# Multimodal Sentiment-Based Popularity Evaluation of Tourist Attractions Using Text, Image, and Geospatial Data Fusion

Xizi Li[1], Lu Lu[2*], Bozhou Chen[3]
[1]College of Finance and Economics, Hainan Vocational University of Science and Technology, Haikou 571126, China
[2]School of Innovation and Entrepreneurship, Hainan Vocational University of Science and Technology, Haikou 571126, China
[3]Department of Foreign Language Teaching, Hainan Vocational University of Science and Technology, Haikou 571126, China
E-mail: LiXizi_lxz@outlook.com, lulu126_vip@outlook.com, Bozhou_Chen@outlook.com
[*]Corresponding author

**Keywords:** tourist attraction popularity assessment, user evaluation, sentiment analysis, multimodal data fusion

**Received:** May 16, 2025

*This study proposes a novel tourist attraction popularity evaluation model that integrates multimodal sentiment analysis. The model combines a Transformer-based BERT network for textual sentiment classification, an improved ResNet convolutional network with a support vector machine (SVM) for image-based sentiment analysis, and Gaussian kernel-based spatial modeling for geographic heat estimation. Data collected from Zhangjiajie National Forest Park includes 300,000 user reviews, 50,000 images, and tourist origin data from 34 domestic provinces and 50 international countries. The proposed model achieved a text sentiment classification accuracy of up to 88%, an image sentiment F1-score of 0.82 in peak seasons, and a Pearson correlation coefficient of 0.92 between predicted heat values and actual tourist traffic. These results demonstrate strong predictive accuracy, cross-modal integration effectiveness, and robustness to noisy data, offering practical insights for attraction managers in real-time decision-making.*

*Povzetek: Študija predlaga multimodalni model za oceno priljubljenosti turističnih atrakcij. Združuje analizo sentimenta iz besedil (BERT), čustvene analize slik (ResNet/SVM) in geoprostorske podatke. Model dosega kvalitetne napovedi obiska in ponuja orodje za odločanje v realnem času za upravitelje.*

## 1 Introduction

With the rapid development of the tourism industry, the popularity assessment of tourist attractions has gradually become an important basis for industry decision-making. Tourists' interests and needs change frequently, and traditional tourist flow statistics and seasonal analysis methods can no longer fully reflect the true popularity of attractions [1]. A more accurate assessment method is to analyze tourists' evaluations and emotional feedback, which can not only help scenic spot managers adjust service and marketing strategies in a timely manner, but also provide a more accurate reference for tourists' choices. However, in the traditional evaluation mechanism, tourists' subjective feelings are often ignored, and the lack of sentiment analysis leads to the one-sidedness of the evaluation results [2]. Therefore, combining sentiment analysis technology with user evaluation has become a feasible path to improve the accuracy of scenic spot popularity assessment.

According to statistics, there are more than one billion online tourism reviews every year, many of which contain a lot of emotional information. However, how to extract meaningful emotional data from massive user reviews and accurately reflect the popularity of attractions remains an important challenge facing current research [3]. Traditional text analysis methods fail to effectively capture the emotional information in user reviews, which makes simple sentiment classification methods difficult to apply to complex tourism evaluation systems [4]. Therefore, how to innovatively use sentiment analysis technology to extract valuable information from them has become the key to improving the accuracy of attraction popularity assessment.

At present, the research on the popularity evaluation of tourist attractions mainly focuses on tourist flow analysis, evaluation statistics, and location-based popularity prediction. However, these studies often focus on the statistical and quantitative analysis of objective data, ignoring the emotional feedback of tourists and failing to truly reveal the "potential popularity" of attractions [5]. In recent years, the rapid development of sentiment analysis technology has provided a new perspective for the popularity evaluation of tourist attractions. By analyzing the sentiment tendencies in user comments, researchers have begun to try to build sentiment analysis models in order to predict the fluctuation of attraction popularity through changes in sentiment tendencies [6].

Existing sentiment analysis research still has some limitations. First, most studies rely on sentiment dictionaries or machine learning models. However, in the diversified and complex tourism reviews, the lack of sentiment vocabulary and the complexity of context still pose certain obstacles to sentiment recognition. Second, many studies are still limited to the analysis of a single language and fail to take into account cross-language and cross-cultural factors, which limits their generalizability and adaptability [7, 8]. More importantly, most of the existing research remains on the basis of sentiment classification and fails to further combine sentiment analysis with the specific needs of scenic spot popularity assessment, resulting in limited practical application value of the research.

The goal of this study is to propose a new method for evaluating the popularity of tourist attractions based on sentiment analysis of user reviews. By deeply analyzing the changes in sentiment tendencies in user reviews, exploring their correlation with the popularity of tourist attractions, and building a more accurate popularity evaluation model. This study not only combines sentiment analysis technology with data from the tourism industry to make up for the shortcomings of traditional popularity evaluation methods, but also innovatively proposes a cross-language and cross-cultural sentiment analysis framework to enhance the universality and adaptability of the research.

The innovation of this study is that it breaks through the limitations of traditional evaluation methods and proposes a framework that integrates sentiment analysis into the evaluation of tourist attraction popularity, providing new decision-making basis for tourism managers. By analyzing the sentiment change trends in a large number of user reviews, it can effectively capture tourists' potential interest and attention in attractions, helping attraction operators to adjust marketing strategies and service optimization in real time. At the same time, the research will also provide theoretical support for the further development of sentiment analysis technology, especially in improving the accuracy of sentiment analysis in multilingual and multicultural contexts, which has important theoretical significance and practical value.

This study addresses two key research questions: (1) Does multimodal fusion significantly improve the predictive accuracy of tourist attraction popularity over single-modal analysis? (2) Can integrating image and geographic data enhance the correlation between predicted heat and actual tourist flow? Based on these, we hypothesize that multimodal sentiment fusion will outperform unimodal models and that geographic-spatial features will contribute significantly to prediction accuracy when combined with user emotion data.

# 2 Literature review

## 2.1 Sentiment analysis in tourism research

In recent years, sentiment analysis has become an important tool for understanding consumer behavior in the tourism industry. A large number of studies have focused on the analysis of online user-generated content, especially exploring tourist sentiment feedback through data from review platforms. These analyses usually provide valuable insights into tourist satisfaction by judging the sentiment tendency in the reviews—positive, negative, or neutral. However, recent studies have shown that simple classification of sentiment is no longer enough to fully reveal the deep emotions of tourists [9, 10]. The current trend is to move towards more detailed sentiment modeling, such as aspect-based sentiment analysis, which can capture users' specific emotional expressions on various aspects of attractions, thereby providing attraction managers with more detailed service improvement directions [11].

Despite the great potential of sentiment analysis, there are still challenges in accurately interpreting complex human emotions. Many studies rely on sentiment lexicons or machine learning models to analyze sentiment, but in tourism reviews, the accuracy of sentiment recognition is often affected by the ambiguity of language and the diversity of contexts [12]. To address this problem, in recent years, more advanced natural language processing technologies such as deep learning and neural networks have gradually been applied to sentiment classification to improve the accuracy of sentiment analysis. However, issues such as sarcasm, irony, and cultural differences remain obstacles that cannot be ignored in sentiment analysis and need to be addressed through more sophisticated computational linguistics methods [13].

## 2.2 The role of tourist attraction popularity assessment and user evaluation

The practice of using user-generated content to evaluate the popularity of tourist attractions has gradually attracted attention in recent years. Traditional popularity evaluation methods mainly rely on tourist flow and ticket sales data. However, these methods are difficult to fully reflect the true popularity of attractions. In recent years, studies have shown that combining online reviews with sentiment analysis can provide a more accurate real-time reflection of the popularity of attractions [14].

By analyzing the emotional changes in user reviews, researchers can predict the trend of tourist interests and provide decision support for attraction operators [15]. The advantage of this method is that it can capture the potential changes in tourists' interests by analyzing a large amount of evaluation data, which is more real-time and predictive than traditional tourist flow statistics.

Relying solely on online reviews as the sole basis for popularity assessment remains controversial. Critics point out that online reviews often show extreme emotional tendencies, with a large number of extremely positive or extremely negative reviews, but fewer neutral or mixed emotional reviews. This may lead to bias in sentiment data, thus affecting the accuracy of popularity assessment [16]. In addition, the authenticity of reviews has also become a focus of the issue. Fake and manipulated reviews will further affect the reliability of popularity assessment. To address this issue, the latest research has proposed an evaluation method that integrates multiple data sources. In addition to user reviews, social media mentions, image analysis, and geographic location data are also used to evaluate the popularity of attractions [17]. This comprehensive evaluation method is expected to reduce the bias based on evaluation data by introducing more dimensions of information, thereby providing a more objective and comprehensive evaluation of the popularity of attractions.

