Research Article

Sentiment Classification of Educational Tourism Reviews Based on Parallel CNN and LSTM with Attention Mechanism

Ying Wang,1,2 Chengxi Chu,3 and Tian Lan4,5

1Tourism College, Sichuan University, Chengdu, Sichuan 610064, China
2The Engineering & Technical College of Chengdu University of Technology, Leshan, Sichuan 614000, China
3Faculty of Electrical Engineering and Computer Science, Ningbo University, Ningbo, Zhejiang 315201, China
4College of Art and Design, Changchun University of Technology, Changchun, Jilin 130000, China
5College of Management, Shih Chien University, Taibei 10462, Taiwan

Correspondence should be addressed to Ying Wang; wy808212@163.com

Received 18 May 2022; Revised 13 July 2022; Accepted 4 August 2022; Published 9 September 2022

Academic Editor: Shadi Aljawarneh

Copyright © 2022 Ying Wang et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

With the rapid development of the Internet and tourism, the Internet has been widely used in the tourism industry. Tourism enterprises and tourists use the Internet to publish and obtain travel-related information. Educational tourism is a new type of tourism activity. As a combination of “tourism + education,” it has gradually attracted the attention of tourists. With its convenience, fast speed, and low barrier, tourism text data provide great convenience for tourists’ sentiment calculation and have become one of the main sources of big data for tourism. However, the reviews of educational tourism have a lot of redundant information and complex sentence patterns, leading to a relatively low classification accuracy of the existing sentiment analysis algorithms. In order to effectively obtain the implicit semantic information of short text reviews for sentiment orientation recognition, a sentiment classification model for educational tourism online reviews based on parallel CNN and LSTM with multichannel attention mechanism is proposed. Firstly, Word2Vec technique is used, and based on noise word filtering, the feature words of educational tourism reviews are extracted to preprocess the input data set. Then, parallel CNN and LSTM are used to extract text local information and contextual features, and a multichannel attention mechanism is used to extract the attention values from the LSTM output. Finally, the output information of the multichannel attention mechanism is fused to effectively extract text features and focus on important words. The experimental results show that compared with other advanced methods, the proposed algorithm achieves improvements in terms of precision, recall, and $F_1$ value and improves the AUC performance. It will help the educational tourism bases to carry out targeted development and construction in response to tourists feedback, enhance the sense of gain and happiness of tourists in educational tourism activities, improve tourism experience, and promote the rapid development of high-quality educational tourism.

1. Introduction

“Big data” refers to massive amounts of information and data. Tourism big data are data generated by tourism practitioners and tourists, including data generated by tourist attractions, hotels, travel agencies, and tour operators as well as data in other fields related to tourism, such as economic data and traffic data [1]. Among them, the data generated by tourists have the greatest application value [2]. Tourism big data come from a wide range of sources. Through the mining of tourism big data, information with research value related to tourism flow, tourism economy, and tourism resources can be obtained. The wide application of big data creates opportunities for the development of experiential tourism [3]. Sentiments are the attitudes and experiences that people have about whether objective things meet their needs. Tourist sentiments refer to the pleasure, excitement, sadness, anger, regret, and other emotional experiences of tourists who are influenced by personal factors or external environment in tourism activities and whether the tourism activities meet the basic needs and social needs of individuals. Therefore, the sentiments of
tourists are diverse and volatile [4]. These sentiments not only constitute an important travel experience for tourists but also have an important impact on travel motivation, satisfaction, behavioral intentions, and interpersonal interactions [5]. During travel, tourists obtain information and share their travel experiences through online platforms and social media. The texts, images, audios, and videos uploaded by tourists have become the main data sources of tourism big data. Among them, the text content has the advantages of convenience, simplicity, intuitiveness, and fast speed with a low barrier for entry, which provides convenience for tourists to express their emotions and exchange information, and occupies an increasingly important position in tourism big data [6]. Through the mining of text data, it can provide decision support for tourism planning and marketing, making sentiment analysis in tourism big data a hot issue in the field of tourism research [7, 8].

As a combination of “tourism + education,” educational tourism combines tourism resources with quality education and has become a new star in today’s global tourism development. Educational tourism is a comprehensive tourism product developed based on certain tourism resources. Compared with general tourism products, it has the characteristics of rich connotation, wide-ranging category, and strong comprehensiveness. The development of educational tourism helps to integrate local tourism resources, promote the optimal combination of tourism products, the transformation and upgrading of tourism formats, and the comprehensive development of tourism destinations, which is bound to promote and deepen the development of all-for-one tourism [9].

