Scholar Publishing Group
International Journal of Educational Innovation and Science
https://doi.org/10.38007/IJEIS.2022.030206
ISSN 2789-7184 Vol. 3, Issue 2: 41-52

Sentiment Analysis of English Text based on LSTM-RNN

Dang Wang*

Leshan Normal University, Sichuan, China

79108572@qq.com

*corresponding author

Keywords: English Text Sentiment Analysis, Feature Selection, Long and Short Time Memory Network, Recurrent Neural Network, Deep Learning

Abstract: With the emergence of a large number of text information on the Internet, not only can we learn the user's personal preferences from the emotional mining and analysis of the text, but also can monitor public opinion. So text sentiment analysis has become a new research hotspot for researchers on Web-based information processing. Aiming at the problem of how to analyze the sentiment class from the English text more accurately, this paper selects some user reviews dataset as the sentiment analysis database of the English text dataset by using the corpus that has already completed emotional annotation on the Internet as corpus, and uses the LSTM-RNN algorithm to classify and analyze. The analysis results of this method in this paper are compared with the traditional classification results of other models, the experimental results show that the LSTM-RNN model proposed in this paper compared with other traditional network models is more effective well in classifying sentiment analysis of English texts. The feature selection method of information gain can reach 88.2%, which proves that the model is effective.

1. Introduction

With the advent of the era of mobile Internet, information resources generated by a large number of network users as the main component of the formation of a new Internet content interaction model. Internet-based content interaction. Internet based content interaction makes information content appear geometric level growth. Most of these information contents are presented through text. It contains a lot of information, emotions and viewpoints of users. How to extract useful emotional information from these valuable texts has gradually become a hot topic in the field of research. Text affective analysis refers to the process of analysis, processing, induction and reasoning of some subjective texts with emotional color through computer technology. Text-based sentiment analysis is the task of automatically determining the author's attitude, emotion or other influential states from the text [1]. Emotional analysis and opinion mining are important for extracting useful subjective information from text documents. Vilares [2] discusses the comparison of three technologies on Twitter and the problem of multilingual polarity classification. On this

Copyright: © 2022 by the authors. This is an Open Access article distributed under the Creative Commons Attribution License (CC BY 4.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited (https://creativecommons.org/licenses/by/4.0/).

basis, the monolingual, synthetic multilingual and code-switching corpuses of English and Spanish on Twitter are evaluated. Litvinova [3] uses the numerical value of linguistic parameters as a feature to design a mathematical model for English texts, mainly using texts which are not obviously dependent on content, and categorized them as texts of suicide or non-suicide individuals.

Deep learning provides an algorithm tool for dealing with natural language tasks, especially in affective analysis tasks. Many researchers have begun to use deep learning methods to analyze text. In recent years, the main research direction of emotion analysis based on deep learning is distributed representation based on word vector and deep learning model in two aspects. Many of the deep neural network models deal with natural language processing, especially text sentiment classification tasks, which are more popular in convolution neural networks and recurrent neural network. Sequence prediction and classification is a ubiquitous and challenging problem in machine learning. It is necessary to identify the complex dependence between time distant inputs. Recurrent neural network (RNN) theoretically has the ability to deal with these time dependent abilities through its recursive (feedback) connection short-term memory [4]. The integration of sentiment analysis and deep learning technology is because the deep learning model has effective automatic learning ability [5]. Giatsoglou [6] uses machine learning to represent text documents with vectors for sentiment detection of text fragments and to train polarity classification models. Beseiso[7] uses a new deep learning structure based on character level convolution neural network (CNN) module and word level recurrent neural network (RNN) module. A hybrid learning structure based on character level analysis and word level analysis is proposed to get the latest learning results. Fern ndez-Gavilanes[8] uses a text categorization method based on unsupervised dependency analysis, which utilizes various techniques to predict emotional features in online text messages, the emotional features are mainly derived from prompt words.