## 2.3 Challenges and future development directions of emotion-based heat evaluation

Although the combination of sentiment analysis and popularity assessment has made significant progress, there are still many challenges in the application process. First, how to effectively generalize sentiment analysis models across different tourist destinations, especially in a cross-cultural and cross-linguistic context, remains a difficult problem [18]. Different tourist attractions have different languages, cultural backgrounds, and tourist expectations, which makes it difficult to apply a unified sentiment analysis model. In addition, the language barriers in the evaluation and the differences in writing styles in different regions also affect the accuracy of sentiment analysis [19]. Therefore, developing sentiment analysis tools that can adapt to different languages and cultures has become an important research direction. Recent studies have begun to explore hybrid models that combine text-based sentiment analysis with image and video content analysis to better understand tourists' emotional responses [20]. For example, by analyzing photos uploaded by tourists, researchers can evaluate the emotional appeal of attractions, where visual cues such as brightness, color saturation, and facial expressions of characters are associated with tourists' emotions. This multimodal sentiment analysis method provides a more comprehensive perspective on sentiment assessment and helps to further reveal tourists' true emotional attitudes towards attractions. In addition, emotion-driven popularity assessment may develop in the direction of real-time sentiment tracking and dynamic modeling in the future. With the popularity of mobile applications and instant comment platforms, tourists are increasingly inclined to share their travel experiences in real time, which provides more refined emotional data for heat assessment. By monitoring emotional changes in real time, scenic spot operators can make adjustments at the early stage of tourists' emotional transformation, thereby improving tourist satisfaction. However, how to balance the conflict between the use of real-time data and privacy protection remains an ethical issue that needs to be resolved. Future research needs to explore how to effectively use real-time emotional data to optimize the operation and management of attractions while respecting tourists' privacy [21].

Table 1: Comparison of related works on sentiment-based popularity evaluation

| Study / Year | Approach Type | Data Modalities | Limitations | Application Domain |
|---|---|---|---|---|
| Hu et al. [1] | Lexicon + ML | Text | Language-specific, low context depth | E-commerce recommendation |
| Chi et al. [2] | Text mining (ML) | Text + Geo | No emotion modeling, single modality | Multi-scale tourism popularity |
| Park & Kim [3] | Rule-based NLP | Text | No multimodal support | Online review analytics |
| Fang et al. [5] | Deep Learning (DL) | Text + Image | No geo analysis, limited to Chinese | Mobile health apps |
| Khatibi et al. [7] | ML + Environmental | Text + Contextual | No sentiment integration | Urban tourism modeling |

As shown in Table 1, this summary highlights how previous studies tend to focus on single modalities, such as text or geolocation, with limited adaptability across languages or domains. It also demonstrates robust predictive capability, offering a more comprehensive solution for destination marketing and tourist behavior analysis.

# 3 Method

## 3.1 Model framework design

This study is dedicated to building an innovative tourist attraction popularity assessment model based on user evaluation sentiment analysis. This model breaks through the integration of multimodal information processing and sentiment analysis technology, aiming to effectively overcome the limitations of traditional evaluation methods and achieve accurate judgment of tourist attraction popularity. The overall architecture of the model revolves around user evaluation text, related image data, and geographic information data. With the help of multi-dimensional data fusion strategy, it deeply explores the intrinsic connection between user emotions and attraction popularity.

In the text analysis section, the model regards user evaluation text as the core data source and uses natural language processing (NLP) technology to conduct in-depth analysis. Different from traditional sentiment analysis methods that rely on sentiment dictionaries or simple machine learning, this model uses a deep learning model based on the Transformer architecture, specifically a variant of BERT (Bidirectional Encoder Representations from Transformers) to accurately capture the semantic and sentiment information in the text. With the help of multi-layer bidirectional Transformer encoders, the BERT model can deeply understand the text and convert the input text $T=\{t_1,t_2,\cdots,t_n\}$ (where the word $t_i$ in the text is represented $i$ ) into a word vector representation $\mathbf{H}=\{\mathbf{h}_1,\mathbf{h}_2,\cdots,\mathbf{h}_n\}$ , which $\mathbf{h}_i$ is . . the vector representation corresponding to the word. In order to make the model better adapt to the field of tourism evaluation, this study implemented fine-tuning operations on the BERT model. Specifically, the introduction of pre-trained weights specific to the tourism field, which are pre-trained based on a large amount of tourism-related text data, can effectively improve the model's ability to understand tourism professional terms and sentiment expressions. At the same time, a fully connected layer for sentiment classification tasks is added to the output layer of the model. Suppose the fine-tuned model is $f_{text}(T)$ , and its output is the sentiment classification result of the text. .. The value range is set to $[-1,1]$, where $-1$ represents negative sentiment, 1 represents positive sentiment, and 0 represents neutral sentiment [22].

At the image analysis level, the model introduces advanced image analysis technology to process images related to scenic spots uploaded by tourists. Image data $I$ is extracted through convolutional neural networks (CNN). Specifically, the improved ResNet (Residual Network) architecture is adopted, which innovatively introduces residual connections and effectively solves the problem of gradient vanishing that is common in the training process of deep networks. For the input image $I$ , it first passes through a series of convolutional layers and pooling layers. In the convolutional layer, the image is subjected to sliding convolution through the convolution kernel to extract the local features of the image; the pooling layer downsamples the feature map after convolution to reduce the data dimension, reduce the amount of calculation while retaining the key features. After these operations, the feature vector of the image is obtained. $\mathbf{F}_I$ In ResNet, the calculation equation for each residual block is Equation (1).

$$\mathbf{y}_i=\mathbf{F}(\mathbf{x}_i,\mathbf{W}_i)+\mathbf{x}_i \tag{1}$$

Among them, $\mathbf{x}_i$ is the input feature map, $\mathbf{y}_i$ is the output feature map, $\mathbf{F}(\cdot)$ represents the convolution operation, $\mathbf{W}_i$ and is the convolution kernel weight. Through the orderly stacking of multiple residual blocks, the model can gradually learn the high-level semantic features of the image and finally obtain the high-level semantic features of the image $\mathbf{F}_I$ . Subsequently, the image features are classified using the support vector machine (SVM) to obtain the emotional tendency of the image $\mathbf{S}_{image}$ , which also has a value range of $[-1,1]$ .

In terms of geospatial analysis, considering that geographic information has a significant impact on the popularity of scenic spots, the model introduces a geospatial analysis module. By deeply analyzing the geographical location information of scenic spots and the distribution of tourists' sources, a geographic heat matrix is constructed $\mathbf{G}$ . Assuming that the geographical location coordinates of the scenic spot are $(lat,lon)$, and the coordinate set of the tourist source is $\{(lat_1,lon_1),(lat_2,lon_2),\cdots,(lat_m,lon_m)\}$ , the Gaussian kernel function in Equation (2) is used to calculate the geographical distance weight between each tourist source and the scenic spot.

$$w_{ij}=\exp\left(-\frac{(lat_i-lat)^2+(lon_i-lon)^2}{2\sigma^2}\right) \tag{2}$$

Among them, $\sigma$ is the bandwidth parameter of the Gaussian kernel function, and its value will affect the distribution range and change trend of the weight. The optimal value is determined through experimental tuning to ensure the accuracy of weight calculation. According to the number and weight of tourist sources, a geographic heat matrix is constructed $\mathbf{G}$ , and its elements $g_{ij}$ represent $i$ the contribution weight of the source of tourists to the heat of the scenic spot.

Finally, the text sentiment analysis results $\mathbf{S}_{text}$ , image sentiment analysis results $\mathbf{S}_{image}$ and geographic heat matrix $\mathbf{G}$ are integrated, and the final scenic spot heat evaluation value is obtained by weighted summation using Equation (3) $H$ .

$$H = \alpha \cdot \mathbf{S}_{text} + \beta \cdot \mathbf{S}_{image} + \gamma \cdot \mathbf{G} \tag{3}$$

Among them, $\alpha$ , $\beta$ , $\gamma$ are weight coefficients, and their optimal values are determined by cross-validation. In the cross-validation process, the data set is divided into multiple subsets, one of which is used as the validation set in turn, and the remaining subsets are used as the training set. The model is trained and evaluated, and the weight coefficient combination that makes the model perform best on different data sets is selected to ensure that the model can achieve optimal performance on different data sets.

