Rural Homestay Spatial Planning and Design Based on Bert BiLSTM EIC Algorithm in the Background of Digital Ecology

Zhibin Qiu, Junghoon Mok*
Department of Space Design, Hanseo University,
Seosan 31962, Korea

***Abstract*—There is a promising development prospect in the digital ecosystem. In this context, the spatial planning of rural homestays has also received widespread attention. The research aims to better explore the advantages and determine the development direction of rural homestays, while providing two-way demand support for consumers and managers. Therefore, this study combines bidirectional long-term and short-term memory networks, pre-trained models, and emotional information attention mechanisms in deep learning. A new emotional analysis model is proposed. Then it is applied to the spatial planning of homestays near Chengdu Normal University and Chengdu Neusoft University. The experimental results show that the accuracy, recall, and F1 values of the new emotion analysis model proposed in this research reach 94%, 93%, and 94%, respectively. In terms of consumer satisfaction with the spatial location of homestays before and after the renovation, the average score of homestays near Chengdu Normal University increases by 21% compared to before the renovation. The average score of homestays near Chengdu Neusoft University increases by 40% compared to before the renovation. In summary, the new emotional analysis model proposed in this research has certain feasibility and effectiveness in the planning of rural homestay spatial location, providing new ideas for homestay spatial location planning.**

Keywords*—Rural homestay; spatial planning; deep learning; emotional analysis; bidirectional long short term memory network

I. INTRODUCTION

As a key support project for the tourism industry, rural homestays have been consistently popular in recent years [1]. With the changing demand for tourism and vacation methods, the market size of the homestay industry is constantly expanding. The rise of new homestays has gradually replaced traditional farmhouses and inns [2]. Featured homestays are constantly emerging. For example, urban-rural integration and distinctive homestay planning meet personalized pursuits and experiences [3]. Abroad, homestay platforms such as Airbnb have established a huge network of accommodation resources worldwide. This specialized operational transformation further improves the quality and image of homestays [4]. In Japan, the homestay policy supports the transformation of private housing into homestays. But the operational safety and good environment of homestays need to be guaranteed [5]. However, there are currently many problems in the planning and design of homestay spaces in the Chinese market, such as unreasonable layout and inability to meet consumer needs. Inappropriate planning methods may cause a series of negative reactions.

In view of this, the research explores in the context of digital ecology and innovatively proposes the integration of Bidirectional Long ShortTerm Memory (BiLSTM), Natural Language Processing and specific elements of spatial planning. Meanwhile, considering the growth and change of the rural lodging industry, a lodging spatial location planning and design model is finally proposed. The aim is to solve the unreasonable spatial location planning in domestic homestays, provide effective advice and guidance for the spatial location planning and design of homestays, and promote the development of rural homestays. This study is divided into six sections. Section I is an introduction to the overall content. Section II outlines the related works. Section IV is about performance testing of rural homestay spatial planning model. Discussion and conclusion is given in Section V and Section VI respectively.

II. RELATED WORKS

BiLSTM is a deep learning algorithm that can effectively process sequence data. In recent years, this algorithm has achieved significant results in fields such as natural language processing and image recognition. Singla P et al. proposed an integrated model that combines wavelet transform with bidirectional long-term and short-term memory deep learning networks to predict solar irradiance at all levels within 24 hours. Wavelet transform decomposes the input time series data into different intrinsic model functions to extract its statistical features. To improve prediction accuracy, the sequences of intrinsic model functions are reduced through comprehensive experimental analysis. Wavelet decomposition components are merged. Next, each intrinsic model function subsequence is assigned a trained BiLSTM network for prediction. Finally, the predicted values of each subsequence reconstructed from the BiLSTM network are used to provide the final solar irradiance prediction at all levels. The research results indicate that the R2 value of the proposed model is 0.94. Compared to the benchmark model, prediction performance increases by 47% [6]. To predict the capacity of photovoltaic energy, Liu B et al. proposed a multi factor installed capacity prediction model based on bidirectional short-term memory grey correlation analysis. The solar photovoltaic installed capacity in China from 2020 to 2035 is predicted using this model. The research results indicate that the constructed model has high prediction accuracy. It can accurately predict China's solar photovoltaic installed capacity from 2020 to 2035, reaching 2833GW by 2035 [7]. Human activity recognition has become an important research field in human behavior analysis, human-computer interaction, and ubiquitous computing. To improve the accuracy of deep learning models in processing time series data, Challa S K et al. proposed a robust classification model for human activity recognition. This model combines convolutional neural networks and bidirectional long-term and short-term memory for design. The proposed multi branch CNN-BiLSTM network can automatically extract features from raw sensor data and minimize data preprocessing. CNN-BiLSTM can learn local features and long-term dependencies in sequence data, which helps improve the feature extraction process. The performance is evaluated using three benchmark datasets. It achieves accuracy of 96.05%, 96.37%, and 94.29% on these datasets, respectively. The experimental results show that the method combining multi-branch CNN-BiLSTM networks outperforms the other compared methods [8].

Spatial planning and design commonly referred to as "spatial design" or "spatial planning". It refers to the process of creating an indoor space that combines functionality and aesthetics in an orderly manner. Space planning and design is the allocation of appropriate functional areas for a specific environment, determining the layout and interaction of these areas. There are currently many related studies. Design space exploration provides intelligent adjustment methods for complex optimization parameters in modern advanced comprehensive space design tools. Due to the long hardware compilation time, advanced comprehensive parameter tuning is a time-consuming process. Gautier Q et al. utilized a design space exploration framework to address multiple conflicting optimization objectives and actively sought Pareto optimal solutions. The research results indicate that the adopted processing framework can achieve optimal spatial design faster. In addition, the design space exploration framework can also customize the regression model based on specific problems, thereby obtaining the model that best reflects the application design space [9]. Wu W et al. proposed a novel data-driven technology. This technology can automatically and efficiently generate floor plans for residential buildings with given boundaries. The proposed data-driven technology can mimic human spatial design processes. Firstly, the room position is determined, and then the wall position is determined, while adapting to the input building boundary. Finally, based on the spatial location of the building, a plan structure diagram is obtained. A large number of experimental results indicate that the proposed data-driven technology is effective. This technology can realistically simulate the floor plans of different floors of a building. In many cases, the floor plan generated by the research method is almost identical to the actual design plan [10].

In summary, many experts have completed a series of studies using BiLSTM, including various data prediction, action recognition, model optimization, etc [11]. For spatial planning and design problems, scholars often start with optimizing complex parameters in space. Various models are used to tune parameters to achieve the expected spatial optimization goals. For buildings like rural homestays, there is a lack of research on integrating deep learning models into the spatial planning and design of homestays. Therefore, based on the BiLSTM algorithm, an improved sentiment analysis model is proposed by combining the Bert model and emotional attention mechanism. The aim is to quickly and accurately understand customer evaluations and emotional tendencies for homestay operators, provide decision-making support, and promote the healthy development of the homestay industry.

III. Rural Homestay Spatial Planning based on Bert-BiLSTM-EIC Algorithm

There are still many problems in the current spatial planning and design of homestays, such as traffic positioning, location planning, and consumer preferences. In response to these issues, the first section constructs a new emotional analysis model through a bidirectional long-term and short-term memory network, Bert model, and emotional information attention mechanism. It effectively extracts relevant suggestions from consumers for the geographical location planning of homestays. The second section applies the model to the spatial planning of homestays near two universities, exploring the feasibility of this method.

