Received 16 May 2024, accepted 1 June 2024, date of publication 5 June 2024, date of current version 14 June 2024.

Digital Object Identifier 10.1109/ACCESS.2024.3409818

RESEARCH ARTICLE

Design of Intelligent Sentiment Classification Model Based on Deep Neural Network Algorithm in Social Media

QINGXIANG ZENG
College of Humanities and Media, Hubei University of Science and Technology, Xianning, Hubei 437100, China
e-mail: 18627091638@163.com

This work was supported by the University-Level Innovative Research Team, Hubei University of Science and Technology "Research on the Development of the Meta-Universe and Media" under Grant 2022T06.

ABSTRACT Aspect-based sentiment classification, as a more fine-grained sentiment analysis task, focuses on predicting the sentiment tendency expressed in a sentence based on specific aspects. However, current text sentiment analysis models face challenges when dealing with long comments posted by users on social media, as users often do not explicitly mention sentiment aspects directly in their comments. This paper focuses on aspect extraction and sentiment classification. By constructing a neural network model that integrates a self-attention mechanism, the model is able to learn word embeddings that incorporate contextual semantic information. Furthermore, the author introduce a self-attention mechanism based on relative position representations, which aims to simulate the order of words and achieve parallelized training of inputs by reducing parameters, while simultaneously extracting aspect and sentiment features. Additionally, the author designed a convolutional neural network model and utilized the ReLu gate to selectively output sentiment features based on the given aspect category, while implementing the K-max pooling technique. Comparative experiments conducted on three standard datasets, SemEval, Tweets, and CVAT, showed that this model achieved average best performance on all three datasets. Specifically, on the SemEval dataset, when predicting valence values, the MSE, MAE, and Pearson correlation coefficient reached optimal values of 1.00, 0.88, and 0.73, respectively. While the overall performance on the CVAT dataset was slightly lower, this model still achieved the best MSE of 0.89, MAE of 0.81, and Pearson correlation coefficient of 0.64 when predicting arousal values. This result demonstrates that this method provides relatively balanced and excellent performance in predicting both valence and arousal, validating its practical application value in the field of sentiment analysis.

INDEX TERMS Sentiment classification, convolutional neural network, self-attention mechanism, relative position, K max pooling.

I. INTRODUCTION

The proliferation of the Internet and the ascendance of social media platforms have engendered a voluminous corpus of user-generated comments and opinions. These contents harbor invaluable insights such as users' sentiments, preferences, and emotions, garnering amplified attention toward automated discernment of textual expressions of emotions.

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

Sentiment analysis [1], a facet of natural language processing, aims to decode embedded sentiment within text, manifesting as discrete categories—positive, negative, neutral—or nuanced emotions such as happiness, anger, and sadness.

Within academic circles, online user reviews on social media platforms furnish scholars with a trove of empirical research material. Analyzing these data facilitates deep exploration into societal phenomena, human behavior, and emotional responses. Presently, deep learning methodologies, notably deep neural network models [2], dominate the sentiment classification of user comment data in social media realms. This classification illuminates users' emotional reactions—preferences, satisfaction, dissatisfaction—towards events, products, or services. This comprehension aids enterprises in gauging user emotional inclinations, steering adjustments in product development and marketing strategies.

Currently, traditional Convolutional Neural Networks (CNN) are widely used for intelligent sentiment classification of textual data to extract emotional features. Originally designed for image recognition, CNN has later been successfully applied to text processing tasks such as sentiment classification. CNN extracts local and global features from text through convolutional layers, pooling layers, and fully connected layers. However, traditional CNNs cannot effectively capture word order and long-distance dependencies when processing text. To address this issue, researchers have introduced techniques such as Partial Sampling Gradient Descent (PSGD) [4], attention mechanisms [5], parallel dual pooling, and Word2vec [6], which have improved feature extraction and the ability to extract local features of target keywords. Among these, attention mechanisms allow the model to focus on different parts of the input sequence and assign different weights based on their relative importance. Additionally, relative position representations can capture the order of words in the sequence, enhancing the stability and classification accuracy of the model.

The traditional approach to discrete sentiment classification in short texts encounters a challenge of feature sparsity in text representation due to their limited length, rendering it unsuitable for lengthy user comments on social media platforms. Furthermore, users typically do not incorporate elemental words directly into their comments when expressing sentiment for a specific element. The nuanced task of element-level sentiment classification aims to predict the sentiment conveyed by a sentence regarding a particular element. This task generally involves two subtasks: extracting elements and conducting sentiment classification based on these elements. Thus, this paper undertakes research on element class extraction and sentiment classification based on elements, enhancing a deep neural network model by leveraging insights from deep learning to improve the sentiment classification performance of user comments on social media platforms.

The primary contributions of this paper are outlined as follows:

- Construct a neural network model with a self-attention mechanism. By introducing the self-attention mechanism into the task of feature extraction, the model achieves the integration of global semantic information through the self-attention mechanism, ensuring that the model can comprehensively understand the dependency relationships between different positions in the review text.
- Introduce relative position representation. A neural network model combining self-attention mechanism based on relative position representation is proposed to simulate the order of words, correcting the defect of the self-attention mechanism overlooking temporal information and enhancing the ability to grasp the textual structure and logical relationships.
- The gate-controlled convolutional neural network model achieves parallelized training, optimizing the number of parameters. Meanwhile, it can extract both feature and emotional characteristics from the review text. Furthermore, it addresses the issue of feature loss through contextual optimization and K-max pooling techniques.

This paper delineates the state-of-the-art research on text sentiment categorization using machine learning and deep learning methodologies in Section II. Section III expounds on the introduction of the self-attention mechanism, relative position representation, and element class embedding to construct a gate-controlled convolutional neural network model. Section IV presents comparative and ablation experiments conducted on three benchmark datasets—SemEVAl, Tweets, and CVAT—describing the experimental results and deliberating on the scheme's performance. Finally, Section V concludes the study with a comprehensive discussion.

II. RELATED WORKS

Text Sentiment Classification [7] serves as the cornerstone for numerous research applications, garnering the attention of experts and scholars. This classification task generally operates across three research levels: chapter, sentence, and word levels. Given that social media user comment data predominantly manifests in sentence form, research into text sentiment classification has primarily focused on sentences.

