Emotion Analysis of COVID-19 Vaccines Based on a Fuzzy Convolutional Neural Network

Dong Qiu¹,²,³ · Yang Yu¹ · Lei Chen¹

Received: 22 December 2021 / Accepted: 16 October 2022
© The Author(s), under exclusive licence to Springer Science+Business Media, LLC, part of Springer Nature 2022

Abstract
COVID-19 created immense global challenges in 2020, and the world will live under its threat indefinitely. Much of the information on social media supported the government in addressing this major public health event. On January 9, to control the virus, the Chinese government announced universal vaccinations. However, due to a range of varied interpretations, people held different attitudes towards vaccination. Therefore, the success of the mass immunization strategy greatly depended on the public perception of the COVID-19 vaccine. This article explores the changes in people’s emotional attitudes towards vaccines and the reasons behind them in the context of the global pandemic in an effort to help mankind overcome this ongoing crisis. For this article, microblogs from January to September containing Chinese people’s responses to the COVID-19 vaccines were collected. Based on fuzzy logic and deep learning, we advance the hypothesis that fuzzy vector adaptive improvements will make it possible to better express language emotion and that fuzzy emotion vectors can be integrated into deep learning models, thus making these models more interpretable. Based on this assumption, we design a deep learning model with a fuzzy emotion vector. The experimental results show the positive effect of this model. By applying the model in analyses of people’s attitudes towards vaccines, we can obtain people’s attitudes towards vaccines in different time periods. We discovered that the most negative emotions about the vaccine appeared in April and that the most positive emotions about the vaccine appeared in February. Combined with word cloud technology and the LDA model, we can effectively explore the reasons for the changes in vaccine attitudes. Our findings show that people’s negative emotions about the vaccine are always higher than their positive emotions about the vaccine and that people’s attitudes towards the vaccine are closely related to the progress of the epidemic. There is also a certain relationship between people’s attitudes towards the vaccine and those towards the vaccination.

Keywords COVID-19 vaccines · Sentiment analysis · Fuzzy logic · Fuzzy convolutional neural network · Fuzzy emotion vector

Introduction
At the end of December 2019, unexplained pneumonia occurred in Wuhan city, Hubei Province, China. The infectious disease ravaged the world in the following year and had a huge impact on China’s health systems and economy [1]. One month after the outbreak, on January 30, China reported 7736 confirmed cases and 12,167 suspected cases, and 82 confirmed cases were identified in 18 other countries [2]. By October 22, 2021, COVID-19 had caused 240 million confirmed deaths worldwide, and the death toll had reached 4.91 million. The mortality rate of critically ill patients with COVID-19 exceeded 16.7% [3]. The Chinese government took a series of measures to control the epidemic, such as requiring people to wear masks, restricting access to cities, and banning parties [4]. However, these policies could not completely contain COVID-19, and the development of a COVID-19 vaccine was critical [5]. Fairly quickly, vaccines produced by Pfizer and Modena were approved for marketing. Many people were still hesitant about the vaccines and expressed doubts about their emergence. A survey showed that only 58% of Europeans are willing to be vaccinated [6]. Previous studies have also shown that people are more likely to refuse vaccinations for newly discovered diseases than for historical diseases [7].

¹ School of Computer Science and Technology, Chongqing University of Posts and Telecommunications, Nanan, Chongqing 400065, China
² College of Science, Chongqing University of Posts and Telecommunications, Nanan, Chongqing 400065, China
³ School of Mathematics and Information Science, Guangxi University, Nanning, China
Due to the existence of vaccine hesitation, even despite the widespread availability of vaccines, it was not possible to ensure that enough people would become vaccinated [8].

Vaccine hesitation refers to the delay in accepting or rejecting vaccination despite the availability of vaccination services. Vaccine hesitation is complex and environment specific [9]. Vaccine hesitation is linked to a reduction in vaccine coverage and an increased risk of vaccine-preventable disease outbreaks and epidemics, and the number of people with vaccine hesitation is increasing [10]. The reasons most often given for refusing vaccines are being against vaccines in general, concerns about safety, thinking that a vaccine produced in a rush is too dangerous, considering the vaccine useless because of the harmless nature of COVID-19, and a general lack of trust [11]. A survey showed that even health care workers were suspicious about vaccines and that women generally refused to become vaccinated [12]. Therefore, it is imperative to evaluate people’s attitudes towards vaccines in the current epidemic normalization period.

The survey [13] shows that to achieve mass immunization, the vaccination rate must reach at least 70%. Therefore, finding ways to guide the masses to actively participate in vaccination and finding ways to attenuate the emotions of the masses are very important issues. The will of individuals has a considerable impact on policy and is positively correlated to policy effectiveness to a certain extent [14]. Therefore, it is necessary for the government to understand people’s public opinion concerning the vaccine and to formulate appropriate policies to improve people’s willingness to vaccinate. Many scholars have already begun to study people’s attitudes towards vaccines [15].

Therefore, in the face of the present epidemic, work on improving the uptake of vaccinations is urgently needed. Finding ways to effectively promote the vaccinations is a serious problem, and emotion analysis plays a vital role in this field. Previous emotion analysis methods were mostly employed based on medium and long texts [16], and did not make full use of the features in the texts. Additionally, few research articles have applied natural language processing methods to epidemic analysis.

Therefore, we propose a fuzzy neural network classification approach based on a fuzzy vector that is suitable for analyzing short texts. The proposed model uses a fuzzy mathematics method to further mine the fuzzy features of the text and to improve text interpretability. In addition, a variety of data analysis methods can be used to explore and analyze the results and effectively consider people’s attitudes and views on vaccines.

Our research focuses on answering the following questions.

1. What is the attitude of Chinese citizens towards the COVID-19 vaccine?

2. What were the differences in the different months of the COVID-19 crisis in 2021?

3. What are China’s main concerns about the promotion of the COVID-19 vaccine?

China has recently started vaccination with booster shots. Understanding these factors influencing the success of these efforts is very important for epidemic prevention and control. This study will help government officials and decision-makers understand the problems that need to be solved before starting a mass vaccination process. The second section of this paper introduces the relevant work, the third section presents the methods employed in this research, the fourth section shows the changes in people’s feelings about the vaccine at different times and their reasons, and the fifth section summarizes the paper.

Related Work

In recent years, many advanced models have emerged in the field of emotion analysis.

