ISSN: 2414-1895 DOI: 10.6919/ICJE.202102_7(2).0020

Weibo Sentiment Analysis based on Double-Layer Attention Mechanism and Bi-LSTM

Qing Xue[1,a], Xichuan Hu[1,b]

[1]College of Information Engineering, Shanghai Maritime University, Shanghai 201306, China.

[a]xueqing-950822@163.com, [b]huxc@shmtu.edu.cn

Abstract

[Objective] Microblog emotion analysis is of great significance to public opinion monitoring and product opinions. In recent years, the depth of the learning method is to combine recurrent neural network and attention mechanism , attention mechanism is used to collect the importance of different sequence of words, but unable to get access to the importance of different sentences, ignoring the sentence on the importance of the micro blog this emotional tendency, so just single layer is difficult to obtain deep emotional attention characteristic information. Based on this, this paper proposes a double-layer attention mechanism and BI-LSTM model to improve the performance of emotion analysis. [Method/process] Microblog word-level and sentence-level features were extracted by BI-LSTM, and the importance of different words and sentences was captured by double-layer attention mechanism. Finally, the emotional classification results were obtained by full connection layer and softmax function. [Results] The experimental results show that the F1 value and accuracy of the model proposed in this paper reach 96.94% and 96.89%, respectively. Compared with other models, this model can achieve better results in the emotional analysis of this microblog.

Keywords

Attention Mechanism; Bi-LSTM; Sentiment Analysis.

1. Introduction

With the continuous rise of the network platform represented by Weibo, users can express their views on a topic and express their views with the help of weibo, instead of just receiving and browsing messages [1]. Due to the convenience of microblog sharing, rapid and real-time sharing can be completed, so a large number of users have been gained. Now, Microblog has become a popular channel for people to communicate and express their emotions. With the continuous release and dissemination of content by users, a large amount of information and data has accumulated on the microblog platform.Various opinions and opinions expressed by users can reveal the user's personal position and emotional tendency. In the face of massive data with emotional tendency, people need the method of data analysis for large-scale data, so the text emotion analysis technology arises at the historic moment. Through the analysis of these data, it is helpful for the study of language application and other issues. For example, it can obtain users' interested directions from the language, track the hot topics of the public, and determine the product opinions [2].

Text emotion analysis is a basic task of natural language processing. In recent years, researchers have paid more and more attention to it. At present, many researchers have made fruitful achievements in the field of emotion analysis on Weibo, but there are still some deficiencies. The traditional emotion analysis method mainly refers to the emotion analysis method based on emotion dictionary and the emotion analysis method based on feature extraction machine learning. The method of emotion

ISSN: 2414-1895
DOI: 10.6919/ICJE.202102_7(2).0020

analysis based on emotion dictionary is to extract the non-institutionalized features of text with the help of emotion dictionary, which depends on the construction of emotion dictionary. The emotion analysis method based on feature extraction machine learning requires manual feature selection, which is liable to lose the grammatical and semantic information of text. With the development of deep learning technology, the method based on deep learning can avoid the work of extracting features manually, which plays an increasingly important role in the analysis. In recent years, the depth of the learning method is to combine recurrent neural network and attention mechanism , attention mechanism is used to collect the importance of different sequence of words, but unable to get access to the importance of different sentences, ignoring the sentence on the importance of the micro blog this emotional tendency, so just single layer is difficult to obtain deep emotional attention characteristic information.

Therefore, in this study, first of all, a crawler was used to take the microblog text, and the text was pre-processed by manual annotation, de-weighting, word segmentation, etc. On the basis of word vector Word2vec processing, the model was constructed by combining the double-layer attention mechanism and Bi-LSTM. In this model, attention mechanism is added at word level and sentence level to capture the importance of different words and sentences respectively for feature weight learning.Overall consideration of the emotional tendency of the whole micro-blog book improves the accuracy of the emotional classification results. In order to verify the performance of the model, SVM, RF and LSTM models were used as the comparison based on the same word vector encoding.

2. Research

With the continuous rise of the network social media represented by sina weibo, users can express their views and opinions on a certain topic on the microblog platform. Microblog analysis is to analyze the data obtained from microblog, which is helpful for the study of language application and other issues.For example, users' information of interest can be obtained from the language, popular topics of popular discussion can be tracked, and product opinions can be determined [3]. Microblog emotion analysis is mainly based on the method of emotion dictionary, machine learning method based on feature information extraction and deep learning method.

