applied sciences

MDPI

Article

Exploring Coastal Tourism Experience Through Social Media Text Mining: Sentiment and Thematic Patterns

Yu Wang [1], Zhiyu Zhang [1] and Zhijun Zhang [2,*]

[1] Department of Urban and Rural Planning, Tianjin University, Tianjin 300072, China; yu_wang2019@tju.edu.cn (Y.W.)
[2] Tianjin Geomatics Research Center, Tianjin 300201, China
* Correspondence: zhangzhijun@whu.edu.cn

Abstract

Research on coastal recreational activities has grown substantially, yet studies focusing on user perceptions of these spaces—critical for optimizing tourism experiences and management—remain fragmented and underdeveloped. This study addresses this gap by examining tourist sentiment in Xiamen, a renowned coastal city in China, using social media data. Text mining tools were utilized to process the Weibo contents through text segmentation, frequency analysis and cluster analysis. The Two-way Neural Network Fusion Model Based on the BERT (TNNFMB) deep learning approach was employed using transfer learning for sentiment analysis, while the Latent Dirichlet Allocation (LDA) model was used to uncover latent thematic patterns. Sentiment polarity analysis revealed that positive comments constituted 56.47%, negative comments only 16.3%, and neutral comments 27.2%, confirming a generally positive perception of visitors' coastal experiences. Tourists' social media posts primarily revolve around five core themes in coastal areas: coastal waters, waterfronts, adjacent environments, culture and creativity, and reputation and expectation. The spatial and temporal changes in sentiment scores were discovered. Areas emphasizing sea–land landscapes, cultural theme reinforcement, and open public activities generally achieved high and stable sentiment scores. Natural and natural–artificial mixed coastlines experienced significant seasonal variations in sentiment. The recommendations of this study, generated from a sentiment perspective, include shaping a harmonious coastal environment by improving coastal management and support services to enhance the comfort of the tourist experience. This study advances understanding of user-centric coastal tourism dynamics, providing evidence-based tools for managers to enhance tourist experiences and spatial quality.

Keywords: coastal tourism; recreation experience; social media text mining; sentiment score; spatiotemporal variations; Xiamen city

check for updates

Academic Editor: Grigorios Beligiannis

Received: 29 July 2025
Revised: 22 August 2025
Accepted: 26 August 2025
Published: 3 November 2025

Citation: Wang, Y.; Zhang, Z.; Zhang, Z. Exploring Coastal Tourism Experience Through Social Media Text Mining: Sentiment and Thematic Patterns. *Appl. Sci.* **2025**, *15*, 11721. https://doi.org/10.3390/app152111721

Copyright: © 2025 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/).

1. Introduction

Amidst the rapid advancement of urbanization, coastal cities, characterized by their unique geographical locations and ecological resources, have demonstrated a significant capacity for factor agglomeration [1]. In contrast to inland cities, coastal areas function as the primary interface for city–sea interactions, integrating urban public functions such as leisure, culture, and socialization, while also serving as corridors for marine activities and ecological perception [2]. Recent studies have identified numerous challenges for coastal cities [3–5]. The utilization of coastlines is limited by natural conditions and disaster

prevention requirements, while land resources are constrained by economic, social, and ownership factors, among others. Heavy reliance on motorized transport systems often results in neglecting walking and cycling routes, leading to poor connectivity between scenic areas and leisure activities, which affects the accessibility and experiences of visitors and residents [6,7]. Ecological space fragmentation, landscape homogenization, environmental stress from poor beach cleanliness, and limited public accessibility are contradictions impeding coastal sentiment attachment and spatial identity [8,9]. Examining coastal experiences from the viewpoint of tourists presents valuable opportunities to explore the links between human activities, marine ecosystems, and built environments.

1.1. The Evaluation of Coastal Spaces from the Perspective of Tourist Experience

Coastal spaces serve as crucial public areas where individuals participate in leisure and social activities. Parasuraman et al. (1988) suggested that perceived value arises from users evaluating and contrasting perceived advantages and disadvantages [10]. Sweeny and Soutar (2001) noted that users have subjective views on the overall impact of product or service features and other elements [11]. According to the waterfront environmental theory by Hanayanagi (1999), coastal tourism relates to the interaction of human activities and waterfront environment, which extends to exploring human awareness and behavior at the waterfront, changes in behavior characteristics over time, and waterfront space value [12]. Consequently, examining coastal spaces through the lens of user perception lays the groundwork for enhancing the current coastal environment, significantly contributing to the creation of more enjoyable and comfortable areas for users. Therefore, user perception and assessment of coastal spaces are key factors for evaluating the quality of these areas.

Conventional tourist experience research predominantly depends on comprehensive interviews, satisfaction evaluations, and questionnaire surveys, supplemented by statistical analysis. For instance, Roca et al. examined user perceptions in a coastal area of Spain using a questionnaire survey and categorized opinion groups [13]. Nonetheless, traditional methods and spatial evaluation techniques are often constrained by quantitative surveys and static data, and their scope of data coverage and capacity to quantify users' sentiment reactions to spatial elements frequently fall short [14]. The low response rate of questionnaire surveys may result in a concentration of samples among "highly loyal users", while on-site observations might not accurately reflect real usage scenarios due to limited case coverage [15]. The recent advancements in natural language processing, combined with the vast wealth of online data resources, offer a novel method for evaluating tourist experiences in coastal areas through data generation and text mining analytics. By using text data mining techniques, especially integrating new technologies such as semantic mining and clustering, sentiment analysis, and spatial analysis, judgment can be conducted on tourist experiences and emotional changes across various spaces and seasons. This allows for a more precise evaluation of the strengths and weaknesses of the coastal spatial environment and offers insights for optimizing coastal zone management.

1.2. The Application of Improved Text Mining Approaches for Understanding Tourism Experiences

In recent years, the significance of online text data in researching images of tourist destinations has considerably increased. With the rapid expansion of the e-economy, tourists' online reviews and opinions have become crucial intelligence sources for both businesses and consumers as they offer a more accurate and genuine reflection of users' true experiences [16]. Taking tourism as an example, tourism industry development and destination choices are increasingly influenced by the voice from social media [17]. The variations in consumer positioning and consumption culture highlighted by social media data can also help in the development of urban commercial areas [18]. Since geotagged

social media data includes time and location information, tourist movement patterns can be further drawn from related spatiotemporal analysis.

Text mining new technology serves as a focused approach for examining the vast amounts of data gathered from online reviews, especially the employment of natural language processing (NLP) to uncover hidden insights within unstructured or semi-structured text data [19]. This involves converting non-quantitative materials into quantitative data and deriving precise meanings through word and sentence inferences. For instance, ROST Content Mining 6 (ROST-CM6) is a widely used content analysis software for the Chinese language that extracts valuable information through methods such as word frequency statistics, social and semantic network analysis, and sentiment analysis [20]. Additionally, network text analysis techniques like Word2vec word vector training, K-means clustering, Term Frequency-Inverse Document Frequency (TF-IDF) keyword extraction, and text sentiment analysis are also in vogue [21]. However, these typical early text representation vector models are often simplistic, lacking contextual information and word-to-word associations. Recently, topic modeling has become a widely utilized tool for document representation. Related modeling methods have evolved from Latent Semantic Analysis (LSA) to Probabilistic Latent Semantic Indexing (PLS), and more recently to Latent Dirichlet Allocation (LDA). In tourism studies, the LDA model has proven powerful in identifying tourism categories. For example, nine tourism categories were identified in the analysis of tourist destinations and preferences in South Korea [22].

By far, text mining still encounters several challenges when handling unstructured data in coastal assessments, such as comments and policies. Studies suggest that the frequent use of text in environmental impact assessments may result in inaccurate risk communication, and text mining has yet to fully address this issue [23]. Additionally, existing research often examines user emotional or environmental data in isolation, lacking a dynamic correlation between the two [24]. For instance, coastal research quantified the supply and demand of beach entertainment services using POIs and social media data, but it has not integrated user preferences with environmental characteristics [25].

Zhang et al. (2011) thoroughly examined the contact relation and document relation in microblogs to aid topic mining by using the model of Micro Blog-Latent Dirichlet Allocation (MB-LDA) [26]. Jiang et al. (2017) gathered travelogs, comments, and other social media data as initial text, utilizing complex network theory for text mining [27]. Qiu et al. (2021) used ROST-CM6 to mine tourist comment data from online commentary texts sourced from Weibo [28]. These studies enable travel recommendation services by extracting keywords from tourists' online texts that reflect their emotional preferences on tourism destinations. With recent technological advancements, increasingly sophisticated tools are being integrated into this research field. Gu et al. (2019) applied a Word2vec-based word vector model for semantic mining of text big data, clustering keywords to effectively extract image perception elements, aiding in the analysis of urban tourism image structure and perception [29]. Wan et al. (2023) employed a deep learning model, TNNFMB (Two-way Neural Network Fusion Model Based on BERT) [30], which leverages advanced language knowledge from large-scale text data through transfer learning to address the data sparsity issue in small-sample sentiment analysis tasks. Their sentiment analyses have uncovered both positive and negative reviews from online users. Additionally, their studies have demonstrated the superior classification accuracy of the deep learning model, which achieves a significantly better classification effect in text mining.

Overall, the number of studies utilizing text mining technology to evaluate the needs and perceptions of users in specific locations remains limited, with most existing research still in the preliminary exploration phase. Compared to traditional methods, such as questionnaires, interviews, and field observations, employing text mining to systematically

investigate the needs and perceptions of place users presents a technical challenge. The processing flow of comment data is overly simplistic; a more professional and in-depth approach can better capture a user's perceived experience [31]. Researchers often directly apply general text analysis models, such as topic models, sentiment analysis, and word frequency statistics [32–34], without thoroughly considering how these technologies can effectively integrate with the core theoretical methods of place user needs and spatial perception. In summary, the research integrating user needs and perceptions based on text mining approach with the evaluation of coastal places is relatively nascent, lacking in-depth designs that adequately meet research needs [35,36].