Basiri et al. [17] proposed an attention-based bidirectional CNN-RNN deep model (ABCDM). By utilizing two independent bidirectional LSTM and GRU layers, [18] proposed a novel iterative network pruning with uncertainty regularization method for lifelong sentiment classification (IPRLS), which leverages the principles of network pruning and weight regularization. Li et al. [19] proposed a dual graph convolutional network (DualGCN) model that considers the complementarity of syntax structures and semantic correlations simultaneously.

Currently, social media plays a large role in people’s lives.1 reported that microblogs have 530 million monthly active Chinese users and 230 million daily active users and that microblog users tend to be younger than non-microblog users. Many NLP scholars research social media and have achieved concrete results. Dong et al. [20] combined BERT semantic extraction and CNN feature extraction to build an emotion classification model for commodity online comments. Neogi et al. [21] used the bag of words approach and TF-IDF to categorize and analyze the sentiments based on a collection of approximately 20,000 tweets about protest.

Social media provide data sources for NLP research. During the epidemic period, microblog traffic soared. As an important social tool, microblogs also provided much of the data worthy of our research during this special period. Many NLP scholars researched microblogs during the epidemic period. In addition, the progress of classifiers and optimization algorithms has laid a foundation for the

---

1 https://www.jiemian.com/article/6113013.html
application of neural networks for COVID-19 [22, 23]. Manguri et al. [24] found a neutral feeling of risk towards the coronavirus and reported that COVID-19 polarization was significantly higher than 50%; additionally, approximately 64% of individuals objected a vaccine.

Alzubi et al. [25] proposed COBERT, a retriever-reader dual algorithmic system for the complex searching of coronavirus-related literature. Naseem et al. [26] believed that negative opinions play an important role in regulating public sentiment and provides a large-scale Twitter sentiment dataset for researchers to study. Rahman and Islam [27] explored the performance of ensemble machine learning classifiers for sentiment analysis of COVID-19 tweets and found that stacking classifier (SC) showed the highest F1-score. Kaur et al. [28] proposed an algorithm referred to as a hybrid heterogeneous support vector machine (h-SVM) which was designed to assess the impact of Twitter on people’s mental health through emotion analysis. Basiri et al. [29] proposed a new method that integrates four deep learning and classical supervised machine learning models. The research shows that common emotion patterns could be observed in all countries during the epidemic.

Fuzzy mathematics has also been applied to epidemic analysis. Akram et al. [30] developed some general aggregation operators (AOs), which were used to accumulate Fermatean fuzzy data in a decision-making environment and applied coronavirus analysis. Garg et al. [31] combined FFS’s valuable properties with Yager operators. Six new operators are proposed and applied to select laboratories for COVID-19 testing to verify the usefulness of the operators. In response to the importance of operators, [32] proposed a multiattribute group decision-making (MAGDM) strategy and applied it to the operators’ selection of respirators for the COVID-19 pandemic. Jeracitano et al. [33] proposed a fuzzy logic–based deep learning (DL) approach to differentiate between CXR images of patients with COVID-19 pneumonia and interstitial pneumonias not related to COVID-19.

Other scholars have analyzed people’s views on vaccines. Ritona et al. [34] assessed the opinion of the Indonesian people through a social network analysis of the COVID-19 vaccine in January 2021, the results show that over 56% of people hold a negative attitude. Shim et al. [35] analyzed Korean tweets and found that when the epidemic was serious, there were more negative tweets about vaccines on Twitter. Praveen et al. [36] discovered that only 35% of Indians have positive viewpoints about the vaccine. Alamoodi et al. [37] confirmed that the anti-vaccination movement is gaining momentum and influencing more people through the internet and social media. [38] concluded through machine learning that most people are skeptical of the vaccine and do not know whether they should be vaccinated.

Although few studies have conducted emotional analysis of vaccines, most studies have suggested that people’s attitudes towards vaccines are not optimistic, which is consistent with the conclusion of this paper. Presently, many studies summarize people’s attitudes towards the epidemic in both the early and follow-up stages of the epidemic, but few studies have specifically explored people’s attitudes towards vaccines in China.

### Materials and Methods

#### Public Dataset and Constructed Vaccine Dataset

In this study, we utilize a total of five datasets: four public datasets, namely, NLPCC2013, NLPCC2014, the simplified Weibo dataset, and a comment dataset and a private vaccine dataset constructed by us. The four public datasets are used to verify the reliability of our proposed model. The corpora of the first three public datasets are obtained from microblogs, and all the public corpor are related to short text. Therefore, the corpora have a very high reference value for verifying our own dataset and predicting emotion. We will combine emotional levels and accuracy to establish a dataset to analyze public opinion and predict people’s emotions about the vaccine.

Through Python crawler technology, we crawled microblogs about vaccines from January 2021 to September 2021 on a microblog platform to build our vaccine dataset and we set the keyword to “vaccine.” In the process of data collection, we manually shielded the interference words that were related to vaccines but unrelated to the purpose of this study, such as “rabies,” “HPV,” and “paediatrics.” Table 1 summarizes the number, average length, and average number of microblog words per month.

| Time | Total Number | Average Length |
|------|--------------|----------------|
| Jan  | 3293         | 29.4           |
| Feb  | 3802         | 27             |
| Mar  | 3222         | 28.7           |
| Apr  | 3036         | 28.8           |
| May  | 4086         | 27.8           |
| Jun  | 3786         | 27.8           |
| Jul  | 3720         | 29.6           |
| Aug  | 3795         | 28.9           |
| Sep  | 3200         | 29.8           |
| Total| 31,335       | 28.6           |
| Average|           | 28.6           |

![Springer Logo]
NLPCC2013 and NLPCC2014 are sorted into a dataset named NLPCC. The division of the public dataset is shown in Table 2.

### Text Preprocessing

As most of the information released by people on social media consists of short texts and the format is not standardized and contains all kinds of invalid text and noise, which has a definite impact on our experimental results, preprocessing the text is a very important step. When calculating the emotion score with the emotion dictionary, a standardized text format should be used to obtain accurate scores. The specific steps are described here. An example of text preprocessing is shown in Table 3.

1. A microblog uses “#” to express a topic, but it has no meaning. The common link address and topic symbol “#” in the microblog were removed, such as “#universal vaccination” and “# new crown.”
2. The deletion of meaningless stop words such as “is,” “I,” and “this” do not impact the meaning of the original text.
3. Microblog texts with a length of less than 5 words, such as “user link” and “ah this,” which are too short and have no reference value, were removed.
4. Text with the character “@” was filtered. In most cases, the text with the character “@” is the content that needs to be forwarded or highlighted for other users to view [39]. For example, “@ want fried chicken,” “reply @ kanifei: #weibolove#//@kanifei: [rabbit]...,” and “@ lovia219a” were processed.
5. The text containing only English was filtered, as the experiment in this paper is based on a Chinese dataset and English text.

