

Article

Sentiment Classification Method Based on Blending of Emoticons and Short Texts

Haochen Zou ^{1,*} and Kun Xiang ²

¹ Department of Computer Science and Software Engineering, Concordia University, Montreal, QC H3G 1M8, Canada

² Department of Science and Engineering, Hosei University, Koganei 184-8584, Tokyo, Japan; kun.xiang.2u@stu.hosei.ac.jp

* Correspondence: haochen.zou@mail.concordia.ca

Abstract: With the development of Internet technology, short texts have gradually become the main medium for people to obtain information and communicate. Short text reduces the threshold of information production and reading by virtue of its short length, which is in line with the trend of fragmented reading in the context of the current fast-paced life. In addition, short texts contain emojis to make the communication immersive. However, short-text content means it contains relatively little information, which is not conducive to the analysis of sentiment characteristics. Therefore, this paper proposes a sentiment classification method based on the blending of emoticons and short-text content. Emoticons and short-text content are transformed into vectors, and the corresponding word vector and emoticon vector are connected into a sentencing matrix in turn. The sentence matrix is input into a convolution neural network classification model for classification. The results indicate that, compared with existing methods, the proposed method improves the accuracy of analysis.

Keywords: sentiment analysis; convolutional neural network; emoticon vectorization algorithm



Citation: Zou, H.; Xiang, K.

Sentiment Classification Method Based on Blending of Emoticons and Short Texts. *Entropy* **2022**, *24*, 398. <https://doi.org/10.3390/e24030398>

Academic Editor: Adam Lipowski

Received: 5 January 2022

Accepted: 11 March 2022

Published: 12 March 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

As an important media platform for spreading social events, the Internet plays a significant role in social events [1]. With the rapid development and maturity of Internet technology, many online social platforms have gradually become the main media for people to obtain information and communicate with each other. Twitter as a social network platform is popular because of its real-time, convenient, and interactive characteristics [2]. The burgeoning increase of Twitter and other social platforms depends on the following two points. First, the short length of tweet text reduces the threshold of information production and reading, catering to the trend of fragmented reading in the current fast-paced life [3]. Second, social network content such as tweets can contain texts, emojis, pictures, videos, and other forms, which makes up for the lack of pure text communication compared with face-to-face communication and makes text communication more immersive and accurate [4]. Users can log in to Twitter and publish information by computers, smartphones, and other terminal devices. There are two striking features of short texts such as tweets. First, they are short, together with a word limitation on tweets [5]. The short content of the tweet means it contains relatively narrow information, which is not conducive to further analysis. Second, short texts such as tweets and comment content contain a wealth of emojis [6]. Emojis have been used frequently on social media, and they have been endowed with rich connotations in the process of use. In addition to basic functions such as expressing actions (e.g., “Go skiing today! 🎿”), objects (e.g., “Sushi 🍣 for lunch.”), weather (e.g., “It’s sunny ☀️ in Montreal this morning.”) or emotions (e.g., “Tonight is a great night! 😊”), emojis can also enhance the emotion and even disambiguate short texts such as tweets with sarcastic words and phrases (e.g., “I love to work overtime! 😞”).

The additional content, such as pictures, videos, and links, has little influence on the emotional inclination of the tweet itself, which is a kind of noise and can be eliminated from study. Tweet short text, as one of the most important elements of Twitter, determines the emotional orientation of tweets in most cases [7]. However, due to the limited number of words, short text sometimes cannot fully express users' emotions and attitudes. Therefore, users add emoticons such as emojis to enrich their emotional leanings. Internet emoticons were born in the 1980s, and the original emoticons were made up of characters [8]. With the advancement of the Internet, emoticons have undergone great changes in form, content, and function. Emojis have moderately become the most popular emoticons on social networks and have become an indispensable chatting tool in today's network communication [9]. Symbolic communication can convey feelings more accurately and change people's communication mode and expression habits. At the same time, emoticons have different extended meanings in different situations [10], which can convey rich semantic information beyond the reach of text expression. Therefore, the importance of emoticons is self-evident. To objectively judge the emotional polarity of short texts such as tweets, it is necessary to study emoticons such as emojis in addition to analyzing short text. By integrating emoticons into the process of short-text sentiment analysis, it can more accurately judge the emotional tendency of short texts in social media such as Twitter and TikTok.

Social networks are not only a medium for people to record their lives and communicate with each other but also a way to express personal feelings and maintain relationships [11]. Therefore, social media such as Twitter is an important carrier for people to express their happiness and sorrow. As an extremely influential news and public opinion platform, short texts from social networks generate huge emotional information from a great number of users, which seems to be chaotic but contains a considerable value. These emotional traits reflect users' interests and preferences and, at the same time, may also have a huge impact on the spread of online public opinion [12]. Therefore, sentiment analysis of short texts can understand users' preferences and their views on some hot events in real society and make trend predictions, providing the scientific basis for government decision making. At the same time, tweets and other short-text data from social media contain a vast majority of users' comments and suggestions on products, services, environment, etc. [13]. Enterprises and institutions can further mine and analyze short-text information to obtain and provide a scientific basis for further research and research or improvement of products [14]. Through the analysis of short-text information, we can not only predict people's personality characteristics and living conditions but also forecast the development trend of new events, which has practical significance for social development.

Previous work has studied the presence of sentiment value in different short texts and attempted to analyze the relevant sentiment characteristics in these cases. Zhao et al. proposed an unsupervised word-embedding method based on large corpora, which utilizes latent contextual semantic relationships and co-occurrence statistical features of words in tweets to form effective feature sets. The feature set is integrated into a deep convolution neural network for training and predicting emotion classification labels [15]. Alharbi et al. proposed a neural network model to integrate user behavior information into a given document and evaluate the data set using a convolutional neural network to analyze sentiment value in short text such as tweets. The proposed model is superior to existing baseline models, including naïve Bayes and support vector machines for sentiment analysis [16]. Sailunaz et al. incorporated tweet responses into datasets and measurements and created a dataset with text, user, emotion, sentiment information, etc. The dataset was used to detect sentiment and emotion from tweets and their replies and measured the influence scores of users based on various user-based and tweet-based parameters. They used the latter information to generate generalized and personalized recommendations for users based on their Twitter activity [13]. Naseem et al. proposed a transformer-based sentiment analysis method, which encodes the representation from the converter and applies deep intelligent context embedding to improve the quality of tweets by removing noise and considering word emotion, polysemous, syntactic, and semantic knowledge. They also designed a

two-way long-term and short-term memory network to determine the sentiment value of tweets [17]. However, these studies focus on methods to improve the accuracy of analyzing the emotional characteristics of short texts, ignoring the effect of emoticons such as emojis on the emotional tendency of the whole text.

Several studies have been conducted to analyze the emotional features and semantic information contained in emoticons and their emotional impact on text content. Barbieri et al. collected more than 100 million tweets to form a large corpus, and distributed representations of emoticons were obtained using the skip-gram model. Qualitative analysis showed that the model can capture the semantic information of emoticons [18]. Kimura et al. proposed a method of automatically constructing an emoticon dictionary with any emotion category. The emotion words are extracted, and the co-occurrence frequency of emotion words and emoticons is calculated. According to the proportion of the occurrence times of each emoticon in the emotion category, a multi-dimensional vector is assigned to each emoticon, and the elements of the vector represent the intensity of the corresponding emotion [19]. However, these studies focused on the emotional characteristics and semantic information contained in emoticons themselves, rather than combining emoticons with textual data. Arifiyanti et al. utilized emoticons to build a model, classified the emotion categories of tweets containing emoticons, and evaluated the performance of the classification model [20]. Helen et al. proposed a method to understand emotions based on emoticons, and a classification model based on attentional LSTM was designed [21]. However, the above studies focus on understanding the emotional characteristics of the whole text by extracting and analyzing emoticons in the text and establishing relevant sentiment labels, without integrating emoticons' emotional information with the emotional features of the text content for further in-depth analysis.

In view of the increasing frequency of people, especially young people, using emoticons such as emojis in text, and the increasingly close relationship between emojis and the emotional tendency of short-text content such as tweets, this paper aims to improve the accuracy of sentiment analysis on text data, especially short-text data, and objectively judge the emotional tendency of short-text content such as tweets. Combined with the features of emojis in short texts, this paper designs and implements a sentiment classification method with the emoji vectorization algorithm based on the blending of emojis and short texts.

