A Microblog Content Credibility Evaluation Model Based on the Influence of Sentiment Polarity

Jinglong Yao, Ling Xing, and Honghai Wu
College of Information Engineering, Henan University of Science and Technology, Luoyang 471023, China
Correspondence should be addressed to Ling Xing; xingling_my@haust.edu.cn
Received 18 October 2021; Accepted 13 May 2022; Published 31 May 2022
Academic Editor: Sikandar Ali
Copyright © 2022 Jinglong Yao et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.
False content in microblogs affects users’ judgment of facts. An evaluation of microblog content credibility can find false information as soon as possible, which ensures that social networks maintain a positive environment. The influence of sentiment polarity can be used to analyze the correlation between sentiment polarity in comments and Weibo content through semantic features and sentiment features in comments, to improve the effect of content credibility assessment. This paper proposes a Weibo content credibility evaluation model, CEISP (Credibility Evaluation based on the Influence of Sentiment Polarity). The semantic features of microblog content are extracted by a bidirectional-local information processing network. Bidirectional long short-term memory (BiLSTM) is used to mine the sentiment features of comments. The attention mechanism is used to capture the impact of different sentiment polarities in comments, and the influence of sentiment polarities is obtained for the credibility assessment of microblog content. The experimental results on real datasets show that the evaluation performance of the CEISP model is improved compared with the comparison model. Compared with the existing Att-BiLSTM model, the evaluation accuracy of the CEISP model is improved by 0.0167.

1. Introduction

With the development of social platforms represented by Weibo, the speed of information exchange among users has continuously improved [1]. However, the spread of false content on the platform will cause great harm to society [2, 3]. Therefore, it is hugely important to evaluate the credibility of microblog content. Sentiment analysis is done by extracting users’ views or opinions about events [4]. To some extent, this reflects the truth of the content, and there is a growing number of studies using sentiment analysis data to evaluate content credibility. However, most of the relevant studies put the sentiment information directly into the model for training, thereby lacking a correlation between the sentiment information and the content to be evaluated, which makes for insufficient use of the sentiment information.

The influence of sentiment polarity refers to the influence degree of different sentiment polarities on text content. Sentiment polarity refers to the attitude or position of users on events obtained by sentiment analysis. The results of sentiment analysis in this paper are derived from the comments. In the microblog, the content published by users allows other users to comment. Events with high attention will get more comments, and the comments contain a great deal of sentiment information, playing an important role in allowing the authenticity of the content to be judged by conducting sentiment analysis of the entire comment set to obtain the overall position of the group users. Related work has begun to use the emotional information of comments for credibility evaluation. Zhang et al. proposed the role of emotional polarity in credibility evaluation [5]. Subsequently, Ajao et al. confirmed the positive role of emotional information in credibility evaluation by hypothetical method [6]. Liu and Yang assigned weights to various pieces of users’ personal information in Sina Weibo comments and completed sentiment analysis with the help of comment content, before finally completing the evaluation of content credibility as the input of a rumor classifier [7].
two different dimensions [8], it is difficult to consider the relationship between them. Therefore, we have made some improvements.

This paper proposes a microblog content credibility evaluation model based on emotional polarity influence. Emotional polarity influence is obtained through the interaction between microblog semantic features and comment emotional features, which are used for microblog content credibility evaluation. The main contributions of this paper include the following: (1) We proposed a credibility evaluation model of microblog content based on the influence of sentiment polarity. This method used a bidirectional-local information processing network to mine microblog text features, used BiLSTM to obtain comment emotional features, and realized credibility evaluation through the fusion of text features and emotional features. (2) We used the attention mechanism to analyze the impact of different emotional tendencies on the emotional characteristics of text and comments and used the full connection layer to generate the evaluation results of microblog credibility. (3) The effectiveness of the model is verified by comparison experiments on real datasets, and the model has good performance.