A. Construction of a Homestay Sentiment Analysis Model Based on Bert-BiLSTM-EIC Algorithm

In recent years, the sentiment analysis method that combines Long Short Term Memory (LSTM) with deep learning has the highest popularity. LSTM updates the time record timing of cell structure through an additive calculation method, avoiding the time step feature retained when the previous state data has a significant impact on this state data, as shown in Fig. 1[12].

In Fig. 1, at time t, a value $f(t)$ between 0 and 1 is output through the internal states of $h(t-1)$ and $x(t)$, representing a trade-off between fully retained or discarded states. The activation function is the sigmoid function. The formula for the input gate is shown in Eq. (1).

$$i_t = \sigma\left(W_i \cdot x_t + U_i \cdot h_{t-1} + b_i\right) \quad (1)$$

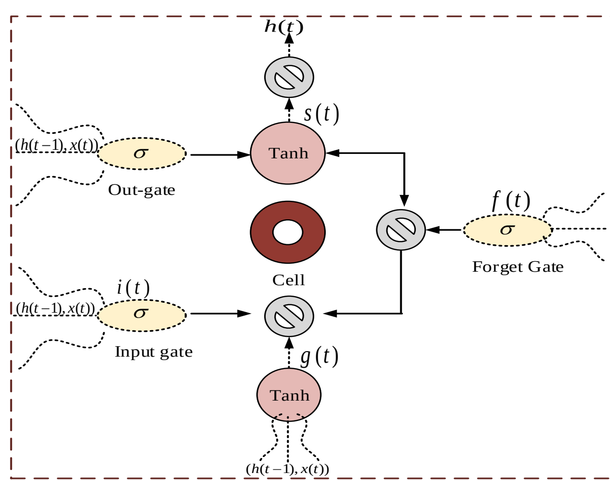


Fig. 1. LSTM network structure diagram.

In Eq. (1), i_t represents the output of the input gate. W_i and U_i represent the weight and bias of the input gate, h_{t-1} denotes the previous state. The input gate represents the information stored in $h(t-1)$ and $x(t)$ within the cell. The output gate is shown in Eq. (2).

$$o_t = \sigma\left(W_o \cdot x_t + U_o \cdot h_{t-1} + b_o\right) \quad (2)$$

In Eq. (2), x_t denotes the output data, σ denotes the sigmoid function. W_o and U_o denote the weight and bias of the output gate, respectively, and b_o denotes the learnable parameters. The information of the output gate is usually obtained by multiplying the result of the tanh activation function with $h(t-1)$ and $x(t)$ through the sigmoid activation function [13]. The forgetting gate is shown in Eq. (3).

$$f_t = \sigma\left(W_f \cdot x_t + U_f \cdot h_{t-1} + b_f\right) \quad (3)$$

In Eq. (3), f_t denotes the output of the forgetting gate, W_f and U_f denote the weight and bias of the forgetting gate, respectively, and b_f denotes the learnable parameter. The error function calculates the partial derivative of the weighted input of the neuron. The backpropagation error transmitted by time to the previous state is shown in Eq. (4).

$$\delta_k^T = \prod_{j-k}^{t-1} \delta_{o,t}{}^{,T} W_{oh} + \delta_{f,t}{}^{,T} W_{fh} + \delta_{i,t}{}^{,T} W_{ih} + \delta_{s,t}{}^{,T} W_{sh} \quad (4)$$

In Eq. (4), δ_k^T refers to the error of the objective function k with respect to T. W_{oh}, W_{fh}, W_{ih} and W_{sh} refer to the weights that go sequentially from unit o, f, i and s to unit h, respectively. $\delta_{o,t}$ refers to the error of the o th unit at the moment of t. $\delta_{f,t}$ refers to the error of the f th unit at the moment of t. $\delta_{i,t}$ refers to the error of the i th unit at the moment of t. $\delta_{s,t}$ denotes the error of the s th unit at the moment of t. The error formula for error propagation upwards is shown in Eq. (5).

$$\frac{\partial E}{\partial net_t^{l-1}} = (\delta_{o,t}{}^{T} W_{ox} + \delta_{f,t}{}^{T} W_{fx} + \delta_{i,t}{}^{T} W_{ix} + \delta_{s,t}{}^{T} W_{sx}) of\,'(net_t^{l-}1) \quad (5)$$

In Eq. (5), W_{ox}, W_{fx}, W_{ix} and W_{sx} refer to the weights from cell o, f, i and s to the formula cell x, respectively. The rest of the algebra is consistent with the previous explanation. However, due to the complexity and massive content of comments on homestays, relying on a single network model often leads to low efficiency and computational errors [14]. Therefore, BiLSTM with bidirectional channels is needed to process sequence data. On the basis of BiLSTM, Bert word embedding and EIC emotional information attention mechanism are introduced. The Bert-BiLSTM-EIC model is proposed. The structure diagram of the Bert-BiLSTM EIC model is shown in Fig. 2.

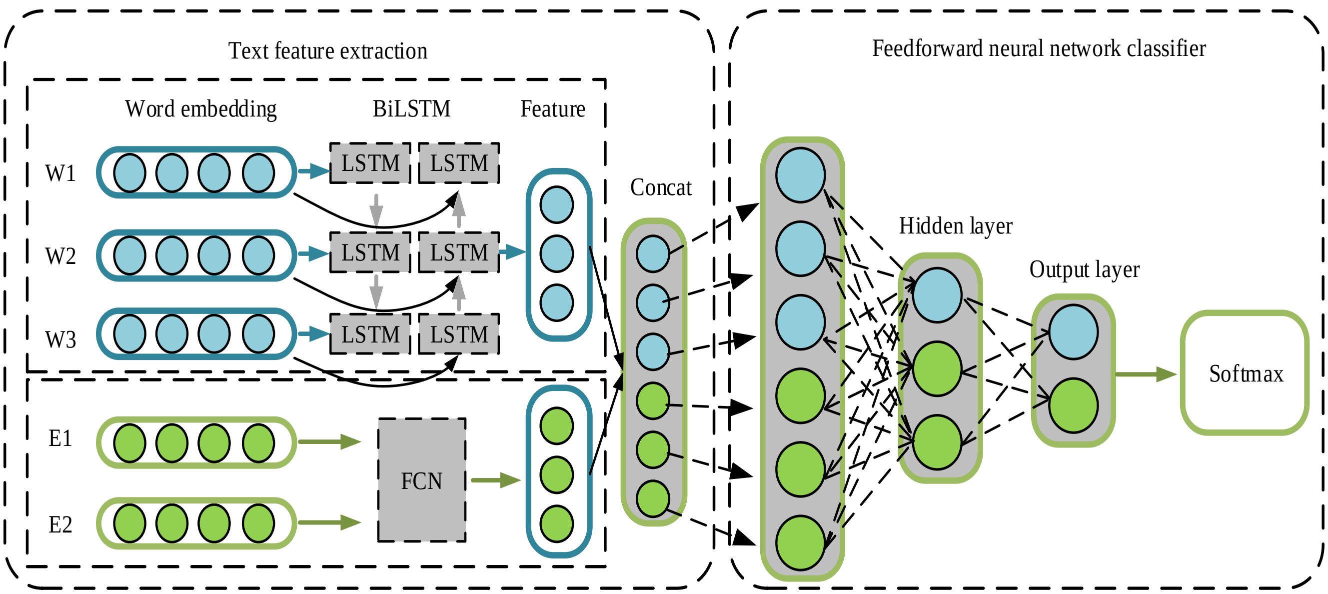


Fig. 2. Bert-BiLSTM-EIC network architecture diagram.