The method based on emotion dictionary mainly uses the emotion dictionary resources to extract the emotion expression keywords contained in the text, so as to realize the classification of the text emotion. In the earlier research, Wu [4] proposed a data-driven emotion dictionary classification method for weibo, which designed a unified framework of emotion knowledge containing three dictionaries. Li Xin [5] extracted a large number of emotional keywords from corpus with a certain scale of emotional tendency, and extended the existing emotional vocabulary knowledge base by using similarity calculation method. The method based on emotion dictionary can reflect the unst-ructured characteristics of text, and the classification results are obvious when the coverage rate and labeling accuracy of emotion words are high in the dictionary. But how to build a high-quality emotional dictionary is difficult.

The machine learning method based on feature information extraction is a monitoring method. It extracts feature information from text, and then uses text classification algorithms such as RF and SVM to judge emotional attributes. Wawre [6] used a feature method. The training set is a film comment set with marked emotional tendency. First, the feature is extracted in the paper, and then the machine learning algorithm is used for training. Pang B [7] used machine learning and minimal cutting framework to show that the application of text classification to subjective analysis of documents can effectively improve the performance of emotion classification. The suitability of feature selection is one of the main factors affecting the classification effect of supervised learning. Therefore, the machine learning method based on feature information extraction relies too much on feature selection.

In recent years, deep learning has become a branch of machine learning with strong learning ability. The method based on deep learning can express the words, sentences and texts in a vectorized way

ISSN: 2414-1895 DOI: 10.6919/ICJE.202102_7(2).0020

to learn the deep semantic information of the text. This method has strong feature learning ability and eliminates the steps of constructing emotion dictionary and selecting features manually. YUAN[8] discussed in detail some models of emotion analysis and proved that deep learning can achieve good results in emotion analysis. Liang [9] presents a based on polarity transfer and length of emotional memory network analysis, to expand LSTM type based on the structure of the chain network to RNN, makes the model can be better access to history information of text, at the same time to dig into the structure of the text information, rich characteristics of learning, but to the structural features of text for learning how to effectively is still a problem to be solved. Guan Pengfei [10] proposed the BiLSTM emotion analysis model with enhanced attention, and adopted the word vector attention mechanism as a parallel structure with BiLSTM. Yang [11] realized sentence level emotion classification based on convolutional neural network based on weibo comment text, and optimized the method according to the sentence length of pooling layer.The method based on deep learning achieves a good effect in the text classification task of microblog emotion analysis [12,14]. In recent years, the depth of the learning method is to combine recurrent neural network and attention mechanism , attention mechanism is used to collect the importance of different sequence of words, but unable to get access to the importance of different sentences, ignoring the sentence on the importance of the micro blog this emotional tendency, so just single layer is difficult to obtain deep emotional attention characteristic information.

To sum up, the emotion analysis method based on deep learning has achieved a lot of research results. However, the current emotion analysis method of Microblog is to study a single microblog text as a whole, ignoring the text structure and emotional structure of microblog.Therefore, this paper proposes to take words and sentences as the minimum units respectively, use BiLSTM network to learn long-term dependence relationship, adjust feature weight by using attention mechanism, and pay attention to the valuable parts of emotion analysis, so as to improve the classification effect of emotion analysis.