2. Related Work

2.1. Sentiment Analysis. Sentiment analysis is an important task in natural language processing, which makes the system of content credibility evaluation more complete by using the views of many users. At first, researchers used the emotional dictionary to match the similarity of word vectors in the corpus for sentiment analysis [9]. However, in real life, the sentiment value of the same word changes over time, and words conveying emotional information are constantly created, resulting in the unsatisfactory effect of sentiment analysis using an outdated emotional dictionary. With the development of deep learning in natural language processing-related research applications, researchers found that the related methods can solve the problem of overreliance on the emotional dictionary in the experiment effectively [10, 11]. Santos et al. used a convolution neural network (CNN) model to obtain the local features of sentences or words by using two convolution layers, and they targeted the mining of semantic information to improve the sentiment analysis effect of short texts [12]. Irsoy and Cardie added sequential characteristics to the model of a recurrent neural network (RNN) to obtain sentence representation on timeseries information, which further improved the accuracy of sentiment analysis [13]. Tai et al. applied long short-term memory (LSTM) to the topology of a tree-structured network; compared with conventional LSTM, the hidden state of a tree-structured LSTM contains more information [14]. Baziots et al. introduced an attention mechanism into LSTM to amplify the role of words with stronger influences in sentiment analysis [15, 16].

As there are many words without sentiment information in the comments, sentiment words exist in the form of sparse distribution. Mining single-directional or local features will often omit sentiment information, leading to certain deviations in the credibility assessment of content. Xu et al. used BiLSTM, thereby solving the problem that LSTM can only mine information in one direction, and they were able to capture contextual information more effectively when obtaining the sentiment features of comments, which have better effects when applied to the evaluation of content credibility [17].

2.2. Deep Learning. Credibility-related fields in social networks have always been a research hotspot. The method based on machine learning uses a classifier to classify the true and false content on the basis of feature engineering [18–20]. Castillo et al. applied the machine learning algorithm to the credibility evaluation task. They proposed four basic features according to the data types, added artificial annotations, and used the optimal priority selection method to select 15 suitable features for classification [21]. However, due to the lack of feature analysis, the overall evaluation effect was unsatisfactory.

In recent years, more and more researchers have begun to use the relevant methods of deep neural networks (DNNs) in deep learning to conduct content credibility evaluation research [22–24]. Ma et al. introduced a DNN into the credibility evaluation task. They used LSTM to model and analyze the text sequence data in the time dimension to obtain the implied features of rumor context changing with time, before completing the evaluation [25]. Xu et al. used LSTM to learn the representation of an original tweet, and they introduced a content attention mechanism to aggregate the keywords in the original tweet [26]. Ghanem et al. used the LSTM model with an attention mechanism to identify false information on the condition of relying only on emotional features [27]. Guo et al. used BiLSTM to process data in two directions, overcoming the limitations of one-way LSTM and obtaining text context information [28]. Alsaeedi and Al-Sarem used a CNN to unearth the most important semantic features in the content locally, which have more symbolic significance when evaluating the credibility of the content [29]. Ajao et al. attempted to combine the characteristics of LSTM and CNN, and they proposed using LSTM-CNN for false news detection. LSTM was used for data sequence classification, and a CNN layer was added after an LSTM layer for local sequence feature extraction [30]. Lv et al. proposed a microblog rumor detection method based on comment emotion and CNN-LSTM, which gauged the truth of the content according to the probability of whether the content involved a rumor or not, as well as the sentiment difference of the comments [31].

Because part of the speech difference between words in microblog content may be too great when using a deep learning method to extract semantic features of microblog content, only single-directional features or local features are considered, which may have limitations and contingencies in credibility assessment.

2.3. Content Credibility Evaluation. The results of sentiment analysis reflect the real position of users, which can be used to evaluate the credibility of microblog content from the
perspective of many users. Sivasangari et al. divided the score value of the emotional dictionary for the selected data to separate false content more accurately [32]. Wang and Guo encoded sentiment information into the time-series division process, and they completed the content credibility evaluation by capturing the changes of context and sentiment information over time [33]. To solve the problem of a lack of positional information in user comments, Tian et al. proposed a method combining CNN and Bert to transfer learning from the data information to predict the user’s position, and they used more complete user sentiment polarity to evaluate the content credibility [34].

Using sentiment analysis to evaluate content credibility can increase the performance of assessment effectively, but the most relevant methods lack the degree of correlation between sentiment information and the content to be evaluated, which makes it difficult to use emotional information effectively.

3. CEISP Model

To apply the influence of different sentiment polarities on the credibility of microblog content to the credibility evaluation with the help of the influence of sentiment polarity, this paper proposes a microblog CEISP model. The model structure is shown in Figure 1.

