Research on Fine-Grained Classification of Rumors in Public Crisis
——Take the COVID-19 incident as an example

Shuaipu Chen

Huazhong Agricultural University Hubei, China

Abstract—[Purpose / Meaning] Rumors are frequent in the COVID-19 epidemic crisis. In order to unite the power of dispelling rumors of various media platforms to help to break the rumors in a timely and professional manner, this article has designed a new fine-grained classification of rumors about COVID-19 based on the BERT model. [Method / Process] Based on the rumor data of several mainstream rumor refuting platforms, the pre-training model of BERT was used to fine-tuning in the context of COVID-19 events to obtain the feature vector representation of the rumor sentence level to achieve fine-grained classification, and a comparative experiment was conducted with the TextCNN and TextRNN models. [Result / Conclusion] The results show that the classification F_1 value of the model designed in this paper reaches 98.34%, which is higher than the TextCNN and TextRNN models by 2%, indicating that the model in this paper has a good classification judgment ability for COVID-19 rumors, and provides certain reference value for promoting the coordinated refuting of rumors during the public crisis.

1 INTRODUCTION

In the era of the Internet, massive amounts of information are spreading at a high speed in human society. At the same time, false rumors can be widely spread through various network channels. Especially on social media platforms, most users have difficulty in discerning the authenticity of information, failing to receive the information of refuting rumors in a timely manner, and directly forwarding the spread of false information, which makes it more difficult to refute rumors online. In the past few years, the widespread blood poisoning incidents of AIDS patients and the domestic salt rush caused by the Japanese nuclear accident have caused certain adverse effects on the normal social order. After the outbreak of the public crisis, all kinds of information and sediments appear, and the network rumors mixed in it even act as a crisis catalyst. If the rumors cannot be effectively dealt with effectively in a timely manner, it will hinder the control of public crisis events and even have an inestimable impact on public sentiment, government image, and social solidarity.

The new Corona Virus Disease 2019 (COVID-19, hereinafter referred to as "COVID-19") outbreak at the end of 2019 is a major public health crisis which is the fastest spread, the most widespread infected, and the most difficult to control in China since the founding of New China. It has had a major impact on the society, economy, politics, and life of our country even all over the world, and the endless network rumors has caused public panic. In this case, the prevention 、 control and governance of rumors is also an important link to test the victory of China's anti-epidemic struggle. According to the "Special Report on" Mobile Epidemic "in the China Mobile Live Broadcasting Industry", during the COVID-19, the amount of time the public spends on the mobile Internet per day has increased by 21.5% compared with the beginning of the year, and people will spend more time on different social platforms to get all kinds of information. Taking Sina Weibo as an example, about 25 million posts are posted every hour. Faced with the impact of massive information, content reviewers will be limited in the timeliness of identifying false information and are prone to misjudgment. At present, several media platforms such as Weibo, Tencent News, Xinhua News Agency and People's Daily have launched rumors prevention and control mechanisms, opened rumors prevention platforms, and tried their best to crack down on rumors. Users can cross-verify the published information of multiple media platforms, but the increase of cross-platform search time cost will still limit the timeliness of rumor refutation information received by users. Whether the fine-grained classification of rumors can be realized by artificial intelligence algorithms is still the key issue to be solved urgently by the rumors platform. When the machine recognizes the rumors and sorts them according to the characteristics of the rumors text, people can quickly find out the official institutions and authoritative experts capable of responding to the rumors, publish and spread the rumors information as soon as possible, and cooperate with the media platforms to clarify the rumors to purify the network space and maintain society and harmony.

774222057@qq.com

In view of the above problems, this paper proposes a new fine-grained classification algorithm based on the BERT pre-training model for rumors of COVID-19, and compares it with the TextCNN and TextRNN models. The experimental results show the effectiveness of the algorithm in this paper in order to provide some technical support for the coordination of dispelling rumors during the public crisis .

2 LITERATURE REVIEW

2.1 Research on the classification of rumors

Rumor topic classification research is a sub-task of Natural Language Processing (NLP) text classification. It is essentially a multi-classification problem of text, that is, to automatically classify rumor topics according to a certain classification system or rule. At present, text classification tasks are widely used in the fields of information indexing, digital book management, and intelligence filtering [1].

The development of text classification has roughly gone through four stages. The first stage was in the 1970s. Salton et al.[2]proposed the vector space model VSM (Vector Space Model) for text representation. The second stage is after the 1980s, Knowledge Engineering (KE) has become the main technology of text classification. It forms logical rules based on the knowledge provided by experts in a field, and manually establishes a classifier, which is used as a computer text classification. The third stage is the late 1990s and early 21st century, and text classification methods based on statistics and machine learning are gradually emerging. The earliest application of machine learning methods is Naive Bayes [3]. After decades of development, almost all important machine learning algorithms have been applied in the field of text classification, such as support vector machines (SVM) [4], neural networks[5]and decision tree [6]etc. Currently in the fourth stage, deep learning technology has also been widely used in text classification tasks.