In Fig. 2, the Bert training model is used as the first stage. BiLSTM-EIC is used as the second stage. BiLSTM-EIC includes semantic information channels and emotional information channels. Firstly, Bert passes each input individual word through a Token embedding layer to convert each word segment into a fixed dimensional vector form. Secondly, BiLSTM-EIC extracts the emotional features of the statement, fuses the feature information and inputs it into the fully connected layer. Finally, softmax calculates the attribution probability. The Bert word embedding layer is shown in Eq. (6).

$$R_w = w_1 \oplus w_2 \oplus \cdots \oplus w_n \tag{6}$$

In Eq. (6), R_w represents the input semantic word vector. w_1, w_2, w_3 represent different semantic participles. The word set after segmentation is used as input for the semantic channel. The Bert model obtains the semantic word input vector. The calculation of emotional channels is similar to the input of semantic channels. The calculation is shown in Eq. (7).

$$R_e = e_1 \oplus e_2 \oplus \cdots \oplus e_m \tag{7}$$

In Eq. (7), R_e represents the input emotional information vector. e_1, e_2, e_3 represents different emotional participles. The semantic and emotional vectors are input into a network composed of forward and backward LSTM for sequence encoding to achieve information extraction and fusion, thereby improving the predictive ability of the model [15]. The calculation for model prediction is shown in Eq. (8).

$$pre = softmax(w_o * V^{s+e} + b_o) \tag{8}$$

In Eq. (8), w_o represents the weight coefficient. pre represents the predicted emotional category. b_o represents the offset coefficient. V^{s+e} denotes the sentiment feature vector. The prediction accuracy of this model is shown in Eq. (9).

$$CV_{(k)} = \frac{1}{k}\sum_{i=1}^{k} P_i \tag{9}$$

In Eq. (9), k represents the number of datasets. i represents the combination method. P_i represents the accuracy indicator. The loss function of this model is shown in Eq. (10).

$$f_{loss} = -\frac{1}{N}\sum_{i=1}^{N}\sum_{j=1}^{M} y_i^j \log \hat{y}_i^j + \lambda \|\theta\|^2 \tag{10}$$

In Eq. (10), N is the number of samples. M is the number of emotional labels. y is the emotional label value. $\hat{y}$ is the predicted value of the model. λ is the normalization term. The preference scoring of subsequent users is shown in Eq. (11).

$$H = \begin{bmatrix} ru_1^n \cdots ru_k^n \\ \vdots \\ ru_1^m \cdots ru_k^m \end{bmatrix} \tag{11}$$

In Eq. (11), H represents the scoring set. n and m represent the regional category of homestays. k represents the number of homestays. At this point, the potential score is shown in Eq. (12).

$$R = \begin{bmatrix} r_1u_n \cdots r_ku_n \\ \vdots \\ r_1u_m \cdots r_ku_m \end{bmatrix} \tag{12}$$

In Eq. (12), R represents the consumer record matrix. r_1u_n represents the number of homestay consumption in n. r_1u_m represents the number of homestay consumption in m. The location similarity of the two homestays is shown in Eq. (13).

$$similarity(h_1, h_2) = \frac{\cos(s_1, s_2 \cdots s_3, s_4)}{e^{dis(h_1, h_2)}} \tag{13}$$

In Eq. (13), $dis(h_1, h_2)$ represents the distance between two places. s_1, s_2, s_3 and s_4 conveniently represents location, transportation, environment, and resource ratings. The sentiment analysis process combined with the BiLSTM-EIC model is shown in Fig. 3.

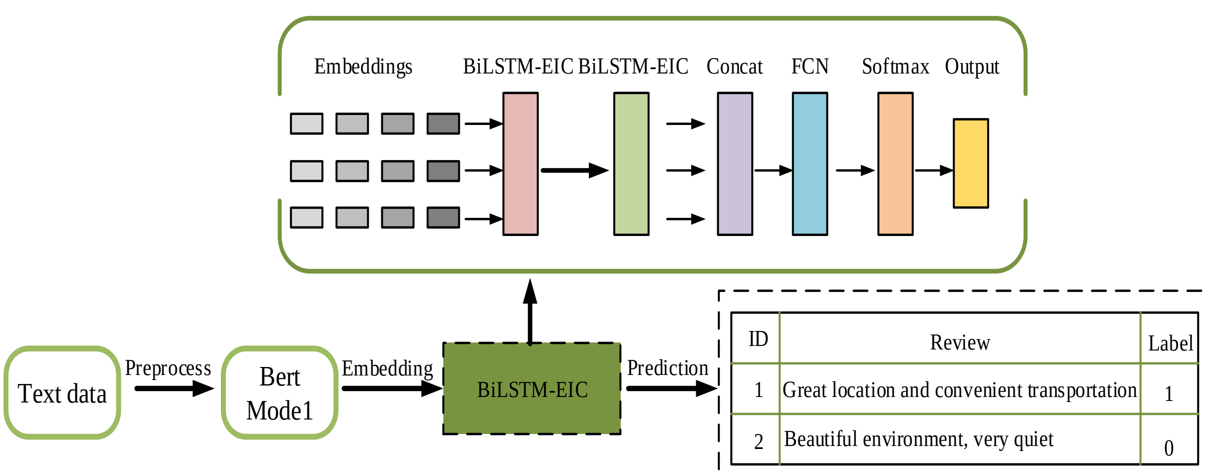


Fig. 3. Emotion analysis flow of BiLSTM-EIC model.

In Fig. 3, step 1 inputs a comment dataset. Step 2 applies the Bert word embedding method to extract and embed sentences from the dataset into the tensor space. Step 3 inputs these different vectors into the BiLSTM model for semantic and feature extraction. Step 4 calculates and predicts feature information to obtain corresponding emotional results.

B. Spatial Location Planning of Rural Homestay Based on Emotional Analysis Model

After the epidemic, the national tourism industry has surged to a new height. The demand for homestays and accommodation requirements are gradually increasing [16]. According to the 2023 Homestay Industry Insight Report released by Feizhu Travel, the keywords for homestays in the hot search are shown in Table I.

In Table I, the best search terms selected by the public include geographical location, homestay environment, and price. Geographical location has become the most perplexing choice factor for consumers. As a homestay manager, the success of site selection almost determines the subsequent operation of homestays [17]. The study used GIS and market research data to identify optimal sites, with relevant location parameters including geographic location, environmental parameters and plot parameters [18]. The location parameters include 10-20 kilometers from the city center, no more than five kilometers from major tourist attractions, and close to major highways or transportation hubs. Environmental factors include noise levels not exceeding 50 dB and an Air Quality Index (AQI) of less than 100 year-round.The plot parameter refers to a slightly sloping or flat, well-drained soil area within the model's applicability range of 2,000 square meters [19]. Aiming at the above hot semantic words, this study compares the spatial locations of B&B rentals near Chengdu Teachers College and Chengdu Neusoft College by combining the proposed sentiment analysis model after the relevant parameters are determined. Chengdu Normal University is located in the university town area of Wenjiang District, Chengdu City. It has convenient subway and public transportation, and commercial districts gather, belonging to the urban area [20]. Chengdu Neusoft College is located in Qingchengshan Town, Dujiangyan Irrigation Project City. There are buses but no subways. There are fewer business districts and more mountain views. It belongs to a rural area [21]. There are nearly 20000 comments on geographical location provided by Meituan APP for homestays in two places. After randomly selecting the comments, they are input into the sentiment analysis model for feature extraction. The data analysis is shown in Fig. 4.