### Building an emotional dictionary

After text preprocessing, we score the text with an emotional dictionary. The emotional dictionary that we use integrates the emotional vocabulary ontology of Dalian University of Technology and the Bosonnlp emotional dictionary, with a total of 132,949 Chinese words. Each word has an emotional polarity score, and the emotion is distributed in the range of $[-1, 1]$. In the emotional dictionary, we denote that a score greater than 0 indicates a positive emotional tendency. The higher the score is, the stronger the tendency. A score less than 0 indicates a negative emotional tendency. The lower the score is, the stronger the tendency. The emotional dictionary is shown in Table 4.

Sometimes, when people express their emotions, they do not just use a single emotion-related word but use degree adverbs and negative words to modify their emotions. The result is referred to as fuzzy emotion in the literature [40]. For fuzzy emotions, our method judges whether there are negative words and degree adverbs between two emotion words. If there are degree adverbs, then the weight of degree adverbs is multiplied by the score of the emotional words. If there are negative words, the opposite score of the result is applied. Different from the previous methods of emotion analysis using an emotion dictionary, our integrated dictionary approach no longer only returns the emotion of the whole sentence but returns the positive emotion score and the negative emotion score in the form of a two-tuple.

The HowNet emotional dictionary is used for degree adverbs. There are 6 kinds of degree words, with intensities ranging from 0.9, 0.7, 0.5, and 0.3 to −0.5, thus including both positive and negative words. The list of degree adverbs is shown in Table 5.

There are often multiple emotional words in a sentence. When there is a degree adverb between two emotional words, we apply the weight of the degree adverb to the second emotional word, as shown in Eqs. 2–3. Similarly, if there is a negative word, we inversely score the second emotional word, such as in Eqs. 4–5.

$W$ represents the weight of degree adverbs, $sen_{word}[i]$ represents the emotional word score at the $i$ position in the

### Tables

| Table 2 | Public dataset partitioning |
|---------|-----------------------------|
|         | NLPCC | Comment | simplified_weibo |
| Train   | 18,001 | 30,000  | 200,000          |
| Dev     | 5000   | 10,000  | 80,000           |
| Test    | 5000   | 10,000  | 80,000           |

| Table 3 | Data cleaning example |
|---------|-----------------------|
|         | Before preprocess | After preprocess |
| Example 1 | @Jony帮我下周一打一针新冠疫苗啦！#新冠# | ['下周一', '一针', '新冠', '疫苗'] |
| Example 2 | 2.4一天过去了，第二针疫苗咋还这么疼 | ['2.4', '针', '疫苗', '疼'] |

| Table 4 | Emotional dictionary |
|---------|----------------------|
| Emotional word | Sentiment score |
| Free      | 0.5          |
| Delight   | 0.9          |
| Stable    | 0.1          |
| Revenge   | −1           |
| Crash     | −0.5         |
| Sorrowful | −0.7         |
sentence, and the getScore(s) function returns Senpos and Senneg.

\[
\text{Score}(\text{sen}) = \text{getScore}(\text{sen})
\]

(1)

\[
\text{Sen}_{\text{pos}} = \text{S}_{\text{pos}} + W \times \text{sen}_{\text{word}[i]}
\]

(2)

\[
\text{Sen}_{\text{neg}} = \text{S}_{\text{neg}} + W \times \text{sen}_{\text{word}[i]}
\]

(3)

\[
\text{Sen}_{\text{pos}} = \text{S}_{\text{pos}} + (-1) \times \text{sen}_{\text{word}[i]}
\]

(4)

\[
\text{Sen}_{\text{neg}} = \text{S}_{\text{neg}} + (-1) \times \text{sen}_{\text{word}[i]}
\]

(5)

### Fuzzy Vector

A fuzzy set is very suitable for addressing fuzziness in real life [41]. A fuzzy vector highly refines people’s expression of emotion on the basis of an emotion dictionary and includes the following five steps: fuzzification, application, implication, aggregation, and defuzzification. The fuzzy membership function is used to convert values into a fuzzy set, and then the fuzzy vector is calculated through fuzzy reasoning. Vashishtha et al. [42] proposed a method based on fuzzy rules for the emotional analysis of social media posts. We will improve the adaptability of this method to calculate fuzzy vectors and then add the obtained fuzzy vector and emotional dictionary vectors to the neural network for calculation.

We use the trapezoidal membership function to express the membership degree of neutral emotions. The S-type membership function and Z-type membership function are employed to represent high emotions degree and low emotions degree, respectively.

\[
\text{trapmf}(x, [a, b, c, d])
\]

(6)

\[
\text{MF}(x) = \begin{cases} 
\frac{x}{b-a} & a < x < b \\
\frac{1}{c-d} & b \leq x \leq c \\
\frac{x}{c-d} + \frac{d}{d-c} & c < x 
\end{cases}
\]

(7)

There are four important key points for the trapezoidal membership function. The numbers we set here are 0, 4, 6, and 10 (a, b, c, and d in Eq. 7). The reason why the trapezoidal membership function is selected is that when emotional polarity exists within this range, the emotional tendency is not obvious. For example, in the phrase “After the vaccine, I went to the market to buy a cabbage,” the positive score was 5.5. However, this sentence is a neutral emotion, so we use the trapezoidal fuzzy function to attain a membership neutrality of 1 for the sentences in this interval. For scores outside this interval, we calculate their membership neutrality according to the primary function. The trapezoidal membership function is expressed as shown in Fig. 1.