2. Materials and Methods

2.1. Data Source and Corpus Construction

The corpus is one of the basic resources of natural language processing [22]. Compared with traditional texts, short texts such as tweets are characterized by short-text content, rich emoticons, more noise, and unstructured language [23]. Taking the short-text data of tweets as an example, it has four distinct characteristics. First, concise language. Although each tweet is limited to 280 words, most users often use only one or two sentences to express their views and opinions, and the number of words is far less than 280 words [24]. In addition, another important reason for concise language is that users often omit sentence components [25]. Short texts such as short tweets resulting in insufficient contextual information and difficult-to-extract evaluation objects due to default sentence elements have brought challenges to sentiment analysis. Second, various forms of expression. Emoticons are widely used in short texts such as tweets. According to the collected tweet data, it is found that the number of tweets containing at least one emoticon accounts for 37.5% of the total, which is enough to show the users' love for emoticons. The reason is not only that emoticons increase the readability and sense of substitution of short text, but also because emoticons can directly and vividly convey users' attitudes and emotions [26]. Third, more noise. Short text, such as tweets, uses specific symbols to indicate a specific role. Links often appear in tweets [27]. These symbols and links do not affect the emotional orientation of tweets and are the noise of sentiment analysis. If not removed, the accuracy of sentiment classification will be affected. Fourth, new words appear frequently on the Internet [28]. With the increasing number of netizens, netizens have created many new words which

are different from traditional language forms in the process of online communication. For example, “TBH”, to be honest, and “amirite”, am I right, etc. Generally, network neologisms also have an emotional tendency. Therefore, in the field of tweet emotion analysis, we also need to analyze network neologisms.

This paper constructs a corpus of short texts containing rich emoticons. Data acquisition tools were used to collect short-text data using Twitter as the source of short texts, and 100,000 non-news tweets published from 15 October 2021 to 31 October, 2021, were selected as backup corpus. The preliminary collected tweet data contained a lot of noise and redundant data, so it was necessary to preprocess the data. First, we deleted tweets without emoticons. The research goal of this paper was to analyze the emotional features of the text by integrating emoticons with the short text. Emoticons can not only directly convey the feelings and opinions of the information publisher but also have an emotional tendency [29]. Therefore, emoticons should be considered as one of the important factors in the emotion analysis of short texts. In this paper, short-text and emoticon co-occurrence tweets are used as candidate corpus. Second, we deleted tweets with less than three words in the corpus. In general, we believe that tweets of less than three characters are not emotionally inclined or do not fully express the opinion holder’s attitude and should be removed from the corpus [30]. Third, we deleted links, usernames, topic names, and retweets from tweets in the corpus. Fourth, we performed word segmentation of Twitter text.

2.2. Annotation of Emotion in Short-Text Corpus

Short texts sent by users on social media can contain colorful emoticons. Among them, emojis are the most popular and most used emoticons. At present, emojis have been widely used in various social networks, and some of them have clear emotional tendencies. For example, 😊 has positive emotional tendencies, while 😞 has negative emotional tendencies. The site “Emojitracker” monitors emoji usage in tweets in real-time, as shown in Figure 1. The site ranks each emoji from highest to lowest in terms of the number of times they have been used in the current time. As can be seen from the figure, many of the emojis that are used frequently have no sentiment value, such as 📷. These emojis have little impact on the sentiment orientation of tweets, so this paper includes them in the neutral category.

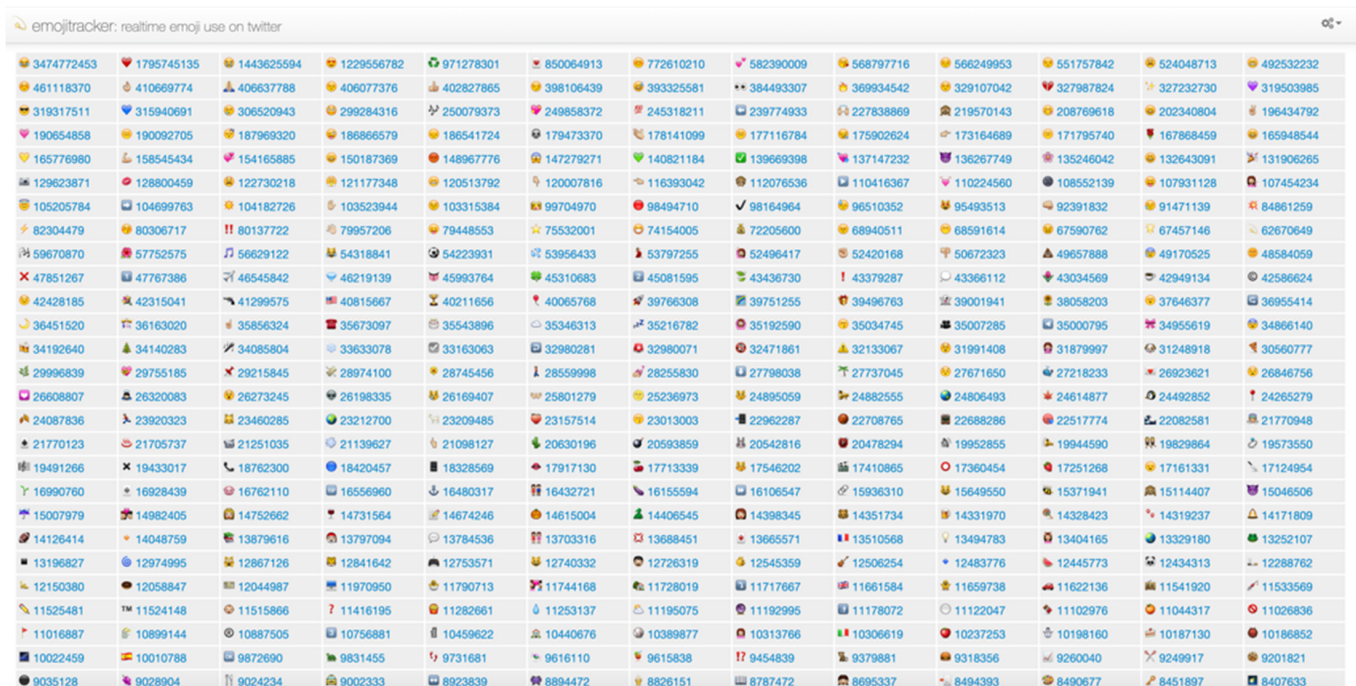




Figure 1. The site “Emojitracker” monitors emoji usage in tweets in real time.

In addition, Cappallo et al. performed statistical analyses of a large amount of data with emoticons and found that the emergence of emoticons conforms to the long-tail distribution [31]. Maximum emojis can be covered by studying only the most frequent ones. Therefore, this paper adopts artificial methods to select the top 300 emojis with clear emotional tendencies from the 819 emoticons in the “Emojitracker” website for research. After screening, 80 emojis were selected for study in this paper, including 40 positive emojis and 40 negative emojis. Emojis and their emotional tendencies are shown in Table 1.

Table 1. Emojis and their emotional tendencies.

Emotional Tendencies	Emojis
Positive	
Negative	

In this paper, two people were organized to annotate the selected short-text corpus data with emojis, and the content met the requirements by manual annotation. To ensure the accuracy of the annotation results, the candidate data could only be tweets with the same annotation results by two people. At the same time, a third person was selected to re-check the candidate data. Finally, a corpus of tweets emotion test was obtained by manual annotation, including 2000 positive and 2000 negative tweets each. All tweets in the corpus contain at least one emoji for analysis.

2.3. Emoji Vectorization Algorithm

2.3.1. Word Vector Training

Emojis are strongly correlated with the sentiment orientation of short texts such as tweets. In the sentimental analysis of tweets, taking emojis as one of the research objects can more objectively judge the sentiment value of tweets. However, one of the questions that need to be addressed is how emojis in the form of pictures should be used to co-operate with the text to improve the accuracy of emotion classification. Eisner et al. proposed the vectorization algorithm called emoji2vec [32]. By transforming emojis into vector forms, emojis can be used in all areas of language processing as well as words.