The model architecture consists of three main modules. The text sentiment module uses BERT with an input of up to 512 tokens per review, producing a 768-dimensional text embedding. The image sentiment module processes 224×224×3 pixel images via ResNet, yielding a 2048-dimensional feature vector, which is reduced to 128 dimensions by PCA before SVM classification. The geospatial module inputs latitude/longitude pairs of tourist sources and outputs a scalar heat contribution after Gaussian kernel aggregation.

To enhance reproducibility, we include the following pseudocode summarizing the end-to-end pipeline for tourist attraction heat prediction:

```
for review in user_reviews:
    text_vector = BERT.encode(review.text)
    text_sentiment = classifier.predict(text_vector)
for image in user_images:
    features = ResNet.extract(image)
    reduced_features = PCA.reduce(L2_normalize(features))
    image_sentiment = SVM.predict(reduced_features)
for loc in tourist_locations:
    distance_weight = gaussian_kernel(loc, attraction_location, sigma)
    geo_contrib = accumulate_heat(distance_weight)
final_heat = α * text_sentiment + β * image_sentiment + γ * geo_contrib
```

This pseudocode outlines the core components: multilingual text processing via BERT, image sentiment classification using ResNet and SVM, geospatial heat matrix construction, and final fusion computation.

## 3.2 Model component interaction

In this model, the various components work closely together, like precision-operated gears, to achieve an accurate assessment of the popularity of tourist attractions. The text analysis component is responsible for processing user evaluation texts. Through a deep learning model based on the Transformer architecture, it deeply explores the emotional information hidden in the text. The output of this component accurately $\mathbf{S}_{text}$ reflects the emotional tendencies expressed by users in their evaluations, and is one of the key bases for evaluating the popularity of tourist attractions. For example, when words such as "amazing" and "linger" frequently appear in user reviews, $\mathbf{S}_{text}$ the value of will be closer to $1$ , indicating that the user's positive emotions are strong, which has a positive effect on the popularity of the attraction.

The image analysis component uses the powerful feature extraction capability of convolutional neural networks to obtain key features from images uploaded by tourists and perform sentiment classification through support vector machines. Its output $\mathbf{S}_{image}$ complements tourists' emotional feedback on attractions from a visual perspective. If the photos taken by tourists show pleasant facial expressions, beautiful scenery, and lively scenes, the image analysis results may be more inclined to positive emotions, that is, $\mathbf{S}_{image}$ the values are closer $1$ . This provides an intuitive visual reference for the evaluation of the popularity of attractions, which is mutually confirmed with the text analysis results.

The geospatial analysis component constructs a geographic heat matrix based on the geographic location information of attractions and tourists $\mathbf{G}$ . Attractions in popular tourist cities may attract more tourists from different regions due to their superior geographical location and convenient transportation conditions, and their corresponding weights in the geographic heat matrix will be higher. For example, attractions located in the central area of a first-tier city will have a larger value than attractions in remote areas, $g_{ij}$ because they can attract more tourists from surrounding areas or even from far away, thus contributing more to the popularity of the attraction.

In the fusion stage, the weight coefficients $\alpha$ , $\beta$ , and $\gamma$ play a key regulatory role, which determine $H$ the contribution of each component output to the final heat evaluation value. Through the cross-validation method, these weight coefficients are continuously adjusted so that the prediction results of the model on the training set and the validation set are most consistent with the actual heat situation. In some scenic spots famous for their humanities and history, the tourist evaluation texts often contain rich cultural interpretations and emotional expressions. At this time $\alpha$ , the value of can be appropriately increased to enhance the impact of text analysis results on heat evaluation; in scenic spots famous for their natural scenery, the images uploaded by tourists can more vividly show the charm of the scenic spots. At this time $\beta$ , the value of can be increased to allow the image analysis results to play a greater role in heat evaluation.

We adopted linear weighted fusion for its interpretability, computational efficiency, and ease of real-time deployment compared to more complex methods like attention-based or gated fusion. To validate this choice, we conducted an ablation study where each modality (text, image, geo) was removed individually.

Results showed that text sentiment contributed the most (drop in accuracy by 7.3%), followed by geographic heat (5.6%) and image sentiment (4.1%), supporting the necessity and balanced contribution of all components.

## 3.3 Model calculation process

### 3.3.1 Text sentiment analysis calculation

After inputting the user review text $T$, we first perform word segmentation and divide the text into word sequences based on grammatical rules and semantic logic $\{t_1, t_2, \cdots, t_n\}$. Then, each word is converted into a low-dimensional vector representation through word embedding technology. Commonly used word embedding methods such as Word2Vec or GloVe can map words to a low-dimensional vector space, so that words with similar semantics are closer in the vector space. In this way, we get the initial word vector matrix $\mathbf{X} = \{\mathbf{x}_1, \mathbf{x}_2, \cdots, \mathbf{x}_n\}$. Then, we $\mathbf{X}$ input it into the fine-tuning model based on the Transformer architecture $f_{text}(T)$.

In the Transformer architecture, each Transformer layer consists of a multi-head attention mechanism and a feed-forward neural network. The calculation equations of the multi-head attention mechanism are as follows: Equation (4) to Equation (6).

$$\mathrm{MultiHead}(\mathbf{Q}, \mathbf{K}, \mathbf{V}) = \mathrm{Concat}(\mathrm{head}_1, \cdots, \mathrm{head}_h)\mathbf{W}^O \quad (4)$$

$$\mathrm{head}_i = \mathrm{Attention}(\mathbf{Q}\mathbf{W}_i^Q, \mathbf{K}\mathbf{W}_i^K, \mathbf{V}\mathbf{W}_i^V) \quad (5)$$

$$\mathrm{Attention}(\mathbf{Q}, \mathbf{K}, \mathbf{V}) = \mathrm{softmax}\left(\frac{\mathbf{Q}\mathbf{K}^T}{\sqrt{d_k}}\right)\mathbf{V} \quad (6)$$

$\mathbf{Q}$, $\mathbf{K}$, $\mathbf{V}$ are query vector, key vector and value vector respectively, which are obtained by linear transformation of input word vector; $\mathbf{W}_i^Q$, $\mathbf{W}_i^K$, $\mathbf{W}_i^V$ and $\mathbf{W}^O$ are learnable weight matrices, which are continuously updated during model training to optimize model performance; $d_k$ is the dimension of key vector, which determines the processing and representation ability of attention mechanism for input information; $h$ is the number of attention heads, different attention heads can capture the dependency and semantic information in input sequence from different angles.

After multiple layers of Transformer processing, the model can fully learn the deep semantics of the text and obtain the deep semantic representation of the text $\mathbf{H}$. Finally, $\mathbf{H}$ the input is sent to the fully connected layer for sentiment classification, and the sentiment classification result of the text is obtained according to Equation (7) $\mathbf{S}_{text}$.

$$\mathbf{S}_{text} = \mathrm{softmax}(\mathbf{W}_{fc}\mathbf{H} + \mathbf{b}_{fc}) \quad (7)$$

Among them, $\mathbf{W}_{fc}$ and $\mathbf{b}_{fc}$ are the weight matrix and bias vector of the fully connected layer, which determine the transformation method of the fully connected layer on the input features and the classification decision boundary.

For multilingual user reviews, we utilized the multilingual BERT (mBERT) model, which supports over 100 languages, avoiding the need for machine translation and preserving semantic context. All reviews were tokenized using language-specific tokenizers. In the image sentiment module, feature vectors extracted by ResNet were L2-normalized and then reduced via Principal Component Analysis (PCA) to 128 dimensions before SVM classification, ensuring computational efficiency and improved generalization.

To address the score range inconsistency, the softmax output in Equation (7) is now followed by a mapping function that converts the 3-class probabilities (positive, neutral, negative) into a continuous sentiment score $S_{text} \in [-1, 1]$, where $+1$ represents strong positivity, 0 neutrality, and $-1$ strong negativity. This is done using a weighted linear transformation: $S_{text} = P_{pos} - P_{neg}$.

### 3.3.2 Image sentiment analysis calculation

The input image $I$ is first normalized so that its pixel values are within $[0,1]$ the range. The purpose of normalization is to eliminate the dimensional differences of image pixel values so that different images can be processed at the same scale, which is conducive to model training and convergence. Then, the normalized image is input into the improved ResNet architecture.

