Tourism Consumer Sentiment Analysis Using a Multi-layer Memory Network Combining Temporal Convolutional and BiLSTM Architectures

Kun Yang[1]*, Wei Yang[2]
[1]School of Tourism Management, Wuhan Business University, Wuhan, 430000, China
[2]Department of Social Sciences, Quanzhou Medical College, Quanzhou, 362000, China
E-mail: Y7yangk@163.com
*Corresponding author

Keywords: multilayer memory networks, text sentiment analysis, tourism, bi-directional long-short term memory, temporal convolutional networks

Received: January 22, 2025

Traditional sentiment analysis methods for tourism review texts often suffer from a strong dependence on vocabulary quality and insufficient ability to capture contextual temporal features, making it difficult to accurately identify sentiment tendencies in complex contexts. To address this issue, this study proposes a multi-layer memory network algorithm that integrates temporal convolutional networks (TCN) and bidirectional long short-term memory (BiLSTM) to improve the depth of feature extraction and the accuracy of emotion classification. A dataset of 30,000 manually labeled tourism review texts was constructed, with sentiment labels categorized into positive, neutral, and negative. Experimental results showed that the proposed model achieved an accuracy of 95.78%, an F1 score of 95.98%, and a training time of only 40.18 seconds in sentiment classification tasks. Ablation studies demonstrated that the model exhibited clear advantages in structural integration and regularization mechanisms, with stronger stability. Sentiment orientation analysis showed that positive reviews account for 74.8% on average, and the model effectively identified multi-emotional expressions, indicating strong practicality and generalization ability. In summary, the proposed model achieves excellent performance in both feature extraction and sentiment classification, and achieves high generalization and training efficiency. Its application to tourism consumer review analysis can provide reliable technical support for the precise management of tourist attractions and user feedback analysis, while also offering new insights for the advancement of sentiment analysis methods.

Povzetek: Članek predstavi novo večplastno spominsko arhitekturo, ki združuje TCN in BiLSTM za učinkovito razpoznavanje čustev v turističnih ocenah z izboljšano kontekstualno zaznavo in računsko učinkovitostjo.

1 Introduction

The booming development of the tourism industry provides consumers with more and more choices in terms of tourism products and tourism services, and consumers often refer to the evaluations and feedbacks of other consumers before making purchase decisions [1]. Therefore, tourism consumer evaluation text has become an important source of information for tourism enterprises to understand consumer needs, improve services, and formulate marketing strategies. The text sentiment analysis (TSA) technology can automatically recognize the emotional tendency of the text and categorize it, thus helping tourism enterprises to quickly understand consumer feedback and make corresponding adjustments [2].

Currently, there are more categories of TSA techniques, and common methods include machine learning, deep learning, etc., in addition to traditional rule-based methods, and methods based on sentiment dictionaries. Among these algorithms, deep learning techniques have a wide range of applications in TSA [3]. Tasks such as text understanding and sentiment analysis that require long-range dependency modeling often benefit from optimized memory network structures. Multi-layer memory networks (MMNs) are an extension of traditional recurrent neural networks, such as long short-term memory (LSTM), by introducing stacked layers to improve the model's ability to capture long-range dependencies [4]. Although these models have strong expressiveness, they often suffer from long training times, high data requirements, and complex hyper-parameter tuning processes [5].

To address these limitations, this study proposes a lightweight MMN architecture that combines a temporal convolutional network (TCN) with a bidirectional long short-term memory (BiLSTM) network. TCN provides strong parallel computational capabilities and effectively extracts local sequential features, while BiLSTM captures global contextual dependencies and improves semantic understanding of complex emotional expressions. The integration of both networks not only mitigates problems such as gradient vanishing, but also improves both model depth and training efficiency. In this research, the TCN-BiLSTM model is applied to the sentiment analysis of tourism consumer reviews, with the aim of improving the classification accuracy and computational efficiency in TSA tasks, and providing a high-performance technical solution for practical applications.

There are four primary sections to the research. An outline of the state of domestic and international research on TSA techniques is provided in the first section. The development of the study's algorithm model constitutes the second section. The third part is the validation of the performance of the algorithm. The fourth section summarizes the research work.

2 Related works

With the continuous development of tourism, tourism enterprises also pay more and more attention to consumer feedback. Consumer reviews are one of the important reference indicators for tourism enterprises to adjust their management strategies. Another crucial area of study in the fields of data mining and natural language processing is the TSA of tourist consumer evaluation [6]. Tourism evaluation TSA aims to extract the emotional tendency as well as the expression of intention in the text by analyzing the text, and MMNs have a high degree of adaptability in this application [7]. The commonly used algorithms for MMNs include LSTM, attention mechanisms, etc., which have many applications in various fields. Most of the researchers have improved MMNs, including optimization of different algorithms as well as combinations. To improve the accuracy of sentiment classification, Ramaswamy S L and Chinnappan J combined convolutional neural network (CNN) algorithm and LSTM algorithm to empirically investigate various ensemble models. It was also compared with other traditional sentiment classification methods to evaluate the accuracy of CNN-LSTM algorithm for sentiment analysis. According to the findings, the CNN-LSTM algorithm performs sentiment analysis with good accuracy [8]. To improve the accuracy and stability of stream flow prediction, Wan W et al. proposed a model combining two-layer TCN and LSTM algorithms while incorporating a multi-head attention mechanism and verifying the model generalization. The experimental results indicated that this algorithm had less error in traffic prediction and had high accuracy and stability [9]. Li C et al. combined TCN with LSTM for the estimation of the health state as well as the remaining lifetime of the battery. The results of simulation experiments indicated that the TCN-LSTM algorithm had a high degree of fit, accurate prediction results, and robustness [10]. He et al. proposed the SSAE-TCN-BiLSTM model in an attempt to capture the temporal nuances of network traffic, which reduced the feature dimensionality by SSAE and performed feature extraction with TCN and BiLSTM algorithms. The results indicated that this algorithm model was computationally efficient, converged quickly, and had a large improvement in accuracy [11]. For accurate groundwater level prediction, Mohammad E et al. proposed a model combining self-attentive TCN and LSTM algorithms to solve the problem of vanishing gradient and overcome the limitations of LSTM. The results showed that this algorithm could effectively capture the long and short-term dependencies in the time series, with a small overall error and good algorithm performance [12].

Numerous studies have been conducted on TSA by researchers both domestically and overseas. To interpret the emotions and opinions in the text, Rengarjan et al. adopted the CNN-RNN algorithm model and utilized AGTO for complex feature extraction and selection. Experimental results indicated that this method could effectively perform text feature extraction and perform sentiment analysis, improving feature recognition [13]. To extract video pop-up information, Zhou X and Li W proposed a local CNN-based pop-up text sub-recognition algorithm, which reconstructed the data into a data lookup table and formatted the content for output. The results indicated that the text classification and recognition of this method was better and advanced [14]. Wang L et al. proposed a multi-feature text data enhancement model, M-DA, in order to learn text information more deeply. The study reconstructed and analyzed text sequences and captured long-term dependencies between sequences with BiLSTM networks. The outcomes revealed that the modeling framework constructed in this way could effectively extract text features, improve TSA, and increase accuracy [15]. By suggesting an NBLex-based method to assess sentiment in Kannada-English code-switched text, Chundi et al. increased TSA's efficiency and decreased the amount of time needed to identify the sentiment. The results of the study indicated that this method was effective in improving the efficiency of TSA and did not require training samples, reducing the cost of TSA [16]. To analyze the comments on social media, Taneja K et al. used SA techniques for language processing. The study proposed an unsupervised learning method using a transformer architecture for performing SA on a dataset. The results of the study revealed that this method had high F1 and AUC scores and was feasible to be used for product recommendation and thus influencing business decisions [17]. Table 1 presents a structured comparison of existing sentiment analysis methods based on different algorithms, application domains, data sets, performance results, and attention limitations. This comparison highlights the research gap that the TCN-BiLSTM model proposed in this study addresses in the context of travel sentiment analysis.

