Journal of Cases on Information Technology
Volume 25 • Issue 1

A Novel Long and Short-Term Memory Network-Based Krill Herd Algorithm for Explainable Art Sentiment Analysis in Interior Decoration Environment

Zhiqiang Gao, Zhengzhou Railway Vocational and Technical College, China*

ABSTRACT

Aiming at the problem that most existing models of art sentiment analysis only consider text encoding from the word level, this paper proposes a novel long and short-term memory network-based krill herd algorithm for explainable art sentiment analysis in interior decoration environment. Firstly, multi-scale convolution is used to capture local correlation of different granularity, so as to obtain more semantic information of different levels and form richer text representation. Then, a gating mechanism is introduced to control the path of sentiment information flowing to the aggregation layer. An improved krill swarm algorithm based on cosine control factor and Cauchy factor is proposed to solve the model. Finally, the full connection layer and argmax function are used to achieve sentiment classification. The experimental results show that compared with other advanced models, the proposed model can improve the accuracy of emotion classification by 2.3% and 0.8% respectively on two public data sets of IMDB and Yelp2014, and obtain the minimum root mean square error (RMSE).

KEYWORDS

Explainable Art Sentiment Analysis, Gating Mechanism, Interior Decoration Environment, Krill Herd Algorithm, Long and Short-Term Memory Network

INTRODUCTION

Design formal language symbols are divided into different forms. First of all, the information in graphic design works generally contains the hierarchy of primary and secondary, first subject and comprehensive, and then local and detailed. No matter the whole or part of information, they need to be displayed through the extension of the medium of design formal language symbols. Generally, the whole information symbol is composed of various language symbols with different or the same part of information through layout and reorganization. For example, points, lines and planes are called basic language symbols, which can form overall information such as images, characters and

DOI: 10.4018/JCIT.324602
*Corresponding Author

colors (Wang et al. 2021; Ji et al. 2021). Secondly, the design of formal language symbols also has personal subjective color. It can not only promote creators to carry out rich emotional expression of the picture, but also build a bridge of communication and understanding between creators and viewers. As a basic element of graphic design, how to artistic creation of formal language symbols should be highly valued.

For graphic design, point, line and surface are the main framework and basis of design. The use of formal language symbols of overall design is based on basic language symbols of point, line and surface. The composition of formal language symbols of overall design is also composed of elements of basic design formal language symbols. It also requires the overall design of formal language symbols to carry out information fusion (Meng et al. 2023). Graphic design graphics, images, text, color is the most important, the most common overall design formal language symbol, the overall formal language symbol is an indispensable element in graphic design, and also has a very important aesthetic significance. In the process of design practice, through the artistic processing of graphics, text, color, etc., the points, lines and surfaces that were originally independent and scattered in the picture are integrated to fill the whole visual space, making the design work more rich and concise, making the scattered layout form a whole object image, and the information of the work becomes integrated from the scattered.

Sentiment Analysis, which analyzes people's emotions or opinions based on the text they generate, has long been one of the most active areas of research in natural language processing. Identifying the underlying emotions expressed in a text is crucial to understanding the full meaning of the text. With the rapid development of social media platforms such as Weibo, Zhihu and Toutiao, people are increasingly sharing their views and opinions online. Sentiment analysis has attracted a lot of attention (He et al. 2022), because the opinions or emotions detected in the text are of great help to product recommendation, public opinion analysis, market prediction and so on.

The goal of document-level emotion analysis is to judge the emotion expressed by the whole document, such as a film review or a comment on a certain hot news. As long as the text to be analyzed exceeds the scope of a sentence, it can be regarded as document-level emotion analysis. A prerequisite for document-level sentiment analysis is the assumption that the opinions expressed in the whole text are directed at a single entity and contain the views of only one opinion holder.

In the traditional task of artistic emotion analysis, most models regard emotion analysis as a classification problem consisting of feature extraction and classifier training (Zehra et al. 2021). Initially, machine-learning-based methods using supervised classification or regression were used to train text models from polarity markers (Castellano et al. 2021). However, the performance of these models is largely dependent on a large number of manually processed features, such as affective dictionaries and other features with specific meanings.

With the proposed deep learning method, the performance of sentiment analysis model has been further improved. The most widely used Neural Network models in the field of emotion analysis include Convolutional Neural Network, (CNN), Bidirectional Long Short-Term Memory (BiLSTM) network (Mei et al. 2021), etc.