The emotional tendency of educational tourism review information is an important basis for travel planning of tourists [10]. In recent years, the research on sentiment analysis in the field of tourism is not deep enough, and a lot of works focus on online product reviews, movie reviews, short texts on social media, etc., without considering the characteristics of educational tourism texts themselves. Educational tourism texts usually include multiple indicators such as scenery, ticket prices, accommodation quality, services, and interests. The complexity and diversity of texts lead to inefficiencies and misclassification of existing algorithms, making it difficult for tourists to obtain effective information [11].

Educational tourism base is the basis for carrying out educational tourism, and the Internet is an important medium for tourists to search for information [12]. In this paper, through the sentiment analysis of educational tourism reviews, it can be determined that whether the educational needs of tourists have been met. On the one hand, it is helpful to understand the changing laws of the tourist source market, and then, it can provide guidance and assistance for the development and construction, marketing management, and actual passenger flow forecasting of the educational tourism base. On the other hand, it is beneficial to grasp the supply and demand relationship between the actual development of the educational tourism base and the needs of tourists and lay the foundation for optimizing its spatial structure. By analyzing the characteristics of online reviews of educational tourism bases, the types of educational tourism bases that are popular with tourists can be concluded. In response to these analysis results, the educational tourism bases can make accurate and timely response measures to the actual passenger flow and provide tourists with high-quality tourism services. In addition, it can also enrich the types of educational tourism bases in the future base construction, in order to enhance people’s sense of acquisition and happiness in tourism activities, improve tourism experience, and promote the rapid development of high-quality tourism [13].

In this paper, a large-scale data set of educational tourism reviews is constructed, and a sentiment classification model for educational tourism reviews based on the multichannel attention mechanism of convolutional neural network (CNN) and long short-term memory (LSTM) is proposed. In the proposed model, CNN can effectively extract local key information when extracting features. Compared with RNN, LSTM can process long text more effectively, alleviate the gradient problem, and can better extract contextual semantic information. The attention mechanism can focus on words that have a greater impact on the final result by setting different weights. As a result, the sentiment classification performance of educational tourism reviews can be further improved, and the most critical and valuable information in text information can be fully utilized. The main contributions of this paper are listed as follows:

1. Using word embedding technique to train word embedding to represent text information as a low-dimensional dense matrix; a method for extracting feature words of educational tourism reviews based on noise word filtering using Word2Vec is proposed, which fully considers contextual semantic information, and can further improve the performance of feature word extraction. By classifying the sentiments of educational tourism reviews at the level of educational tourism characteristics, it is beneficial to analyze the sentiment tendencies of educational tourism.

2. CNN and LSTM are used to extract text local information and contextual features, and the output information is used as the input of multichannel attention to extract attention scores.

3. The output information of the multichannel attention model is fused to obtain the final text information vector representation, and then, the sentiment classification of educational tourism reviews is performed.

The proposed model makes full use of the advantages of CNN and LSTM to extract text features, and on this basis, a multichannel attention mechanism is introduced and different weights according to the impact of different words on the classification results are assigned so that words play different roles in classification tasks, thereby achieving the purpose of improving the classification performance of educational tourism reviews.

The rest of this paper is organized as follows. The related methods on the sentiment classification of educational tourism reviews are introduced in Section 2. The knowledge
background of educational tourism reviews is explained in Section 3. Section 4 describes in detail the proposed sentiment classification method for education tourism reviews based on CNN and LSTM-Attention. Section 5 presents the experimental results and discussions. Finally, Section 6 summarizes the full text.

2. Related Methods

Sentiment analysis aims to identify emotional tendencies through effective analysis and mining of information [14]. Medhat et al. [15] consider sentiment analysis as the focus of research in the field of text mining, that is, the computational processing of texts. The early sentiment analysis methods are mainly based on text data to perform sentiment calculation of word semantics and text sentiment calculation. With the deepening of research, sentiment analysis has become more refined, and research such as sentiment calculator, sentiment summarization, and product attribute mining have appeared. In recent years, with the development of big data, researchers have proposed many sentiment analysis models and software, providing strong support for sentiment research [16].

Choi et al. [17] utilize text data to study the Macau tourism image and verified that text analysis methods can not only conduct qualitative research but also quantitative research. Govers et al. [18] adopt artificial neural network to analyze the text content related to the images of seven tourist destinations. Radiojevic et al. [19] propose a sentiment analysis method based on more than 2 million online review data from more than 6,000 hotels in Europe and conclude that reviews are the most significant factor affecting hotel satisfaction.

Sentiment computing is the process of analyzing texts with emotions and classifying them into positive, negative, and other emotional types. If sentiment is divided into positive, neutral, and negative sentiment, then sentiment calculation is a classification problem; if sentiment is a specific numerical value or an ordered value in a given interval (such as 1–5 points), then sentiment values can be calculated by regression methods. Based on the granularity of text, sentiment computing can essentially be viewed as a multilevel hierarchy of words, phrases, sentences as well as articles. In the existing tourist sentiment research, high-frequency feature word statistics, content analysis software analysis, sentiment dictionary analysis, machine learning, and deep learning methods are mainly used to study the images of the destinations and the image differences of different destinations [20].