Rumor topic classification research has gradually emerged in the third stage of the development of text classification. Early research used machine learning algorithms to process various network information topic classification tasks. Ding Shengchun et al.[7] used ontology to integrate domain knowledge and domain text features into the classification process, and used a weighted naive Bayes model to classify network information topics. For the topic classification task on Sina Weibo, Jia Longjia et al.[8]proposed to convert features from word-based representations to category-based representations, and then use the support vector machine method. With the appearance of online news and online socialization, the identification and classification of false rumors has aroused widespread concern in academia. Zubiaga et al[9]used the decision tree model to establish a rumor classification model consisting of four parts: rumor detection, rumor tracking, rumor stance classification, and rumor accuracy classification. Zong Qianjin et al.[10]classified the topics of rumors generated and reported through the implicit Dirichlet distribution. Jiang Ying et al.[11]used text

sentence feature analysis method to identify rumors. Mikolov et al. proposed the Word2ve model based on the calculation of word vectors in deep learning in 2013[12], which triggered the upsurge of applying deep learning technology to the word embedding model (Word Embedding). After Yoon Kim[13]proposed to apply Convolutional Neural Network (CNN) to text classification tasks in 2014 and achieved good results, the academic community tried to use deep learning techniques to deal with rumors-related tasks. Scholars such as Liao Xiangwen[14]proposed a new hierarchical neural network (HSA-BLSTM) that combines social information to detect rumors and integrate the attention mechanism and social context into the network. Scholars such as Jin[15]proposed to use Attention-Recurrent Neural Network (Att-RNN) to fuse multi-modal features to achieve effective rumor detection. At the same time, many scholars have put Z.Yang[16]and others proposed to integrate the layered attention mechanism (Attention) into their own rumor detection.

The BERT model based on the layered attention mechanism[17]has achieved better results in processing text classification tasks, but there are still few studies on the application of the BERT model to the fine-grained classification of COVID-19 incidents. This article will use the BERT pre-training model to fine-tune the research task to capture the text features of the COVID-19-related rumors to achieve fine-grained classification, and integrate it into our country 's rumors system for handling public crisis rumors, reflecting the practical significance of its application to the refutation of public crisis rumors.

2.2 earch on network refuting rumors

1) Definition and spread of rumors

There are various ways of defining rumors in the academic world. For example, in 2007, DiFonzo[18] defined rumors as "unverified and relevant instrumental information statements in circulation", and in 2014 Cai[19]defined rumors as "some False information ". With reference to scholars such as Ding Haiyan[20]and Miao Junfu[21], rumors are defined as "unconfirmed information", and authoritative dictionaries such as Oxford define rumors as "currently circulating stories about uncertain or suspicious facts or Report".This article will continue the above definition, and in the context of the public crisis, the public crisis rumors are defined as some of the truth at the time of publication that is still to be verified during the public crisis(this article refers to the COVID-19 incident.

The public's stimulation by public crisis events will produce an adaptive response of psychological feelings and emotions, which breeds public crisis rumors[22]. Affected by emergencies, public opinion and the spread of rumors often show a short-term exponential growth. The rumor propagation model generally accepted by the academic community is the model proposed by Allport and Postman[23]——rumor = importance of event × ambiguity of event. In recent years, based on this model, scholars have continuously tried to add new elements through research. On the basis of this formula, foreign scholars like Prasad et

al.[24-26] and Wu Jian and other domestic scholars[27-31]incorporated personal anxiety and other emotions, critical consciousness, etc. into the model to better predict the speed of rumors. From the perspective of the measurement of these rumors spreading models, rumors are more tenacious under the new spreading pattern and spreading environment.The refutation of public crisis events is easy to fall into the unfavorable situation of "spreading rumors moves the mouth, refuting rumors breaks the leg".

2) Research status of Internet refuting rumors

Refuting rumors of Internet rumors refers to the release of authoritative and scientific information, refuting the wrong rumors, and hedging the rumors spread in the cyberspace, and setting the order anyway[32]. Some constructive theories and methods have been developed in China for the research of refuting Internet rumors. At present, the academic circle focuses on two aspects of refuting Internet rumors: the first aspect is to carry out simulation experiments on the rumor propagation model with the rumor refutation mechanism, so as to discover the key node of the rumor refutation mechanism. Wang xiaoli[33]et al. concluded that refuting rumors against the ignorant can effectively reduce the maximum influence of rumor propagation by establishing a rumor propagation model that takes into account the refuting mechanism of rumors in heterogeneous networks. The second aspect is to explore the mechanism of refuting rumors through the construction of relevant theoretical framework, so as to realize the network refuting rumors. The two types of online rumor refuting methods have a common feature, that is, they refute rumors by seeking the influence of the subject, channel and audience in rumor spreading. At present, the academic circle is more widely recognized by the establishment of theoretical framework to realize the research of rumor refutation. With the rapid change of the communication form and communication environment, these theoretical studies analyze its internal mechanism and continuously innovate its realization form.

The first mode of applying the theoretical framework to refute rumors is self-cleaning. Self-cleaning refuting rumors means that users use the online platform to reprint and cite spontaneously to identify the authenticity of the network information, and widely spread the verification information to refute the rumors, and realize the self-purification of online rumors through the exchange of information between users rather than official media. Academia's attention to self-cleaning rumors stems from the rise of social platforms represented by Weibo. There are two representative views of "shadowless lamp" and "failure puzzle" in its understanding of its utility. Scholar Yu Guoming[34]called Weibo rumors as a false proposition. In his view, as long as the mechanism for publishing information on Weibo is sufficiently open, Weibo will supplement, correct, and verify the information, and restore the truth to the "shadowless lamp effect." , wich also achieved online refuting rumors. Scholars such as Wang Guohua[35]believe that the self-purification effect of Weibo is often limited. Fragmented information is difficult to be effectively spliced under the conditions of poor flow, insufficient discussion, and unequal public voice,which is a "failure puzzle" effect . Therefore, this kind of rumor-dispelling mode is often regarded as the result of the instrumental role of the social platform, but its actual effect is limited by the actual communication environment and individual psychological factors[36], that is, the in-depth rumor-dispelling function, media literacy education function and rumor self-purification function of the platform need to be strengthened.