Early scholars in this domain initially employed machine learning [8] for text classification, conducting multiple comparative experiments to verify model efficacy. Subsequent studies [9] explored diverse feature extraction methods, training text sentiment classification models by integrating plain Bayes [10] and SVM [11] algorithms separately, revealing superior results through SVM utilization. Further investigations [12] in microblog text classification employed different feature weights in tandem with SVM, finding optimal accuracy via information gain as a feature selection criterion. Another study [13], drawing from Twitter data, combined multiple lexical features in Simple Bayes and EM algorithms [14], demonstrating the efficacy of these algorithms in text sentiment classification tasks.

In recent endeavors to enhance text sentiment classification efficiency, researchers delved into integrating deep neural network technology [15] atop machine learning frameworks. This approach autonomously learns internal statistical structures from extensive data samples via deep neural network structures, primarily emphasizing textual feature extraction. However, concerns about slow textual feature extraction surfaced. Consequently, a study [16] introduced a hierarchical structure in neural networks, enhancing both rapid feature extraction capability and mitigating challenges related to training deep neural networks.

Utilizing CNNs [17] for text categorization significantly outperformed traditional machine learning algorithms, signifying a new avenue for addressing natural language processing challenges. Subsequent work [18] leveraged recurrent neural network models for sentiment feature extraction in microblog sentiment classification. However, its applicability was limited to short text sentiment classification. In contrast, another study [19] proposed a semi-supervised convolutional neural network sentiment analysis model, leveraging distinct feature information across various information channels, confirming its efficacy.

The aforementioned sentiment categorizations predominantly rely on rule-based methodologies. These approaches involve assigning varying weights to emotion words based on an emotion lexicon encompassing negative terms, degree adverbs, domain-specific language, and other words indicative of emotional tendencies. Subsequently, sentiment scores are weighted, summed using basic statistical methods, and compared against predetermined thresholds for sentiment classification.

Presently, the rapidly evolving landscape of internet lexicon introduces challenges such as multiple word meanings and the inadequacy of existing sentiment lexicons to encompass new sentiment-laden vocabulary encountered in user comments on social media. To address this, recent studies have explored novel approaches to expand sentiment lexicons from existing seed sentiment words. One such study [20], [21], [22] leveraged convolutional neural networks (CNNs) for sentence-level sentiment classification, exploring variants including static CNNs based on random and pre-trained Word2vec word embeddings. Results indicated that pre-trained word embeddings notably enhanced model performance [23], [24].

Alternatively, a distinct study [25] introduced a dynamic Convolutional Neural Network (DCNN) model, leveraging pre-trained embeddings to bolster the performance of deep neural networks in sentiment categorization, with a keen focus on semantic modeling of sentences. In parallel, the CharSCNN model [26], which incorporates character embeddings and utilizes dual convolutional layers, exhibited remarkable gains in performance, though its application is primarily confined to sentiment classification of short texts.

Alternatively, a study [27] utilized Long Short-Term Memory (LSTM) networks [28] for Twitter text sentiment classification, simulating word interactions in the text formation process, enabling sentiment classification in longer texts. However, its performance in recognizing elemental words within the text was suboptimal. Finally, another study [28] devised a recurrent random wandering network model to conduct sentiment classification experiments on user-posted tweets.

III. METHODOLOGY

In this research, the author design an efficient deep neural network model tailored to the distinctive characteristics of user comment data on social media platforms. Recognizing the pivotal role of word co-occurrence distribution information in comment text for element extraction, the author begin by introducing the self-attention mechanism. This addresses the limitation of the general attention mechanism, which fails to comprehensively consider semantic and syntactic relationships among different words across sentences. Furthermore, the author incorporate relative position representation between inputs to capture word order, building upon the foundation of the self-attention mechanism.

To circumvent limitations inherent in the sequential nature of LSTM models, hindering parallelized training of text, the author pivot towards constructing a CNN model grounded in a gating mechanism for elemental class sentiment classification. Enhancements are made to the pooling operation method within the GCAE (Gated Convolutional Autoencoder) model, culminating in a novel proposal: a gated convolutional neural network model featuring K-max pooling integrated with element class embedding. This amalgamation aims to optimize the performance of sentiment classification tasks pertaining to elemental classes.

A. SELF-ATTENTION MECHANISM

The author employ an encoder founded on the attention mechanism to generate vector representations for sentences. Additionally, the author utilize a linear combination of reconstructed sentences, utilizing the element class embedding matrix for element class embedding. Initially, the author establish a unified model, illustrated in Figure 1.

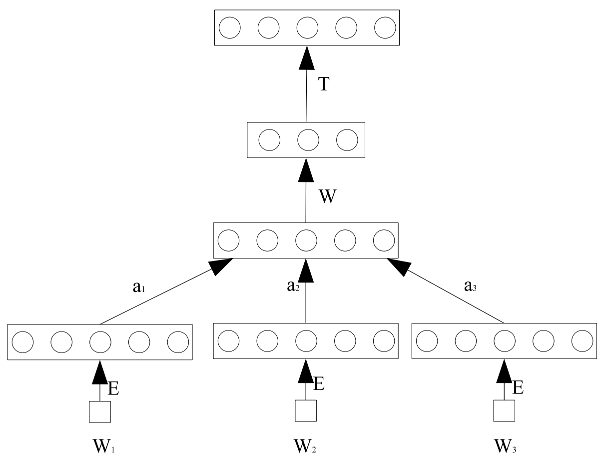


FIGURE 1. Model of factor extraction task.

Given that word-based embeddings frequently position words sharing contextual similarity in proximate regions within the word embedding space, encoding word co-occurrence distributions involves transforming words within the sentence into word vectors e_i within the word embedding matrix $E = (e_1, e_2, \ldots, e_v) \in \mathbb{R}^{v \times d} (e_i \in \mathbb{R}^d)$.

Subsequently, leveraging the attention mechanism, each word is assigned an associated weight. This weighting mechanism serves to filter out irrelevant words within the sentence, enhancing the coherence of the extracted elemental classes.