Long and short time memory network is a recursive neural network architecture, which is designed to solve the problem of the disappearance and explosion gradient of traditional RNNs. Researchers have begun to use the long-term memory network as a special RNN for emotional classification tasks. They have been successfully used in sequence labeling and sequence prediction tasks, such as handwriting recognition, language modeling, and voice frame speech marking [9]. Eyben [10] proposes an online speech emotion recognition method based on long-term and short-term memory recurrent neural network. Zhang[11] uses the attention model to construct the semantic representation of the text and emotional symbols. Finally, the emotion classification model is trained on the basis of semantic representation, combining the attention model and the enhancement of emotion symbols. The model provides the best accuracy for emotional classification. Bhaskar [12] proposes a new method of voice and text-based voice conversation emotion classification. It improves the accuracy of voice emotion classification by selecting features and generating a single feature vector through support vector machine, WordNet and OuthWordNet. Wang [13] proposed a regional CNN-LSTM model: the VA (valence arousal (VA) space rating of regional CNN and LSTM prediction texts. Unlike traditional CNN, the proposed model uses a single sentence as a region and inputs text into multiple regions to extract and weigh useful emotional information in each region according to the contribution of each region to the prediction of VA.

In view of the characteristics of sentiment analysis in English text, this paper preprocesses the English text firstly, and selects the features of the English text by means of the words in the English text. The LSTM recursive neural network with deep learning is used for text representation learning. According to the syntactic structure of the text, RNN is used to recursively combine to learn the structured information of the text. This paper proposes an emotion analysis model based on LSTM recursive network and RNN network to analyze the sentiment of the English text.

2. Method

2.1. Text Emotional Analysis

As the basis of the field, text sentiment analysis is a process of analyzing, processing, summarizing and calculating subjective texts with emotional color. This process is also called text based tendentious analysis and opinion mining. Its basic task is to use computer technology to analyze and process natural language texts to identify users 'opinions and standpoints on things or people. Finally, these opinions and positions can be classified into subjective and objective evaluation, praise and derogation evaluation, or emotional intensity rating of the text. The emotional analysis process of text is divided into two parts generally: training process and classification process, as shown in Figure 1. In the training process, the labeled text set to be trained is simply pre-processed before feature selection. Different training algorithms are used to classify and get the classification model through these feature sets. In the process of classification, the input text set is pre-processed and feature selection is carried out, and then the output results are obtained through the established classification model according to these feature representations.

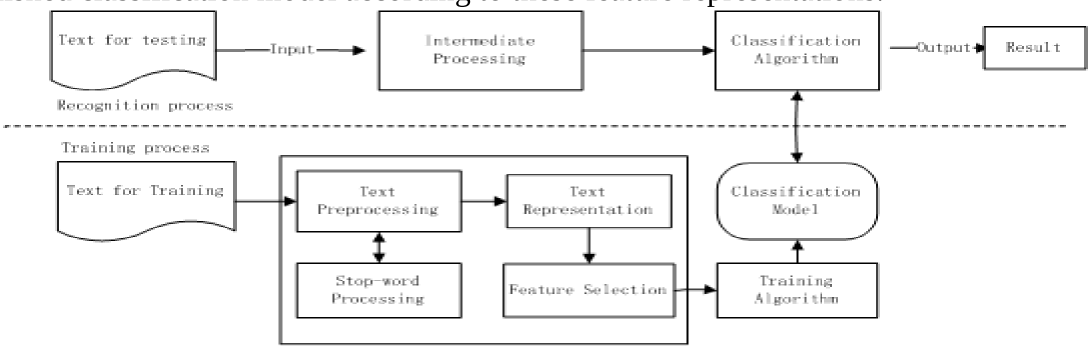

Figure 1. General process of text emotion analysis

With the size of text granularity as the reference value, the text sentiment analysis can be divided into four levels: word-level sentiment analysis, phrase-level sentiment analysis, sentence-level sentiment analysis and text-level sentiment analysis. Lexical Emotional analysis is not only the basis of text emotional analysis and judgment, but also the prerequisite of emotional analysis at phrase, sentence and text levels. Phrase-level affective analysis refers to the identification of emotional phrases based on emotional words and the determination of emotional polarity and intensity. Sentence-level affective analysis is the key to text affective analysis, which mainly identifies the holders of opinions and evaluates the text tendency of the object in sentence judgment.

2.2. Preprocessing of English Texts

In order to reduce the computational complexity of the system and reduce the dimension of feature selection attributes, we need to preprocess the text before entering the model to improve the efficiency of the analysis. The results of text pre-processing will also directly affect the final analysis and classification results. The purpose of text pre-processing is to remove some noise data in the text and get only the data related to the final analysis and classification. The preprocessing of text mainly includes the removal of irregular characters, the removal of stop words, and word segmentation.

Remove irregular characters. Processing some text data or characters that are not related to text content in English text data to eliminate irrelevant information and make the text data more standardized.

Stop word processing. In order to save storage space and improve query speed, some words or words in English sentences need to be automatically shielded to filter out some frequently used but not practical words, thereby reducing the noise impact caused by stop words. There are articles, prepositions, interjection, and numeral and so on.