Word2vec is a tool that Google open-sourced in 2013 to turn words in text into a data format that computers can understand [33]. It learns hidden information between words unsupervised in unlabeled training sets and obtains word vectors that can preserve syntactic and semantic relationships between words. The emoji2vec emoji vector algorithm uses the word2vec tool to train the word vector. The word2vec tool was used for training in the processed corpus, and 765,285 five-dimensional word vectors were obtained after the training.

2.3.2. Construct Sample Set

Emoticons and languages interact semantically. This paper constructs a sample set of emoticons according to the actual requirements of the vectorization algorithm. The sample set constructed consists of three parts: emoji, emoji name, and CLDR short name, such as: { 😊, Cute, Smiling Face with Smiling Eyes}.

The first component of the sample is the emoji picture. The emoji name is the second component of the sample. In the naming process, the emoji name must meet two conditions.

First, the name can reflect the basic meaning of the emoji. Second, names are relatively unique [34]. The third component of the sample is CLDR. The Common Locale Data Repository (CLDR) project is a project of the Unicode Consortium to provide locale data in XML format for use in computer applications. CLDR contains locale-specific information that an operating system will typically provide to applications [35]. The Full Emoji List is a list of emojis that details the code and CLDR short name of each emoji. Sample information on emoticons can be found in this list. Through the above methods, we construct the positive samples in the sample set.

2.3.3. Algorithm Flow

The emoji vectorization algorithm maps an emoji to a point in a high-dimensional vector space such that an emoji can be transferred into an N -dimensional vector with the format of $(\text{dim}_1, \text{dim}_2, \text{dim}_3, \dots, \text{dim}_N)$. The algorithm steps are as follows.

Initialize the emoji vector x_i . Each sample contains the name of the emoji, and we take the word vector corresponding to the emoji name w_{name} as the initial vector of the emoji vector. If the emoji name is an unregistered word, the emoji vector will be randomly initialized. The emoji vector $x_i = w_{name}$. It can be seen from the sample set information that the name of an emoji is a simple description of the meaning of the emoji, so the initial emoji vector already contains the basic part of the semantic information of the emoji, which will be conducive to the formation of the emoji vector.

Construct the description the vector v_j . w_1, w_2, \dots, w_N is a set of word vector sequences, which, respectively, correspond to the word sequences in the descriptive sentences in the sample. In this paper, these word vectors are added together as description vectors of emoticons. Then, the formula to describe the vector is:

$$v_j = \sum_{k=1}^N w_k \quad (1)$$

Description vector is the sum of the corresponding word vectors of each word in a descriptive statement, which synthesizes the syntactic and semantic information of all words in a descriptive statement [36].

Establish the mathematical model. The dot product of emoji vector x_i and description vector v_j can indicate the similarity between the two vectors. The sigmoid function is used to model the similarity probability of emoji vector x_i and description vector v_j , and the formula is:

$$P(y) = h(x_i^T v_j)^y (1 - h(x_i^T v_j))^{1-y} \quad (2)$$

$$h(x) = \frac{1}{1+e^{-x}}$$

Calculate the emoji vector x_i . The sample dataset $D = \{(v_j, y_{ij}) | v_j \in R_n, y_{ij} \in \{0, 1\}\}$ consists of every description vector v_j . When the description sentences j match with the emoji i , then $y_{ij} = 1$. Otherwise, $y_{ij} = 0$.

For all description vectors v_j in the sample dataset D , the logarithmic loss function of Equation (2) is calculated, which is:

$$-\sum_{i,j} y_{ij} \log h(x_i^T v_j) - \sum_{i,j} (1 - y_{ij}) \log(1 - h(x_i^T v_j)) \quad (3)$$

The batch gradient descent algorithm is used to find the best emoji vector x_i . The emoji vector obtained in this paper is a five-dimensional vector, and each emoji in the sample set has a corresponding emoji vector. Table 2 shows the vector of four emojis.

Table 2. Vectors of four emojis.

Emoji	Emoji Vector
😄	1.253765409217565 −0.926587569876967 1.698378103032509 −0.527834892346427 0.847561023874692
😺	1.157409314569509 −0.375025790846135 0.746948348694133 −1.047891534675927 0.287905347982658
❤️	−2.219247091347005 0.345388574098972 −1.613782945782691 0.547893128730935 −0.789132897543139
😞	−0.441708935387097 1.134897523487950 −1.824560823954166 0.078318930571228 −0.927943898316451

Visualization of five-dimensional emoji sentiment vectors in a two-dimensional space is displayed in Figure 2. Emojis include positive emojis and negative emojis.

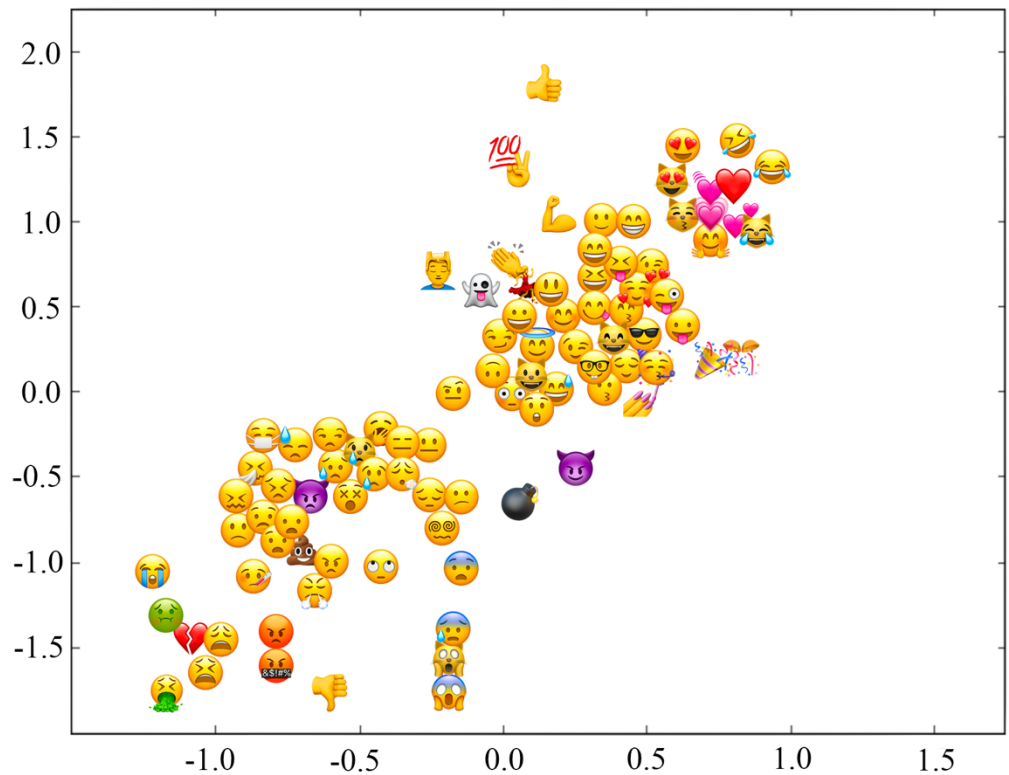


Figure 2. Visualization of five-dimensional emoji sentiment vectors in a two-dimensional space.

2.4. Naïve Bayes

Naïve Bayes is a classification method based on Bayes’ theorem, which assumes conditional independence among features. When the naïve Bayes algorithm is applied to

text classification, it assumes that the words above and below the text are independent of each other. The training set is counted and the prior probability of text category C_i is calculated:

$$P(C_i) = \frac{N_i}{N} \quad (4)$$

where N_i represents the total number of documents whose document category is C_i , and N represents the total number of all documents in the training set. Then, the conditional probability of the characteristic attributes of document d with classification is calculated:

$$P(d|C_i) = P((t_1, t_2, t_3, \dots, t_n)|C_i) = \prod_{j=1}^n P(t_j|C_i) \quad (5)$$

where t_j represents the j features of document d . $P(t_j|C_i)$ represents the probability that feature t_j appears in text category C_i . Finally, the formula for calculating the probability of all categories of documents to be classified is as follows:

$$P(C_i|d) = \frac{P(d|C_i) \cdot P(C_i)}{P(d)} \quad (6)$$

Therefore, document d is in the category with the highest probability. Naïve Bayes is a common text classification method with stable classification efficiency, can handle multiple classification tasks, and performs well on small-scale data.