The objective is to capture the most pertinent information concerning the elemental classes inherent within the sentence. The weights a_i is calculated as follows:

$$y_s = \frac{1}{n}\sum_{i=1}^{n} e_i \tag{1}$$

$$d_i = e_i^T \cdot M \cdot y_s \tag{2}$$

$$a_i = \frac{\exp(d_i)}{\sum_{j=1}^{n}\exp(d_j)} \tag{3}$$

where n denotes the sentence length, y_s is the average word embedding representation of the sentence and reflects the global contextual information of the sentence, $M \in \mathbb{R}^{d\times d}$ is the transformation matrix to be learned, and after obtaining the attentional weights of the individual words in the sentence, the corresponding word vectors are weighted for linear summation, and the embedding representation of the sentence can be obtained from z_s:

$$z_s = \sum_{i=1}^{n} a_i e_i \tag{4}$$

Next, the author perform dimensionality reduction with softmax normalization on z_s to obtain the normalized weight vector of the element class p_t:

$$p_t = soft\max(W \cdot z_s + b) \tag{5}$$

where W denotes the weight matrix and b denotes the bias vector. Ultimately, let T represent the element class embedding matrix. At this stage, the sentence can be effectively depicted as a linear combination of element class embeddings. This computation r_s is accomplished through the utilization of the normalized weight vector p_t associated with the element class.

$$r_s = T^T \cdot p_t \tag{6}$$

In the preceding process, the absence of inter-word attention computation within the input sequence results in degraded performance. To address this limitation, the author introduce enhancements to the self-attention mechanism. This upgraded mechanism facilitates direct computation of dependencies among words, enabling the model to learn the intrinsic structure of sentences. Consequently, each position's representation encompasses comprehensive semantic information at a global level, thereby acquiring word embeddings imbued with contextual semantic nuances. The model's structural framework is delineated below.

According to Fig. 2, the author use three initialization matrices W^Q, W^K, and W^V of all sizes $d_x \times d_x$ but with different values to get Q, K, and V to enhance the representation of the model when perform a linear transformation on the input $X = [e_1, e_2, \ldots, e_n] \in \mathbb{R}^{n\times d_x}$ of the self-attention module, and then the scaled dot product function is used as the alignment score function to compute the similarity between query vectors and key vectors $s(Q, K)$:

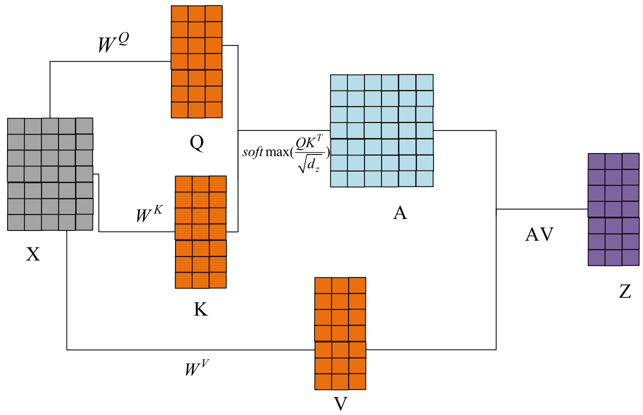


FIGURE 2. Improved self-attention mechanism.

$$s(Q_i, K_i) = \frac{Q_i K_j^T}{\sqrt{d_z}} = \frac{(e_i W^Q)\cdot(e_j W^K)^T}{\sqrt{d_z}} \tag{7}$$

The normalization of the chi-score function yields the attention weight matrix, denoted as A. The calculation for each element within matrix A is derived as follows:

$$a_{ij} = \frac{\exp(s(Q_i, K_j))}{\sum_{j=1}^{n}\exp(s(Q_i, K_j))} \tag{8}$$

The output of the self-attention module, constituting the vector representation of the sentence, is computed as the weighted sum of the attention weights assigned to individual words within the sentence and their corresponding word vectors. This calculation, achieved by the product of matrix A and matrix V, serves as the output of the self-attention module, encapsulating the vector representation of the sentence $Z = [z_1, z_2, \ldots, z_n] \in \mathbb{R}^{n\times d_z}$:

$$z_i = \sum_{j=1}^{n} a_{ij} V^j = \sum_{j=1}^{n} a_{ij}\cdot(e_j W^V) \tag{9}$$

B. SELF-ATTENTION MECHANISM BASED ON RELATIVE POSITION REPRESENTATION

In order to avoid the problem of low model performance caused by not taking the location feature of the input into account, the calculated weights only depend on the relevance of the query vector Q_i and the key vector K_j. The author add the input position encoding information to correct the self-attention mechanism.

As shown in Fig. 3, the author model the input as a labeled, directed, and fully connected graph. In Fig. 3, $a_{ij}^V, a_{ij}^K \in \mathbb{R}^{d_a}$ represents the edge between the inputs e_j and e_i.

The author introduce information pertaining to relative position representation into the computation of the similarity function and module outputs. This integration occurs through an additive combination mechanism based on the

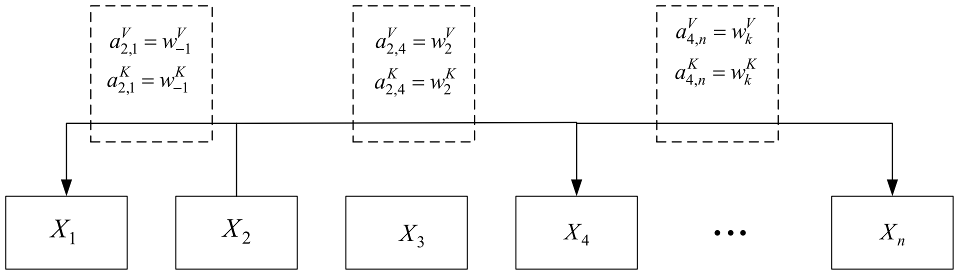


FIGURE 3. Schematic representation of relative positions.

self-attention mechanism. Check the alignment score function $s(Q, K)$ for the similarity between the query vector and the key vector at this point:

$$s(Q_i, K_i) = \frac{Q_i(K_j + a^K_{ij})^T}{\sqrt{d_z}} = \frac{(e_i W^Q) \cdot (e_j W^K + a^K_{ij})^T}{\sqrt{d_z}} \tag{10}$$