The S-shaped membership function Eq. (8) here represents the obvious degree of emotional membership. The formula of the S-shaped membership function is devised as follows Eq. (9). Through a derivation, we find that the closer to the center the function is, the greater the slope and the faster the rate of change. When we calculate the score through the emotional dictionary, we find that the closer to both ends the function is (that is, the higher the score and the lower the score). When the emotional polarity of a sentence is obvious, the score of the dictionary is more accurate, and when the emotion of a sentence is more ambiguous, the score is also fuzzier, which conforms to the distribution of the S-type membership function. The slope is gradual at both ends and steep in the middle. The S-shaped membership function is shown in Fig. 2.

\[
\text{smf}(x, [a, b])
\]

(8)

| Degree adverb | Weight |
|---------------|--------|
| One hundred percent | 0.9    |
| Strikingly    | 0.9    |
| Surprisingly  | 0.7    |
| Rather        | 0.7    |
| A little      | 0.3    |

Table 5  Degree adverb vocabulary

**Fig. 1** Trapezoidal membership function

**Fig. 2** S-shaped membership function
The Z-shaped membership function Eq. (10) here represents the membership degree of an emotion that is not obvious. When the emotion is 0, there is no emotion. There are two important key points for the Z-shaped membership function, which we set to 0 and 5. The reason for choosing the Z-shaped membership function is the same reason for choosing the S-shaped membership function, which conforms to our emotional distribution. The formula of the S-shaped membership function is presented as Eq. (11). The Z-shaped membership function will be expressed as shown in Fig. 3.

$$MF(x) = \begin{cases} 0 & x \leq a \\ 2 \times \left(\frac{x-a}{b-a}\right)^2 & a \leq x \leq \frac{a+b}{2} \\ 1 - 2 \times \left(\frac{x-b}{a-b}\right)^2 & \frac{a+b}{2} \leq x \leq b \end{cases}$$ (9)

$$MF(x) = \begin{cases} 1 - 2 \times \left(\frac{x-a}{b-a}\right)^2 & a \leq x \leq \frac{a+b}{2} \\ 2 \times \left(\frac{x-b}{a-b}\right)^2 & \frac{a+b}{2} \leq x \leq b \\ 1 & b \leq x \end{cases}$$ (11)

The Z-shaped membership function Eq. (10) here represents the membership degree of an emotion that is not obvious. When the emotion is 0, there is no emotion. There are two important key points for the Z-shaped membership function, which we set to 0 and 5. The reason for choosing the Z-shaped membership function is the same reason for choosing the S-shaped membership function, which conforms to our emotional distribution. The formula of the S-shaped membership function is presented as Eq. (11). The Z-shaped membership function will be expressed as shown in Fig. 3.

$$zmf(x, [a, b])$$ (10)

After aggregation, the fuzzy membership function will be as shown in Fig. 4.

When the $S_{pos}$ and $S_{neg}$ are calculated for a microblog, we can obtain six membership values through fuzzy calculation: $S_{pos\_high}$, $S_{pos\_mid}$, $S_{pos\_low}$, $S_{neg\_high}$, $S_{neg\_mid}$, and $S_{neg\_low}$.

Fuzzy rules Eqs. (12–20) are used for reasoning with these six fuzzy values.

$$R_1 = S_{pos\_high} \land S_{neg\_high}$$ (12)

$$R_2 = S_{pos\_high} \land S_{neg\_low}$$ (13)

$$R_3 = S_{pos\_high} \land S_{neg\_mid}$$ (14)
The intensity of the nine fuzzy rules ($R_1$-$R_9$) is obtained by calculating the intersection of the six process degrees of the two emotions. In operations, fuzzy logic is slightly different from deterministic logic. In fuzzy mathematics, we use $\wedge$ to represent the small operator and $\lor$ to represent the large operator.

When the subordinate degree of the sentence is higher for $POS$ than for $NEG$, we believe that the emotion of the sentence is positive, and vice versa; additionally, when the two are equal, the emotion is neutral. Therefore, we divide the above nine fuzzy rules into three parts for aggregation. Next, we obtain the intensity corresponding to the three emotions and apply it as our fuzzy vector $[s_{pos}, s_{mid}, s_{neg}]$.

\begin{align}
R_4 &= S_{pos\_mid} \land S_{neg\_high} \\
R_5 &= S_{pos\_mid} \land S_{neg\_mid} \\
R_6 &= S_{pos\_mid} \land S_{neg\_low} \\
R_7 &= S_{pos\_low} \land S_{neg\_high} \\
R_8 &= S_{pos\_low} \land S_{neg\_mid} \\
R_9 &= S_{pos\_low} \land S_{neg\_low}
\end{align}

The intensity of the nine fuzzy rules ($R_1$-$R_9$) is obtained by calculating the intersection of the six process degrees of the two emotions. In operations, fuzzy logic is slightly different from deterministic logic. In fuzzy mathematics, we use $\land$ to represent the small operator and $\lor$ to represent the large operator.

When the subordinate degree of the sentence is higher for $POS$ than for $NEG$, we believe that the emotion of the sentence is positive, and vice versa; additionally, when the two are equal, the emotion is neutral. Therefore, we divide the above nine fuzzy rules into three parts for aggregation. Next, we obtain the intensity corresponding to the three emotions and apply it as our fuzzy vector $[s_{pos}, s_{mid}, s_{neg}]$.

\begin{align}
s_{pos} &= R_2 \lor R_3 \lor R_6 \\
s_{mid} &= R_1 \lor R_5 \lor R_9 \\
s_{neg} &= R_4 \lor R_7 \lor R_8
\end{align}

**Fuzzy Neural Network Model**

The first step in modelling is word embedding. In this step, unstructured text is processed into structured data so that we can extract features from the text and complete the downstream NLP task. As the corpus that we built is very small, we use the Tencent_AI_lab’s pretraining model to construct the word vectors. These Chinese word vector data contain more than 8 million Chinese words, and each word corresponds to a 200-dimensional vector. Compared with existing Chinese word corpora, our corpus improves the quality and availability in terms of coverage, freshness, and accuracy. After the word-segmentation, each microblog is expressed as a [32,200] dimensional vector. We define the maximum length of a sentence as 32, intercept the redundant part, and fill in 0 for the lesser part. After conversion to a word vector, we add an expression fuzzy vector and semantic fuzzy vector to construct the word embedding, which is input into the model.

Since CNN was applied to emotion analysis and achieved great success [43], CNN models have been widely employed in NLP emotion classification tasks. This model has become an important in-depth learning model in the NLP field.

The most important feature of CNN is its convolution layer, which mainly constructs a feature map from a word embedding. First, each feature is locally perceived, and then the local information is comprehensively evaluated at a higher level to obtain the global information.

In the maximum pooling layer, pooling is also referred to as downsampling. The feature map obtained by convolution generally requires downsampling to make the features more obvious and to reduce the amount of data, as shown in Figs. 5 and 6.