3.1. Problem Definition

Definition 1. Raw text sequence. The S text content published in the microblog is \( T = \{X_1, X_2, \ldots, X_n, \ldots, X_s\} \), where any text \( X_n = \{w_1, w_2, \ldots, w_l\} \), that is, text \( X_{\text{input}} \) is a sequence of \( s \) words.

Definition 2. Comment text sequence. The content set of \( s \) comments is under the Chinese text of microblogs \( T' = \{x'_1, x'_2, \ldots, x'_{m}, \ldots, x'_d\} \), where any comment \( x'_m = \{w'_1, w'_2, \ldots, w'_{l'}, \ldots, w'_j\} \), that is, comment \( x'_{\text{input}} \) is a sequence of \( j \) words.

Definition 3. The influence of sentiment polarity. The influence of sentiment polarity is defined as \( q_{\text{pol}} \), used to indicate the influence of positive or negative emotion in a comment, sentiment polarity \( i = \begin{cases} 1, & \text{positive} \\ 0, & \text{negative} \end{cases} \).

Definition 4. Credibility evaluation. Given a message text \( X_p \) in Weibo, \( \{O', S', S''\} \) is a triple, representing the semantic feature, comment sentiment feature, and the influence of sentiment polarity. Then, the credibility of the information text \( p_s = f(O', S', S'') \), that is, target \( p_s \) is learning mapping function \( f \), which evaluates the credibility of microblog content.

3.2. Raw Data Processing. Due to the lack of large open and complete datasets in the credibility assessment of microblog content, most relevant studies use the application program interface provided by the platform to obtain data and conduct experiments [35]. Therefore, this paper builds a content credibility evaluation of experimental data based on Sina Weibo, which collects 1500 real contents, 1500 fake contents, and about 600000 comments on those contents, and the formed dataset is named CSA Dataset. Untrustworthy Weibo data is derived from the result of the efforts of the publicity office of Sina Weibo’s Community Management Center in reporting false information, and the content here has been confirmed as false by Sina Weibo officially. To ensure that the obtained content has enough comment data, we selected fake Weibo content where the number of comments is greater than 50 to build the dataset. Some fake content was also excluded from the dataset because the publisher controlled the browsing permission or deleted the original information. In addition, the emoticon in comments in the form of "[]" and "." comprised Weibo’s basic forms of written expression, such as "[开心]" and "[哭泣]."

Before the experiment, text information in the dataset needs to be preprocessed. In this paper, a Jieba toolkit is primarily used for related operations. First, the Jieba word participle is used for word segmentation, and then the Chinese stop-word “table,” the Baidu Chinese stop-word “table,” the Hit stop-word “table,” and the machine intelligence laboratory of Sichuan University’s stop-word database are combined to create a general stop-word database to delete the stop-word. The word embedding layer in this paper mainly extracts the features of words from the corpus after word segmentation and word filtering. The GloVe model is used to train the dataset and obtain the word vector as the input of the model. However, because the length of each vector is different, we use truncation or filling operations in natural language processing to set the vector dimension reasonably according to the situation of the dataset and to obtain the vector with the same dimension. For the content that needs to be filled, we fill it uniformly at the front of the content. The steps of data preprocessing are shown in Figure 2.

In addition, since this paper involves sentiment analysis of the comment content, spam or useless comments should be filtered and deleted. Generally, spam comments are less like the original Weibo content, and most of them are statements with a high proportion of nouns. Therefore, filtering is based on the characteristics of similarity and the proportion of nouns. The comments without text content (the forwarding operation in microblogs and the comments may appear without content in the dataset) are filtered and processed directly.

3.3. Semantic Representation of Microblog Text. A semantic representation of microblog text extracts the semantic features of microblog content by using a bidirectional-local information processing network. The network can be divided into an embedding layer, a BiLSTM layer, a CNN layer, and an output layer. The implementation of each layer is as follows.