Word segmentation. Text segmentation in English is mainly based on the space between words and words in English text generally, and the punctuation symbol is recognized separately by punctuation data set and regular expression. The majority of English text can be segmenting by space, but it is necessary to use predefined regular expressions for word segmentation when the common English abbreviations of "is' t" and "don" t.

2.3. Feature Selection in English Text

The process of feature sampling of data samples through a series of methods is called feature selection of text. The main purpose of feature selection of text is to filter out all feature dimensions of samples, discarding features that are useless for analysis and classification tasks and retaining features that are useful for classification. The feature items of the text have many representations. In order to reduce the feature dimension and reduce the amount of calculation, the feature selection of the text needs to select the feature item that best characterizes the text, so as to reduce the interference of the unrelated features and finally improve the accuracy of the data analysis classification. .

The feature selection for English text mainly lies in the method of scoring and sorting the feature items in the original feature set of English text, and finally obtaining the feature item subset that best represents the text category. Commonly used feature selection methods include feature selection based on document frequency, feature selection based on information gain method, cross entropy method, mutual information method, word strength method, χ^2 statistic method and so on.

(1) Feature extraction based on document frequency:

Feature extraction based on the frequency at which a feature appears in all documents is referred to as feature extraction based on document frequency. This method is to count the frequency of occurrence of a document containing a specified feature in all training documents, and then set a minimum threshold and a maximum threshold according to the statistical frequency value. When the statistical frequency of the selected feature item is lower than the minimum threshold, it indicates that the feature item is not representative in the document, and when the selected feature item is greater than the maximum threshold, the frequency of occurrence in the document is too high, and the feature item is There is no obvious distinction, and both cases need to be rejected. The feature extraction method based on document frequency has low computational complexity and can effectively filter out some noise feature items. However, if some feature items containing valid information are too low in frequency, the feature extraction method based on the document frequency will still discard it, so its accuracy may be affected.

(2) Feature extraction based on information gain method

The feature extraction method determined by the feature item to determine the amount of information or contribution provided by the analysis classification is called feature extraction based on the information gain method. When the feature item provides a large amount of information or contribution to the analysis classification, the feature item is selected, and when the information quantity or contribution degree of the characteristic classification is small, the feature item is discarded.

The contribution of the feature item to the analysis classification is mainly based on the difference between the information entropy of the document after considering the feature item and considering the feature item:

$$\mathrm{Gaint}(t_i) = \mathrm{Entropy(S)} - \mathrm{ExpectedEntropy}(S_{ti}) = \{-\sum_{j=1}^{M} P(C_j) * logP(C_j)\} - \{P(t_i) * [-\sum_{j=1}^{M} P(C_j|t_i) * log\, P(C_j|t_i)] + P(\overline{t_i}) * [-\sum_{j=1}^{M} P(C_j|\overline{t_i}) * log\, P(C_j|\overline{t_i})]\} \quad (1)$$

$P(C_j)$ indicates the probability that the document class C_j appears in the corpus, $P(t_i)$ indicates the probability that the document contains the feature item t_i in the corpus, and $P(C_j|t_i)$ indicates the conditional probability when the document belongs to the class C_j and contains the feature item t_i. $P(\overline{t_i})$ indicates the probability that the document does not contain the feature item t_i in the corpus, $P(C_j|\overline{t_i})$ indicates that the document belongs to the class C_j and the document does not contain the probability of the feature item t_i, and M indicates the number of categories.

(3) Feature extraction based on mutual information method

The feature selection method using the mutual information method indicates that the greater the mutual information, the greater the degree of association between the feature items and the categories. The mutual information method is calculated as follows:

$$\mathrm{I}(t_i, C_j) = \log\frac{P(t_i,C_j)}{P(t_i)P(C_{j)}} = log\frac{P(C_j|t_i)}{P(t_i)} \quad (2)$$