2.5. Support Vector Machine

Support vector machine (SVM) is a kind of classifier whose core idea is to determine an optimal hyperplane that can correctly divide samples into two classes by maximizing the interval of the nearest samples in different classes of samples in the training set [37].

Given the i training sample in a sample set $(x^{(i)}, y^{(i)})$, where x represents the eigen-vector, $y = \{-1, 1\}$ represents the class tag. When the linear is separable, the hyperplane can be expressed as:

$$w^T x + b = 0 \quad (7)$$

Hence, for any sample set $(x^{(i)}, y^{(i)})$, when $y^{(i)} = 1$, $w^T x + b > 0$, and when $y^{(i)} = -1$, $w^T x + b < 0$. We define:

$$\begin{cases} w^T x + b \geq 0, y^{(i)} = 1 \\ w^T x + b \leq 0, y^{(i)} = -1 \end{cases} \quad (8)$$

The sum of the distances between the two support vectors belonging to different categories and the hyperplane is:

$$\gamma = \frac{2}{\|w\|} \quad (9)$$

In order to determine the optimal hyperplane, it is necessary to satisfy the parameters w and b in Equation (8), such that the interval γ is the largest, namely:

$$\begin{cases} \min \frac{1}{2} \|w\|^2 \\ y^{(i)} (w^T x^{(i)} + b) \geq 1, i = 1, 2, 3, \dots, n \end{cases} \quad (10)$$

In this paper, the support vector machine (SVM) algorithm, which is widely used in the classification field, is selected as the classification algorithm to construct an SVM emotion classifier. It is crucial to select suitable features to get a better SVM emotion classifier. According to the characteristics of short texts, this paper selects the following text features.

First, the frequency of emotional words. Counting the frequency of emotion words requires an emotion dictionary. The emotion dictionary used in the experiment is Word-Emotion Association Lexicon [38]. Each short text in the experimental data is traversed, and

the number of positive emotion words and negative emotion words is counted according to the emotion dictionary. Due to the different lengths of each short text, normalization is needed. The number of emotional words obtained by statistics is divided by the total number of words in the short text to obtain the frequency of emotional words.

Second, negative words and adverbs of degree. When people communicate, they habitually use negative words and adverbs of degree. Although negative words and adverbs of degree do not have emotional polarity, when they are combined with emotional phrases, they will affect the original emotional tendency of emotional words [39]. Specifically, the combination of degree adverb plus emotional words will enhance or weaken the original emotional tendency of emotional words, such as “really fancy”. The combination of negative words plus emotion words will reverse the polarity of emotion words, such as “not into”. The combination of degree adverb plus negative word plus emotion word will make the emotion degree and polarity of emotion word change, such as “really not into”. Therefore, special attention should be paid to negative words and adverbs of degree when analyzing the emotional tendency of short texts.

Third, the number of exclamation marks and question marks. In short-text content, there are often multiple question marks or multiple exclamation marks used together. The combination of punctuation marks indicates the strengthening of the original emotional tendency. For example, the combination of multiple exclamation marks indicates the strengthening of surprise, anger, and other emotions. The combination of multiple question marks indicates the strengthening of doubts and puzzles. Therefore, counting the number of exclamation marks and question marks in short-text content is helpful to analyze the emotional tendency of short texts.

Fourth, the number of emoticons such as emojis. Emoticons have their own emotional tendency, which affects the whole process of short text to a certain extent. The emotional intensity of the body is even the emotional tendency. Therefore, the number of emoticons is one of the characteristics of this paper.

2.6. Convolutional Neural Network

Deep learning has been widely used in image recognition, speech recognition, computer vision, and other fields since it was proposed and has made remarkable achievements [40]. Compared with traditional machine learning algorithms, deep learning has advantages in feature expression and model building [41]. Therefore, we use the convolutional neural network (CNN) to analyze short-text emotion. To give full play to the role of emoticons in promoting short-text emotional tendency, this paper adds emoticons vector to the short-text emotional analysis based on the CNN classification model.

The convolution layer and pooling layer play an important role in the convolution neural network. The convolution layer can extract local features and semantic combinations from input data. The pooling layer selects local features and semantic combinations based on the convolution layer and then filters out unimportant local features and semantic combinations with low confidence [42]. The alternating superposition of multiple convolution layers and pooling layers can extract highly abstract features from text data and improve the accuracy of emotion classification. Figure 3 is the structure diagram of the classification model based on the convolutional neural network adopted in this paper.

The model has the following four layers:

Input layer. The input of the classification model is a matrix. The matrix is formed by connecting the word vectors corresponding to all words in the sentence after word segmentation. If the word vector corresponding to the i th word in the input sentence with length n is $X_i \in R^5$, then the matrix is $X = X_1 \oplus X_2 \oplus \dots \oplus X_n$, and \oplus is the connector.

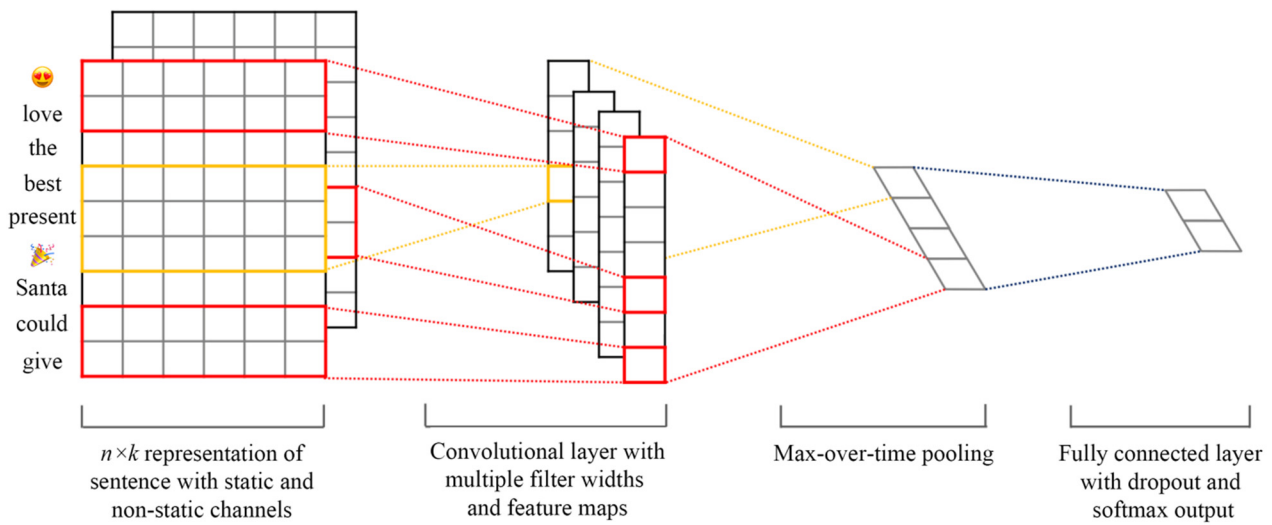


Figure 3. Structure diagram of the classification model.

Convolutional layer. The classification model based on the convolution neural network uses convolution filters with different window h lengths to extract the local features of the input layer. In the research, we implement the parallel convolution layer with multiple convolution kernels of different sizes to learn short-text features. Multiple convolution kernels are defined to acquire features in the short-text content and reduce the degree of fortuity in the feature extraction process. We define the filter size of convolution kernels as h_1Xk, h_2Xk, h_3Xk , where k is an integer and the dimension of word embeddings, and h_1 is the stride value. The feature obtained by using the convolution filter w as the input layer is:

$$c_i = f(w \times X_{i:i+h-1} + b) \tag{11}$$

where f is the non-linear activation function. The rectified linear unit (ReLU) is used in the research:

$$c_i = \max(0, w \times X_{i:i+h-1} + b) \tag{12}$$

In the equation, b is the bias term, $w \in R^{hk}$ is the shared weight, $X_{i:i+h-1}$ represents the connection of the word embedding which is from the i word of short text X to the $i + h - 1$ word ordered from top to bottom, and $b \in R$ is an offset term. A characteristic graph can be obtained by applying the convolution filter to all adjacent word vectors with length h in the input matrix $C = [c_1, c_2, \dots, c_{n-h+1}]$, $c \in R^{n-h+1}$. Therefore, $n - h + 1$ feature maps are used for each convolution kernel to obtain a feature vector t whose dimension is $X(n - h + 1)$.