3.3.1. Embedding Layer. To transform the data into a matrix form, it is necessary to use the embedding layer. In
microblogs, the content set $T = \{X_1, X_2, \ldots, X_a, \ldots, X_s\}$, where $X_a$ represents any text content and $s$ represents the amount of text. For any text $X_a$ in $T$, $X_a \in \mathbb{R}^{w_1, w_2, \ldots, w_c}$, after data preprocessing, each text information vector $X_a$ is obtained. $X_a = \{x_1, x_2, \ldots, x_i, \ldots, x_L\}$, and $X_a \in \mathbb{R}^L$, where $L$ is the maximum length of the sequence. Through the embedding layer, each $x_i \in X_a$ is transformed into statement embedding matrix, as shown in the following equation:

$$x'_i = W_{X}x_i,$$

where $W_{X}$ is the word embedding matrix. After processing $x'_i \in \mathbb{R}^{K_x}$, the statement matrix $\{x'_i\}_{i=1}^L$ is entered into the BiLSTM layer.

3.3.2. BiLSTM Layer. In the one-way LSTM layer, the LSTM unit only considers the information of the former unit and ignores the information of the latter unit. To solve this problem, the BiLSTM layer can be used to process text encoding and obtain information with richer context. The BiLSTM layer contains a forward LSTM layer and a backward LSTM layer. At each time point $t$, each LSTM cell corresponds to an input gate $i$, a forget gate $f$, an output gate $o$, and a state cell $c$. The forget gate $f$ decides at the last minute to forget information in the memory unit. The input gate $i$, based on the current emphasis on input level, determines updating the information in the memory unit. The output gate $o$ determines the LSTM unit’s output value. The forward LSTM layer connects two adjacent unit processing sequences from left to right, such as the current cell input $x_1$ and the hidden state $h_{t-1}$ of the previous cell input. For the given input sequence $x_1, x_2, \ldots, x_{L-1}$, the forward LSTM layer produces output sequence $\tilde{h}_t$. The calculation formula of the forward LSTM layer is shown in the following equation:

$$\tilde{h}_t = W_{\tilde{h}}x_t,$$

where $W_{\tilde{h}}$ is the word embedding matrix. After processing $\tilde{h}_t \in \mathbb{R}^{K_{\tilde{h}}}$, the statement matrix $\{\tilde{h}_t\}_{t=1}^L$ is entered into the BiLSTM layer.
\[
\begin{aligned}
  f_t &= \sigma(W_f x_t' + U_f h_{t-1} + b_f), \\
  i_t &= \sigma(W_i x_t' + U_i h_{t-1} + b_i), \\
  c_t &= f_t \odot c_{t-1} + i_t \odot \tanh(W_c x_t' + U_c h_{t-1} + b_c), \\
  o_t &= \sigma(W_o x_t' + U_o h_{t-1} + b_o), \\
  h_t &= o_t \odot \tanh(c_t),
\end{aligned}
\]

where \( W_f, W_i, W_o, \) and \( W_c \) represent the input gate weight matrices. \( U_f, U_i, U_o, \) and \( U_c \) represent the output gate weight matrices. \( b_f, b_i, b_o, \) and \( b_c \) represent bias vectors. They are all parameters to be learned in the BiLSTM layer. \( \sigma \) is the sigmoid activation function. \( \tanh(g) \) is the hyperbolic tangent, and \( g = H' \) is the Hadamard product.

\( h \) and \( h' \) represent the forward and backward context representations, respectively, which are then merged into the new state matrix \( H = [h_1, h_2, \ldots, h_L] \), and \( H \in \mathbb{R}^{L \times k \times s} \). The forward and backward outputs are combined through the following equation:

\[
h = h' \oplus h.
\]

The statement matrix is then sent to the CNN layer.

3.3. CNN Layer. To obtain locally important features of the whole data sequence based on obtaining context information, it is necessary to go through the CNN layer. The convolution operation involves a statement matrix and a filter matrix, assuming a convolution filter \( W' \in \mathbb{R}^{k \times s} \), where \( k \) is the filter window size. To generate the eigenmap, each independent word region in the output matrix \( H \) obtained from the BiLSTM layer is put into the convolution filter as the input matrix of CNN, and the convolution operation is performed to obtain the eigenvalue \( v_h \), as shown in the following equation:

\[
v_h = g'(W' \odot H_{j,m} + b),
\]

where \( j \) represents the local features within the partial range and \( m \) represents the local features within the partial range in the opposite direction; the value range of \( j \) and \( m \) starts from 1 and cannot exceed the limit of filter matrix. \( b \) represents bias vectors. \( \odot \) represents the convolution operator. \( g'(g) \) is a nonlinear activation function, which could be \( \tanh \) or ReLU and so forth, and we use ReLU. After performing the convolution operation, matrix \( V \) is shown in the following equation:

\[
V = [v_{1,1}, v_{1,2}, \ldots, v_{j,m}].
\]