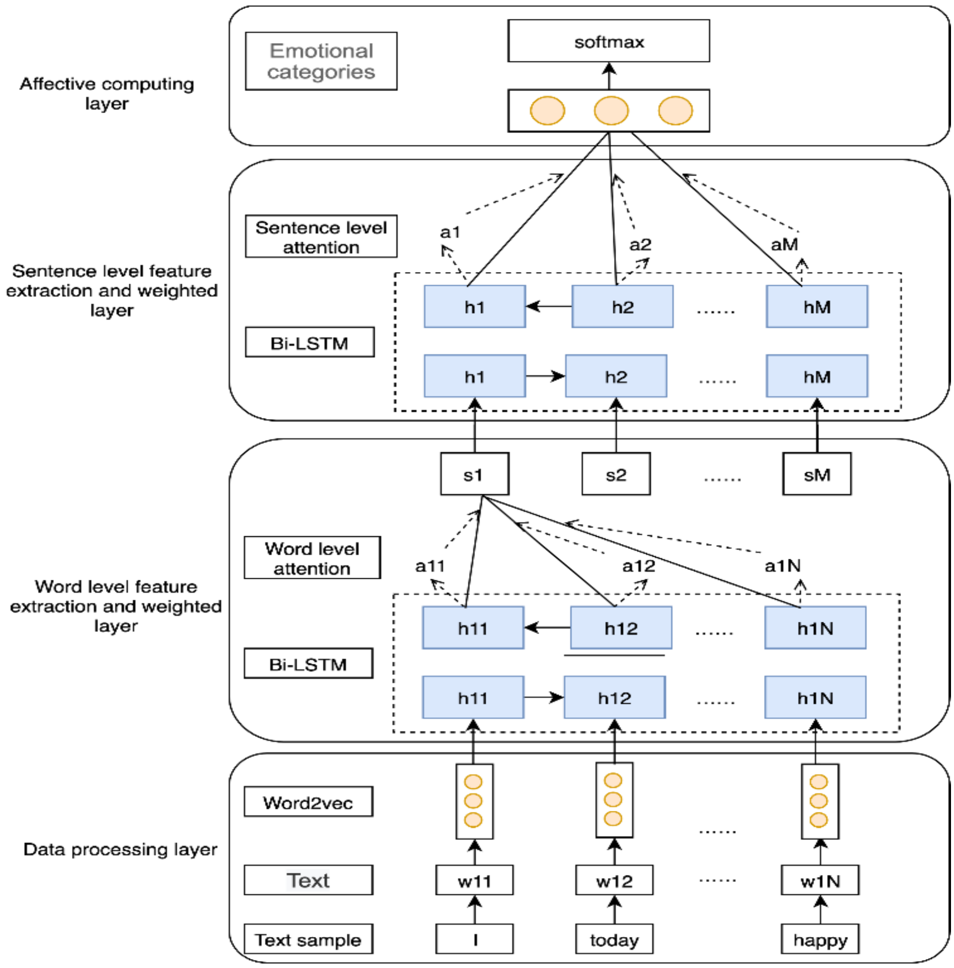

Figure 1. bi-LSTM model architecture based on two-layer attention

ISSN: 2414-1895 DOI: 10.6919/ICJE.202102_7(2).0020

3. Microblog emotion analysis model

This section focuses on the details of the Bi-LSTM model based on double-layer attention, see Figure 1. Suppose the microblog training set has K microblog corpus, each microblog has M sentences, and any sentence has N words. The emotion analysis model constructed in this paper mainly includes data processing layer, word feature learning and feature weight learning layer, sentence feature learning and feature weight learning layer, and emotion classification layer.

3.1 Data Processing

In this study, first of all, the crawler technology was used to crawl the microblog data and conduct pre-processing such as manual annotation, emoji removal, Jieba segmentation and so on.Then the word is embedded to get the word vector representation of each word.

Before deep learning was applied to NLP, the traditional expression of words was one-hot [15]. However, one-hot method does not consider the relationship between words, so it cannot well express the similarity between words. Moreover, the features obtained by one-hot are discrete and sparse, which is easy to cause the disaster of dimension. To solve these problems, Word2Vec is adopted in this study. Mikolov proposed the Word2Vec model [16], in which word vectors can better learn the meaning information contained in words in a low-dimensional space. The data input and output of this model are generally divided into CBOW (Continuous Bag-of-words) and Skip-Gram. The training input of CBOW model is the word vector corresponding to the context-dependent word of a particular feature, while the output is the word vector of a particular word. Skip-gram model is opposite to CBOW's idea, that is, the input is the word vector of a specific word, while the output is the context word vector corresponding to a specific word. CBOW is more suitable for small databases and skip-gram performs better in large corpora.

3.2 Bi – LSTM

Long and Short term memory network (LSTM) is a kind of cyclic neural network RNN[17], which contains a loop and allows the persistence of information. The long - and short-term memory network is a special type of RNN that can learn to rely on information over long periods of time.As first proposed by Hochreiter and Schmidhuber [18], the long-term dependence problem can be avoided through deliberate design, and the problems of RNN gradient disappearance and gradient explosion can be overcome at the same time. The LSTM unit is shown in Figure 2, mainly consisting of 4 interaction layers and 3 gates. The short-term and long-term memory networks control the passage of information through a "gate" structure, which consists of a Sigmoid neural network and a point-by-point multiplication operation, as shown in Figure 3.The LSTM has three gates to protect and control the cell state, namely the forget gate, the input gate and the output gate.