Document-level sentiment analysis mainly focuses on generating rich document representation and individuation to improve the performance of classification model. A person writing a critical text is subjectively biased towards his/her own perception. Forgiving users tend to give higher ratings than critical users, even if they review the same products, and popular products may receive more praise than less popular ones. Therefore, the model must consider the influence of user and product information on text sentiment classification. (Wan et al. 2022) used CNN as the basic encoder and combined user and product information into the emotion classification neural network model for the first time, which greatly improved the performance of the emotion analysis model.

In recent years, researchers have begun to build attention-based models to enhance document presentation, so as to highlight important words or sentences in a paragraph of text. It has become mainstream to construct different models of attention using information embedded in the text, including user information, product information, and review text data in the local context. (Cheng et al. 2020) proposed a hierarchical neural network to model document semantics through hierarchical structure, and introduced an attention mechanism to integrate user product information into attention to propose user product attention. The introduction of user product attention had greatly improved the performance of document-level sentiment analysis model. (Gaye et al. 2021) used basis vectors to integrate user product information into the classifier of the model. When the number of classification features was large, the use of basis vectors significantly reduced the number of parameters. (Sarkar et al. 2020) also used hierarchical structure to model document semantics, and by integrating user information and product information into multi-attention, the model obtained the influence of different users and products on emotional score in multiple sub-spaces.

In order to assign appropriate labels to the text, the model should also capture the core semantic units of more advanced information than the word level information in the source text, and then assign text labels based on its understanding of the semantic units. Since traditional attention mechanisms only focus on extracting word-level information containing redundant and irrelevant details, it is difficult to extract more critical information from semantic units. To solve this problem, this paper proposes a novel long and short term memory network-based krill herd algorithm for explainable art sentiment analysis in interior decoration environment. In this paper, the model uses global user preferences and product features to learn the review text, and extracts text representations of different scales for emotion classification. To verify the validity of the model, three evaluation data sets from IMDB and Yelp(including Yelp2013 and Yelp2014) were evaluated. Experimental results show that this model can outperform the benchmark model by a large margin. Compared with the most advanced model in the benchmark model, the Accuracy of the proposed model is improved by 2.3% on IMDB and 0.8% on Yelp2014 data sets, respectively, and the minimum root-mean-square error is obtained on IMDB and Yelp2013 data sets.

The main work of this paper consists of the following three points:

1. Using multi-scale convolution attention to encode the text. Firstly, multi-scale convolution is used to extract multi-granularity short-range local semantic information between document words, and then richer document representation at different levels is obtained through user product attention.
2. We propose an improved krill swarm algorithm based on cosine control factor and Cauchy factor to accurately extract features from the model.
3. Gated Unit is introduced to construct a new gated unit GTUU(Gate Tanh Update Unit) to control the path of emotion information flowing to the aggregation layer. Experiments prove that GTUU is more effective in document-level emotion analysis.

PROPOSED MODEL

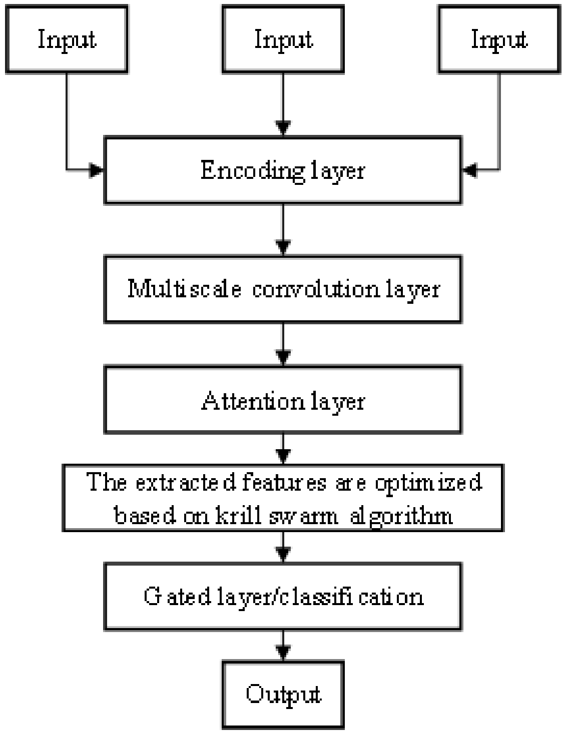
The framework of art emotion analysis model based on multi-scale convolution and gating mechanism proposed in this paper is shown in Figure 1. The model in this paper consists of six parts: coding layer, multi-scale convolution layer, feature extraction layer, attention layer, gating layer and classification layer.