The methods based on content analysis software are based on the statistical method of high-frequency feature words and sentiment dictionary, and by writing logic codes, word segmentation, word frequency statistics, clustering, co-occurrence analysis, co-citation analysis, semantic network, co-occurrence analysis matrix, and other analysis methods are integrated into one software to realize the processing of texts. The commonly used software are CATPAC II, ICT-CLAS, ROST, and so on. The sentiment lexicon analysis methods are mainly based on one or more dictionaries containing labeled sentiment words, sentiment phrases with corresponding intensities, combined with degree adverbs, negative words, conjunctions, and syntactic structures to construct sentiment computing models for sentiment analysis. Some researchers use machine learning to perform tourist sentiment calculations. A sentiment calculator is obtained by manually annotating a portion of texts that express positive or negative sentiments and training a machine learning algorithm with these texts. The sentiment calculator is used to perform positive and negative sentiment calculations on text data and finally gives a specific score of 0 or 1 or gives the positive and negative probabilities of the texts [21].

In machine learning research, methods such as logistic regression, decision tree, and support vector machine (SVM) are commonly used to perform sentiment analysis. The short text sentiment classification methods based on machine learning design features and use multiple classifiers for sentiment analysis. Based on the word collocation features of dependent syntax and the deep features of combined semantics, Li et al. [22] propose a semi-Markov conditional random field text sentiment analysis model with phrases as the main clue, which plays an important role in solving the problem of implicit sentiment analysis. Gurkhe et al. [23] propose a naive Bayesian sentiment classification algorithm based on the weight features of knowledge semantics, in which the features are fused into the Naive Bayes classifier based on the correlation between the dictionary polarity distribution information and document sentiment classification to improve the accuracy of document-level sentiment classification.

Kalaivanani et al. [24] propose a generic algorithm that incorporates the information gain for feature selection and combined with KNN to improve sentiment classification performance. Li et al. [25] propose an SVM classifier based on multiple features and resources such as sentiment lexicon and word embedding. At the same time, an iterative method is used to assign different weights to the probability outputs, which improves the classification accuracy.

The above methods have achieved good sentiment classification results, but manual feature selection is a time-consuming and labor-intensive process. In recent years, with the rise and rapid development of deep learning techniques, these problems can be effectively solved. Deep learning methods use large-scale corpus to allow the model to actively learn the potential syntactic and semantic features in the text to achieve better understanding, effectively making up for the lack of artificial extraction features in information representation, thereby achieving better flexibility, robustness. Deep learning also has many applications in the field of short text processing. The convolution and pooling operations of CNN can be well applied to local feature extraction. Attardi et al. [26] use convolutional neural networks for sentiment classification and achieve good results on three-category data sets. Deriu et al. [27] exploit large amounts of data for remote supervision, train a convolutional neural network model, and combine it with a random forest classifier to optimize performance in polarity classification. Xu et al. [28] propose an LSTM neural network, by
introducing storage unit and gating mechanism to capture long-term dependencies in the sequence to decide how to use and update the information in the storage unit, and then obtain a more durable memory, expanding the advantages of deep computing. Tang et al. [29] adopt the LSTM model to combine target information, which significantly improves the accuracy of target-dependent sentiment analysis. Hao et al. [30] propose a novel sentiment-analysis model based on the parallel-CNN architecture which concatenates the representation of each segment to obtain the final representation of the text. Wang et al. [31] propose a regional CNN-LSTM model consisting of regional CNN and LSTM to predict the sentiments of texts. By combining the regional CNN and LSTM, both local information within sentences and long-distance dependency across sentences can be considered in the prediction process.

3. Research Background

3.1. Definition of Educational Tourism. Educational tourism originated in Japan after World War II and became popular in developed countries such as Europe and the United States in the 1970s. It is considered to be an important part of contemporary quality education. However, there is no unified concept at home and abroad, and scholars have three main views:

(1) Educational tourism refers to a tourism activity with students as the main body and education as the purpose. Prakapienė et al. [32] define educational tourism as a tourism product with students as the main body and the purpose of visiting educational institutions and learning knowledge. The development of educational tourism can not only broaden the horizons and increase skills but also strengthen the communication between teachers and students during the journey and increase the feelings of teachers and students.

(2) Educational tourism is a tourism activity characterized by "study as the mainstay and travel as the supplement." Tomasi et al. [33] describe educational tourism as tourists visiting and studying in historical sites, scientific research institutes, or well-known universities so as to improve cognition of the tourists.

(3) Educational tourism refers to the tourism activities in which tourists "promote their studies through travel and learn from each other." Hales et al. [34] take the meaning and purpose of tourism as the starting point and draw the conclusion that educational tourism relies on specific tourism resources, combining tourism products and education content to improve personal ability, learn knowledge, and achieve physical and mental cultivation.