Table 1: Summary of related sentiment analysis methods

Study	Method	Domain	Dataset	Performance	Limitations
Ramaswamy & Chinnappan [8]	CNN-LSTM	Sentiment analysis	General review texts	Good accuracy for sentiment classificationt	Limited modeling of long-term dependenciest
Wan et al. [9]	TCN+LSTM+Multi-head Attention	Streamflow prediction	Time series data	High accuracy, low prediction error, good stability	Not applied in text sentiment domain
Li C et al. [10]	TCN+LSTM	Streamflow prediction	Battery lifecycle data	High fit, accurate and robust predictions	Non-textual domain, no NLP validation
He Z et al. [11]	SSAE-TCN-BiLSTM	Network traffic analysis	Time series intrusion data	High computation efficiency, fast convergence	Not designed for textual sentiment
Mohammad E et al.[12]	Self-attention TCN+LSTM	Groundwater level prediction	Environmental time series	Models short- and long-term dependencies, low error	Not applicable to sentiment analysis
Rengarjan et al. [13]	CNN-RNN+AGTO	Literary sentiment analysis	Novel "Immortals of Meluha"	Effective feature extraction and sentiment recognition	Limited generalization beyond literary texts
Zhou X & Li W [14]	Local CNN	Video popup text classification	Popup texts from video	Better classification performance	Not focused on emotion polarity
Wang L et al. [15]	M-DA+BiLSTM	Chinese TSA	Multi-feature Chinese datasets	Improved accuracy through feature enhancement	Model complexity increases computation load
Chundi R et al. [16]	NBLex (Lexicon-based)	Code-switched sentiment analysis	Kannada-English mixed corpus	Efficient, cost-effective, no training data needed	Dependent on lexicon quality, low adaptability
Taneja K et al. [17]	Transformer (unsupervised)	Social media review analysis	E-commerce reviews	High F1 and AUC, usable in recommendation	Complex architecture, weaker interpretability

In summary, MMNs exhibit strong performance in feature classification and analysis tasks, and show high applicability in tourism sentiment analysis. However, existing TSA methods often face challenges such as limited modeling depth and inefficient training. This study proposes a novel MMN architecture that integrates TCN with BiLSTM. Rather than replacing LSTM or attention mechanisms, the proposed model builds on their ability to capture long-term dependencies while leveraging TCN to improve local feature extraction and parallel processing efficiency. As a result, the TCN-BiLSTM framework provides a lighter and more efficient solution for sentiment classification. Given its relatively limited application in the tourism domain, the model demonstrates both novelty and feasibility. Moreover, it is expected to provide new methodological insights for TSA research.

3 TSA approach based on multi-layer memory networks

3.1 Algorithmic modeling of multi-layer memory networks

To address the challenges of complex emotional expressions, diverse sentence structures, and frequent sentiment shifts in tourism review texts, this study proposes a TSA method based on a MMN. Based on the hypothesis that local features should be extracted in parallel, while global semantics require bidirectional modeling, the method integrates TCN with BiLSTM, providing improved contextual awareness and sentiment classification accuracy compared to existing approaches. When dealing with noisy or ambiguous sentiment expressions, the model leverages BiLSTM to capture contextual dependencies from both directions, improving robustness under uncertain emotional content. Compared to Transformer-based models, the proposed method achieves a better trade-off between computational efficiency and classification accuracy without requiring pre-training, making it suitable for real-time applications or resource-constrained environments.

MMNs is a network architecture that adds memory mechanisms to traditional neural networks, which is suitable for tasks requiring long-term memory and reasoning ability, such as question-answer systems in natural language processing, text understanding, and so on. Commonly used algorithms for MMNs include recurrent neural networks, LSTM, and attention mechanisms [18]. However, TSA needs to analyze and identify the sentiment features (SFs) contained in the text, and then classify them. Based on this consideration, this study takes BiLSTM network as the basis, combines TCN, constitutes MMNs algorithmic model, and carries out sentiment analysis on the text of tourism consumer evaluation. BiLSTM provides bi-directional modeling capabilities, allowing it to capture both past and future contextual dependencies. This makes it particularly suitable for analyzing complex tourism reviews that contain emotional shifts, negations, or mixed sentiments, enabling more accurate extraction and identification of key SFs.

To analyze the sentiment of the text, it is first necessary to vectorize the text content. Text vectorization is the process of transforming text data into digital signals, and commonly used algorithms include BERT dynamic word vector language model, Glove, Word2vec, and so on. Among them, Word2vec can capture the complex relationship between words and includes two architectures, CBOW and Skip-gram [19-20]. The computational process of CBOW is shown in Equation (1).

$$p(W \mid context) = p(w_t \mid w_{t-c}, w_{t-c+1}, \cdots, w_{t-1}, w_{t+1}, w_{t+c+1}, \cdots, w_{t+c}) \tag{1}$$

$$p(context \mid w_t) = p(w_{t-c}, w_{t-c+1}, \cdots, w_{t-1}, w_{t+1}, \cdots, w_{t+c+1}, \cdots, w_{t+c} \mid w_t) \tag{2}$$

$$p(w_{t+n} \mid w_t) = soft\max(v_{w_t}^T v_{t+n}^{'}) = \frac{\exp(v_{w_t}^T v_{t+n}^{'})}{\sum_{w' \in v} \exp(v_{w_t}^T v_{w'}^{'})} \tag{3}$$

In Equation (1), W represents the vocabulary. w_t represents the target word of the current time step t. c represents the context window size. w_{t-c} through w_{t+c} are the context terms. $context$ context refers to the sequence of words composed of c words on both sides of the target word, which is used to predict the conditional probability of the current word. As demonstrated in Equation (2), Skip-gram predicts the context by the target word, whereas CBOW primarily predicts the target word by the provided context.

In Equation (2), the Skip-gram can make a judgment on the probability of the occurrence of the glossary content in the text before and after the word based on the provided target word, as shown in Equation (3).