(1) "Forget gate": Deciding what information to discard from the cellular state. The gate reads $h_{t-1}$and x_t, Output (values between 0 and 1) is given to each number in the cell state C_{t-1}. The expression of f_t is shown in formula (1).

$$f_t = \sigma\left(W_f \cdot [h_{t-1}, x_t] + b_f\right) \tag{1}$$

(2) "Input gate": Determines what new information is stored in the cell state. There are two parts: Firsta Sigmoid layer is called an "input gate" layer to determine which values to update. Then, a tanh layer creates a new candidate value vector $\tilde{C}_t$, It might be added to the cellular state. Next, the cell state is updated by combining these two parts of information. The discarded value i_t is shown in formula (2), the candidate content $\tilde{C}_t$ is shown in formula (3), and the new state C_t is shown in formula (4).

$$i_t = \sigma(W_i \cdot [h_{t-1}, x_t] + b_i) \tag{2}$$

$$\tilde{C}_t = tanh(W_c \cdot [h_{t-1}, x_t] + b_c) \tag{3}$$

$$C_t = f_t * C_{t-1} + i_t * \tilde{C}_t \tag{4}$$

ISSN: 2414-1895
DOI: 10.6919/ICJE.202102_7(2).0020

(3) "Output gate": Determines the output value based on the current cell state. First, a Sigmoid layer determines which parts of the cell state will be output to o_t, as shown in Formula (5). Secondly, the cell state C_t is processed through tanh to obtain a value, which is multiplied by the output of Sigmoid gate. The final output gate determines the output information h_t, as shown in Formula (6).

$$o_t = \sigma(W_o \cdot [h_{t-1}, x_t] + b_o) \tag{5}$$

$$h_t = o_t * tanh(C_t) \tag{6}$$

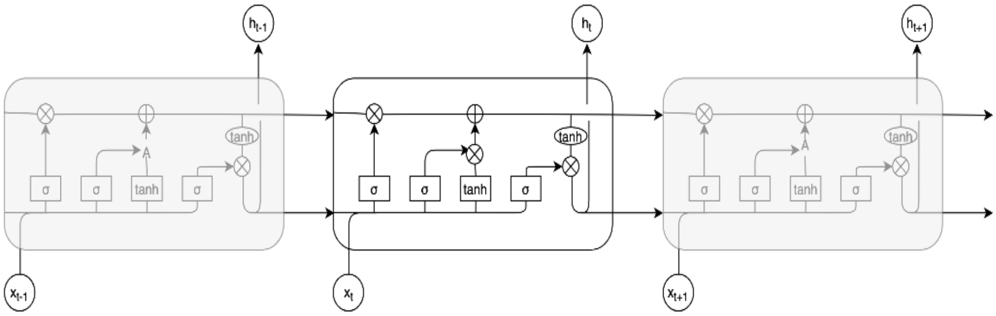

Figure 2. LSTM neural unit

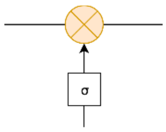

Figure 3. Gates in the long and short term memory network

Because the length of the micro-blog text is different, the specification of the text is not standard and other characteristics, micro-blog text in the characteristics of interdependence may be at a relatively long distance.Considering that LSTM is a forward propagation algorithm, in order to better capture the semantic information contained in the context, backward propagation algorithm needs to be integrated for learning. Therefore, bi-directional long and Short term memory network (Bi-LSTM) is used in this study [19], so as to learn the characteristics of long distance dependence that may exist in texts.As shown in Formula (7)- Formula (9).

$$\overrightarrow{h_t^f} = \mathrm{LSTM}(x_t) \tag{7}$$

$$\overleftarrow{h_t^b} = \mathrm{LSTM}(x_t) \tag{8}$$

$$H_t = [\overrightarrow{h_t^f}; \overleftarrow{h_t^b}] \tag{9}$$

Where, h in formula (7) represents the hidden layer state from front to back for each time step t; h in formula (8) represents the hidden layer state from back to front for each time step t.The splicing of the two is H_t representing Bi-LSTM extracting text feature vector. x_t represents the input node at time t.

3.3 Attention

Attention mechanism is similar to the attention mechanism of human brain, which is a mechanism of resource allocation.For example, when we look at something, we pay special attention to some characteristics of that thing, so it is impossible to pay attention to all characteristics, namely, on this thing, people's attention distribution is different.The core goal of the attention mechanism in deep learning is to select the information that is more important to the current task goal from the numerous information.At present, attention mechanism has been widely used in various types of deep learning tasks such as natural language processing [20], image recognition and speech recognition.