According to the above different viewpoints, it can be seen that the concept of educational tourism can be divided into narrow and broad concepts. In a narrow sense, it specifically refers to tourism activities in which students are the main body of tourism, with the purpose of learning knowledge and cultivating skills. In a broad sense, educational tourism refers to tourism activities that rely on certain tourism resources to achieve the goal of education and study through tourism products; this study adopts the latter concept.

3.2. Educational Tourism Base. As a special tourism base, an educational tourism base must have certain tourism resources and at the same time provide a high level of comfort. As a learning place, it should have facilities to meet the teaching needs and be equipped with corresponding academic staff to complete the teaching tasks. According to the types of tourism resources that the educational tourism bases rely on, these bases can be divided into natural landscape educational tourism bases and human landscape educational tourism bases.

The former refers to understanding natural science and cultural knowledge while enjoying the natural scenery, such as Xishuangbanna Tropical Botanical Garden and Hexigten Global Geopark; the latter refers to understanding the history, culture, customs, literature, and art of the tourist destination during the tour, such as Xibaipo Memorial and Xuanzhi Culture Park.

3.3. Educational Tourism Review Data. In terms of data sources, this paper selects travel e-commerce websites and social platforms with abundant domestic travel reviews, including Baidu Travel.com, Qunar.com, Ctrip.com, and Sina Weibo. The web crawler tool is used to capture about 60,000 comments related to different educational tourism bases. For details, please refer to Section 5.2.

4. Sentiment Classification of Educational Tourism Reviews Based on CNN and LSTM-Attention

Both convolutional neural networks and long-short-term memory neural networks have their own advantages in sentiment classification tasks. CNN uses multiple convolution kernels to perform convolution operations on the word embedding of texts to effectively mine the potential semantic information of texts, while LSTM networks can better predict the semantics of text sequences. Combining these two types of neural networks, a neural network model based on CNN-LSTM-Attention is proposed. The structure is shown in Figure 1.

4.1. Word Embedding Layer Based on Word2Vec. Before the classification task, the text needs to be converted into a digital matrix that can be recognized by the computer and represented by a fixed-length real number. This representation is called word embedding. Early research mostly used one-hot encoding for conversion, representing each word as a digital matrix with corresponding dimensions according to the size of the vocabulary. Although this method can uniquely identify each word, it cannot reflect the correlation between words, and the dimension of the
embedding is positively correlated with the size of the vocabulary, which can easily lead to the dimension disaster problem.

Word2Vec is a word semantic computing technique proposed by Google in 2013. Through Word2Vec training, the processing of text content can be simplified into vector operations in the K-dimensional vector space, and the similarity in the vector space can be used to represent the semantic similarity of the text. Word2Vec provides two classic language models for training: CBOW model and Skip-gram model. For these two models, Word2Vec gives two frameworks which are designed based on Hierarchical Softmax and Negative Sampling, respectively [35]. This paper adopts the Skip-gram model based on Hierarchical Softmax.

Hierarchical Softmax uses a Huffman tree structure to represent the words of the output layer, where the words of the output layer exist as leaf nodes, and each node represents the relative probability of its child nodes. In a Huffman tree, there is always an optimal path from the root node to each leaf node. The Skip-gram model consists of a three-layer network model, namely the input layer, the projection layer, and the output layer, as shown in Figure 2. The training goal of the Skip-gram model is to find word representations that help predict similar words in a sentence or document.

During the training process of Skip-gram, the conditional probability value of the intermediate word embedding $W_i$ is used to solve the contextual word embedding, which can be calculated as

$$P(W_i | W_t) = \frac{P(W_t \ast W_i)}{P(W_t)},$$

$$i = t - 1, t - 2, t + 1, t + 2.$$  

Let the number of words in a sentence to be input into the model be $N$, and the two-dimensional matrix $Z = [W_1, W_2, ..., W_N]$ is used to represent this text. After the word embedding layer, the text representation is converted into $X = [x_1, x_2, ..., x_n]$, $x_i \in \mathbb{R}^d$, where $d$ denotes the dimension of the word embedding.

4.2. CNN Layer. The CNN network consists of several convolutional layers, pooling layers, and fully connected layers, and it has strong feature extraction capabilities. By setting convolution kernels with different sizes, local key information can be effectively extracted, and then, the input feature map is compressed through the Pooling layer to make the feature map smaller and simplify the computational complexity of the network. Finally, all the features are connected by the fully connected layer, and the output values are sent to the classifier.