2.4. LSTM-RNN Model

Recurrent neural network has the advantage of "memory", which is mainly used in voice sequence processing and Natural Language Processing. The network structure of RNN belongs to feed forward network. When the network structure is expanded, there is no circulation and there is definite information flow. Therefore, the gradient of parameters can be calculated by back propagation algorithm. The dimensions of input and output in RNN networks are determined according to specific tasks. The long term and short term memory network is a variant of LSTM, which is mainly referred to as the gradient disappearance of RNN. The LSTM network is formed by replacing the RNN hidden layer with LSTM cells. It has a chain structure similar to that of a cyclic neural network, but the structure of the cyclic module is different. The special structure of the LSTM can effectively avoid the problem of "long-distance (long-term) dependence." A standard LSTM network typically consists of a basic three-layer structure, the input layer, the output layer, and the hidden layer. The hidden layer is mainly composed of an input gate, a forgetting gate, an output gate and a memory unit. Controlling the flow of information into the memory unit through the state of the three gates for selective addition, forgetting or deleting. The state of the cell is updated at time t by the following formula:

Forgotten Gate: The LSTM network first solves which information in the memory unit needs to be discarded through the Forgotten Gate, and selects the duration status information and the part of the new input information that needs to be forgotten. Taking h_{t-1} and X_t as inputs, a value between 1 and 0 is output in the C_{i-1} memory unit, 1 means that all information is retained, and 0 means that all information is forgotten.

$$f_t = \sigma(W^f X_t + V^f h_{t-1} + b^f) \quad (3)$$

Input gate: Determine the information to be stored in the new memory unit. The information of the sigmoid function layer of the input gate is used to determine which information needs to be updated; then a new candidate vector can be added through the tanh function layer to update the old LSTM unit state and enter the new unit state.

$$i_t = \sigma(W^i X_t + V^i h_{t-1} + b^i) \quad (4)$$

$$C_t = f_t \otimes C_{t-1} + i_t \otimes \tanh(W^c X_t + V^c h_{t-1} + b^c) \quad (5)$$

Output gate: Multiply the state of the cell normalized by the tanh function through the output gate to obtain the cell state information to be output.

$$O_t = \sigma(W^o X_t + V^o h_{t-1} + b^o) \quad (6)$$

3. Construction of the LSTM-RNN Model

This experiment mainly includes three parts: experimental data preprocessing stage, model training stage, and model testing stage. The training phase of the model is the process of training the model by using the training data; the experimental test phase is the process of evaluating the effect of the model by using the test data and the corresponding evaluation index.

3.1. Data Set Analysis

In this paper, three data sets of word vector training data set, sentiment dictionary data set and sentiment analysis data set are used in the experiment.

(1) Word vector training data set

In order to complete the training of the word vector, this paper uses the IMDB movie comment data set and the NRC tweet data set to complete the data set of two completed emotional annotations. The IMDB Movie Review data set consists of English movie reviews, which contain 50,000 reviews from the main source of the Movie Review Network. When training the word vector, some infrequently appearing words are excluded and part of them are selected as the emotional word vector training set in the comment field. The NRC Tweet Data Set contains 140w tweets. This data set uses the emoticons in the tweet to mark the emotional category of the tweet. The positive class data is expressed by including the symbol ":)" indicating happiness, and the negative class data is expressed by the symbol ":(", etc. The emoticons in the annotated tweet are removed completely which are shown as table 1.

Table 1. Number statistics of data sets

Data Set	Positive Samples	Negative Samples	Total Samples
IMDB Movie Review Dataset	25000	25000	50000
NRC Tweet data set	700000 (Contains symbols ":)")	700000 (Contains symbols":(")	1400000

(2) Emotional dictionary data set

In order to improve the accuracy of the sentiment analysis task, the emotional word vector can be obtained through data set training, and the emotional word vector is combined with the emotional information. Semeval 2013 Task2 data set is used for training to construct an emotional dictionary data set containing emotional word vectors in this paper, which is shown as table 2.

Table 2. Text sentiment analysis task data distribution based on Semeval 2013 Task2

	Positive	neutral	Negative	Total
Training Set	3213	4016	1243	8472
Development Set	521	732	345	1598
Test Set	1567	1636	600	3803

(3) Sentiment analysis task data set

This paper introduces two standard sentiment analysis task data sets on the sentiment analysis task, one is the large movie review data set of the two-category comments obtained from the IMDB annotation, and the other is the Stanford Sentiment Treebank data set. The standard deviation of the Stanford Sentiment Treebank (SST) data set is 20, which has 5 kinds of labeled data: very good, good, neutral, bad, very bad. Therefore, the neutral comment data in the SST is removed to obtain the data set SST' for the two classification.