In Fig. 4, in the context of massive review data, consumers maintain independent opinion on the geographical location characteristics of homestays. After data processing in the emotional semantic analysis model, four main emotional analysis focuses are identified, convenient transportation, superior location, quiet environment, and rich supporting facilities. Based on the above online comment sentiment analysis data, the spatial location planning of the two homestays is designed before and after improvement, as shown in Fig. 5.

In Fig. 5(a) and Fig. 5(c) are the spatial maps of homestays before the renovation design of two universities. Fig. 5(b) and Fig. 5(d) show the spatial planning of homestays after the renovation of two universities. From Fig. 5, the original homestays of Chengdu Normal University are relatively concentrated. They are located in the downtown area. Although the transportation is convenient and shopping is convenient, the living environment and comfort are slightly poor [22]. Therefore, after the renovation, green circles represent quiet and comfortable homestays. The red circle indicates homestays with convenient transportation and superior location. Due to the unique geographical location, Chengdu Neusoft College is mostly located near mountains [23]. Quiet and comfortable homestays have become a trend. Considering the previous trend of loose distribution of homestays, and the difficulties of commercial and transportation areas, this renovation mainly focuses on bus stops and the surrounding areas of rare commercial areas. Consumers can balance both.

TABLE I. HOMESTAY HOT SEARCH KEYWORD LIST

First level keywords	Secondary keywords	Hot search words
Geographical position	Convenient transportation	Convenient
	Superior location	Shortcut
	Security environment	Secure
	Surrounding facilities	Richness
Price distribution	Below 200 yuan	Cheap
	200-500 yuan	Moderate price
	500-1000 yaun	Slightly expensive
	Above 1000 yuan	Too expensive
Residential environment	Lighting	Good indoor lighting
	Ventilate	Good indoor air
	Soundproof	Sleep undisturbed
	Water quality	Clean
Homestay quality	General homestay	Not so bad
	Boutique homestay	Cost performance
	Trendy homestay	Generally not selected
Home style	Minority	Seldom choose
	Rural	Acceptable
	Neo-chinese style	Acceptable
	Simple style	Most popular
Star standard	Five star	Don't have to
	Four-star	Maybe
	Three-star level	Consider moving in
	Two-star level	Consider moving in
	One-star level	Will not check in
Characteristics of homestay	Landscape experience	Family or couple
	Rural experience	Relax
	Artistic experience	Experience
	Sports experience	Recreation
	Cultural experience	For fun

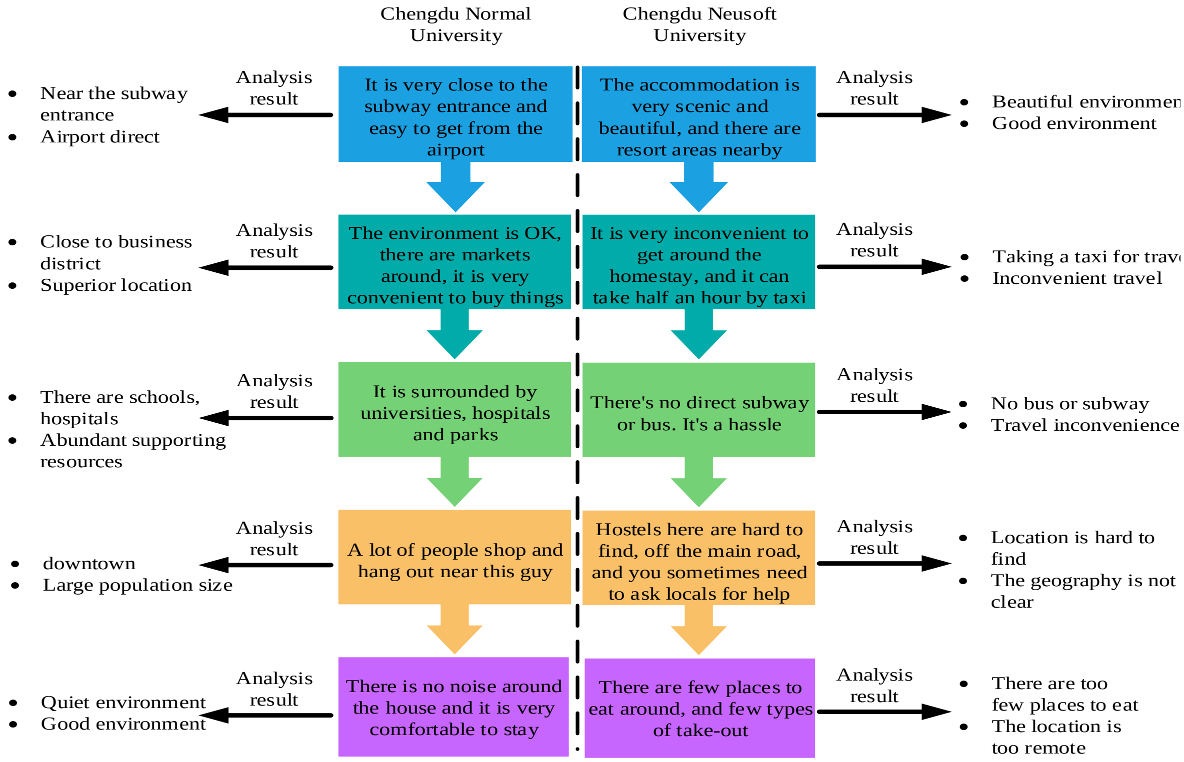


Fig. 4. Analysis of review data of homestays near two universities.

