Journal of Computers Vol. 34 No. 1, February 2023, pp. 1-11
doi: 10.53106/199115992023023401001

An Emotional Analysis Method Based on Multi Model Ensemble Learning

Chang-You Zhang, Jing-Jing Wang*, Li-Xia Wan, Ruo-Xue Yu

College of Artificial Intelligence, Jiangxi University of Engineering, Xinyu, China
{448754885, 495567725, 1240147131, 1154550825}@qq.com

Received 9 March 2022; Revised 5 June 2022; Accepted 7 July 2022

Abstract. Traditional machine learning models generally use weak supervision model, which is difficult to adapt to the scene of multi classification for emotional text. Therefore, a multi model ensemble learning algorithm for emotional text classification is proposed. The algorithm takes the labeled emotional text data as the training sample, uses the improved TF-IDF algorithm to train the word vector space model, selects three weakly supervised machine learning algorithms, linear SVC, xgboost and logistic regression, to construct the base classifier, and uses the random forest algorithm to construct the meta classifier. It realizes the function of dividing emotional text into three categories: positive, neutral and negative. From the simulation and test results, the AUC values of the multi model ensemble learning algorithm model for each category are 0.93, 0.94 and 1.00, and the AP values are 0.87, 0.86 and 1.00, and the indicators of accuracy and recall are better than the single machine learning model, which realizes the high performance and high accuracy for emotional text classification.

Keywords: text multi classification, ensemble learning algorithm, word vector space model, emotional analysis

1 Introduction

With the wide use of various network platforms, more and more emotional information is published on social media and e-commerce websites, such as comments on commodities, views on hot events, political events, etc. These information can reflect citizens' experience of using com-modities, reflections on hot events, and attitudes towards national policies. Emotional analysis of comments on various network platforms, on the one hand, is of great significance for traditional consumers and enterprises to collect opinions on products or services, on the other hand, it also plays an important role in national security and public opinion monitoring. How to realize automatic and efficient mining and analysis of these comments and the emotions contained in the text has become a hot spot, and text emotion analysis technology came into being.

Emotional analysis is the process of analyzing, processing, summarizing and reasoning subjective texts with emotional color, which also known as opinion mining and propensity analysis. Generally, emotional texts are divided into three categories: positive, neutral and negative [1]. Emotional analysis is a research field with a wide range of application scenarios. According to the granularity of text processing, emotional analysis can be divided into three research levels: word level, sentence level and text level [2]. A basic step of text emotional analysis is to classify the emotional polarity of known words in the text, which may be sentence level or function level. This paper mainly studies the three polar classification for emotional text, that is, comment texts are divided into positive, neutral and negative categories. How to classify text accurately and efficiently is the problem to be solved in this paper.

In recent years, many scholars have done a lot of research work in the field of text emotional analysis. Ji Shen and Fangbi Tan proposed a deep neural network based sentiment analysis model for MOOC course reviews, which uses Bidirectional Long Short-Term Memory Network (BiLSTM) to analyze Chinese semantic and introduces two methods to balance it and adds dropout mechanism to prevent the over fitting of the model [3]. However, all the review data sets for the emotion analysis model are simply divided into positive and negative ones, which may have a certain impact on the model's actual application effect. Qingqing Wang proposes a hybrid neural network model based on CNN-BGRU to solve the problem of accurate classification, which can obtain more semantic information between texts and better capture the dependence of specific emotions in the whole text [4]. Angel Deborah presented a Multi Kernel Gaussian Process (MKGP) regression model for emotion analysis on news headlines and fine-grained sentiment analysis of financial microblogs and news [5]. However, the model

* Corresponding Author

has the disadvantages of complex design and long prediction time.

In order to accurately identify the emotional polarity and emotional state contained in the specified comment text, an emotional analysis algorithm based on multi model ensemble learning is proposed in this paper. The algorithm mainly includes two parts: constructing word vector space model and training classification model. The word vector space model takes the manually labeled emotional text as the input sample, segments the text part through the Jieba word segmentation library, and uses the improved TF-IDF algorithm to complete the construction of the word vector space model. The word vector space model can convert the emotional text into feature vectors; The classification model takes the feature vector as the input, uses three weakly supervised machine learning algorithms, such as linear SVC, xgboost and logistic regression algorithm, as the base classifier, and constructs the meta classifier with random forest algorithm to complete the training of the whole ensemble learning classification model. The algorithm can divide emotional texts into three categories: positive, neutral and negative.