In Equation (3), v_w represents the vector representation of the input word in the input embedding matrix. v_w' represents the vector representation of the prediction target word in the output embedding matrix. n denotes the proximity between the specified word and the current input word. The training sample w is contained in the vocabulary list v. The Skip-gram model can represent the relationship between different words with fewer word vector dimensions, and at the same time, it can take into account the contextual information to effectively extract the semantic information in the text. Then, the text data is processed by convolutional network. The convolutional kernel of the CNN can extract the data features of the text within the convolutional layer (CL), as shown in Equation (4).

$$x_i^k = f(\sum_{j \in N_i} x_j^{k-1} * w_{j_i}^k + b_i^k) \quad (4)$$

In Equation (4), x_i^k and x_j^{k-1} represent the output of the i neuron in layer k and the j neuron in layer $k-1$, respectively. $w_{j_i}^k$ is the weight parameter connecting these two neurons, b_i^k represents the bias term, and $f(\cdot)$ is the activation function. N_i represents the index set of the neurons of the previous layer connected by the i-th neuron in the k layer. $j \in N_i$ represents the traversal and weighted summation of all the neurons of the previous layer that have connections with this neuron. After the CL computes the features, the next step is the pooling layer (PL) to fuse the features. The PL simultaneously collects the best features and minimizes feature dimensions to prevent overfitting. The formula for processing text data by PL is shown in Equation (5).

$$x_i^k = f(\beta_i^k D(x_i^{k-1}) + b_i^k) \quad (5)$$

In Equation (5), $D(x_i^{k-1})$ represents the pooling operation, and according to the task needs, the maximum pooling method or average pooling method can be selected to process the text information. After the data is processed by the CL and PL, the fully connected layer will combine the local information extracted from them to map the features into the sample labeling space, highly purify the features, and input them to the classifier. The TCN algorithm used in the study has an additional feature compared to the normal CNN, which is to deform the latter so that it can handle temporal problems. This improvement can give the CNN the ability to perform text sequence prediction. In this case, the causal convolution enables the algorithm to have a causal property, which is calculated as shown in Equation (6).

$$P(x) = \prod_{t=1}^{T} P(x_t \mid x_1, x_2, \cdots, x_{t-1}) \quad (6)$$

In Equation (6), x_t represents the current input value. T represents the total length of the input sequence. That is, the number of terms in the text or word sequence, which is used to control the modeling range of the model in the time dimension. It follows that the current information can only be predicted by the previous input values. Since the feature extraction is limited to a certain range, the study adds dilation convolution to compensate for this limitation as shown in Equation (7).

$$f_{k\_d} = (d-1) \times (f_{k-1}) + f_k \quad (7)$$

In Equation (7), d is the number of voids. Dilation convolution is equivalent to inserting d zeros into the convolutional network, and the text sequence is sampled and computed every certain d. The addition of dilation convolution increases the number of network layers. However, to prevent gradient vanishing due to excessive depth, residual links are necessary, as illustrated in Equation (8).

$$R = x + F(x) \quad (8)$$

Equation (8) is the process of linear transformation of the inputs to each layer. The TCN network algorithm proposed in the study is shown in Equation (9).

$$\begin{cases} y_i = Conv(W_i \cdot F_j + b_i) \\ \{Y_0, Y_1, \cdots, Y_t\} = Hi - norm(\{y_0, y_1, \cdots, y_t\}) \\ \{T_0, T_1, \cdots, T_t\} = \mathrm{Re}\,Lu(\{Y_0, Y_1, \cdots, Y_t\}) \end{cases} \quad (9)$$

In Equation (9), y_i represents the state values. W_i is the word matrix. F_j represents the convolution kernel of the j th layer. b_i is the bias matrix. The final resulting text sequence features are then $\{T_0, T_1, \cdots, T_t\}$. In order to make the overall algorithm able to learn the before and after associations, the study uses BiLSTM algorithm to optimize the TSA process. Its network structure is shown in Figure 1.

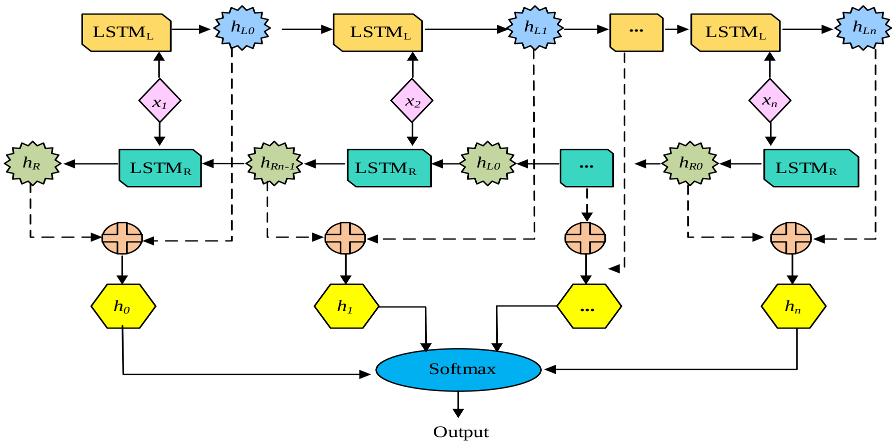


Figure 1: Schematic diagram of BiLSTM network structure

In Figure 1, BiLSTM constitutes an acyclic graph, which is able to obtain the output value (OV) through the text information before and after, and has higher stability. Splicing the text features in both directions can extract the dependencies of the text sequences and obtain the SFs. The text sequence feature in the forward direction is $\vec{H} = \{h_{L1}, h_{L2}, \cdots, h_{Lt}\}$, and the text sequence in the reverse direction is $\overleftarrow{H} = \{h_{R1}, h_{R2}, \cdots, h_{Rt}\}$. The vector of splicing the two directions is shown in Equation (10).

$$H = \{\vec{H}, \overleftarrow{H}\} = \{h_1, h_2, \cdots, h_t\} \tag{10}$$

The text features captured by BiLSTM are further fused as shown in Equation (11).

$$H(x) = Temp(x) + Bi(x) \tag{11}$$

In Equation (11), $Temp(x)$ is the OV of TCN. $Bi(x)$ is the OV of the BiLSTM network layer. The combination of the two is the final output. Summarizing the algorithms used in the study, the MMNs algorithm model is constructed as shown in Figure 2.

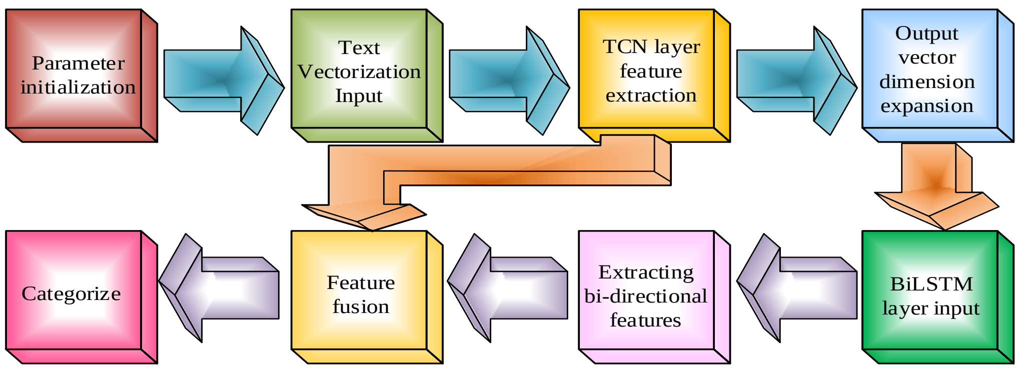


Figure 2: Model of TCN-BiLSTM multi-layer memory network algorithm

In Figure 2, the first step of the proposed TCN-BiLSTM MMNs algorithmic process lies in the initialization of the text data. Moreover, then the data is transmitted to the input layer of the TCN network, which outputs the vector information. Next, the data is used for final classification and TSA after the features are further extracted and fused using the BiLSTM method.

3.2 TSA method based on TCN-BiLSTM algorithm

TSA is a kind of computer technology to recognize and analyze the expression of emotion in text. It mainly includes four aspects: extraction of sentiment information, identification of subjectivity and objectivity of text, determination of sentiment tendency, and measurement of sentiment intensity. The analysis methods include methods based on deep learning as well as methods based on emotion features and emotion lexicon [21-22]. Figure 3 displays the details.