ISSN: 2414-1895 DOI: 10.6919/ICJE.202102_7(2).0020

In micro blog this sentiment analysis, to consider what words in the sentence can express emotion more and more text which sentence in the text of the emotions, thus giving the key to express emotion words higher weight, so this research adopts the double attention mechanism study respectively the different word in the sentence of weight distribution and different sentences in the text of the weight distribution. The basic calculation of attention mechanism [21] is shown in formula (10)-(12)

$$u_t = sigmoid(W_c H_t + b_w) \tag{10}$$

$$a_t = softmax(u_t) \tag{11}$$

$$\mathrm{v} = \sum a_t H_t \tag{12}$$

Where, H_t in formula (10) is the text feature extracted by Bi-LSTM, W_c is the weight parameter,b_w is the bias parameter, and u_t is the output of attention layer. a_t in Formula (11) is the attention-weight matrix calculated by softmax function.In formula (12), v is the attention weight feature obtained by combining the weight matrix and text features.

3.4 Affective computing

The emotion analysis in this study is essentially a classification problem.In order to carry out the final emotion classification, the emotion classification layer inputs the output vector v of word-level attention and sentence-level attention into softmax function to calculate the probability of emotion category. The calculation process is shown in Formula (13). In this study, the binary cross entropy loss function was used as the objective function to reduce the gap between the real result and the predicted result, so as to train and update the parameters in the model by using the back propagation mechanism.

$$\mathrm{Y} = \mathrm{softmax}(W_c \mathrm{v} + \mathrm{b}) \tag{13}$$

Where, W_c and b in formula (13) are weight parameters and bias parameters respectively, and Y is the result matrix of emotion prediction.

4. Experimental results and analysis

4.1 Experimental data

The experimental data set of this study is obtained through the crawler micro-blog book, which climbs down a total of 30,000 micro-blog books in chronological order.After deleting invalid text and duplicate text, a Boolean value was used to manually mark the emotional tendency of the mic-roblog text, with 1 representing positive and 0 representing negative.In the end, a total of 10,000 micro-blog books are retained, including 5,000 positive text data and 5,000 negative text data.

In the experiment, the data set is divided into training set and test set. 80% of the data is selected as the training set and 20% of the data as the test set.Table 1 shows a sample of the microblogs with good polarity.

Table 1. Sample microblog data set

Emotional polarity	text
1(positive)	Xiao hei, a stray dog, insisted on illegal invasion every day in order to meet his friend Sooni. The dog's persistence impressed the owner, who eventually adopted xiao hei from Sooni's family.Dogs may not understand death, but they are so affectionate...[Sadness][heart]{% dog's second video per day %}
0(negative)	Gouren, really angry to run into a fake money [anger] is also made of paper money (true or false) but one is recognized by the state, when can cancel the paper money is good [sweat][HMMMM] tear don't pit others...

ISSN: 2414-1895 DOI: 10.6919/ICJE.202102_7(2).0020

After manually marking the emotions of the microblog text, this study firstly uses the regular method to delete all symbols except letters, numbers and Chinese characters. The text is then segmented using Jieba, a third-party Python library. In this study, Google open source Word2vec tool was used to train the word vector. The CBOW model was selected. The context window size was set to 7, the word vector dimension size was set to 150, the Alpha learning rate was set to 0.025, and other parameters were set by default. When using pre-trained word vectors, we will use special symbols such as '<OOV>' instead. Then the deep learning model was trained using tensorFlow open source deep learning framework.

4.2 Evaluation Criteria

In order to evaluate the performance of the algorithm, four indexes were evaluated, namely Precision (P), Recall (R), F1-measure and Accuracy (Acc). The relevant calculation is shown in formula (14)-Formula (17).