Encoding Layer

First it defines a document $doc = \left\{x_1, x_2, \cdots, x_j, \cdots, x_n\right\}$. x_j is the j-th word in doc. n is the length of the statement. They are then initialized with pre-trained word embedding vectors, which are fine-tuned during the training phase. All words are embedded in the corresponding vector w_j through a word embedding matrix.

In this paper, BiLSTM, an encoder with good performance in sentiment analysis of long documents, is used to learn basic document representation. BiLSTM obtains the feature representation of the word by summarizing the information from both directions of the word, and combines the

Figure 1. Proposed structure for sentiment analysis



context information into the feature representation. Since the forward and backward Long Short-Term Memory (LSTM) networks (Wang et al. 2022) look similar, only the calculation process of the forward LSTM is given for brevity, as shown in equations (1)~(3):

$$\begin{bmatrix} i_i \\ f_i \\ o_i \\ \hat{c}_i \end{bmatrix} = \begin{bmatrix} \sigma \\ \sigma \\ \sigma \\ tanh \end{bmatrix} \left(W_{en} \left[\vec{h}_{i-1}, w_i \right] + b_{en} \right) \tag{1}$$

$$c_i = f_i \odot c_{i-1} + i_i \odot \hat{c}_i \tag{2}$$

$$\vec{h}_{i-1} = o_i \odot tanh\left(c_i\right) \tag{3}$$

Where i_i, f_i and o_i are the gate activation function. $\odot$ indicates the dot product operator. σ is the sigmoid function. $tanh$ represents the tangent function of Hyperbolic. W_{en} and b_{en} are parameters that need to be trained.

Then, the forward and backward hidden states are spliced together to form a new word vector $h_i = \left[\vec{h}_i, \overline{h}_{n-i}\right]$. The document matrix is expressed as $H = \left[h_1, h_2, \cdots, h_n\right]$ after the encoding layer.

Multiscale Convolution Layer

Inspired by the idea of global coding of text summaries, this paper uses convolutional neural networks (CNN) to capture local interactions between words and generate information representations of a higher level than words, such as phrases or short sentences. MultiScale Convolution (MSC) uses multiple one-dimensional convolution filters with convolution kerns of different widths to capture local dependencies of different sizes, in order to obtain more semantic information of different levels.

Specifically, in the q-th convolution filter, a single feature m_i^q is formed by the convolution operation of each word with k^q step length using the filter in the receptive field, as shown in equation (4):

$$m_i^q = f^q\left(H_{i:i+k^q} \times W_c^q + b_c^q\right) \tag{4}$$

Where $i = \left[0, k^q, 2k^q, \cdots, \left(n / k^q - 1\right)k^q\right]$. $f^q()$ is a nonlinear function. W_c^q and b_c^q are the parameters to be learned.

Finally, a new document matrix $M^q = \left[m_1^q, m_2^q, \cdots, m_{L^q}^q\right]$ ($L^q = n / k^q$) is formed by the feature vectors of the convolved semantic units.

Attention Layer

Although the multi-scale convolution layer is used to condenses the document at different scales, not all semantic units contribute equally to the emotional expression meaning of the document. Attention calculation must be carried out to assign different weights to different semantic units (Guo et al. 2022; Qi et al. 2022). Moreover, for different users and different products, There are also differences in the use of emotional words. Therefore, it is necessary to integrate user information and product information into the calculation of attention, so as to extract emotional information in the text more accurately.

The attention layer calculates user product attention for each document, and the specific calculation process is shown in Equations (5)~(7):

$$e_i^q = v^T tanh\left(W_{att}^q m_i^q + W_u^q u + W_p^q p + b_{att}^q\right) \tag{5}$$

$$\alpha_i^q = \frac{exp\left(e_i^q\right)}{\sum_{k=1}^{L^q} exp\left(e_i^q\right)} \tag{6}$$

$$d^q = \sum_{i=1}^{L^q} \alpha_i^q m_i^q \tag{7}$$

Where m_i^q represents the i-th semantic unit of the q-th document. e_i^q is its scoring function. W_{att}^q, W_u^q, W_p^q, b_{att}^q and v^T are the corresponding weight matrices and vectors to be trained. α_i^q is the weight of attention after considering user product information, which measures the importance of m_i^q in the whole document.

A weighted sum of α_i^q and the corresponding m_i^q gives the characteristic representation of the entire document d^q. Since the contribution of document representation d^q of different scales to the final document representation cannot be the same, simple stitching cannot be carried out directly, so different weights should be assigned:

$$h = \sum_{q=1}^{N_q} W_d^q d^q + b_d \tag{8}$$

Where N_q represents the number of document representations of different scales. W_d^q represents the weight matrix to be trained for the q-th scale document. b_d is bias.