The results of the normalization of the chi-score function form the attention weight matrix A. Each element of the matrix A is calculated as follows:

$$a_{ij} = \frac{\exp(s(Q_i, K_j))}{\sum_{j=1}^{n} \exp(s(Q_i, K_j))} \tag{11}$$

The output of the self-attention module, constituting the vector representation of the sentence, is computed as the weighted sum of the attention weights allocated to individual words within the sentence along with their corresponding word vectors. This computation involves the product of matrix A and matrix, that is $Z = [z_1, z_2, \ldots, z_n] \in \mathbb{R}^{n \times d_z}$:

$$z_i = \sum_{j=1}^{n} a_{ij}(V^j + a^V_{ij}) = \sum_{j=1}^{n} a_{ij} \cdot (e_j W^V + a^V_{ij}) \tag{12}$$

C. CNN MODEL BASED ON GATING MECHANISM

In this section, the author develop a K max pooled gated convolutional neural network model, integrating a context-optimized gating mechanism for elemental class embedding along with a self-attention mechanism based on relative position representation outlined in 3.2. Illustrated in Fig. 4, the text input undergoes transformation using pre-trained word vectors. Two independent convolutional layers are configured on the word embedding layer and subsequently linked via the Tanh gating unit and ReLU gating unit. These units are merged, followed by the derivation of a fixed-dimension feature vector through maximum value pooling. Finally, a softmax process is applied.

The Gated Tanh-ReLU Unit (GTRU) governs the flow path of sentiment features to the pooling layer, enabling the model to selectively produce sentiment features based on a specified element class. The author set the ReLu gate to receive the output of the convolutional layer along with information about

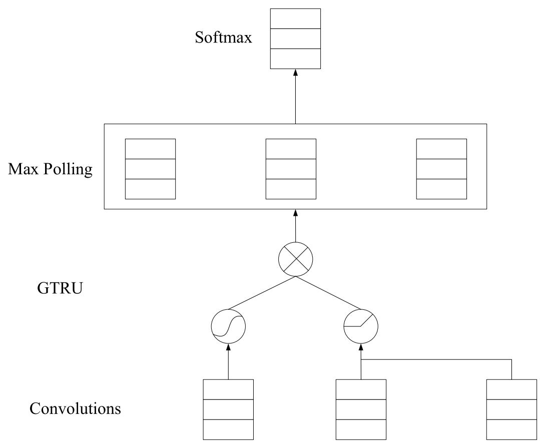


FIGURE 4. Convolutional neural networks for emotion classification.

a specific element class to generate element class-specific features a_i:

$$a_i = relu(X_{i:i+k} \cdot W_a + V_a \cdot v_a + b_a) \tag{13}$$

where $X_{i:i+k}$ represents the output of the convolutional layer, v_a represents the embedding vectors of a particular lineage, W_a, v_a, and b_a represent the parameters to be trained.

The Tanh gate exclusively accepts the output originating from the convolutional layer, generating the sentence's sentiment feature. The activation function employed is Tanh, yielding values within the range of -1 to 1. The sentiment feature of the sentence s_i is:

$$s_i = \tanh(X_{i:i+k} \cdot W_s + b_s) \tag{14}$$

where W_s and b_s represent the parameters to be trained. Next, the outputs of the ReLU gate and the Tanh gate are multiplied element by element so that the element class features and the sentiment features can be corresponded to obtain the sentiment features of a particular element class c_i:

$$c_i = s_i \cdot a_i \tag{15}$$

Subsequently, the author acquire a fixed-dimensional feature vector through maximum pooling and execute softmax processing. When a sentence of length 'n' undergoes a convolution operation utilizing a convolution kernel of dimension 'k,' it yields 'n - k + 1'features. The resultant convolutional feature map 'c' corresponds to the sentence.

$$c = [c_1, c_2, \ldots, c_{n-k+1}] \tag{16}$$

The extraction of the most crucial sentiment features related to specific elemental classes within the sentence involves preserving the maximum value within each convolutional outcome. Subsequently, less pertinent sentiment features are discarded.

$$c' = \max(c) \tag{17}$$

Given the typical presence of multiple convolutional kernels within a convolutional layer, the feature information extracted from the maximum pooling of 'N' convolutional kernels individually within the pooling layer is concatenated or spliced together. Hence, the resulting outcome is:

$$C' = [c'_1, c'_2, \ldots, c'_N] \tag{18}$$

The concatenated spliced result is utilized as an input for the fully connected layer and subjected to normalization via the softmax activation function to derive a probability value. Denoting 'W' as the weight matrix of the fully connected layer and 'B' as the bias, the predicted classification result is as follows:

$$y' = soft\max(W \cdot C' + B) \tag{19}$$

IV. EXPERIMENTS AND ANALYSIS

In this section, the author perform a comparative analysis, evaluating the enhanced convolutional neural network model against DCNN [25], CharSCNN model [26], and LSTM model [27] to assess its performance. The evaluation encompasses three distinct datasets: SemEVAl, Tweets, and CVAT. CVAT stands as a Chinese dataset, while the remaining two are English datasets. These datasets categorize each text with discrete sentiment categories such as positive, negative, etc., or with continuous/multipoint scale labels [28]. The English datasets label text with discrete emotion categories like positive, negative, or utilize continuous/fine-grained emotion intensity magnitudes, all of which serve as valid datasets for this experimental analysis [29].

A. EXPERIMENTAL INDICATORS

The model's predictive performance is assessed by evaluating the disparity between predicted and actual values. A set of evaluation metrics was employed to validate the predictions, encompassing Mean Square Error (MSE), Mean Absolute Error (MAE), and Pearson correlation coefficient (Pearson correlation coefficient 'r'). The Pearson correlation coefficient 'r' values range within [−1, 1], where values closer to 1 denote a positive correlation, values near -1 signify a negative correlation, and those closer to 0 indicate a weaker linear correlation.

While both MSE and MAE gauge the magnitude of the gap between the true and predicted values, MSE penalizes larger errors more than MAE. Consequently, the simultaneous use of both metrics aids in comprehensively analyzing the method's validity. where n denotes the number of samples in the test data, and y_i and $\widetilde{y_i}$ denote the true and predicted values, respectively. avg(y_i) represents the average value of y_i.

