IEEE Access
Multidisciplinary : Rapid Review : Open Access Journal

Received 12 April 2024, accepted 21 June 2024, date of publication 24 June 2024, date of current version 8 July 2024.

Digital Object Identifier 10.1109/ACCESS.2024.3418847

RESEARCH ARTICLE

Sentiment Analysis Based on Improved Transformer Model and Conditional Random Fields

LISHA YAO AND NI ZHENG

School of Big Data and Artificial Intelligence, Anhui Xinhua University, Hefei, Anhui 230088, China

Corresponding author: Ni Zheng (sciencejsj@163.com)

This work was supported in part by the Key Research Project of Natural Science in Universities of Anhui Province under Grant KJ2020A0782, in part by the Key Scientific Research Project of Anhui Provincial Research Preparation Plan in 2023 under Grant 2023AH051806 and Grant 2023AH051807, in part by Anhui Xinhua College Quality Engineering Project under Grant 2020sysxx01 and Grant 2020ylzyx02, in part by the National First-Class Professional Construction Project Software Engineering First-Class Professional Construction Project under Grant 2020ylzy01, in part by Anhui Province College Ideological and Political Work Ability Improvement Plan Project Young and Middle-Aged Backbone Team Construction Project under Grant sztsjh2019-8-39, in part by Anhui Xinhua University Level Research Project under Grant 2022zr003, and in part by Anhui Province Quality Engineering Construction Project under Grant 2022xsxx089 and Grant 2023bgyx040.

ABSTRACT With the rapid development of the Internet, people independently write comments with emotional characteristics on e-commerce platforms, which express consumers' emotional tendencies towards products or services from multiple perspectives. The sentiment analysis technology of product reviews has attracted more and more attention. In recent years, the Transformer model has performed well in the field of text sentiment analysis. However, in the process of text emotion classification, the Transformer basic model cannot be well obtained when the distance between words is relatively long, and there are problems such as a low accuracy rate and recall rate and too much time spent on the overall training of model construction. Using conditional random field (CRF) as a classifier can effectively solve the problem of too long text segmentation distance, reduce model training time, and increase the accuracy of text sentiment analysis. Therefore, this paper proposes a new method for text sentiment analysis combining the improved Transformer model and a conditional random field. First of all, this paper enhances the Transformer model by adapting the decoder to better suit the task of sentiment classification, and introduces an enhanced version of the Transformer model. It is then classified by combining long-short-term memory (LSTM) and CRF. The experimental results show that in the IMDB data set, the Transformer CRF model has higher accuracy, recall rate, and F1 value, reaching 85.51%, 83.77%, and 85.06%, respectively. Compared with other methods, the results of evaluation indicators further verify that the text method has better recognition performance and generalization ability, and at the same time, it has a better understanding of customers' emotions towards a specific aspect, provides users with accurate services, effectively improves user satisfaction, and has a high use value for enterprises' business decisions.

INDEX TERMS Transformer model, conditional random field, sentiment analysis, deep learning, LSTM, attention mechanism.

I. INTRODUCTION

The study of natural language processing includes the discipline of text sentiment analysis. It will detect, analyze and mine subjective texts containing users' views, preferences and feelings [1], [2]. People are starting to voice their arbitrary opinions about goods, news, subjects, events, and other things as the Internet develops further. People's emotional analysis has recently become a prominent topic on the Internet [3], [4]. Text sentiment analysis has very high

The associate editor coordinating the review of this manuscript and approving it for publication was Bing Li.

use value. E-commerce reviews help e-commerce companies to understand the needs of users, intuitively understand their attitudes towards products, and analyze these attitudes to help businesses improve product design and marketing strategies. News comments help new news media to understand users' preferences, promote new news personally, and help government agencies to monitor public opinion and avoid major public opinion events. Film community comments can help relevant creators understand the public's attitude towards the film creation, and at the same time let the audience know the word-of-mouth of the film and avoid minefields. Sentiment analysis is widely used in the business field, especially in customer satisfaction analysis, market trend prediction, marketing strategy optimization, brand image monitoring, product research and development reference, competitive intelligence analysis, etc. It can help enterprises better understand the market and consumer demand, optimize products and services, improve marketing effect, and enhance competitiveness. At the same time, enterprises can also monitor brand image and reputation through sentiment analysis, timely detect and deal with potential crises, and maintain brand image.

Natural language processing and deep learning have recently achieved a breakthrough in the realm of emotional analysis, but it is still facing a huge challenge. These challenges are mainly manifested in:

(1) The types of commodities are complex. Comments are made on different products in the website, and the focus of comments is different.

(2) Comment on the nonstandard data. The irregular grammar of language organization leads to difficulties in sentiment analysis based on grammar rules.

(3) Multi-semantic problems. The accuracy of word vectors obtained from multiple semantic times will be affected, which will bring inconvenience to subsequent processing.

In recent years, Transformer model has performed well in the field of text sentiment analysis. For the Transformer model, if the distance between two words is long, the original model can't be obtained well, and the Transformer basic model needs to consume a lot of resources to train the whole model, which makes its training input cost too high. The motivation of this paper is to improve the decoder structure of the Transformer model to make it more suitable for the task of sentiment analysis, in order to solve the problems of relatively low accuracy and recall in the process of text sentiment classification, and too long training time for building the model as a whole, and to ensure the correctness of the effect of text sentiment analysis.

The main contributions of this paper are as follows:

(1) Transformer model is a classic Seq2Seq model. The input and output of this model are word sequences, which is not suitable for emotion classification task. Therefore, this paper improves the Transformer model, changes the Transformer decoder to make it more suitable for emotion classification, and proposes an improved Transformer model.

(2) This paper proposes a new idea for text sentiment analysis by combining the improved Transformer model and conditional random field to propose Transformer-CRF model. The model uses the improved Transformer model to find out the correlation and semantic information of words in the text, and uses LSTM to further excavate the contextual semantic relations. Finally, the conditional random field is used as the classifier, which can effectively solve the problem of long text segmentation distance and reduce the training time of the model. Increase the accuracy of text sentiment analysis. The Transformer CRF model proposed in this paper can obtain more sufficient features, improve the training speed of the model, and achieve better training results.

(3) Experiments show that the model in this paper achieves good results in emotion analysis.

The organizational structure of this paper is as follows: section I briefly introduces the background significance of sentiment analysis; Section II introduces the related work in the field of text sentiment analysis and conditional random fields; Section III introduces the principle and structure of the model proposed in this paper; Section IV is experimental analysis; Section V is the summary and prospect of the main work of this paper.

II. RELATED WORK

A. TEXT SENTIMENT ANALYSIS

The three categories of text sentiment analysis in natural language processing are as follows: sentiment analysis in dictionaries, sentiment analysis in machine learning algorithms and methods for deep learning that analyze sentiment [5].

The essential components of dictionary-based text sentiment analysis are the extraction, analysis, and creation of sentiment terms and a sentiment dictionary, which serves as the foundation for evaluating text sentiment. Different word types are given varying weights in text classification. The accuracy of emotion vocabulary is a key factor in emotion analysis based on it [6]. The process of gathering words and phrases for emotional dictionaries typically involves a lot of physical labor on the part of experts. As a result, sentiment analysis based on sentiment dictionaries has significant drawbacks and is unable to produce results with high accuracy [7]. The benefit of using a dictionary is that a lot of emotive words may be accessed fast and effortlessly. Yet, each word stands alone, the dictionary-based approach ignores word connotations, and the emotional connotations of some words change depending on the article. Corpus-based techniques can address this shortcoming [8].

The typical machine learning approach for emotion analysis views emotion analysis as a classification process. The processes of text data processing that are more crucial than simple data cleaning include text vectorization and feature extraction. Some other issues with natural language processing have involved the study of the text vectorization problem. The selection of features and the quality of the corpus and dictionary are the main considerations in sentiment analysis using machine learning algorithms. For sentiment analysis, popular machine learning methods include Deep Forest (DF) [9], Support Vector Machine (SVM) [10], Naive Bayes (NB) [11], and others. The SVM algorithm has the highest accuracy, according to numerous research. However, these algorithms have strict corpus requirements and need to train a lot of corpora. The corpus, however, lacks a strong extensibility. In other disciplines, like commodities reviews, corpus applied to film reviews cannot produce satisfactory results.