Gating Layer

The gating mechanism Tanh-ReLU proposed in language modeling has achieved good results. The two most famous types of gated units are Gate Tanh Unit (GTU) and Gate LinerUnit (GLU). GTU is represented by $tanh\left(Wx+b\right)\odot\sigma\left(Vx+b\right)$, while GLU is replaced by $\left(Wx+b\right)\odot\sigma\left(Vx+b\right)$, so that the gradient is not reduced and more information can be transmitted. In this paper, a new Gate Tanh Update Unit (GTUU) is constructed to control the path of emotional information flowing to the aggregation layer, and experiments prove that the gating mechanism is effective in artistic emotion analysis.

$$g = GTUU\left(h\right) \tag{9}$$

Specific calculation process of GTUU gated unit is as follows:

$$\hat{h} = tanh\left(W_{\hat{h}}h + b_{\hat{h}}\right) \tag{10}$$

$$\lambda = \sigma\left(W_{\lambda}h + b_{\lambda}\right) \tag{11}$$

$$\mathrm{g} = \mathrm{ReLU}\left(\lambda\odot\hat{h} + \left(1-\lambda\right)\odot h\right) \tag{12}$$

Where $W_{\hat{h}}$, W_{λ} are trainable weights. $b_{\hat{h}}$, b_{λ} are the deviation value. $\odot$ indicates the dot product operator. Rectified Linear Unit (ReLU) is a linear rectified function that has a linear path allowing gradients to pass easily through active units (Jung et al. 2021).

Improved Krill Swarm Algorithm

Gandomi et al. (2012) first proposed the krill swarm algorithm. It is a swarm intelligent optimization algorithm designed by observing the daily behavior characteristics of krill populations. The initial population distribution is random, and the individual iteration process is mainly influenced by the induction of nearby krill, foraging behavior, and random physical diffusion.

The mathematical model of krill movement can be expressed by Lagrange model in n-dimensional decision space.

$$\frac{dx_i}{dt} = N_i + F_i + D_i \tag{13}$$

Where N_i is the influence of nearby krill population. F_i is foraging behavior. D_i is the random diffusion behavior of the i-th krill individual, then

$$N_i = N^{max}\alpha_i + w_n N_i^o \tag{14}$$

$$F_i = V_f\beta_i + w_f F_i^o \tag{15}$$

$$D_i = D^{max}\left(1-\frac{t}{t_{max}}\right)\delta \tag{16}$$

Where N^{max}, V_f and D^{max} are the maximum induction velocity, maximum foraging velocity and maximum diffusion velocity, respectively. α, β and δ are induction direction, foraging direction and diffusion direction, respectively. w_n and w_f are foraging weight and induction weight respectively. t and t_{max} are the current and maximum iterations, respectively. The position updating formula of krill individuals in the interval from t to $t+\Delta t$ is:

$$x_i\left(t+\Delta t\right)=x_i\left(t\right)+\frac{dx_i}{dt}\left(\Delta t\right) \tag{17}$$

$$\Delta t=C_t\sum\nolimits_{j=1}^{N_V}\left(UB_j-LB_j\right) \tag{18}$$

Where Δt is the scaling factor of the velocity vector. C_t is the step scaling factor, which is a constant in the interval [0,2]. N_V is the number of variables. UB_j, LB_j are the upper and lower bounds of the j-th variable respectively.

The reactive power optimization model is dynamic and multi-constraint. Conventional krill population algorithm can not make full use of the historical population and the historical optimal solution information when the optimization time domain changes, and can not quickly build the initial population with the change of the environment. Therefore, we improve the krill swarm algorithm through the following three aspects.

1. Population initialization based on Sobol sequence

Li et al. (2018) proposed to solve the problem of uncertain distribution of feasible solutions. The distribution of initial population data should be evenly distributed in the data space as far as possible, which not only enables the population to have a high diversity, but also enhances the reliability of the solution set. The traditional krill swarm algorithm generates initial krill individuals in the form of random numbers, which cannot evenly distribute initial krill individuals in the data space. The random distribution may make the population distribution near the optimal solution sparse, thus affecting the calculation effect and efficiency, and the over-aggregation of the population will also lead to the local optimal situation. The use of Sobol sequence can produce ultra-uniform distribution particles, improve the initial population diversity of krill, and make krill individuals evenly distributed in the space.

The Sobol sequence generates initial krill colonies as follows:

$$x_i=x_{min}+\lambda\left(x_{max}-x_{min}\right) \tag{19}$$

Where x_{max} and x_{min} represent the upper and lower limits of krill position respectively. λ is the random number in the range [0,1] generated by the Sobol sequence.