Then, matrix \( V \) obtained through the convolution layer is inputted into the pooling layer using max-pooling to obtain the most important features, as shown in the following equation:

\[
\bar{v} = \max[V].
\]

3.3.4. Output Layer. Finally, the embedding matrix of semantic features is obtained from the Weibo dataset, as shown in the following equation:

\[
O' = [o_{1,1}, o_{1,2}, \ldots, o_{j,m}], \quad O' \in \mathbb{R}^{L \times s \times k}.
\]

3.4. Sentiment Representation of Microblog Comments. The BiLSTM model can be used to mine the sentiment features of words in comments, and the cellular state in the LSTM unit can ensure the learning of sequence dependencies between contexts. For the extraction of sentiment features, comment information was obtained through the channel supervised by the emotion dictionary and the original channel without an emotion dictionary marker, before being put into the forward LSTM and the backward LSTM for feature extraction. A collection of comment text content in a comment dataset is as follows: \( T' = \{x'_1, x'_2, \ldots, x'_n\} \), where \( x'_n \) represents any text content and \( s \) represents the amount of comment text for any comment \( x'_n = [w_1, w_2, \ldots, w_s, w'_1] \). Through the embedding layer, the statement embedding matrix was obtained, and then the sentiment feature embedding matrix representation \( S' = \{s_{1,1}, s_{1,2}, \ldots, s_{j,m}\} \), \( S' \in \mathbb{R}^{L \times s \times k} \), was extracted by the BiLSTM.

Then, the k-means++ algorithm was used to divide the sentiment features [36]; the sentiment feature embedding matrix representation \( S' \) was used as the input of the clustering algorithm. Finally, emotional polarity \( C = \{c_1, c_2\} \) is obtained, so the number of clusters is two.

In the algorithm, two initial clustering centers need to be selected. The k-means algorithm determines the initial clustering center through random selection, and the resulting clustering results depend on the selection of the initial clustering center. If the selection of the initial clustering center is not good, the problem of local optimal solution will occur. To solve this problem, this paper uses an improved k-means++ algorithm.

First, we should make the cluster center set \( C' \) be initialized randomly. In the input set \( S' \) we can select a sample point randomly as the first cluster center and then calculate the shortest distance between the remaining sample points and all the current cluster centers, in order to make possible any point being selected as the next cluster center. According to this probability, we can select the cluster center and put it
into $C$. Then, using the standard k-means algorithm, we can obtain the emotional polarity $C$ by clustering, and, in the cluster center set $C = \{c_i^0\}_{i=1}^l$, for each $c_i^0 \in C$, $c_i^0$ is used as the feature vector of sentiment polarity $C^i$, $c_i^0 \in \mathbb{R}^{K_x}$, and $C^i \in \mathbb{R}^{K_x \times K_x}$.

3.5. Credibility Evaluation Based on the Influence of Sentiment Polarity. For the content posted by users in microblogs, the sentiment characteristics contained in the comments can reflect the authenticity of the original content to some extent. The model proposed in this paper uses the influence of sentiment polarity to capture the reduction of sentiment features in comments to the content credibility, to obtain the evaluation results of content credibility. In the process of computing the influence of sentiment polarity, the microblog semantic feature embedding matrix $O'$ and the feature vector of sentiment polarity $c_i^0$ will participate in the training of the model. The input dimension is $O' \in \mathbb{R}^{L \times K_x}$, presented for convenience as "$L = 50, K_x = 128$" in this article.

To facilitate the calculation of the influence of sentiment polarity, the idea of average pooling in a CNN is used to transform the semantic feature embedding matrix $O' \in \mathbb{R}^{L \times K_x}$ into a vector representation according to the following equation:

$$o' = \frac{Z_i^o O'}{L},$$

(9)

where $Z_i^o \in \mathbb{R}^{1 \times L}$ is a vector whose element is one. After transformation, the eigenrepresentation is transformed from an eigenmatrix to vector form.