The second mode of applying the theoretical framework to refute rumors is organized refuting rumors. Self-cleaning refuting rumors are highly susceptible to personal influence, which provides reasonableness for the intervention of official subjects such as governments, media, and platform companies. Organized refuting rumors refers to the behavior of the official subject to refute rumors by virtue of its authority, through its rigorous and scientific investigation. The main body of organized refutting rumors are government, media and well-recognized authoritative organizations represented by the World Health Organization. Compared with traditional rumor-refuting, the advantage of organized refuting rumors is that it is based on credible information that has good timeliness, fast speed, rapid interaction, and attribute dissemination structure. The official subject can play the role of rumor manager in emergencies, but there are also difficulties in organized refuting rumors. Taking the level of government rumours removal as an example, Tang Xuemei's scholars[37]analyzed the government's rumor-breaking behaviors of Taifu Middle School from the aspects of time, subject and content. After the rumors spread quickly, the government's repelling effect would be counterproductive, causing a strong sense of distrust in the public. Regarding media refuting rumors, the research focused on tracking hot events. Scholars such as Yao Qi[38]used the questionnaire method to analyze timely and effective official reports, which not only increased public confidence in the victory of the epidemic, but also promoted the autonomy of public online rumors, thereby achieving the diversification of the main body of online rumors. However, if some self-predicted self-media carry out irresponsible investigation under the banner of official media, it will cause the public to lose media trust. In practice, the government, the media and other official subjects have always been the backbone of the army of rumors. The result of organized refuting rumors not only relates to the overall efficiency of online rumors, but also involves the construction of organizational image and social credibility. It can be seen that the administrative logic of the government and the professional qualities of the media are the starting points for analyzing organized refuting rumors. Only scientific and reasonable investigations can gain the trust of the public. With the transformation and development of the new government media and the traditional mainstream media, the government media should focus more on the social platform to refuting rumors.

The third mode of applying the theoretical framework to refute rumors is social repulsion of rumors. Some rumors are not only accompanied by appealing words , but also packed with eye-catching pictures. There may be interest groups behind this fine rumors. For such rumors, it is necessary to integrate the power of society, so that individuals, media, governments and others can use different forms of division of labor to coordinate rumors,

which is also the main form of social repulsion of rumors. Jin Jianbin[39]regards "socialization" and "synergy" as the key to this coping mode, in which "socialization" emphasizes the mobilization and integration of various rumor-dispersing forces scattered in complex social systems, while in essence "synergy" is understood as formal and informal effective collaboration and information sharing among various forces. social repulsion of rumors can essentially be understood as a collective action. This kind of collective action is also the construction of a coordination mechanism. Most scholars regard the information exchange mechanism as the basis for the linkage and coordination of various subjects. For example, Li Biao[40] emphasized the data sharing in the collaborative process, and believed that a national-level network rumor data platform should be built, and on this basis, a mature rumor trigger mechanism should be formed through the construction of media integration. In recent years, a number of rumor-disruption platforms have been formed in China: one is a special rumour-removal platform launched on behalf of platforms such as Sina, Netease and Tencent; The second is the guoku network rumor mill, guoku network gossip encyclopedia and science squirrels will be represented by the professional folk refuting website; The third is based on the rumor platform established by the State Central Cyberspace Administration. However, as far as the current situation is concerned, many rumors-dispelling platforms and organizations have not effectively classified all kinds of rumors, and are still in their respective "divided operations". There is no effective coordination yet, which limits the development of social collaborative defamation.

From the above research, it can be found that a large number of machine learning algorithms have been applied to the rumors topic classification task, and the emerging deep learning technology has also achieved good results in this task. Self-cleaning rumors are based on individuals, and organized refuting rumors are based on official subjects. Both need to work together and strengthen internal ties to better address the negative impact of current rumors. From this point of view, social repulsion of rumors is the main battlefield for effectively refuting rumors. A large amount of valuable rumor-dispelling information is spread on many different media platforms. Some rumors and rumor-dispelling information have not been sorted in time. The time cost for the public to obtain real information is relatively high, which hinder the effect of social repulsion of rumors. Domestic scholars have used machine learning algorithms to automatically classify rumors. During the COVID-19 incident, false information on social platforms has increased rapidly. It is urgent to use deep learning techniques to quickly mine massive data features to achieve fine-grained rumors classification. Therefore, this article uses the BERT model to fine-tune the fine-grained classification task of the COVID-19 rumors, so that the model learns the characteristics of different types of rumors, and classifies the rumors and rumors information on multiple platforms. The fast and accurate rumors classification algorithm will provide certain technical support for combining the rumors-repellent platforms with the power to refute rumors. This article draws on the framework of social repulsion of rumors, integrates rumor-

refutingl information into the three-stage: intelligence flow of pre-warning, decision-making in-process, and post-mortem evaluation, and uses the fine-grained rumor classification algorithm in this paper to gather information from multiple rumor-removal platforms to build a commonly recognized rumor classification guidelines, so as to achieve fast and accurate matching of the rumor categories and the main body of refuting rumor, thereby achieving the positive feedback effect of the rumor intelligence system.