The formula for calculating the MSE is as follows:

$$MSE = \frac{1}{n}\sum_{i=1}^{n}(\text{avg}(y_i) - \widetilde{y_i})^2 \tag{20}$$

The formula for the MAE is given below:

$$MAE = \frac{1}{n}\sum_{i=1}^{n}|\text{avg}(y_i) - \widetilde{y_i}| \tag{21}$$

The Pearson correlation coefficient is calculated as follows:

$$r = \frac{\sum_{i=1}^{n}(\widetilde{y_i} - \frac{1}{n}\sum_{j=1}^{n}\widetilde{y_i}) \cdot (y_i - \frac{1}{n}\sum_{j=1}^{n}y_i)}{\sqrt{\sum_{i=1}^{n}(\widetilde{y_i} - \frac{1}{n}\sum_{j=1}^{n}\widetilde{y_i})^2} \cdot \sqrt{\sum_{i=1}^{n}(y_i - \frac{1}{n}\sum_{j=1}^{n}y_i)^2}} \tag{22}$$

B. PARAMETERS SETTING

In this segment, the pertinent hyperparameters are uniformly established based on prior experience and pre-experimental adjustments. The skip-gram model within the word2vec method is invoked via the gensim package to pre-train word vectors for words across the three datasets. This paper utilizes pre-trained word2vec word vectors, employing a dimensionality of 300. For unregistered words absent from the word lists and not included in the word vector dictionary, a uniform distribution ranging from [-0.25, 0.25] is utilized for random initialization, generating random vectors of 300 dimensions.

Moreover, in the element class extraction experiment, within the Keras experimental framework, Adam serves as the optimizer. The hyperparameters entail setting the number of negative samples as 6, an initial learning rate of 0.01, 50 training rounds, a batch size of 256, a dropout rate of 0.5 to mitigate overfitting, and a regularization term weight of 0.001. In the element class sentiment classification experiment employing PyTorch as the experimental framework, the optimizer remains Adam, employing 100 filters with window sizes of 2, 3, and 4. Further specifications include an initial learning rate of 0.01, 30 training rounds, a batch size of 256, and a dropout rate of 0.5 to counteract overfitting.

C. COMPARATIVE RESULTS

Valence typically denotes the extent of positive or negative emotional states. A higher value signifies a positive emotion, while a lower value signifies a negative emotion. Arousal describes the degree of emotional excitement, where higher values indicate more intense emotions and lower values denote calmer emotions. Hence, this paper compares the model against comparison models in the VA (valence and arousal) prediction experiment. To ensure a fair evaluation of different methods' performance, valence and arousal values

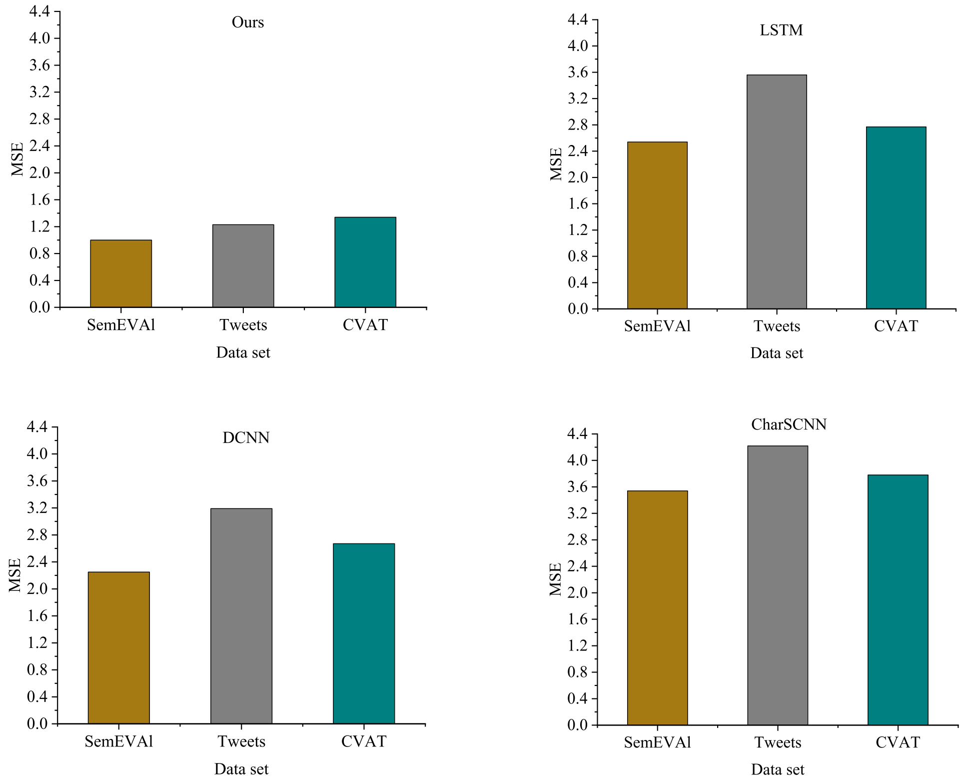


FIGURE 5. **Comparison of MSE results of the models in predicting Valence values on the three data sets.**

at both sentence and word levels are mapped within the intervals [1], [9].

The model's hyperparameters are selected through held-out data and subsequently retrained using the chosen optimal hyperparameters. To avert overfitting due to excessive iterations, the network training epoch parameters are established using the Early Stop strategy. This strategy involves training on the training set and obtaining test results from the validation set. As the epochs progress, if it's observed that the test error ceases to decrease in the validation set, the training process is halted. The weights after this point are utilized as the final parameters for the network.

Figure 5 illustrates the MSE outcomes of various models in predicting valence across the aforementioned three datasets. The experimental analysis for valence prediction spans all three datasets. The findings reveal that the model devised in this paper outperforms nearly all benchmark methods, particularly showcasing its prowess in predicting valence. Specifically, on the SemEVAl dataset, the MSE for valence prediction reaches an impressive low of 1.00, while the comparison models LSTM, DCNN, and CharSCNN exhibit MSE values of 2.54, 2.25, and 3.54, respectively. On the Tweets dataset, the MSE values peak at 3.56, 3.19, and 4.22 for the LSTM, DCNN, and CharSCNN models, respectively.

Overall, this paper's model exhibits an average improvement of 45.23% in MSE performance across all three datasets in comparison to the alternative schemes.

Figure 6 displays the outcomes of Mean Absolute Error (MAE) and Pearson's correlation coefficient of various models in predicting VALENCE. The comparison is conducted across individual datasets to assess the performance of each model. On the SemEVAl dataset, this model achieves an MAE of 0.88 and a Pearson's correlation coefficient of 0.73, showcasing notable performance. Interestingly, the CharSCNN model exhibits similar performance with an MAE of 0.98 and a Pearson's correlation coefficient of 0.72, where the model's Pearson's correlation coefficient stands out at 0.79, representing the best performance achieved. In contrast, the CharSCNN model attains a Pearson correlation coefficient of 0.69.