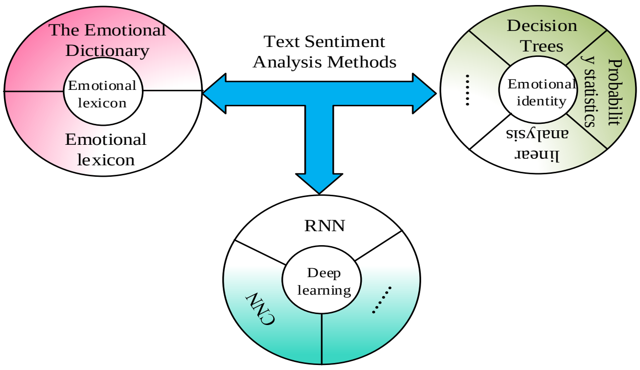


Figure 3: TSA method

In Figure 3, each of the three general categories into which TSA can be classified has its own subdivided procedures. The study adopts the analysis method based on SFs to structure the text data and put it into the training and testing of the algorithm model. The TCN-BiLSTM algorithm is used to classify the text features and predict the sentiment pointing of the text. Its algorithm flow is shown in Figure 4.

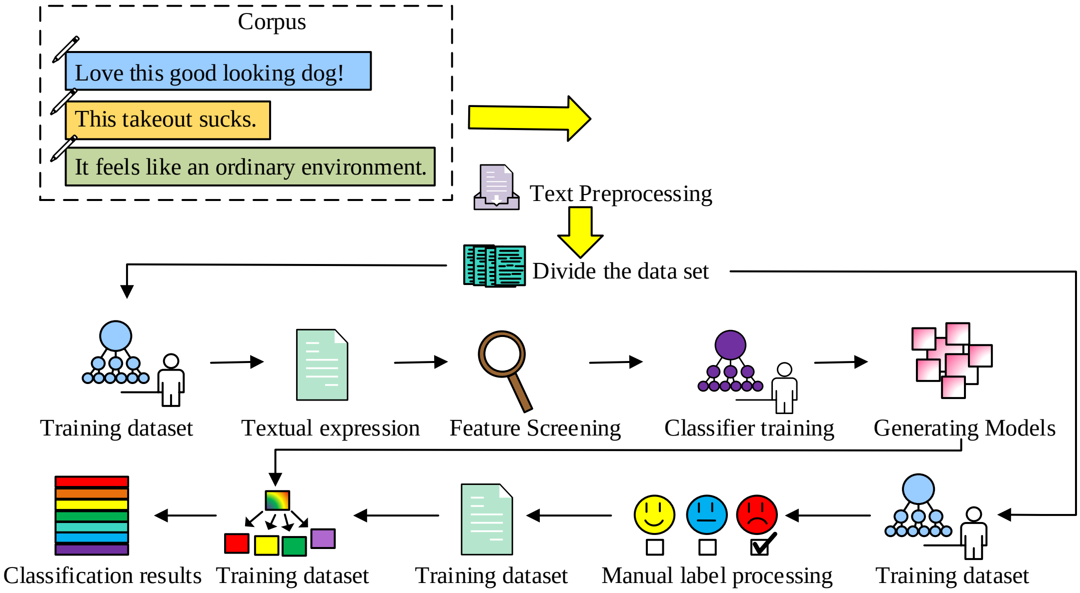


Figure 4: Flow of text analysis based on sentiment features

In Figure 4, the first step of text analysis based on SFs lies in preprocessing the text information and filtering and processing the features of the text by dividing the two datasets, training and testing. Finally, the classification model is generated to achieve the effect of classifying the textual SFs. For this purpose, the text needs to be preprocessed to represent the text in a computerized way, and the features are screened and classified. Among them, there are many methods commonly used for the screening of text features, such as the mutual information method. The mutual information approach may sort the frequency of texts based on the size of the MI value and measure the text's emotional inclination. Equation (12) displays the calculating formula.

$$MI(w_i,c_j)=\sum_{w\in W}\sum_{c\in C}P(w_i,c_j)\log\frac{P(w_i,c_j)}{P(w_i)P(c_j)} \quad (12)$$

In Equation (12), C represents the category of the document. W denotes all words in the dataset. P represents the probability. w_i is the feature item. c_j represents the text category of the item. $MI(w_i,c_j)$ then represents the mutual information value between the feature item and the category. Low-frequency features are more informative while high-frequency features are less informative. Another method used is information gain, which evaluates the discriminative power of textual features based on their information gain with respect to the target sentiment labels. By ranking features according to their computed information gain scores, features with low scores-indicating limited contribution to classification-are removed, while more informative features are retained for subsequent modeling. Equation (13) illustrates the computation procedure.

$$IG(t_j)=-\sum_{i=1}^{n}P(c_i)\log_2^{P(c_i)}+P(t_j)\sum_{i=1}^{n}P(c_i\mid t_j)\log_2^{P(c_i|t_j)}$$
$$+P(\bar{t}_j)\sum_{i=1}^{n}P(c_i\mid \bar{t}_j)\log_2^{P(c_i|\bar{t}_j)} \quad (13)$$

In Equation (13), c_i is the text type and the feature term is t_j. The probability that t_j exists in c_i is represented by $P(c_i\mid t_j)$ and $P(c_i\mid \bar{t}_j)$, respectively.

Through the information gain method, the important feature terms can be extracted quickly, and it has a wider application in dealing with problems that use the same set of features. It is worth noting that in the feature selection stage, this study employs mutual information and information gain methods, which are primarily used to preliminarily filter and rank the high-dimensional term features in the original review texts. These two methods are commonly used supervised feature evaluation techniques in text classification, capable of measuring the correlation between individual features (terms) and the target labels (sentiment categories). Therefore, it can effectively reduce input dimensionality, eliminating redundant information, and improving training efficiency. Considering that the focus of this study is on constructing a high-performance sentiment classification model architecture, the purpose of feature selection is to improve modeling efficiency and generalization ability, rather than model interpretability. Therefore, feature evaluation methods that focus on interpretability, such as SHapley Additive exPlanations (SHAP) and local interpretable model-agnostic explanations (LIME), are not introduced in this study.

To effectively classify the extracted features in the TCN-BiLSTM algorithm model, the study uses Softmax function for classification as shown in Equation (14).

$$P(y=k \mid x)=\frac{\exp(w_k^T x+b_k)}{\sum_{k=1}^{k}\exp(w_k^T x+b_k)} \tag{14}$$

In Equation (14), the number of categories is denoted by k. w is the weight. b is the bias. The larger the calculation result is, the greater the possibility that the feature item belongs to the category. Through this method, text features can be effectively classified, which is conducive to further TSA. Combining the TCN-BiLSTM algorithm and the computational flow of TSA, the proposed MMNs algorithm can be summarized as Figure 5. The proposed MMNs algorithm can be summarized as Figure 5.