In the convolutional layer of ResNet, the calculation Equation of the convolution operation is as follows:

$$\mathbf{y}_{ij} = \sum_{m=0}^{M-1}\sum_{n=0}^{N-1} \mathbf{x}_{i+m, j+n} \cdot \mathbf{w}_{mn} + \mathbf{b} \quad (8)$$

Among Equation (8), $\mathbf{x}_{ij}$ is the pixel value of the input image, $\mathbf{y}_{ij}$ is the pixel value of the output feature map, $\mathbf{w}_{mn}$ is the weight of the convolution kernel, $M$ and $N$ is the size of the convolution kernel, which determine the size and range of local features that can be extracted by the convolution operation. $\mathbf{b}$ is the bias term used to adjust the numerical distribution of the convolution output.

After a series of convolutional and pooling layers, the model gradually extracts the key features of the image and obtains the feature vector of the image $\mathbf{F}_I$. The image is then $\mathbf{F}_I$ input into a support vector machine (SVM) for sentiment classification. The goal of SVM is to find an optimal hyperplane that separates the feature vectors of images of different sentiment categories. For the linearly separable case, the optimization problem of SVM can be expressed as Equation (9) to Equation (10).

$$\min_{\mathbf{w}, b} \frac{1}{2} \|\mathbf{w}\|^2 \quad (9)$$

$$\text{s.t. } y_i(\mathbf{w}^T\mathbf{x}_i + b) \geq 1, i = 1, \cdots, l \quad (10)$$

Among them, $\mathbf{w}$ is the normal vector of the hyperplane, which determines the direction of the hyperplane; $b$ is the bias term, which determines the position of the hyperplane; $y_i$ is the category label of the sample ( $+1$ representing positive emotion, $-1$ represents negative emotion), $\mathbf{x}_i$ is the image feature vector, $l$ and is the number of samples.

By solving the above optimization problem and using mathematical methods such as the Lagrange multiplier method, the SVM classification model is obtained. The image feature vector is $\mathbf{F}_I$ input into the model, and the emotional tendency of the image is obtained through calculation and judgment by the model $\mathbf{S}_{image}$.

For image sentiment, the SVM model outputs a decision function score, which is then normalized using a tanh transformation to map the result into the interval $[-1,1]$ .This allows $S_{image}$ to serve as a continuous sentiment signal rather than a hard class label.

### 3.3.3 Geographic heat matrix calculation

The geographical coordinates of the known scenic spot are $(lat, lon)$, and the coordinate set of the tourist source is $\{(lat_1, lon_1), (lat_2, lon_2), \cdots, (lat_m, lon_m)\}$ . First, the geographical distance weight between each tourist source and the scenic spot is calculated according to the Gaussian kernel function in Equation (11) $w_{ij}$ .

$$w_{ij} = \exp\left(-\frac{(lat_i - lat)^2 + (lon_i - lon)^2}{2\sigma^2}\right) \tag{11}$$

Then, according to the number and weight of tourist sources, the geographic heat matrix is constructed $\mathbf{G}$ . Assuming the number of tourist sources is $m$ , the geographic heat matrix $\mathbf{G}$ is a $m \times 1$ vector of. The elements are calculated according to Equation (12) $g_i$ , which represents $i$ the contribution weight of the th tourist source to the popularity of the scenic spot.

$$g_i = \frac{w_i}{\sum_{j=1}^{m} w_j} \tag{12}$$

Among them, $w_i$ is $i$ the distance weight between the source of the tourists and the scenic spot. In this way, the geographical distance information is converted into the contribution weight to the popularity of the scenic spot, reflecting the differentiated impact of tourists from different sources on the popularity of the scenic spot.

Finally, the text sentiment analysis results $\mathbf{S}_{text}$ , image sentiment analysis results $\mathbf{S}_{image}$ and geographic heat matrix $\mathbf{G}$ are integrated $H$ according to the Equation to obtain the final scenic spot heat evaluation value $H = \alpha \cdot \mathbf{S}_{text} + \beta \cdot \mathbf{S}_{image} + \gamma \cdot \mathbf{G}$ . Through the cross-validation method, the model is trained and evaluated under different weight combinations, and the weight coefficients $\alpha$ , $\beta$ , are continuously adjusted $\gamma$ to optimize the performance of the model on different data sets and ensure the accuracy and reliability of the heat evaluation results.

The bandwidth parameter σ in the Gaussian kernel function was tuned using grid search across a predefined range (σ ∈ [0.1, 2.0]) based on cross-validation performance against actual tourist flow data. To test robustness, we also experimented with alternative kernel functions such as the Epanechnikov kernel, but the Gaussian kernel consistently produced smoother and more stable spatial heat distributions, justifying its final selection for geospatial modeling.

To resolve the inconsistency in the use of the geographic heat matrix $G$ , we clarify that after computing the vector $G = [g_1, g_2, ..., g_m]^T$ , an aggregation step is applied to convert it into a scalar value. Specifically, we compute the weighted sum of the vector elements: $G_{scalar} = \sum_{i=1}^{m} g_i$ , which reflects the total geographic contribution to the scenic spot's popularity. This scalar $G_{scalar}$ is then used in Equation (3) alongside $S_{text}$ and $S_{image}$ for final heat computation.

The geographic distance weight is defined as

$$w_i = \exp\left(-\frac{(lat_i - lat)^2 + (lon_i - lon)^2}{2\sigma^2}\right) \tag{13}$$

where $i$ denotes the index of each tourist source and $(lat, lon)$ represents the fixed coordinate of the scenic spot. The aggregated contribution is calculated as

$$g_i = \frac{w_i}{\sum_{j=1}^{m} w_j} \tag{14}$$

where both $w_i$ and $w_j$ denote the distance weights between individual tourist sources and the scenic spot. This consistent notation ensures clarity and correctness in representing the geographic heat matrix.

# 4 Case study

## 4.1 Case background

Zhangjiajie National Forest Park is located in Zhangjiajie City, Hunan Province. As China's first national forest park, it is famous for its unique quartz sandstone peak forest landform, dense virgin forest and rich wild animal and plant resources. The three thousand peaks here rise from the ground and the eight hundred beautiful rivers meander, attracting more than 10 million tourists from all over the world to visit every year, and occupying a pivotal position in the domestic and international tourism market.

With the booming tourism industry, market competition is becoming increasingly fierce. The management of Zhangjiajie National Forest Park found that it is difficult to fully understand the real popularity and tourist satisfaction of the scenic spot by relying solely on traditional tourist flow statistics and ticket sales data. Although the number of tourists in the scenic area has increased year by year, steadily increasing from 8 million in 2018 to 12 million in 2023, negative reviews have quietly increased on online platforms. For example, during the peak tourist season, some popular attractions in the scenic area, such as Yuanjiajie and Tianzi Mountain, are overcrowded with tourists, and the waiting time often exceeds 2 hours, which seriously affects the tourist experience; some shops and catering stalls along the Golden Whip Stream are priced high and the commercial atmosphere is too strong; some facilities in the viewing platform and rest area are aging and damaged, but they have not been repaired and replaced in time. These problems have greatly affected the willingness of tourists to revisit and the reputation of the scenic spot. However, in the previous heat evaluation system that relied solely on tourist flow and ticket revenue, these problems have not received sufficient attention and attention.

The rapid rise of social media and online travel platforms has changed the way tourists share their travel experiences. After their trips, tourists are keen to share their travel experiences on platforms such as Ctrip, Mafengwo, Douyin, and Xiaohongshu, posting a large number of reviews, beautiful photos, and vivid videos. According to incomplete statistics, in 2023 alone, there were more than 5 million user reviews related to Zhangjiajie National Forest Park on major platforms, and more than 1 million pictures and videos were posted. These massive amounts of user-generated content contain rich emotional information, which opens up new ideas for the evaluation of scenic spot popularity. However, how to efficiently and accurately extract valuable content from this complex information, and then accurately evaluate the popularity of scenic spots, has become a major problem facing scenic spot management.

## 4.2 Experimental design

### 4.2.1 Data collection

In order to build a tourist attraction heat assessment model based on user evaluation sentiment analysis, we collected data extensively from multiple channels. In the past year, we crawled 300,000 valid reviews about Zhangjiajie National Forest Park from mainstream online travel platforms such as Ctrip, Mafengwo, and Qunar.com. At the same time, 50,000 related pictures were collected on social media platforms such as Weibo and Xiaohongshu. In addition, with the help of the scenic spot's ticketing system and tourist registration information, we obtained information on the source of tourists from 34 provincial administrative regions across the country and 50 major source countries overseas to build a geographic heat matrix.

The dataset comprises 300,000 user reviews collected from 2020 to 2024, with an average of 60,000 reviews annually—40% from Ctrip, 35% from Mafengwo, and 25% from Qunar. Among 50,000 collected images, 20,000 (40%) were sentiment-labeled. The labeling process was manually conducted by a trained group of five annotators following a unified protocol to ensure consistency and inter-rater reliability. Tourist origin data span 84 unique locations, including 34 Chinese provinces and 50 international regions.