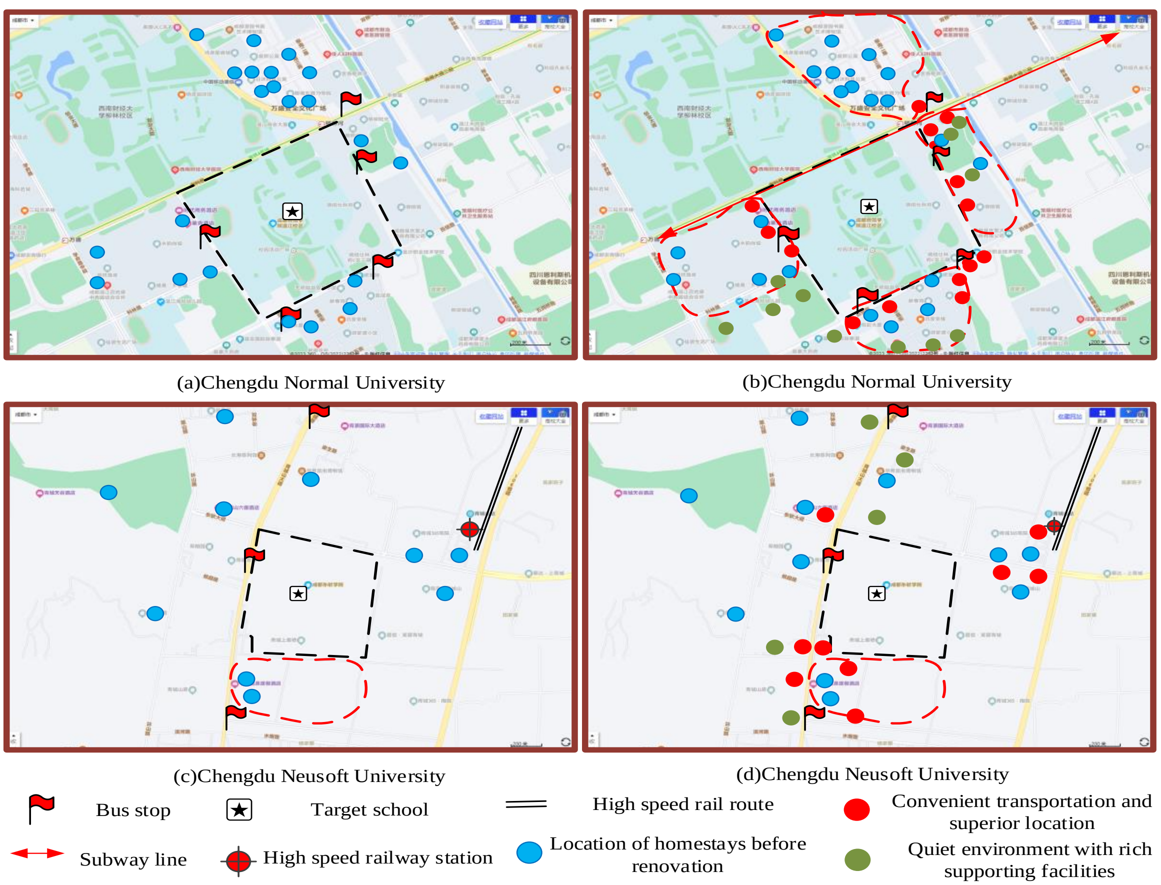


Fig. 5. Emotional analysis model of rural residential space location planning and design.

IV. Performance Testing of Rural Homestay Spatial Planning Model Based on Bert-BiLSTM-EIC Algorithm

The Bert module is used as the first part of the model training to improve its ability to be used to understand and extract contextually relevant features from the input text. This is followed by the BiLSTM layer, which is used to process the features extracted by the Bert layer to capture the long-term dependencies in the text. Finally the EIC layer is trained which is used for used for final sentiment classification based on the output of the first two layers. During the training process, the average length of the text in the training set is too long, while the amount of training data is large. Therefore, the study selectively set the Bert sequence length to 256, BiLSTM hidden layer unit to 128, and the EIC layer parameters to 64. Also, the learning rate of AdaGrade optimizer was set to 0.001, the batch size was set to 32, and the training period was set to five weeks.

Precision is the most intuitive performance metric, which indicates the ratio of instances correctly predicted by the model to the total number of instances. Recall measures the ratio of positive instances correctly identified by the model to all actual positive instances. Recall is particularly important in sentiment analysis because missing instances of positive or negative sentiment can lead to poor business decisions or misinterpreted user feedback. The F1 score is the reconciled average of accuracy and recall, which provides a balance between the two. The F1 score is particularly useful when classes are unevenly distributed or when the penalties for false positives and false negatives are the same. Compared to other metrics, these three metrics are commonly used in sentiment analysis models because they provide a comprehensive view of model performance and are easy to understand and communicate.

Firstly, a suitable experimental environment is established. Accuracy, recall, and F1 are used as evaluation indicators to test the sentiment analysis model fused with the Bert-BiLSTM-EIC algorithm. Secondly, time series, geographic similarity of homestays, and user satisfaction are used as indicators for actual performance testing.

A. *Performance Testing of Sentiment Analysis Model Based on Bert-BiLSTM-EIC Algorithm*

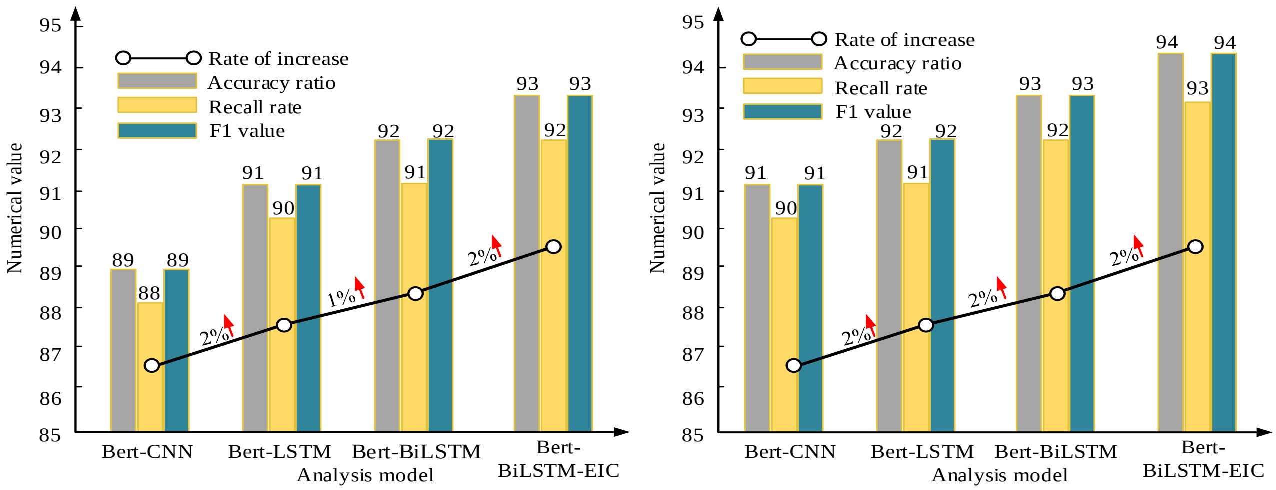
To verify the performance of the emotion analysis model proposed in this experiment, a suitable experimental environment platform is established. A guesthouse comment corpus, ChnSensiCorp, containing both Chinese and English corpora, is used as the dataset. At the same time, to ensure sufficient experimental data and the reliability of the results, a new dataset consisting of approximately 8000 positive comments and 3000 negative comments, totaling 11000, is developed by combining the review data of numerous travel softwares such as Feizhu, Ctrip, and Meituan. It is divided into a training set and a testing set according to 8:2. The specific experimental environment is shown in Table II.

Fig. 6(a) shows the accuracy, recall, and F1 values of four different analysis models in the training set. Fig. 6(b) shows the accuracy, recall, and F1 values of four different analysis models in the test set. From Fig. 6, the CNN model has the worst comprehensive experimental performance, and LSTM is good at capturing temporal features. It is slightly better than the CNN model. BiLSTM has better accuracy in capturing semantic features than LSTM. Based on the above results, the Bert-BiLSTM-EIC proposed in this study performs the best in all three aspects. Its accuracy, recall, and F1 values can reach up to 94%, 93%, and 94%, respectively. The overall score has increased by six percentage points compared to the CNN network. Therefore, the EIC emotional information attention mechanism can extract emotional word sets from text information in the set order, thereby improving the accuracy of semantic and emotional information retrieval and better capturing potential semantic features. To further verify the performance status of the Bert BiLSTM EIC analysis model proposed in this research, the loss function and accuracy are used as reference indicators during the training iteration process. The Bert-CNN model, Bert-LSTM model, Bert-BiLSTM, and Bert-BiLSTM-EIC model are used as experimental objects for testing. The specific test results are shown in Fig. 8.

Accuracy, recall, and F1 value are used as evaluation indicators. The emotion analysis model based on BiLSTM-EIC proposed in this research and three different analysis models, including Bert-CNN model, Bert-LSTM model, and Bert-BiLSTM, are compared and tested. The specific test results are shown in Fig. 6.