If the number of convolution kernels is p , then p feature vectors can be obtained through feature mapping, and $T = [t_1, t_1, \dots, t_p]$. If q parallel convolution kernels of different types are used; for example, $h_1Xk, h_2Xk, \dots, h_qXk$, and the number of each type of convolution kernel is p , then pXq feature vectors that can be obtained after feature mapping, and $S = [t_1, t_1, \dots, t_p]$. Therefore, S is the output from the convolution layer. The output will be sent to the pooling layer of the CNN.

Pooling layer. The role of the pooling layer is to screen out the optimal local features. The pooling layer performs a max-pooling operation on all the characteristic graphs obtained by the convolution layer, which is $\hat{C} = \max\{C\}$.

Fully connected layer. The output of the pooling layer is connected to the output node of the last layer by full connection, and the tweet emotion is classified by the SoftMax classifier. In the final implementation, the dropout technique is used on the fully connected layer to prevent the hidden layer neurons from self-adapting and to reduce overfitting. The output layer is a fully connected SoftMax layer with the dropout technique. The output

layer outputs the classification accuracy and loss of the method. The proportion of dropout starts from 0.5 and gradually decreases until the model performs best, which is 0.2. The number of training epochs is 8. The optimizer used for training is AdamOptimizer.

2.7. Recurrent Neural Network

The recurrent neural network (RNN) is one of the artificial neural networks. It is a neural network that can model sequence data and process sequence data of any length. The difference between it and the convolutional neural network (CNN) is that the cyclic neural network can consider the sequence characteristics in the text [43]. In traditional neural networks, nodes in the same layer are not connected. Traditional neural networks assume that elements do not affect each other. However, this assumption does not accord with the realistic logic. Therefore, this network structure will appear powerless in dealing with many problems. The nodes between the hidden layers of the recurrent neural network are connected. The input of the current time and the output of the hidden layer of the previous time jointly determine the output of the current time. In other words, the recurrent neural network can remember the previous information. This sequence characteristic is that when RNN processes the current input information, it will calculate the current text together with the previously memorized information.

The RNN includes the input layer, the hidden layer, and the output layer, as displayed in Figure 4 [44]. It can be clearly seen from the figure that the nodes of the hidden layer can not only be self-connected but also interconnected.

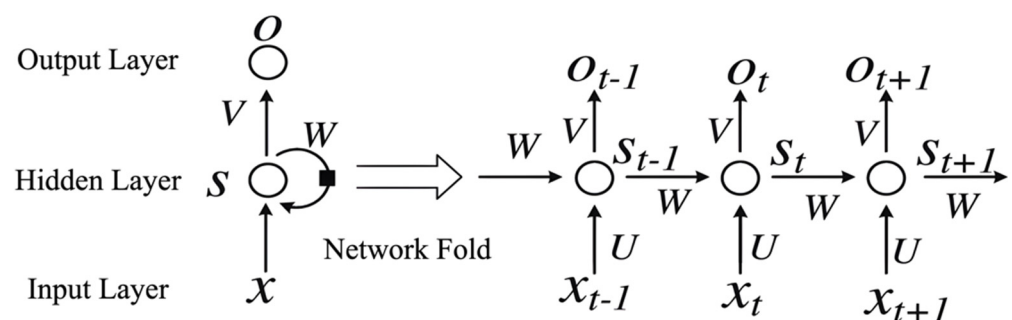


Figure 4. Structure of RNN.

In the RNN, the hyperbolic tangent activation function \tanh is used to determine the output of the current network unit and transfer the current network unit state to the next network unit:

$$\tanh(x) = 2\sigma(2x) - 1 \quad (13)$$

$$f(x) = \frac{1 - e^{-2x}}{1 + e^{-2x}} \quad (14)$$

2.8. Long Short-Term Memory

The backpropagation through time (BPTT) algorithm is used in the training of the RNN. In the training process of this algorithm, there will be the problems of gradient explosion and gradient disappearance, which makes the RNN unable to deal with long sequences. The long short-term memory (LSTM) network is an improvement of cyclic neural networks and solves the above problems [45]. LSTM solves the problem of gradient explosion encountered in the RNN, and it is accurate to extract text long-distance dependent semantic features when processing text information [46]. LSTM solves the problem of text long-distance dependence through the gating system in the network unit. LSTM contains

three such gating systems, namely input gate, output gate, and forgetting gate. These gating systems are realized by the sigmoid function:

$$\sigma(x) : f(x) = \frac{1}{1 + e^{-x}} \quad (15)$$

The output of the sigmoid function is a value between 0 and 1. The closer the value is to 1, the more information the door opens and retains. On the contrary, the closer it is to 0, the more information it needs to forget. The single-cell unit of LSTM is composed of three sigmoid functions, two tanh activation functions and a series of operations.

The structure of the diagram LSTM cell is displayed in Figure 5 [47] below.

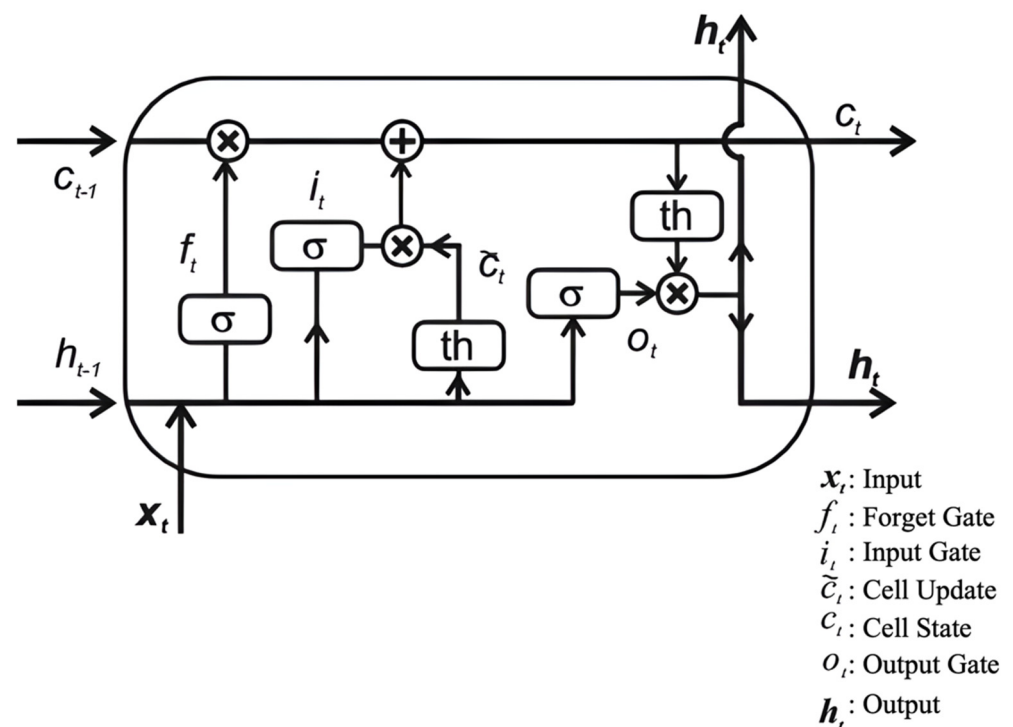


Figure 5. Structure diagram of LSTM cell.

In the LSTM neural network, the first step is to process the current input information and the information transmitted from the previous state through the forgetting gate to determine which information will be lost from the cell state. The second step is to use the input gate to control the input of useful information into the current state and obtain the latest state of the current state through the tanh activation function. The output determines the state of the output and passes it to the next door.

3. Results

We introduced the convolutional neural network in deep learning to extract hidden sentiment features from short-text data and find the best analysis method for the fusion of short-text content and emoticons by comparing different classification methods.