Against the backdrop of growing attention to coastal recreational spaces in urban quality enhancement, understanding how tourists perceive and utilize these spaces has become critical for evidence-based management and experience optimization. Drawing on social media data from coastal tourists, rich in real-time, user-generated insights into on-site experiences, this study focuses on the spatial classification and temporal dynamics of coastal tourism activities. This paper aims to use text mining techniques to identify tourist experiences related to coastal areas, classify themes, and calculate sentiment values, from which the sentiment dynamics, thematic patterns, and spatial–temporal variations of coastal tourism can be captured. Ultimately, these efforts aim to provide actionable insights for refining coastal management strategies, tailoring space optimization measures to specific coastline types and enhancing the overall quality of coastal tourist experiences.

2. Materials and Methods

2.1. Case Study City

Xiamen is a pivotal city on China's southeast coast, serving as an international transportation hub and significant gateway for external engagement. Its distinctive island-land–sea spatial structure, coupled with its densely packed coastal recreational resources, diverse coastal geography, and concentrated tourism functions, makes it an ideal subject for examining coastal recreational activities, spatial demands, and their interactions. The climate is characterized by a southern subtropical marine monsoon type, with an average annual temperature of 21 °C (annual range < 15 °C), virtually no climatological winter, and an average annual precipitation of 1200 mm.

2.1.1. Selection of the Research Area

The coastal spaces discussed in this study were derived from satellite image data of Xiamen Island over various years and were selected using high-definition satellite maps from Google Earth. Images and coastline data were chosen from 1976, which marked the early stages of reclamation, to 2018, when reclamation projects had largely ceased. The results are shown in Figure 1. Over the years, the sinuosity of Xiamen Island's coastline has steadily decreased, highlighting the significant straightening effect of recent reclamation projects on the shoreline morphology. However, this development is spatially uneven: the northwest coast accounts for 83.7% of the island's reclaimed area, forming a large artificially modified shoreline, while the southern coast has remained virtually undeveloped since 1976, preserving its original natural state. This stark north–south contrast makes the southern coast a key research focus, retaining the characteristic bedrock-beach composite landscape unique to southeastern Fujian Province and, due to minimal human disturbance over nearly 40 years, provides a valuable natural reference sample for investigating natural coastline evolution. Consequently, this study selected the area from Baicheng to the White Stone Fort on the southern coast of Xiamen Island as the study area.

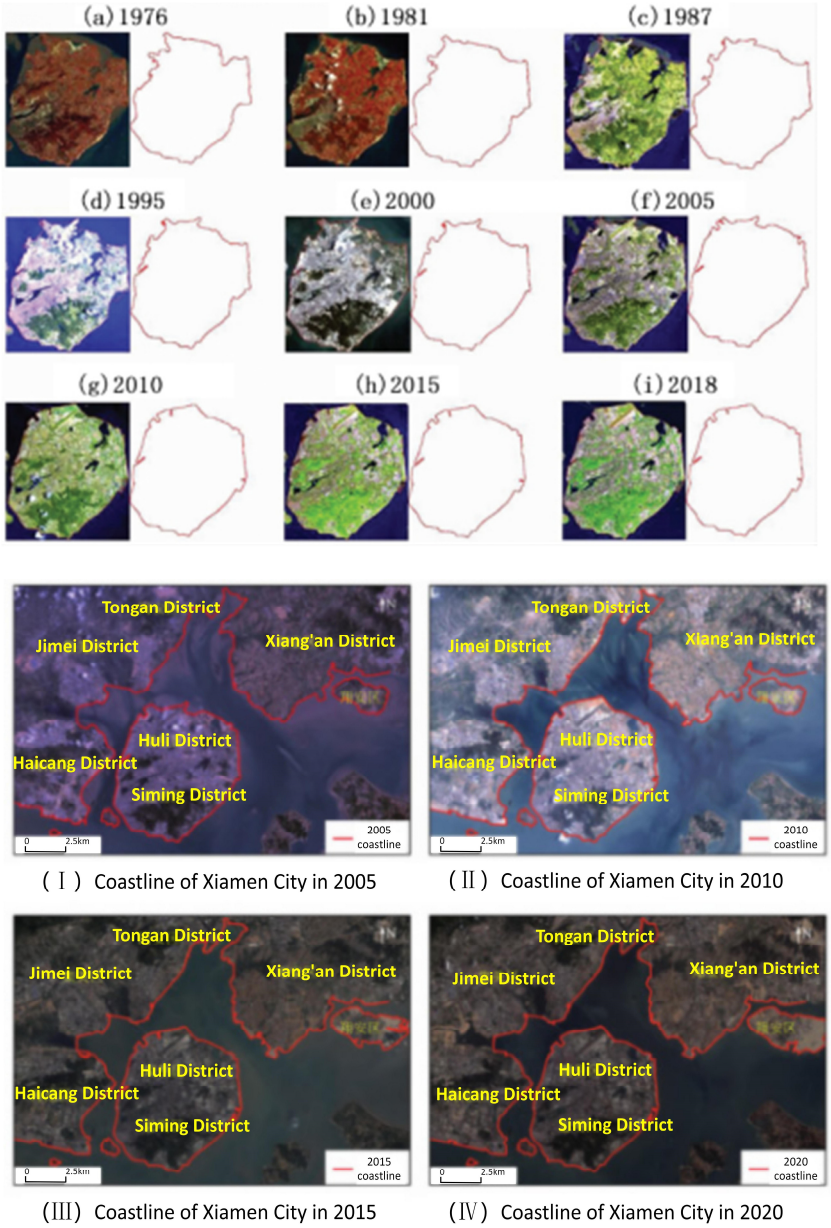


Figure 1. Diagram illustrating Xiamen City's coastline changes.

2.1.2. Spatial Features of the Southern Coastline of Xiamen

This study examines the central section of the coastal recreation area in southern Xiamen. Geographically, this region faces the Taiwan Strait to the south and is connected to the higher education district of Xiamen University and Zengcuoan Historical District to the north. As a high-density tourism hotspot, it functions as both a transportation hub and ecologically sensitive zone. Coastal spaces can be classified into three typical structural types based on their form and function (Table 1).

(1) Natural: Most are natural coastline dominated by ecological function. Represented by Baicheng-Huli Mountain, Pearl Bay, and Zengcuoan Beach, these areas form sandy intertidal zones as land–sea transition areas with a slope of $\leq 5°$, offering accessible and interactive recreational waterfronts.

(2) Hybrid: Many are vertical cliff walkways with mixed-use functions combing natural and built environment. For example, the section from Baicheng-Huli Mountain to White Stone Fort features vertical cliffs (with a relative height difference of 8–15 m)

developed along the bedrock coast, complemented by elevated wooden walkway systems for ecologically sensitive transit.

(3) Artificial: Many are enclosed bays or plazas with clear function and specific urban design. Typical examples include the Calligraphy Square and Music Square, where embankments create semi-enclosed bays or landscape plazas that serve as venues for cultural performances and gatherings. Through morphological diversity (natural proportion 62.3% + artificial proportion 37.7%) and functional complementarity, these three types of spaces together form a graded recreation spectrum that meets a range of differentiated needs, including sightseeing, sports, and social interaction.

Table 1. Spatial Characteristics of the Southern Coastal Sections of Xiamen Island.

Coastline Type	Spatial Characteristics	Typical Areas	Proportion of Landscape Elements	Core Recreational Functions
Natural Coastline	Gentle sandy coastline (slope $\leq 5°$), gradual natural transition between land and sea (intertidal zone width: 40–80 m)	Pearl Bay, Zengcuoan	Natural elements: 95% (beach: 85%, vegetation belt: 10%); artificial facilities: 5%	Water-related activities
Mixed Coastline	Vertical bedrock cliffs (height difference: 8–15 m), abrupt boundary transition between land and sea (land area directly adjacent to sea cliffs)	Baicheng-Huli Mountain, White Stone Fort	Natural elements: 90% (rock reefs: 65%, native shrubs: 25%); artificial facilities: 10%	Ecological sightseeing, geological popularization
Artificial Coastline	Semi-enclosed bays or squares (curvature radius > 100 m), artificial interface cutting transition between land and sea (hard revetment proportion > 70%)	Calligraphy Square, Music Square	Natural elements: 45% (lawn: 30%, artificial water body: 15%); artificial facilities: 55%	Cultural experience, festival gatherings

The coastal spaces of southern Xiamen City were divided into natural, artificial, and mixed categories. The coastal space in the southern part of Xiamen Island can be divided into several sections (Figure 2). They are Baicheng-Huli Mountain, Pearl Bay, Calligraphy Square, Zengcuoan, Music Square, and White Stone Fort, distributed from west to east. Each section features a variety of forms with rich and diverse expressions: Baicheng-Huli Mountain, Pearl Bay and Zengcuoan are centered around beaches; Calligraphy Square and Music Square are mainly plazas; The White Stone Fort is primarily a historical heritage. The sections are interspersed and connected by the inner ring road and outer coastal boardwalk. With short distances, indistinct boundaries, and close relationships, different areas can be reached quickly by walking.