2. Introducing the cosine control factor

It can be seen from equations (14) and (15) that the induced motion weight w_n and foraging action weight w_f have great influence on the movement of krill individuals. A cosine decreasing strategy of foraging weight and induced weight is proposed to improve the global search power of the algorithm.

$$w_n = w_f = \left(w_{max} - w_{min}\right) cos \frac{\pi t}{2N} + w_{min} R \tag{20}$$

Where t is the current iteration number and N is the maximum iteration number. w_{max} and w_{min} are the maximum and minimum values of induction weight and foraging weight respectively. If the weight value of w in the early stage is too large, the algorithm can carry out global search better. In addition, with the increasing number of iterations, the weight value of w gradually decreases, and a smaller range of fine search will be carried out at this time. The introduction of random number R can not only improve the randomness of w, so that the improved algorithm can effectively adapt to the constant change of search conditions over time, but also improve the global search ability.

3. Introducing Cauchy variation factor

The traditional krill population optimization algorithm has the disadvantages of prematurity and local optimum. To solve these problems, we introduced Cauchy variation factors to reduce the possibility of krill individuals entering the local optimum when studying the position renewal of krill individuals. The main principle is as follows: large area local search of potential optimal krill individuals is carried out through Cauchy variation factor, and random interference is generated within the range of certain potential optimal solutions, which can effectively increase the local search ability of krill swarm algorithm. The Cauchy variation probability density function is as follows:

$$g\left(x; x_0, \gamma\right) = \frac{1}{\pi\gamma\left[1 + \left(\left(x - x_0\right) / \gamma\right)^2\right]} \tag{21}$$

Where x_0 is a position parameter. γ is a random variable greater than 0, and x is a real number. x_0 =0, γ =1, which is the Cauchy distribution rule. Through the analysis of its probability density function, we know that its mean and variance are not specific values. But both the mode and the median are equal to the positional parameter, which is x_0. Its distribution function is as follows.

$$\mathrm{G}\left(\mathrm{x}\right) = 0.5 + \frac{1}{\text{À}} \arctan\left(\mathrm{x}\right) \tag{22}$$

Using Cauchy variation factor to improve the perturbation formula as follows:

$$x_{ibest}\left(t\right) = x_i\left(t\right) + x_i\left(t\right) G\left(x\right) \tag{23}$$

$$x_i\left(t+1\right) = \begin{cases} x_{ibest}\left(t\right), f\left(x_{ibest}\left(t\right)\right) > f\left(x_i\left(t\right)\right) \\ x_i\left(t\right), \qquad f\left(x_{ibest}\left(t\right)\right) \le f\left(x_i\left(t\right)\right) \end{cases} \tag{24}$$

Where $f\left(x_i\left(t\right)\right)$ is the fitness value of i-th krill in the t-th iteration. By introducing Cauchy variation factor, the improved krill swarm algorithm can jump out of the local optimal.

Classification Layer

The document vector g is a high-level representation of the document and can be used as a feature of the document sentiment classification. It maps the document representation g to the target space of class C using a linear layer:

$$p = argmax\left(W_c g + b_c\right) \tag{25}$$

Where C is the number of emotional categories. p is the final affective prediction. $argmax()$ is used to obtain the affective category that maximizes $f\left(\cdot\right)$. W_c and b_c are weight to be trained and bias respectively.

EXPERIMENT AND RESULT ANALYSIS

The experiment was conducted to verify the validity of the model on three public sentiment analysis data sets (IMDB, Yelp2013, Yelp2014) with user product information constructed by Tang et al. (2015), in which IMDB is the movie review data set, Yelp2013 and Yelp2014 are the commodity review data sets.

The statistics of the three data sets are shown in Table 1. The data set was divided into training set, verification set and test set in a ratio of 7:2:1. Stanford CoreNLP was used for tokenization and sentence splitting. Accuracy and root mean square error (RMSE) are used to measure the reliability of the model. Accuracy measures overall emotion classification performance. The root mean square error describes the degree of dispersion between the predicted emotion and the true emotion.