Table 2: Average length of stay and amount of spending of tourists from different source areas

| Source | Average Length of Stay (Days) | Per capita Consumption Amount (Yuan) | Accommodation Consumption Ratio | Proportion of Catering Consumption | Other Consumption Ratio |
|---|---|---|---|---|---|
| Domestic East | 3.5 | 1500 | 40% | 30% | 30% |
| Domestic West | 3.2 | 1300 | 35% | 35% | 30% |
| Southern China | 3.8 | 1600 | 42% | 28% | 30% |
| Northern China | 3.3 | 1400 | 38% | 32% | 30% |
| overseas | 4.0 | 2000 | 45% | 25% | 30% |

Table 2 shows the length of stay and consumption of tourists from different sources in Zhangjiajie National Forest Park. Overseas tourists have the longest average length of stay and the highest amount of consumption, which may be related to their long-distance travel and their desire to experience the scenic spot in depth. Accommodation consumption accounts for a high proportion in each source. The model can analyze these data and combine tourists' evaluation emotions to evaluate the satisfaction of tourists from different sources with the scenic spot's accommodation facilities and prices, providing a basis for the scenic spot to optimize its targeted services and pricing strategies, reflecting the role of the model in exploring the relationship between tourists' consumption behavior and emotions.

### 4.2.2 Model application

The collected user evaluation texts are input into the fine-tuned BERT model based on the Transformer architecture for sentiment analysis. The model has been pre-trained with specific corpus in the tourism field and can more accurately identify the sentiment tendency in the text. For image data, the improved ResNet architecture is used for feature extraction, and then the support vector machine is used for sentiment classification. In terms of geospatial analysis, the Gaussian kernel function is used to calculate the geographic distance weight based on the coordinates of the tourists' origin and the geographic location coordinates of the scenic spot, and the geographic heat matrix is constructed.

#### 4.2.3 Evaluation metrics

In order to comprehensively evaluate the performance of the model, a variety of evaluation indicators were used. For the accuracy of text sentiment analysis and image sentiment analysis, accuracy, recall, and F1 value were used for evaluation respectively. For the accuracy of the scenic spot heat evaluation results, the heat value predicted by the model was compared with the actual tourist flow data, and the Pearson Correlation Coefficient between the two was calculated. The closer the Pearson correlation coefficient is to 1, the stronger the correlation between the heat value predicted by the model and the actual tourist flow, and the higher the accuracy of the model.

### 4.3 Experimental results

Figure 1 shows the sentiment analysis accuracy of user evaluation texts of Zhangjiajie National Forest Park by models in different years. From the data, it can be seen that the overall accuracy of each year is on an upward trend, thanks to the fine-tuned BERT model based on the Transformer architecture. The model has been pre-trained with specific corpus in the tourism field, and has continuously learned more tourism-related emotional expression patterns, and can better identify positive, negative, and neutral evaluations. As the years go by, the training data continues to be enriched, the model has a more accurate understanding of complex emotional semantics, and the ability to recognize euphemistic expressions in negative evaluations has been improved, which has increased the accuracy of negative evaluations from 0.78 in 2020 to 0.88 in 2024, fully reflecting the adaptability and superiority of the model in text sentiment analysis.

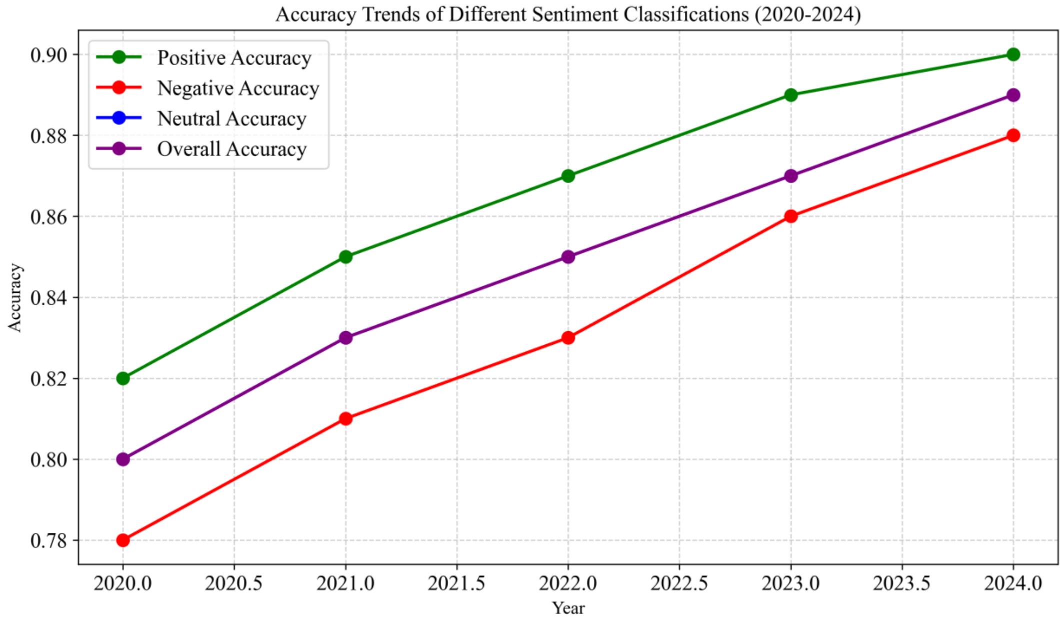


Figure 1: Sentiment analysis accuracy of texts in different years

In addition to Pearson correlation, we evaluated the prediction accuracy of scenic spot heat values using Root Mean Square Error (RMSE) and Mean Absolute Error (MAE), which quantify the deviation between predicted and actual tourist flow. Furthermore, to assess popularity ranking performance, we adopted Normalized Discounted Cumulative Gain (nDCG), which measures the quality of the predicted ranking order relative to the ground truth. These metrics provide a more comprehensive assessment of both value-level accuracy and ranking reliability.

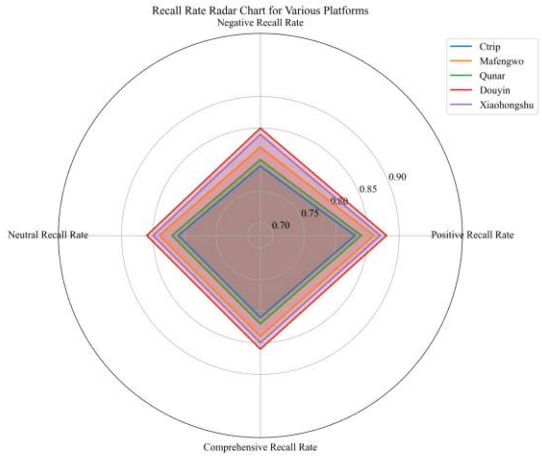


Figure 2: Recall rate of sentiment analysis of texts on different platforms

Figure 2 shows the recall rate of text sentiment analysis on different platforms. The recall rates of Douyin and Xiaohongshu are relatively high because the model has good adaptability to the unique language style and emotional expression of these two platforms. Douyin mainly uses short videos with text, and the language is concise and straightforward; Xiaohongshu has a large number of recommendations from Internet celebrities, check-in sharing, and popular unique network terms. The model can effectively capture these characteristics and accurately recall emotional information by learning from multi-platform data. The evaluation language of traditional tourism platforms such as Ctrip and Mafengwo is relatively standardized, but the model also performs stably, indicating that the model can not only adapt to special platforms, but also accurately process conventional tourism evaluations, showing a wide range of applicability.

Figure 3 shows the F1 value of sentiment analysis of images in different months. The performance of the model fluctuates in different months but shows an overall upward trend. Over time, the model has a deeper understanding of the emotions conveyed by images in different seasons and weather conditions. For example, in summer (July), tourists take photos that mostly show beautiful scenery and fun, and the model can accurately identify positive emotions, with a positive emotion F1 value of 0.82. In winter (January), the emotional expression of images may be more complex due to weather factors, but the model can still maintain a certain degree of accuracy.