2 Related Work and Our Contributions

The purpose of emotional analysis is to identify the overall emotion expressed in the text, which can be network comments, articles, microblogs and so on [6]. At present, there are three schemes in the field of text emotional classification: emotion dictionary matching, machine learning and deep learning [7].

The emotional dictionary matching scheme usually needs to segment the text to be tested, then match the word segmentation words with the emotional dictionary, find the emotional words, degree words, negative words, etc., and calculate the emotional tendency score of the text to be tested, so as to realize emotional analysis [8]. However, this scheme needs to manually construct an emotional dictionary composed of emotional words and design a series of correlation matching rules to complete the classification of polarity for emotional text, which is not efficient; In addition, simple emotional dictionary matching is difficult to reflect the implied semantics and context of the text, so this scheme is less applied [9].

Machine learning schemes are usually supervised learning schemes, which mainly include three steps: acquiring text data sets, extracting text emotional features and training classification models [10]. The analysis process is shown in Fig. 1. Li Wang presented a probabilistic semi-supervised learning (SSL) framework based on sparse graph structure learning, and proposed a simple inference approach for the embeddings of unlabeled data based on point estimation and kernel representation, which will get promising results in the setting of SSL compared with many existing methods and significant improvements on small amounts of labeled data [11]. Wang Hong proposed a form of supervised discrete hash algorithm to learn more stable hash codes by learning mutual similarity and using the relationship between different semantic tags [12]. Xiaodong Jia propose a semi-supervised multi-view deep discriminant representation learning (SMDDRL) approach, which comprehensively exploits the consensus and complementary properties as well as learns both shared and specific representations by employing the shared and specific representation learning network [13]. However, most of the above traditional machine learning models are weakly supervised models, which have a selective preference for the categories of text classification, and are difficult to adapt to the multi classification scenario of emotional analysis [14].

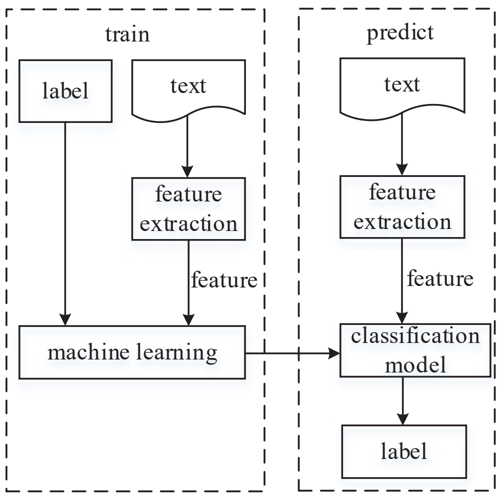

Fig. 1. Flow of machine learning text classification algorithm

Compared with the traditional machine learning scheme, deep learning technology can automatically and efficiently extract text features and realize the task of emotion classification [15]. At present, there are three mainstream deep learning models: feedforward neural network model, recurrent neural network model and convolutional neural network [16]. Feedforward neural network models are represented by Word2vec and Doc2vec. These models have simple structure and can get reliable results on specific tasks, but this model fails to make full use of training data, so the classification effect needs to be improved [17]. The recurrent neural network models mainly include RNN and LSTM models, which can capture the time sequence information of the text, but the memory of long text disappears in the process of model training, so it is difficult to identify the relationship between process words [18]. Convolutional neural networks are represented by DCNN and textcnn. These models can identify process semantics and key paragraph features, but some information will be lost in the pooling operation in the training process, which may lead to errors in emotional analysis [19].

The main contributions of this paper are as follows.

(1) Based on the traditional TF-IDF algorithm, the cross-category feature component is introduced to construct the word vector space model, which is used to convert text into word vector (feature vector).

(2) Three weak supervised models, linearsvc, xgboost and logistics expression, are used as base classifiers, random forest algorithm model is used as meta classifier, and stacking ensemble learning algorithm is used to train strong supervised model, so as to complete the construction of stacking ensemble learning model, which can realize the polarity classification of emotional text.