Table 3. Standard sentiment analysis task data set

	Stanford Sentiment Treebank Data Set	IMDB Data Set
Training Set	8544	25000
Development Set	1101	0
Test Set	1101	25000
Standard deviation	20	199

3.2. Construction of the Evaluation System

Common performance metrics used to evaluate classification performance are precision and recall. The sentiment analysis research studied in this paper belongs to the two-category problem of positive and negative, so the evaluation indicators of this paper are defined as follows:

Positive rate of precision: $PP = a/c_1$

Positive recall rate: $RP = a/d_1$

Negative face precision: $NP = b/c_2$

Negative recall rate: $RN = a/d_2$

Comprehensive accuracy rate: $F = (a + b)/(c_1 + c_2)$

a represents the number of positive class test documents that are correctly classified; b represents the number of passive class test documents that are correctly classified; c_1 represents the total number of positive class documents judged by the model; c_2 represents the total number of the negative class documents judged by the model; d_1 represents the total number of actual documents; d_2 represents the total number of documents that are actually negative..

3.3. Experimental Steps

The training data to train the LSTM-RNN model established by experiment in this paper, which store historical information by the LSTM unit, and use RNN recursive network to represent the characteristics of English text.

(1) Pre-processing stage of experimental data

The experimental data is mainly read and processed at this stage,, and the original text sentiment analysis data set is preprocessed, including 5000 labeled texts in the semeval data set used for the classification model and text used to construct the word vector, 60% of the text in the data set is used as a training set to train the parameters of the experimental model. The process is as follows:

Step1: Extract the corresponding text and labels;

Step2: Remove the stop words and punctuation marks in the text;

Step3: Perform word segmentation on the text after removing the stop word;

Step4: Take the intersection of the key value of the word vector and the result of the word segmentation, and then delete the word that exists in the sentence and the key of the word vector does not exist;

Step5: All the words are made into a dictionary after the intersection is completed, and the words are indexed according to the order in the dictionary;

Step6: Disrupt the data.

(2) Model training phase

The training data is used to train the LSTM-RNN model established by the experiment. The 10% text in the data set is used as the verification set to verify the effect of the training phase model, which is used to intuitively understand the training situation of the model and select the optimal model parameters.

(3) Model testing stage

Thirty percent of the text in the text dataset is taken as the test set to test the effect of the experimental model. The test data set is used to validate the effect of the model trained in the previous stage, and the accuracy of the model is used to evaluate the model. In order to ensure the accuracy and independence of the model evaluation, the experiment selects a plurality of different corpora for experiment, wherein the data used for training and testing are all a subset of the original data set, and the test data and the training data are independent of each other.

3.4. Optimization Loss Function

The essence of sentiment analysis for English texts is a two-class problem. The loss function of the two-class problem is realized by the cross-entropy function generally. In order to obtain a better classification effect, a threshold greater than 0.5 is selected. The threshold of this experiment is 0.6. This article adjusts the loss function:

$$L = -\sum_y \lambda(y_{true}, y_{pred}) y_{true} log y_{pred} \tag{7}$$

Unit step function is:

$$\theta(x) = \begin{cases} 1 & x > 0 \\ \frac{1}{2} & x = 0 \\ 0 & x < 1 \end{cases} \tag{8}$$

Where y_{true} is the actual output; y_{pred} is the expected value, and the following relationship exists:

$$\lambda(y_{true}, y_{pred}) = 1 - \theta(y_{true} - m)\theta(y_{pred} - m) - \theta(1 - m - y_{true})\theta(1 - m - y_{pred}) \tag{9}$$

$y_{true} = 1$, $\lambda(1, y_{pred}) = 1 - \theta(y_{true} - m)$ when a positive sample enters the model , if $(y_{pred} - m) > 0$, then $\lambda(1, y_{pred}) = 0$,the cross entropy reaches minimum 0 . If $(y_{pred} - m) < 0$, then $\lambda(1, y_{pred}) = 1$ and keep the original cross entropy.

This paper combines the characteristics of sentiment analysis of English text to preprocess English texts, and then uses the different features of English texts to select the words in English texts as feature selection units. According to the characteristics of English texts, the LSTM neural network is introduced into the RNN neural network to get the LSTM-RNN neural network model, and it is applied to sentiment analysis of English texts.

4. Experimental Results and Analysis

In this experiment, the LSTM-RNN method is applied to the sentiment analysis of English texts. In the experiment process, the word is used as the characteristic unit of the document, and the text set prepared for training is preprocessed as de-stop words; Different network models are used for comparison experiments firstly. Then, the characteristics of 1000~5500 dimensions are used to test the training sets respectively using information gain method, χ^2 statistics method and mutual information method.