Texts are divided between positive and negative texts using emotion classification. As a result, the study of words and sentences that convey emotion is given more consideration in the classification of emotions. The emotional categorization of Twitter was researched by Ren et al. [12]. Every tweet is viewed as a complete thought. Together with general characteristics, tags, emojis, and punctuation are employed in supervised learning algorithms as special features of twitter [13]. To advance their understanding of emotion analysis, many researchers employ supervised learning. Li and Huang [14] published for the first time the Stacking combination algorithm which classifies four emotional texts based on supervised learning. Combining different classification methods can achieve better classification results than the best base classification method. Li et al. [15] extracted the text sentiment abstract of supervised learning by extracting the features of multi-document text, sentiment features, pagerank features and comment quality features, and the rouge value was significantly higher than that of unsupervised learning. Although the experiment of unsupervised learning algorithm was relatively easy, it was difficult to accurately calculate the emotional fit between words and to determine the seed words, so the research on unsupervised learning algorithm was less and less. It has been recognized that the SO-PMI algorithm [16] can be used to forecast the emotional part-of-speech tendency of text.

Deep learning neural network-based models have seen considerable success in many domains recently. Some researchers report that these models have powerful ability to learn distributed knowledge representation and often give good results. A technique for effectively training a huge amount of data based on shallow neural networks called word2vec was released as open source by google in 2013 and has since gained interest from both the business and academic communities [17]. In 2015, Lei et al. [18] introduced a convolutional neural network model for text sentiment analysis and carried out tests using data sets from tasks involving news categorization and sentiment classification as well as standard tasks. The experimental findings support the convolutional neural network's applicability to text processing. Angeli [19] et al. used deep convolutional neural networks to analyze the texts in cancer pathology reports. The experiment proves that TextCNN model achieves good effect in clinical text classification. Ma et al. introduced a perceptron approach based on word vector and focused on the viability of Chinese word segmentation and word marking by deep learning [20]. Deep learning-based emotion analysis has significantly improved both the accuracy of emotion analysis and the vectorization of text. Recurrent neural network approach is more ideal for text sentiment analysis. LSTM model is a kind of neural network with recursive characteristics, and has obtained good emotion analysis results. Using the grammatical structure of sentences, sentiment analysis is only one of the numerous NLP tasks for which LSTM has been shown to be effective. Zhang et al. employed the Long and Short Memory Model (LSTM) to address the issue of text sentiment analysis following the creation of sentiment analysis [21]. Recurrent neural networks like LSTM have several unique characteristics. The LSTM model performs better than the recurrent neural network in many tasks. Change the proportion of its own loop by increasing the input, forgetting and output thresholds of LSTM. Single convolutional neural networks (CNNs) are unable to employ text context information to solve problems, and basic recurrent neural networks (RNNs) cannot resolve long-term dependencies. Guo et al. [22] suggested an approach for sentiment analysis of online comments that was based on CNN-BiLSTM. The problem with the sequence model was that it decoded the input text sequence into a vector of a specified length, and if the vector length was set too low, the input text information would be lost. This was because the emotion analysis job did not make full use of the linguistic and emotion resources already available. An attention-based bidirectional LSTM emotion categorization approach was put forth by Li et al. [23] to construct BiLSTM_ATT to complete the recognition of relation and tail entity in sentences. Mewada et al. [24] proposed SA-BERT, a bidirectional encoder representation of a transformer, to address the issue of synthetic concern. Furthermore, they utilized XGBoost, an extreme gradient boosting classifier, for accurate emotional polarity classification within a review dataset. Yang et al. [25] proposed an aspect-level emotion classification model based on graph neural networks, utilizing the graph convolutional network to enhance node representation and enable the model to fully capture global semantic and syntactic structure information of sentences. He et al. [26] introduced a bidirectional LSTM with a trapezoidal structure, which demonstrated similar performance as the normal structure model in experiments but with fewer parameters. Zhang et al. [27] proposed the SA-Model, a poetic emotion analysis model that integrated multiple encoders to eliminate text features at various levels, thereby augmenting textual feature information, enhancing the accuracy of text semantics, and improving the learning and generalization capabilities of the model. Rahman et al. [28] introduced a multi-layer classification model for supervised machine learning on social media texts. Mohamed et al. [29] employed LexDeep, a hybrid emotion analysis technology, to analyze emerging social trends on Twitter. The findings demonstrated that all simple embedded LexDeep models outperform support vector machines, affirming the superiority of the proposed LexDeep model over classical machine learning classifiers in domain-specific sentiment analysis tasks. The proposed model by Xu et al. [30] leveraged BERT and a hypergraph dual attention mechanism for text emotion analysis. By exploiting the powerful representation learning capability of the pre-trained BERT model, dynamic feature extraction was performed on emotional text, followed by aggregation of the aforementioned correlation information through the dual graph attention mechanism. This approach significantly enhanced sentiment analysis accuracy for short Chinese online texts. Liu et al. [31] proposed an emotion analysis method that leveraged the dynamic characteristics of words and sentences, as well as self-attention. This approach dynamically encoded comments using a pre-training model, integrated sentence fusion features based on a feature recombination method utilizing the self-attention mechanism, and optimizes weight parameters to reduce computational complexity and training time.

In 2017, Google team proposed a new deep learning model, namely Transformer model. However, there are some problems, such as relatively low precision and recall rate, and too long training time to build the model. In a 2020 paper on techniques for fine-grained sentiment analysis based on CRF and ATAE-LSTM, Xue and Liu [32] and coauthors investigated the associated CRF methodologies. However, there is a problem of extracting only explicit attribute words but not implicit attribute words. The sentiment analysis model proposed by Gao and Huang [33] employed the TF-IDF (Term frequency-inverse Document Frequency) approach and a multi-head attention Transformer model. The TF-IDF algorithm is utilized to initially screen words in the text, identifying those with significant emotional tendencies. Subsequently, the encoder of the multi-head attention Transformer model is employed for extracting features that capture crucial semantic information within the text, thereby enhancing both semantic analysis and generalization capabilities of the model. Based on the Transformer model, Zhang and Zhou [34] designed a Transformer-Encoder-GRU (T-E-GRU) model for sentiment analysis of product reviews, which solves the problems of information loss and poor culture in existing text sentiment classification models. The training difficulty and time of the model are reduced, and the overall performance of the model is improved. Koshy and Elango [35] conducted a study where they proposed a disaster response system that employs a bidirectional attention model based on transformers for multi-modal tweet classification. The system integrates diverse models including fine-tuned RoBERTa for text analysis, Vision Transformer for image processing, BiLSTM, and an attention mechanism.

The pre-trained transformer model can effectively capture comprehensive linguistic knowledge and contextual information through extensive corpus learning, thereby enhancing its ability to comprehend the interrelationships between sentences and texts, facilitating the extraction of essential features for emotion classification tasks. Firstly, owing to the utilization of a substantial amount of data during the pre-training phase for parameter optimization, the pre-trained model typically exhibits robust generalization capabilities. Consequently, it maintains satisfactory performance even when applied to novel domains or tasks, thus reducing reliance on extensive annotated data. Moreover, in transfer learning scenarios, we can employ transformer models that have undergone pre-training on relevant tasks as underlying networks and subsequently fine-tune their parameters accordingly. This approach not only expedites model convergence and diminishes the requirement for labeled data but also leverages knowledge acquired from other tasks to enhance emotion classification performance.

Therefore, this paper takes Transformer model as the base model. However, because of the multiple meanings of the text, it still can not effectively solve the problems of poor performance, low accuracy, and unable to accurately understand the semantic information expressed in the text. This paper makes improvements based on Transformer model, optimizes its encoder and decoder, and makes use of the vector after the fusion of all words in the sentence to better represent the whole sentence and accurately understand the semantic information of the text.

B. CONDITIONAL RANDOM FIELD

The most significant area of application for CRF, where text and audio data are processed, is natural language processing. CRF performs well in text segmentation and tagging, information identification and extraction, text classification and matching, etc. because it is primarily employed to address the issue of sequence tagging.