$w$ represents the size of the input matrix, $k$ represents the size of the convolution kernel, $s$ represents the step size, and $p$ represents the number of layers of zero filling. After feature extraction in the convolution layer, we add the fuzzy vector and then enter the fully connected layer for further operations. The formula of the feature graph is given in Eq. (24).

\begin{align}
w' &= \frac{(w + 2p - k)}{s} + 1
\end{align}

The neural network outputs a vector with a dimension equal to the number of categories, where $z_i$ is the output value of the $i$th node and $c$ is the number of output nodes, that is, the number of classified categories. Through the soft-max function (25), the output values of multiple categories can be converted to probability distributions in the range of [0, 1].
Experimental Steps

To make our analysis more comprehensive, we use unsupervised machine learning methods, supervised deep learning methods, and our improved deep learning method to compare the different datasets. The machine learning methods are the support vector machine (SVM), k-nearest neighbor (KNN) algorithm, random forest (RF), and naive Bayes (NB) methods. The deep learning methods are a convolutional neural network (CNN) and a bidirectional long-term memory based on attention mechanism (Bi_lstm_ATT). The learning rate and number of epochs are set to \(10^{-3}\) and 10, respectively. Bidirectional encoder representation from transformers (BERT) is a pretrained language representation model launched by Google that has become a popular pretraining model in recent years. RoBERTa (a robustly optimized BER) is an improved version of Bert, in which the above two models are applied as the baseline model. The parameters of all the methods are shown in Table 6.

All the experiments were conducted using the following procedure:

1. Step 1: Crawl relevant text from microblogs.
2. Step 2: Perform text cleaning and preprocessing to reduce the noise in the text and to make features easier to extract.
3. Step 3: Calculate the emotion score and fuzzy vector.
4. Step 4: Conduct feature extraction through word embedding.
5. Step 5: Use the public dataset for model training.
6. Step 6: Input the text features into the pretrained model for prediction.

The flow chart of our model is shown in Fig. 7. The analysis of results in the next section indicates that the effect of the fused fuzzy vector input into our neural network for analysis is better than that achieved with a neural network alone.

![Model flow chart](image_url)

**Table 6** Classifier parameters

| Classifier | Parameters |
|------------|------------|
| SVC        | Kernel = linear, C = 1.0 |
| KNN        | n_neighbors = 5 |
| NB         | - |
| RF         | n_estimator = 100, max_depth = 300 |
| CNN        | filter_size=(2,3,4), num_filter=256 |
| Bi_lstm(att)| hidden_size=128, hidden_size2=64 |
| BERT       | batch size = 64 |
| RoBERTa    | batch size = 64 |
Experimental Results

In this section, we conduct a word-level analysis with the established vaccine dataset, explore it through word frequency and topic modelling, and analyze it at the macro-level.

Evaluating Indicator

The indicators used to evaluate the proposed method include accuracy, precision, recall, and F1. The values of precision, recall, and F1 are calculated using Eqs. (26)–(29). TP, FP, FN, and TN indicate that the positive sample we predict is positive, that the actual negative sample we predict is positive, that the positive sample we predict is actually negative, that the negative sample we predict is actually negative, and that the negative sample we predict is actually negative, respectively, as shown in Table 7. Train-dur refers to the training time in minutes.

\[
\text{Accuracy} = \frac{TP + TN}{TP + TP + FP + FN} \tag{26}
\]

\[
\text{Precision} = \frac{TP}{TF + FP} \tag{27}
\]

\[
\text{Recall} = \frac{TP}{TP + FN} \tag{28}
\]

\[
F_1 = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \tag{29}
\]

Model Results

NLPCC is a seven-category short-text emotion data set with distinct emotions. The accuracy, precision, recall, and F1 score for NLPCC dataset are shown in Table 8. The method in this paper yields the best result for this dataset, as the fuzzy vector abstracts the emotion scores for microblog short texts and integrates them into the convolution neural network in the form of fuzzy probability distributions, thus enhancing emotion analysis and improving the accuracy of emotion classification. We note that RoBERTa has the highest accuracy, which may be attributed to the longer training time and longer training sequence than those required by BERT and the ability to dynamically adjust the masking mechanism.

However, the performance of the machine learning models was generally not as good as that of the deep learning models, so we did not apply them in the vaccine emotion analysis.

As a deep learning method, the long-term and short-term memory network performs poorly because the microblog data are short texts, while the RNN has advantages in sequence recognition modelling due to its memory function; therefore, bidirectional long-term memory is more suitable for the analysis of long texts. For a task with a short sentence length, the CNN can summarize the overall structure of the sentence with its convolution function. However, when a sentence is long, the CNN can only process the information in the established window, and the information in adjacent windows can only be fused with the latter convolution layer. The CNN can grasp multiple different n-gram features of the text better than can the LSTM model, and for an n-gram feature, there are many different filters that can be used to extract useful information from different perspectives. The CNN has a good effect and generally avoids overfitting.

Although RoBERTa has the best effect on nlpcc dataset, the training time for the Bert and RoBERTa is very long, and is more than ten times that of the other methods, resulting in low efficiency. Therefore, we opt to use the FCNN as our model for epidemic public opinion prediction. Similarly, the method proposed in this paper obtains the best effect for the comment dataset and simplified_weibo dataset, as shown in Table 8.

Exploratory Analysis

As the experimental data that we crawled are not labelled, they are impossible to use for training, and deep learning models need a large number of datasets for feature learning and training. Chen et al. [44] points out that good results can be obtained by training with other similar corpora when the sample size is small. Therefore, we use the data provided by the public corpora NLPCC2013 and NLPCC2014 for training as these two standard datasets and the vaccine corpora that we have established involve short microblog texts with high accuracy and strong references and can support fine-scale emotion analysis. The analysis results are shown in Table 9.

We obtain the specific emotional distribution in each month, as shown in Fig. 8. In January, February, April, July, August, and September, disgust is the most prominent emotion. Like emotions are the most prominent emotions in March and May, and happy emotions are the most prominent emotions in June.

The accuracy obtained for the comment dataset is the highest. Through training based on the comment dataset, we obtain the proportions of positive and negative emotions...
from January to September. We discover that the proportion of people’s negative attitudes towards the vaccine is greater than that for positive attitudes in almost all periods and reaches a peak in April. Through LDA modelling and word clouds, we discover that “allergy” is a very frequent topic discussed by people at this time, which is likely an important reason for the negative orientation of public opinion, and that negative emotion reached an extreme value in July. We find that there are many words related to “outbreak” in that month. Notably, China did experience a brief outbreak of the epidemic in July, which may explain these negative emotions.