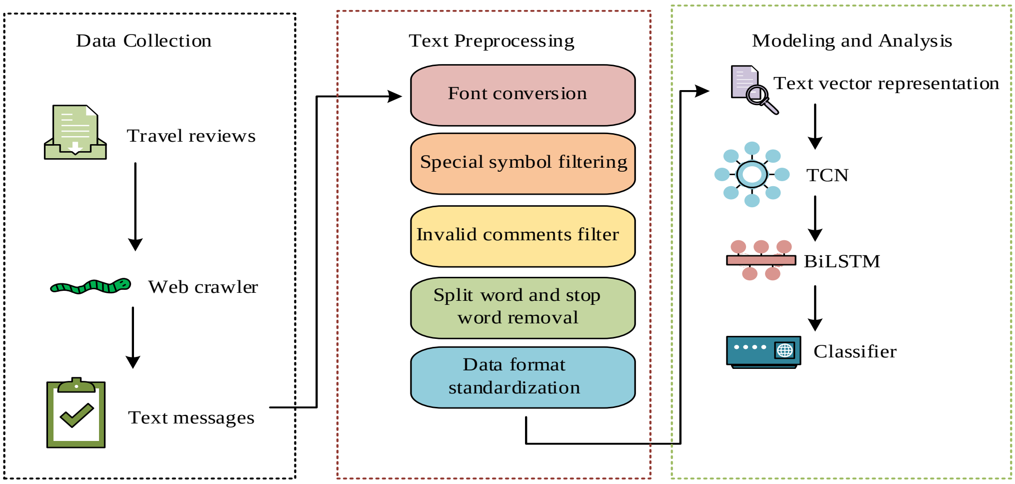


Figure 5: TSA model based on TCN-BiLSTM

The study's suggested MMNs algorithm model is broken down into three sections in Figure 5: text processing, data collecting, and memory network-based text analysis. The evaluation text of tourism consumers is extracted and converted into standard text data information by formatting the text information. Moreover, it is introduced into the TCN-BiLSTM MMNs model to represent the text as vectors. Then the vector information is output from the TCN network, and the text features are extracted and analyzed by the BiLSTM network, and finally output to the classifier to realize the classification of text SFs.

4 Tourism consumer evaluation TSA results

4.1 Validation of modeling effects of multi-layer memory networks algorithms

The study initially creates an experimental platform to confirm the overall performance of the algorithm model before examining its impact on the TSA of tourism assessment in order to confirm the impact of the suggested algorithm on the sentiment analysis of consumer evaluation. The dataset used in this study is collected using a Python-based web scraping tool from major tourism platforms such as Ctrip and Mafengwo, focusing on user reviews of popular tourist attractions in Hunan Province. To ensure data quality, preprocessing steps includes HTML tag removal, noise filtering, Chinese word segmentation, stopword removal, and word vector transformation. A total of 30,000 review samples are collected and sentiment labeled using a combination of keyword-based rules and manual correction. Of these, 15,000 are labeled positive, 4,000 are labeled neutral, and 11,000 are labeled negative. To reduce the impact of class imbalance, the sample proportions in the training set are adjusted accordingly, while the test set retained the original distribution. Given the subjective nature of user reviews, the consistency of manual annotations is verified through cross-validation to increase label reliability. Table 2 displays the specifications of the experimental platform.

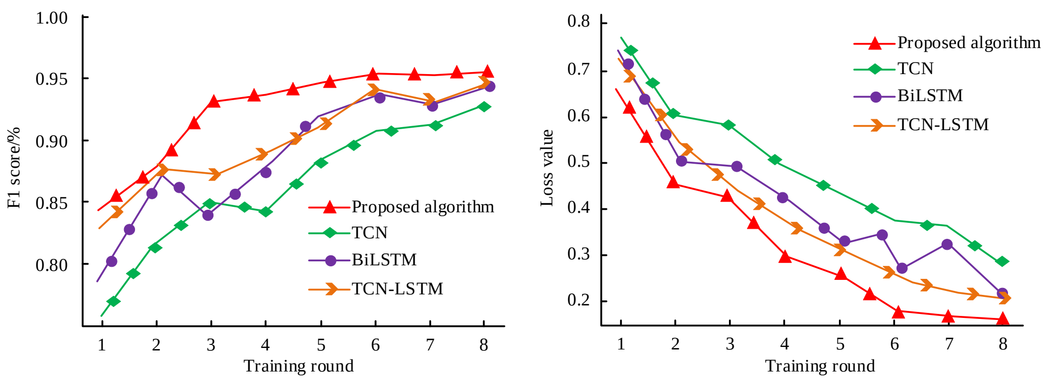
Table 2: Parameter settings of the experimental platform

Factors	Setting
Operating system	Windows 10
Processor	IntelXeon(R) E5
Graphics card	RTX3090
Host memory	64GB
Experimental frame	TensorFlow
Learning rate	0.001
Dropout value	0.5
Word vector size	210
Number of hidden layer units	128
BatchSize	128
Epoch	20

In Table 2, during the experiment, the model is run on Windows 10 operating system, the hardware platform includes Intel Xeon E5 series processor, RTX 3090 graphics card, and 64GB host memory, and the deep learning framework adopts TensorFlow. To ensure the performance stability and convergence efficiency of the model, the main hyper-parameters are experimentally pre-tuned in this study. The learning rate is initially set to 0.001 and tested in the range {0.0005, 0.001, 0.002}, and 0.001 is finally selected as the optimal value of the model for this task. The dropout rate is set to 0.5 to avoid overfitting. The word vector dimension is set to 210, chosen after comparing the performance in {128, 210, 300}. The number of hidden layer units is 128, the batch size is 128, and the number of training rounds (EPOchs) is 20. The above parameters show stable performance in multi-group combined tests, taking into account training efficiency and model effect. Considering that this study focuses on model structure design and comparative evaluation, the selection of hyper-parameters follows the principle of combining performance first with empirical verification.

The performance of the algorithms suggested in the study is confirmed using the experimental parameter settings mentioned above, and the trends of F1 and loss values of several methods are contrasted. Figure 6 displays the findings.

In Figure 6(a), although all the algorithmic models are able to stabilize the F1 scores as the training rounds increases, the TCN-BiLSTM algorithm proposed in the study is able to achieve the results faster. According to the change trend, it can be realized that the TCN-BiLSTM algorithm, besides being able to achieve a higher F1 score in a shorter time, also has stronger stability, faster convergence and higher accuracy. Figure 6(b) shows that the loss value of the proposed algorithm is the lowest when the iterations reaches 8, which indicates that the proposed algorithm has a higher fit and feature learning ability. Table 3 shows the average performance indicators of each model in the emotion classification task of travel reviews.



(a)Trend of F1 scores for different algorithms (b)Trend of loss value for different algorithms

Figure 6: Trend of F1 score and loss value of different algorithms

Table 3: Performance data for each algorithm model

Algorithm	Accuracy/%	Recall rate/%	F1 score/%	Accuracy 95% CI	Recall 95% CIt	F1 95% CI
TCN-BiLSTM	95.46±0.17	95.34±0.19	95.25±0.21	[95.25, 95.67]	[95.12, 95.56]	[95.04, 95.46]
TCN	92.36±0.22	94.26±0.23	94.89±0.24	[92.09, 92.63]	[94.00, 94.52]	[94.61, 95.17]
BiLSTM	93.28±0.19	94.17±0.21	94.41±0.23	[93.07, 93.49]	[93.96, 94.38]	[94.16, 94.66]
TCN-BiLSTM (without Dropout)	94.63±0.21	94.51±0.22	94.47±0.23	[94.40, 94.86]	[94.28, 94.74]	[94.24, 94.70]

TCN-BiLSTM (without BatchNorm)	94.81±0.19	94.33±0.20	94.59±0.22	[94.61, 95.01]	[94.13, 94.53]	[94.37, 94.81]
TCN-LSTM	94.59±0.18	93.29±0.22	94.65±0.20	[94.39, 94.79]	[93.03, 93.55]	[94.45, 94.85]

Table 3 presents a comparison of the model performance under different structural configurations and regularization settings. The full TCN-BiLSTM model achieves the best overall results, with an accuracy of 95.46%, recall of 95.34%, and an F1 score of 95.25%. In contrast, the TCN-only model achieves a recall of 94.26% but has a lower accuracy of 92.36%, indicating strong local feature extraction but weaker classification precision. The BiLSTM-only model improves accuracy to 93.28% and achieves an F1 score of 94.41%, showing better performance in modeling global context, but with limited ability to capture short-term dependencies.