Experiments in this paper are carried out on NVIDIA 2080Ti GPU based on PyTorch deep learning framework. In the experiment, word vectors were initialized using the pre-trained GloVe word embedding vector. The dimension of GloVe, the input dimension of bidirectional LSTM (150 dimensions in the hidden state forward and backward) and the number of input and output channels of one-dimensional convolution filter are set to 300 dimensions. The multi-scale convolution layer adopts three parallel one-dimensional convolution operations, and the convolution kernels are respectively 1, 2 and 4. The convolution step size is the same as the convolution kernels size. To prevent over-fitting, one Dropout is made after each layer, and the Dropout rate is 0.1. The Batch Size of the experiment is set to 64. The optimization of model hyper-parameters is completed on the Adadelta optimizer. The model is trained and adjusted on the training set, the model parameters with the best performance are selected on the verification set, and finally tested on the test set.

In order to verify the validity of the model proposed in this paper, the model will be compared with the following benchmark models, ABCDM (Basiri et al. 2021), AKEG (Liang et al. 2022), BiERU (Li et al. 2022), WWE (Onan et al. 2021), DGCNN (Li et a. 2021). The Accuracy value and RMSE value were calculated on the test set in the experiment, and the comparative experimental results of each model are shown in Table 2. In order to verify the effectiveness of multi-scale convolution and gated units, ablation experiments were conducted in this paper, and the experimental results

Table 1. Statistics of IMDB, Yelp2013, and Yelp2014 data sets

Data	Category Number	Comment Number	Number of Users	Number of Products	Comments/ Users	Comments/ Product
IMDB	10	84921	1312	1637	64.81	51.93
Yelp2013	5	231165	4820	4196	48.41	48.35
Yelp2014	5	78968	1633	1635	47.96	55.12

are shown in Table 3. The model proposed in this paper is MSC-{GTU,GLU,GTUU}, where MSC represents multi-scale convolution, NoGate means no gated unit is used in the gated layer, NoMSC means no multi-scale convolution layer is removed, and no up means no user product information is fused into the model. GTU and GLU represent the two gated units proposed by Yann et al., while GTUU represents the gated unit proposed in this paper. In this paper, three different gating units are used in the gating layer to do comparative experiments.

As can be seen from Table 2, the accuracy of MSC-GTU and MSC-GLU in IMDB and Yelp2014 both exceed the value with the best performance in the benchmark model. MSC-GTU increases by 0.2% and 0.5%, and MSC-GLU increases by 0.4% and 0.6%, respectively. Compared with MSC-NoGate, the accuracy of MSC-GTU on Yelp2013 and Yelp2014 data sets is increased by 0.2%. The accuracy of MSC-GLU in IMDB, Yelp2013 and Yelp2014 data sets increases by 0.2%, 0.1% and 0.3%, respectively. The accuracy of MSC-GTUU in IMDB, Yelp2013 and Yelp2014 data sets increases by 1.0%, 0.8% and 0.4%, respectively. The introduction of gated unit is helpful to improve the performance of sentiment analysis.

In Table 2, comparing the performance of models using gated units on three data sets, the model MSCGTUU using GTUU gated units proposed in this paper is significantly better than MSC-GTU and MSC-GLU. Compared with MSC-GTU, the accuracy of MSC-GTUU on IMDB, Yelp2013 and Yelp2014 data sets increased by 1.0%, 0.6% and 0.2%, respectively.

The results of comparison experiments between MSC-GTUU (no up) and MSC-GTUU showed that the model integrating user product information has higher accuracy than the model considering only comment text information. Compared with MSC-GTUU (no up) without considering user product information, the accuracy of MSC-GTUU model on IMDB and Yelp2014 increases by 6.3%, 3.0% and 3.8%, respectively. It shows the importance of integrating user product information to text sentiment analysis.

Compared with MSC-GLU, the accuracy of MSC-GTUU on IMDB, Yelp2013 and Yelp2014 data sets is increased by 0.8%, 0.7% and 0.1%, respectively. Experimental results show that the proposed GTUU unit is more effective in document-level sentiment analysis.

At the same time, in order to verify the effectiveness of multi-scale convolution joint coding, different convolution filters are also used in the multi-scale convolution layer as substitutes, and comparison experiments are conducted on IMDB data sets. The experimental results are shown in Table 4. In Table 4, Conv(n) represents the use of a one-dimensional convolution filter with convolution kernel size and convolution step size of n in the multi-scale convolution layer to extract semantic information of adjacent n words as an emotional semantic unit. Conv(a,b,c) means that three one-dimensional convolution filters with convolution widths of a,b and c are used simultaneously to extract information of different scales for text joint encoding.