3 MODEL INTRODUCTION

This paper uses the BERT model to classify fine-grained texts of public crisis rumors. The input of the traditional pre-trained model is to input the text sequence from left to right or from right to left. The BERT model uses the two-way text representation function of the Transformer model. Studies by Devin and other scholars[17]also show that Transformer-based BERT has strong feature extraction capabilities, which makes the language model of two-way training a deeper understanding of context than that of one-way language model, and when applied to text classification tasks, it can achieve better results than previous deep learning models.

3.1 Transformer model

As an important part of the BERT model, the Transformer model solves the problem that the traditional Word2vec word vector cannot express different word meanings in different contexts, It is different from the sequential execution structure of RNN, and has better parallel nature.

Transformer uses the structure of Encoder-Decoder, as shown in Figure 1. The Encoder side and Decoder side are based on the self-attention mechanism for feature selection, so the working principle of the two is basically the same. But the difference is that the Decoder side uses the Multi-Headed Attention mechanism with a "Masked" mechanism. Because the task of the Decoder side is to predict data, it adds this mechanism to prevent the premature disclosure of predicted data. In addition, it also adds the Encoder-Decoder-attention mechanism, which takes into account the information obtained on the Encoder side. After fully extracting the features, it will add a fully connected layer (Linear Layer) ,and then carry out the softmax function to find the probability, so as to realize the data prediction.
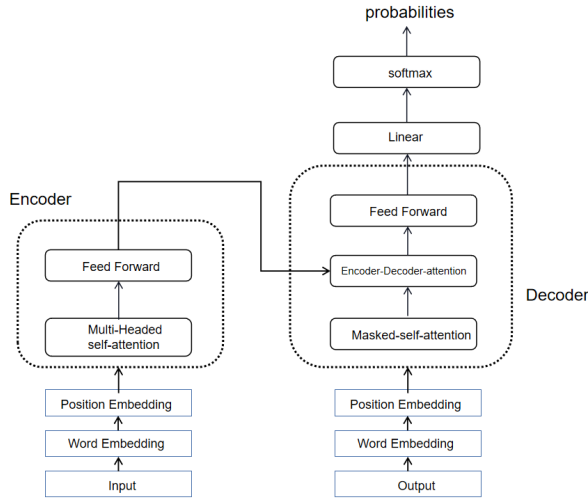
Fig. 1.   Overall architecture diagram

1)Self-Attention mechanism

In order to pay attention to the valuable information in the text for the computer, the Transformer model introduces a self-attention mechanism. This mechanism calculates the degree to which other related words are related to the word being processed by capturing the syntactic and semantic characteristics between words in the same text.

Taking the "it" in "The family didn't go to the hotel because it was bankrupt" as an example, the results are shown in Figure 2. The darker the text and lines, the greater the degree of correlation. Later, when feature extraction is performed on "it", more attention will be paid to these highly related words.
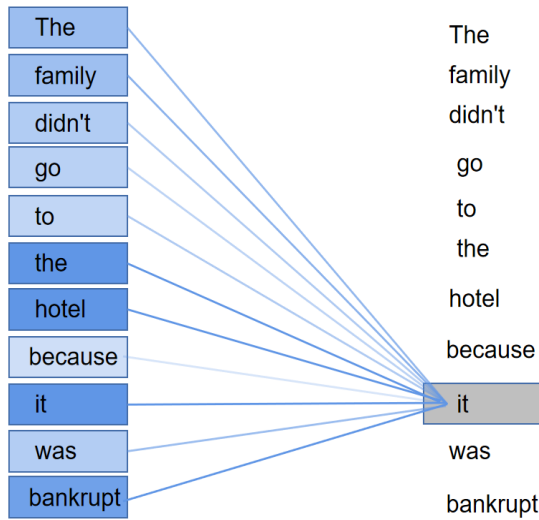

Fig. 2.   The degree of "it" attention

Next, the calculation process of the self-attention mechanism is introduced. First, each word in the text is Embedding, and then the relationship between the degree of attention of each word and other words is established, which is a weighted operation. This step is achieved by constructing Query, Key, and Value matrices. Finally, the corresponding results are obtained through the softmax function, and features are also allocated based on the result scores. In order to explain the calculation process in detail, this article introduces related formulas and legends for explanation.