The main function of the convolution layer is to use the convolution kernels to perform convolution operations on the word embedding matrix from the input layer to obtain deeper text features. Each convolution kernel corresponds to the extraction of a certain part of the feature. In this paper, the number of convolution kernels is set to 128.
following convolution operations are performed on each sentence matrix $X$ output by the embedding layer:

$$S = \zeta(\Omega X + b),$$

where $S$ represents the feature matrix extracted from the convolution operation; weight $\Omega$ and bias $b$ are the learning parameters of the network. To facilitate computation, a nonlinear mapping of the convolution results of each convolution kernel is required:

$$\zeta = \text{relu} = \max(0, x),$$

where the relu function is one of the excitation functions commonly used in neural network models. In order to extract features more comprehensively, this paper simultaneously uses convolution windows of size 2 and 3 to extract binary features and triplet features of sentences, respectively.

The main function of the pooling layer is to perform feature selection and information filtering on the text embedding extracted by the convolutional layer, reduce the parameters and computation of the next layer while preserving the main features, and prevent overfitting. The K-Max pooling operation is used in this paper, which selects the Top-$K$ maximal values of each filter to represent the semantic information of the corresponding filters. The expression for $K$ value is

$$K = \left\lfloor \frac{l - \zeta_{\text{size}} + 1}{2} \right\rfloor,$$

where $l$ is the length of the sentence vector, which is set to 50 in this paper. $\zeta_{\text{size}}$ is the convolution window size. After the pooling operation, the number of feature embedding extracted by each convolution kernel is significantly reduced, and the core semantic information of the sentence is retained.

Through the convolution operation and the pooling operation respectively, the convolutional layers and the pooling layers of CNN perform feature extraction on short text sentences and obtain the generalized binary and ternary feature vectors. After the fusion layer, the two types of feature vectors are merged together and used as the input matrix of LSTM-Attention.

4.3 LSTM-Attention Layer. LSTM is an improved version of RNN. By adding input gate $i$, forget gate $f$, output gate $o$, and internal memory unit $c$ to neurons, it has more advantages in processing long sequences of text and alleviates the problems of gradient disappearance and explosion so that LSTM can extract textual context information more effectively than RNN. The input gate $i$ controls how much of the input $X_t$ of the network at the current moment is saved to the unit state $C_t$, the forget gate $f$ determines how much of the unit state $C_{t-1}$ at the previous moment is kept to the current moment $C_t$, and the output gate $o$ controls how much of the unit state $C_t$ is output to the current output value $H_t$ of the LSTM. The model structure is shown in Figure 3.

When the input text word embedding matrix is $X = [x_1, x_2, ..., x_k]$, the LSTM can be updated as

$$i_t = \sigma(\Omega_i \cdot [H_{t-1}, x_t] + b_i),$$

$$o_t = \sigma(\Omega_o \cdot [H_{t-1}, x_t] + b_o),$$

$$f_t = \sigma(\Omega_f \cdot [H_{t-1}, x_t] + b_f),$$

$$C_t = f_t \otimes C_{t-1} + i_t \otimes \tanh \big(\zeta \cdot [H_{t-1}, x_t] + b_c\big),$$

$$H_t = o_t \otimes \tanh(C_t),$$

where $\sigma(\cdot)$ denotes the Sigmoid activation function, $\tanh(\cdot)$ is the hyperbolic tangent function, $\Omega$ denotes the corresponding weight, $b$ is the bias term, and $H_t$ is the final output. The output $H_t$ obtained by LSTM after extracting textual context information is used as the input of the multichannel attention mechanism, and the model structure is shown in Figure 4.

After $H_t$ is processed by the LSTM-Attention model, the final vector $V$ is obtained, which not only contains the textual context information but also focuses on important words, which better represents the semantic information. After that, the embedding matrix output by the LSTM-Attention model is input into the Dropout layer to prevent data overfitting. Then, the embedding matrix is input into the fully connected layer for dimensionality reduction. Finally, through the excitation function, the sentiment classification probability is obtained.

5. Experiments and Discussion

5.1 Experimental Environment. The experiments are performed on a computer running Windows 10 system with Intel i5-7400 CPU, the deep learning framework is TensorFlow 2.1.0, and Python 3.6.0 is used as the programming language. In order to better represent semantic information, Skip-gram is used to train word embedding, and CUDA10.1 is used for computational acceleration. The specific experimental environment is shown in Table 1.
5.2. Data Collection and Processing. In order to build a scientific tourist sentiment analysis model, it is necessary to have a data reference system that can be verified, and at the same time, an exclusive data set for educational tourism reviews should be established, and the factors of semantic logic and emotional expression tendencies should be considered. Eight educational tourism bases in China are selected as data collection points, namely Shanghai Science and Technology Museum, Chengdu Dujiangyan irrigation system, Shaanxi History Museum, Anyang Red flag canal, Chifeng Hexigten Global Geopark, Xuancheng City Chinese Xuanzhi Culture Park, Xishuangbanna Tropical Botanical Garden, and Shijiazhuang Xibaipo Memorial.