Table 2. Comparison of accuracy and RMSE of different models

Model	IMDB		Yelp2013		Yelp2014	
	Accuracy/%	RMSE	Accuracy/%	RMSE	Accuracy/%	RMSE
ABCDM	44.6	1.613	60.7	0.759	61.9	0.757
AKEG	48.7	13.99	63.4	0.7125	64.6	0.709
BiERU	47.6	1.403	63.4	0.725	64.6	0.701
WWE	49.9	1.462	65.0	0.705	65.0	0.699
DGCNN	54.4	1.292	66.1	0.703	67.8	0.665
MSC-GTU	55.8	1.202	67.4	0.684	69.2	0.654
MSC-GLU	56.1	1.211	67.3	0.679	69.3	0.651
MSC-GTUU	58.8	1.176	67.9	0.668	69.4	0.660

Table 3. Ablation experiment of Multi-scale convolution, gated units and user product information

Model	IMDB		Yelp2013		Yelp2014	
	Accuracy/%	RMSE	Accuracy/%	RMSE	Accuracy/%	RMSE
noMSC-GTUU	55.2	1.244	66.9	0.694	68.3	0.678
MSC-noGate	55.8	1.272	67.2	0.689	69.0	0.657
MSC-GTUU(no up)	50.5	1.392	65.0	0.708	65.6	0.691
MSC-GTUU	56.8	1.176	68.0	0.668	69.4	0.661

Table 4. Comparison of experimental results of different convolution filters

Convolution Scale	Accuracy/%	RMSE
Conv(1)	55.8	1.209
Conv(2)	55.4	1.241
Conv(4)	54.8	1.278
Conv(1,2,4)	56.8	1.176

The convolution width of the convolution layer will directly affect the basic granularity of the text semantic unit (word level or n-bit phrase) for the attention layer to compute the attention of the document. As shown in experimental Conv(1), Conv(2) and Conv(4) in Table 4, when a single convolution wave is used in the convolution layer, the accuracy rate decreases with the increase of the convolution width. The accuracy of Conv(1,2,4) in experiment (1,2,4) is increased by 1.0% by using the multi-scale convolution joint encoding with convolution kernel sizes of 1, 2, and 4, respectively, compared with the best word-level encoding in the corresponding single-scale coding experiment (1). The experimental results show that the generation of rich documents with different granularity through multi-scale convolution is effective for the improvement of accuracy.

CONCLUSION

In this paper, an attention art emotion analysis network model based on multi-scale convolution and gating mechanism is proposed. Three convolutional layers with different convolution sizes are used to model text information respectively, and richer contextual information with different granularity is obtained. Then, semantic units with higher relevance are selected by user product attention to generate document representation. Gating mechanism is introduced into document-level sentiment analysis, and a new gating unit, GTUU, is proposed to achieve better performance. We will then consider how best to use user product information to improve the generalization of the model.

CONFLICTS OF INTEREST

The authors declare that there is no conflict of interest regarding the publication of this paper.

REFERENCES

Basiri, M. E., Nemati, S., Abdar, M., Cambria, E., & Acharya, U. R. (2021). ABCDM: An attention-based bidirectional CNN-RNN deep model for sentiment analysis. *Future Generation Computer Systems*, *115*, 279–294. doi:10.1016/j.future.2020.08.005

Castellano, G., & Vessio, G. (2021). Deep learning approaches to pattern extraction and recognition in paintings and drawings: An overview. *Neural Computing & Applications*, *33*(19), 12263–12282. doi:10.1007/s00521-021-05893-z

Cheng, K., Yue, Y., & Song, Z. (2020). Sentiment classification based on part-of-speech and self-attention mechanism. *IEEE Access : Practical Innovations, Open Solutions*, *8*, 16387–16396. doi:10.1109/ACCESS.2020.2967103

Gandomi, A. H., & Alavi, A. H. (2012). Krill herd: A new bio-inspired optimization algorithm. *Communications in Nonlinear Science and Numerical Simulation*, *17*(12), 4831–4845. doi:10.1016/j.cnsns.2012.05.010

Gaye, B., Zhang, D., & Wulamu, A. (2021). Improvement of support vector machine algorithm in big data background. *Mathematical Problems in Engineering*, *2021*, 1–9. doi:10.1155/2021/5594899

Guo, M. H., Xu, T. X., Liu, J. J., Liu, Z.-N., Jiang, P.-T., Mu, T.-J., Zhang, S.-H., Martin, R. R., Cheng, M.-M., & Hu, S.-M. (2022). Attention mechanisms in computer vision: A survey. *Computational Visual Media*, *8*(3), 331–368. doi:10.1007/s41095-022-0271-y

He, M., Ma, C., & Wang, R. (2022). A Data-Driven Approach for University Public Opinion Analysis and Its Applications. *Applied Sciences (Basel, Switzerland)*, *12*(18), 9136. doi:10.3390/app12189136