The calculation formula for defining the influence of sentiment polarity is shown in the following equation:

$$q_i = W_C^{i} (\tanh(W_C^{o} o') \circ \tanh(W_C^{o} o')),$$

(10)

where $q_i$ represents the influence of different emotions in the comments. In equation (10), the parameters are $W_C^{i} \in \mathbb{R}^{1 \times 128}$, $W_C^{o} \in \mathbb{R}^{128 \times 128}$, and $W_C^{o} \in \mathbb{R}^{128 \times 128}$. Thus, the correlation between the two types of sentiment polarities and the semantic features of content and the sentiment features of comments can be obtained. The influence of sentiment polarity is calculated through semantic features of content and sentiment features of comments, and then the influence of different sentiment polarities on content credibility is explored.

Next, the influence of sentiment polarity is assigned a probability distribution $a^C$, as shown in the following equation:

$$a^C = \text{Softmax}(q^C).$$

(11)

Based on equation (11), the influence expression of sentiment polarity based on different attention distribution weights is obtained, as shown in the following equation:

$$s'' = \sum_{i=1}^{k} a_i^C \cdot c_i^0.$$

(12)

So far, the inputs of the model are Weibo’s semantic feature $o'$ and the emotional polarity influence $s''$, which are combined to obtain input $m'$ of the full connection layer, as shown in the following equation:

$$m' = a^O s' + a^S s''.$$

(13)

Then, two full connection layers and the softmax function are used to generate the final evaluation results. The calculation process is shown in the following two equations:

$$y' = W_M^{f} g' (W_M^{m} m' + b_M^{m}) + b_M^{m'}.$$  

(14)

$$p_i = \text{softmax}(y'),$$

(15)

where $W_M^{f} \in \mathbb{R}^{128 \times 128}$ and $b_M^{m} \in \mathbb{R}^{128}$ are learning parameters of the first connection layer, $W_M^{m} \in \mathbb{R}^{128 \times 128}$ and $b_M^{m} \in \mathbb{R}^{128}$ are learning parameters of the second connection layer, and $g' (\cdot)$ is the ReLU activation function. The loss function is the cross-entropy loss function, which is shown in the following equation:

$$\mathcal{L} = - \sum_{i=1}^{n} y_i \log (p_i),$$

(16)

where $y_i$ represents the actual probability that the sample belongs to class $i$. If the prediction is true, the value of $i$ is 1; otherwise, it is 0.

4. Experiment Results and Evaluation

4.1. Experimental Environment and Dataset Division. In this paper, the experimental hardware platform is Intel Xeon (2.20 GHz), with 12 G memory and 16 GB NVIDIA Tesla P100. The experimental software platform is Ubuntu 18.04 operating system and development environment is Python 3.6 programming language. As for the experimental parameters, in addition to the parameter settings mentioned in relevant literature in the model listed in Section 4.2, the adjustable parameter settings of the CEISP are shown in Table 1.

To evaluate the prediction results of the model, the dataset was divided into a training set, validation set, and test set. Almost 70% of the datasets were used as training datasets. A third of the remaining datasets was randomly selected as verification datasets to determine the optimal parameters, and two-thirds were used as test datasets to evaluate the accuracy. Table 2 shows the division of the microblog dataset, CSA Dataset, constructed in this paper.

4.2. Evaluation Index. At present, most of the relevant research on content credibility evaluation is aimed at the dataset with a certain amount of data. During learning, the training set and verification set are used for feature mining and parameter learning, and then the data obtained by the model are used to evaluate the data in the test set. The model ability is reflected by the evaluation index. The higher the accuracy, precision, recall, and F1 in the evaluation index, the better the model ability for credibility evaluation. The evaluation index is calculated according to the confusion matrix in Table 3. The calculation formulas of accuracy, precision, recall, and F1 are shown in the following equation:
Accuracy  
\[ \text{Accuracy} = \frac{TP + TN}{TP + FP + FN + TN} \]  
\[ \text{Precision} = \frac{TP}{TP + FP} \]  
\[ \text{Recall} = \frac{TP}{TP + FN} \]  
\[ F1 = \frac{2 \times \text{precision} \times \text{recall}}{\text{precision} + \text{recall}} \]  

In the confusion matrix, TP represents the data evaluated as true, and it is actually true; TN indicates that the data is evaluated as false, and it is actually false; FN indicates that the data is evaluated as false, but it is actually true; FP represents the data that is evaluated as true, but it is actually false.