4.1. Contrast Experiment of Emotion Analysis Fusion Model Based on LSTM-RNN

The sigmoid function is used as the output activation function in this paper, and the model is compiled by the binary cross entropy loss function in the dataset training, and the softmax activation function is used as the output layer in the SST' dataset training, and the multi-class cross entropy loss function is used as the loss function, which makes the error of each output neuron exactly equal to the error between its output and the labeled real data at the time of error propagation.

In this paper, the accuracy, loss and time performance analysis on the dataset are used as the evaluation criteria. The proposed model LSTM-RNN is compared with the classification experiments. The classification performance of different network models under the same conditions is compared and analyzed. The word bag model of Na ïve Bayes (BiNB), recurrent neural increment network (RNTN), convolutional neural network (CNN) and deep recurrent neural network DRNN etc., the results are shown in Table 4.

Table 4. Comparison of different classification results on different network models

network model	SST		SST’		IMDB		AUC	Time(s)
	Accuracy rate	Loss value	Accuracy rate	Loss value	Accuracy rate	Loss value		
LSTM-RNN	52.5	0.43	84.3	0.23	88.9	0.23	0.88	2264
CNN	42	0.51	84.6	0.31	88.6	0.24	0.92	4473
BiNB	42.2	0.48	84.3	0.28	87.8	0.28	0.88	2345
RNTN	45.1	0.45	85.4	0.24	87.1	0.25	0.89	3155
DRNN	43.3	0.46	84.4	0.26	86.2	0.26	0.91	2543

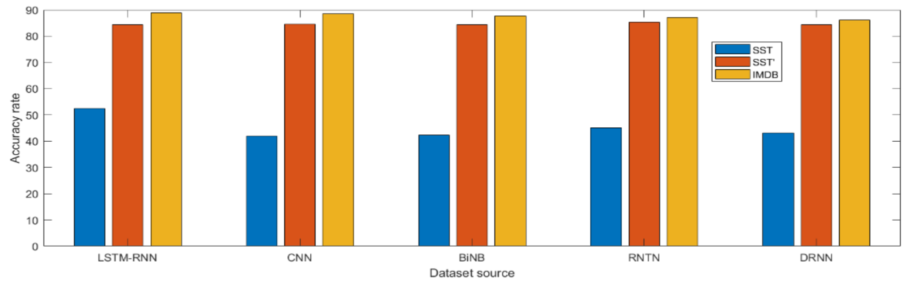

Figure 2. Classification accuracy rate of three data sets by different network structures

The three different data sets of SST, STT' or IMDB which compared with other deep learning models, the accuracy rate proposed in the two-category sentiment orientation analysis of English texts based on the LSTM-RNN model is 52.5%, 84.3% and 88.9% respectively, and the loss values are 0.43, 0.23 and 0.23 respectively. It can be seen from the accuracy and loss values that the LSTM-RNN model can achieve better classification results for the two classification of English text, which proves that the model is effective. The classification accuracy is shown in Figure 2 and the loss value is shown in Figure 3.

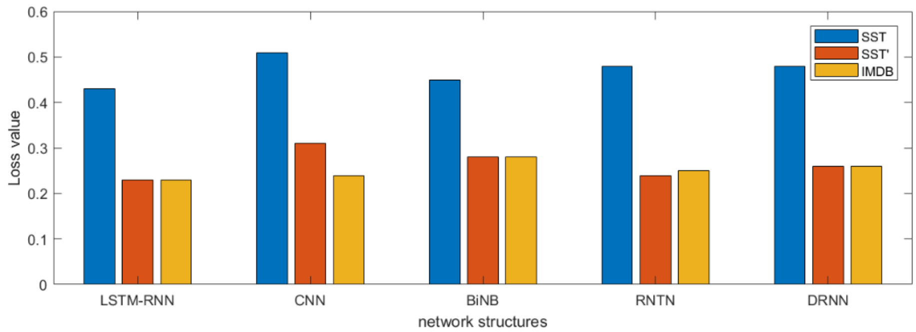

Figure 3. Classification loss rate of three data sets by different network structures

4.2. Comparison Experiment of Different Feature Dimensions

Based on the best dimension obtained from different feature selection methods in English texts of this stage, tests the sentiment classification will experiment based on the experimental data sets of English texts through different feature selection methods.

(1) When the dimension is 3000 by using the feature selection method of information gain, the positive precision rate, negative recall rate, negative precision rate, negative recall rate and comprehensive accuracy rate F tend to be balanced and stable. On the premise that F is the main indicator, the highest comprehensive accuracy rate is obtained when the dimension is 3000, and then the F value of the dimension tends to decrease.