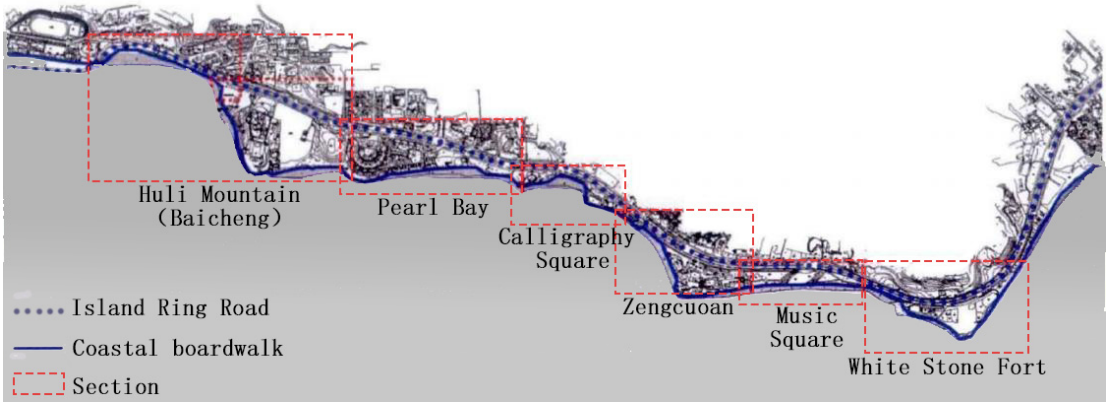


Figure 2. The divided coastal areas of Xiamen city in this study.

We reviewed the spatial characteristics of coastlines and functionally listed coastline types such as port wharves, coastal parks, coastal squares, tourist attractions, and beach bathing areas. Representative coastal frontages were analyzed by incorporating morphological parameters such as coastline curvature and hinterland depth.

2.2. Data Collection and Preprocessing

This study employed the location-based service dynamic API from the Chinese social media platform Weibo to gather online posts pertinent to the specified area in Xiamen. The time period was set to be between 1 January and 31 December 2024 based on the Weibo user check-in date. To collect user comment data for the target area, relevant keywords were searched on Weibo to extract comments from the designated regions. The collected data encompassed various types of text information: user ID, gender, check-in latitude and longitude, posting time, registration location, and text content, resulting in 5780 text entries. All original online posts were in Chinese. To improve the accuracy of topic extraction, our study also excluded some irrelevant text. This includes instances where users repeatedly extract comment data due to quoting other tourists, as well as marketing accounts that replicate a large volume of low-quality comment data and unrelated topics for profit. The data clean process was only upon the duplicate, irrelevant or wrong location data (referencing other users' original text, marketing accounts with advertisement, and some attached text that did not belong to the defined locations). The process did not delete any irrelevant text regarding any themes. This data-cleaning process led to a final valid data sample of 3124 entries.

2.3. Text Mining Procedure

We adopted an integrated approach to improve text mining accuracy through deep learning, constructing two core modules, namely, the TNNFMB model and LDA topic model, and adopting a progressive analytical framework (Figure 3). Utilizing the Weibo API and request library, large-scale extraction of Weibo posts was conducted from the search of related locations. Searches encompassed both core regions and sensitive suburban areas, ensuring spatial and temporal representativeness of the dataset.

At the semantic segmentation level, our investigation was conducted on the high-dimensional sparse characteristics of Weibo texts. Based on text mining, the ROSTCM6 word segmentation system was used as the basic word library. Through manual observation and systematic collection, a semantic network repository containing event-related professional terms and popular internet slang words was constructed. The repository was imported into the ROSTCM6 Chinese text analysis tool to optimize the segmentation performance. Our analysis was conducted within a Chinese language environment, and the results were presented using an English translation.

The LDA model determines the number of topics extracted, while the content of these topics is evaluated and derived based on the theory of human activity—waterfront environment interaction, in conjunction with the distribution of frequently occurring cloud words. The first focus was on introducing LDA topic modeling as a core analysis method for clustering and extracting topics from online posts. By generating a "comment topic word" probability matrix, deep semantic mining of the text set can be achieved. Thus, topic sets can be derived from high-frequency word clusters, key indicators can be screened, and core research topics can be identified. Compared to high-frequency word statistics, keyword extraction can more effectively express the semantic core of text, and by eliminating irrelevant text interference during the preprocessing phase, it allows for a deeper investigation into the cognitive structure of tourists.

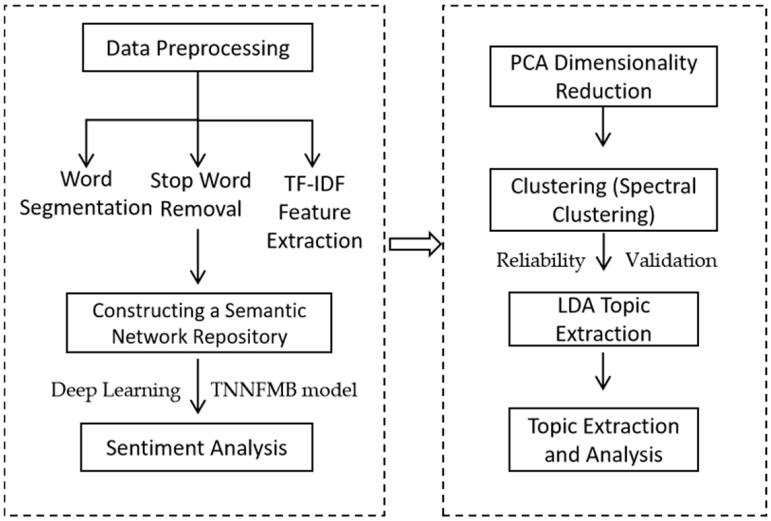


Figure 3. Text Mining Technical Roadmap.

At the sentiment analysis level, building on semantic network relation diagrams and using the ROSTCM6 sentiment analysis system, the TNNFMB was proposed through transfer learning combined with existing large-scale text data to analyze the sentiments of relevant Weibo posts in the study area. This was based on a 'dual pathway fusion' design (where one pathway focuses on surface emotional word features and the other pathway explores deeper semantic associations), the multi-dimensional features are fused through neural network layers to enhance sentiment classification accuracy.

When the ROSTCM6 software (v.5.8.0.603) conducts sentiment analysis, it calculates the weights of sentiment words within each clause of the text and then sums these weights to derive the sentiment score of the text. This yielded the distribution of tourists' sentiment tendencies toward coastal spaces, providing effective strategies for investigating the features and issues reported by tourists in these areas.

Finally, cluster analysis and multivariate regression models revealed differential patterns in tourists' sentiments across multiple dimensions. Using spatial statistics and kernel density estimation approaches on the ArcGIS platform, the spatial and temporal changes in tourists' sentiment was visually presented, offering rational coastal management recommendations.

2.3.1. Text Segmentation and Removal of Stop Words

Text segmentation is the process of splitting sentences in documents into independent word units according to specific rules and is a crucial step in building semantic network relationships. Chinese text segmentation was performed using the Jieba segmentation toolkit in Python (PC—243.26053.29), yielding a segmented dataset. Some words or characters, such as modal particles and special symbols, offer little actual semantic value in subsequent analyses, and have weak relevance to the main topic. These stop words were removed.

2.3.2. Textual Semantic Analysis

The tourists' overall impressions of coastal recreational spaces revealed distinct preference patterns. The term "not bad" emerged as the most frequently used term, representing the largest semantic share. Following closely, the words "relaxing" and "fun" highlight tourists' primary concerns about the navigation system and the recreational experience offered by the space. Contextual semantic analysis indicates that tourists also vividly recall significant performances within the space, such as entertainment events, as well as the quality of the environment, including sanitation facilities and landscape resources, such as coastal scenery. Notably, the issue of "crowded" conditions, stemming from the space's car-

rying capacity, is frequently mentioned, underscoring tourists' concerns regarding queuing times and congestion. In summary, tourists' main interests are concentrated on four core dimensions: the experiential quality of recreational programs, the uniqueness of landscape resources, environmental comfort, and value for money.

2.3.3. Destination Word Frequency Analysis

Destination features were analyzed by extracting keywords from tourist comments, revealing the core attributes of destinations and areas of interest. Unlike simple word frequency statistics, keyword extraction techniques effectively capture the semantic essence of text. By filtering out irrelevant data during preprocessing, these techniques delve deeply into the cognitive structures of tourists. This approach establishes a "keyword–destination characteristic" mapping, enabling a systematic analysis of tourists' perceived impressions and providing quantitative evidence for identifying destination traits. To mitigate cognitive bias from extreme ratings, a stratified sampling strategy was employed, which was created based on comprehensive evaluation scores. The representative destinations closest to the average in each layer were selected, and the top four high-frequency characteristic words from each segment were extracted for the micro-level analysis. Simultaneously, a comprehensive aggregated corpus was constructed for macro-level analysis, forming a dual-dimensional research framework of "micro-segmented attribute analysis–macro-level recognition", which ensures the representativeness of the sample while achieving multi-scale cognitive mining.

2.4. Thematic Mining and Modeling Methods

2.4.1. Thematic Mining and Clustering

The optimal number of clustering topics was determined using the clustering evaluation metric, that is, the silhouette coefficient (S(i)). In spectral clustering, if the value of K is extremely small, the data may not be effectively distinguished, thereby affecting the clustering outcome. Conversely, if K is too large, the silhouette coefficient tends to decrease, indicating poor clustering. To assess the performance of the chosen spectral clustering method, it was compared with two other common clustering algorithms: K-means and hierarchical clustering. As all three are unsupervised learning algorithms, the silhouette coefficient was used as a unified evaluation standard to measure clustering quality. As shown in Figure 4, the optimal silhouette coefficient under this method is S = 0.521, with the number of topics K to be 5.