From a structural integration perspective, the TCN-BiLSTM model effectively combines local and global representations, resulting in superior performance in sentiment classification tasks. Further analysis on regularization shows that removing the dropout layer reduces the F1 score to 94.47%, while removing the batch normalization yields an F1 score of 94.59%. Both results are lower than the full model, suggesting that regularization plays a key role in improving stability and generalization. Overall, the ablation study confirms the effectiveness of each module and supports the rationality of the proposed model design.

The performance of TCN-LSTM is relatively balanced, the accuracy rate is 94.59%, and the F1 score is 94.65%. However, due to the unidirectional LSTM structure, there is still a problem of insufficient understanding of bidirectional semantics. Compared with TCN-LSTM, TCN-BiLSTM performs better in emotion classification tasks, mainly due to the bidirectional modeling ability of BiLSTM, which can simultaneously capture the dependency between the preceding and following sentences. However, the LSTM in TCN-LSTM is a one-way structure that can only be modeled forward, and its comprehension ability is limited for texts with emotional transition, contrast, or summary sentence patterns. For example, when confronted with comments such as "the scenic area environment is good, but there are too many people and too much noise," TCN-LSTM is susceptible to being dominated by the positive emotions present in the first paragraph and disregarding the negative tendency exhibited in the second half of the sentence. This results in an emotional judgment bias. In addition, in texts containing negative constructs or mood reversals, such as "The hotel was supposed to be good, but the service attitude let me down," the CN-LSTM often struggled to effectively identify the dominance of the subsequent emotional intensity. Therefore, in terms of structural design, TCN-BiLSTM is more suitable for dealing with common complex sentences, multi-emotional sentences, and semantic contrast sentences in tourism texts. Moreover, it has stronger discriminative ability and adaptability. In addition, the training efficiency of the comparison algorithms is shown in Figure 7.

In Figure 7, the average training time (ATT) of the TCN-BiLSTM algorithm model is 43.8 seconds. The ATT for the TCN and BiLSTM algorithms, on the other hand, is 45.32 seconds and 48.29 seconds, respectively. The ATT for TCN-LSTM is 45.12 seconds. In comparison, TCN-BiLSTM has higher algorithmic performance.

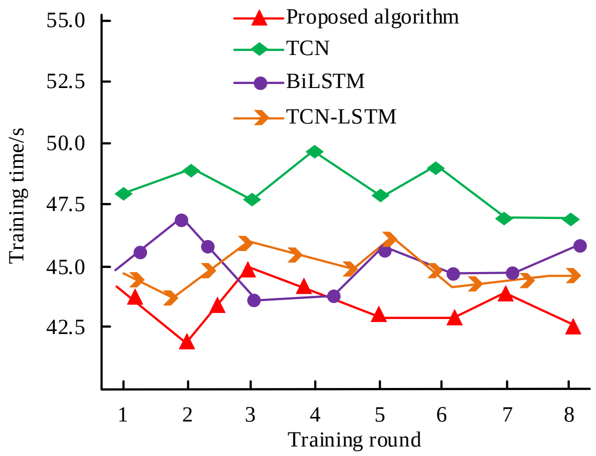


Figure 7: Comparison of model training efficiency of different algorithms

This is because the TCN-BiLSTM algorithm can utilize the CL of TCN to achieve parallel computing, which has a higher utilization of computational resources, thus increasing the speed of the model to process data. From the above data, it can be concluded that the TCN-BiLSTM algorithm model shows higher training efficiency and training accuracy on the training data, and better performance on text data processing, which can be applied to tourism evaluation TSA. Next, the study will verify the effect of this algorithm for TSA.

The pre-trained language models based on the Transformer architecture, such as BERT, RoBERTa, and XLNet, have made significant progress in TSA tasks. Moreover, these are widely considered to be the current mainstream SOTA models. To further verify the effectiveness of the proposed TCN-BiLSTM model, this study selects three representative models, BERT-base, RoBERTa-base, and XLNet-base. It also conducts comparative experiments on the same tourism consumer review dataset to evaluate their classification performance and training efficiency. The results are shown in Table 4.

Table 4: Performance comparison between the proposed model and the Transformer model

Algorithm	Accuracy/%	Recall rate/%	F1 score/%	Average training time (s)
TCN-BiLSTM	95.78	95.62	95.98	40.18
BERT-base	94.80	94.63	94.62	89.14
RoBERTa-base	95.10	94.81	94.08	93.46
XLNet-base	94.48	94.72	94.13	97.02

In Table 4, although Transformer-based models perform well in sentiment classification tasks, their training times are generally longer and they require more computational resources. For example, the training time of XLNet-base reaches 97.02 seconds, while RoBERTa-base and BERT-base require 93.46 seconds and 89.14 seconds, respectively. In contrast, the proposed TCN-BiLSTM model achieves higher accuracy (95.78%) and F1 score (95.98%) with a significantly lower training time of only 40.18 seconds, demonstrating superior computational efficiency and practical value.

Furthermore, Transformer models typically rely on large pre-trained corpora and complex hyper-parameter tuning, which pose higher barriers to domain adaptation. In comparison, the MMN architecture proposed in this study requires no pre-training and can achieve excellent classification performance on medium-sized tourism review datasets. This makes it particularly suitable for resource-constrained application environments, such as mobile tourism applications and real-time review monitoring systems on online travel platforms. Thus, the TCN-BiLSTM model achieves a well-balanced trade-off between performance and computational cost, offering strong potential for real-world applications.

4.2 TSA validation

Sentiment analysis of tourism review text typically involves three aspects: sentiment polarity classification, sentiment intensity analysis, and sentiment dimension identification. Sentiment polarity refers to the overall emotional orientation expressed in the text-positive, negative, or neutral-and serves as a key indicator of the consumer's overall tourism experience. Sentiment intensity reflects the degree or strength of the emotion conveyed, while sentiment dimension refers to the specific aspects being evaluated, such as service quality, scenic environment, or price experience. The study uses Python technology for web crawling to obtain reviews of famous tourist attractions in Hunan, and analyzes the evaluation text sentiment with TCN-BiLSTM model. The results are shown in Figure 8.