3 System Design of Multi-Model Ensemble Learning

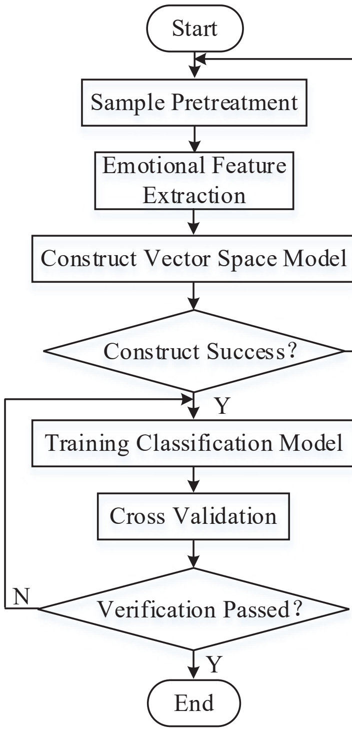

Fig. 2. Algorithm flow of multi model ensemble learning classification

In the field of traditional machine learning, the goal of supervised learning algorithm is to train a model with stable and good performance in all aspects, but the actual situation is often difficult to meet the requirements. More often, only multiple preferred models can be obtained, that is, weak supervised models with better classification performance in some categories [20]. Ensemble learning is to combine multiple weak supervised models in order to obtain a strong supervised model with better classification effect. In other words, even if a weak classifier gets wrong prediction, other weak classifiers can correct the error. In the field of ensemble learning, there are three schemes: bagging, boosting and stacking [21]. The stacking method obtains a new model by combining other models. As long as appropriate model combination strategy is adopted, the bagging and boosting methods can be expressed, which has better adaptability and generalization ability. Therefore, this paper designs a multi model ensemble learning model based on stacking.

The process of multi model ensemble learning classification algorithm is shown in Fig. 2. The whole process includes three steps: sample preprocessing, building word vector space model and training classification model. Firstly, the sample data is preprocessed, including filtering interference words and invalid characters, text word segmentation and removing stop words, so as to obtain labeled text word segmentation data; Then, taking the word segmentation data as the input, and the word vector space model is constructed by using the improved TF-IDF algorithm to calculate the feature vector of each text word segmentation; Finally, taking the feature vector of labeled text as the training set, the multi model ensemble learning algorithm is used to train the classification model by integrating three basic classification algorithms: linear SVC, xgboost and logistics expression.

4 Construct Word Vector Space Model

4.1 Training Set Acquisition and Preprocessing

The training samples mainly come from the comment texts of major portals, SNS community and microblogging sites that contain noise data. Before feature extraction, the original text data needs to be preprocessed, which is the key to feature extraction. A good preprocessing process can improve the quality of feature extraction and the performance of the classification algorithm [22]. The sample preprocessing group in this article includes the following steps.

(1) Jieba thesaurus is used to segment the text, and then a stop word dictionary including adverbs, adjectives and conjunctions without actual semantics is constructed;

(2) Part of speech tagging of word segmentation data, in other words, Judging the verb, noun, adjective and other attributes of word segmentation data.

4.2 Improved TF-IDF Algorithm

TF-IDF is a statistical method used to evaluate the importance of a word to a document in a corpus. TF (term frequency) represents word frequency [23]. The definition of probability representation for TF is as follows.

$$\mathrm{TF}(t_i, d_i) = \frac{n_{i,j}}{\sum_k n_{k,j}}. \tag{1}$$

Where $n_{i,j}$ is the appear times of word t_i in the document d_i, and the denominator $\sum_k n_{k,j}$ is the sum of all words appear times in the document d_i. IDF (inverse document frequency) represents the reverse document frequency, which is used to describe the importance of a word [24]. Its definition is shown as follows.

$$\mathrm{IDF}(t_i, D) = \log\left(\frac{|D|}{\left|\{d \in D : t_i \in d\}\right|}\right). \tag{2}$$

Where $|D|$ is the total number of documents in the corpus, $n_{i,j}$ represents the number of documents containing words in the corpus, and the definition of TF-IDF is shown as follows.

$$\text{TF-IDF} = \mathrm{TF} \times \mathrm{IDF}. \tag{3}$$

However, the word frequency can only be simply connected with the inverse document frequency in the traditional TF-IDF scheme, which can not reflect the distribution of words among different categories and is difficult to represent the cross category characteristics of documents [25]. Therefore, the training set is divided into j-class and non-j-class to improve the traditional IDF algorithm in this paper, and considers the distribution of words in different classes. The improved IDF formula of words in j-class is shown as follows.

$$\text{IMIDF} = \log\left(\frac{n_{i,j}}{n_{i,j}+m_{i,j}}|\text{D}|\right). \tag{4}$$

Where $n_{i,j}$ is the appear times of documents containing word t_i in j-class, and $m_{i,j}$ is the appear times of documents containing word t_i in non-j-class. The improved IDF value increases with the increase of $n_{i,j}$ while decreases with the increase of $m_{i,j}$. If the number of documents containing word t_i in j-class is more and the number of documents containing word t_i in non-j-class is less, then the word t_i can better represent j-class. In the task of text emotion classification, the word have the ability to distinguish between categories, and the word t_i in j-class should be given higher weight. At this time, the improved TF-IDF algorithm is shown as follows.