The most current applications of CRF have been expanded in recent years to include legal text abstract extraction, dialogue behavior recognition (DAR), user activity recognition, and other domains. DAR can replicate the speaker's purpose, making it a fundamental step in helping computers comprehend spoken language dialogue. The typical machine learning models are different in how they adapt to DAR tasks, nevertheless, because of the varied features, the statistical reliance between conversation behavior tags, and the complex link between characteristics and dialogue behavior tags. To address this issue, Zhou et al. [36] suggested a novel model that combines heterogeneous deep neural network (HDNN) with CRF. The HDNN model is employed in this framework to handle heterogeneous data, and there is a maximum statistical dependence between conversational behavior tags and CRF. The model has a high degree of categorization accuracy, according to experiments. Understanding people's intentions and interests also requires being able to identify activities from their behavior. Yet, most of the time, users' behaviors are only partially seen or recorded, which makes replicating users' activation relationships extremely difficult. Based on this, Wen et al. [37] coupled context information with the internal properties of the instance to use PRM-CRF to extract user actions from the behavior flow. This model has the advantage of integrating existing domain knowledge through the use of a regularization term, which lessens the impact of missing tags. In order to extract key sentences, it is believed that the text segmentation of judicial decisions is ideal for CRF. It is useful for identifying vocabulary trends and usage patterns.

CRF was initially employed to address fundamental issues, such Chinese word segmentation (and POS marking), and it did so successfully. At present, based on the demand of large-scale document processing and deep understanding of the model, CRF is widely used in subtitle segmentation, text marking and other issues. The essential nature of POS tagging tasks was demonstrated by Zhang et al. [38] who evaluated their newly suggested end-to-end neural CRF automatic encoder (NCRF-AE) model using POS tagging tasks in eight distinct languages. Additionally, the model may make use of the hidden data information in both big and small amounts of unlabeled and labeled data.

The main goal of CRF's use in natural language processing is to extract and separate information from text or audio. Due to the popularity of social platforms and Q&A communities, many people today use the Internet to communicate with one another and learn new things. As a result, there are many different types of text natural language that need to be efficiently handled to aid in decision-making and suggestion. CRF helps users to intelligently analyze users' emotions and behaviors by effectively processing these texts.

Therefore, this paper introduces a CRF classifier, combines the improved Transformer model and LSTM structure, and proposes a new Transformer-CRF model to solve the sentiment analysis task. The model combines the powerful presentation capabilities of Transformer with the contextual information capture capabilities of CRF to enable more accurate sentiment analysis.

Compared with traditional sentiment analysis models, the Transformer-CRF model has the following advantages:

1. Powerful presentation capabilities: The improved Transformer model is better able to capture contextual information in input sequences through a self-attention mechanism to provide more accurate classification predictions. By improving the traditional Transformer model in this paper, it can better adapt to the needs of emotion analysis tasks and provide richer contextual information and more accurate emotion expression.

2. Efficient training and reasoning: The improved Transformer model has faster training and reasoning due to its attention mechanism and feedforward neural network structure.

3. Capture long-term dependencies in text: Use the improved Transformer in conjunction with the LSTM model to capture long-term dependencies in text. In addition, you can also consider using attention mechanisms to highlight important parts of the text, thereby improving the performance of sentiment analysis.

4. Better feature extraction capability: the improved Transformer model can better extract features from text and pass them to the CRF layer for category prediction. CRF can provide stronger label sequence constraints, help models better understand context information and improve the accuracy of label predictions.

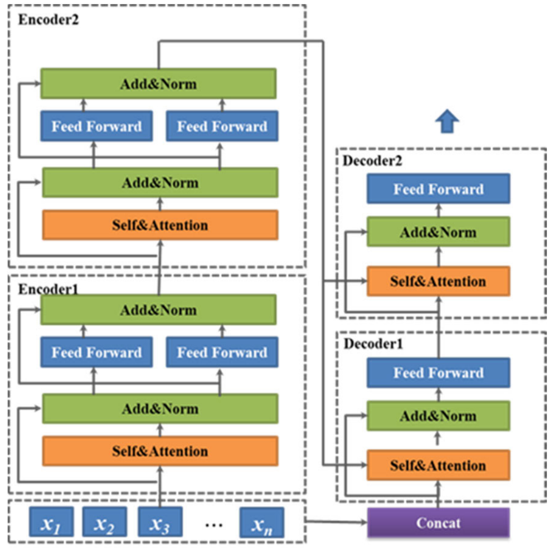


FIGURE 1. Transformer model structure diagram.

III. METHODOLOGY

A. TRANSFORMER MODEL

Transformer model is a new modeling method of sequence information. This model still uses the classical Encoder-Decoder coding structure. The difference is that the Attention algorithm mechanism is used in the Transformer model, which has the advantage of not being affected by the relative distance between words, successfully addressing the issue of long-distance information loss and enhancing the model's generalizability. FIGURE 1 depicts the internal organization of the Transformer model.

Encoder and Decoder are included in the Transformer model [39]. The encoder has layers for a feed-forward neural network and a multi-head attention layer. The self-feedforward neural network layer has two layers and is fully connected. Relu performs the first layer's activation, the second layer does not employ the activation function. Formula (1) displays the Feed Forward calculation formula.

$$F = \max(0, X \cdot W_1 + b_1) \cdot W_2 + b_2 \tag{1}$$

W_1 and W_2 are the network weight matrices, and X is the input. b_1 and b_2 are the offset. The self-attention layer employs the multi-head self-attention mechanism. As a key feature of the Transformer model, the Self-Attention mechanism can establish the association of the words in the current position and the words related to the context, and calculate the relevance degree of the context. The application method is as follows: Firstly, the input vector X is multiplied by three randomly initialized quality matrices W^Q, W^K and W^V to calculate three new vector query vectors Q, key vector K and value vector V. Formula (2)-(4) displays the Feed

Forward calculation formula.

$$Q = X \cdot W^Q \tag{2}$$
$$K = X \cdot W^K \tag{3}$$
$$V = X \cdot W^V \tag{4}$$

Q vectors encode active words. It should be noted that other K vectors in the text have searchable active word information. V represents the actual content of the current word. To query every candidate position, a Q vector is employed. The dot product of the Q vector and the K vector for each candidate position is the query process. The dot product results are weighted to their respective V vectors after passing through the softmax function, and the final self-attention score value of self-attention results is obtained by summing up. When coding the words in a sentence, focus on other parts of the sentence.

Here is the precise procedure:

(1) Determine how closely the Q and K vectors resemble one another. As shown in formula (5):

$$f(Q, K_i) = Q \cdot K_i \tag{5}$$

(2) Calculate the weight coefficient: perform softmax operation on the similarity in formula (6).

$$\alpha_i = f(Q, K_i) \Big/ \sum\nolimits_{j=1}^{n} e^{f(Q,K_i)} \tag{6}$$

(3) Calculate attention vector: sum all V vectors with weight coefficient. As shown in formula (7):

$$Attention(Q, K, V) = \sum_{i=1}^{n} \alpha_i \cdot V_i \tag{7}$$

(4) Scale the attention vector to improve the convergence speed. As shown in formula (8):

$$Attention\,(Q, K, V) = \frac{Attention\,(Q, K, V)}{\sqrt{d}} \tag{8}$$

The purpose of dividing by $\sqrt{d}$ is to avoid the calculation result being too big to enter the saturation interval of softmax function, and d is the dimension of Q and K vectors.

Multi-head self-attention integrates *num_heads* different self-attention, and the value of *num_heads* in this paper is 8.

In addition, both Feed Forward and Multi-Head Attention adopt residual connection (Add) and batch standardization (Norm) to boost the network's capacity for generalization and quicken convergence.

Decoder is similar to Encoder, except that it contains two self-attention layers. The first attention layer uses Masked operation to mask future input in the derivation process. The second attention layer is similar to feedforward neural network and Encoder, the only difference is that the input of Decoder not only contains the output of the previous layer, but also increases the output of Encoder.

B. CRF MODEL

The CRF model has become more crucial as artificial intelligence has advanced. Researchers have taken notice of CRF, a significant class of machine learning model based on the Maximum Entropy Markov Model (MEMM). As a matter of fact, in the past few years, there have been many researches on CRF, some of which focus on the invention of applications, while others are committed to the enhancement of fundamental models. The advancement of CRF in the aforementioned two areas has sped up the model's spread across a variety of felts in addition to deepening our grasp of it.

CRF is a probabilistic graphical model, which combines probability and graph theory. It provides a global probabilistic modeling framework based on local functions and represents the distribution on labels. CRF is a discriminant model, which considers the connections across observational sequences with identified labels, in contrast to MRF, a generation model that aims to represent the probability distribution. The discriminatory nature of CRF allows any interaction between the results to be observed without independent assumptions.