Based on the convolutional neural network classification model, this paper analyzes three classification models. The first classification model only considers short texts in the tweet corpus, removes emoticons, divides each tweet in the experimental data into words, and takes the sentence matrix connected by the corresponding word vector of all words as the input of the convolutional neural network classification model to classify short texts. The second classification model first converts emoticons in the short-text corpus into named texts corresponding to emoticons. For example, it would convert the tweet: “Getting everybody together for the start of the Christmas tour! 😊” into “Getting

everybody together for the start of the Christmas tour! Smiling face with sunglasses". Then, the transformed tweets are segmented into sentence matrices, which are trained and tested by the classification model of the convolutional neural network. The third classification model is to transform emoticons into emoticons vectors using the emoji vectorization algorithm, and then connect the corresponding word vector and emoticons vector into the sentence matrix according to the lexical order of the tweet corpus. Finally, the sentence matrix is input into the convolutional neural network classification model for classification.

We conducted comparative experiments on the previously established corpus of tweet data and used naïve Bayes, LSTM, RNN, and SVM as the baseline methods. In the experiments, the Python programming language and the TensorFlow platform were utilized for implementation. The experimental results of the corpus with positive sentiment value and negative sentiment value are shown in Tables 3 and 4.

Table 3. Experimental results with positive-sentiment-value corpus.

Analysis Method	Identify Quantity	Accuracy
Naïve Bayes	1447	72.35%
CNN (Word2Vec)	1632	81.60%
CNN (Emoji2Word, Word2Vec)	1649	82.45%
CNN (Emoji2Vec, Word2Vec)	1704	85.15%
LSTM	1660	83.00%
RNN	1651	82.55%
SVM	1558	77.90%

Table 4. Experimental results with negative-sentiment-value corpus.

Analysis Method	Identify Quantity	Accuracy
Naïve Bayes	1351	67.55%
CNN (Word2Vec)	1473	73.65%
CNN (Emoji2Word, Word2Vec)	1385	69.25%
CNN (Emoji2Vec, Word2Vec)	1596	79.80%
LSTM	1479	73.95%
RNN	1470	73.50%
SVM	1402	70.10%

According to the comparative analysis of experimental results, it can be concluded that the model based on deep learning performs better, namely in accuracy, because the deep learning method can extract deep-seated data features in short texts such as tweets. In the short-text data set containing rich emoticons, the best experimental model is the third one, which converts emoticons and text into vectors and analyzes them. By comparing the experimental outcomes of the analysis of positive emotions and negative emotions, we can find that the accuracy of all models on the analysis of negative short-text emotions is reduced, especially the second model. When emoticons are converted into text messages and then combined with short-text content analysis, the reduction range is the largest, and the instability is the highest. This is because many short texts with negative emoticons may not express negative emotions but express things such as surprise, emotion, etc. In this case, the conversion of emoticons into words to analyze the emotional tendencies of the text will have the opposite effect. In general, converting emoticons and texts into vectors to achieve the highest accuracy in analyzing emotional value, which reflects that emoticons vectorization algorithm can play a significant role in the emotional analysis of short texts.

Novak et al. provide a mapping to positive, negative, and neutral occurrence information for 751 emojis, also available on Kaggle [48]. Based on the dataset with 70,000 tweets and 969 different emojis, we designed a contrast experiment. We first extract the tweets from the dataset that contained the emojis in Table 1. A total of 37,810 tweets were selected from the dataset. Then, we proceeded with the accuracy analysis experiment with different methods based on the corpus with extracted data.

The evaluation indices of the experiment included accuracy (P), recall (R), positive and negative class F1 values. For the overall performance, the overall correction rate accuracy is used, and the calculation formula is:

$$Accuracy = \frac{TP + TN}{TP + FP + TN + FN} \quad (16)$$

In this formula, TP , FP , TN , and FN represent the correctly classified positive tweets, the misclassified positive tweets, the correctly classified negative tweets, and the misclassified negative tweets, respectively.

The experiment compared the proposed method CNN (Emoji2Vec, Word2Vec) with traditional classification methods, including naïve Bayes, CNN (Word2Vec), CNN (Emoji2Word, Word2Vec), LSTM, RNN, and SVM. The parameter setup is as follows. The parameters of c and g of SVM are obtained by grid search. The parameter setup of naïve Bayes is the default parameter setup of sklearn. RNN is MV-RNN of reference, and its parameter setup is the same. LSTM is Tree-LSTM of reference, and its parameter setup is the same. The emotional classification results are displayed in Table 5.

Table 5. Comparison of emotional classification results.

Analysis Method	Positive P	Negative P	Positive R	Negative R	Positive F1	Negative F1	Accuracy
Naïve Bayes	79.15%	82.58%	82.37%	78.66%	80.27%	79.36%	80.75%
CNN (Word2Vec)	86.38%	87.75%	87.20%	85.52%	87.05%	86.97%	86.91%
CNN (Emoji2Word, Word2Vec)	85.60%	89.95%	87.42%	85.10%	87.83%	86.65%	87.60%
CNN (Emoji2Vec, Word2Vec)	89.23%	91.60%	92.16%	88.33%	90.59%	89.70%	90.35%
LSTM	86.55%	87.10%	87.05%	88.42%	86.98%	88.15%	87.20%
RNN	84.92%	86.03%	87.13%	84.88%	86.04%	84.79%	85.33%
SVM	83.78%	85.35%	85.41%	83.88%	84.79%	84.32%	84.66%

The experimental results indicate that CNN (Emoji2Vec, Word2Vec) has a better performance with higher overall accuracy and two types of F1 values than the traditional classification methods. The naïve Bayes algorithm has the lowest overall performance indices. The experimental results show that the CNN (Emoji2Vec, Word2Vec) method is effective for the emotion classification of short texts with emoticons such as emojis.

4. Discussion

Entropy refers to the degree of the chaos of a system. A system with a low degree of chaos has low entropy, while a system with a high degree of chaos has high entropy [49]. In the absence of external interference, entropy increases automatically [50]. In information theory, entropy is the average amount of information contained in each piece of information received, which is also called information entropy [51]. In the information world, the higher the entropy, the more information can be transmitted, and the lower the entropy, the less information can be transmitted [52]. The booming development of social media such as Twitter and TikTok with the increasingly wide range of short-text communication has reduced the difficulty of information dissemination. At the same time, the extensive use of emoticons such as emojis in short texts increases the amount of information covered in

short-text content and boosts the entropy value of short-text information, which makes the prediction of short-text information content represented by sentiment characters difficult.

The topic of this paper is to improve the accuracy of identifying sentiment features of short texts with emoticons, such as emojis. Based on this research goal, we first established a corpus containing rich emoticons and short texts and identified their sentiment tendencies in an artificial way, which are used as a data source for subsequent analysis. Second, we screened the emojis commonly used in 819 social media and selected 40 emojis with positive emotions and 40 emojis with negative emotions, respectively. Third, we built an algorithm to convert emoticons into vector information and analyzed the emojis we selected. Fourth, we combined the vector information transformed by emojis with the vector information transformed by characters in the short-text content and analyzed the sentiment tendency of the short-text content by using the CNN model. The results were compared with simple text analysis and emoticon conversion, and the proposed method improved the accuracy of identifying positive and negative emotional tendencies.

Existing analysis methods of short-text emotion tendency adopt analysis methods such as combining with context. For example, Wan et al. proposed an ensemble sentiment classification system of Twitter data [53]. Although they accurately analyzed the sentimental characteristics of the content of the short text, they removed the punctuation, symbols, emoticons, and all other non-alphabet characters from the short text and hence ignored the important factor of emoticons in the emotional tendency of the short text. Some studies have analyzed the emotional value of emoticons. For example, Mohammad et al. designed an algorithm and method for sentiment analysis using texts and emoticons [54]. Matsumoto et al. developed an emotion estimation method based on emoticon image features and distributed representations of sentences [55]. However, nowadays, the majority of people use emojis to express their sentiment in short texts, and emojis dominate the use of emoticons. The above-mentioned methods only focus on emoticon symbolic expression tokens and text-based emoticons, which face difficulties in analyzing short texts with emojis. Therefore, based on the existing text vectorization algorithm and emoticon vectorization algorithm, this paper designs a sentiment classification method that integrates emoticons and characters. This method has been proved to be effective. Compared with analyzing the text content of short texts or emoticons of short texts, the method proposed in this paper has higher sentiment character recognition accuracy.