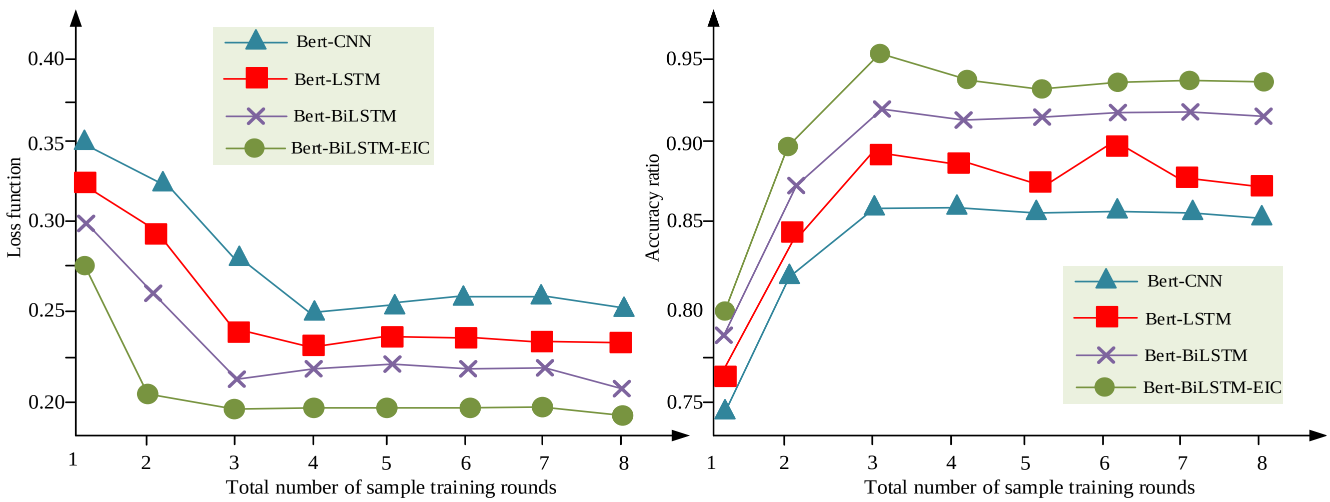
TABLE II. EXPERIMENTAL ENVIRONMENTAL PARAMETERS

Equipment environment and parameter items	Model and specific parameters
CPU	3.6GHz Intel Core i9
GPU	NVIDIA GeForce GTX 3066Ti 8G
Operating system	Windows10 64
Memory	32G
Deep learning library	TensorFlow
Amount of data for one training session	1 article
The total number of rounds required for training samples	5 rounds
Iterations	45000 times
Learning rate	0.001
Optimizer	AdaGrade



(a) Training set (b) Test set

Fig. 6. Graph of test results for four models.



(a) Loss rate comparison chart (b) Accuracy variation chart

Fig. 7. Comparison of loss function and accuracy of different models.

Fig. 7(a) shows the loss function curve for the four analytical models. Fig. 7(b) shows the accuracy curve changes of the four analysis models. In Fig. 7, the accuracy curve of the proposed Bert-BiLSTM-EIC analysis model has always been higher than the other three models. The loss rate is the lowest among the four analysis models. Throughout the entire iteration process, the performance of all four models decreases. At this point, the loss function is negatively correlated with accuracy. Although overfitting training can cause slight data fluctuations, the Bert BiLSTM-EIC analysis model tends to be more stable than the other three models. This model has a dual channel construction of semantic emotions. Therefore, the convergence speed during training is faster than the other three models.

B. Application Test of Emotional Analysis Model in Rural Homestay Spatial Planning

Generally speaking, the emotional consumption patterns of users are influenced by changes in time series. Therefore, based on the spatial location planning maps of Chengdu Normal University and Chengdu Neusoft University, the application effect of the sentiment analysis model in spatial planning is further demonstrated by drawing the emotional time series curves of consumers in the two places over the year. The detailed results are shown in Fig. 8.

Fig. 8(a) shows the changes in the number of reserved homestays, quiet environment homestays, and convenient transportation homestays in the vicinity of Chengdu Normal University during the year before the renovation. Fig. 8(b) shows the changes in the number of reserved homestays, quiet environment homestays, and convenient transportation homestays near Chengdu Normal University during the year after renovation. Fig. 8(c) shows the changes in the number of reserved homestays, quiet environment homestays, and convenient transportation homestays near Chengdu Neusoft College during the year before the renovation. Fig. 8(d) shows the changes in the number of reserved homestays, quiet environment homestays, and convenient transportation homestays near Chengdu Neusoft College after renovation during the year. According to Fig. 8(a) and 8(b), the demand is highest near Chengdu Normal University after the Spring Festival, from June to August, and in November. It may be due to the urgent need to rent a house after the New Year, the summer vacation from June to August for tourism, and an increase in traveling or taking the postgraduate entrance exam in November. From Fig. 8(c) and Fig. 8(d), near Chengdu Neusoft College, due to geographical location factors before the renovation, the average annual total rental volume does not fluctuate significantly. After the renovation, the total number of rented houses has increased by nearly 60. After analyzing consumer emotions, the designed homestay housing location is more favored by the public.

To better understand consumers' preferences for the geographical location of homestays, Eq. (12) is used to calculate the similarity of housing locations. The total rental volume and consumer selection evaluation of homestays after the renovation of the two places are used as the horizontal and vertical coordinates. The scatter plot of homestays in both locations is shown in Fig. 9.

Fig. 9(a) shows the location similarity of homestays near Chengdu Normal University. Fig. 9(b) shows the location similarity of homestays near Chengdu Neusoft College. From Fig. 9, public consumers near Chengdu Normal University prefer homestays with high similarity, that is, homestays with similar geographical locations. At the same time, the higher the similarity of homestay locations, the greater the rental volume is. Consumers have higher evaluation of their choices. However, near Chengdu Neusoft College, the geographical similarity of homestays is generally not high, and their geographical locations are relatively far apart.

To enable homestay operators to quickly and accurately understand customer evaluations and emotional tendencies, provide decision support, and improve service quality, the CSI (Customer Satisfaction Index) is used for evaluation. A satisfaction survey is conducted on consumers in both regions, with a score of 1-5 on the Likert scale from dissatisfied to satisfied. The data results are shown in Table III.

From Table III, the two homestays before and after the renovation have significantly improved in convenient transportation, superior location, quiet environment, and rich resource allocation. Especially in the vicinity of normal colleges, the popularity of homestays with quiet environment, convenient transportation, and superior location has generally increased. The average score increases by 21% compared to before the renovation. After planning the geographical location of homestays near the Soft College, the scores for convenient transportation and location increased significantly. The average score has increased by 40% compared to before the renovation. The proposed sentiment analysis model that integrates the Bert BiLSTM EIC algorithm has shown high feasibility in actual rural homestay spatial transformation testing. It has played a certain reference role in the subsequent spatial location planning of rural homestays.