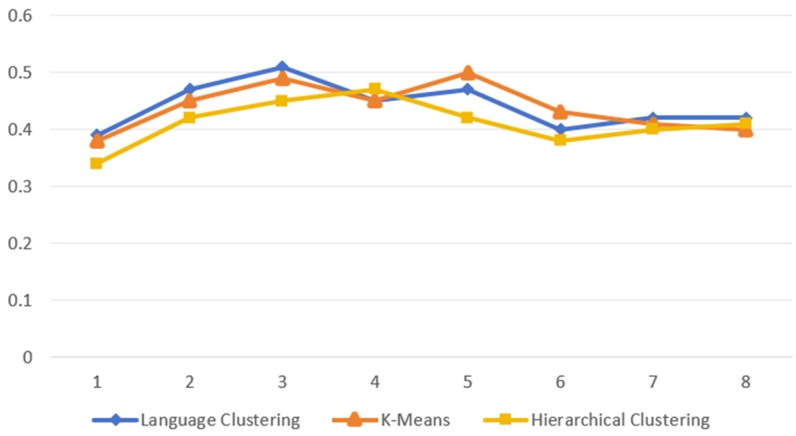


Figure 4. Comparison Results of Clustering Algorithms.

2.4.2. Reliability Examination and Validation

We employed two spectral clustering algorithms to determine the number of topics. The probabilistic latent semantic analysis (PLSA) topic model was selected as a comparing baseline with the LDA topic model. PLSA is a classic topic modeling method in the fields of text mining and topic identification. The clustering performance was assessed using two quantitative metrics. The F-score, which is the harmonic mean of Precision and Recall, reflects the balance of the clustering results. The equation is as follows:

$$\text{F-score} = \frac{\left(1+\beta^2\right)\text{Precision} \times \text{Recall}}{\beta^2\text{Precision} + \text{Recall}} \quad (1)$$

Here, β is the harmonic coefficient of the F-score, which is generally set as β = 1. In the cluster validity evaluation, the overall F-score was obtained by calculating the weighted average of the F-scores between classes, using the following equation:

$$\text{F-scare} = \sum_i \frac{n_i}{n} \max_j \left(\frac{2\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}}\right) \quad (2)$$

The purity of clustering results (purity) is a simple and intuitive cluster evaluation metric. This was calculated as follows:

$$Purity = \sum_{i=1}^{k} \frac{m_i}{m} p_i \quad (3)$$

$$p_i = \max\left(p_{ij}\right), p_{ij} = \frac{m_{ij}}{m_i} \quad (4)$$

where m_i is the number of all members in cluster i, and m_{ij} is the number of members belonging to both cluster i and category j.

Purity assesses the proportion of the dominant class within each cluster and then averages these proportions across all clusters, with lower values indicating better performance. The F-score, which represents the harmonic mean of the precision and recall for each topic, is more favorable when it is higher. The clustering effects of the two models are shown in Figure 5. The results indicate that the LDA model demonstrates lower error and higher value, underscoring its robustness.

After employing the clustering model to extract five key topics from valid Weibo tourism posts, high-frequency keywords for each topic were identified. The results are displayed in Table 2. Table 3 lists some typical tourist comments that represent these topics. Drawing from these findings, the five key topics of tourism posts were summarized by objective and subjective dimensions. Coastal waters, waterfronts, and adjacent environments pertain to physical perception, while culture and creativity and culture and creativity relate to non-physical perception, as detailed in Table 4.

Specific thematic items were further refined. For instance, in coastal waters, this includes factors such as the cleanliness of seawater, its color, clarity, unusual odors, and auditory experience of the sea, among others. Regarding the waterfronts, it pertains to the ease of accessing the seawater, convenient transportation to the area, the ease of reaching the water's edge, the variety of activities available, and the overall approachability of the location. The adjacent environment includes urban landscapes, coastal vegetation coverage, shoreline maintenance, coastal congestion, openness of view by the shore, and the integration of coastal scenery with nearby buildings. Regarding culture and creativity, people often feel joy, novelty and feelings of locality when enjoying spaces. Regarding reputation and expectation, texts often include whether visitors wish to return, whether they would recommend the place to friends and family, and their overall satisfaction with the coastline.

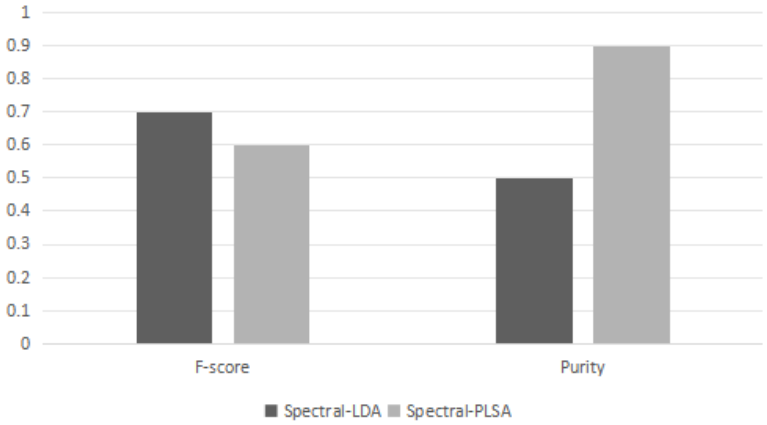


Figure 5. Examination of the validation of LDA model by F-score and Purity.

Table 2. Examples of the high-frequency keywords.

Topic Number	High-Frequency Keywords			
1	Photography	Sunset	Healing	Sunset Glow
2	Price	Service	Signpost	Public Facilities
3	Parent–child	Tide Pooling	Romance	Coast
4	Satay Noodles	Old Villa	Music	Culture
5	Queuing	Crowded	Waiting	Noisy

Table 3. Examples of the social media comments in relation to the five topics.

Topic No.	Text Examples	
1	"The sunset along Huandao Road is breathtaking! The sea water looks like a spilled palette, gradient from sapphire to gold. Sitting on the reefs listening to the waves, all worries are blown away by the sea breeze~ #XiamenHealingMoments#"	"Took the photo of my life at Yanwu Bridge Observation Deck! The sunset glow dyed the sea pink-purple, egrets flying close to the waves. At this romantic moment, I finally understand why it's called the 'Garden on the Sea'."
2	"The old villa café behind Gulangyu Pipe Organ Museum, listening to vinyl jazz under a century-old banyan tree, suddenly understood the true meaning of Minnan's 'tea ancient' (leisurely sitting). Time should be wasted like this!"	Shapowei Art West District is too cool! After browsing the vintage store and sitting on the embankment, watching the fishing boats swaying in and out of the sheltered dock, the air was filled with the smoke and fire of seafood stalls. This is the soul of Xiamen!"
3	"The intelligent shower at Huangcuo Beach is amazing! Scanning the code gives hot water with parent–child cubicles, but the trash can overflow in half an hour. . .Environmental protection details need to keep up #XiamenCoastalService#"	"Ripped off renting a sunshade at Yefengzhai! 150 yuan/hour is daylight robbery? But the lifeguards are responsible, always reminding of rip current dangers. Security points plus!"
4	"The live performance at Music Square was so exciting! The lights and music combined in the evening created an amazing atmosphere, and the children danced happily. Unfortunately, the sound system was too loud and a bit noisy, which was unbearable for the ears. But overall, it was a complete joy! #Music Square #Family Fun"	"The serene atmosphere of the Calligraphy Square is fantastic! I enjoy taking a walk alone, painting sketches while watching the sea, and the cultural inscriptions make me feel calm. The facilities are a bit old, but overall, it's very relaxing."
5	"Weekend Baicheng Beach is like a human can! Can't even squeeze in to dig sand with the kids, queued 25 min to rinse feet. . .Discouraged! Better come during off-peak hours #XiamenAvoid#"	"Zengcuoan entrance blocked by tour groups! Can't move an inch with a stroller, the smell of food stalls' oil smoke mixed with sweat, so annoyed want to escape immediately. . .Management should limit the number of visitors!"

Table 4. The five key topics summarized from text mining and high-frequency key words.

Dimension	Topic
Physical Perception	Coastal Waters
	Waterfronts
	Adjacent Environment
Non-physical Perception	Culture and Creativity
	Reputation and Expectation

2.4.3. Sentiment Detection and Semantic Transfer

During text preprocessing, it was discovered that some Weibo users incorporated emoticons or emojis into their posts. Consequently, for visitor sentiment analysis, the data can be categorized into two types: pure text and text mixed with emojis. Once preprocessed, pure text data can be directly utilized for subsequent sentiment detection, as outlined previously. The text-emoji mixed data were processed using a dual-channel neural network model, Bidirectional Encoder Representations from Transformers (BERT), trained on a combination of balanced social media corpora numbering in millions and news corpora numbering in hundreds of thousands. This model employs semi-supervised learning techniques, refined with industry-standard annotations, achieving a recognition accuracy for positive and negative sentiments of 85–90%, which is both more accurate and efficient than traditional computational models. By utilizing the Python Software Development Kit interface and inputting text, an original sentiment score ranging from 0 to 1 was obtained (with higher values indicating more positive sentiments). In this study, it was linearly mapped to a standard score of 0–5. For mixed-type text, the final sentiment score was the arithmetic mean of the calculated text and quantized emoji values. Emojis were categorized into five levels based on sentiment intensity (very positive = 4.5, positive = 3.5, neutral = 2.5, negative = 1.5, and very negative = 0.5). The highest sentiment score was obtained when multiple emojis were present. For easier visualization, the SnowNLP sentiment analysis module in Python's NLP toolkit was employed to calculate the sentiment scores for the 5780 blog posts by dividing the text into positive and negative categories. The sentiment words in the text were further analyzed with weighted scoring; scores close to 1 indicated positive sentiment, while scores close to 0 indicated negative sentiment.

