

Article

Attribute Correlation and Hotel Recommendation Based on Online Reviews

Yanyan Chen ¹, Sining Chen ¹, Zhuoying Wu ¹, Sumin Yu ^{1,*} and Zhijiao Du ^{2,3}¹ College of Management, Shenzhen University, Shenzhen 518060, China² School of Foreign Languages and Business, Shenzhen Polytechnic University, Shenzhen 518055, China³ Business School, Sun Yat-sen University, Shenzhen 518107, China

* Correspondence: yusumin@szu.edu.cn

Abstract

Online travel website reviews contain abundant hotel attribute information. Nevertheless, evaluations of such hotel attributes are not completely objective, as they are affected by subjective perceptions from different types of travelers. This study explores the correlations of attribute perceptions across different traveler groups and optimizes hotel recommendation models from the perspective of attribute correlation. Hotel attributes are analyzed based on online reviews collected from TripAdvisor, and the TOPSIS method is used to identify the core attributes. The SentiWordNet sentiment lexicon and PolarityRank algorithm are utilized to calculate attribute sentiment values, which are then applied to OLS regression analysis to explore the relationships between attributes. Finally, a hotel recommendation model based on the collaborative filtering algorithm is constructed, incorporating traveler types and attribute correlation. The research results show core attributes for different types of travelers, verify the existence of attribute perception correlation, and demonstrate that a model considering attribute correlation enhances recommendation accuracy, with an accuracy improvement of 3.16 percentage points.

Keywords: attribute correlation; hotel recommendation; online reviews**MSC:** 90B50, 68T05

Academic Editor: Pasquale De Meo

Received: 5 May 2026

Revised: 25 June 2026

Accepted: 1 July 2026

Published: 9 July 2026

Copyright: © 2026 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 \(CC BY\)](https://creativecommons.org/licenses/by/4.0/) license.

1. Introduction

With the development of information technology, the “Internet plus tourism” model has revolutionized how travelers share experiences and complete accommodation bookings. User-generated content on social platforms has become a core source for travelers to acquire tourism information and support decision-making [1–3].

Among various types of tourism information, hotel attributes occupy a crucial position, as they directly reflect the actual service quality of hotels and significantly affect travelers’ overall satisfaction [2,3]. Potential tourists usually make consumption decisions according to the performance of hotel attributes they focus on. Nevertheless, travelers’ evaluations of hotel attributes are far from fully rational. For example, although the location attribute features inherent geographical objectivity and is theoretically independent of other attributes, existing research confirms that travelers’ experience of other attributes can influence their evaluation of hotel location [4,5], which proves the

irrational nature of travelers' hotel assessments. It is evident that hotel attributes, as major evaluation criteria for travelers, cannot be simply classified as purely objective or subjective. Travelers' attribute evaluations are rooted in their personal subjective feelings, which are further shaped by individual characteristics including travel experience, gender, age and traveler categories. Among all influencing factors, the tourism industry attaches the greatest importance to traveler types, covering business travelers, family travelers, couple travelers, group travelers and solo travelers [6,7], all of whom hold differentiated preference tendencies. On this basis, this study focuses on exploring the dominant hotel attributes favored by different groups of travelers.

As for hotel attributes, we can define them as distinct characteristics of hotel-provided products and services [8]. Multi-attribute theory holds that various attributes exert differentiated impacts on travelers' perceived experience [9]. The ring model proposed by Levitt [10] illustrates that inner-ring core attributes of products and services are fundamental to satisfying customer expectations, whereas outer-ring attributes help boost customer satisfaction by supporting and optimizing core attributes. Compared with auxiliary services, hotels' core products and services exert a more pronounced influence on customer satisfaction. To sum up, core attributes refer to essential products and services that fulfill travelers' basic expectations and exert greater effects on their overall travel experience. Accordingly, different traveler groups prioritize distinct hotel attributes, so their perceived core attributes also differ. Furthermore, how do core attributes influence other attributes? When core attributes deliver satisfactory performance, will travelers form positive perceptions of other auxiliary attributes? Conversely, if core attributes fail to meet preset expectations, will travelers develop biased generalizations and thus generate negative attitudes toward other attributes?