Based on statistics, the monthly vaccination situation is reported on the official website of the China Health

| Table 8 | The result of comment, NLPCC, and simplified_weibo dataset |
|---------|-----------------------------------------------------------|
| comment | KNN 0.65 0.67 0.65 0.65 0:32 | NB 0.9 0.9 0.9 0.9 0:57 |
|         | RFC 0.89 0.89 0.89 0.89 0:31 | SVM 0.9 0.9 0.9 0.9 1:24 |
|         | Bi_lstm(att) 0.91 0.91 0.91 0.91 1:58 | CNN 0.91 0.91 0.91 0.91 1:02 |
|         | FCNN 0.93 0.93 0.93 0.93 1:02 | RoBERTa 0.92 0.92 0.92 0.92 280:51 |
|         | BERT 0.7 0.7 0.7 0.7 280:16 | |
| NLPCC   | KNN 0.45 0.76 0.68 0.7 1:04 | NB 0.69 0.64 0.6 0.61 1:04 |
|         | RFC 0.73 0.75 0.69 0.71 1:00 | SVM 0.73 0.8 0.66 0.71 1:24 |
|         | Bi_lstm(att) 0.68 0.64 0.6 0.61 2:19 | CNN 0.73 0.73 0.67 0.69 1:04 |
|         | FCNN 0.74 0.76 0.68 0.7 1:04 | RoBERTa 0.76 0.76 0.76 0.75 47:32 |
|         | BERT 0.65 0.65 0.62 0.63 44:47 | |
| simplified_weibo | KNN 0.51 0.44 0.51 0.46 6:14 | NB 0.57 0.51 0.57 0.43 3:36 |
|         | RFC 0.57 0.51 0.57 0.52 3:21 | SVM 0.59 0.5 0.59 0.51 4:13 |
|         | Bi_lstm(att) 0.59 0.5 0.59 0.52 2:38 | CNN 0.59 0.52 0.59 0.52 2:06 |
|         | FCNN 0.59 0.52 0.59 0.51 2:11 | RoBERTa 0.55 0.55 0.55 0.55 169:08 |
|         | BERT 0.55 0.55 0.55 0.55 156:56 | |

| Table 9 | Sentiment distribution |
|---------|------------------------|
|         | Disgust | Happiness | Like | Surprise | Sadness | Anger | Fear |
| Jan    | 1060 | 786 | 812 | 111 | 315 | 178 | 32 |
| Feb    | 1311 | 636 | 1030 | 131 | 331 | 146 | 66 |
| Mar    | 708 | 661 | 818 | 110 | 456 | 143 | 49 |
| Apr    | 912 | 583 | 569 | 95 | 603 | 146 | 29 |
| May    | 643 | 696 | 967 | 190 | 672 | 215 | 53 |
| Jun    | 822 | 962 | 826 | 233 | 631 | 204 | 42 |
| Jul    | 1016 | 777 | 777 | 216 | 547 | 214 | 58 |
| Aug    | 987 | 885 | 790 | 137 | 538 | 232 | 49 |
| Sep    | 701 | 697 | 639 | 98 | 396 | 141 | 25 |
Commission\(^2\). Due to the lack of official data in January, February, and March, starting from April and comparing Fig. 9 with Fig. 10, we find that the vaccination increment in the months other than September is positively correlated with the proportion of positive emotions in that month. According to the official data, as of mid-September, the total number of vaccinations in China exceeded 1.1 billion, accounting for 78% of the total population in China. Then, the vaccination rate decreased significantly in September. At this time, the impact of emotion on vaccination became very small.

**Word Cloud Analysis**

When analyzing keywords, we use the method of creating a word cloud, which intuitively shows the topics discussed by people. Word clouds are also referred to as text clouds or tag clouds. In the generation of word clouds, we generally stipulate that the more times a specific word appears in a text data source, the more important the word is, and the larger the area it occupies in the word cloud.

Our dataset is divided according to time, so we can also observe the changes in people’s discussion topics in different time periods. For example, in January, people were very concerned about the “America” and “Virus” topics. In March, people were concerned about “pain,” “the first needle,” and “the second needle.” A very extensive discussion was carried out. People often discussed queuing in June. From the above rate chart, we determine that the vaccination rate reached its peak in June. According to people’s discussions in September, we discover that many people were vaccinated at this time, as reflected by the data. Note that when the word cloud is used, we can manually screen certain words strongly related to the COVID-19 vaccine, which better shows what people truly care about in the battle against the epidemic. For example, we screened the words “vaccine,” “COVID-19,” and “epidemic” in the word cloud. The word clouds at different times are shown in Fig. 12. A word cloud table can be obtained from the figure as shown in Table 10, which represents several topics that are highly discussed by people in different months.

**LDA Model**

LDA uses the word bag model to consider the semantics of a document and judge the relevance of the document\cite{45}, consequently, semantic mining and topic modelling can be performed. In LDA, each word in a document is obtained by selecting a topic with a certain probability or by selecting a word from the topic with a certain probability. If we want to generate a document, the probability of each word being in it is calculated based on Eq. (30), where \(p(\text{word/doc})\) represents the probability of a word appearing in the document, and \(p(\text{theme/doc})\) represents the probability of a topic appearing in the document. We set the number of topics to 6.

---

\(^2\) [http://www.nhc.gov.cn/](http://www.nhc.gov.cn/)
and use 7 words to describe the topic. The distribution of the first six topics in LDA is shown in Table 11. Notably, LDA helps us divide microblogs from January to September into 6 topics. Each topic is discussed around seven topics. LDA can be used to identify the topics discussed by the masses during the entire vaccination period at the macro level.

The analysis results reveal that topics such as the epidemic situation, time of vaccination, and feeling of vaccination are discussed.

\[
p(\text{word/doc}) = \sum_{\text{theme}} p(\text{word/doc}) \times p(\text{theme/doc})
\]

Discussion

In this section, we discuss the experimental results and collected data and answer the questions raised in the first section. In 2020, COVID-19 wreaked havoc on the world. People passed and discussed information through the internet. Thus, we judged and predicted people’s emotional trends by collecting text data. We crawled Chinese microblogs about the vaccine in the different months of 2021. We used the CNN model combined with fuzzy vectors to assess the emotional components of these statements.

For the first question, generally, there are more microblogs dominated by negative emotions than by positive emotions. According to Fig. 9, the negative attitude is higher than the positive attitude in almost every month, indicating that people’s overall attitude towards vaccines is pessimistic. The negative emotions peaked in April. According to the vaccination data, most people became vaccinated after April (Fig. 11). According to the word cloud chart (Fig. 12), most people expressed hope and had a positive expectation, but as the number of vaccinations increased, negative emotions gradually dominated.