Details of data collection from several tourism bases in China are shown in Table 2. Firstly, some meaningless educational tourism reviews are cleaned, and finally, a total of 60,365 educational tourism reviews are obtained after filtering. Secondly, the jieba word segmentation package in the Python language is used to tokenize the review texts in the review set and label the parts of speech. Finally, based on the common stopword list, by performing word frequency statistics on the words in the educational tourism review set after word segmentation, and selecting words with high frequencies that have nothing to do with educational tourism characteristics and sentiments, a stopword list suitable for educational tourism reviews is constructed, and it is used to remove stopwords from the segmented data set. The educational tourism review data set collected in this paper can scientifically and comprehensively reflect the perceptions of tourists on tourism resources, products, and services. It contains objective reference data, with a certain scientific nature.

5.2.1. Feature Words of Educational Tourism Reviews. The preprocessed educational tourism review data set is used for the word embedding training corpus, and the build-in function in the Word2Vec library in the gensim package of the Python language is used to train the word embedding and build a word embedding model. The parameter settings of Word2Vec are shown in Table 3.

Firstly, the word frequency statistics analysis is performed on the nouns and noun phrases in the preprocessed educational tourism review data set, and the top 200 high-frequency nouns and noun phrases related to educational tourism features are extracted as initial feature words. Classify these feature words according to the categories of educational tourism features to form an initial feature word list. Secondly, the build-in function most_similar() of Word2Vec in the gensim package of the Python language is used to calculate the similarity between the nouns and noun phrases in the educational tourism review data set and the initial feature words. From the top 50 words with the highest similarities with the educational tourism feature words, the words that are truly related to the characteristics of educational tourism are selected to expand the initial feature word list; then, a noise word list is built on the basis of the feature word list.

With the help of the trained word embedding model, the feature words contained in each educational tourism review are extracted to form a list of feature words corresponding to the educational tourism review data set. Then, the noise word list is used to filter out the noise words in the educational tourism review feature word list to generate a noise-free feature word list.

In order to facilitate the comparative analysis with the existing feature word extraction methods, this paper adopts the precision ($P_{\text{extract}}$), the recall rate ($R_{\text{extract}}$), and the F1-score ($F1_{\text{extract}}$) as the evaluation metrics of the feature word extraction. $P_{\text{extract}}, R_{\text{extract}},$ and $F1_{\text{extract}}$ can be calculated as (6), where $c_i$ denotes the feature word set extracted by the feature extraction method from the $i$th educational tourism review, $d_i$ denotes the feature word set attached to the $i$th educational tourism review itself, and $Q$ indicates the number of reviews in the educational tourism review set to be processed.

$$P_{\text{extract}} = \frac{\sum_{i=1}^{M}|c_i \cap d_i|}{Q}$$

$$R_{\text{extract}} = \frac{\sum_{i=1}^{M}|c_i \cap d_i|}{|d_i|}$$

$$F1_{\text{extract}} = \frac{2 \times P \times R}{P + R}$$

500 educational tourism reviews are selected from the data collection and are processed using the TF-IDF method [36], the TextRank method [37], and the proposed feature word extraction method, respectively, and the results are shown in Figures 5(a)–5(c). It can be seen from Figure 5 that when using the proposed extraction method for feature word extraction with settings, the precision ($P_{\text{extract}}$) and F1-score ($F1_{\text{extract}}$) are better than that of the TF-IDF method and Text Rank method. The recall rate ($R_{\text{extract}}$) is also on par with the TF-IDF method and the Text Rank method when the number of feature words is set to 8–10.
Therefore, for short texts such as educational tourism reviews, the proposed feature extraction method based on noise vocabulary filtering is more effective than other methods.

5.3. Training Parameter Settings. Firstly, semantic segmentation is performed on each review in the educational tourism review set, then the segmented set is preprocessed and divided into three parts: training set, validation set, and test set according to the ratio of 6:2:2. The role of the training set is to fit the model and train the classification model by setting the parameters of the classifier. The function of the validation set is to use the trained model to predict the validation set data, adjust the model parameters, and select the parameters corresponding to the model with the best performance. The role of the test set is to use the trained CNN-LSTM-Attention model to perform sentiment classification for educational tourism reviews.

In order to prevent the occurrence of overfitting, the Drop_out value is set to 0.5 in the CNN and LSTM network layers, and 50% of the neural units are randomly deactivated. Using the ReLu activation function can speed up the convergence speed and further prevent overfitting. The cross-entropy loss function commonly used in multiclassification tasks is adopted. The optimizer is Adam, the Batch_size is 256, and the Epoch is 10. The specific parameters are shown in Table 4.