The parameterized form (Formula 9) is one of the various ways that CRF can be stated. In this form, the local recognition functions of t_k and s_j depend on the position, and t_k is that the recognition function represented by the edge of the recognition function moves between each node in turn, depending on the roles held in the past and present; The attributes of a node are represented by the s_j status identifier, which is based on the node's current location. The characteristic functions t_k and s_j and the related weights λ_k and μ_j fully govern the model's output, and the parameter form of CRF intuitively displays the factors and parameters injected into the output of the model.

$$\begin{aligned} p\,(y\,|\,x) = \frac{1}{z(x)} exp(&\sum\nolimits_{i,k} \lambda_k t_k\,(y_{i-1}, y_i, x, i) \\ &+ \sum\nolimits_{i,j} \mu_j s_j(y_i, x, i)) \end{aligned} \tag{9}$$

$$\begin{aligned} z(x) = \sum\nolimits_{y} exp(&\sum\nolimits_{i,k} \lambda_k t_k\,(y_{i-1}, y_i, x, i) \\ &+ \sum\nolimits_{i,j} \mu_j s_j(y_i, x, i)) \end{aligned} \tag{10}$$

The process of CRF modeling faces the same three fundamental issues as HMM (Hidden Markov Model): parameter estimation, model reasoning, and probability calculation.

C. IMPROVED TRANSFORMER MODEL

In this paper, firstly, the data set of comment text is preprocessed, and the comment sentiment is analyzed by Transformer-CRF model. Transformer-CRF model uses Transformer model to extract text features, and then injects the text into CRF classifier for text sentiment analysis. FIGURE 2 displays the flowchart's overall structure.

Transformer model is a classic Seq2Seq model. Machine translation jobs make good use of the multi-head self-attention mechanism. The model's input and output are all word sequences, which are not suitable for emotional

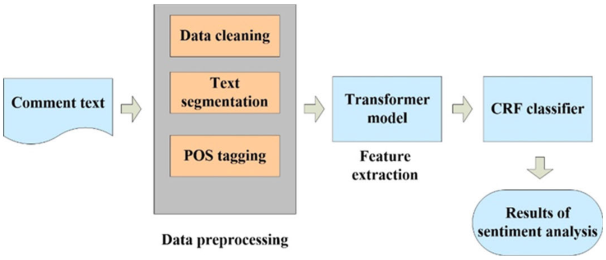


FIGURE 2. Overall flow chart.

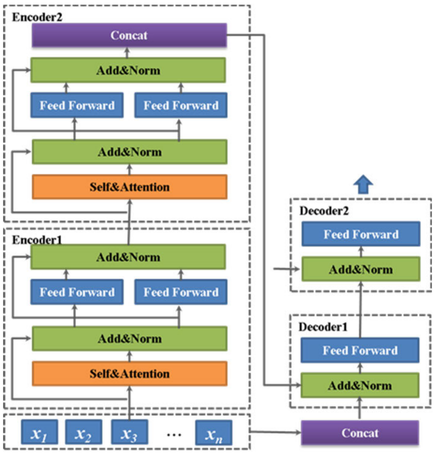


FIGURE 3. Improved transformer model structure diagram.

classification tasks. Therefore, this paper improves the model and changes the decoder of Transformer to make it more suitable for emotion classification. The improved Transformer model is shown in FIGURE 3, and its work mainly includes:

(1) The Transformer model's encoder module is capable of extracting text's semantic information. However, multi-attention is mainly about feature extraction of a single word vector, lacking the overall semantic information of the sentence. Add a Concat module to this text to splice all the word vectors and construct the global semantic information of the sentence. For example, the outputs of the Encoder module are $e_1, e_2, \cdots, e_n$, where n represents the number of word segments in the text and the word vector's dimension is indicated by the letter d. The sentence vector after Concat is $E = \{e_1, e_2, \cdots, e_n\} \in R^{n\times d}$.

(2) Decoder module of Transformer model decodes the information of Encoder. It is inappropriate for the goal of classifying emotions. Therefore, this paper retains Feed Forward layer and Add&Norm layer in Decoder module of Transformer model, and removes Multi-Head Attention layer of Decoder module. The Feed Forward layer is used to further quadratic map the sentence vector E. The Add&Norm layer enables stacked Decoder modules to solve multi-layer network training problems.

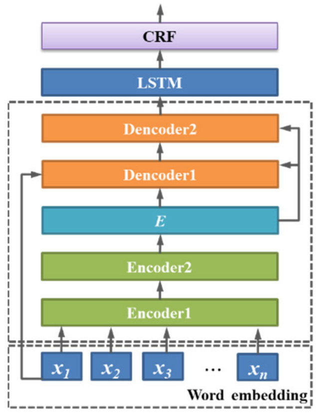


FIGURE 4. Transformer-CRF model calculation flow chart.

D. WORKING MECHANISM OF TRANSFORMER-CRF MODEL

This research suggests the Transformer-CRF model to address the issue of sentiment analysis based on the enhanced Transformer model. This module is divided into five main modules, and FIGURE 4 displays the calculating flow frame for it.

(1)Word embedding module. Using Word2Vec tool, training in pre-training corpus, quantizing words, also known as word embedding.

Words are the basic unit of text analysis. First, the input text is preprocessed, and then the words need to be vectorized. Represented by the word embedding layer as a vector as shown in Formula (11):

$$X = [x_1, x_2, x_3, \cdots, x_n]^T \in R^{n\times d} \tag{11}$$

where n is the number of words in the text that have been segmented, and d is the size of the word vector, and $x_i(i = 1, 2, 3, \cdots, n)$ reflects the ith word's place in the word vector.

(2)Encoders module. The word vector is sent to the Encoders module of Transformer model to identify the key elements of the text and obtain the corresponding coding vector. Encoders modules can be superimposed and coded many times. In this paper, two Encoders modules are used for superposition.

In addition to the semantic information of words, the position information of words is equally important, which represents the text's overall characteristics. So the first step in the Transformer layer is position embedding. The position vector is introduced into the construction of Transformer model, which is added to the word vector to better express the word's semantic information of the word combined with the position information. The model mainly uses sine function and cosine function to learn the position of the words in a sentence quickly. Formula for calculating position vector is shown in Equations (12) and (13):

$$PE_{(pos,2i)} = sin(pos/10000^{2i/d}) \tag{12}$$

$$PE_{(pos,2i+1)} = cos(pos/10000^{2i/d}) \tag{13}$$

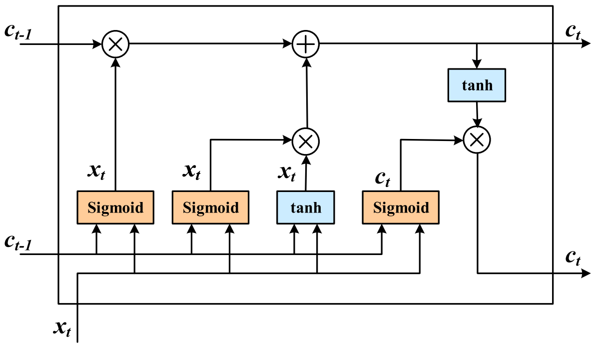


FIGURE 5. Spatial attention.

Among them, *pos* represents the amount of words in each location in the phrase, and *i* reflects the word embedding vector's dimension index. Finally, the word vector representation of the embedded position vector is shown in Formula (14):

$$RE_i = WE_i + PE_i \tag{14}$$

where WE_i represents the sentence's ith word's word vector, PE_i represents the position vector of the ith word in the sentence, and RE_i represents the vector encoded by Encoder.

The word vectors added with position information are respectively input into various self-attention modules to acquire *num_heads* (the value of *num_heads* in this paper is 8) weighting matrices Z_i, $i \in \{1, 2, \cdots, 8\}$, and then the total text matrix Z is obtained by column splicing. The calculation is shown in Formula (15):

$$Z = (Z_1, Z_2, \cdots, Z_8) \tag{15}$$

(3) Decoders module. The input of Decoder module includes two parts: one is the sentence vector E encoded by Encoder, and the second part is the output E' of the previous layer of Decoder. The E' output by the first layer Decoder is the spliced vector $X' = \{x_1, x_2, x_3, \cdots, x_n\} \in R^{n\times d}$ of the original word vector. Connect the coding vectors and input them into Decoder module for decoding. After many feature mappings, the semantic vector of the whole sentence is obtained. Decoder modules can also be superimposed. In this paper, two Decoder modules are used for superposition.