Ji, C., Wu, F., Zhu, Z., Chang, L.-P., Liu, H., & Zhai, W. (2021). Memory-efficient deep learning inference with incremental weight loading and data layout reorganization on edge systems. *Journal of Systems Architecture*, *118*, 102183. doi:10.1016/j.sysarc.2021.102183

Jung, H. C., Maly, J., Palzer, L., & Stollenwerk, A. (2021). Quantized compressed sensing by rectified linear units. *IEEE Transactions on Information Theory*, *67*(6), 4125–4149. doi:10.1109/TIT.2021.3070789

Li, R., Chen, H., & Feng, F. (2021). Dual graph convolutional networks for aspect-based sentiment analysis. *Proceedings of the 59th Annual Meeting of the Association for Computational Linguistics and the 11th International Joint Conference on Natural Language Processing (Volume 1: Long Papers),* 6319-6329.

Li, W., Shao, W., Ji, S., & Cambria, E. (2022). BiERU: Bidirectional emotional recurrent unit for conversational sentiment analysis. *Neurocomputing*, *467*, 73–82. doi:10.1016/j.neucom.2021.09.057

Li, Y., Yang, Z., Li, G., Zhao, D., & Tian, W. (2018). Optimal scheduling of an isolated microgrid with battery storage considering load and renewable generation uncertainties. *IEEE Transactions on Industrial Electronics*, *66*(2), 1565–1575. doi:10.1109/TIE.2018.2840498

Liang, B., Su, H., Gui, L., Cambria, E., & Xu, R. (2022). Aspect-based sentiment analysis via affective knowledge enhanced graph convolutional networks. *Knowledge-Based Systems*, *235*, 107643. doi:10.1016/j.knosys.2021.107643

Mei, S., Li, X., & Liu, X. (2021). Hyperspectral image classification using attention-based bidirectional long short-term memory network. *IEEE Transactions on Geoscience and Remote Sensing*, *60*, 1–12.

Meng, X., Wang, X., Yin, S., & Li, H. (2023). Few-shot image classification algorithm based on attention mechanism and weight fusion. *Journal of Engineering and Applied Sciences (Asian Research Publishing Network)*, *70*(1), 14. doi:10.1186/s44147-023-00186-9

Onan, A. (2021). Sentiment analysis on product reviews based on weighted word embeddings and deep neural networks. *Concurrency and Computation*, *33*(23), e5909. doi:10.1002/cpe.5909

Qi, J., Liu, X., Liu, K., Xu, F., Guo, H., Tian, X., Li, M., Bao, Z., & Li, Y. (2022). An improved YOLOv5 model based on visual attention mechanism: Application to recognition of tomato virus disease. *Computers and Electronics in Agriculture*, *194*, 106780. doi:10.1016/j.compag.2022.106780

Sarkar, M., Aggarwal, M., & Jain, A. (2020). Document structure extraction using prior based high resolution hierarchical semantic segmentation. In *Computer Vision–ECCV 2020: 16th European Conference, Glasgow, UK, August 23–28, 2020, Proceedings, Part XXVIII 16*. Springer International Publishing.

Tang, D., Qin, B., & Liu, T. (2015). Document modeling with gated recurrent neural network for sentiment classification. *Proceedings of the 2015 Conference on Empirical Methods in Natural Language Processing,* 1422-1432.

Wan, C. X., & Li, B. (2022). Financial causal sentence recognition based on BERT-CNN text classification. *The Journal of Supercomputing*, *78*(5), 1–25. doi:10.1007/s11227-021-04097-5

Wan, L., Huebl, A., Gu, J., Poeschel, F., Gainaru, A., Wang, R., Chen, J., Liang, X., Ganyushin, D., Munson, T., Foster, I., Vay, J.-L., Podhorszki, N., Wu, K., & Klasky, S. (2021). Improving I/O performance for exascale applications through online data layout reorganization. *IEEE Transactions on Parallel and Distributed Systems*, *33*(4), 878–890. doi:10.1109/TPDS.2021.3100784

Wang, L., Yin, S., & Hashem, A. (2022). A novel deep learning-based single shot multibox detector model for object detection in optical remote sensing images. *Geoscience Data Journal*. Advance online publication. doi:10.1002/gdj3.162

Zehra, W., Javed, A. R., Jalil, Z., Khan, H. U., & Gadekallu, T. R. (2021). Cross corpus multi-lingual speech emotion recognition using ensemble learning. *Complex & Intelligent Systems*, *7*(4), 1–10. doi:10.1007/s40747-020-00250-4