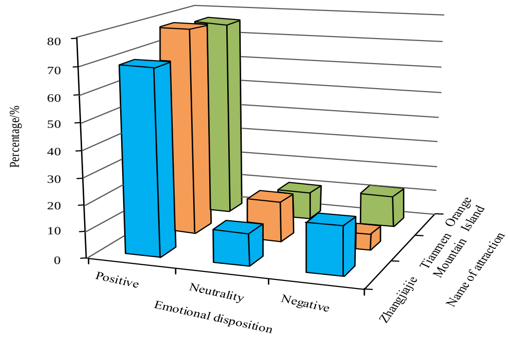
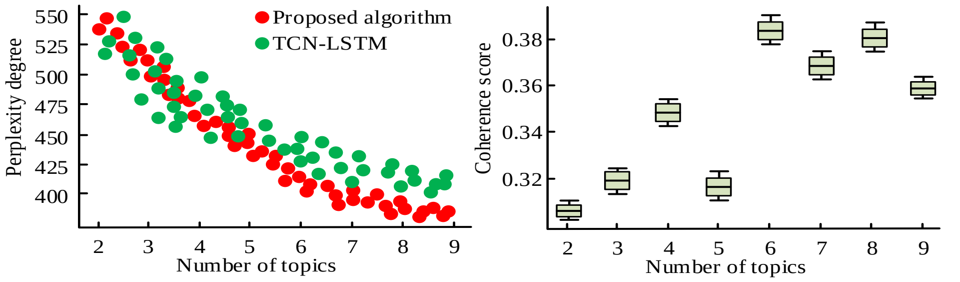


Figure 8: Sentiment analysis results of tourism text

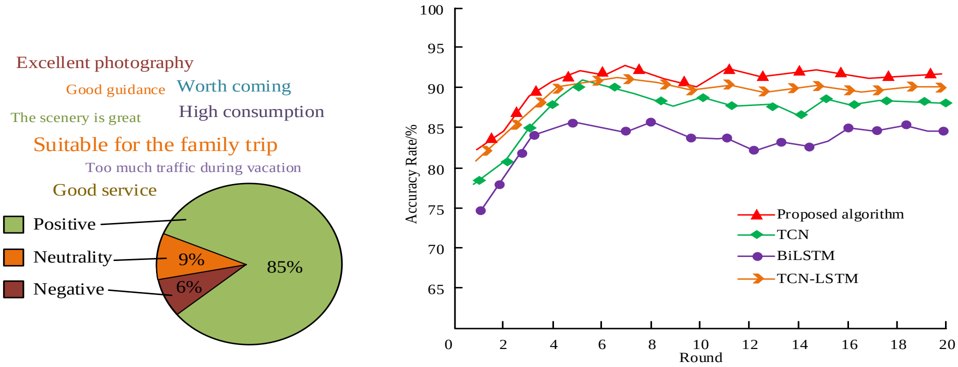
In Figure 8, the ratings of the three famous tourist attractions are generally dominated by positive ratings. The average proportion of positive evaluations is 74.8%, the average proportion of neutral evaluations is 12.9%, and the average proportion of negative evaluations is 12.3%. The sentiment tendency of the evaluations may be accurately analyzed by the TCN-BiLSTM algorithm model, it can be concluded. This is because the TCN algorithm can effectively extract local as well as global features of the time series to capture the emotional changes of the text, while the BiLSTM algorithm is able to do bi-directional information processing, which is able to more comprehensively understand the emotional expression of tourism evaluation. According to TSA results, the analysis effect is superior and the TCN-BiLSTM model has a reasonable level of validity. The extraction effect is further verified, as shown in Figure 9.

In Figure 9(a), the text theme extraction effect is mainly indicated by the perplexity degree. The lower the perplexity degree, the more accurate the effect of theme extraction. A thorough comparison of the text theme extraction effects of various algorithms reveals that the TCN-BiLSTM model has the best text theme extraction impact when the number of themes approaches nine. This is because the model has the lowest perplexity. However, if the number of themes is too high, the model is prone to overfitting. Figure 9(b) illustrates the relationship between the number of topics and topic coherence. The coherence score reaches its peak when the number of topics is 6, suggesting that six is the optimal number of latent topics in the review dataset. Summarizing the Hunan tourism consumer reviews, the proportion of their evaluation sentiment categories and the algorithm fitting effect are shown in Figure 10.



(a)Plot of number of topics against perplexity (b)Plot of number of themes against coherence

Figure 9: Topic number selection based on perplexity and coherence analysis



(a)Emotional cloud maps and statistics (b)Algorithm accuracy comparison

Figure 10: Sentiment analysis results and algorithm accuracy

In Figure 10(a), the high-frequency words calculated by the TCN-BiLSTM model constitute a sentiment cloud map, and its judgments of positive, negative, and neutral for evaluations are summarized as a fan statistic map. Combined with the model accuracy in Figure 10(b) for a comprehensive judgment, the TCN-BiLSTM model proposed in the study has a certain degree of accuracy in sentiment analysis, and can effectively differentiate between tourism evaluations of different sentiments. On the final data derived, the OV of the TCN-BiLSTM model is close to the real value. The tourism evaluation algorithm built on the basis of TCN-BiLSTM can effectively analyze the sentiment trend of tourism evaluation, which facilitates the efficient and accurate management of tourist attractions.

5 Discussion

The experimental results demonstrated that the proposed TCN-BiLSTM model outperformed traditional deep learning architectures in sentiment classification tasks for tourism review texts. This performance gain was mainly attributed to the combination of TCN's ability to extract local and hierarchical temporal features and BiLSTM's strength in modeling bidirectional contextual dependencies. This fusion enabled the model to better capture emotional shifts, polarity reversals, and nuanced sentiments that were commonly found in tourism-related user-generated content.

Despite its advantages, the model had certain limitations. First, while the TCN-BiLSTM architecture was more efficient than transformer-based models, it still required relatively long training times compared to simpler models such as standard CNN or LSTM. Second, the model was less effective when dealing with long texts where dependencies span multiple paragraphs, or when sentiment was implicit or highly contextual. In addition, scalability to very large datasets or multilingual scenarios remained an open challenge.

With regard to domain applicability, the TCN-BiLSTM model exhibited potential for transfer to other sentiment analysis tasks, including product reviews, social media comments, and service feedback, provided that the text manifested analogous sequential and emotional characteristics. However, domain adaptation and retraining would be necessary to maintain accuracy.

Compared to the most advanced transformer-based models such as BERT, RoBERTa and XLNet, the TCN-BiLSTM model achieves higher classification accuracy with significantly lower training costs and computational requirements. This makes it particularly suitable for real-time, resource-constrained applications such as mobile travel applications or online review monitoring systems. However, for tasks involving complex sentiment structures or fine-grained sentiment classification, Transformer models may still offer advantages due to their stronger global attention mechanisms and pre-trained knowledge representations.

6 Conclusion

To enhance the accuracy of tourism evaluation TSA and to realize the sustainable development of tourism consumption program, the study proposed a multi-layer neural network model based on TCN-BiLSTM algorithm. This model mainly processed the input data information by TCN algorithm and then outputs the vector information. The features were extracted by BiLSTM algorithm, and feature fusion and classification were performed, and finally TSA was performed. The outcomes of the study indicated that the TCN-BiLSTM algorithm had a test accuracy of 95.46% and an F1 score of 95.25%. Compared with other traditional algorithms, its accuracy and F1 score were improved by 2.05% and 0.6%, respectively. In terms of the efficiency of model training, the ATT of TCN-BiLSTM was 43.8 s, which was more efficient than the traditional TCN and BiLSTM algorithms with an average improvement of 2.44 s. The ATT of TCN-BiLSTM was 43.8 s, which was more efficient than the traditional TCN and BiLSTM algorithms. Meanwhile, the TCN-BiLSTM algorithm also had a lower loss rate with higher fit and feature learning ability. The TCN-BiLSTM algorithm also had the lowest perplexity when analyzing the topics of tourism evaluation, which further proved the superiority of the TCN-BiLSTM algorithm in the analysis of tourism evaluation. Meanwhile, after the actual analysis of the tourism evaluation text, it can be concluded that the TCN-BiLSTM algorithm could effectively distinguish the emotional tendency of the text with high accuracy. In summary, TCN-BiLSTM has good performance in sentiment analysis of tourism evaluation text. With its high accuracy rate and high algorithm training efficiency, it has good performance in practical applications. Nevertheless, further enhancements can be made to this model. In subsequent studies, the incorporation of an attention mechanism may be a valuable avenue for exploring the dynamic allocation of attention to data of varying importance, with the aim of further optimising the model's performance.

