclusion

This paper integrated LDA topic information and BERT semantic information with K-Means clustering using a topic clustering extension model, extracted topics from the comment data in two stages, constructed a lexicon of aspect words, and realized aspect-level sentiment analysis of anti-epidemic policy comments using the BERT-pair-ABSA model of auxiliary sentence semantic expansion. In addition, this paper explored the influence of the anti-epidemic policies on the attention hotspots of netizens in the different stages of dynamic zero-COVID and orderly relaxation of epidemic control, as well as the characteristics of the sentiment feedback and sentiment evolution trends of different aspects. It was found that (1) The anti-epidemic policies and government work were the most important discussion contents in both stages. However, the negative feedback was overwhelming in the dynamic zero-COVID stage, while the proportion of positive sentiment exceeded that of negative sentiment in the orderly relaxation stage of epidemic control. Meanwhile, with the adjustment and optimization of the anti-epidemic policies, the positive sentiment in both aspects showed a short peak followed by a gradual disappearance or a stronger trend. (2) In the dynamic zero-COVID stage, netizens paid attention to specific anti-epidemic measures, such as grassroots management and lockdown scope, and their concern and negative sentiment peaked with the implementation of regional lockdown measures. (3) During the orderly relaxation stage of epidemic control, netizens were concerned about drug guarantee, medical care guarantee, personal health protection and health protection of special groups, etc. Negative sentiment always dominated in drug guarantee, medical care guarantee and health protection of special groups. (4) With the rapid removal of nucleic acid testing sites, a variety of scenarios in daily life still retained the original nucleic acid testing standards, and the disconnect of the anti-epidemic policies caused much inconvenience to the life of the public, which provoked strong discontent of netizens. This revealed that policy makers should pay attention to the synchronization and coordination between policies with "complementary relationships" when adjusting policies in order to ensure the persistence and sustainability of epidemic prevention measures.

In the aspect-based sentiment analysis, this paper only used the auxiliary sentence construction form of "pseudosentence", and made no further attempts and explorations on the more standardized question and answer sentence form. More diverse and realistic auxiliary sentence construction forms can be used to improve the classification effect of the model in the future. Only positive, negative and neutral categories were used to calculate the polarity of netizens' emotions, but the fine-grained emotion calculation method was not used to explore the influence of different aspect words on fine-grained emotions, such as fear, anger and disgust and their evolution.

Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Funding

References

Ahmad, A. R. & Murad, H. R. (2020). The impact of social media on panic during the COVID-19 pandemic in Iraqi Kurdistan: online questionnaire study. *J. Med Internet Res.*, *22*(5), Article ID: e19556. https://doi.org/10.2196/19556.

Campbell, A. L. (2012). Policy makes mass politics. *Annu Rev. Polit Sci.*, *15*, 333-351. https://doi.org/10.1146/annurev-polisci-012610-135202.

De Landsheer, C. & Walburg, V. (2022). Links between rational and irrational beliefs, trait anxiety and fear of COVID 19. *Psychol. Fr.*, *67*(3), 305-316. https://doi.org/10.1016/j.psfr.2022.06.005.

Dillard, A. J. & Meier, B. P. (2023). Getting COVID-19: Anticipated negative emotions are worse than experienced negative emotions. *Soc Sci. Med.*, *320*, Article ID: 115723. https://doi.org/10.1016/j.socscimed.2023.115723.

Du, C. Y. & Liu, P. Y. (2020). Helical attention networks for aspect-level sentiment classification. *J. Chin Inf. Process.*, *34*(9), 70-77. http://dx.chinadoi.cn/10.3969/j.issn.1003-0077.2020.09.010.

Han, X., Wang, J., Zhang, M., & Wang, X. (2020). Using social media to mine and analyze public opinion related to COVID-19 in China. *Int J. Environ Res. Public Health*, *17*(8), Article ID: 2788. https://doi.org/10.3390/ijerph17082788.