The abundance of online reviews can effectively reflect hotel attributes. However, the explosive growth of review data has resulted in severe information overload, which further drives the demand for intelligent hotel recommendation systems. Previous research on hotel recommendation usually analyzes travelers' preferences for each attribute separately, on the premise that all hotel attributes are mutually independent. As mentioned above, travelers' attribute evaluations are affected by subjective factors, and travelers from different groups show distinct priority tendencies toward hotel attributes. Attributes that travelers attach great importance to can significantly affect people's judgment of other attributes. In fact, hotel attributes are mutually complementary and closely interdependent. Hence, incorporating attribute correlations into the design of hotel recommendation models can help travelers make more rational accommodation choices. Accordingly, this study puts forward the following three research questions:

1. What are the core attributes valued by different traveler groups?
2. How does the perception of core attributes influence the evaluation of other attributes?
3. Can hotel recommendation systems be further optimized by integrating attribute correlation mechanisms?

To address the above research questions, this study categorizes traveler groups by travel purposes, analyzes online reviews using the Word2vec algorithm and sentiment analysis, converts text reviews into data, and uses regression analysis to study how core attributes influence other attributes. Finally, we propose a hotel recommendation model that takes both traveler types and inter-attribute relationships into account, so as to provide reliable decision support for travelers' hotel selection.

The remainder of this study is organized as follows. Section 2 reviews the relevant literature for study motivation. Section 3 introduces an attribute correlation framework. Section 4 presents and discusses the empirical results of attribute correlation analysis. Section 5 details the hotel recommendation model's construction and case analysis.

Section 6 concludes with theoretical and practical implications, limitations, and future research directions.

2. Literature Review

Hotel attributes shape travelers' accommodation experiences, among which core attributes exert the most prominent impacts. Considering the relationships between attributes can optimize hotel recommendations. This study reviews literature on hotel attributes, traveler perceptions, and hotel selection recommendations to discuss the current research in these areas.

2.1. Hotel Attributes and Traveler Perceptions

The scholar Lewis introduced the concept of hotel attributes in 1984, referring to the decisive features of products or services that influence consumers' choice of one hotel over another [8]. Hotel attributes have long been a focal point in tourism research, with scholars making significant efforts to identify them to assist hotel managers in developing clear strategic plans.

Xu et al. [11] categorized hotel attributes into product and service dimensions: product attributes cover guest rooms, supporting facilities and surrounding environments, while service attributes involve staff attitudes, service behaviors and service efficiency. Jokelainen et al. [12] identified hotel attributes as value for money, uniqueness, service quality, room facilities, breakfast experience, and responsibility. Francesco et al. [13] classified hotel attributes as hotel amenities, budget, service, location, and general satisfaction. Sharma et al. [14] further divided 27 specific hotel attributes into five major dimensions: relative advantage, compatibility, complexity, observability and communication. Additionally, many existing studies use evaluation dimensions from travel websites as hotel attributes, with the six common attributes on TripAdvisor being value, location, service, room, cleanliness, and sleep quality. This study aims to investigate attribute relationships using online reviews from tourism websites, where the evaluation dimensions from these websites serve as the content source. Following the approach of scholars like Li and Bi [4,5,15,16], we use these evaluation dimensions as hotel attributes.