(2) When the dimension is 3000 by using the feature selection method of χ^2 statistics method, the performance indicators are basically optimal and tend to be balanced.

(3) When using the feature selection method of mutual information, the highest comprehensive accuracy F value is 52.1% when the dimension is 3000, and the three comprehensive accuracy ratios of different dimensions are shown in Figure4.

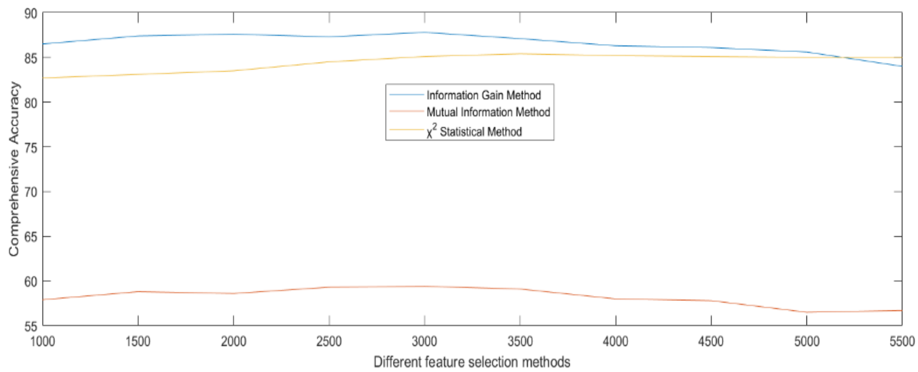

Figure 4. Accuracy comparison of different feature dimensions

It can be concluded from Figure 4 that the feature dimension has a certain influence on the classification performance of the model. According to the best feature obtained by the experiment, the dimension is selected, and then the above three feature selection methods are used respectively to test, and the different feature selection methods are obtained. The classification results are shown in Figure 5.

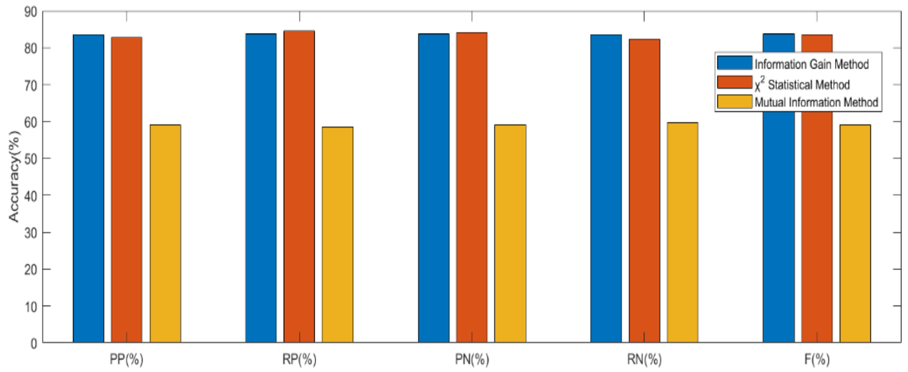

Figure 5. Classification results obtained by different feature selection methods

It can be seen from Figure 5 that the use of information gain and χ^2 statistical methods is better than the cross-information method. Compared with theχ^2 statistical method, each of the sentiment analysis using the information gain feature selection method The evaluation indicators are nearly equal, which means the method is more stable.

5. Conclusion

By introducing memory cells, the LSTM-type RNN network can learn more historical information of English text without causing a gradient explosion or attenuation of the network. At the same time, we need to further study the sentiment analysis of English text in the future. We also need to consider how to change the emotion when the time changes, and how to evaluate the emotion which is only related to the object of evaluation. In order to realize the sentiment analysis of the English text, this paper introduces the LSTM memory unit into the RNN neural network to obtain the LSTM-RNN neural network, stores the historical information of the English text through the LSTM unit, and uses the RNN recursive network to represent the characteristics of the English text. The network model can not only learn the structure information of English text but also learn the semantic information of English text. The learned English text features can better represent English sentences to achieve emotional analysis of English text. The word in this paper is used as the feature selection method of the unit and the de-stop word processing is added to the emotional processing analysis process. In this paper, the LSTM-RNN model is used to classify and analyze the emotional relationship of the existing words. In the process of experiments, different feature selection methods are used to classify the emotions, and compare the results of different traditional networks. Experiment results show that the effect is obviously better than other deep learning models, which proves the validity of the model.