Sentiment feature analysis can help organizations and enterprises collect and analyze users' attitudes towards their products or services in public opinion and help improve them, which is an effective means to improve the efficiency of data analysis. An increasing number of systems and methods are designed and applied for this. Our method, which blends short texts with emoticons, still has room for improvement in its accuracy in identifying negative sentiment tendencies. The analysis shows that emoticons with negative emotions contain more complex emotional information than emoticons with positive emotions. Not only can it express negative emotions, but it can also express emotions such as movement, surprise, and even pleasantness. Therefore, our future research direction and improvement is to improve the efficiency and accuracy of methods to identify texts with negative emotions.

5. Conclusions

Analyzing and understanding the sentimental characteristics of short texts is conducive to obtaining the data information of the deep value of public opinion and helping organizations and companies optimize their services and products. Airline services have been improved and enhanced after implementing the ensemble sentiment classification method [53]. Fast-moving consumer goods (FMCG) brands such as P&G and hotels such as Marriott Corporation mine and analyze consumers' sentiment opinions from short texts in social media [56,57]. These implementations effectively improve the production efficiency of enterprises and boost users' satisfaction, which achieves a win-win situation. We advocate further study in this research field.

In this paper, we propose a short-text sentiment analysis method that combines emoticons such as emojis with short-text content. A great number of tweets containing emojis are processed and analyzed to obtain the sentimental characteristics of short texts such as tweets. This paper first classifies popular emoticons, converts emoticons together with characters into vector representations, and analyzes them using the convolutional neural network method. Experimental results show that the proposed method is more accurate than the existing method.

Author Contributions: Conceptualization, H.Z. and K.X.; methodology, H.Z.; software, H.Z.; validation, K.X.; formal analysis, K.X.; investigation, K.X.; resources, H.Z.; data curation, H.Z.; writing—original draft preparation, H.Z.; writing—review and editing, H.Z.; visualization, H.Z.; supervision, K.X.; project administration, K.X.; funding acquisition, K.X. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Tsou, M. Research challenges and opportunities in mapping social media and Big Data. *Cartogr. Geogr. Inf. Sci.* **2015**, *42* (Suppl. 1), 70–74. [[CrossRef](#)]
2. Gupta, H.; Jamal, M.S.; Madisetty, S. A framework for real-time spam detection in Twitter. In Proceedings of the 2018 10th International Conference on Communication Systems & Networks (COMSNETS), Bangalore, India, 3–7 January 2018; pp. 380–383.
3. Chatzakou, D.; Kourtellis, N.; Blackburn, J.; De Cristofaro, E.; Stringhini, G.; Vakali, A. Mean birds: Detecting aggression and bullying on twitter. In Proceedings of the 2017 ACM on Web Science Conference, Troy, NY, USA, 25–28 June 2017; pp. 13–22.
4. Baym, N.K. *Personal Connections in the Digital Age*; John Wiley & Sons: Hoboken, NJ, USA, 2015.
5. Wang, X.; Liu, Y.; Sun, C.J.; Wang, B.; Wang, X. Predicting polarities of tweets by composing word embeddings with long short-term memory. In Proceedings of the 53rd Annual Meeting of the Association for Computational Linguistics and the 7th International Joint Conference on Natural Language Processing, Beijing, China, 26–31 July 2015; Volume 1, pp. 1343–1353.
6. Na’aman, N.; Provenza, H.; Montoya, O. Varying linguistic purposes of emoji in (Twitter) context. In Proceedings of the ACL 2017, Student Research Workshop, Vancouver, BC, Canada, 30 July–4 August 2017; pp. 136–141.
7. Ji, X.; Chun, S.A.; Wei, Z.; Geller, J. Twitter sentiment classification for measuring public health concerns. *Soc. Netw. Anal. Min.* **2015**, *5*, 13. [[CrossRef](#)] [[PubMed](#)]
8. Venter, E. Bridging the communication gap between Generation Y and the Baby Boomer generation. *Int. J. Adolesc. Youth* **2017**, *22*, 497–507. [[CrossRef](#)]
9. Kejriwal, M.; Wang, Q.; Li, H.; Wang, L. An empirical study of emoji usage on Twitter in linguistic and national contexts. *Online Soc. Netw. Media* **2021**, *24*, 100149. [[CrossRef](#)]
10. Highfield, T.; Leaver, T. Instagrammatics and digital methods: Studying visual social media, from selfies and GIFs to memes and emoji. *Commun. Res. Pract.* **2016**, *2*, 47–62. [[CrossRef](#)]
11. Velten, J.C.; Arif, R. The influence of snapchat on interpersonal relationship development and human communication. *J. Soc. Media Soc.* **2016**, *5*, 5–43.
12. Barberá, P.; Jost, J.T.; Nagler, J.; Tucker, J.A.; Bonneau, R. Tweeting from left to right: Is online political communication more than an echo chamber? *Psychol. Sci.* **2015**, *26*, 1531–1542. [[CrossRef](#)] [[PubMed](#)]
13. Sailunaz, K.; Alhajj, R. Emotion and sentiment analysis from Twitter text. *J. Comput. Sci.* **2019**, *36*, 101003. [[CrossRef](#)]
14. Cai, L.; Zhu, Y. The challenges of data quality and data quality assessment in the big data era. *Data Sci. J.* **2015**, *14*. [[CrossRef](#)]
15. Zhao, J.; Gui, X.; Zhang, X. Deep convolution neural networks for twitter sentiment analysis. *IEEE Access* **2018**, *6*, 23253–23260.
16. Alharbi, A.S.M.; de Doncker, E. Twitter sentiment analysis with a deep neural network: An enhanced approach using user behavioral information. *Cogn. Syst. Res.* **2019**, *54*, 50–61. [[CrossRef](#)]
17. Naseem, U.; Razzak, I.; Musial, K.; Imran, M. Transformer based deep intelligent contextual embedding for twitter sentiment analysis. *Future Gener. Comput. Syst.* **2020**, *113*, 58–69. [[CrossRef](#)]
18. Barbieri, F.; Ronzano, F.; Saggion, H. What does this emoji mean? A vector space skip-gram model for twitter emojis. In Proceedings of the Tenth International Conference on Language Resources and Evaluation (LREC 2016), Portorož, Slovenia, 23–28 May 2016; Calzolari, N., Choukri, K., Declerck, T., Eds.; European Language Resources Association (ELRA): Paris, France, 2016; pp. 3967–3972.