Furthermore, existing studies have demonstrated that hotel attributes exert heterogeneous effects on traveler satisfaction. Scholars have adopted classic theoretical frameworks, including the three-factor theory and importance-performance analysis, to classify and rank hotel attributes, confirming that various attributes impose asymmetric influences on travelers' perceptual judgments and overall experience [15,17,18]. Table 1 presents typical studies on hotel attributes. As fundamental components determining travel experience, core attributes vary substantially across traveler groups, with traveler type serving as a decisive factor in identifying and classifying core attribute preferences. Rhee and Yang [6] concluded that business and solo travelers focus on sleep quality, family and friend travelers emphasize value and couples prioritize room. Wang et al. [7] noted that business travelers prioritize rooms, service, desks, lobbies, and restaurants, while couples focus on bars, beds, and floors. Families value ample breakfast options and spacious rooms, friends prioritize location and cleanliness, and solo travelers emphasize location. Based on a large-scale empirical analysis of 136,413 valid review data retrieved from Booking.com, Athanasopoulou et al. [19] further validated that the salience and influence of hotel attributes differ significantly depending on traveler types and hotel star levels. Ervianty et al. (2024) further demonstrated that for family-friendly hotels, cleanliness is the most critical factor [20].

Table 1. Research on hotel attributes.

Literature	Data Source	Traveler Type	Method	Identified Core Attributes	Key Findings
Rhee & Yang [6]	TripAdvisor	Business, Couples, Family, Friends, Solo	Conjoint analysis	Overall: value, room; Business/Solo: sleep quality; Family/Friends: value; Couples: room	Attribute importance varies considerably across trip purposes
Wang et al. [7]	TripAdvisor	Business, Couples, Family, Friends, Solo	TF-IDF + Word2Vec + Picture Fuzzy TODIM	Business: room/service/desk, etc.; Couples: bar/bed/floor; Family: breakfast/spacious room; Friends: location/cleanliness; Solo: location	Key factors and selection outcomes differ across traveler types
Athanasopoulou et al. [19]	Booking.com	Solo, Groups, Couples, Family	LDA + Three-factor theory	Location, Staff, Comfort, Facilities, Cleanliness, Room, Price (effects vary by traveler type and star rating)	Attribute effects are contingent on both traveler type and hotel star rating
Ervianty et al. [20]	TripAdvisor	Family	Cluster analysis + MCDM-VIKOR	Cleanliness is the most critical factor	Attribute priority varies across different clusters

In summary, the existing literature fails to form a unified consensus regarding the identification of core hotel attributes across diverse traveler groups. To fill this research gap, this study integrates multiple natural language processing techniques, including the Word2vec algorithm and word frequency statistics, combined with the TOPSIS method, to systematically analyze massive traveler online reviews and quantify the relative importance of various hotel attributes. This will help identify core attributes for different traveler groups.

2.2. The Relationship Between Hotel Attributes

Travelers are exposed to multiple hotel attributes during their stay, and these attributes jointly shape their experience. For a business traveler under work pressure, a well-located hotel with exceptional service can reduce stress and enhance their overall impression of the hotel. This illustrates that hotel attributes are interconnected rather than independent. Doğan et al. [21] explored the relationship between hotel attributes and price-performance (PP), and found that individual hotel attributes do not directly impact PP, but the combined effect of two or more attributes does. Some attributes may not affect PP initially but become significant when other attributes meet certain criteria. Although their study focuses on attributes affecting the cost-effectiveness of hotels, it also reflects the potential interrelationship between these attributes.

In consumer psychology research, the halo effect is frequently discussed. The halo effect, first proposed by Thorndike [22], is a cognitive bias that refers to the subjective perceptions formed by generalizing overall judgments from partial cues. Travelers' hotel experiences result from weighing multiple attributes, and their satisfaction with one attribute may stem from their high satisfaction with another attribute or even from the positive halo effect generated by overall satisfaction [5]. This phenomenon has been documented in existing scholarship. Mellinas et al. [4] found that for hotels with similar locations, location attribute assessments are influenced by other attributes, especially negative ratings. Nicolau et al. [5] examined the attribute halo effect on location perception

by tracking long-term rating changes of individual hotels. This highlights that travelers' evaluations are not purely rational, and assessments of certain attributes can affect those of others.