The accuracy rate P refers to the correct proportion of the sample judged as positive emotion by the training model

$$\text{Precision} = \frac{tp}{tp + fp} \tag{14}$$

Recall rate R refers to the proportion of correct judgment by the trained model in the sample with actually positive emotion.

$$\text{Recall} = \frac{tp}{tp + fn} \tag{15}$$

The value of F1 is a composite metric and is the harmonic average of P and R.

$$\text{F1} = \frac{2 \times \text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \tag{16}$$

Acc is the proportion of correct prediction in all samples.

$$\text{Accuracy} = \frac{tp + tn}{tp + fp + tn + fn} \tag{17}$$

Among them, *tp* represents the number of records of the microblog data that are actually positive emotion and judged as positive emotion by the training model, *tn* represents the number of records of the microblog data that are actually negative emotion and judged as negative emotion by the training model, both of which are correct records predicted. *fp* represents the number of records of the microblog data that are actually negative emotion and judged as positive emotion by the training model, *fn* represents the number of records of the microblog data that are actually positive emotion and judged as negative emotion by the training model, both of which are the number of records that are wrong in prediction

4.3 Experimental parameter setting

See Table 2 for the value and description of model super parameters in this experiment. The training of the model is optimized by the Adam [22] optimizer. Some of the parameters are based on model training and hardware conditionals, such as the size of each batch of data, the number of LSTM hidden layer units, and the choice of the optimizer.

Table 2. Model hyperparameter Settings

parameter	instructions	value
Optimizer	Optimizer to calculate update model parameters	Adam
Hatch_size	The size of each batch of data	32
Lstm_units	Number of Lstm hidden layer units	128
Maxlen	Maximum sequence length, long is truncated, short is filled	100
Learning_rate	Learning rate	0.01
Dropout	The ratio of randomly disconnected input neurons	0.5

ISSN: 2414-1895 DOI: 10.6919/ICJE.202102_7(2).0020

4.4 Experimental results and analysis

This study choose the SVM, RF and LSTM as a classifier to set three control group, use the same word vector model output as input of each classifier, the sample data is divided into 10, of each evaluation index data calculation, then the calculation results averaged 10 times to get the final evaluation results. Therefore, the results of microblog emotion analysis model (AT-Bi-LSTM) proposed in this paper are compared and the performance of the model is evaluated. The experimental results are shown in Table 3.

Table 3. Compares the final experimental results of the model

model	P	R	F1	Acc
SVM	88.14%	75.15%	80.13%	83.45%
RF	92.46%	93.45%	93.01%	92.67%
LSTM	95.36%	95.92%	95.63%	95.81%
AT-Bi-LSTM	97.59%	96.89%	97.29%	96.92%

As can be seen from Table 3, the trend of F1 value in the four groups of experiments is basically consistent with Acc, and the final result shows that the model based on deep learning is more accurate than the model based on machine learning. The AT-Bi-LSTM proposed in this paper combined with the emotional structure characteristics of microblog texts, F1 value and Acc improved by 1.66% and 1.11% compared with LSTM model of deep learning, and Acc increased by 13.47% and 4.25% compared with SVM and RF of machine learning. This shows that the algorithm proposed in this paper has good classification effect and performance.

Since various machine learning algorithms and deep learning algorithms were set up in the comparison experiment, in order to fully explore the performance of each model, this study will explore the influence of data set size change on each model with data volume as the dependent variable under the premise that other processing steps remain unchanged. Assuming that the specified amount of data in a certain experiment is M, the positive and negative emotion corpora are randomly selected M/2. $M \in [1000,20000]$ was extracted at the interval of 1000. See Figure 4 and Figure 5 for the changes of F1 and Acc values of the four models.