For the second question, people had different emotional attitudes towards the vaccine in each month, overall, negative attitudes were more common higher than positive attitudes, and the emotional components significantly differed. First, the emotions fluctuated greatly, such as the most obvious emotion, disgust, which was identified in 35.91% in February but only 18.71% of texts after 3 months, with a fluctuation greater than 15%. Similar trends were observed for sadness and happiness. Second, two negative emotions, fear and anger, were rarely experienced, which may be due to the limited training data or a certain coincidence between these emotions and disgust and sadness. Thus, the classifier did not successfully recognize these emotions.

For the third question, we applied the Baidu Index, which is based on a big data statistical platform and can reflect the attention of internet users to a topic in different periods. We discovered that pessimism peaked in April. Combined with the word cloud and LDA model analysis, we believe that a
cause of negative emotions was allergies. In June, people’s negative emotions were still high, but the number of vaccinations peaked, and at the end of May and the beginning of June, the Baidu search index (Fig. 13) also peaked. People paid close attention to the vaccine and had been vaccinated, but negative emotion was still prevalent. As indicated by the Baidu Index on May 25, based on the highest number of searches, the most common search was “stop the first injection of new crown vaccine on June 10.” At that time, it was rumored that the distribution of the free Crown-19 vaccine would stop or that a charging system would be adopted. Words such as “queuing” and “pain” in the word cloud chart are also likely the sources of negative emotions.

The research results presented in this article apply not only to China but also to other countries. Moreover, the results can be applied by medical research institutions to improve vaccines or by government agencies to monitor public opinion, better promote vaccination, and help decision-makers establish better policies to stabilize public sentiment and maintain social stability.

**Conclusion**

The goal of this research was to explore how people’s views and emotional attitudes towards vaccination changed over time and whether the changes were accompanied by a certain rationale within the context of the new crown vaccine, as evidenced through the proposed fuzzy CNN model.

Our proposed model performed better than most of the baseline models. For the simplified Weibo and comment datasets, our proposed method achieved the highest accuracy, but it did not perform as well as RoBERTa for the NLPCC datasets. However, considering the time requirements, the FCNN is still the most suitable model for classifying short texts from microblogs.

In this study, we used the public datasets as the training set. Based on the emotion model, we combined word cloud and LDA methods so that we could macroscopically observe the reasons behind the changes in epidemic data; this approach could support the government in formulating more accurate publicity strategies and policies for epidemic prevention and control.

The limitations of this study are as follows: noise in building the dataset; ambiguity, which is minimally resolved; and confusion with the use of “vaccine” which may refer to the smallpox or HPV vaccine rather than the COVID-19 vaccine. Additionally, if run time is not considered, the accuracy obtained with the proposed model is not as good as that achieved with other models in all cases.

**Funding** This work was supported by the National Natural Science Foundations of China (Grant Nos. 12171065 and 11671001 and 61901074) and the Science and Technology Research Program of Chongqing Municipal Education Commission (Grant No. KJQN201900636).

**Data Availability** The data that support the findings of this study are available from the corresponding author upon reasonable request.
References