(4) LSTM layer. When the LSTM method in the Transformer model classifies text emotions concretely, the LSTM neural network combines the core words in order, and each word can get the data information of the previous word, and finally get the sentence vector covering the semantic relationship of the whole sentence. FIGURE 5 is the internal structure diagram of LSTM.

Here is the output matrix of the Decoders module, where X is the input data, $X \in R^{n\times d}$. The LSTM's cell output is h, and c is the value of the memory cell in LSTM. The current input data X_t, output value h_{t-1} and unit value will jointly affect the current h value. Formulas (16)–(17) display the input layer calculation formula:

$$i_t = \sigma(W_i * [h_{t-1}, X_t] + b_i) \tag{16}$$

$$c_t = tanh(W_c * [h_{t-1}, X_t] + b_c) \tag{17}$$

The formula of the output layer is shown in Formula (18)-(19):

$$o_t = \sigma(W_o * [h_{t-1}, X_t] + b_o) \tag{18}$$

$$h_t = o_t * tanh(c_t) \tag{19}$$

(5) CRF layer: The global optimal result is predicted by the CRF layer using the output result of the LSTM layer as the input data. The maximum likelihood label is obtained after classification by CRF classifier, and the output layer outputs the label corresponding to the final classification result.

The CRF model in this model is variant high-order CRF, which can decompose high-order text attributes into low-order attribute dependencies implied in singles and pairs, and construct nonparametric high-order potential to quickly capture high-order text attribute dependencies. High-order CRF model can capture the content of context message well, and calculate the conditional maximum probability of all possible state sequences by inputting the observation sequence, that is, the output result $X = \{x_1, x_2, \cdots, x_n\}$ of LSTM layer, and take the maximum probability $Y = \{y_1, y_2, \cdots, y_n\}$ as the output state of the sequence. The formula demonstrates the computation process (20).

$$p(y\,|\,x) = \frac{1}{z(x)} \prod\nolimits_{t=1}^{T} exp \sum\nolimits_{k=1}^{K} \theta_k f_k\,(y_t, y_{t-1}, x_t) \tag{20}$$

Among them,

$$z(x) = \prod\nolimits_{t=1}^{T} \exp\left\{\sum_{k=1}^{K} \theta_k f_k\,(y_t, y_{t-1}, x_t)\right\} \tag{21}$$

IV. EXPERIMENT PREPARATION

A. DATA SET

The data set in this paper uses four English-published data sets: IMBD movie review data set [40], MR movie review data set [41], Yelp Polarity merchant review data set [42] and Amazon Polarity commodity review data set [41]. Unbalanced category distribution will affect the accuracy of model classification, so the imbalanced data set is undersampled first to balance the data and reduce the impact on model training. The ratio of the training set to the test set is 8:2. The 50-fold crossover method is used for many experiments to improve the robustness of the model. The information tables corresponding to the above four data sets are shown in Table 1.

TABLE 1. Data set information table.

Data set	IMBD	MR	Yelp P.	Amz. P.
Amount of positive data	25,000	5331	299,000	2000000
Amount of negative data	25,000	5331	299,000	2000000
Average length	232	20	135	74
category	2	2	2	2

B. DATA PREPROCESSING

To assess how well these three models for text sentiment analysis perform, this study evaluates the effectiveness of the models using three different metrics, namely Precision (P), recall rate (R) and F1 value.

- TP: number of supportive comments in the appropriate category.
- FP: number of positive comments in misclassification.
- TN: number of comments classified as negatives.
- FN: number of negative comments in misclassification.
- P: the ratio of projected positive case data to expected positive case data that turned out to be accurate.

$$P = TP/(TP + FP) \tag{22}$$

- R: the ratio of correctly anticipated to actually observed positive case data.

$$R = TP/(TP + FN) \tag{23}$$

- F1 value (F1 score): harmonic average.

$$F = 2/(1/P + 1/R) = 2 * P * R/(P + R) \tag{24}$$

C. EXPERIMENTAL PARAMETERS

Firstly, this paper uses IMBD data set to do basic experiments. By analyzing the text data set, it is found that the sentence length of 280 words can cover most of the data, so the sentence length input by the model is vectorized to 280 words, and 80% of the data set in the sub-corpus is used as the training set and 20% as the test set.

The combination of Transformer and conditional random field model parameters has a significant impact on the accuracy and durability of the experimental model. Therefore, LSTM+Attention method is used to continuously optimize and update the experimental parameters. It significantly enhances the model's universality. In addition, in order to prevent the model from being overfit, the Dropout parameter is also introduced into the model. After optimizing the experimental parameters for many times, Table 2 displays the major characteristics of the Transformer-CRF model.

TABLE 2. Setting of main experimental parameters.

Parameter	Value	Explain
n	200	sentence length
d	512	word vector dimension
num_heads	8	head number of multi-head attention mechanism
LSTM_hidden	128	LSTM hidden layer unit number
batch_size	32	minimum batch size
hidden_units	512	hidden layer dimension
dropout	0.1	discard rate, prevent over-fitting
learning_rate	0.0001	learning rate
optimizer	Adam	optimizer
batch size	128	batch size
epoch	3000	training times

The sentence length is 280. The word vector has a 512-dimensional size. Four levels make up the multi-head attention mechanism, in which the number of num_heads is 8. In the LSTM layer, there are 128 hidden layer units. Dropout is set to 0.1, and the learning rate is 0.0001. The Adam optimizer is employed, with a batch size of 128, and the model is trained for 3000 epochs.

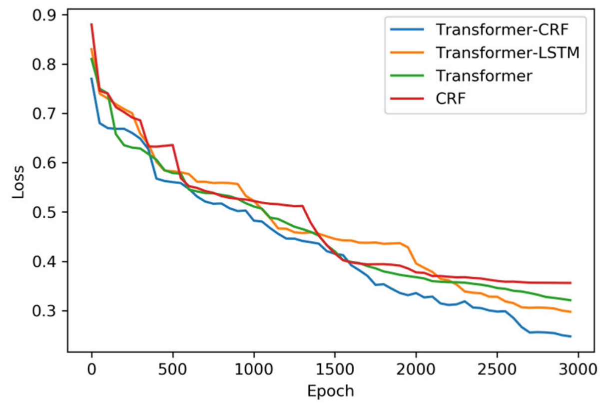


FIGURE 6. Comparison chart of loss value.

V. EXPERIMENTAL RESULT ANALYSIS

A. ABLATION EXPERIMENTS

In order to verify the effectiveness of this model, the following three control models are set for the Transformer-CRF model proposed in this paper for comparison while maintaining the experimental parameters:

- Transformer model: The original Transformer Model.
- CRF model: a variant of the higher-order CRF model with the same structure as the CRF layer in the Transformer CRF model proposed in this paper.
- Transformer-LSRM model: The original Transformer model is combined with the LSTM layer.

The cross entropy loss function is employed in this study to determine the model's loss value. The experiment demonstrates that as the model's iteration times are increased during the training process, the loss value gradually reduces, as illustrated in FIGURE 6, and the loss value gradually approaches to stability after 1500 steps. The accuracy and loss value of the 4 models are contrasted in this study. FIGUREs 6 and 7 respectively illustrate this.

It is not difficult to see that with the increase of the number of iterations, the loss decreases and the accuracy rate increases. The loss value and accuracy rate corresponding to the same number of iterations are different. The accuracy rate of the model gradually rises as the loss value decreases. In this paper, the Transformer-CRF model converges faster at the loss value, and the accuracy is improved.

The comparison results of the four models are shown in Table 3.

In order to see the comparison results more directly, the visualization results obtained in this paper according to the data in the above table are shown in FIGURE 8.

The experimental results show that the Transformer model and CRF model are relatively low, indicating that the accuracy of word segmentation using CRF model to extract

TABLE 3. Ablation experiment comparison.