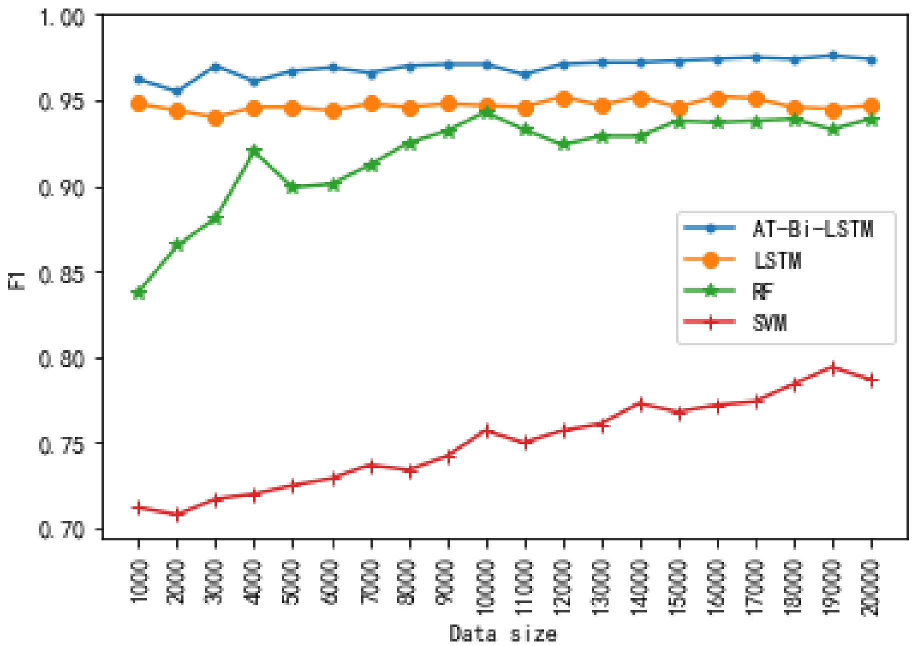

Figure 4. Influence of data volume on model F1 value

ISSN: 2414-1895
DOI: 10.6919/ICJE.202102_7(2).0020

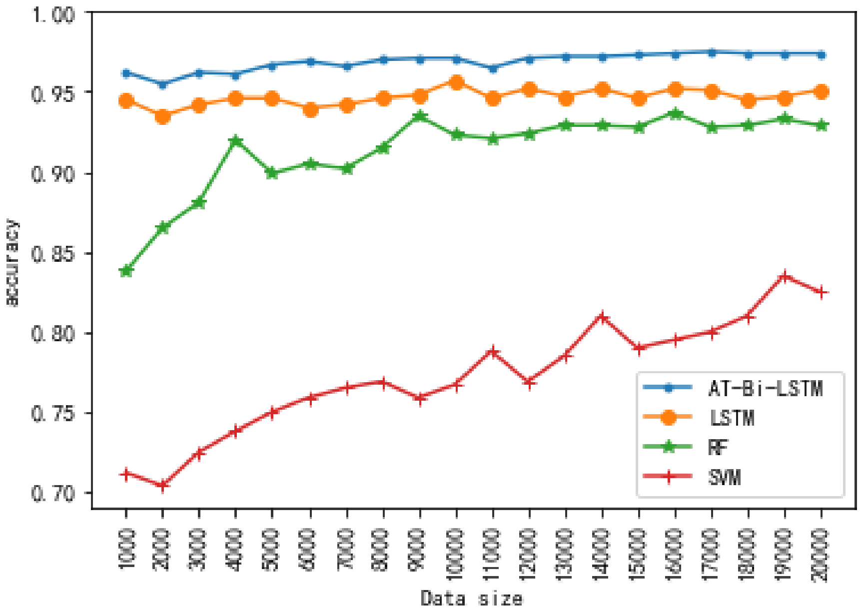

Figure 5. Influence of data volume on model Acc

As can be seen from Figure 4 and Figure 5, the F1 value and Acc of SVM are most affected by the size of the data set, and compared with the other three emotion classification models, SVM has the worst performance. When RF data volume reaches 8000, classification F1 and Acc tend to be stable, and the index value remains above 90%. The two algorithms based on deep learning are least affected by the data volume, and the variation range of F1 and Acc is within 4%, among which AT-Bi-LSTM is the most stable, the average value of F1 and Acc is 96.94% and 96.89%, respectively.In general, among the two machine learning algorithms, RF model has a little better performance than SVM model, and both F1 value and Acc of these two models tend to be stable with the increase of data volume.In the case of data volume change, the classification performance of deep learning model is better than that of machine learning classification model. On the basis of LSTM, the classification index F1 value and Acc have been improved in the AT-Bi-LSTM.

5. Conclusion

This paper proposes a Bi-LSTM emotion analysis model based on the two-layer attention mechanism, with the aim of more accurately mining the emotional tendency of microblog.First of all, this study use time-series modeling characteristic of Bi - LSTM algorithm to extract the micro blog this word level, sentence level text characteristic, by combining the characteristics of attention mechanism, put forward the double attention to strengthen respectively microblogging word level, sentence level characteristics of semantic focus, at last, through the whole connection layer fusion feature weights, and the final emotional calculation.

The experimental results show that the microblog emotion analysis model proposed in this paper shows superior performance and exceeds the existing models in many indicators. Finally, the influence of data set size on the five algorithms is further explored and analyzed. Although the model presented in this paper shows superior performance in the experimental results, the minimum granularity of the model is the word level. Considering the short size of microblog text, this study can make further refinement in the aspect of text granularity, and the next step will be to study the use of attention mechanism at word level.