1. Ciotti M, Ciccozzi M, Terrinoni A, Jiang W, Wang C, Bernardini S. The COVID-19 pandemic. Crit Rev Clin Lab Sci. 2020;57(6):365–88.
2. Burki T. Coronavirus in China. Lancet Respir Med. 2020;8(3):238.
3. Yao W, Wang T, Jiang B, Gao F, Wang L, Zheng H, et al. Emergency tracheal intubation in 202 patients with COVID-19 in Wuhan, China: lessons learnt and international expert recommendations. Br J Anaesth. 2020;125(1):e28–37.
4. Baloch S, Baloch M, Zheng T, Pei X. The coronavirus disease 2019 (COVID-19) pandemic. Tohoku J Exp Med. 2020;250(4):271–8.
5. El Zowalaty M, Järhult J. From SARS to COVID-19: a previously unknown SARS-related coronavirus (SARS-CoV-2) of pandemic potential infecting humans—call for a one health approach. One Health. 2020;9.
6. Marcce R, Majta M, Likic R. Will vaccination refusal prolong the war on SARS–CoV-2? Postgrad Med J. 2021;97(1145):143–9.
7. Chou W, Budenz A. Considering emotion in COVID-19 vaccine communication: addressing vaccine hesitancy and fostering vaccine confidence. Health Commun. 2020;35(14):1718–22.
8. Omer S, Salmon D, Orenstein W, Doherty M, Halsey N. Vaccine refusal, mandatory immunization, and the risks of vaccine-preventable diseases. N Engl J Med. 2009;360(19):1981–8.
9. MacDonald N. Vaccine hesitancy: definition, scope and determinants. Vaccine. 2015;33(34):4161–4.
10. Dube E, Laberge C, Guay M, Bramadat P, Roy R, Bettinger J. Vaccine hesitancy: an overview. Hum Vacc Immunother. 2013(8):1763–73.
11. Troiano G, Nardi A. Vaccine hesitancy in the era of COVID-19. Public Health. 2021;194:245–51.
12. Dror A, Eisenbach N, Taiber S, Morozov N, Mizrachi M, Zigron A, et al. Vaccine hesitancy: the next challenge in the fight against COVID-19. Eur J Epidemiol. 2020;35(8):775–9.
13. Giardina J, Bilinski A, Fitzpatrick M, Kendall E, Linas B, Salomon J, et al. Model-estimated association between simulated us elementary school-related SARS-CoV-2 transmission, mitigation interventions, and vaccine coverage across local incidence levels. JAMA Netw Open. 2022;5(2):e2147827–e2147827.
14. Page B, Shapiro R. Effects of public opinion on policy. Am Polit Sci Rev. 1983;77(1):175–90.
15. Agustiningsih K, Utami E, Al Fatta H. Sentiment analysis of COVID-19 vaccine on Twitter social media: systematic literature review. In 2021 IEEE 5th International Conference on Information Technology, Information Systems and Electrical Engineering (ICITISSE), pages 121–126. IEEE, 2021.
16. Fei H, Ren Y, Wu S, Li B, Ji D. Latent target-opinion as prior for aspect-based sentiment analysis: a variational approach from fine-grained perspective. In Proceedings of the Web Conference. 2021;2021:553–64.
17. Basiri M, Nemati S, Abdar Mand, Cambria E, Acharya U. ABCDM: an attention-based bidirectional CNN-RNN deep model for sentiment analysis. Futur Gener Comput Syst. 2021;115:279–294.
18. Geng B, Yang M, Yuan F, Wang S, Ao X, Xu R. Iterative network pruning with uncertainty regularization for lifelong sentiment classification. In Proceedings of the 44th International ACM SIGIR conference on Research and Development in Information Retrieval, pages 1229–1238, 2021.
19. Li R, Chen H, Feng F, Ma Z, Wang X, Hovy E. Dual graph convolutional networks for aspect-based sentiment analysis. In Proceedings of the 59th Annual Meeting of the Association for Computational Linguistics and the 11th International Joint Conference on Natural Language Processing (Volume 1: Long Papers), pages 6319–6329, 2021.
20. Dong J, He F, Guo Y, Zhang H. A commodity review sentiment analysis based on BERT-CNN model. In 2020 5th International Conference on Computer and Communication Systems (ICCCS), pages 143–147. IEEE, 2020.
21. Neogi A, Garg K, Mishra R, Dwivedi Y. Sentiment analysis and classification of Indian farmers’ protest using twitter data. International Journal of Information Management Data Insights. 2021;1(2):100019.
22. Alzubi O, Alzubi J, Alweshah M, Qiiqueh I. An optimal pruning algorithm of classifier ensembles: dynamic programming approach. Neural Comput Appl. 2020;32(20):16091–107.
23. Movassagh A, Alzubi J, Ghieisari M, Rahimi M, Mohan S, Abbasi A, et al. Artificial neural networks training algorithm integrating invasive weed optimization with differential evolutionary model. J Ambient Intell Humaniz Comput. 2021;1–9.
24. Manguri K, Ramadhan R, Amin P. Twitter sentiment analysis on worldwide COVID-19 outbreaks. Kurdistan Journal of Applied Research. 2020;54–65.
25. Alzubi J, Jain R, Singh A, Parwekar P, Gupta M, Cobert: COVID-19 question answering system using BERT. Arab J Sci Eng. 2021;1–11.
26. Naseem U, Razzak I, Khushi M, Eklund P, Kim J. COVIDSenti: a large-scale benchmark Twitter data set for COVID-19 sentiment analysis. IEEE Transactions on Computational Social Systems. 2021;8(4):1003–15.
27. Rahman M and Islam M. Exploring the performance of ensemble machine learning classifiers for sentiment analysis of COVID-19 tweets. In Sentimental Analysis and Deep Learning, pages 383–396. Springer, 2022.
28. Kaur H, Ahsaan S, Alankar B, Chang V. A proposed sentiment analysis deep learning algorithm for analyzing COVID-19 tweets. Inf Syst Front. 2021;23(6):1417–29.
29. Basiri M, Nemati S, Abdar M, Asadi S, Acharayya U. A novel fusion-based deep learning model for sentiment analysis of COVID-19 tweets. Knowl-Based Syst. 2021;228.
30. Akram M, Shahzadi G, Ahmadini A. Decision-making framework for an effective sanitizer to reduce COVID-19 under Fermatean fuzzy environment. Journal of Mathematics. 2020.
31. Garg H, Shahzadi G, Akram M. Decision-making analysis based on Fermatean fuzzy Yager aggregation operators with application in COVID-19 testing facility. Math Probl Eng. 2020.
32. Shahzadi G, Akram M. Group decision-making for the selection of an antivirus mask under Fermatean fuzzy soft information. J Intell Fuzzy Syst. 2021;40(1):1401–16.
33. Ieracitano C, Mamnone N, Versaci M, Varone G, Ali A, Armentano A, et al. A fuzzy-enhanced deep learning approach for early detection of COVID-19 pneumonia from portable chest x-ray images. Neurocomputing. 2022;481:202–15.
34. Vitonja M, Al Ihsan M, Anjar A, Rambe F. Sentiment analysis of COVID-19 vaccine in Indonesia using Naive Bayes Algorithm. In IOP Conference Series: Materials Science and Engineering, volume 1088, page 012045. IOP Publishing, 2021.
35. Shim J, Ryu K, Lee S, Cho E, Lee Y, Ahn J. Text mining approaches to analyze public sentiment changes regarding COVID-19 vaccines on social media in Korea. Int J Environ Res Public Health. 2021;18(12):6549.
36. Praveen S, Ittamalla R, Deepak G. Analyzing the attitude of Indian citizens towards COVID-19 vaccine—a text analytics study. Diabetes Metab Syndr Clin Res Dev. 2021;15(2):595–9.
37. Alamoodi A, Zaidan B, Al-Masawa M, Taresh S, Noman S, Ahmaro I, et al. Multi-perspectives systematic review on the applications of sentiment analysis for vaccine hesitancy. Comput Biol Med. 2021;139: 104957.
38. Amjad A, Qaiser S, Anwar A, Ali R. Analysing public sentiments regarding COVID-19 vaccines: a sentiment analysis approach. In 2021 IEEE International Smart Cities Conference (ISC2), pages 1–7. IEEE, 2021.
39. Yu Y, Qiu D, Yan R. A multi-modal and multi-scale emotion-enhanced inference model based on fuzzy recognition. Complex & Intelligent Systems. 2022;8(2):1071–84.
40. Phan H, Tran V, Nguyen N, Hwang D. Improving the performance of sentiment analysis of tweets containing fuzzy sentiment using the feature ensemble model. IEEE Access. 2020;8:14630–41.
41. Chakraborty K, Bhatia S, Bhattacharyya S, Platos J, Bag R, Hassanien A. Sentiment analysis of COVID-19 tweets by deep learning classifiers-a study to show how popularity is affecting accuracy in social media. Appl Soft Comput. 2020;97.
42. Vashishtha S, Susan S. Fuzzy rule based unsupervised sentiment analysis from social media posts. Expert Syst Appl. 2019;138: 112834.
43. Liao S, Wang J, Yu R, Sato K, Cheng Z. CNN for situations understanding based on sentiment analysis of twitter data. Procedia Computer Science. 2017;111:376–81.
44. Chen Z, Cao Y, Lu X, Mei Q, Liu X. Sentimoji: an emoji-powered learning approach for sentiment analysis in software engineering. In Proceedings of the 2019 27th ACM Joint Meeting on European Software Engineering Conference and Symposium on the Foundations of Software Engineering, pages 841–852, 2019.
45. Blei D, Ng A, Jordan M. Latent dirichlet allocation. J Mach Learn Res 2003;3(Jan):993–1022.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.