According to the input text, the initial Emdedding of the input words is shown in Equation 1:

$$X = (x0, x1, x2, x3, \ldots, xn) \qquad （1）$$

In order to construct three matrixes to be trained for the input Embedding, it requires linear mapping for each word vector x and multiplying by the W^Q weight matrix to obtain the corresponding Query vector. The meaning of this vector is to Query the relationship between this word and other words. When all the word vectors complete this operation in parallel, the Query matrix is obtained. Similarly, the Key matrix and Value matrix can be obtained by multiplying by W^K and W^V respectively. The relevant calculation formulas are shown in Equations 2, 3, and 4:

$$key = linear\_k(x) \qquad （2）$$

$$query = linear\_q(x) \qquad （3）$$

$$value = linear\_v(x) \qquad （4）$$

After obtaining the matrix, the q vector of each word and each k vector in the text are calculated for the inner product, that is QK^T, we use the softmax function to process the score obtained by the inner product and the V vector of the word. And we use pair $\sqrt{d_k}$ to scale QK^T, $\sqrt{d_k}$ is actually the K vector dimension of the square root, so as to exclude the influence of vector dimension. The formula is shown in Equation 5:

$$Attention(Q, K, V) = Soft\max(\frac{QK^T}{\sqrt{d_k}})V \qquad （5）$$

The "The" and "family" examples above show the overall calculation process of the self-attention mechanism. The calculation results of "disease" are shown in equation 6:

$$z1 = soft\max(\frac{q1 \times k1}{\sqrt{d_k}})v1 + soft\max(\frac{q1 \times k2}{\sqrt{d_k}})v2 \qquad （6）$$

Where z1 represents "The", and q1, k1, k2, v1, and v2 represent q, k, and v vectors corresponding to "The" and "family" respectively.

2) Multi-Headed-Attention mechanism

A set of Query, Key, and Value matrices can get the characteristic expression of a set of current words, and the Transformer model introduces a multi-headed-attention mechanism (Multi-Headed-Attention) mechanism. Multiple sets of Query, Key, and Value matrices are used to obtain multiple feature expressions (Transformer models generally have 8 attention heads), which is somewhat similar to the convolution kernel in convolutional neural networks. After all the features are spliced together, dimension reduction can be carried out through a full connection layer, which is conducive to more accurate feature expression. The calculation formula of the multi-headed mechanism is shown in equation 7, and the process is shown in figure 3:

$$Z = Concat(head_0, head_1, \ldots\ldots, head_7) \bullet Connected$$

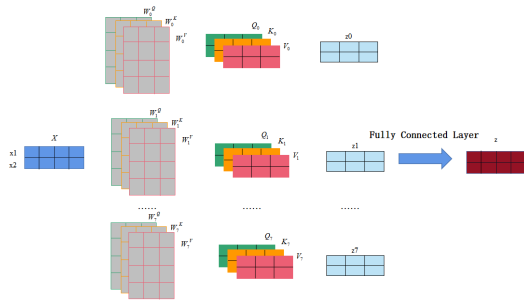$$Head_i = Attention(Q_i, K_i, V_i) \qquad （7）$$

Fig. 3. Multi-Headed mechanism

The Encoder side does not only perform self-attention based on the Multi-Headed mechanism once, it will repeat the calculation by stacking multiple layers, and after performing a self-attention, it will also use residual connection and normalize the layers (layer-normalize).The residual connection is to connect the result of the input word Embedding with the result of self-attention, the purpose is to prevent some layers from bad learning the effect, so as to ensure that the result is not worse than the original. The normalization is to limit the output result to a certain range, and the calculation process is shown in Equation 8:

$$Z_{new} = Layernormdize(concat(Z, X)) \qquad (8)$$

The residual connection is shown in Figure 4, and the layer normalization is shown in Figure 5:
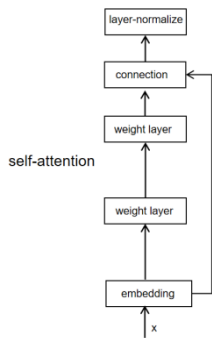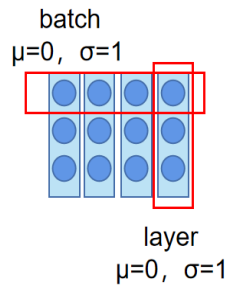

Fig. 4. The residual connection   Fig 5 . Ihe layer normalization

3.2BERT model

The BERT (Bidirectional Encoder Representations from Transformers) model is equivalent to the Encoder part of the Transformer model. The input of the model is the word vector of each word after embedding and the position vector of each word in the sentence, and this position vector is calculated by the sine function and the cosine function. The calculation formula is shown in equations 9 and 10:

$$PE_{(pos,2i)} = \sin(\frac{pos}{10000^{\frac{2i}{d_{Model}}}}) \qquad (9)$$

$$PE_{(pos,2i+1)} = \cos(\frac{pos}{10000^{\frac{2i}{d_{Model}}}}) \qquad (10)$$

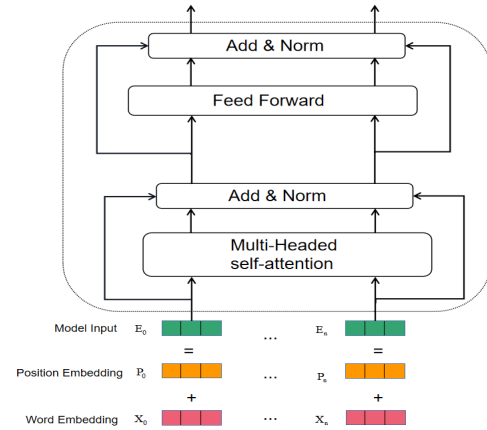The final BERT model is shown in Figure 6:

Fig 6. BERT model structure

3.3 Fine-grained text classification model based on BERT

This article uses the BERT model for fine-grained classification of rumors, and fine tune (fine tune) based on the pre-trained model officially released by Google to achieve better results. The classification model of BERT is shown in Figure 7:
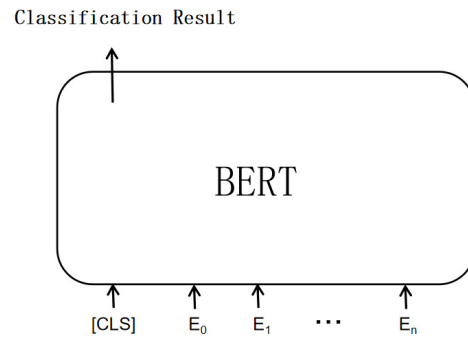

Fig 7. BERT classification model structure

4 ANALYSIS OF RESEARCH RESULTS

4.1 Research data

The data set of this paper is from the Chinese Internet joint rumor refuting platform, tencent comparative truth platform, sina news rumor hunt, baidu rumor refuting, 360 rumor refuting information public platform, after the elimination of invalid data to obtain a total of 3737 rumors related data.

In order to ensure the accuracy of the fine-grained classification of rumors in the COVID-19 Public Crisis, this article combines Yu Guoming[41]Wang Fang[42]and other related research and the rumors category of the Chinese Internet joint rumor refuting platform to classify the rumors as three categories:Scientific, healthy and social categories. The scientific rumors in the public crisis are represented by scientific research on the epidemic situation. The specific drugs for the treatment, such as COVID-19, can be transmitted through the conjunctiva, and chloroquine phosphate is the specific drug for the treatment of COVID-19.Scientific rumors in public crises manifest as

scientific research on the epidemic. Healthy rumors are manifested in aspects requiring attention to food and diet as well as prevention methods that are easy to refutable. For example, eating fruits during the epidemic period requires peeling, and drinking alcohol can prevent COVID-19.Social categories are manifested in social events, conspiracy theories, etc .For example: "Xiaotangshan" in Wuhan will be suspended and relocated Guangzhou will soon be closed. After sorting the data set, a total of 1428 rumors for health, 405 rumors for science, and 1904 rumors for society were obtained.

4.2 Parameter setting

In this paper, the BERT model is used for experiments, and the network structure is constructed through the TensoFlow framework. The network structure has a total of 12 layers, the hidden layer has 768 dimensions, the attention head number is 12, the Dropout value is set to 0.1, that is, 10% of the parameters are omitted each time, the final model has a total of 110M parameters, and other training parameters are adjusted as shown in Table 1.

TABLE I.   BERT MODEL TRAINING PARAMETERS

BERT model parameters	Model parameter values
Data size of each batch of training set	16
Learning rate	$2e-5$
Sample iteration round	10

4.3 Experimental design

1) Evaluation index

In this paper, Accuracy, Precision, Recall, and comprehensive index value of F_1 (f1-score) are used as evaluation indicators for the rumor classification model[43]. The accuracy rate refers to the correct classification of rumor information divided by the total amount of rumor information in the entire data set; The accuracy rate refers to the ratio of the amount of rumor information that the classifier correctly judges to be in this category and the total amount of rumor information that the classifier judges to belong to this category; The recall rate refers to the ratio of the amount of rumor information that the classifier correctly judges to be in this category to the total number of rumor information belonging to this category. The value of F_1 is the harmonic mean value of the precision rate and the recall rate. The maximum value is 1 and the minimum value is 0, as shown in Equation 11:

$$F_1 = \frac{2PR}{P+R} \times 100\% \qquad (11)$$

Among them P (Precision) represents the precision rate, R (Recall) represents the recall rate.

2) Comparative experiment

In order to compare with the fine-grained rumor classification method based on the BERT model in this paper, the TextCNN model and TextRNN model based on

word vectors are selected as control models, and the same pre-processed data is selected for comparative experiments. The main parameter design of the two models is shown in Table 2:

TABLE II.   OTHER MODEL TRAINING PARAMETERS

TextCNN		TextRNN	
Model parameters	Parameter value	Model parameters	Parameter value
Number of convolution kernels	256	LSTM hidden layer	128
Convolution kernel size	(2, 3, 4)	LSTM layers	2
Learning rate	$1e-3$	Learning rate	$1e-3$
Iteration round	10	Iteration round	10

4.4 Experimental results and analysis

After a series of classification experiments, the results obtained on the test set are shown in Table 3:

TABLE III.   EXPERIMENTAL RESULTS

Model	Accuracy rate/%	Precision rate /%	Recall rate/%	Value of F1/%
BERT	99.20	99.17	98.13	98.34
TextCNN	98.40	97.03	97.46	97.24
TextRNN	98.40	97.43	96.82	97.12

From the experimental results in Table 3, it can be seen that the prediction accuracy of the BERT, TextCNN, and TextRNN models has all reached more than 90%, which shows that the deep learning model is applied to the rumor fine-grained classification task of the COVID-19 event. In these four types of models, based on the performance of BERT's evaluation indicators, the accuracy rate is 0.8 percentage points ahead of the TextCNN and TextRNN models; the effect on accuracy rate is also nearly two percentage points ahead of the other two models. In comparing the two indicators of recall rate and value, the BERT model is still in a leading position compared with other models. From the overall comparison, the effects of TextCNN and TextRNN models are basically similar, while the BERT model achieves the optimal effect for this text classification task.