2.4.4. Regional Spatial Analysis

To display the spatial distribution of sentiment scores more intuitively and remove duplicate check-ins in the same area or mass-copied comments from public accounts to boost view counts, a grid network with grid sizes of 50 m × 50 m was constructed in ArcGIS. The sentiment score for each grid is the arithmetic mean of the sentiment scores of all checked-in Weibo posts within that grid and is calculated as follows:

$$P_i = \frac{m_1 + m_2 + \ldots + m_x + m_n}{n} \quad (5)$$

where P_i represents the sentiment score in the ith grid that covers the park, m_x refers to the sentiment score of the xth Weibo post within the grid, and n is the total number of Weibo posts within the grid. Using the natural break classification method, the sentiment scores for these grids were divided into 10 categories, and the average sentiment score in each grid was distinguished using color shading.

2.4.5. Kernel Density Analysis

Coastal visitors' sentiments are diverse, and their Weibo posts about coastal recreation experiences can express a variety of feelings. An analysis identifying positive and negative

sentiments allows coastal recreation spaces to be categorized accordingly. Based on the sentiment score, the geographical locations of the posts were classified as follows: 0–2 for negative sentiment, 2–3 for neutral sentiment, and 3–5 for positive sentiment. The classified data were then analyzed in ArcGIS using spatial analysis to compute clusters of point features; the equation used is as follows:

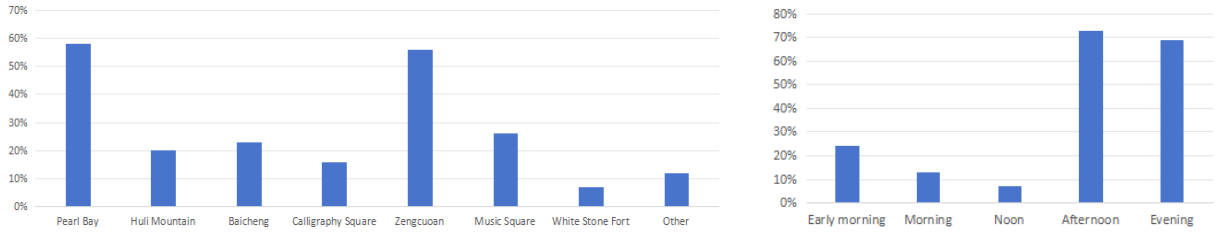
$$D = \frac{3(1 - k_{scale}^2)^2}{\pi r^2} \quad (6)$$

The k_{scale}^2 represents the ratio of the distance from the grid center to other points within the search radius to r^2. The intensity of clustering is indicated by color depth.

3. Results

3.1. Tourists Preferences

According to observations of the flow of people in different sections of these coastal areas in Xiamen city, people gather from the west, north, and east, and the flow varies depending on the popularity of each section. Statistical analysis of the gathered social media posts (Figure 6a) reveals that Baicheng and Zengcuoan are the most popular sections and, therefore, have the highest pedestrian flow, while the least popular section across all routes is White Stone Fort, with only 8% of visitors choosing to go there. The visit periods were mainly in the afternoon and evening (Figure 6b). After performing a word frequency analysis of all preprocessed Weibo texts, the top 50 most frequent characteristic words in tourists' online reviews of coastal spaces were obtained (Table S1). The high-frequency characteristic words are shown in Figure 7, which was produced using word cloud generator.



(**a**) (**b**)

Figure 6. (**a**) Popularity feature of the southern coastal spaces of Xiamen. (**b**) Preferences of visiting period to the southern coastal spaces of Xiamen.



Figure 7. Word cloud of coastal visitor's online Weibo posts.

Recreational activities in the southern coastal spaces of Xiamen are rich and diverse, reflecting the different preferences and purposes of visitors (Figure 8). The overall survey results indicated that chatting and scenic viewing were the most popular leisure activities, accounting for 64% and 62%, respectively. Strolling was also one frequently chosen activity (41%). Regarding leisure activities, watching performances and beach activities were relatively popular, accounting for 48% and 42%, respectively, in addition to chatting and sightseeing. Among the Weibo participants, 20% took part in marine sports activities, while 14% participated in running. The different age groups exhibited different activity preferences. For example, those who chose activities popular among seniors, such as martial arts, qigong, and chess, accounted for 2%, while those who opted for activities such as beach ball games and camping, which are favored by young people, accounted for 6%. This indicates that visitors have distinct preferences for recreational activities.

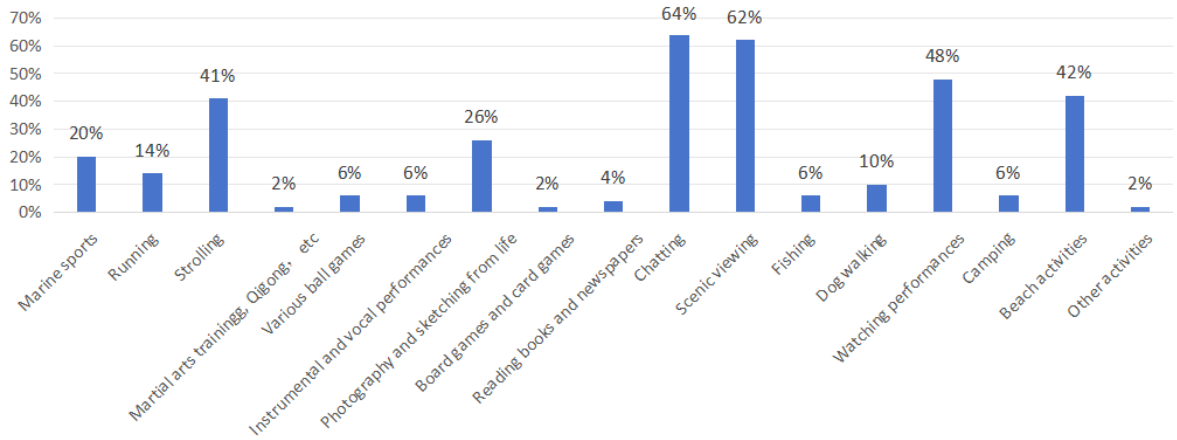


Figure 8. The statistical chart of popular activities among coastal tourists.

3.2. Sentiment Distribution

In ROSTCM6, visitors' text sentiments were categorized into three types: positive, neutral, and negative. The sentiment score intervals for these categories were (5, +∞), [−5, 5], and (−∞, −5), respectively. Using clustered data as input, the sentiment analysis function module in ROSTCM6 was employed to determine the sentiment score for each Weibo text, resulting in 1764 texts being classified as positive, accounting for 56.47%. There were 509 neutral texts, comprising 16.29%. Figure 9 illustrates the comparative distribution of Weibo user sentiments, showing that 851 texts (27.24%) were classified as negative. Statistical analysis of the sentiment polarity proportion for each topic was conducted, and the results for the three categories are presented in Table 5.

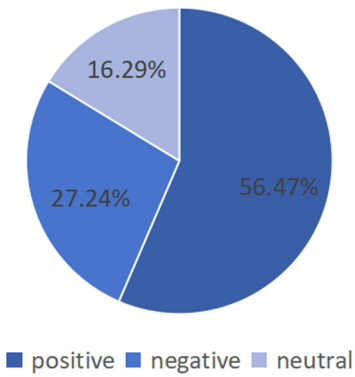


Figure 9. Summary of the tourist sentiment from social media posts.

Table 5. The results of sentiment categories by topic.

Theme	Type	Quantity	Proportion %
Topic 1 Coastal Waters	Positive	813	49.18
	Neutral	364	22.02
	Negative	476	28.80
Topic 2 Waterfronts	Positive	428	73.67
	Neutral	57	9.81
	Negative	96	16.52
Topic 3 Adjacent Environment	Positive	402	64.22
	Neutral	65	10.38
	Negative	159	25.40
Topic 4 Culture and Creativity	Positive	96	87.27
	Neutral	10	9.09
	Negative	4	3.64
Topic 5 Reputation and Expectation	Positive	25	16.23
	Neutral	13	8.44
	Negative	116	75.33

3.3. The Spatial Changes in Tourist Sentiments

Figure 10 shows an example of the spatial distribution of tourist sentiment: one surveyed coastal zone from Baicheng-Huli Mountain to White Stone Fort section. The overall sentiment level of tourists in the southern coastal spaces of Xiamen Island was generally high. According to GIS data, the study area can be divided into 132 grids, with 109 grids falling within the positive sentiment range of 3.0–5.0, and only five grids showing negative sentiments with scores below 2.0. The distribution of negative sentiment spaces is extremely limited. High sentiment score nodes (average $\geq$ 4.0) were mainly found in Music Square (an open-air stage blending performance with sea views), Calligraphy Square (a cultural landmark and viewing platform), Pearl Bay Beach (a premium waterfront activity area), and Huli Mountain Fortress (a historic site and coastal lookout). In contrast, low-sentiment score nodes (average $\leq$ 1.8) were concentrated at the Baicheng scenic area ticket office (queuing congestion area) and parking lots on the west side of Zengcuoan (traffic chaos area). The spatial distribution pattern of sentiments indicates that areas with high sentiment scores generally have the triple characteristics of "integration of sea-land landscapes + cultural theme reinforcement + open public activities", while areas with low sentiment scores are directly related to bottlenecks in service facilities such as ticket purchasing and parking.