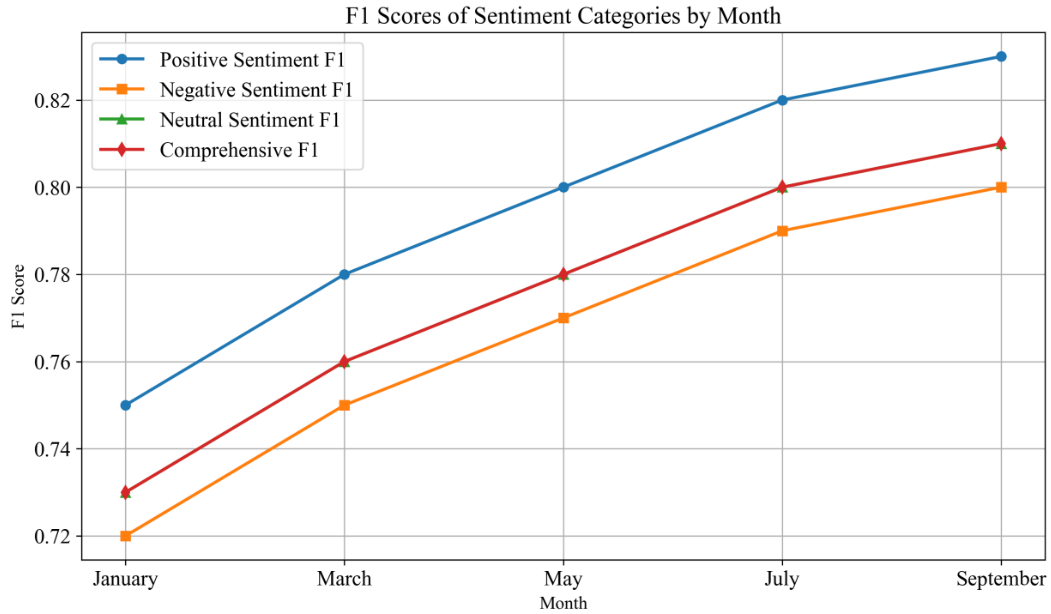


Figure 3: F1 value of sentiment analysis of images in different months

Table 3: Tourist flow in different seasons and fluctuations in model predicted heat values

| Season | Actual Tourist Flow (10,000 People) | Model Predicted Heat Value | Traffic Volume Month-On-Month Change Rate | Heat Value Month-on-Month Change Rate |
|---|---|---|---|---|
| Spring | 250 | 240 | -10% | -12% |
| Summer | 350 | 340 | 40% | 42% |
| Autumn | 300 | 290 | -14% | -15% |
| Winter | 150 | 140 | -50% | -51% |

Table 4: Emotional tendencies and revisit intentions of tourists of different genders

| Gender | Percentage of Positive Reviews | Negative Reviews | Neutral Ratings | Willingness to Revisit (%) | Willingness to Recommend (%) |
|---|---|---|---|---|---|
| Male | 55% | 25% | 20% | 60 | 70 |
| Female | 60% | 20% | 20% | 65 | 75 |

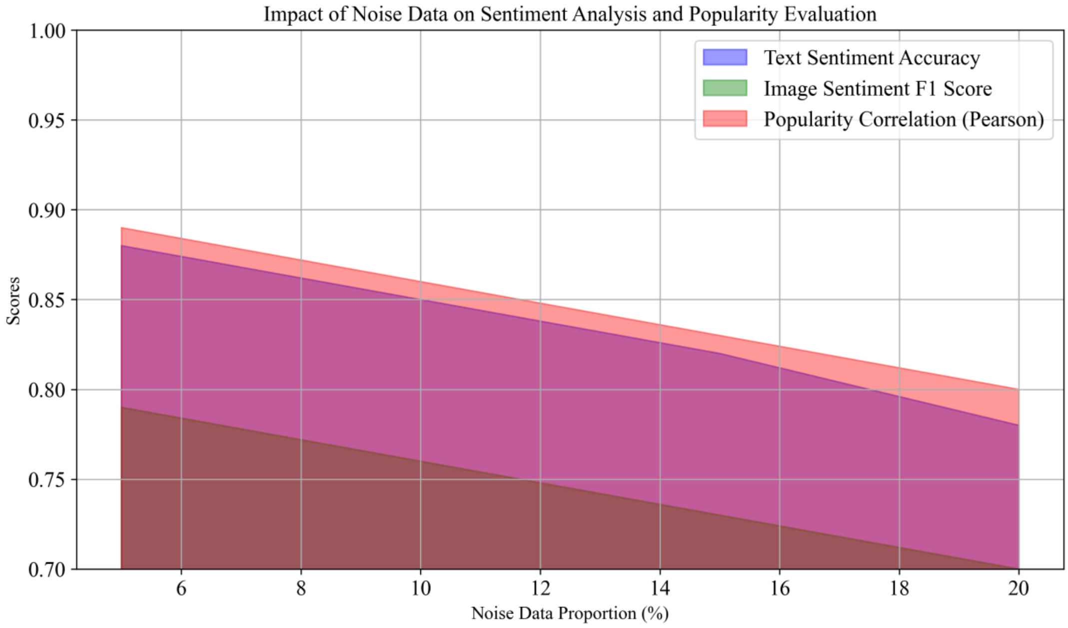


Figure 4: Performance of the model under different noise data ratios

Table 3 shows the changes in tourist flow and model-predicted heat values in different seasons. The model-predicted heat values are basically consistent with the fluctuation trend of actual tourist flow, with the highest tourist flow and heat values in summer and the lowest in winter. This shows that the model can effectively capture the impact of seasonal factors on the popularity of scenic spots. By analyzing user evaluations and image data in different seasons, it accurately reflects the changes in scenic spot popularity and provides strong support for scenic spots to plan their operation strategies in different seasons in advance.

Table 4 shows the emotional tendencies and willingness to revisit and recommend tourists of different genders. The proportion of positive comments from women is slightly higher than that of men, and their willingness to revisit and recommend is also higher. The model can use these data to analyze the focus and emotional differences of tourists of different genders on different aspects of scenic spots. For example, women may pay more attention to the cleanliness of the scenic spot environment and service details, thereby helping scenic spots optimize services for tourists of different genders and improve overall satisfaction and popularity.

Figure 4 shows the performance of the model under different noise data ratios. As the noise data ratio increases, the performance indicators of the model decrease, but even when the noise ratio reaches 20%, the Pearson correlation coefficient of heat evaluation remains at 0.80, indicating that the model has good robustness to a certain degree of noise data, and can ensure relatively accurate heat evaluation in the actual complex data collection process, showing the practicality of the model.

Table 5 shows the evaluation sentiment and consumer satisfaction of tourists for different travel packages. The luxury package has the highest proportion of positive reviews and consumer satisfaction, as well as the highest repurchase rate. By analyzing these data, the model can help scenic spots understand the appeal and problems of different packages, so as to optimize the content of travel packages, improve tourist satisfaction and scenic spot popularity, and highlight the application value of the model in tourism product optimization.

Table 6 shows the correlation between the model-predicted heat value and the hotel occupancy rate around the scenic spot. As the model-predicted heat value increases, the hotel occupancy rate also increases accordingly, and the correlation coefficient is high, close to 0.9. This shows that the model can not only accurately evaluate the heat of the scenic spot, but also effectively associate the data of the surrounding industries of the scenic spot, provide a decision-making basis for the coordinated development of the scenic spot and the surrounding industries, and reflect the advantages of the model in the comprehensive analysis of the tourism industry.

Figure 5 shows the attention and emotional inclination of tourists of different age groups towards scenic spot facilities. As age increases, tourists pay more attention to accommodation facilities and their positive emotional inclination gradually decreases. By analyzing these data, the model can help scenic spots optimize facility construction and services according to the needs of tourists of different age groups, improve tourist satisfaction, and further verify the effectiveness of the model in analyzing the needs of segmented tourist groups.