The above experimental results show that the three models have achieved good results on the test data set. TextCNN extracts text features better through convolution operations, thus achieving better classification results; The TextRNN model can not only extract text features, but also capture the relationship between semantics and context, so it is superior to the TextCNN model in some evaluation indicators; The comparison shows that the model in this paper is superior to the other two models in three indicators. This is because the BERT model introduces a multi-head attention mechanism. For a certain word, it can pay attention to the corresponding word in the sentence from different angles, and in After adding the position vector, it has better ability to capture distance features than the TextRNN model.The TextCNN model is limited by the

length of the CNN convolution kernel and the number of convolution kernels, and its ability to extract semantic features is not as good as the multi-head attention mechanism of the BERT model.Since CNN was originally used to extract image features, the image was composed of pixels, and its purpose was to detect a certain overall structural feature of the pixels. The pooling operation was an important step in its implementation process and would not change the image. The essence of the content is understood, but for text information, the difference in detail in the text may cause the entire sentence meaning to change.BERT's fine-tuning classification model learns word embedding features with contextual semantics through unsupervised methods in the pre-training stage, absorbs a lot of complex language knowledge, and can better express semantics. In the fine-tuning stage, it adapts to the relevant context, so its performance is better than the above model, and it can obtain a good solution to the fine-grained rumor classification of new coronary pneumonia.

5 CONCLUSION

In this paper, the BERT model is used for the fine-grained classification task of rumors in the COVID-19 event in response to the problem of complex and large differences in semantic information between rumors in public crisis events, and the comparison of other classification models in deep learning confirms that the multi-head attention mechanism and the mechanism for fine-tuning downstream tasks of the BERT model have strong semantic expression and feature extraction capabilities.

This article sorts out and analyzes the three most widely used models of online refuting rumors.

In order to deal with the spreading rumors in the public crisis, China should establish a social cooperative rumor refuting mechanism with diversified cooperation of rumor refuting subjects, and integrate the early-warning mechanism and governance system of rumor into the information system dealing with public crisis events. Through the fine-grained classification model of rumors used in this article, private rumors, field experts and government official media can finally obtain different types of rumors in a timely manner, combine professional knowledge and news data to screen information and release rumors in a timely manner.To some extent, the credibility of multi-subject joint refuting rumors can be improved, which can help to accelerate the construction of a complete collaborative rumor mechanism for individuals → media → government → experts to dispel information → information release.

REFERENCES

1. Zheng Jie. Principles and Practice of NLP Chinese Natural Language Processing [M]. Beijing: Electronic Industry Press, 2017.
2. Salton G.A vector space model for automatic indexing [J]. Communications of the Acm, 1974, 18 (11): 613-620.
3. CUI W. A chinese text classification system based on naive bayes algorithm [C] // MATEC Web of Conferences. 2016: 1015.
4. ZHANG M Y, AI X B, HU Y Z. Chinese text classification system on regulatory information based on SVM [C] // IOP Conference Series: Earth and Environmental Science. 2019: 252.
5. SAHA D. Web text classification using a neural network [C] // 2011 Second International Conference on, 2011.
6. Lei Fei. Text classification and its application based on neural network and decision tree [D]. Chengdu: University of Electronic Science and Technology, 2018.
7. Ding Shengchun, Wang Xiaoying, Liu Menglu. Classification of Internet public opinion topics based on ontology and weighted naive Bayes [J]. Modern Information, 2018, 38 (8): 12-17.
8. Jia Longjia, Zhang Bangzuo. Research on the topic classification method in the network public opinion security of colleges and universities--Taking Sina Weibo data as an example [J]. Data Analysis and Knowledge Discovery, 2018 (7): 55-62.
9. Zubiaga, A., Aker, A., Bontcheva, K., Liakata, M., & Procter, R. (2018). Detection and Resolution of Rumours in Social Media. ACM Computing Surveys, 51 (2), 1–36.
10. Zong Qianjin, Huang Zifeng, Shen Hongzhou. Research on rumors and rumors of social media users based on gender perspective [J]. Modern Intelligence, 2017, 37 (7): 25-29, 34.
11. Jiang Ying, Zhang Jing, Zhu Lingxuan, et al .. Network rumor text sentence feature analysis and monitoring system [J]. Electronic Design Engineering, 2017, 25 (23): 7-10, 15.
12. Mikolov T, Sutskever I, Chen K, et al. Distributed Representations of Words and Phrases and their Compositionality [J]. Advances in Neural Information Processing Systems, 2013: 3111-3119.
13. KIM Y. Convolutional neural networks for sentence classi- fication ［C］ / / Proceedings of the 2014 Conference on Em- pirical Methods in Natural Language Processing. Doha: EMNLP, 2014: 1746-1751.
14. Liao Xiangwen, Huang Zhi, Yang Dingda, Cheng Xueqi, Chen Guolong. Social media rumors detection based on layered attention network [J]. Science in China: Information Science, 2018, 48 (11): 1558-1574.
15. Zhiwei Jin, Juan Cao, Han Guo, Yongdong Zhang, Jiebo Luo: Multimodal Fusion with Recurrent Neural Networks for Rumor Detection on Microblogs. ACM Multimedia 2017: 795-816.
16. Yang Z, Yang D, Dyer C, etal. Hierarchical attention networks for document classification [C] // Proceedings of the 2016 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies. 2016: 1480- 1489.
17. Devlin J, Chang MW, Lee K, et al. BERT: Pre-training of Deep Bidirectional Transformers for

Language Understanding [C] // Proceedings of the 2019 Conference of the North American Chapter of the Association for Com- putational Linguistics : Human Language Technologies, Volume 1 (Long and Short Papers), 2019: 4171-4186.