However, the performance across the Chinese dataset CVAT diminishes for each model. Notably, the CharSCNN model obtains an MAE and Pearson's correlation coefficient of 0.63 and 0.6, respectively, displaying relatively better performance compared to this model and the other comparative models, which maintain MAEs around 0.7 and Pearson's correlation coefficients around 0.3. Notably, the LSTM model exhibits the poorest performance with an MAE of 0.71 and a Pearson correlation coefficient of 0.26.

By further combining the experimental results in Figures 6 and 7, it can be seen that the model constructed in this paper can effectively identify the positive or negative degree of emotional states when predicting valence. Therefore, it can classify emotions such as happiness, joy,

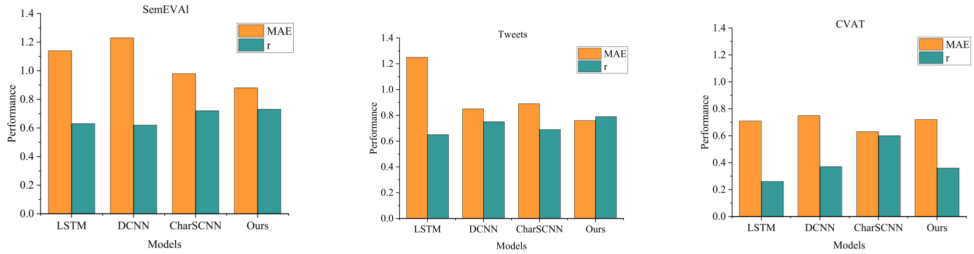


FIGURE 6. Comparison of MAE and Pearson correlation coefficient results for each model in predicting valence.

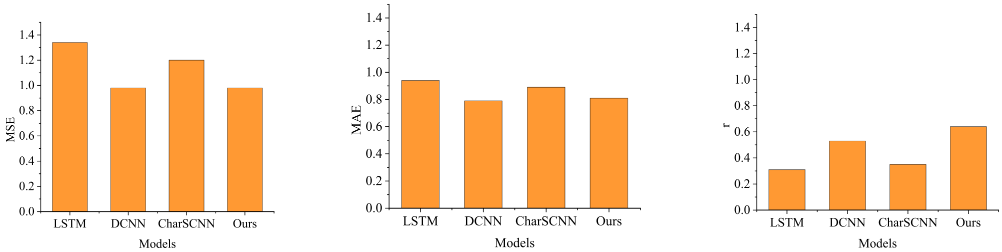


FIGURE 7. Comparison of MSE, MAE and Pearson's correlation coefficient dismissal in predicting arousal by models.

love, satisfaction, and self-confidence based on the positive emotions expressed in the review text, or classify emotions such as sadness, anger, fear, anxiety, and depression based on the negative emotions expressed in the review text, thus improving the accuracy of sentiment classification at a fine-grained level.

Figure 7 exhibits MSE, MAE, and Pearson correlation coefficient results for each model in predicting arousal. The experiments for arousal prediction are exclusively conducted on the CVAT dataset. Notably, the constructed convolutional neural network showcases identical MSE performance to the DCNN model, both registering an MSE of 0.98. In contrast, the LSTM model and CharSCNN model yield MSEs of 1.34 and 1.20, respectively.

Furthermore, a comparative evaluation between the proposed model in this paper and the DCNN model for predicting arousal includes an in-depth analysis of both MAE and Pearson's correlation coefficient. While the Pearson correlation coefficient of the current model exhibits a slight decrement compared to other datasets, it is noteworthy that the DCNN model outperforms the LSTM and CharSCNN models, demonstrating superior results. However, this model surpasses the DCNN model by a margin of 0.11 in Pearson's correlation coefficient, demonstrating its robustness and effectiveness. Moreover, MAE metrics further confirm this superiority, with this model achieving a value of 0.81, while the LSTM, DCNN, and CharSCNN models record scores of 0.94, 0.79, and 0.89, respectively.

These arousal prediction results demonstrate that the model leverages the degree of positive and negative emotions to ascertain the degree of relevance in cognition and expression. It effectively detects the level of agitation within emotional expression, discerning positive emotions like affirmation, self-confidence, happiness, or negative emotions such as denial, inferiority complex, and pain. The comparative experiments highlight that this method offers relatively balanced scores in both valence and arousal prediction, achieving high accuracy in emotion classification.

D. ABLATION EXPERIMENTS

In this section, the author perform ablation experiments on the proposed convolutional neural network model to enhance the control over individual variants and verify the impact of improving modules or adjusting parameters on the model's performance. Let's denote E1 as the final model presented in this paper. In E1, both the element class vectors of the input model and the standard word vectors are initialized using pre-trained word vectors from word2vec. The pooling layer follows the maximum pooling operation.

Moving on to E2, it involves initializing the element class vectors of the model with the element class embedding matrix extracted from the neural network model combined with the attention mechanism. Here, the standard word vectors are also initialized from the model's training for consistency in the word embedding space. The pooling layer again adopts the maximal pooling operation.

Next, E3 initializes based on E2 but enhances it by utilizing the element class embedding matrix extracted from the neural network model, which incorporates an improved self-attention mechanism.

Finally, E4 evolves from E3 by initializing the element class vectors of the input model using the element class embedding matrix extracted by introducing the relative

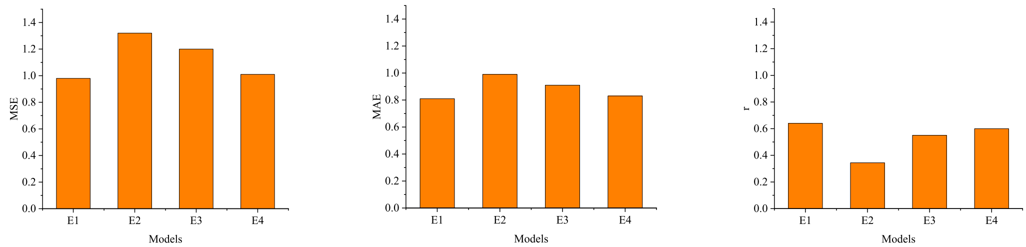


FIGURE 8. Results of ablation experiments.

position representation. Notably, in E4, the pooling layer operation does not involve the maximum pooling operation.