$$\text{TF-IMIDF} = \text{TF} \times \text{IMIDF}. \tag{5}$$

4.3 Construction of Word Vector Space Model

Most texts use natural language, and the information contained is unstructured and difficult to be processed by computer. Therefore, how to accurately represent text features is the main factor affecting the performance of text emotion classification [26]. In recent years, researchers have proposed many text representation models such as Boolean model, spatial vector model, latent semantic model and probability model to express the semantics of the text with a specific structure [27]. For the sake of classification efficiency, the multi model ensemble learning model proposed in this paper uses the space vector model to represent text. The space vector model is proposed by G Salton of Harvard University. The model converts a given text into a vector with high dimension, and takes the featured item as the basic unit of text representation. Each dimension of the vector corresponds to a feature item in the text, and each dimension itself represents the weight of its corresponding feature item in the text. The weight represents the importance of the feature item to the text, that is, the weight can reflect the degree of its document category [28].

Word vector space model is an algebraic model that represents a document as a vector. The similarity between documents is compared with the included angle between vectors [29]. Assuming that the number of all words in the corpus is T, the j-th document is d_j, and the document to be queried is q_i, then the vectors of the two document are expressed as follows.

$$d_j = \left(w_{1,j}, w_{2,j}, \cdots, w_{T,j}\right) \in \text{R}^T,\ q_i = \left(w_{1,i}, w_{2,i}, \cdots, w_{T,i}\right) \in \text{R}^T. \tag{6}$$

Where R^T represents the word vector space model, then the calculation formula of similarity between d_j and q_i is shown as follows.

$$\text{sim}(q_i, d_j) = \log\left(\frac{q_i \bullet d_j}{|q_i| \bullet |d_j|}\right). \tag{7}$$

5 Training Ensemble Learning Model

5.1 Selection of Base Classifier

Ensemble learning is a scheme to combine multiple base classifiers according to a certain algorithm to improve the overall classification performance of the model. It requires that the implementation principles of each base classifier are different, applicable to the same data set, and the classification effect is similar [30]. LinearSVC is a supervised machine learning algorithm that uses the hinge loss function as the kernel method to realize linear classification. The trained classifier has the characteristics of sparsity and robustness [31], and the overfitting control mechanism is introduced to ensure the generalization performance of the classifier. It is widely used in pattern recognition fields such as portrait recognition and text classification; Xgboost is a machine learning algorithm based on a gradient boosting classifier. It improves the existing gradient lifting algorithm by sparse feature processing and L1/L2 regularization, which has the advantages of reducing over fitting and improving computational efficiency [32]; Logistics regression is a generalized linear regression algorithm based on logistic function, which is commonly used in data mining, economic prediction, automatic disease diagnosis and other fields [33]. In view of the above reasons, three machine learning algorithms of linear SVC, xgboost and logistics regression are selected as base classifiers in this paper.

5.2 Training of Meta Classifier

The meta classifier is a two-level classifier, which takes the output combination of each base classifier as the input feature set. The selection of training algorithm needs to consider the randomness of training set sampling and the estimation variance of the model. Therefore, random forests is selected as the meta classifier in this paper. Assuming that the number of samples is N and the number of attributes of each sample is M, then the training steps of the classifier are as follows.

(1) N samples are randomly selected to train a decision tree as samples at the root node of the decision tree.

(2) m attributes (m << M) are randomly selected from M attributes, and a specified strategy, such as information gain, is used to select an attribute as the splitting attribute of the subordinate node to obtain the splitting node.

(3) Each child node in the decision tree is split according to step (2). If the attributes of the current node are the same as those used by the parent node, it means that the node has reached the leaf node, and then the splitting is completed.

(4) Repeat steps (1) ~ (3) to establish a large number of decision trees to complete the training process of random forest.

Assuming that the n-th decision tree after training is f_n , for the sample x of an unknown category, the prediction output of the classifier can be obtained according to the prediction mean of all decision trees, i.e

$$\hat{f} = \frac{1}{N}\sum_{n=1}^{N} f_n(x) . \tag{8}$$

The correlation calculation formula between each decision trees is shown as follows.

$$\sigma = \sqrt{\frac{\sum_{n=1}^{N}[f_n(x) - \hat{f}]^2}{N-1}} . \tag{9}$$

According to the calculation results, the multi decision tree model of random forest can reduce the correlation between models and reduce the noise of the base classifier.