Model	Precision%	Recall%	F1-score%
Transformer	81.78	81.79	82.02
CRF	82.23	81.98	82.11
Transformer-LSTM	83.16	82.29	83.01
Transformer-CRF	**85.51**	**83.77**	**85.06**

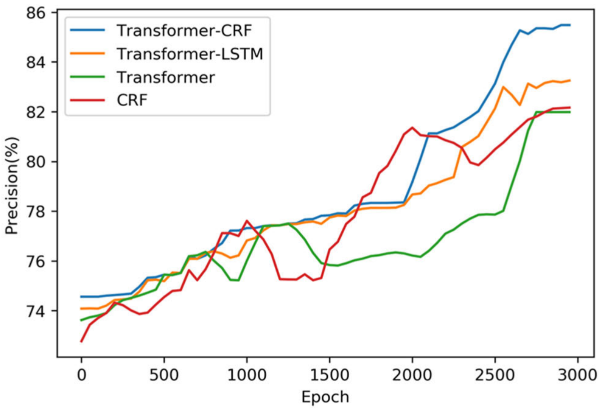


FIGURE 7. Comparison chart of precision.

emotional feature fragments and directly extract emotional tendency from comments is low. The Transformer model is used to extract emotional feature fragments and then classify them, which is insensitive to segmentation errors. The word features used in this paper only include the words themselves, but do not use lexical, syntactic, semantic features. When classifying the emotions expressed in online comments with words that are segmented incorrectly, the words in the comments will be segmented in the same way regardless of the segmentation method used. In other words, the weight of the words in the online review is the same as the weight of the same words in the training model, which does not change the accuracy of the segmentation. Therefore, the method is not sensitive to the accuracy of word segmentation. Compared with the basic Transformer model, the Precision, Recall and F1-score of Transformer-LSTM model are improved by 1.38%, 0.5% and 0.99% respectively. This shows that combining the parallel computing power of Transformer with the sequential processing power of LSTM can improve the efficiency and performance of the model. Compared with the Transformer-LSTM model, the Precision, Recall and F1-score of the Transformer-CRF model proposed in this paper are respectively improved by 2.35%, 1.48% and 2.05%. This shows that the improved Transformer model proposed in this paper can Concat all word vectors of words and carry out secondary mapping at the Decoder stage, which can make the semantics contained in feature vectors more comprehensive. At the same time, the introduction of CRF model can improve the effect of each model to a certain extent, improve the accuracy rate, recall rate and F1 value. When the improved Transformer model is combined with the CRF model, the emotion effect of text is significantly increased, and the F1 value is much higher than other models, which makes up for the error of Transformer model in text emotion analysis. In addition, this paper also compares the time of text sentiment analysis of the four models, and the comparison shows that the time benefit of the text sentiment analysis model of the Transformer CRF model proposed in this paper is significantly higher than that of the other three models. The comparison effect diagram is shown in FIGURE 9.

B. COMPARATIVE EXPERIMENT OF CRF CORRELATION ALGORITHM

In 2020, Xue and Liu [32] studied the association rules, ATAE-LSTM, LSTM and other related methods in CRF model using CRF and ATAE-LSTM, a fine-grained sentiment analysis technique. In this paper, the variant high-order CRF method was used for text classification. Here, the P, R, F1 obtained by each algorithm of CRF model for the same data set are compared. In Table 4, comparison outcomes are displayed. From the results in Table 4, it can be seen that the Precision, Recall and F1 values of the variant high-order CRF method adopted in this paper on the IMBD data set reached 85.51%, 83.77% and 85.06%. Compared with the ATAE-LSTM method in [32], the Precision, Recall and F1 values are improved by 6.61%, 2.51% and 6.41% respectively. Experiments verify the superiority of CRF method in this paper.

TABLE 4. Performance comparison table of CRF related algorithms.

Algorithm	Precision%	Recall%	F1%
Association rule	48.65	14.89	21.76
ATAE-LSTM	78.90	81.26	78.65
LSTM	71.84	79.70	73.76
Algorithm in this paper	**85.51**	**83.77**	**85.06**

C. COMPREHENSIVE CONTRAST EXPERIMENT

In order to further verify the validity, superiority and generalization ability of this model, the following models are used as control groups and compared with the other three data sets (MR, Yelp Polarity and Amazon Polarity). The comparison models are TextCNN model [19], LSTM model [21], CNN-BiLSTM model [22], BiLSTM_ATT model [23] and T-E-GRU [34]. The sentence length of MR data set is n = 30, Yelp P. data set is n = 170, and AMZ P. data set is n=100. Other experimental parameters are the same as those of IMDB data set described in Section IV-C. The comparison results are shown in Table 5 -7.

TABLE 5. Precision (%) comparison.

Algorithm	MR	Yelp P.	Amz. P.
TextCNN [19]	74.36	90.25	88.79
LSTM [21]	76.85	91.67	90.88
CNN-BiLSTM [22]	81.27	92.11	92.76
BiLSTM_ATT [23]	81.89	92.19	93.10
T-E-GRU [34]	82.26	93.12	94.17
Algorithm in this paper	**83.28**	**93.88**	**94.23**

As can be seen from Table 5 -7, TextCNN is weaker than other networks. The performance of LSTM-based networks

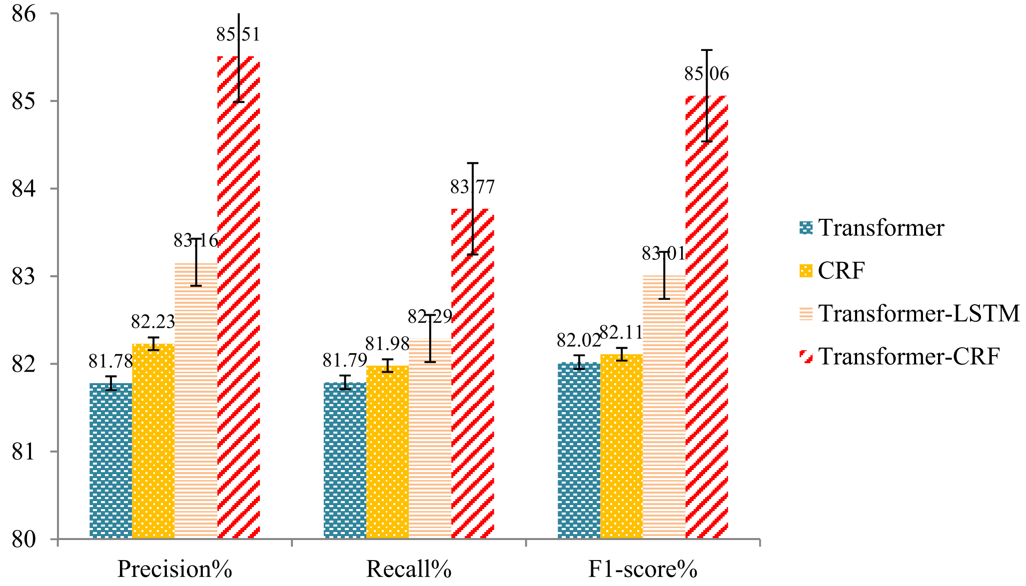


FIGURE 8. Comparison of ablation experiments.

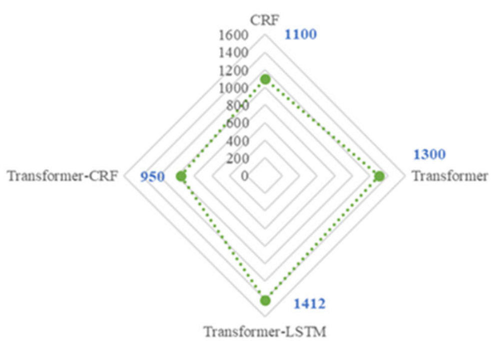


FIGURE 9. Comparison chart of model consumption time.

TABLE 6. Recall (%) comparison.

Algorithm	MR	Yelp P.	Amz. P.
TextCNN [19]	79.66	89.77	87.19
LSTM [21]	80.52	90.08	90.72
CNN-BiLSTM [22]	81.18	90.36	91.75
BiLSTM_ATT [23]	81.73	91.25	92.81
T-E-GRU [34]	82.28	92.16	93.18
Algorithm in this paper	**82.34**	**93.77**	**93.37**

TABLE 7. F1 -score (%) comparison.

Algorithm	MR	Yelp P.	Amz. P.
TextCNN [19]	75.81	90.52	90.12
LSTM [21]	78.15	91.18	91.72
CNN-BiLSTM [22]	80.47	90.36	91.75
BiLSTM_ATT [23]	81.56	92.25	92.11
T-E-GRU [34]	81.78	92.71	93.43
Algorithm in this paper	**82.03**	**93.28**	**93.56**

(LSTM, CNN-BiLSTM, BiLSTM_ATT) is better than that of TextCNN model, which shows that the memory ability and contextual information reasoning ability of recurrent neural networks show strong advantages in text sentiment analysis tasks. The performance of T-E-GRU model is better than that of LSTM-based networks (LSTM, CNN-BiLSTM, BiLSTM_ATT), because Tranformer model realizes parallel computing through self-attention mechanism, and it does not need to consider the distance between any two positions in the sequence when calculating the association, so it can better capture long-distance dependence. The Transformer-CRF model used in this paper has the best effect, and the above three indicators have been significantly improved in the three data sets, reaching the best in the comparison model. Experiments show that the model has strong generalization ability.