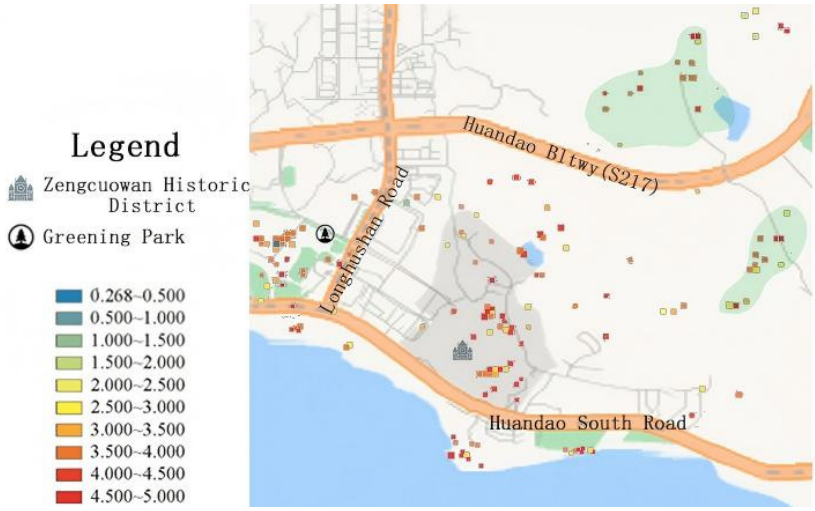


Figure 10. The spatial distribution of tourist sentiment scores in Zengcuoan historical district.

Through the above spatial analysis, hotspots of sentiment expression were identified. Data clustering revealed that tourists' sentiment expressions were unevenly distributed, and coastal service zones were the main nodes for sentiment expression. Typical average sentiment scores are listed in Table 6. Landscape-dominated spaces have received higher sentiment scores. Local architectural style may also drive sentiment premiums, and some cultural and creative shops in renovated southern Fujian red-brick houses (Zengcuoan) have sentiment scores 23% higher than modern steel-structure shops. The penetration of natural features determines the sentiment baseline (grids directly facing the sea have an average value of 4.2, which is 35% higher than that of the secondary areas at 3.1). The accounts of the number of posts and the average sentiment scores in the different zones are shown in Table 7.

Table 6. Common sentiment expression nodes in the southern coastal region of Xiamen.

Hotspot Ranking	Land Use Type	Representative Location	Mean Sentiment Score
1	Tourism Attractions	Huli Mountain Fort	4.6
2	Commercial areas	Zengcuoan Historical District	3.8
3	Open spaces	Calligraphy Square	4.3
4	Transportation Hubs	White Stone Fort Bus Station	2.9

Table 7. The statistics of hotspots of sentiment posts across the surveyed coastal areas.

Serial No.	Area	Number of Posts	Mean Sentiment Score
1	Zengcuoan and Yuqiao Entrance	1746	3.695
2	Music Square	799	3.641
3	Baicheng Beach (Near Xiamen University)	495	3.673
4	Calligraphy Square	267	3.514
5	Pearl Bay Beach and Bathing Area	245	3.502
6	Huandao Road Wooden Plank Road	144	3.451
7	Reef Area near Sun Bay	89	3.442
8	Beach near Haiyuntai	50	3.397

The results also reflect the spatial differentiation of areas favored by tourists when expressing their sentiments. The spatial distributions of sentiment types, positive and negative, were analyzed separately. Dense zones of positive sentiments were mainly located in Zengcuoan's cultural and creative streets (clusters of dining and creative shops), the Huli Mountain Fortress viewing platform (historic landmark with sea views), Pearl Bay Music Square (street performance and beach activity zone), and Calligraphy Square rest area (seafront promenade and art installations). These areas have become cores for sentiment expression, with the number of interactive experiences (measured by volume) being 40% higher than in purely sightseeing areas. Secondary positive sentiment zones, such as the Baicheng-Huli Mountain, rocky shore exploration area (intertidal organism observation), the Ring Road cycling section (sea breeze experience), and Xitouxia wedding photography spot (romantic landscapes), are characterized by dynamic coastal activities, primarily triggering tourists' instant sentiment sharing. By contrast, purely natural landscape areas generated fewer online posts.

3.4. The Temporal Changes in Tourist Sentiments

Tourist sentiments were examined monthly, with a focus on the temporal characteristics of these sentiments displayed over time (Figure 11). In 2024, the average annual sentiment score for tourists in the surveyed coastal area was 3.114, reflecting a moderately positive trend. Although there were notable monthly fluctuations, no significant correlation

was found between the sentiment score and the number of social media posts (Pearson's r = 0.21, $p > 0.05$).

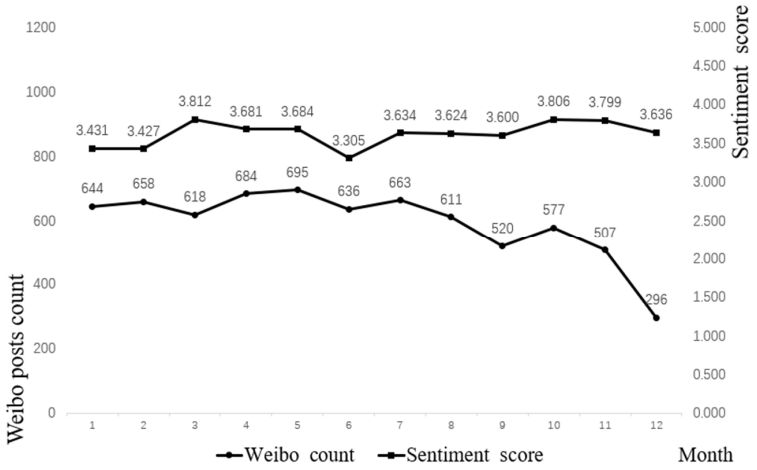


Figure 11. Monthly variation of tourists' overall sentiment scores in the whole study area.

The sentiment peak in March had an average score of 3.812. Its high-frequency words (such as "bougainvillea" accounting for 12.7%, "sea breeze" for 9.3%, "flower corridor" for 8.1%) reflect the positive impact of recreational activities (such as the Island Ring Road Bougainvillea Flower Exhibition) and the suitable seasonal climate in the study area during the spring of March. The sentiment trough in June had an average value of 3.305, and its high-frequency words (such as "muggy" at 15.2%, "heavy rain" at 11.6%, "typhoon" at 7.4%) indicate that the high temperature and humidity during this period have a significant impact on tourists' sentiment expressions.

Significant fluctuations were also observed during transitional periods, which is noteworthy. This could be attributed to the calm sentiments in January and February, influenced by cold air (3.42 $\pm$ 0.11), and a rebound in sentiments driven by traditional holidays such as Tomb Sweeping Day and the International Labor Day Holiday in April and May (3.68 $\pm$ 0.09). The impact of climate was further validated in June, when the mean sentiment score dropped significantly, possibly because of rising temperatures. However, cultural festivals can mitigate these negative effects; during the Dragon Boat Festival, sentiment scores increased by 17.3% compared to the same period.

We employed the Mean Absolute Deviation (Mad) and Coefficient of Variation (CV) to assess sentiment fluctuations across the three types of coastlines, as detailed in Table 8. Our findings indicate that artificial coastline sections, such as Calligraphy Plaza and Music Plaza, which are entirely man-made and of high quality, have exhibited the most stable sentiment levels (Figure 12). These plazas are well-developed and enjoy a strong reputation among travelers. In contrast, natural coastlines, including beaches, experienced greater fluctuations in sentiment owing to climate and weather factors (Figure 13). Mixed coastlines showed the highest fluctuations (Figure 14) according to both the Mad and CV volatility tests. These semi-natural and semi-artificial areas are influenced by both climate and weather, and their fluctuations may be linked to varying levels of satisfaction with the surrounding man-made facilities, which can differ between areas.

Table 8. A comparison of sentiment fluctuation in three types of coastlines.

Ranking	Coastline Type	Mean Absolute Deviation	Coefficient of Variation
1	Mixed	0.164	6.79%
2	Natural	0.150	6.18%
3	Artificial	0.082	2.72%



Figure 12. Monthly variation of sentiment scores in artificial coastlines.

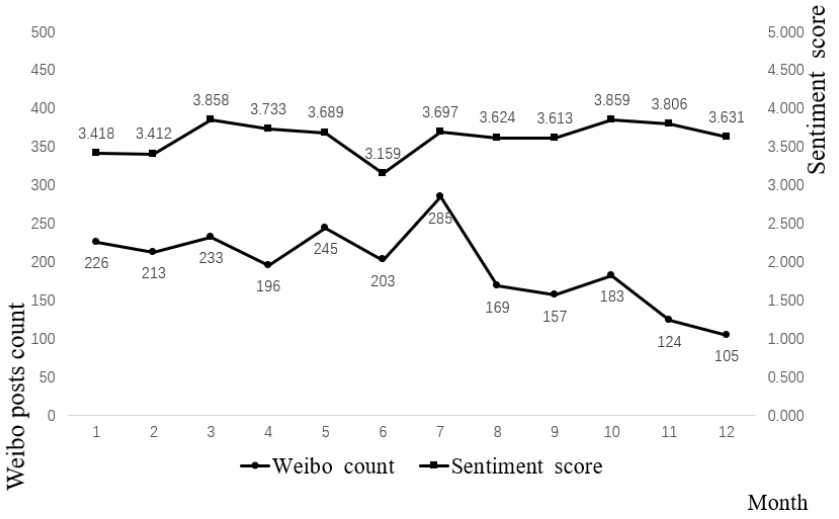


Figure 13. Monthly variation of sentiment scores in natural coastlines.