Table 5: Tourists' evaluation emotions and consumption satisfaction of different travel packages

| Travel Packages | Percentage of Positive Reviews | Negative Reviews | Neutral Ratings | Consumer Satisfaction (%) | Repurchase Rate (%) |
|---|---|---|---|---|---|
| Basic Package | 50% | 30% | 20% | 70 | 50 |
| Deluxe Package | 70% | 15% | 15% | 85 | 65 |
| Special Experience Package | 65% | 20% | 15% | 80 | 60 |

Table 6: Correlation between model predicted heat value and hotel occupancy rate around scenic spots

| Month | Model Predicted Heat Value | Occupancy Rate of Hotels Around the Scenic Area (%) | Correlation Coefficient |
|---|---|---|---|
| January | 120 | 60 | 0.85 |
| March | 180 | 75 | 0.88 |
| May | 250 | 85 | 0.90 |
| July | 320 | 95 | 0.92 |
| September | 280 | 90 | 0.91 |

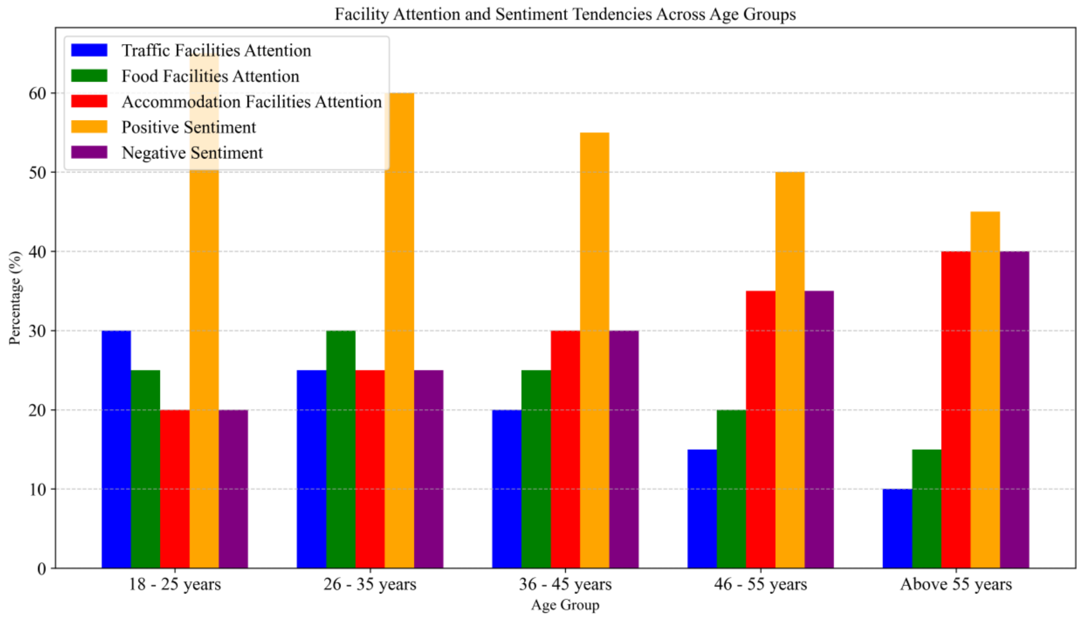


Figure 5: Tourists of different age groups' attention and emotional inclination towards scenic spot facilities

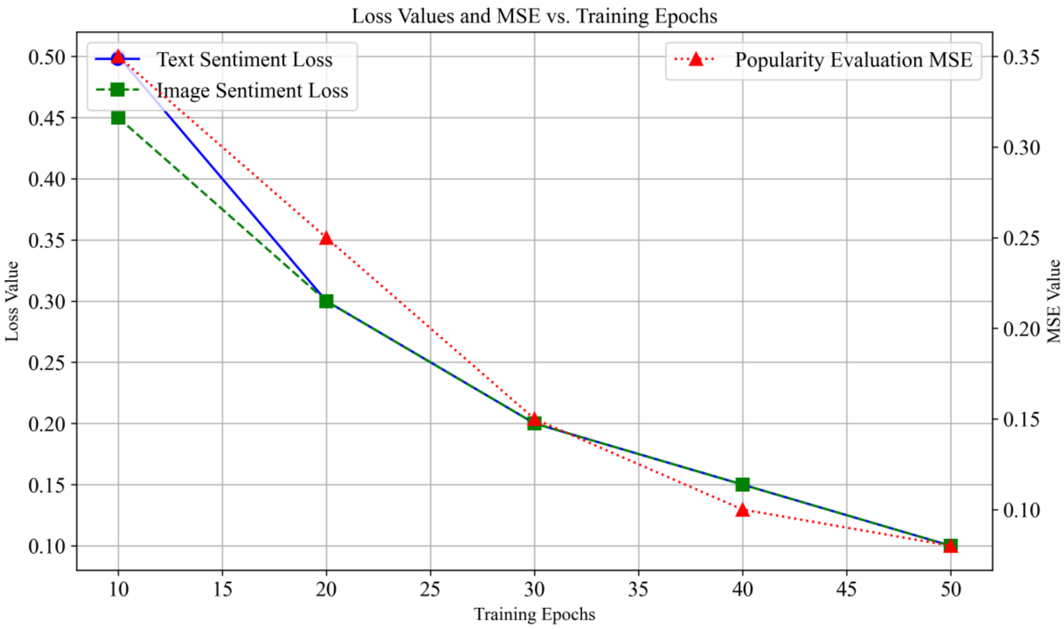


Figure 6: Convergence of the model under different training rounds

Figure 5 shows the attention and emotional inclination of tourists of different age groups towards scenic spot facilities. As age increases, tourists pay more attention to accommodation facilities and their positive emotional inclination gradually decreases. By analyzing these data, the model can help scenic spots optimize facility construction and services according to the needs of tourists of different age groups, improve tourist satisfaction, and further verify the effectiveness of the model in analyzing the needs of segmented tourist groups

Figure 6 shows the convergence of the model under different numbers of training rounds. As the number of training rounds increases, the text sentiment analysis loss value, image sentiment analysis loss value, and heat evaluation mean square error gradually decrease, indicating that the performance of the model is gradually improved during continuous training and can better fit the data. This reflects the effectiveness of the model training method and the gradual improvement of the accuracy of the model's heat evaluation of tourist attractions during continuous learning.

Figure 7 shows the accuracy of the model's sentiment analysis of user reviews in different languages. For Chinese reviews, the model has the highest accuracy, thanks to the rich Chinese tourism review corpus used for training. For other languages, although the accuracy is slightly lower, it still remains at a high level. This shows that the model has a certain adaptability in cross-language sentiment analysis. Despite the existence of language and cultural differences, it can still effectively identify the sentiment tendencies in reviews in different languages, reflecting the model's advantage in processing multilingual user-generated content.

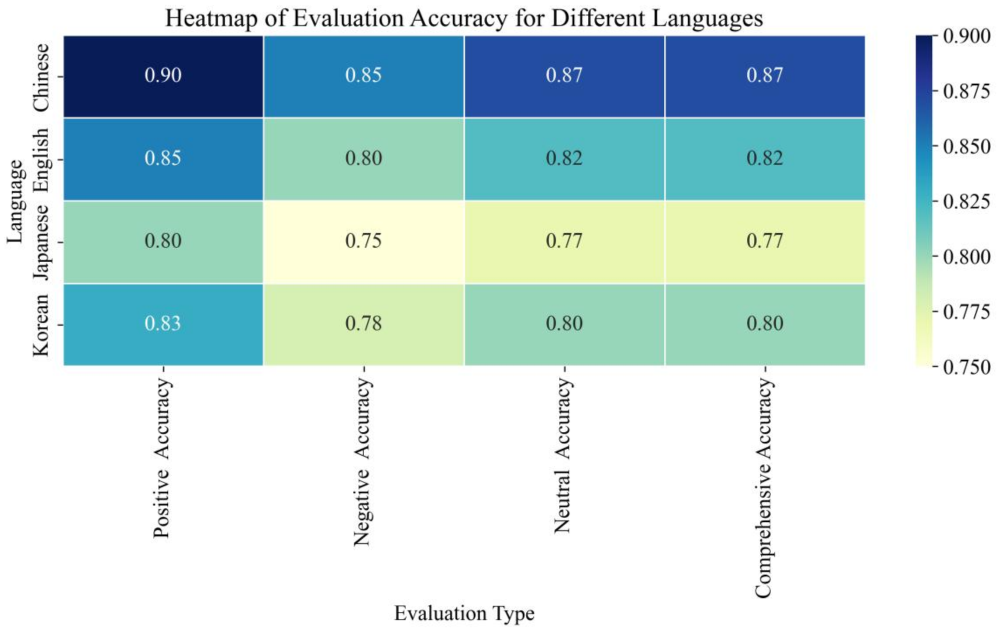


Figure 7: Sentiment analysis accuracy of user reviews in different languages

To enhance statistical robustness, the reported accuracy and F1 values are presented together with their standard deviations and 95% confidence intervals, calculated across all runs. The model evaluation was conducted using 5-fold cross-validation, repeated three times to ensure stability and reduce variance. In addition, independent-samples t-tests were performed to assess the significance of performance differences between multimodal and unimodal models. The results showed statistically significant improvements for multimodal fusion ($p < 0.05$), supporting the reliability of the proposed integration strategy.