However, this study specifically targets sentiment analysis in the tourism domain, and the dataset and model design are tailored to the linguistic and emotional characteristics of tourism-related user reviews. As a result, no external validation has been performed on reviews from other domains, such as e-commerce or social media. This is a limitation in assessing the cross-domain generalizability of the model. In future work, the proposed framework will be evaluated for its robustness and transferability to broader sentiment analysis applications by incorporating external datasets from different domains.

Fundings

The research is supported by: 2023 Ministry of Education Industry-University Cooperation Collaborative Education Project, Research on Curriculum Teaching Reform of Tourism Consumer Behavior in the Digital Economy Era. (NO. 230806642301615).

References

[1] Hung L P, Alias S. Beyond sentiment analysis: A review of recent trends in text-based sentiment analysis and emotion detection. Journal of Advanced Computational Intelligence and Intelligent Informatics, 2023, 27(1): 84-95. https://doi.org/10.20965/jaciii.2023.p0084

[2] Liu H, Chen X, Liu X. A study of the application of weight distributing method combining sentiment dictionary and TF-IDF for text sentiment analysis. IEEE Access, 2022, 10(2): 32280-32289. https://doi.org/10.1109/ACCESS.2022.3160172

[3] Zhu T, Li L, Yang J, Zhao S, Liu H, Qian J. Multimodal sentiment analysis with image-text interaction network. IEEE transactions on multimedia, 2022, 25(1): 3375-3385. https://doi.org/10.1109/TMM.2022.3160060

[4] Jiang W, Zhou K, Xiong C, Du G, Ou C, Zhang J. KSCB: A novel unsupervised method for text sentiment analysis. Applied Intelligence, 2023, 53(1): 301-311. https://doi.org/10.1007/s10489-022-03389-4

[5] Alwehaibi A, Bikdash M, Albogmi M, Roy K. A study of the performance of embedding methods for Arabic short-text sentiment analysis using deep learning approaches. Journal of King Saud University-Computer and Information Sciences, 2022, 34(8): 6140-6149. https://doi.org/10.1016/j.jksuci.2021.07.011

[6] Rahman H, Tariq J, Masood M A. Multi-tier sentiment analysis of social media text using supervised machine learning. Comput. Mater. Contin, 2023, 74(1): 5527-5543. http://dx.doi.org/10.32604/cmc.2023.033190

[7] Mittal D, Agrawal S R. Determining banking service attributes from online reviews: text mining and sentiment analysis. International Journal of Bank Marketing, 2022, 40(3): 558-577. https://doi.org/10.1108/IJBM-08-2021-0380

[8] Ramaswamy S L, Chinnappan J. Review on positional significance of LSTM and CNN in the multilayer deep neural architecture for efficient sentiment classification. Journal of Intelligent & Fuzzy Systems: Applications in Engineering and Technology, 2023, 45(4):6077-6105. https://doi.org/10.3233/JIFS-230917

[9] Wan W, Zhou Y, Chen Y. Interpretable and explainable hybrid model for daily streamflow prediction based on multi-factor drivers. Environmental Science and Pollution Research, 2024, 31(23):34588-34606. https://doi.org/10.1007/s11356-024-33594-2

[10] Li C, Han X, Li M, Rao Z, Liu X, Zhang Q, Liao W, Li G. State-of-health and remaining-useful-life estimations of lithium-ion battery based on temporal convolutional network-long short-term memory. Journal of Energy Storage, 2023, 74(Dec. Pt.B):1.1-1.18. https://doi.org/10.1016/j.est.2023.109498

[11] He Z, Wang X, Li C. A Time Series Intrusion Detection Method Based on SSAE, TCN and BiLSTM. Computers, Materials & Continua, 2024, 78(1):845-871. https://doi.org/10.32604/cmc.2023.046607
[12] Mohammad E, Elham G A. Self-attention (SA) temporal convolutional network (SATCN)-long short-term memory neural network (SATCN-LSTM): an advanced python code for predicting groundwater level. Environmental Science and Pollution Research, 2023, 30(40):92903-92921. https://doi.org/10.1007/s11356-023-28771-8
[13] Rengarjan M, Bindu M R, Ponnuswamy V F G. Revolutionizing sentiment analysis in literary text 'the immortals of meluha' through a hybrid cnn-rnn architecture and advanced feature techniques. Journal of theoretical and applied information technology, 2023, 101(23):7866-7879.
[14] Zhou X, Li W. Local Convolutional Neural Network Based Pop-Up Text Recognition and Sentiment Analysis. International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems, 2024, 32(04):409-431. https://doi.org/10.1142/S0218488524400014
[15] Wang L, Xu X, Chen L Z. M-DA: A Multifeature Text Data-Augmentation Model for Improving Accuracy of Chinese Sentiment Analysis. Scientific programming, 2022, 2022(Pt.9):3264378.1-3264378.13. https://doi.org/10.1155/2022/3264378
[16] Chundi R, Hulipalled V R, Simha J B. Lexicon-based sentiment analysis for Kannada-English code-switch text. IAES International Journal of Artificial Intelligence, 2023, 12(3):1500-1507.
[17] Taneja K, Vashishtha J, Ratnoo S. Transformer Based Unsupervised Learning Approach for Imbalanced Text Sentiment Analysis of E-Commerce Reviews. Procedia Computer Science, 2024, 235(2):2318-2331. https://doi.org/10.1016/j.procs.2024.04.220
[18] Imran A S, Yang R, Kastrati Z. The impact of synthetic text generation for sentiment analysis using GAN based models. Egyptian Informatics Journal, 2022, 23(3): 547-557. https://doi.org/10.1016/j.eij.2022.05.006
[19] Yan X, Yao L, Zhou D. Optimizing Tourism Service Quality in 5G Multimedia Environments Using Deep Learning: A Model-Based Empirical Study. Informatica, 2024, 48(22): 147-161. https://doi.org/10.31449/inf.v48122.6806
[20] Chakravarthi B R, Priyadharshini R, Muralidaran V. Dravidiancodemix: Sentiment analysis and offensive language identification dataset for dravidian languages in code-mixed text. Language Resources and Evaluation, 2022, 56(3): 765-806. https://doi.org/10.1007/s10579-022-09583-7
[21] Mathew L, Bindu V R. Efficient transformer-based sentiment classification models. Informatica, 2022, 46(8). https://doi.org/10.31449/inf.v46i8.4332
[22] Ahmad H O, Umar S U. Sentiment analysis of financial textual data using machine learning and deep learning models. Informatica, 2023, 47(5). https://doi.org/10.31449/inf.v47i5.4673