5.4. Model Evaluation Metrics. Commonly used evaluation metrics for classification tasks including accuracy, precision, recall, and F1-score are used for the evaluation of the proposed model. Accuracy (Acc) represents the ratio of correctly predicted samples to the total samples, precision (Pre) represents the ratio of correctly predicted positive samples to all positive sample predictions, recall rate (Rec) represents the ratio of correctly predicted positive samples to all actual positive samples, FPR (False Positive rate) is the negative-positive rate, which represents the ratio of falsely predicted positive samples to all actual negative samples, and F1-score is the weighted harmonic mean of precision and recall rate. The confusion matrix is shown in Table 5, and the metrics can be calculated as

\[
\text{Acc} = \frac{TP + FN}{TP + TN + FP + FN},
\]

\[
\text{Pre} = \frac{TP}{TP + FP},
\]

\[
\text{Rec} = \frac{TP}{TP + FN},
\]

\[
F1 = \frac{2 \times \text{Pre} \times \text{Rec}}{\text{Pre} + \text{Rec}},
\]

\[
Fpr = \frac{FP}{FP + TN},
\]

where TP is the number of positive samples correctly identified, TN is the number of negative samples correctly identified, FP is the number of positive samples incorrectly identified, and FN is the number of negative samples incorrectly identified.

The number of positive and negative samples in the experimental data set is unbalanced. The ROC curve is commonly used as the performance evaluation criterion to resolve the data imbalance problem, where Rec represents the vertical coordinates and Fpr represents the corresponding horizontal coordinates. By adjusting the threshold used in the classification of the classifier, a curve passing through points (0, 0) and (1, 1) will be obtained, that is, the ROC curve of this classifier. Usually, this curve should be above the line connecting (0, 0) and (1, 1). Because the ROC curve formed by the connection of (0, 0) and (1, 1) actually represents a random classifier, the area under the ROC curve is the AUC value, and the larger the AUC value, the better the classifier performance.

5.5. Sentiment Classification Results of Educational Tourism Reviews. In order to verify the prediction performance of the proposed model, the classical machine learning algorithms SVM [25], KNN [24], classical CNN algorithm [26], LSTM model [28], parallel CNN [30], and CNN + LSTM
method [31] are compared with the proposed method under the same experimental environment. The results on the educational tourism reviews data set are provided in Table 6.

It can be seen from the table that the proposed model has the best performance on the experimental data set and greatly improves the performance of sentiment classification of educational tourism reviews. The conventional machine learning methods SVM and KNN have poor classification performance and are not suitable for the sentiment classification task of educational tourism reviews. CNN and LSTM methods achieve significantly better performance than machine learning methods. Compared with the CNN+LSTM model, the Acc, Pre, Rec, and $F_1$ results of the proposed method are improved by 2.33%, 3.34%, 2.61%, and 3.3%, respectively. This is because the CNN-LSTM model uses a progressive structure. Although CNN can effectively extract local key information when extracting features, it will lead to the loss of some information. The semantic information extracted by CNN is incomplete when transmitted backwards. The proposed method not only extracts the local key information but also effectively extracts the context information. The information is complete when it is passed backwards, so the classification effect will be better. In addition, the proposed model introduces the Attention

![Figure 5: Comparison of feature word extraction. (a) Precision $P_{\text{extract}}$. (b) Recall rate $R_{\text{extract}}$. (c) $F_1$-score $F_{1\text{extract}}$.](image-url)

<table>
<thead>
<tr>
<th>Table 4: Network parameters.</th>
<th>Table 5: Confusion matrix.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Network parameters</td>
<td>CNN</td>
</tr>
<tr>
<td>Word embedding dimensions</td>
<td>256</td>
</tr>
<tr>
<td>Learning ratio</td>
<td>0.001</td>
</tr>
<tr>
<td>Kernel size</td>
<td>128</td>
</tr>
<tr>
<td>Convolution window size</td>
<td>2, 3</td>
</tr>
<tr>
<td>Hidden layers</td>
<td>—</td>
</tr>
<tr>
<td>Drop_out</td>
<td>0.5</td>
</tr>
<tr>
<td>Activation function</td>
<td>ReLU</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Table 6: Comparison of sentiment classification results.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Model</td>
</tr>
<tr>
<td>SVM [25]</td>
</tr>
<tr>
<td>CNN [26]</td>
</tr>
<tr>
<td>LSTM [28]</td>
</tr>
<tr>
<td>Parallel CNN [30]</td>
</tr>
<tr>
<td>CNN + LSTM [31]</td>
</tr>
<tr>
<td>Proposed method</td>
</tr>
</tbody>
</table>
Figure 6: Performance comparison using different models. (a) Precision. (b) Recall rate. (c) F1-score.