At the same time, the text sentiment analysis model based on LSTM recursive network depends on the syntactic structure of training text data, which increases the workload of labeling training data. Further research in the future depends on the following work:

(1) Realize the complexity of the neural network, and complete the text analysis and classification task on the basis of reducing the complexity of the model;

(2) Using the deep learning method to automatically acquire features from the original text, reduce the learning characteristics from the annotation information, and reduce the dependence of the network model on the annotation data.

(3) How to analyze the emotional changes caused by changes over time

(4) How to evaluate some of the issues that are only slightly more granular in relation to the evaluation object in text analysis.

In the field of natural language processing, deep learning is still in its infancy, so in order to better adapt to natural language processing tasks, it is necessary to conduct in-depth research and explore deep learning methods.

Funding

This article is not supported by any foundation.

Data Availability

Data sharing is not applicable to this article as no new data were created or analysed in this study.

Conflict of Interest

The author states that this article has no conflict of interest.

References

[1] Mohammad, S. M. (2016). Sentiment analysis: Detecting valence, emotions, and other affectual states from text. In Emotion measurement (pp. 201-237). Woodhead Publishing. https://doi.org/10.1016/B978-0-08-100508-8.00009-6

[2] Vilares, D., Alonso, M. A., & Gómez-Rodríguez, C. (2017). Supervised sentiment analysis in multilingual environments. Information Processing & Management, 53(3), 595-607. https://doi.org/10.1016/j.ipm.2017.01.004

[3] Litvinova, T. A., Seredin, P. V., Litvinova, O. A., & Romanchenko, O. V. (2017). Identification of suicidal tendencies of individuals based on the quantitative analysis of their internet texts. Computación y Sistemas, 21(2), 243-252. https://doi.org/10.13053/cys-21-2-2721

[4] Koutník, J., Greff, K., Gomez, F., & Schmidhuber, J. (2014). A clockwork rnn. Computer Science, 1863-1871.

[5] Ain, Q. T., Ali, M., Riaz, A., Noureen, A., Kamran, M., Hayat, B., & Rehman, A. (2017). Sentiment analysis using deep learning techniques: a review. Int J Adv Comput Sci Appl, 8(6), 424. https://doi.org/10.14569/IJACSA.2017.080657

[6] Giatsoglou, M., Vozalis, M. G., Diamantaras, K., Vakali, A., Sarigiannidis, G., & Chatzisavvas, K. C. (2017). Sentiment analysis leveraging emotions and word embeddings. Expert Systems with Applications, 69, 214-224. https://doi.org/10.1016/j.eswa.2016.10.043

[7] Beseiso, M. (2019). Word and Character Information Aware Neural Model for Emotional Analysis. Recent Patents on Computer Science, 12(2), 142-147. https://doi.org/10.2174/2213275911666181119112645

[8] Fernández-Gavilanes, M., Álvarez-López, T., Juncal-Martínez, J., Costa-Montenegro, E., & González-Castaño, F. J. (2016). Unsupervised method for sentiment analysis in online texts. Expert Systems with Applications, 58, 57-75. https://doi.org/10.1016/j.eswa.2016.03.031

[9] Sak, Haşim, Senior, Andrew, & Beaufays, F. (2014). Long short-term memory based recurrent neural network architectures for large vocabulary speech recognition. Computer Science, 338-342. https://doi.org/10.21437/Interspeech.2014-80

[10] Eyben, F., Wöllmer, M., Graves, A., Schuller, B., Douglas-Cowie, E., & Cowie, R. (2010). On-line emotion recognition in a 3-d activation-valence-time continuum using acoustic and linguistic cues. Journal on Multimodal User Interfaces, 3(1-2), 7-19. https://doi.org/10.1007/s12193-009-0032-6

[11] Zhang, Y., Zheng, J., Huang, G., & Jiang, Y. (2018). Microblog sentiment analysis method based on a double attention model. Journal of Tsinghua University, 58(2), 122-130.

[12] Bhaskar, J., Sruthi, K., & Nedungadi, P. (2015). Hybrid approach for emotion classification of audio conversation based on text and speech mining. Procedia Computer Science, 46, 635-643. https://doi.org/10.1016/j.procs.2015.02.112

[13] Wang, J., Yu, L. C., Lai, K. R., & Zhang, X. (2016, August). Dimensional sentiment analysis using a regional CNN-LSTM model. In Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics (Volume 2: Short Papers) (pp. 225-230). https://doi.org/10.18653/v1/P16-2037