5.3 Cross Validation

Since the multi model ensemble learning algorithm proposed in this paper adopts the scheme of two-level classifier, using the same sample to train the base classifier and meta classifier may lead to over fitting. In order to train a reliable and stable classification model, the k-fold cross validation method is usually used to evaluate the accuracy of the model. Then a 5-fold cross validation scheme is adopted in this paper. Firstly, the sample set is randomly divided into five groups, four groups of samples are selected to train each base classifier, and then the remaining group of samples is input into the trained base classifiers to obtain the prediction label as the feature vector for training the meta classifier; Finally, five groups of full sample sets are used to train each base classifier. At this time, the newly trained base classifier and meta classifier constitute the final multi model ensemble learning model. Algorithm 1 is the specific implementation process of cross validation.

Algorithm 1. Cross validation algorithm

```
Input: training data set D={Xj,Xi}^m_i=1 (xj ∊ R^T,yi ∊Y)
Output: stackingEnsemble learning classifier
a) Adopt cross validation approach in preparing a training set
Randomly split D into 5 equal-size subsets: D = {D1,D2,...,D5}
   for k ← 1 to 5
       // Train base classifiers
   for t ← 1 to T do
       Train a classifier hkt from Dk ∊ D
   end for
    // Construct a training set for meta classifier
    for xi ∊ Dk do
        Get one record{x'i,yi}, where x'i ={hk1(xi),hk2(xi),...,hk5(xi),}
    end for
  end for
b) Train a meta classifier
    Learn a new classifier h' from the collection of {x'i,yi}
c) Re- train base classifiers
    for t ←1 to T do
    Train a classifier ht, based on D
   end for
return H(x)=h'(h1(x),h1(x),...,h1(x))
```

6 Experimental Analysis

6.1 Experimental Data Set and Environment

The user comments of major portal websites, e-commerce platforms, microblog platforms, etc are used as the model evaluation data set, of which 8000 are positive comments and 7000 are negative comments. Before the training, the data of the two data sets were pretreated to remove stop words. The average length of the pretreated data of the hotel review data set is about 37, and the average length of the meituan catering review data set is about 48. The positive and negative comments of the data set were randomly disordered. 70% are selected as the training set and 30% as the test set for model evaluation. The experimental environment is shown in Table 1 below.

Table 1. Experimental environment

SN	Experimental environment	Details
1	Operating system	Windows 10
2	CPU	Intel i5-7300HQ 2.50GHz
3	GPU	GTX 1050
4	Memory	16GB
5	Open source library	sklearn, xgboost, mlxtend

6.2 Performance Indicator

Precision and recall are common indicators for evaluating machine learning models [34]. Assuming that the number of positive samples predicted by the model is TP, the number of positive samples predicted by the model is FP, the number of positive samples predicted by the model is FN, and the number of negative samples predicted by the model is TN, then the calculation method of accuracy rate and recall rate is shown in formula (10).

$$\text{Precision} = \frac{\text{TP}}{\text{TP+FP}},\ \text{Recall} = \frac{\text{TP}}{\text{TP+FN}}. \tag{10}$$

In addition, the true positive rate TPR and false positive rate FPR indicators are usually introduced to evaluate the sensitivity and misjudgment of the machine learning model [20], which is shown in formula (11).

$$\mathrm{TPR} = \frac{\mathrm{TP}}{\mathrm{TP+FN}}, \ \mathrm{FPR} = \frac{\mathrm{FP}}{\mathrm{FP+TN}}. \tag{11}$$

6.3 Result Analysis

ROC curve and PR curve are usually used to intuitively evaluate the classification effect of machine learning model. In this paper, 1000 labeled word segmentation samples are selected as the training data. The ROC curve obtained from the simulation test is shown in Fig. 3, in which FPR is the abscissa, TPR is the ordinate, and AUC (area under curve) is the area under the ROC curve, indicating the probability that the predicted positive example is ahead of the negative example. According to the simulation diagram, the AUC values corresponding to each category are more than 0.9, indicating that the model performance is good. The PR curve corresponding to the classification model is shown in Fig. 4. The curve takes recall as the abscissa and precision as the ordinate. The average accuracy value (AP) corresponding to each category is more than 0.85, indicating that the multi classification accuracy of the model is high.