References

[1] Ran li, Lin, Helen lam, Wang Weiping, Meng Dan. A review of textual emotion analysis [J]. Computer research and development, 2018,55(01):30-52. (In Chinese)

ISSN: 2414-1895 DOI: 10.6919/ICJE.202102_7(2).0020

[2] Lei Zhang, Shuai Wang, Bing Liu. Deep learning for sentiment analysis: A survey[J]. Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery,2018,8(4).

[3] Zhou Ying, Liu Yue, Cai Jun. Emotional analysis of Weibo based on attention mechanism [J]. Information theory and practice, 2018,41(03):89-94. (In Chinese)

[4] Fangzhao Wu, Yongfeng Huang, Yangqiu Song, Shixia Liu. Towards building a high-quality microblog-specific Chinese sentiment lexicon[J]. Decision Support Systems, 2016,87.

[5] Li xin, LiYu sunny, Han Xu, song dandan, Lejian liao. A multiclass emotion analysis method and system for bilingual micro-blog book [P]. Beijing:CN104331506A,2015-02-04. (In Chinese)

[6] Wawre S V, Deshmukh S N. Sentiment classification using machine learning techniques[J].International Journal of Science and Research (IJSR)2016, 5(4):819-821.

[7] Pang B. A Sentimental Education: Sentiment Analysis Using Subjectivity Summarization Based on Minimum Cuts[C]// Meeting of the Association for Computational Linguistics. 2004.

[8] YUAN JianHua, WU Yang, LU Xin, ZHAO YanYan, QIN Bing, LIU Ting. Recent advances in deep learning based sentiment analysis[J].Science China(Technological Sciences), 2020,63(10):1947-1970.

[9] Liang Jun, Yumei chai, huibin Yuan, Gao Minglei, Hongying zan. Emotion analysis based on polarity shift and LSTM recursive network [J]. Chinese Journal of Information technology, 2015, 29(05):152-159. (In Chinese)

[10] GuanPengFei, Baoan li, Lv Xueqiang, Zhou jianshe. The bidirectional LSTM emotion analysis of enhanced attention [J]. Chinese Journal of Information technology,2019,33(02):105-111. (In Chinese)

[11] Xiaoyilei Yang, Shuaijing Xu, Hao Wu, Rongfang Bie. Sentiment Analysis of Weibo Comment Texts Based on Extended Vocabulary and Convolutional Neural Network[J]. Procedia Computer Science, 2019, 147.

[12] Chen xin, Yang xiaobin, Yao yuhong. A text emotion classification method based on BLSTM algorithm and word fusion [J]. Journal of China Metrology University, 2020,31(02):225-232. (In Chinese)

[13] Chen Ling, Song yanxin. Research on LSTM Emotion analysis based on public emotion context——Take Typhoon Lekima [J]. Modern intelligence,2020,40(06):98-105. (In Chinese)

[14] Shaoxiu Wang, Yonghua Zhu, Wenjing Gao, Meng Cao, Mengyao Li. Emotion-Semantic-Enhanced Bidirectional LSTM with Multi-Head Attention Mechanism for Microblog Sentiment Analysis[J]. Information, 2020,11(5).

[15] Tedesco, Joe. Return of Code Clusters: One Hot Nest[J]. EC & M,2014.

[16] Mikolov T, et al. Efficient estimation of word representations in vector space[C]//Proceedings of the International Conference on Learning Representations, 2013.

[17] Mohammed Farsi. Application of ensemble RNN deep neural network to the fall detection through IoT environment[J]. Alexandria Engineering Journal,2020.

[18] Hochreiter S, Schmidhuber J. Long Short-Term Memory[J]. Neural Computation, 1997, 9(8):1735-1780.

[19] Schuster M, Paliwal K K. Bidirectional recurrent neural networks[J]. IEEE Transactions on Signal Processing, 2002, 45(11):2673-2681.

[20] Bahdanau, Dzmitry, Cho, Kyunghyun, Bengio, Yoshua. Neural Machine Translation by Jointly Learning to Align and Translate[J].,2014.

[21] Ling L, Zhihao Y, Pei Y, et al. An Attention-based BiLSTM-CRF Approach to Document-level Chemical Named Entity Recognition[J]. Bioinformatics (8):8.

[22] Dokkyun Yi, Jaehyun Ahn, Sangmin Ji. An Effective Optimization Method for Machine Learning Based on ADAM[J]. Applied Sciences,2020,10(3).