Inam, S. Y., Rehman, A., & Gulshan, T. (2022). Effects of Covid-19 pandemic on the working performance of journalists: Perception of swat based journalists. *Hum Nat. J. Soc Sci.*, *3*(2), 42-60.

Jang, H., Rempel, E., Roth, D., Carenini, G., & Janjua, N. Z. (2021). Tracking COVID-19 discourse on twitter in North America: Infodemiology study using topic modeling and aspect-based sentiment analysis. *J. Med Internet Res.*, 23(2), Article ID: e25431. https://doi.org/10.2196/25431.

Li, C., Zhou, B., Lin, W., Tang, Z., Tang, Y., Zhang, Y., & Cao, J. (2023). A personalized explainable learner implicit friend recommendation method. *Data Sci. Eng.*, 8, 23-35. https://doi.org/10.1007/s41019-023-00204-z.

Liang, Y., Zhao, Y., Ji, W., Bian, X., & Xi, J. (2023). Latent profiles of psychological status among populations cumulatively exposed to a flood and the recurrence of the COVID-19 pandemic in China. *Int J. Disaster Risk Reduct.*, 85, Article ID: 103520. https://doi.org/10.1016/j.ijdrr.2022.103520.

Lin, C. Y., Broström, A., Griffiths, M. D., & Pakpour, A. H. (2020). Investigating mediated effects of fear of COVID-19 and COVID-19 misunderstanding in the association between problematic social media use, psychological distress, and insomnia. *Internet Interv.*, 21, Article ID: 100345. https://doi.org/10.1016/j.invent.2020.100345.

Nazir, A., Rao, Y., Wu, L., & Sun, L. (2020). Issues and challenges of aspect-based sentiment analysis: A comprehensive survey. *IEEE Trans. Affect Comput.*, *13*(2), 845-863. https://doi.org/10.1109/TAFFC.2020.2970399.

Sukhwal, P. C. & Kankanhalli, A. (2022). Determining containment policy impacts on public sentiment during the pandemic using social media data. *Proc. Natl Acad. Sci.*, *119*(19), e21172. https://doi.org/10.1073/pnas.2117292119.

Wen, H. & Zheng, H. (2022). On differences about public sentiment feedback of public policies making in major public health emergencies-An investigation about the policy of 'delay of school opening' based on social media data. *J. Beijing Univ. Technol.*, 22(6), 49-66.

Yu, J., Chen, K., & Xia, R. (2022). Hierarchical interactive multimodal transformer for aspect-based multimodal sentiment analysis. IEEE Transactions on Affective Computing. https://doi.org/10.1109/TAFFC.2022.3171091.

Zhang, W., Li, L., Zhu, Y., Yu, P., & Wen, J. (2022a). CNN-LSTM neural network model for fine-grained negative emotion computing in emergencies. *Alex Eng. J.*, *61*(9), 6755-6767. https://doi.org/10.1016/j.aej.2021.12.022.

Zhang, W., Li, X., Deng, Y., Bing, L., & Lam, W. (2022b). A survey on aspect-based sentiment analysis: tasks, methods, and challenges. IEEE Transactions on Knowledge and Data Engineering, 1-20. https://doi.org/10.1109/TKDE.2022.3230975.

Zhang, W., Wang, M., & Zhu, Y. C. (2020). Does government information release really matter in regulating contagion-evolution of negative emotion during public emergencies? From the perspective of cognitive big data analytics. *Int J. Inf Manage.*, *50*, 498-514. https://doi.org/10.1016/j.ijinfomgt.2019.04.001.

Zhang, W., Yang, X., Tao, K., Zhu, Y., & Zhang, M. (2022c). Influence of media framing on cognition of COVID-19 vaccine effectiveness: The mediating role of emotions. *Am J. Health Behav.*, *46*(4), 423-441. https://doi.org/10.5993/AJHB.46.4.4.