In the research of credibility evaluation task, relevant methods are proposed, hoping to identify all malicious data as much as possible. What is reflected in the evaluation index is that the recall rate should be high. In some cases, there will be some contradictions between accuracy and recall rate. When one is higher, the other will be lower. In addition, the real data cannot be identified as false, which will increase the later investigation work and may also miss the investigation and cause impact. Therefore, the precision of credibility evaluation should also reach a certain level. Considering these problems, most researchers usually use the harmonic average F1 of precision and recall rate to measure the relationship between the two values in credibility research. The higher the F1 value, the better the overall effect of the evaluation method or model.

4.3. Experimental Comparison Model. To test the performance of the CEISP model for content reliability evaluation, experimental comparisons will be made with the following model methods, including classifier methods as support vector machines (SVMs) and deep learning methods (CNN, H-BLSTM, and Att-BiLSTM).

SVMs have been widely used in content credibility classification tasks and have achieved good results. We use the related parameters of machine learning in literature and use the classification based on the semantic features of microblogs to conduct a comparative experiment [37].

For CNN [38], by inputting the text sequence into the convolutional neural network, the semantic features of the local key points of the data are extracted to evaluate the content credibility.

For H-BLSTM [28], the context content representation is encoded by a bidirectional LSTM, and the relevant contextual semantic features are obtained to evaluate the content credibility.

For attention-based bidirectional long short-term memory [39] (Att-BiLSTM), the attention mechanism is added based on the BiLSTM model. Text words are embedded and encoded by BiLSTM and inputted into the attention layer to learn a context vector containing attention information for content credibility assessment.

4.4. Experimental Results and Analysis

4.4.1. Parameter Analysis of the CEISP Model. To test the influence of parameter settings on the accuracy of the model, different parameter combinations are set in this paper and applied to the microblogging dataset, CSA Dataset, for experiments to evaluate the results of content credibility assessment. Table 4 shows the settings of the seven groups of parameters of the model, and the accuracy of the test results is shown in Figure 3.

By setting different parameters, when the number of filters is set to eight, the BiLSTM cell size is set to 230, the filter size is set to three, and the better evaluation accuracy is 0.8850. By increasing the size of BiLSTM cells and reducing the number of filters, the performance of the model can be improved effectively.

4.4.2. CEISP Model Classification Comparison Experiment. In this paper, under the CSA Dataset, multiple models are used to conduct classification and comparison experiments to verify the validity of the microblog content credibility evaluation model based on emotional polarity influence. To ensure the objectivity and fairness of the experiment, the data preprocessing methods are the same as the model in this paper, and the final experimental results use evaluation indexes related to the reliability assessment task. The accuracy comparison results of each model under the microblog dataset, CSA Dataset, are shown in Figure 4. The experimental results of the accuracy are shown in Figure 5.
The experimental results of the recall rate are shown in Figure 6. The experimental results of the F1 values are shown in Figure 7.

Since the credibility evaluation of microblog content is related to the research of natural language processing, the results of the assessment are more dependent on the selection of data and the processing method of the data in the model. The SVM model is limited by the difficulty to obtain high-dimensional and complex feature data, so the evaluation effect is slightly weak. The CNN model makes use of local key features for evaluation, while the H-BLSTM model makes use of relevant features of context information for evaluation. These two models have different focuses on data mining, and the H-BLSTM model has a slightly better evaluation effect in this experiment. The Att-BiLSTM model adds an attention mechanism based on the H-BLSTM model, and it obtains context vectors that have a greater influence on the model, so the evaluation effect is better.

The primary reason for the improved performance of the CEISP model is that it utilizes the influence of sentiment polarity. It also uses the attention mechanism to obtain the correlation degree between semantic features of content and sentiment features of comments, and it integrates the influence degree of a difference in the sentiment polarity of comments on content into the credibility evaluation. In addition, a bidirectional-local information processing network is used. Feature vectors are first entered in the BiLSTM layer, and the obtained data with contextual information are sent to the CNN layer. Then, the features that are more important for credibility evaluation are extracted. Thus, the overall evaluation effect is better.