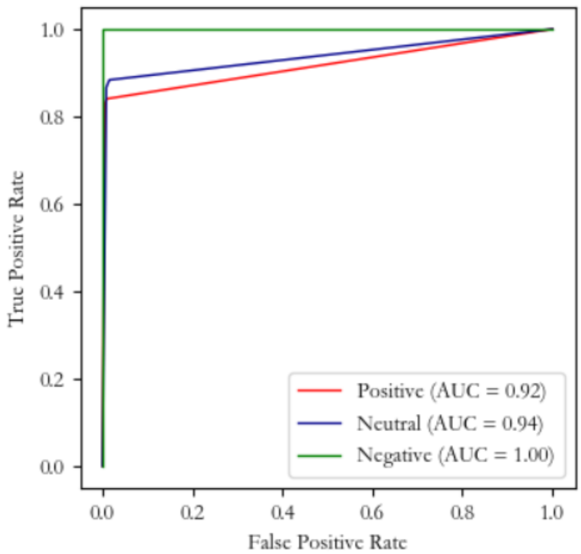

Fig. 3. ROC curve of multi model ensemble learning algorithm

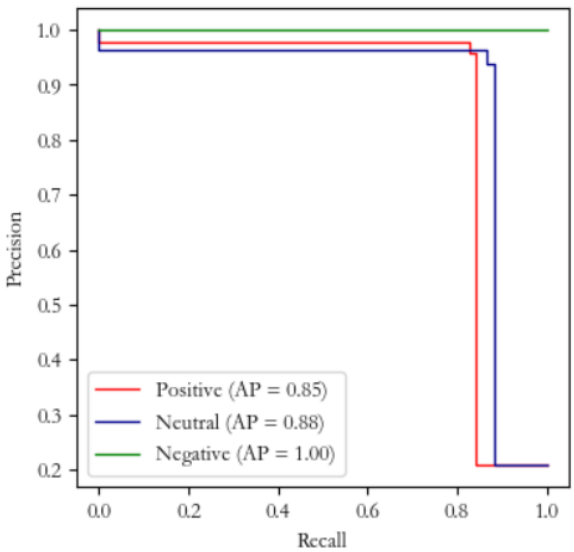

Fig. 4. PR curve of multi model ensemble learning algorithm

In addition, in order to compare the performance difference between single base model and multi model ensemble learning algorithm model, 3000 labeled samples is selected in this paper as test data and inputs them into linear SVC, xgboost, logistic regression and multi model ensemble learning algorithm models respectively. Taking precision and recall as evaluation indexes, the test results are shown in Table 2, Table 3, Fig. 5 and Fig. 6.

Table 2. The precision of each classification model

Id	Model	Precision		
		Positive	Neutral	Negative
1	LinearSVC	0.95	0.92	0.94
2	xgboost	0.94	0.89	0.96
3	logistic regression	0.94	0.90	0.92
4	ensemble learning	0.98	0.96	0.99

Table 3. The recall of each classification model

Id	Model	Recall		
		Positive	Neutral	Negative
1	LinearSVC	0.92	0.91	0.94
2	xgboost	0.95	0.94	0.89
3	logistic regression	0.90	0.88	0.92
4	ensemble learning	0.98	0.95	0.97

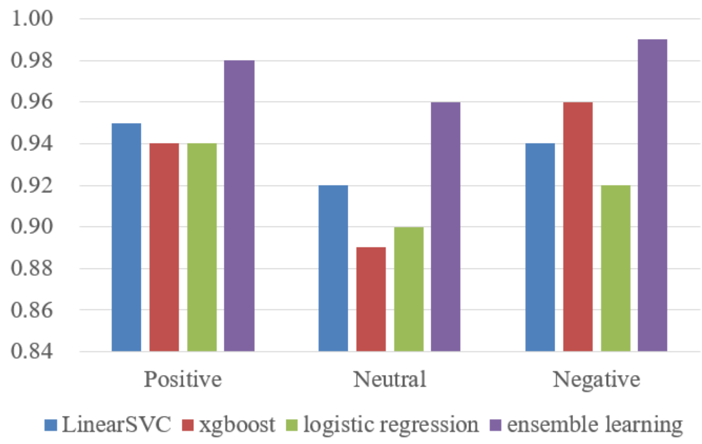

Fig. 5. Histogram of accuracy rate for each classification model

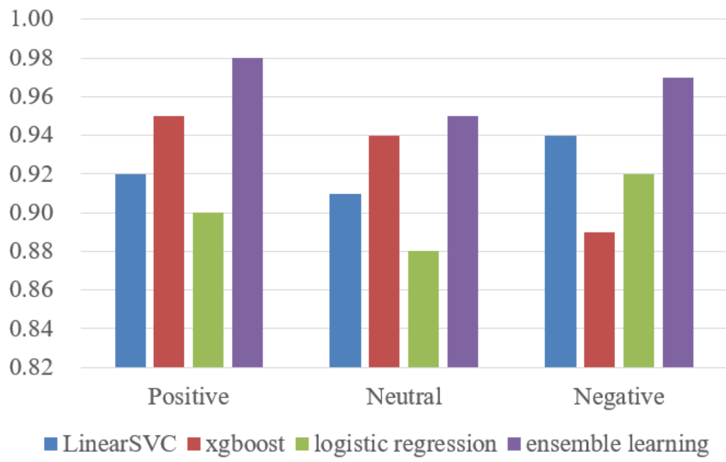

Fig. 6. Histogram of recall rate for each classification model

According to the test results, compared with a single machine learning model, the accuracy and recall rate of the multi model ensemble learning algorithm model in each category of data set are higher than that of the single machine learning model, which shows that the algorithm has better multi classification performance than the single machine learning model.