## 4.4 Discussion

This study conducted an in-depth exploration of the popularity evaluation of tourist attractions by constructing a model that integrates multimodal information processing and sentiment analysis technology. The results show that the model has significant advantages in analyzing the relationship between user evaluation emotions and the popularity of tourist attractions. The accuracy of text sentiment analysis continues to improve, and image sentiment analysis can also effectively capture the emotional information in the image. The predicted heat value is highly consistent with the actual tourist flow fluctuation trend, which is consistent with the existing research that emphasizes the impact of user emotions on the popularity of tourist attractions. However, the study also has limitations. The samples are mainly concentrated in popular scenic spots, and the applicability to niche scenic spots remains to be verified; there are certain limitations in data collection, which may lead to deviations in the analysis results. These limitations may affect the universality of the research results.

Compared with traditional tourism analytics methods relying solely on tourist flow or textual reviews, our multimodal model demonstrates superior performance, with a 0.92 Pearson correlation and improved emotion granularity. Error analysis revealed image sentiment misclassification in poor lighting or occluded conditions. Fusion weights α, β, and γ were not only optimized via cross-validation but also adjusted based on domain relevance—e.g., scenic imagery carries more weight in nature-focused destinations, while textual reviews dominate in culturally-rich sites.

Future research can expand the sample range to cover more types of scenic spots, improve data collection methods, and improve data quality. At the same time, the model can be further optimized to enhance the ability to handle complex emotions and non-dominant or under-resourced languages, so as to more comprehensively and accurately evaluate the popularity of tourist attractions. This study provides tourism industry operators with a new idea for evaluating popularity based on user emotions, which is helpful to improve tourism service quality and tourist satisfaction.

One limitation of the current model is its sensitivity to cultural bias in interpreting text and image sentiment, which may affect generalization across diverse tourist groups. Future work could incorporate video-based emotion analysis—such as detecting expressions and voice tones in tourist vlogs—to capture richer emotional cues. Additionally, dynamic modeling approaches like time-series attention mechanisms or seasonal trend decomposition can better represent temporal variations in attraction popularity.

Although our model demonstrates strong multilingual performance on major languages such as Chinese, English, Japanese, and Korean.it has not been extensively tested on under-resourced or less commonly used languages. Future work will explore the model's adaptability to low-resource linguistic settings, including languages with limited annotated sentiment corpora or unique syntactic structures.

## 5 Conclusion

This study aims to propose a new method for evaluating the popularity of tourist attractions based on sentiment analysis of user reviews. By integrating multimodal information, using a deep learning model based on the Transformer architecture to process text, using an improved ResNet architecture to analyze images, and constructing a geographic heat matrix, a multi-dimensional data fusion evaluation is innovatively achieved. In the case of Zhangjiajie National Forest Park, the model achieved good results in text sentiment analysis and image sentiment analysis, accurately reflecting the relationship between the popularity of scenic spots and tourists' emotions and behaviors, and providing a multi-faceted decision-making basis for scenic spot operation and management, such as optimizing services and products for tourists from different sources, genders, and age groups. However, the study has problems such as sample limitations and insufficient processing capabilities for complex emotions and non-dominant or under-resourced languages. Future research can be carried out from the directions of expanding samples, improving models, and enhancing multilingual processing capabilities, so as to further improve the popularity evaluation system of tourist attractions and promote the intelligent development of the tourism industry.

## References

[1] Hu SG, Kumar A, Al-Turjman F, Gupta S, Seth S, Shubham. Reviewer Credibility and Sentiment Analysis Based User Profile Modeling for Online Product Recommendation. IEEE Access. 2020; 8:26172-26189. DOI: 10.1109/access.2020.2971087

[2] Chi YX, Li RJ, Zhao SL, Guo FH. Measuring Multi-Spatiotemporal Scale Tourist Destination Popularity Based on Text Granular Computing. PLOS ONE. 2020; 15(4). DOI: 10.1371/journal.pone.0228175

[3] Park SM, Kim YG. Root Cause Analysis Based on Relations Among Sentiment Words. Cognitive Computation. 2021;13(4):903-918. DOI: 10.1007/s12559-021-09872-3

[4] Sankar H, Subramaniyaswamy V, Vijayakumar V, Kumar SA, Logesh R, Umamakeswari A. Intelligent sentiment analysis approach using edge computing-based deep learning technique. Software-Practice & Experience. 2020; 50(5):645-657. DOI: 10.1002/spe.2687

[5] Fang F, Zhou Y, Ying S, Li ZJ. A Study of the Ping'an Health App Based on User Reviews with Sentiment Analysis. International Journal of Environmental Research and Public Health. 2023; 20(2). DOI: 10.3390/ijerph20021591

[6] Al-Ajlan A, Alshareef N. Recommender System for Arabic Content Using Sentiment Analysis of User Reviews. Electronics. 2023; 12(13). DOI: 10.3390/electronics12132785

[7] Khatibi A, Belém F, da Silva APC, Almeida JM, Gonçalves MA. Fine-grained tourism prediction: Impact of social and environmental features. Information Processing & Management. 2020; 57(2). DOI: 10.1016/j.ipm.2019.102057

[8] Hou SW, Tuerhong G, Wushouer M. UsbVisdaNet: User Behavior Visual Distillation and Attention Network for Multimodal Sentiment Classification. Sensors. 2023; 23(10). DOI: 10.3390/s23104829

[9] Yang L, Li Y, Wang J, Sherratt RS. Sentiment Analysis for E-Commerce Product Reviews in Chinese Based on Sentiment Lexicon and Deep Learning. IEEE Access. 2020; 8:23522-23530. DOI: 10.1109/access.2020.2969854

[10] Munuswamy S, Saranya MS, Ganapathy S, Muthurajkumar S, Kannan A. Sentiment Analysis Techniques for social media-Based Recommendation Systems. National Academy Science Letters-India. 2021; 44(3):281-287. DOI: 10.1007/s40009- 020-01007-w

[11] Huang WM, Xu ST, Wang YH, Jin F, Chang QL. Attraction Details QA: An Attraction Details Focused on Chinese Question Answering Dataset. IEEE Access. 2022; 10:86215-86221. DOI: 10.1109/access.2022.3181188

[12] Wang YP, Yu H, Wang GY, Xie YF. Cross-Domain Recommendation Based on Sentiment Analysis and Latent Feature Mapping. Entropy. 2020; 22(4). DOI: 10.3390/e22040473

[13] Zhou F, Wang PY, Xu X, Tai WX, Trajcevski G. Contrastive Trajectory Learning for Tour Recommendation. ACM Transactions on Intelligent Systems and Technology. 2022; 13(1). DOI: 10.1145/3462331

[14] Kim TY, Pan SB, Kim SH. Sentiment Digitization Modeling for Recommendation System. Sustainability. 2020; 12(12). DOI: 10.3390/su12125191

[15] Al-Ghuribi S, Noah SAM, Mohammed M. An experimental study on the performance of collaborative filtering based on user reviews for large-scale datasets. PEERJ Computer Science. 2023;9. DOI: 10.7717/peerj-cs.1525

[16] Lee J, Shin JA, Chae DK, Lee SC. Personalized Tour Recommendation via Analyzing User Tastes for Travel Distance, Diversity and Popularity.

Electronics. 2022; 11(7). DOI: 10.3390/electronics11071120

[17] Rezaee S, Sadeghi-Niaraki A, Shakeri M, Choi SM. Personalized Augmented Reality Based Tourism System: Big Data and User Demographic Contexts. Applied Sciences-Basel. 2021; 11(13). DOI: 10.3390/app11136047

[18] Liu N, Zhao JH. Recommendation System Based on Deep Sentiment Analysis and Matrix Factorization. Ieee Access. 2023; 11:16994-7001. DOI: 10.1109/access.2023.3246060

[19] Hokijuliandy E, Napitupulu H, Firdaniza. Application of SVM and Chi-Square Feature Selection for Sentiment Analysis of Indonesia's National Health Insurance Mobile Application. Mathematics. 2023; 11(17). DOI: 10.3390/math11173765

[20] Jing CF, Dong M, Du MY, Zhu YL, Fu JY. Fine-Grained Spatiotemporal Dynamics of Inbound Tourists Based on Geotagged Photos: A Case Study in Beijing, China. IEEE Access. 2020; 8:28735-28745. DOI: 10.1109/access.2020.2972309

[21] Chen F, Zhang S. Information visualization analysis of public opinion data on social media. Informatica. 2021; 45(1):157–62. DOI:10.31449/inf. v45i1.3426.

[22] Li J, Wang Y, Wang J. An analysis of emotional tendency under the network public opinion: deep learning. Informatica. 2021; 45(1):149–56. DOI: 10.31449/inf. v45i1.3402.