These variations (E1, E2, E3, and E4) allow us to examine how different initializations and mechanisms impact the performance of the model, contributing to a more nuanced understanding of the model's behavior under various settings and enhancements.

The ablation experiments conducted on the SemEVAl dataset yielded results for E1, E2, E3, and E4, showcased in Fig. 8. Notably, the model's performance demonstrated continual improvement with each enhanced mechanism in the model. The E2 model exhibited the poorest performance, indicating a Pearson's correlation coefficient of only 0.334. However, as the author progress to E3 and E4, there's a gradual enhancement in performance. Particularly, the utilization of the improved self-attention mechanism in E3 led to a remarkable 23% improvement in MSE.

It's important to note that a higher Pearson correlation coefficient implies a stronger correlation between words in the sentiment dictionary and the target sentiment category, thereby favorably impacting sentiment classification performance. Analyzing the shift in Pearson's correlation coefficient between E1 and E4 reveals that the enhanced self-attention mechanism significantly elevates the deep neural network algorithm's effectiveness in sentiment classification.

In E4, where the maximum pooling operation along the pooling layer is omitted, the model's MSE, MAE, and Pearson correlation coefficient are 1.01, 0.83, and 0.6, respectively. Despite this performance, it falls slightly behind E1. This analysis from the ablation experiments highlights the crucial role of the introduced and improved self-attention mechanism, as well as the adoption of maximum pooling operations along the pooling layer, in the model's construction. These mechanisms significantly contribute to the model's effectiveness and overall performance in sentiment classification tasks.

E. DISCUSSION

The experimental analyses conducted in Sections IV-C and IV-D underscore the substantial enhancements introduced in this paper. By integrating the self-attention mechanism with relative positional representation into the elemental class extraction task, the author successfully imbued each position in comment text with comprehensive global semantic information. This integration not only enhances the self-attention mechanism but also addresses the inherent drawback of disregarding temporal information in self-attention models.

Moreover, my incorporation of the self-attention mechanism and relative position representation into a gate-controlled convolutional neural network model compensates for the limitations in sentiment analysis of short texts. This integration facilitates element-level sentiment classification with a focus on fine granularity. Experimental results across three datasets affirm the improvements observed in MSE, MAE, and Pearson's correlation coefficient. Notably, on the SemEVAl dataset, the model achieves an MSE of 1.00, while MAE and Pearson correlation coefficient stand at 0.88 and 0.73, respectively, in valence prediction. For arousal prediction on the CVAT dataset alone, the model attains optimal MSE, MAE, and Pearson's correlation coefficient results of 1.00, 0.88, and 0.73, respectively.

The efficiency of this model becomes evident when applied to sentiment analysis of user comments on social media platforms. It adeptly comprehends users' emotional stances towards specific events, products, or topics, thereby assisting companies in better understanding user needs and feedback. This understanding fosters product improvements and aids in fostering stronger user relationships. Additionally, the model's capability to identify user emotional states enables the provision of personalized services and solutions. Consequently, this tool helps companies decipher consumer emotional feedback about their brands, identify potential issues, delineate improvement strategies, and ultimately augment user satisfaction.

V. CONCLUSION

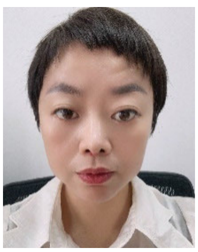
In this study, the author designed a convolutional neural network (CNN) tailored specifically for class extraction and sentiment classification of elemental classes. The model achieved contextual semantic learning in word embeddings by integrating a self-attention mechanism into the elemental class extraction task, along with enhancements to this mechanism. Additionally, the author introduced innovative elements such as relative position representation and gating mechanisms. Ultimately, a CNN model incorporating K-max value pooling was developed, effectively addressing issues of context detachment and significant feature loss in elemental class embeddings. Experimental results validate the model's efficacy.

On the SemEVAl dataset, the model exhibited commendable performance in valence prediction, achieving an MSE of 1.00, an MAE of 0.88, and a Pearson correlation coefficient of 0.73. Similarly, on the CVAT dataset for arousal prediction, the model delivered notable results with an MSE of 0.89, an MAE of 0.81, and a Pearson correlation coefficient of 0.64. These findings underscore the model's ability to maintain balanced performance in both valence and arousal predictions. Ablation experiments further confirm the significance of the introduction and enhancement of the self-attention mechanism, as well as the utilization of maximum pooling operations at the pooling layer. These mechanisms significantly boost the model's performance, enabling accurate sentiment categorization of user comments on social media platforms.

Beyond sentiment analysis, the model's applications extend to public opinion monitoring and management, personalized services, market research, product improvement, and emotional intelligent assistance. As technology continues to advance, this emotion classification model holds the promise of bringing further convenience and intelligence to people's lives and work in the future.