7 Summary and Prospect

This paper designs an emotional analysis model of multi model ensemble learning, which can divide emotional texts into three categories: positive, neutral and negative. Based on the word vector space model constructed by the improved TF-IDF algorithm, a two-level classification architecture composed of base classifier and meta classifier is designed, and a 5 fold cross validation scheme is introduced to eliminate the over fitting phenomenon between each sub model. From the simulation results, the AUC value of each category for multi model ensemble learning are more than 0.9 and AP value of the emotion classification algorithm for multi model ensemble learning are more than 0.85. In addition, the indicators of accuracy and recall are better than that of a single machine learning model, which realizes the high precision and high performance in emotional text multi classification.

In the follow-up research, we will study the emotional analysis of long text and fine-grained text in the future, and enrich the network vocabulary and cross domain sample set, which will improve the prediction accuracy of text emotional analysis.

References

[1] R. Chatterjee, S. Mazumdar, R.-S. Sherratt, R. Halder, T. Maitra, D. Giri, Real-Time Speech Emotion Analysis for Smart Home Assistants, IEEE Transactions on Consumer Electronics 67(1)(2021) 68-76.

[2] A. Drif, K. Hadjoudj, An Opinion Spread Prediction Model with Twitter Emotion Analysis during Algeria's Hirak, The Computer Journal 64(3)(2021) 358-368.

[3] J. Shen, F.-B. Tan, Emotion Analysis Model of MOOC Course Review Based on BiLSTM, International Journal of Emerging Technologies in Learning 16(8)(2021) 93-105.

[4] Q.-Q. Wang, J.-L. Luo, J.-W. Song, Emotion analysis method for elderly living alone based on CNN-BGRU neural network, International Journal of Wireless and Mobile Computing 20(4)(2021) 352-362.

[5] S.A. Deborah, T.T. Mirnalinee, S.M. Rajendram, Emotion Analysis on Text Using Multiple Kernel Gaussian, Neural Processing Letters 53(2)(2021) 1187-1203.

[6] C.-C. Li, J.-L. Wang, H.-W. Wang, M. Zhao, W. Li, X. Deng, Visual-Texual Emotion Analysis with Deep Coupled Video and Danmu Neural Networks, IEEE Transactions on Multimedia 22(6)(2020) 1634-1646.

[7] G.-H. Hu, G.-Y. Meng, Q.-S. Zhao, X. Zheng, Emotion feature optimisation based on PCA-GRA analysis, International Journal of Computing Science and Mathematics 12(4)(2020) 339-349.

[8] W.-W. Shen, S.-P. Zhang, Emotional tendency dictionary construction for college teaching evaluation, International Journal of Emerging Technologies in Learning 13(11)(2018) 117-129.

[9] Y.-H. Rao, J.-S. Lei, W.-Y. Liu, Q. Li, M.-L. Chen, Building emotional dictionary for sentiment analysis of online news, World Wide Web 17(4)(2014) 723-742.

[10] I. Psychoula, A. Gutmann, P. Mainali, S.H. Lee, P. Dunphy, F. Petitcolas, Explainable Machine Learning for Fraud Detection, Computer 54(10)(2021) 49-59.

[11] L. Wang, R. Chan, T.-Y. Zeng, Probabilistic Semi-Supervised Learning via Sparse Graph Structure Learning, IEEE Transactions on Neural Networks and Learning Systems 32(2)(2021) 853-867.

[12] H. Wang, X.-B. Liu, X.-S. Nie, Supervised discrete hashing through similarity learning, Multimedia Tools and Applications 80(11)(2021) 16215-16227

[13] X.-D. Jia, X.-Y. Jing, X.-K. Zhu, S.-C. Chen, B. Du, Z.-Y. Cai, Z-Y. He, D. Yu, Semi-supervised multi-view deep discriminant representation learning, IEEE Transactions on Pattern Analysis and Machine Intelligence, 43(7)(2021) 2496-2509.