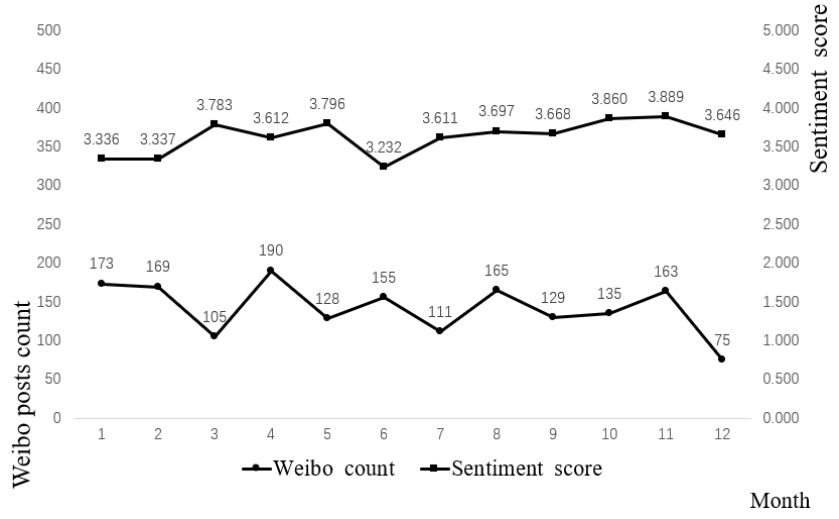


Figure 14. Monthly variation of sentiment scores in mixed coastlines.

4. Discussion

4.1. Constructing a New System of Understanding Coastal Recreation Quality from the Perspective of Tourist Perception

This study found the existence of influencing relationships between physical and non-physical themes and identified the main problems existing in the coastal spaces of Xiamen. By integrating the classification approach of coastline typology with feedback from public opinion and drawing on the study and evaluation of Xiamen's unique coastal recreational areas, the coastal regions can be categorized into high, moderate, and low quality levels. Through a spatial satisfaction survey based on the dimensions of tourist perception, the main issues with local coastal environments can be identified. In response to these reported issues, suggestions can be provided towards the improvement of coastal space. Work on spatial elements (such as the uncovered core thematic impressions on coastal waters, waterfronts, and adjacent environments) often positively correlated with the improvement of visitor experiences. These physical attributes provide the backdrop for various activities, from leisurely strolls along the shore to water-based recreational pursuits. The visual appeal, ecological diversity, and sensory experiences offered by these spatial elements contribute to the overall satisfaction and memorable encounters of visitors. Efficient management of non-spatial elements, such as streamlining ticketing processes to reduce queuing times, can significantly improve visitor satisfaction by minimizing frustration and maximizing time spent enjoying the coastal environment.

According to the results of the sentiment analysis, although blogs posted by netizens show a range of sentiments from positive to neutral to negative, the overall sentiment tendency is positive. The content of each topic varies, and so do the sentiment tendencies displayed in online comments. This research indicates that tourists demonstrate a significantly positive sentiment tendency in the coastal spaces of the study area. The proportion of positive sentiments in the first four topics exceeded that of negative sentiments, likely because of the superiority of its coastal recreation landscape resources (e.g., the visual appeal created by the blue sea and golden sands along with historical landmarks such as the Huli Mountain Fort), the joy brought by the design of activities in the coastal recreation space (e.g., immediate joy triggered by scenarios such as intertidal explorations and open-air concerts), and the comfort experience brought to tourists by the facilities within its space (e.g., the comfort foundation constituted by the ring-road cycling system and shaded rest points). Meanwhile, it can be observed that negative sentiments are significantly higher than positive sentiments in topic 5, with words like "heard", "friends advised against", and "congestion" appearing in tourist comments. It is speculated that the negative sentiments mainly stem from issues such as an incomplete parking management system and inappropriate spatial scale in some areas within the study section, or possibly from false advertising by some tourists rather than intrinsic flaws in the space itself. The findings also indicate that tourists are more active on social media in various functional zones that combine natural landscapes and sociocultural activities. In addition, sentiment sharing activities is primarily triggered in those areas associated with positive emotions.

4.2. Recommendations for the Southern Coastal Areas of Xiamen City

The analysis of the text related to more negative sentiments shows that coastal recreation spaces may still have unreasonable arrangements, such as weaknesses in cycling tour paths, current spatial settings lacking effectively organized commercial spaces on pedestrian roads, lack of good seascape-viewing spaces, and shortcomings in functional settings. Simultaneously, the supporting facilities of the coastal spaces within the study section are not sufficiently abundant, with the most needed supplements being footwash basins and showers, followed by public toilets, storage areas, and rescue facilities. For

landscape installations, indoor, outdoor, and semi-outdoor rest spaces that can be shielded from the sun and rain can be supplemented. Current coastal buildings may be improperly planned, as restaurant buildings lack public openness to the sea interface and the utilization of historical buildings is low. The above analysis indicates possible seasonal flexible management mechanisms in response to sentiments in coastal areas. Areas with higher user satisfaction are often linked with the quality of the coastal area, such as water bodies and nearby environments, leading to positive developments in visitors' perceptions and increasing their willingness to revisit and promote coastal recreation activities.

By tracking how visitor sentiments change over time, managers can develop more nuanced and effective strategies for different types of coastal areas. For natural coastlines, this understanding allows for the implementation of seasonal management approaches that can help mitigate the impact of climate-driven fluctuations on visitor experiences. This might involve adjusting access during certain periods, enhancing protective measures during vulnerable seasons, or providing educational programs that align with natural cycles. In mixed zones, where both natural and artificial elements coexist, the focus shifts to balancing the appeal of natural features with the stability offered by man-made structures. By analyzing sentiment trends, managers can identify when and how to enhance cultural and service elements to complement the natural environment. This could involve introducing seasonal events, improving amenities, or creating interpretive experiences that highlight the unique characteristics of these hybrid coastal areas. For artificial zones, the emphasis is on maintaining high satisfaction levels through optimized facility maintenance. Understanding sentiment patterns helps in predicting wear and tear, scheduling renovations during low-impact periods, and continuously improving infrastructure to meet evolving visitor expectations.

4.3. Contributions to the Literature and Implications to Other Coastal Cities

It is reported that text representation derived from TF-IDF models are superior to those obtained through traditional techniques like the Hownet word similarity (Zhao et al., 2021) [37]. In this study, the adoption of the deep learning TNNFMB approach overcomes the constrains of previous text mining methods and solves the problem of fine-grained classification under small sample conditions by integrating transfer learning and neural network algorithms into tourism sentiment analysis. Compared with previous research results, the effectiveness of this algorithm was further verified by the LDA topic model. In the study on tourism services by Lin et al. (2021), the authors constructed an emotion orientation and used comments from tourism platforms, which was mainly on the spatiotemporal distribution of tourism information in attractions at small site scale [16]. Our study indicates that superior natural environment, rich entertainment activities, and distinctive historical and cultural features can have a positive impact on tourists' emotions, which are all verified at large city scale.

This study's roadmap, while grounded in a local context, provides a foundational framework for understanding coastal tourism experiences that can be extrapolated to other coastal settings. The findings from Xiamen offer implications for other coastal cities globally. By understanding tourist sentiment patterns, thematic experiences, and spatial–temporal rhythms, stakeholders can identify both universal and context-specific factors shaping tourist perceptions. Conversely, differences emerging from unique cultural attributes, urban infrastructure, or environmental management practices highlight critical areas for localized intervention. For instance, a strategy that successfully enhanced the visitor experience in Xiamen's harbor area might be adapted to a historic port city, but must be tailored to its distinct heritage and spatial layout. By developing a data-driven coastal experience interactive system, the inherent laws and universal characteristics of

the landscape quality of coastal cities are revealed, providing decision-making basis and improvement strategies for the planning and optimization of coastal cities.

5. Conclusions

This study utilizes text mining, sentiment detection, and classification to analyze tourist sentiments in coastal areas using social media data. The main findings are as follows: (1) five core topics were identified: coastal waters, waterfronts, adjacent environments, culture and creativity, and reputation and expectation; (2) tourist sentiment is predominantly positive; (3) tourist sentiment exhibits both spatial and temporal variation. Areas that emphasize sea–land landscapes, enhance cultural themes, and support open public activities generally achieve high and stable sentiment scores. In addition, the natural and mixed coastlines showed significant seasonal sentiment fluctuations. This study has achieved several innovative breakthroughs. The employment of the deep learning Two-way Neural Network Fusion Model, TNNFMB, has overcome the limitations of traditional text mining approaches by integrating transfer learning and neural network algorithm into the tourism sentiment analysis, solving the fine-grained classification problem under small sample conditions. By evaluating social media posts, issues within coastal spaces and recreational environments can be identified more effectively. The development of coastal spaces is crucial for enhancing the quality of coastal cities, and their planning must also consider improvements to the environment and transport of physical elements in response to social needs to ensure sustainable development that integrates a good visit experience and well-being.

6. Limitations and Future Prospects

Our objective was to create a scientific, data-driven system for assessing urban landscapes, aiming to uncover the fundamental principles and shared characteristics of landscape quality in coastal cities. This system forms a basis for decision-making and offers strategies for enhancing and planning urban areas. While coastal spaces encompass numerous factors, this study concentrated on physical perceptions, such as waterfronts and their surrounding environments, as well as non-physical perceptions, including reputation and expectations. There remain gaps in understanding the full perceptions of coastal visitors from their social media posts. While we employed data filtering techniques (e.g., removing posts with excessive commercial keywords, duplicate content, or suspicious patterns) and focused on natural language indicative of personal experiences, it remains impossible to definitively verify the authenticity and organic origin of every single post analyzed. We acknowledge this inherent limitation of user-generated content research and its implications for the absolute veracity of urban big data points. Future research will broaden the scope of the investigation, allowing for a more precise identification of shared and region-specific issues by comparing similarities and differences across various coastal cities. Coastal spaces are intricate systems, and by creating comprehensive evaluation models using diverse data sources, we can further analyze changes in coastline spaces through remote sensing images and GIS and assess ecological sensitivity and spatial utilization efficiency.

Supplementary Materials: The following supporting information can be downloaded at https://www.mdpi.com/article/10.3390/app152111721/s1, Table S1: Top-50 High-Frequency Sentiment Words in Weibo Posts in Xiamen Southern Coastal Areas.