VI. DISCUSSION

The sentiment analysis is applied in this study by enhancing the Transformer model and proposing an improved version of it, which is further classified through a combination of Long Short-Term Memory (LSTM) and conditional random field (CRF). Consequently, a novel approach for text sentiment analysis that integrates the enhanced Transformer model with conditional random field is introduced in this paper. The key assumptions underlying this method are as follows:

1.The Transformer model effectively captures comprehensive language knowledge and context information by leveraging a wide corpus for learning. It employs a self-attention mechanism to compute the interdependencies among words, encoding the input sequence based on these associations. This approach enables the Transformer model to consider simultaneous interactions between all words in a sentence, facilitating a deeper understanding of sentence and text interrelationships. Consequently, this study enhances the conventional Transformer model by modifying its decoder specifically for emotion classification purposes and presents an improved version of the Transformer model.

2.CRF (Conditional Random Fields) enhances prediction accuracy by taking into account the dependencies between labels. In sentiment analysis tasks, labels often represent different emotional categories such as positive, negative and neutral. Traditional classification methods only consider the probability of each label appearing independently and ignore possible correlations between them. However, in practical applications, emotional expression is often complex and diverse with a certain degree of dependence among individual labels. By modeling these dependencies using CRF technology, we can better understand how individual labels are connected in emotional expression and how they affect each other during prediction process. This approach leads to more detailed and accurate sentiment analysis results. For instance, when words like "very" or "like" are combined to indicate strong positive emotions in a sentence, adding a negation prefix changes it to negative emotion.

3.This paper employs an enhanced Transformer model to capture contextual information, thereby augmenting the comprehension of the interplay between sentences and texts. Conversely, CRF enhances prediction outcomes by considering label dependencies, leading to more intricate and precise sentiment analysis outputs.

The potential limitations and challenges inherent in this approach are as follows:

1.Label inconsistency: Variations in emotion classification among different annotators can lead to inconsistent labeling of the same text, thereby impacting the accuracy of model training and evaluation results.

2.Model overfitting: The combination of multiple deep learning components, such as enhanced Transformer, LSTM, and CRF, into a complex model may result in overfitting if there is insufficient diverse training data and inadequate regularization measures are employed. Consequently, the performance of the complex model on the test set might be compromised.

3.Limited explanatory capacity: The prediction results lack detailed explanations or interpretability, which hinders users from understanding specific reasons behind certain emotion classifications. As a consequence, user trust in the system is diminished.

VII. CONCLUSION

This paper proposes a text sentiment analysis model combining improved Transformer and conditional random field, and proves the feasibility and effectiveness of Transformer-CRF model for text sentiment analysis through experiments. The improvement of Transformer model and the introduction of LSTM and CRF models can improve the original effect of this model to a certain extent, and significantly improve the accuracy rate, recall rate and F1 value of text sentiment analysis, making up for the shortcomings of Transformer model in text sentiment analysis. At the same time, the Transformer model has been improved to make it more suitable for handling sentiment analysis tasks.

Improve the accuracy of sentiment analysis of large data sets. By analyzing consumers' online comments, social media interactions and other data, we can understand consumers' emotional tendency and specific evaluation of products and service evaluation and demand, so as to find out the trend and characteristics of consumers' preferences, formulate more targeted marketing strategies, improve marketing effect and customer satisfaction, and improve product user experience. Provide strong support for product or service improvement. Enterprises can also monitor market trends and competitive situation, timely discover new business opportunities and competitors' trends, timely discover and deal with potential brand crises and market risks, avoid or reduce the negative impact on brand image, maintain brand image and reputation, adjust their strategy and product positioning, and maintain competitive advantage.

In short, this model improves the feasibility and effectiveness of text sentiment analysis, which can help enterprises better understand the market and consumer demand, optimize products and services, improve marketing effect and enhance competitiveness. Improvements in these aspects can help enterprises better adapt to market changes and seize business opportunities, improve brand value and market share, and have great commercial value and application prospects.

However, there are many factors that affect the accuracy of text sentiment analysis, among which sample size is an important reason. For deep learning, increasing the number of training samples can fully learn text emotional features and achieve better recognition results. Therefore, the follow-up work of this experiment needs to be tested on a richer data set.

DATA AVAILABILITY

Due to commercial use, the models, code, and data generated by this experiment cannot be released to the public for the time being. The datasets used in the experiments have been marked with sources and can be obtained as needed. Some or all data, models, or code generated or used during the study are available from the corresponding author by request.

REFERENCES

[1] L. Wang, X. Xu, C. Liu, and Z. Chen, "M-DA: A multifeature text data-augmentation model for improving accuracy of Chinese sentiment analysis," *Sci. Program.*, vol. 2022, pp. 1–13, Apr. 2022.

[2] R. Ghasemi, S. A. A. Asli, and S. Momtazi, "Deep Persian sentiment analysis: Cross-lingual training for low-resource languages," *J. Inf. Sci.*, vol. 48, no. 4, pp. 449–462, Aug. 2022.

[3] H. T. Phan, N. T. Nguyen, and D. Hwang, "Convolutional attention neural network over graph structures for improving the performance of aspect-level sentiment analysis," *Inf. Sci.*, vol. 589, pp. 416–439, Apr. 2022.

[4] H. Zhang, Z. Chen, B. Chen, B. Hu, M. Li, C. Yang, and B. Jiang, "Complete quadruple extraction using a two-stage neural model for aspect-based sentiment analysis," *Neurocomputing*, vol. 492, pp. 452–463, Jul. 2022.

[5] B. Yang, B. Shao, L. Wu, and X. Lin, "Multimodal sentiment analysis with unidirectional modality translation," *Neurocomputing*, vol. 467, pp. 130–137, Jan. 2022.

[6] C. Yang, S. Feng, D. L. Wang, N. Yang, and G. Yu, "Analysis on Web public opinion orientation based on extending sentiment lexicon," *J. Chin. Comput. Syst.*, vol. 31, no. 4, pp. 691–695, 2010.

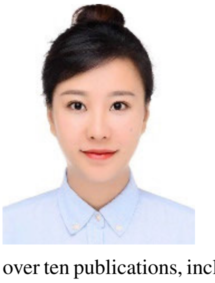
[7] Z. Yanyan, Q. Bing, S. Qiuhui, and L. Ting, "Large-scale sentiment lexicon collection and its application in sentiment classification," *J. Chin. Inf. Process.*, vol. 31, no. 2, pp. 187–193, 2017.

[8] L. Xu, K. Ding, N. Chen, and B. Li, "Corpus construction for citation sentiment in Chinese literature," *J. China Soc. Sci. Tech. Inf.*, vol. 39, no. 1, pp. 25–37, 2020.