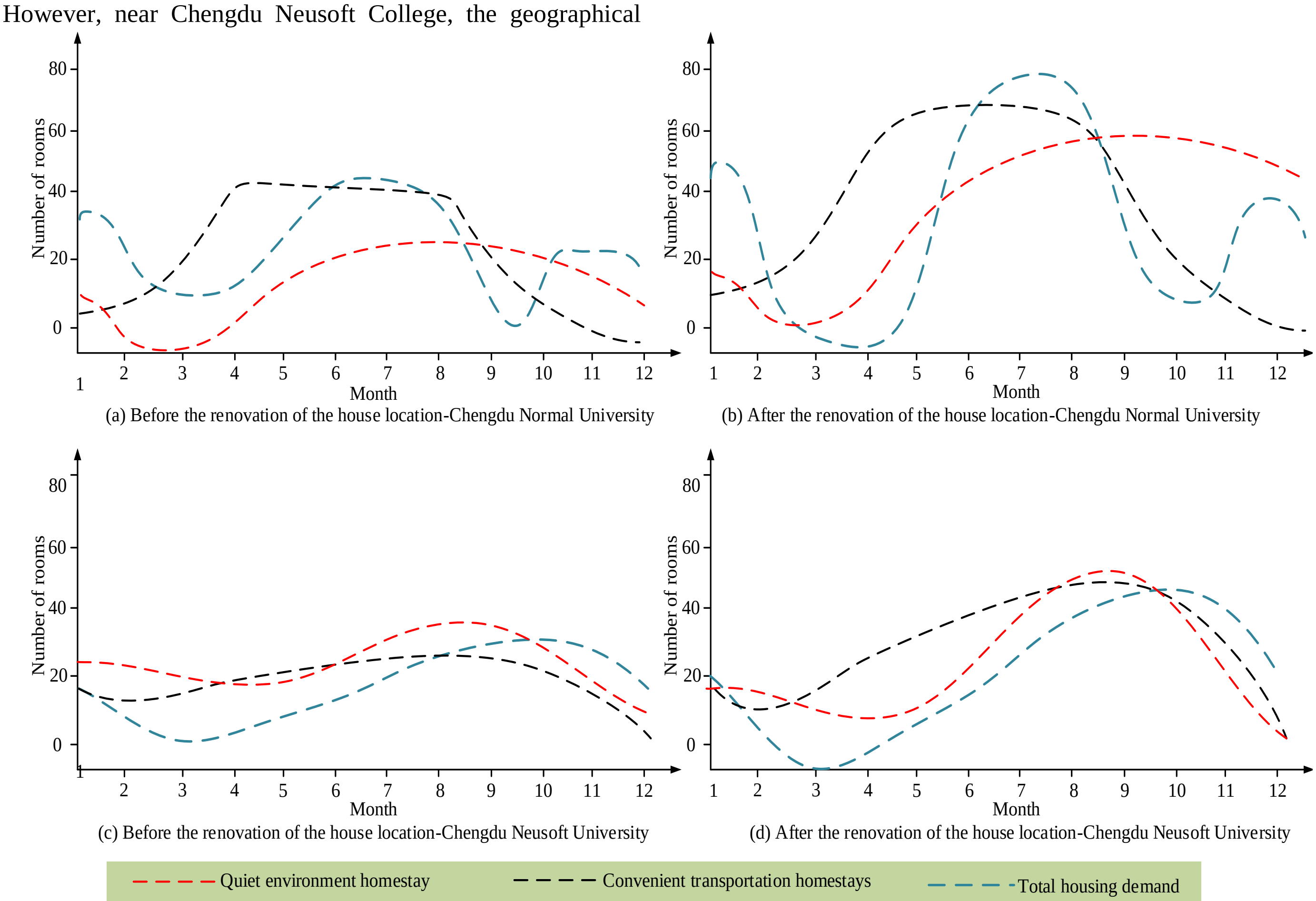


Fig. 8. The change of consumer sentiment in two places in one year.

TABLE III. CONSUMER SATISFACTION RATING

Regional variable			Convenient transportation	Superior location	Quiet environment	Abundant supporting resources	Average score
Chengdu University	Normal	Before modification	4.0	4.0	2.0	3.0	3.3
		After transformation	5.0	5.0	3.0	3.0	4.0
Chengdu University	Neusoft	Before modification	1.0	2.0	5.0	2.0	2.5
		After transformation	3.0	3.0	5.0	3.0	3.5

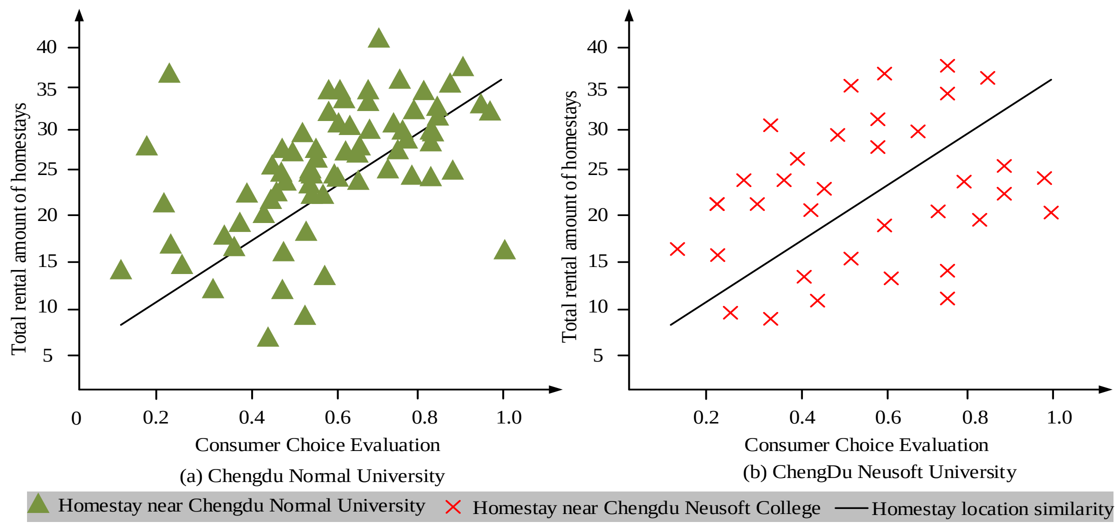


Fig. 9. Scatter map of homestays near two universities.

V. DISCUSSION

For this study, the data results are further expanded through the following three areas of discussion. The first is the feasibility of the algorithm effect and application. This study demonstrates the potential of digital technology in rural revitalization and tourism development by combining the Bert-BiLSTM-EIC algorithms in the spatial planning and design of rural B&Bs. The Bert module shows superiority in understanding the preferences and feedbacks of the B&B customers, while the BiLSTM performs excellently in dealing with the time-series data such as the booking trend and the flow of guests. The EIC module is introduction further optimizes the decision-making process of spatial planning, making it more accurate and efficient. Secondly, the impact of digital ecology, in the context of digital ecology, this study highlights the importance of data-driven decision making. As more and more data becomes available, it enables the public to understand the needs and challenges of the rural lodging industry more fully. The digital ecology not only provides rich data resources, but also offers new perspectives on the sustainable development of rural B&Bs through the application of algorithms. The result exists a consistent recognition with the optimal row design of neighborhood houses proposed by Teng Y et al. [24]. Finally, the innovation of spatial planning and design, the study is able to solve the complex problems in spatial planning of rural lodging more effectively by utilizing the Bert-BiLSTM-EIC algorithm. For example, the algorithm helps identify the optimal spatial layout and spatial location distribution of the B&B, while taking into account the emotional needs of different users. This approach brings innovation to traditional spatial planning and design, making it more data-driven and result-oriented. The result also has similarities with the algorithmic application in landscape classification by Zhang D et al. [25].

However, despite the positive results of this study, there are some challenges and limitations. First, access to high quality and relevant data remains a challenge. Second, the algorithmic models need to be continuously adjusted and optimized to adapt to changing market and environmental conditions. Finally, applying these techniques to rural lodging in different regions may require targeted adaptation to specific local needs and conditions. For example, sentiment analysis may be more focused on community services and quality of life feedback in rural areas, while in urban areas it may be more focused on business services and the experience of city life. Regardless of the application scenario or geographic location, the main means of studying the proposed model is to obtain key information through public and customer sentiment analysis, and then classify and identify it through algorithmic modeling, as well as give key solutions for spatial planning and positioning. In terms of expansion, the multilingual utility and scalability of the model needs to be further enhanced.