18. Nicholas DiFonzo and Prashant Bordia. 2007. Rumor, gossip and urban legends. Diogenes 54, 1 (2007), 19–35.

19. Guoyong Cai, Hao Wu, and Rui Lv. 2014. Rumors detection in chinese via crowd responses. In Proceedings of the 2014 IEEE / ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM'14). IEEE, 912 – 917.

20. Ding Haiyan. The government's strategy to deal with online rumors [D]. Guangzhou University, 2013.

21. Miao Junfu, Shao Wen. Looking at the psychological mechanism of gossip from the salt robbing farce [J]. Journal of Qilu Teachers College, 2011, 26 (05): 25-27.

22. Tian Xiaorui, Wang Danchen, Dong Anbang- Research on Public Crisis Information Dissemination Mechanism under the Influence of Psychological Stress [J]. Library and Information Work, 2014, 58 (2): 59-65

23. Allport, G.W. & Postman, L. (1947). The Psychology of Rumor. NewYork: HenryHolt.

24. Prasad J. The psychology of rumour: A study relating to the great Indian earthquake of 1934 [J]. British Journal of Psychology. General Section, 1935, 26 (1): 1-15.

25. Chorus, A. The basic law of rumor [J]. The Journal of Abnormal and Social Psychology, 1953, 48 (2): 313-314.

26. Anthony, S. Anxiety and rumor. Journal of Social Psychology. 1973, 89, 91-98.

27. Hu Yu. Mass Communication Effect: Problems and Countermeasures [M]. Beijing: Xinhua Publishing House, 2000.

28. Kuang Wenbo, Guo Yufeng. "The Spread and Resolution of Rumors in the Age of Weibo-Taking the" 7.23 "Yongwen Line High-speed Railway Accident as an Example", "International Press", 2012, Issue 2, pages 64-69

29. Wu Jian, Ma Chao. Rumor spread formula, traceability + revision and development [J]. Journalism. 2015 (13): 20-23.

30. Yuan Hui, Xie Yungeng. Research on the rumors of Internet rumors of public events——Based on the content analysis of 118 public events Internet rumors with greater influence [J]. Journalist, 2015 (05): 58-65.

31. Wang Qian, Yu Feng. Improvement and verification of the rumor spreading formula of Allport and Postman: Analysis of rumors of Sina Weibo based on

the casualties caused by Siberian tigers [J]. International Press. 2017, 39 (11): 47-67.

32. Jin Xuan, Zhao Yuxian. Analysis of Internet rumors from the perspective of co-governance——WeChat platform rumors governance practice [J]. News and Writing, 2017 (6): 41-44

33. Wang Xiaoli, Zhao Laijun, Wu Zhong. Rumor propagation model considering rumor dispelling mechanism in non-uniform network [J]. System Engineering, 2015, 33 (12): 139-145.

34. Yu Guoming. "Weibo rumors" is a false proposition [N]. China Economic Times, 2012-01-06 (012).

35. Wang Guohua, Wu Dan, Wang Ge, et al. Research on False News Communication from the Perspective of Frame Theory——Based on the "Content Analysis of Shanghai Girl Escape from Jiangxi Rural Metallurgical Event [J]. Information Magazine, 2016, 35 (6) : 56-64.

36. Tong Wensheng, Yi Baihui. Internet rumors: domestic research progress and theoretical analysis framework [J / OL]. Information Magazine: 1-8.

37. Tang Xuemei, Lai Shengqiang. Research on the Government 's Strategies for Dispelling Internet Rumors in Emergencies——Taking the Taifu Middle School Incident as an Example [J]. Information Magazine, 2018, 37 (9): 95-99

38. Yao Qi, Cui Lijuan, Wang Yan, Yang Ying. The influence of social media trust on the autonomy of public network rumors in major public health emergencies [J]. Psychological Science, 2020, 43 (02): 481-487.

39. Jin Jianbin, Jiangsu Jia, Yang Hongyan. Socialized collaborative repelling rumors: actor network and operating mechanism [J]. News and Writing, 2019 (8): 33-39.

40. Li Biao, Yu Guoming. "Study on the Discourse Space and Dissemination Field of Internet Rumors in the Post-truth Era ———— Based on the Analysis of 4160 Rumours on WeChat Circle of Friends [J]. News University, 2018 (2): 103 112.

41. Yu Guoming. Text structure and expression characteristics of online rumors——Analysis of 6000+ rumors based on Tencent big data screening and identification [J]. News and Writing, 2018 (02): 53-59.

42. Wang Fang, Lian Zhixuan. Calculation of rumor truth degree in public crisis and its confrontation with positive information [J]. Library and Information, 2020 (01): 34-50.

43. Duan Dandan, Tang Jiashan, Wen Yong, Yuan Kehai. Research on Chinese short text classification algorithm based on BERT [J / OL]. Computer Engineering: 1-12.