Author Contributions: Conceptualization, Y.W. and Z.Z. (Zhiyu Zhang); methodology, Y.W., Z.Z. (Zhiyu Zhang) and Z.Z. (Zhijun Zhang); software, Z.Z. (Zhiyu Zhang); validation, Y.W., Z.Z. (Zhiyu Zhang) and Z.Z. (Zhijun Zhang); formal analysis, Z.Z. (Zhiyu Zhang); investigation, Y.W. and Z.Z. (Zhiyu Zhang); resources, Y.W. and Z.Z. (Zhijun Zhang); data curation, Z.Z. (Zhiyu Zhang); writing—

original draft preparation, Y.W. and Z.Z. (Zhiyu Zhang); writing—review and editing, Y.W. and Z.Z. (Zhijun Zhang); visualization, Z.Z. (Zhiyu Zhang); supervision, Y.W.; project administration, Y.W.; funding acquisition, Y.W. and Z.Z. (Zhijun Zhang). All authors have read and agreed to the published version of the manuscript.

Funding: This study was funded by Tianjin Association for Science and Technology (Grant No. 2024100D).

Data Availability Statement: The raw data supporting this study can be available upon reasonable request to the authors.

Acknowledgments: The authors express their gratitude to the anonymous reviewers for their valuable suggestions that helped enhance this paper.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Wang, F.; Zhu, D. The study on strategy of sustainable development in coastal tourism resources from the global change perspective. *J. Resour. Ecol.* **2014**, *5*, 32–41. [CrossRef]
2. Zhang, S.; Xue, H.; Zhang, T.; Ju, H. The Spatial–Temporal Evolution and Driving Factors of the Coastal Tourism Economy in China. *Land* **2024**, *13*, 1542. [CrossRef]
3. Hsiung, A.; Hartanto, R.; Bhatia, N.; Morris, R. Challenges and Opportunities for Implementing Nature-Based Coastal Protection in an Urbanised Coastal City Based on Public Perceptions. *J. Environ. Manag.* **2024**, *370*, 122620. [CrossRef]
4. Chatterjee, P.; Dasgupta, R.; Paul, A.K. Beach Beauty in Bengal: Perception of Scenery and Its Implications for Coastal Management in Purba Medinipur District, Eastern India. *Mar. Policy* **2022**, *139*, 105034. [CrossRef]
5. Chen, C.-L.; Teng, N. Management Priorities and Carrying Capacity at a High-Use Beach from Tourists' Perspectives: A Way Towards Sustainable Beach Tourism. *Mar. Policy* **2016**, *74*, 213–219. [CrossRef]
6. Cavallaro, F.; Irranca Galati, O.; Nocera, S. A Tool to Support Transport Decision-Making in Tourist Coastal Areas. *Case Stud. Transp. Policy* **2019**, *7*, 540–553. [CrossRef]
7. Das, S.; Bhattacharya, S. Factors Affecting Beach Walkability—Tourists' Perception Study at Selected Beaches of West Bengal, India. *J. Outdoor Recreat. Tour.* **2021**, *35*, 100423. [CrossRef]
8. Ser, J.; Yang, B. How GeoAI Improves Tourist Beach Environments: Micro-Scale UAV Detection and Spatial Analysis of Marine Debris. *Land* **2025**, *14*, 1349. [CrossRef]
9. Pafi, M.; Flannery, W.; Murtagh, B. Coastal tourism, market segmentation and contested landscapes. *Mar. Policy* **2020**, *121*, 104189. [CrossRef]
10. Parasuraman, A.; Zeithaml, V.; Berry, L. SERVQUAL: A Multiple-Item Scale for Measuring Consumer Perceptions of Service Quality. *J. Retail.* **1988**, *64*, 12–40.
11. Sweeney, J.; Soutar, G. Consumer Perceived Value: The Development of a Multiple Item Scale. *J. Retail.* **2001**, *77*, 203–220. [CrossRef]
12. Hanayanagi, K. *Waterfronts of Cities and Human Behavior*; Gihōdō Shuppan Co., Ltd.: Kanagawa, Japan, 1999.
13. Roca, E.; Villares, M.; Ortego, M.I. Assessing Public Perceptions on Beach Quality According to Beach Users' Profile: A Case Study in the Costa Brava (Spain). *Tour. Manag.* **2009**, *30*, 598–607. [CrossRef]
14. Zhao, Y. Text Mining Method for Network Public Opinion Based on Big Data. *China Comput. Commun.* **2021**, *33*, 184–186.
15. Lave, R.; Lane, S. *The Field Guide to Mixing Social and Biophysical Methods in Environmental Research*; Open Book Publishers: Cambridge, UK, 2025.
16. Lin, Z.; Xie, J.; Yang, T.; Liu, Z.; Zhao, J. Construction Method of Multi-Theme Sentiment Dictionary for Tourism. *Geogr. Geo-Inf. Sci.* **2021**, *37*, 22–27.
17. Oliveira, T.; Araujo, B.; Tam, C. Why Do People Share Their Travel Experiences on Social Media? *Tour. Manag.* **2020**, *78*, 104041. [CrossRef]
18. Zhang, Y.; Wang, M.; Yang, X.; Zhang, R. Urban Commercial Space Vitality Evaluation Method Based on Social Media Data: The Case of Shanghai. *Land* **2025**, *14*, 697. [CrossRef]
19. Zuo, M.; Li, C. Review of Network Public Opinion Research: From Theoretical Exploration to Practical Application. *J. Intell.* **2017**, *36*, 71–78.
20. Li, R.; Zhang, Y. Research on Big Data Aided Scientific Research Management Decision—Based on ROST Text Mining System. In Proceedings of the CIBDA 2022, 3rd International Conference on Computer Information and Big Data Applications, Wuhan, China, 25–27 March 2022; pp. 1–5.

21. Jiang, Y.H.; He, Y. Perception Study of Tourist Destination Image based on Online Comment Data: A Case Study of Solidary Beauty Peak & Royal City Scenic Area in Guilin City. *Archit. Cult.* **2024**, *05*, 258–261.
22. Lee, H.; Kang, Y. Mining tourists' destinations and preferences through LSTM-based text classification and spatial clustering using Flickr data. *Spat. Inf. Res.* **2021**, *29*, 825–839. [CrossRef]
23. Hileman, J.D.; Angst, M.; Scott, T.A.; Sundström, E. Recycled text and risk communication in natural gas pipeline environmental impact assessments. *Energy Policy* **2021**, *156*, 112379. [CrossRef]
24. Cao, H.; Wang, M.; Su, S.; Kang, M. Explicit quantification of coastal cultural ecosystem services: A novel approach based on the content and sentimental analysis of social media. *Ecol. Indic.* **2022**, *137*, 108756. [CrossRef]
25. Li, T.; Xiang, Z.; Li, Y.F. POIs-based public preferences mapping on imbalanced supply-demand of recreation services can support sustainable coastal beach management. *Front. Mar. Sci.* **2024**, *11*, 1288112. [CrossRef]
26. Zhang, C.Y.; Sun, J.L.; Ding, Y.Q. Topic Mining for Microblog Based on MB-LDA Model. *J. Comput. Res. Dev.* **2011**, *48*, 1795–1802.
27. Jiang, J.; Ma, R. Analysis of Personalized Tourism Preference Characteristics Based on Text Mining. *Enterp. Econ.* **2017**, *36*, 129–133.
28. Qiu, Z.; He, B. Network Public Opinion Clustering and Sentiment Evolution Analysis Based on PCA-Spectral-LDA: A Microblog Text Mining Study. *J. Syst. Sci. Math. Sci.* **2021**, *41*, 2906–2918.
29. Gu, J.P.; Wang, Y.B.; Liu, G.W.; Tian, Z. Tourism Image Mining Based on User Generated Big Data. Taking Chongqing as An Example. *Mod. Urban Res.* **2019**, *12*, 117–125.
30. Wan, Y.; Zhang, C.; Wang, D.; Li, X.; Hao, X. Tourist impression mining and sentiment analysis based on online review data. *Inf. Eng.* **2023**, *9*, 15–29.
31. Wang, G.; Qiu, Y. Text Clustering Research Based on Ontology and Similarity. *Appl. Res. Comput.* **2010**, *27*, 2494–2497.
32. Huang, B.; Yang, Y.; Mahmood, A.; Wang, H. Microblog Topic Detection Based on LDA Model and Single-Pass Clustering. In Proceedings of the International Conference on Rough Sets and Current Trends in Computing, Chengdu, China, 17 August 2012; Springer: Berlin/Heidelberg, Germany, 17 August 2012; pp. 166–171.
33. Peng, Y.; Wan, C.; Jiang, T.; Liu, D.; Liu, X.; Liao, G. Extraction of Product Features and Sentiment Words Based on Semantic-Constrained LDA. *J. Softw.* **2017**, *28*, 676–693.
34. Chi, M.M.; Pan, M.Y.; Wang, W.J. A Cross-platform Comparative Study of Reviews on Sharing Accommodation and Hotels Reservation Platform: Combined with LDA- SNA and Sentiment Analysis. *Libr. Inf. Serv.* **2021**, *65*, 107–116.
35. Li, P.; Chen, T.; Wang, F.Y.; Wang, X.G. Urban tourism community image perception and differentiation based on online comments: A case study of Beijing. *Geogr. Res.* **2017**, *36*, 1106–1122.
36. Wang, X.; Sheng, S.; Lu, Y. A Microblog Public Opinion Analysis Model Based on Topic Clustering and Sentiment Intensity. *Data Anal. Knowl. Discov.* **2018**, *2*, 37–47.
37. Zhao, Z.; Liu, Y.; Zhang, Y.; Zhou, W.; Meng, L. Hotel user preference model based on Lasso-LDA. *Comput. Appl. Softw.* **2021**, *38*, 19–26.

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.