Follow-up research could explore the application of this algorithmic framework to broader areas of rural revitalization, such as agricultural production, rural education and health services. Further research on the adaptability and effectiveness of the algorithm in different cultural and geographical contexts is also necessary.

VI. Conclusion

The current spatial location planning of rural homestays lacks systematicity and scientificity. Therefore, a sentiment analysis model integrating Bert BiLSTM EIC algorithm is proposed based on deep learning networks. This model is applied to the spatial planning of homestays near Chengdu Normal University and Chengdu accuracy, recall, and F1 values of the model can reach 94%, 93%, and 94%. The comprehensive score has increased by 6% compared to the traditional CNN model. In the loss function test, the model has the lowest loss rate, the best stability, and the fastest convergence speed among the four analytical models. In the statistics of the homestay housing resources in the two regions over the year, the demand is highest after the Spring Festival, June August, and November. In the consumer satisfaction scores before and after the renovation of the two places, the average score of homestays near Chengdu Normal University increases by 21% compared to before the renovation. The average score of homestays near Chengdu Neusoft College increases by 40% compared to before the renovation. In summary, the proposed sentiment analysis model that integrates the Bert BiLSTM EIC algorithm has high accuracy and feasibility in the spatial location planning of rural homestays, which can effectively improve the spatial layout optimization and service quality of rural homestays. However, there are many limiting factors in the actual planning of homestays, such as local policies, population size, etc. Therefore, this is also an area for further improvement in subsequent experiments. For example, a detailed analysis of local policies, such as Neusoft University. For the Bert-BiLSTM-EIC algorithm, the land use regulations, building codes, and tourism promotion policies, was included in the experiment. Simulate the impact of different policy changes on B&B spatial planning in order to evaluate planning options under different policy environments. Or integrate data such as population size, population density, and population movement into the model to better understand the needs and constraints of the target area. Use the model to simulate the impact of different population dynamics on B&B planning, such as population movement during peak and off-season, and long-term population trends.

References

[1] Che L, Zhou L, Xu J. Integrating the ecosystem service in sustainable plateau spatial planning: A case study of the Yarlung Zangbo River Basin. Journal of Geographical Sciences, 2021, 31(2):281-297.

[2] Sugandi Y B W, Paturusi S A, Wiranatha A S. Community Based Homestay Management in The Village Tourism of Tete Batu, Lombok. E-Journal of Tourism, 2020, 7(2): 369-383.

[3] Guoqing Z. Study on the influence of over development and construction of traditional villages on the development of B & B and its planning strategy: A case study of Jinlin Village in Zhaoqing, Guangdong Province. Journal of Landscape Research, 2021, 13(2): 36-46.

[4] Karur K, Sharma N, Dharmatti C, Siegel J E. A survey of path planning algorithms for mobile robots[J]. Vehicles, 2021, 3(3): 448-468.

[5] Rhanoui M, Mikram M, Yousfi S, Barzali S. A CNN-BiLSTM model for document-level sentiment analysis. Machine Learning and Knowledge Extraction, 2019, 1(3): 832-847.

[6] Singla P, Duhan M, Saroha S. An ensemble method to forecast 24-h ahead solar irradiance using wavelet decomposition and BiLSTM deep learning network. Earth Science Informatics, 2022, 15(1): 291-306.

[7] Liu B, Song C, Wang Q, Wang Y. Forecasting of China's solar PV industry installed capacity and analyzing of employment effect: based on GRA-BiLSTM model. Environmental Science and Pollution Research, 2022, 29(3): 4557-4573.

[8] Challa S K, Kumar A, Semwal V B. A multibranch CNN-BiLSTM model for human activity recognition using wearable sensor data. The Visual Computer, 2022, 38(12): 4095-4109.

[9] Gautier Q, Althoff A, Crutchfield C L, Kastner R. Sherlock: A multi-objective design space exploration framework. ACM Transactions on Design Automation of Electronic Systems (TODAES), 2022, 27(4): 1-20.

[10] Wu W, Fu X M, Tang R, Wang Y H, Qi Y H, Liu L G. Data-driven interior plan generation for residential buildings. ACM Transactions on Graphics (TOG), 2019, 38(6): 1-12.

[11] Roshani S, Koziel S, Yahya S I, et al. Mutual coupling reduction in antenna arrays using artificial intelligence approach and inverse neural network surrogates. Sensors, 2023, 23(16): 7089.

[12] Fang Y, Luo B, Zhao T, He D, Jiang B, Liu Q. ST-SIGMA:Spatio-temporal semantics and interaction graph aggregation for multi-agent perception and trajectory forecasting. CAAI Transactions on Intelligence Technology, 2022, 7(4):744-757.

[13] Barlow D R, Torres L G. Planning ahead: Dynamic models forecast blue whale distribution with applications for spatial management. Journal of Applied Ecology, 2021, 58(11):2493-2504.

[14] Abdurasulovich T J. Shaxs Yuz Tasvirini Klassifikatsiyalashda Bilstm Tarmog'idan Foydalanish. Models and Methods for Increasing the Efficiency of Innovative Research, 2023, 2(24): 66-68.

[15] Dileep P, Rao K N, Bodapati P, Gokuruboyina S, Peddi R, Grover A, Sheetal A. An automatic heart disease prediction using cluster-based bi-directional LSTM (C-BiLSTM) algorithm. Neural Computing and Applications, 2023, 35(10): 7253-7266.

[16] Conrow L, Mooney S, Wentz E A. The association between residential housing prices, bicycle infrastructure and ridership volumes. Urban Studies, 2021, 58(4): 787-808.

[17] Guo Y, Mustafaoglu Z, Koundal D. Spam Detection Using Bidirectional Transformers and Machine Learning Classifier Algorithms. Journal of Computational and Cognitive Engineering, 2023, 2(1): 5-9.

[18] Guo P, Yuan Y, Peng Y. Analysis of Slope Stability and Disaster Law under Heavy Rainfall. Geofluids, 2021, 2021(3):1-17.DOI:10.1155/2021/5520686.

[19] Parsamehr K, Gholamalifard M, Kooch Y. Impact of Land Cover Changes on Reducing Greenhouse Emissions: Site Selection, Baseline Modeling, and Strategic Environmental Assessment of REDD plus Projects. Land Degradation and Development, 2023, 34(10):2763-2779.

[20] Ray P, Chattaraj S, Bandyopadhyay S, Jena R, Singh S, Ray S. Shifting cultivation, soil degradation and agricultural land use planning in the North-eastern hill region of India using geo-spatial techniques. Land Degradation & Development, 2021,32(14):3870-3892.

[21] Almutairi K. Determining the appropriate location for renewable hydrogen development using multi-criteria decision-making approaches. International Journal of Energy Research, 2022, 46(5):5876-5895.

[22] Joao S, Linehan D. The Colonial Hotel: spacing violence at the Grande Hotel, Beira,Mozambique. Journal of planning literature, 2023, 38(1):118-119.

[23] Che L, Zhou L, Xu J. Integrating the ecosystem service in sustainable plateau spatial planning: A case study of the Yarlung Zangbo River Basin. Journal of Geographical Sciences, 2021, 31(2):281-297.

[24] Teng Y, Yang S, Huang Y. Research on space optimization of historic blocks on Jiangnan from the perspective of place construction. Applied Mathematics and Nonlinear Sciences, 2021, 6(1): 201-210.

[25] Zhang D, Leng J, Li X. Three-Stream and Double Attention-Based DenseNet-BiLSTM for Fine Land Cover Classification of Complex Mining Landscapes. Sustainability, 2022, 14(19): 12465.