19. Kimura, M.; Katsurai, M. Automatic construction of an emoji sentiment lexicon. In Proceedings of the 2017 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining, Sydney, Australia, 31 July–3 August 2017; pp. 1033–1036.
20. Arifiyanti, A.A.; Wahyuni, E.D. Emoji and emoticon in tweet sentiment classification. In Proceedings of the 2020 6th Information Technology International Seminar (ITIS), Surabaya, Indonesia, 14–16 October 2020; pp. 145–150.
21. Helen, A.; Suryani, M.; Fakhri, H. Emotional context detection on conversation text with deep learning method using long short-term memory and attention networks. In Proceedings of the 2021 9th International Conference on Information and Communication Technology (ICoICT), Yogyakarta, Indonesia, 4–5 August 2021; pp. 674–678.
22. Zeroual, I.; Lakhouaja, A. Data science in light of natural language processing: An overview. *Procedia Comput. Sci.* **2018**, *127*, 82–91. [[CrossRef](#)]
23. Afyouni, I.; Al Aghbari, Z.; Razack, R.A. Multi-feature, multi-modal, and multi-source social event detection: A comprehensive survey. *Inf. Fusion* **2022**, *79*, 279–308. [[CrossRef](#)]
24. Piao, Z.; Park, S.M.; On, B.W.; Choi, G.S.; Park, M.S. Product reputation mining: Bring informative review summaries to producers and consumers. *Comput. Sci. Inf. Syst.* **2019**, *16*, 359–380. [[CrossRef](#)]
25. Liang, J.; Tsou, C.H.; Poddar, A. A novel system for extractive clinical note summarization using EHR data. In Proceedings of the 2nd Clinical Natural Language Processing Workshop, Minneapolis, MN, USA, 7 June 2019; pp. 46–54.
26. Zhang, Y.; Shao, B.J. Influence of service-entry waiting on customer's first impression and satisfaction: The moderating role of opening remark and perceived in-service waiting. *J. Serv. Theory Pract.* **2019**, *29*, 565–591. [[CrossRef](#)]
27. Samuel, J.; Ali, G.G.; Rahman, M.; Esawi, E.; Samuel, Y. COVID-19 public sentiment insights and machine learning for tweets classification. *Information* **2020**, *11*, 314. [[CrossRef](#)]
28. Zhang, Z.; Robinson, D.; Tepper, J. Detecting hate speech on twitter using a convolution-gru based deep neural network. In *European Semantic Web Conference*; Springer: Cham, Switzerland, 2018; pp. 745–760.
29. Kim, Y.; Jun, J.W. Factors affecting sustainable purchase intentions of SNS emojis: Modeling the impact of self-presentation. *Sustainability* **2020**, *12*, 8361. [[CrossRef](#)]
30. Shah, P.V.; Swaminarayan, P. Sentiment analysis—An evaluation of the sentiment of the people: A survey. In *Data Science and Intelligent Applications*; Springer: Singapore, 2021; pp. 53–61.
31. Cappallo, S.; Svetlichnaya, S.; Garrigues, P.; Mensink, T.; Snoek, C.G. New modality: Emoji challenges in prediction, anticipation, and retrieval. *IEEE Trans. Multimed.* **2018**, *21*, 402–415. [[CrossRef](#)]
32. Eisner, B.; Rocktäschel, T.; Augenstein, I.; Bošnjak, M.; Riedel, S. emoji2vec: Learning emoji representations from their description. *arXiv* **2016**, arXiv:1609.08359.
33. Deghani, M.; Johnson, K.M.; Garten, J.; Boghrati, R.; Hoover, J.; Balasubramanian, V.; Parmar, N.J. TACIT: An open-source text analysis, crawling, and interpretation tool. *Behav. Res. Methods* **2017**, *49*, 538–547. [[CrossRef](#)] [[PubMed](#)]
34. Jaeger, S.R.; Roigard, C.M.; Jin, D.; Vidal, L.; Ares, G. Valence, arousal and sentiment meanings of 33 facial emoji: Insights for the use of emoji in consumer research. *Food Res. Int.* **2019**, *119*, 895–907. [[CrossRef](#)] [[PubMed](#)]
35. Wright, S.E. The creation and application of language industry standards. *Perspect. Localization* **2006**, 241–278. [[CrossRef](#)]
36. Anderson, A.J.; Kiela, D.; Binder, J.R.; Fernandez, L.; Humphries, C.J.; Conant, L.L.; Lalor, E.C. Deep artificial neural networks reveal a distributed cortical network encoding propositional sentence-level meaning. *J. Neurosci.* **2021**, *41*, 4100–4119. [[CrossRef](#)] [[PubMed](#)]
37. Zheng, Q.; Tian, X.; Yang, M. The email author identification system based on support vector machine (SVM) and analytic hierarchy process (AHP). *IAENG Int. J. Comput. Sci.* **2019**, *46*, 178–191.
38. Kwok, S.; Wai, H.; Sai, K.V.; Guanjin, W. Tweet topics and sentiments relating to COVID-19 vaccination among Australian Twitter users: Machine learning analysis. *J. Med. Internet Res.* **2021**, *23*, e26953. [[CrossRef](#)] [[PubMed](#)]
39. Chen, S.; Lv, X.; Gou, J. Personalized recommendation model: An online comment sentiment-based analysis. *Int. J. Comput. Commun. Control* **2020**, *15*. [[CrossRef](#)]
40. Liu, W.; Wang, Z.; Liu, X.; Zeng, N.; Liu, Y.; Alsaadi, F.E. A survey of deep neural network architectures and their applications. *Neurocomputing* **2017**, *234*, 11–26. [[CrossRef](#)]
41. Li, C.; Ding, Z.; Zhao, D.; Yi, J.; Zhang, G. Building energy consumption prediction: An extreme deep learning approach. *Energies* **2017**, *10*, 1525. [[CrossRef](#)]
42. Hou, Z.; Ma, K.; Wang, Y.; Yu, J.; Ji, K.; Chen, Z.; Abraham, A. Attention-based learning of self-media data for marketing intention detection. *Eng. Appl. Artif. Intell.* **2021**, *98*, 104118. [[CrossRef](#)]
43. Liang, G.; Hong, H.; Xie, W. Combining convolutional neural network with recursive neural network for blood cell image classification. *IEEE Access* **2018**, *6*, 36188–36197. [[CrossRef](#)]
44. Gao, M.; Shi, G.; Li, S. Online prediction of ship behavior with automatic identification system sensor data using bidirectional long short-term memory recurrent neural network. *Sensors* **2018**, *18*, 4211. [[CrossRef](#)] [[PubMed](#)]
45. Ribeiro, A.H.; Tiels, K.; Aguirre, L.A.; Schön, T. Beyond exploding and vanishing gradients: Analysing RNN training using attractors and smoothness. *Int. Conf. Artif. Intell. Stat.* **2020**, *108*, 2370–2380.
46. Zhang, Y.; Zheng, J.; Jiang, Y.; Huang, G.; Chen, R. A text sentiment classification modeling method based on coordinated CNN-LSTM-attention model. *Chin. J. Electron.* **2019**, *28*, 120–126. [[CrossRef](#)]

47. Hrnjica, B.; Bonacci, O. Lake level prediction using feed forward and recurrent neural networks. *Water Resour. Manag.* **2019**, *33*, 2471–2484. [[CrossRef](#)]
48. Kralj Novak, P.; Smailović, J.; Sluban, B.; Mozetič, I. Sentiment of emojis. *PLoS ONE* **2015**, *10*, e0144296. [[CrossRef](#)] [[PubMed](#)]
49. Neill, C.; Roushan, P.; Fang, M.; Chen, Y.; Kolodrubetz, M.; Chen, Z.; Martinis, J.M. Ergodic dynamics and thermalization in an isolated quantum system. *Nat. Phys.* **2016**, *12*, 1037–1041. [[CrossRef](#)]
50. Eskov, V.M.; Eskov, V.V.; Vochmina, Y.V.; Gorbunov, D.V.; Ilyashenko, L.K. Shannon entropy in the research on stationary regimes and the evolution of complexity. *Mosc. Univ. Phys. Bull.* **2017**, *72*, 309–317. [[CrossRef](#)]
51. Delgado-Bonal, A.; Marshak, A. Approximate entropy and sample entropy: A comprehensive tutorial. *Entropy* **2019**, *21*, 541. [[CrossRef](#)] [[PubMed](#)]
52. Osamy, W.; Salim, A.; Khedr, A.M. An information entropy based-clustering algorithm for heterogeneous wireless sensor networks. *Wirel. Netw.* **2020**, *26*, 1869–1886. [[CrossRef](#)]
53. Wan, Y.; Gao, Q. An ensemble sentiment classification system of twitter data for airline services analysis. In Proceedings of the 2015 IEEE international Conference on Data Mining Workshop (ICDMW), Atlantic, NJ, USA, 14–17 November 2015; pp. 1318–1325.
54. Ullah, M.A.; Marium, S.M.; Begum, S.A.; Dipa, N.S. An algorithm and method for sentiment analysis using the text and emoticon. *ICT Express* **2020**, *6*, 357–360. [[CrossRef](#)]
55. Fujisawa, A.; Matsumoto, K.; Yoshida, M.; Kita, K. Emotion Estimation Method Based on Emoticon Image Features and Distributed Representations of Sentences. *Appl. Sci.* **2022**, *12*, 1256. [[CrossRef](#)]
56. Yadav, M.L.; Dugar, A.; Baishya, K. Decoding Customer Opinion for Products or Brands Using Social Media Analytics: A Case Study on Indian Brand Patanjali. *Int. J. Intell. Inf. Technol. (IJIIIT)* **2022**, *18*, 1–20. [[CrossRef](#)]
57. Aydin, G.; Uray, N.; Silahtaroglu, G. How to Engage Consumers through Effective Social Media Use—Guidelines for Consumer Goods Companies from an Emerging Market. *J. Theor. Appl. Electron. Commer. Res.* **2021**, *16*, 768–790. [[CrossRef](#)]