In general, the evaluation ability of the model is related to many factors, such as the selection of dataset. Generally, Chinese datasets contain more semantic features than English datasets, and Chinese datasets perform better when evaluated separately using the same model. In this paper, the dataset we constructed contains a large number of comments, and the CEISP model uses k-means++ algorithm to obtain the sentiment features of comments, so that the final sentiment polarity can always be obtained in the comments, so it has achieved good evaluation results.
4.4.3. Simplified Model Testing. To verify the effectiveness of the CEISP model’s characteristics and the influence of sentiment polarity, the following models are constructed in this paper: (1) CEISP-NO: experiments were conducted solely using sentiment polarity influences, and semantic features were removed; (2) CEISP-NS: the semantic characteristics of content and sentiment characteristics of comments were used in the experiment, and the influence of sentiment polarity was not considered; (3) CEISP-S: an experiment with only sentiment traits was conducted. These models all take the fusion vector of the data used as input, and the experimental results are shown in Table 5.

Compared with the CEISP-S model, the F1 value of the CEISP-NO model is increased by 0.0233, indicating that the sentiment polarity influence can obtain more accurate classification results in the evaluation of content credibility than the sentiment characteristics. This is because the method of obtaining sentiment polarity in this paper summarizes the sentiment characteristics of the whole comment set, which is more concise and representative. The most accurate classification results can be obtained by combining the influence of sentiment polarity based on semantic features. The effectiveness of the influence of each feature and sentiment polarity of the model was verified.
Table 5: Experimental results of model simplification test.

<table>
<thead>
<tr>
<th>Model</th>
<th>Classification</th>
<th>Accuracy</th>
<th>Precision</th>
<th>Recall</th>
<th>F1-measure</th>
</tr>
</thead>
<tbody>
<tr>
<td>CEISP-NO</td>
<td>Untrusted</td>
<td>0.8567</td>
<td>0.8408</td>
<td>0.8800</td>
<td>0.8599</td>
</tr>
<tr>
<td></td>
<td>Trusted</td>
<td></td>
<td>0.8741</td>
<td>0.8333</td>
<td>0.8532</td>
</tr>
<tr>
<td>CEISP-NS</td>
<td>Untrusted</td>
<td>0.8633</td>
<td>0.8682</td>
<td>0.8567</td>
<td>0.8624</td>
</tr>
<tr>
<td></td>
<td>Trusted</td>
<td></td>
<td>0.8586</td>
<td>0.8700</td>
<td>0.8642</td>
</tr>
<tr>
<td>CEISP-S</td>
<td>Untrusted</td>
<td>0.8333</td>
<td>0.8408</td>
<td>0.8533</td>
<td>0.8366</td>
</tr>
<tr>
<td></td>
<td>Trusted</td>
<td></td>
<td>0.8472</td>
<td>0.8133</td>
<td>0.8299</td>
</tr>
<tr>
<td>CEISP</td>
<td>Untrusted</td>
<td>0.8850</td>
<td>0.8787</td>
<td>0.8933</td>
<td>0.8840</td>
</tr>
<tr>
<td></td>
<td>Trusted</td>
<td></td>
<td>0.8881</td>
<td>0.8767</td>
<td>0.8860</td>
</tr>
</tbody>
</table>

5. Conclusions
This paper puts forward a kind of sentiment polarity based on an influential Weibo content CEISP credibility evaluation model. The model uses a bidirectional-local information processing network to determine Weibo semantic characteristics. It also uses BiLSTM access to comment on sentiment characteristics and then an attention mechanism to examine the influence of different sentiment polarities for text content, obtaining sentiment polarity influence. Finally, semantic features and sentiment polarity influence are used to conduct modeling learning through two fully connected layers to evaluate the credibility of microblog content. Experiments have proved the validity of the proposed model.

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

Conflicts of Interest
The authors declare that they have no conflicts of interest.

Acknowledgments
This research was supported by the National Natural Science Foundation of China (Grants nos. 62171180, 62071170, and 62072158), the Henan Province Science Fund for Distinguished Young Scholars (222300420006), the Program for Innovative Research Team in University of Henan Province (21IRTSTHN015), the Key Science and Research Program in the University of Henan Province, under Grant 21A510001, and the Science and Technology Research Project of Henan Province, under Grant 22210221001.

References