[14] X. Han, Z.-Y. Zhang, Z.-Y. Liu, Knowledgeable machine learning for natural language processing, Communications of the ACM 64(11)(2021) 50-51.

[15] S.M. Mohamed, E.H. Mohamed, M.A. Belal, Polarity Detection of Dialectal Arabic using Deep Learning Models, International Journal of Advanced Computer Science and Applications 12(11)(2021) 212-218.

[16] C. Pelta, Emotional Cascade Model and Deep Learning, International Journal of Advanced Computer Science and Applications 12(8)(2021) 363-367.

[17] A. Haldorai, A. Ramu, Canonical Correlation Analysis Based Hyper Basis Feedforward Neural Network Classification for Urban Sustainability, Neural Processing Letters 53(4)(2021) 2385-2401.

[18] X.-M. Liu, F.-Z. Qi, Research on advertising content recognition based on convolutional neural network and recurrent neural network, International Journal of Computational Science and Engineering 24(4)(2021) 398-404.

[19] Y.-M. Hsieh, T.-J. Wang, C.-Y. Lin, L.-H. Peng, F.-T. Cheng, S.-Y. Shang, Convolutional Neural Networks for Automatic

Virtual Metrology, IEEE Robotics and Automation Letters 6(3)(2021) 5720-5727.
[20] P. Pham, P. Do, The approach of using ontology as a pre-knowledge source for semi-supervised labelled topic model by applying text dependency graph, International Journal of Business Intelligence and Data Mining 18(4)(2021) 488-523.
[21] C.-W. Ju, E.J. French, N. Geva, A.W. Kohn, Z. Lin, Stacked Ensemble Machine Learning for Range-Separation Parameters, Journal of Physical Chemistry Letters 12(39)(2021) 9516-9524.
[22] R.M. Bommisetty, A. Khare, T. Siddiqui, P. Palanisamy, Fusion of gradient and feature similarity for Keyframe extraction, Multimedia Tools and Applications 80(10)(2021) 15429-15467.
[23] H.-L. Yu, Y.-N. Ji, Q.-L. Li, Student sentiment classification model based on GRU neural network and TF-IDF algorithm, Journal of Intelligent and Fuzzy Systems 40(2)(2021) 2301-2311.
[24] Z. Zhang, Y. Lei, X.-G. Mao, X. Chang, J.-X. Xue, Q.-Y. Xiong, Fault Localization Approach Using Term Frequency and Inverse Document Frequency, Journal of Software 31(11)(2020) 3448-3460.
[25] J.-H. Qin, Z. Zhou, Y. Tan, X.Y. Xiang, Z.-B. He, A big data text coverless information hiding based on topic distribution and tf-idf, International Journal of Digital Crime and Forensics 13(4)(2021) 40-56.
[26] R. Mahajan, M. Zaveri, Harnessing Emotive Features for Emotion Recognition from Text, International Journal of Advanced Computer Science and Applications 12(7)(2021) 166-175.
[27] M. Lei, H.-Y. Huang, C. Feng, Multi-granularity semantic representation model for relation extraction, Neural Computing and Applications 33(12)(2021) 6879-6889.
[28] A. Nazir, R.-N. Mir, S. Qureshi, Idea plagiarism detection with recurrent neural networks and vector space model, International Journal of Intelligent Computing and Cybernetics 14(3)(2021) 321-332.
[29] V.-K. Kotte, S. Rajavelu, E.-B. Rajsingh, A Similarity Function for Feature Pattern Clustering and High Dimensional Text Document Classification, Foundations of Science 25(4)(2020) 1077-1094.
[30] X.-H. Zhu, Z.-Y. Zou, B. Qiao, K. Fang, Y. Chen, A named entity recognition model based on ensemble learning, Journal of Computational Methods in Sciences and Engineering 21(2)(2021) 475-486.
[31] Y. Wan, Extended SVC modeling for frequency regulation, IEEE Transactions on Power Delivery 36(1)(2021) 484-487.
[32] S. Zhang, X. Zhu, B. Anduv, X. Jin, Z. Du, Fault detection and diagnosis for the screw chillers using multi-region XGBoost model, Science and Technology for the Built Environment 27(5)(2021) 608-623.
[33] F. Abramovich, V. Grinshtein, T. Levy, Multiclass classification by sparse multinomial logistic regression, IEEE Transactions on Information Theory 67(7)(2021) 4637-4646.
[34] R. Bhavani, V. Prakash, K. Chitra, A Sly Salvage of Semantic Web Content with Insistence of Low Precision and Low Recall, Wireless Personal Communications 117(4)(2021) 2757-2780.