REFERENCES

[1] M. Wankhade, A. C. S. Rao, and C. Kulkarni, "A survey on sentiment analysis methods, applications, and challenges," *Artif. Intell. Rev.*, vol. 55, no. 7, pp. 5731–5780, Oct. 2022.
[2] A. S. M. Alharbi and E. de Doncker, "Twitter sentiment analysis with a deep neural network: An enhanced approach using user behavioral information," *Cognit. Syst. Res.*, vol. 54, pp. 50–61, May 2019.
[3] Y. Cheng, M. Lin, J. Wu, H. Zhu, and X. Shao, "Intelligent fault diagnosis of rotating machinery based on continuous wavelet transform-local binary convolutional neural network," *Knowl.-Based Syst.*, vol. 216, Mar. 2021, Art. no. 106796.
[4] J. Jiang and L. Hu, "Decentralised federated learning with adaptive partial gradient aggregation," *CAAI Trans. Intell. Technol.*, vol. 5, no. 3, pp. 230–236, Sep. 2020.
[5] Z. Niu, G. Zhong, and H. Yu, "A review on the attention mechanism of deep learning," *Neurocomputing*, vol. 452, pp. 48–62, Sep. 2021.
[6] G. Di Gennaro, A. Buonanno, and F. A. N. Palmieri, "Considerations about learning Word2Vec," *J. Supercomput.*, vol. 77, no. 11, pp. 12320–12335, Nov. 2021.
[7] X. Yang, S. Feng, D. Wang, and Y. Zhang, "Image-text multimodal emotion classification via multi-view attentional network," *IEEE Trans. Multimedia*, vol. 23, pp. 4014–4026, 2021.
[8] J. G. Greener, S. M. Kandathil, L. Moffat, and D. T. Jones, "A guide to machine learning for biologists," *Nature Rev. Mol. Cell Biol.*, vol. 23, no. 1, pp. 40–55, Jan. 2022.
[9] S. Rani and N. Gill, "Hybrid model for Twitter data sentiment analysis based on ensemble of dictionary based classifier and stacked machine learning classifiers-SVM, KNN and C5.0," *J. Theor. Appl. Inf. Technol.*, vol. 98, no. 4, pp. 624–635, 2020.
[10] I. Wickramasinghe and H. Kalutarage, "Naive Bayes: Applications, variations and vulnerabilities: A review of literature with code snippets for implementation," *Soft Comput.*, vol. 25, no. 3, pp. 2277–2293, Feb. 2021.
[11] V. K. Chauhan, K. Dahiya, and A. Sharma, "Problem formulations and solvers in linear SVM: A review," *Artif. Intell. Rev.*, vol. 52, no. 2, pp. 803–855, Aug. 2019.
[12] H. Li, Z. Ma, H. Zhu, Y. Ma, and Z. Chang, "An ensemble classification algorithm of micro-blog sentiment based on feature selection and differential evolution," *IEEE Access*, vol. 10, pp. 70467–70475, 2022.
[13] B. H. Back and I. K. Ha, "Comparison of sentiment analysis from large Twitter datasets by Naïve Bayes and natural language processing methods," *J. Inf. Commun. Convergence Eng.*, vol. 17, no. 4, pp. 239–245, 2019.
[14] C. Mari and C. Baldassari, "Unsupervised expectation-maximization algorithm initialization for mixture models: A complex network-driven approach for modeling financial time series," *Inf. Sci.*, vol. 617, pp. 1–16, Dec. 2022.
[15] N. D. Al-Shakarchy, H. K. Obayes, and Z. N. Abdullah, "Person identification based on voice biometric using deep neural network," *Int. J. Inf. Technol.*, vol. 15, no. 2, pp. 789–795, Feb. 2023.
[16] M. M. Taye, "Theoretical understanding of convolutional neural network: Concepts, architectures, applications, future directions," *Computation*, vol. 11, no. 3, p. 52, Mar. 2023.
[17] S. Soni, S. S. Chouhan, and S. S. Rathore, "TextConvoNet: A convolutional neural network based architecture for text classification," *Appl. Intell.*, vol. 53, no. 11, pp. 14249–14268, Jun. 2023.
[18] Y. Qin, Y. Shi, X. Hao, and J. Liu, "Microblog text emotion classification algorithm based on TCN-BiGRU and dual attention," *Information*, vol. 14, no. 2, p. 90, Feb. 2023.
[19] B. Haq, S. M. Daudpota, A. S. Imran, Z. Kastrati, and W. Noor, "A semi-supervised approach for aspect category detection and aspect term extraction from opinionated text," *Comput., Mater. Continua*, vol. 77, no. 1, pp. 115–137, 2023.
[20] J. Mutinda, W. Mwangi, and G. Okeyo, "Sentiment analysis of text reviews using lexicon-enhanced BERT embedding (LeBERT) model with convolutional neural network," *Appl. Sci.*, vol. 13, no. 3, p. 1445, Jan. 2023.
[21] F. K. Khattak, S. Jeblee, C. Pou-Prom, M. Abdalla, C. Meaney, and F. Rudzicz, "A survey of word embeddings for clinical text," *J. Biomed. Informat.*, vol. 100, Jan. 2019, Art. no. 100057.
[22] Z. Yuan, B. Wu, G. Sun, Z. Liang, S. Zhao, and W. Bi, "S2DNAS: Transforming static CNN model for dynamic inference via neural architecture search," in *Proc. 16th Eur. Conf. Comput. Vis.* Glasgow, U.K.: Springer, Aug. 2020, pp. 175–192.
[23] E. M. Mercha, H. Benbrahim, and M. Erradi, "Graph convolutional network for multilingual sentiment analysis," in *Modern Artificial Intelligence and Data Science*. Cham, Switzerland: Springer, 2023, pp. 111–119.
[24] C. Sitaula and T. B. Shahi, "Multi-channel CNN to classify Nepali COVID-19 related tweets using hybrid features," *J. Ambient Intell. Hum. Comput.*, vol. 15, no. 3, pp. 2047–2056, Mar. 2024.
[25] E. Elsayed and D. Fathy, "Semantic deep learning to translate dynamic sign language," *Int. J. Intell. Eng. Syst.*, vol. 14, no. 1, pp. 316–325, Feb. 2021.
[26] S. Voghoei, J. Byars, J. A. Miller, K. Rasheed, and H. A. Arabnia, "Decoding the alphabet soup of degrees in the United States postsecondary education system through hybrid method: Database and text mining," 2023, *arXiv:2309.13050*.
[27] T. Li, P. Quinn, and S. Zhai, "C-PAK: Correcting and completing variable-length prefix-based abbreviated keystrokes," *ACM Trans. Comput.-Hum. Interact.*, vol. 30, no. 1, pp. 1–35, Feb. 2023.
[28] H. Zhong, Z. Wei, Y. Man, S. Pan, J. Zhang, B. Niu, X. Yu, Y. Ouyang, and Q. Xiong, "Prediction of instantaneous yield of bio-oil in fluidized biomass pyrolysis using long short-term memory network based on computational fluid dynamics data," *J. Cleaner Prod.*, vol. 391, Mar. 2023, Art. no. 136192.
[29] M. Davoudi, M. R. Moosavi, and M. H. Sadreddini, "DSS: A hybrid deep model for fake news detection using propagation tree and stance network," *Expert Syst. Appl.*, vol. 198, Jul. 2022, Art. no. 116635.



QINGXIANG ZENG received the Ph.D. degree in literature from Wuhan University, in June 2014. She currently works with the School of Humanities and Media, Hubei University of Science and Technology. Her research interests include new media, film and television, and internet communication.