[9] Z.-H. Zho and J. Feng, "Deep forest: Towards an alternative to deep neural networks," 2017, *arXiv:1702.08835*.
[10] G. Wang, X. Huang, X. Wu, and X. Hu, "Algorithm of Chinese sentiment classification of SVM based on optimization information gain feature selection method," *J. Chengdu Univ. Technol.*, vol. 46, no. 1, pp. 105–110, 2019.
[11] Y. Su, Y. Zhang, P. Hu, and X. Tu, "Sentiment analysis research based on combination of naive Bayes and latent Dirichlet allocation," *J. Comput. Appl.*, vol. 36, no. 6, pp. 1613–1618, 2016.
[12] Y. Ren, R. Wang, and D. Ji, "A topic-enhanced word embedding for Twitter sentiment classification," *Inf. Sci.*, vol. 369, pp. 188–198, Nov. 2016.
[13] S. Liu and J. Liu, "Public attitudes toward COVID-19 vaccines on english-language Twitter: A sentiment analysis," *Vaccine*, vol. 39, no. 39, pp. 5499–5505, Sep. 2021.
[14] S. Li and C. R. Huang, "Chinese sentiment classification based on stacking combination method," *J. Chin. Inf. Process.*, vol. 24, no. 5, pp. 56–62, 2010.
[15] Y. C. Li, L. Y. Lin, and G. D. Zhou, "Multi-document opinion summarization based on supervised learning," *J. Chin. Inf. Process.*, vol. 28, no. 6, pp. 143–149, 2014.
[16] J. Liu, M. Yan, and J. Luo, "Research on construction of microblog sentiment lexicon based on the smooth SO-PMI algorithm," *J. Hunan Univ. Technol.*, vol. 29, no. 5, pp. 77–81, 2015.
[17] K. Hu, H. Wu, K. Qi, J. Yu, S. Yang, T. Yu, J. Zheng, and B. Liu, "A domain keyword analysis approach extending term frequency-keyword active index with Google Word2Vec model," *Scientometrics*, vol. 114, no. 3, pp. 1031–1068, Mar. 2018.
[18] T. Lei, R. Barzilay, and T. Jaakkola, "Molding CNNs for text: Non-linear, non-consecutive convolutions," *Indiana Univ. Math. J.*, vol. 58, no. 3, pp. 1151–1186, 2015.
[19] K. De Angeli, S. Gao, I. Danciu, E. B. Durbin, X.-C. Wu, A. Stroup, J. Doherty, S. Schwartz, C. Wiggins, M. Damesyn, L. Coyle, L. Penberthy, G. D. Tourassi, and H.-J. Yoon, "Class imbalance in out-of-distribution datasets: Improving the robustness of the text CNN for the classification of rare cancer types," *J. Biomed. Informat.*, vol. 125, Jan. 2022, Art. no. 103957.
[20] W. Ma, H. Z. Yu, and J. Ma, "Tibetan text classification based on word vector features," *Basic Clin. Pharmacol. Toxicology*, vol. 125, no. 2, pp. 76–77, 2019.
[21] C. Jia, J. Ma, X. Yang, and X. Lv, "Multi-modal emotion recognition based on multi-LSTMs fusion," *J. Chin. Inf. Process.*, vol. 2022, no. 5, pp. 145–152, 2022.
[22] X D. Guo, N. Zhao, and S. Z. Cui, "Consumer reviews sentiment analysis based on CNN-BiLSTM," *Syst. Eng.-Theory Pract.*, vol. 40, no. 3, pp. 653–663, 2020.
[23] D. Y. Li, Z. L. Li, and L. Yan, "Research on chinese-oriented entity relation joint extraction method," *J. Chin. Comput. Syst.*, vol. 43, no. 12, pp. 2479–2486, 2022.
[24] A. Mewada and R. K. Dewang, "SA-ASBA: A hybrid model for aspect-based sentiment analysis using synthetic attention in pre-trained language BERT model with extreme gradient boosting," *J. Supercomput.*, vol. 79, no. 5, pp. 5516–5551, Mar. 2023.
[25] J. Yang, Y. Li, H. Zhang, J. Hu, and R. Bai, "Aspect-level sentiment analysis incorporating semantic and syntactic information," *J. Comput. Commun.*, vol. 12, no. 1, pp. 191–207, 2024.
[26] Z. He, "Text sentiment analysis based on multi-layer bi-directional LSTM with a trapezoidal structure," *Intell. Autom. Soft Comput.*, vol. 37, no. 1, pp. 639–654, 2023.
[27] L. Zhang, Y. Wu, Q. Chu, P. Li, G. Wang, W. Zhang, Y. Qiu, and Y. Li, "SA-Model: Multi-feature fusion poetic sentiment analysis based on a hybrid word vector model," *Comput. Model. Eng. Sci.*, vol. 137, no. 1, pp. 631–645, 2023.
[28] H. Rahman, J. Tariq, M. Ali Masood, A. F. Subahi, O. Ibrahim Khalaf, and Y. Alotaibi, "Multi-tier sentiment analysis of social media text using supervised machine learning," *Comput., Mater. Continua*, vol. 74, no. 3, pp. 5527–5543, 2023.
[29] A. Mohamed, Z. Muhammad Zain, H. Shaiba, N. Alturki, G. Aldehim, S. Sakri, S. Farik Mat Yatin, and J. Mohamad Zain, "LexDeep: Hybrid lexicon and deep learning sentiment analysis using Twitter for unemployment-related discussions during COVID-19," *Comput., Mater. Continua*, vol. 75, no. 1, pp. 1577–1601, 2023.
[30] G. X. Xu, L. Y. Liu, J. C. Wang, and Z. Chen, "Text sentiment analysis based on BERT and hypergraph with dual attention network," *Appl. Res. Comput.*, vol. 41, no. 3, pp. 1–786, 2024.
[31] S. Wang, Y. Zhu, W. Gao, M. Cao, and M. Li, "Emotional analysis approach based on dynamic word-sentence features and self attention," *J. Data Acquisition Process.*, vol. 39, no. 1, pp. 193–203, 2024.
[32] X. Fuliang and L. Lifang, "Fine-grained sentiment analysis with CRF and ATAE-LSTM," *Data Anal. Knowl. Discovery*, vol. 4, no. 2, pp. 207–213, 2020.
[33] J. X. Gao and H. Y. Huang, "Text emotion analysis based on TF-IDF and multihead attention transformer model," *J. East China Univ. Sci. Technol.*, vol. 50, no. 1, pp. 129–136, 2024.
[34] B. Zhang and W. Zhou, "Transformer-encoder-GRU (T-E-GRU) for Chinese sentiment analysis on Chinese comment text," *Neural Process. Lett.*, vol. 55, no. 2, pp. 1847–1867, Apr. 2023.
[35] R. Koshy and S. Elango, "Multimodal tweet classification in disaster response systems using transformer-based bidirectional attention model," *Neural Comput. Appl.*, vol. 35, no. 2, pp. 1607–1627, 2023.
[36] Y. Zhou, Q. Hu, J. Liu, and Y. Jia, "Combining heterogeneous deep neural networks with conditional random fields for Chinese dialogue act recognition," *Neurocomputing*, vol. 168, pp. 408–417, Nov. 2015.
[37] W. Wen, R. Cai, Z. Hao, X. Yang, and Y. Li, "Recognizing activities from partially observed streams using posterior regularized conditional random fields," *Neurocomputing*, vol. 260, pp. 294–301, Oct. 2017.
[38] X. Zhang, Y. Jiang, H. Peng, K. Tu, and D. Goldwasser, "Semi-supervised structured prediction with neural CRF autoencoder," in *Proc. Conf. Empirical Methods Natural Lang. Process.*, 2017, pp. 1–12, doi: 10.18653/V1/D17-1179.
[39] A. Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A. N. Gomez, Ł. Kaiser, and I. Polosukhin, "Attention is all you need," 201, *arXiv.1706.03762*.
[40] A. L. Maas, R. E. Daly, P. T. Pham, D. Huang, A. Y. Ng, and C. Potts, "Learning word vectors for sentiment analysis," in *Proc. 49th Annual Meeting Assoc. Comput. Linguistics, Hum. Lang. Technol.*, 2011, pp. 142–150.
[41] *Movie Review Data*. Accessed: Jan. 10, 2024. [Online]. Available: https://www.cs.cornell.edu/people/pabo/movie-review-data/
[42] X. Zhang, J. Zhao, and Y. LeCun, "Character-level convolutional networks for text classification," in *Proc. Adv. Neural Inf. Process. Syst.*, 2015, pp. 1–9.

LISHA YAO was born in Anhui, China, in 1986. She received the master's degree in applied computer technology from Anhui University, in 2011. She is currently pursuing the Ph.D. degree in computer science with National University, Philippines, in 2020.

She has been a Teacher with Anhui Xinhua University, since 2011, and is also an Associate Professor. She presided over six scientific research projects, published more than 20 papers in domestic and foreign academic journals and international conferences, and obtained one national invention patent.



NI ZHENG was born in Anhui, China, in 1985. She received the Bachelor of Science degree from Anhui Normal University, in 2007, the Master of Management degree from Hefei University of Technology, in 2016, and the Ph.D. degree in management from the University of Perpetual Help System DALTA, Philippines, in 2023.

Since 2018, she has been successfully overseen five provincial or higher-level teaching and research projects while actively participating in over ten publications, including articles and appearance patents. Her research interests include management principles along with data management and application techniques. Notably, she was awarded the Provincial Second Prize at the 5th Anhui University Teachers Entrepreneurship Guidance Course Teaching Competition.