Figure 7: ROC results. (a) SVM, ROC curve (area = 0.61). (b) CNN, ROC curve (area = 0.72). (c) Proposed, ROC curve (area = 0.85).
mechanism on the basis of CNN + LSTM and assigns different weights to each word by calculating the attention score so that the words that have a greater impact on the classification result can be effectively identified, therefore obtaining significantly better classification performance than the CNN + LSTM model.

In order to further verify the effectiveness of the proposed model compared with common machine learning algorithms, basic neural network models LSTM, and parallel CNN models, a significance test experiment is designed. On the experimental data set, using word frequency as the feature, KNN classifier [24], SVM classifier [25], LSTM model [28], parallel CNN model [30], and the proposed model are used for 10-fold cross-validation, and the precision, recall rate, and F1-score are used for evaluation. The results are shown in Figure 6.

It can be seen from Figure 6 that for the experimental dataset, when only word frequency is considered as features, the values of the three evaluation indicators obtained with the machine learning method KNN and SVM are mostly distributed between 0.5 and 0.7, and the performance is relatively low. The LSTM model and the parallel CNN model have better classification results, and the values of the three indicators are distributed between 0.7 and 0.9. The classification accuracy of the proposed model is better than the other four models, and the three evaluation index values are in the range of 0.8 to 0.9, which is significantly better than the classification results of KNN and SVM. The improvement in classification performance is also larger than that of the LSTM model and the parallel CNN model, and the proposed method achieves higher stability.

When using SVM [25], CNN [26], and the proposed method for sentiment classification of the educational tourism reviews, the ROC curves of each model are shown in Figure 7, where the corresponding AUC values are given.

As can be seen from Figure 7, the classification performance of the convolutional neural networks is better than the classical machine learning algorithm SVM. In deep learning, the quantity and quality of the data sets are the key factors for classification and discrimination. In the big data environment, the better the learning effect, the closer the classification results of the classifier will be to the ground truth.

In the proposed method, based on the convolutional neural network CNN, the features of the short comment text are extracted, and the convolution windows of different sizes are used to extract the binary features and triplet features of the sentences, respectively. The long- and short-term memory neural network LSTM combined with the attention mechanism is used to predict the sentimental tendency of the review texts, and finally, the outputs from CNN and LSTM are merged to realize the classification of positive and negative sentiments, thereby improving the classification performance. The experimental results verify the effectiveness of the proposed method.

5.6. Hyperparameter Analysis. The setting of hyperparameters has an important impact on the final experimental results. In order to further improve the performance of the proposed model, the embedding dimensions and the number of LSTM hidden layers are further explored. Keeping other hyperparameters unchanged, the embedding dimension is set to 50, 100, 200, and 300, and the experimental result is shown in Figure 8(a). In addition, set the number of LSTM hidden layers to 64, 128, 256, and 512 and keep other parameters unchanged. The result is shown in Figure 8(b).

It can be seen from Figure 8(a) that when the embedding dimension is 300, it is close to the optimal classification performance, but considering the model complexity and classification efficiency, this paper chooses to set the embedding dimensions to 256. In Figure 8(b), the classification performance of the proposed model on the experimental data set increases with the increase of the number of LSTM hidden layers. When the number of hidden layers is greater than 256, the growth rate slows down, and the classification performance when the number of hidden layers is 512 is close to that when the number of hidden layers is 256. This is because the number of words in the text of the experimental data set is large, and the increase in the number of hidden layers enables better extraction of semantic information. When the number of hidden layers is greater than the number of text words, the classification performance will decrease as the number of hidden layers increases. So, the number of LSTM hidden layers should be set to 256.
6. Conclusion

The educational tourism industry is gradually prospering. It is of great practical significance to provide accurate recommendations to tourists and to provide tourists’ feedback to educational tourism bases. It is helpful to make full use of tourism resources, improve the protection of architectural facilities and relics, promote the development of scenic spots at different levels, enrich the types of educational tourism bases, and meet the diversified education and tourism needs of tourists. This paper proposes a text classification model with multichannel attention mechanism based on CNN and LSTM for the classification of educational tourism reviews. Firstly, the input text is represented as a low-dimensional dense word vector matrix with word embedding, then the local key information and contextual semantic information are extracted by CNN and LSTM, and the attention score of the output information of LSTM is extracted by the multichannel attention mechanism. Finally, the output information of the multichannel attention mechanism is fused, which realizes the effective extraction of text features and focuses on important words and improves the performance of text classification. Through comprehensive experiments, the advantages of the proposed model are further demonstrated by comparing with other models. Since most of the tourists gave positive reviews, there is an imbalance in the data set, resulting in a relatively low AUC value of the proposed model. In the future, we will focus on using an improved oversampling algorithm to improve the classification accuracy of minority samples.

Data Availability

The data used to support the findings of this study are included within the article.

Conflicts of Interest

The authors declare that there are no conflicts of interest regarding the publication of this paper.

References