Based on previous studies, we further argue that diverse preferences across traveler groups allow any attribute to become a core attribute for specific travelers. Core attributes exert a greater impact on travelers' accommodation experience, and may shape evaluations of other logically distinct attributes—a pattern consistent with the halo effect—thereby influencing overall hotel impressions.

2.3. Hotel Attributes and Recommendation Model

Early studies relied on single ratings to understand traveler preferences [23], but this method fails to capture their full range of preferences. Recent research adopts multi-criteria ratings and text reviews to explore traveler preferences more comprehensively. Multi-criteria ratings cover aspects including facilities, services, and cleanliness, while text reviews uncover travelers' emotional responses toward different attributes.

In terms of multi-criteria rating-based hotel recommendation research, Kwok et al. [24] introduced the Vague Set TOPSIS algorithm by incorporating factors such as location, reputation, and room quality. Shambour et al. [25] selected seven hotel attribute ratings, including value for money, hotel location, room quality, overall service quality, check-in quality, and specialized business services, and developed a novel fusion-based multi-criteria collaborative filtering model to offer more personalized hotel recommendations. Le et al. [26] collected rating criteria including value, location, cleanliness, service ratings, and overall rating, integrated matrix factorization with deep learning to predict multi-criteria ratings, and applied evidential reasoning and Dempster's rule for rating aggregation.

For hotel recommendation studies that further explore attribute characteristics from text reviews, Nie et al. [27] analyzed review data to evaluate hotel attributes and used evidence theory for comprehensive hotel ranking. Wu et al. [28] employed the RFMP model to extract reviews from opinion leaders, applied Word2vec for attribute extraction, and determined the objective weight from opinion leaders' reviews using Word2Vec and the subjective weight using the best-worst method (BWM). These two types of weights were combined to calculate attribute importance, and the TOPSIS method was adopted for hotel ranking. Tao et al. [29] proposed a hotel selection model that considers traveler types and attribute interdependence.

Overall, hotel attributes play a vital role in constructing recommendation models. Previous studies combined multi-standard ratings and textual review attributes to capture traveler preferences, while generally assuming independent relationships among hotel attributes [24–27]. Although several recent studies have accounted for attribute correlations [7,28,29], the internal mechanism of such interdependencies remains underexplored. In particular, existing studies have not explicitly identified the psychological mechanism (e.g., the halo effect) driving attribute interactions, nor has it incorporated such causal mechanisms into collaborative filtering recommendation frameworks. This study targets the exploration of interactive attribute relationships, explicitly models the influence mechanism of core attributes on other attributes, and develops a recommendation model incorporating traveler characteristics, attribute scores, and supplementary data to improve hotel decision-making. Specifically, the novelty of this study lies in three aspects: (1) identifying core attributes for different traveler groups and verifying their directional influence on other attributes through regression analysis, which aligns with the halo effect theory; (2) correcting the user-hotel rating matrix based on verified attribute correlations; (3) embedding the corrected matrix into the

collaborative filtering framework to generate fine-grained personalized hotel recommendations.

3. Research Design of Attribute Correlations

To explore the interactive relationships among hotel attributes, this study collected and preprocessed online reviews from 320 three-star and above hotels in London on TripAdvisor. Core attributes were identified by integrating Word2vec analysis, attribute word frequency statistics and the TOPSIS method. The influences of core attributes on other attributes were examined using the SentiWordNet sentiment lexicon, PolarityRank algorithm and OLS regression model. Word frequency reflects travelers' attention salience, namely the frequency of attribute mentions, while sentiment values represent their evaluative valence, namely the degree of traveler satisfaction. The TOPSIS method identifies core attributes based on attention salience, and regression analysis verifies whether the sentiment of these highly salient attributes affects travelers' evaluations of other attributes. This observed influential pattern is consistent with the halo effect. Figure 1 presents the overall research framework of this study.

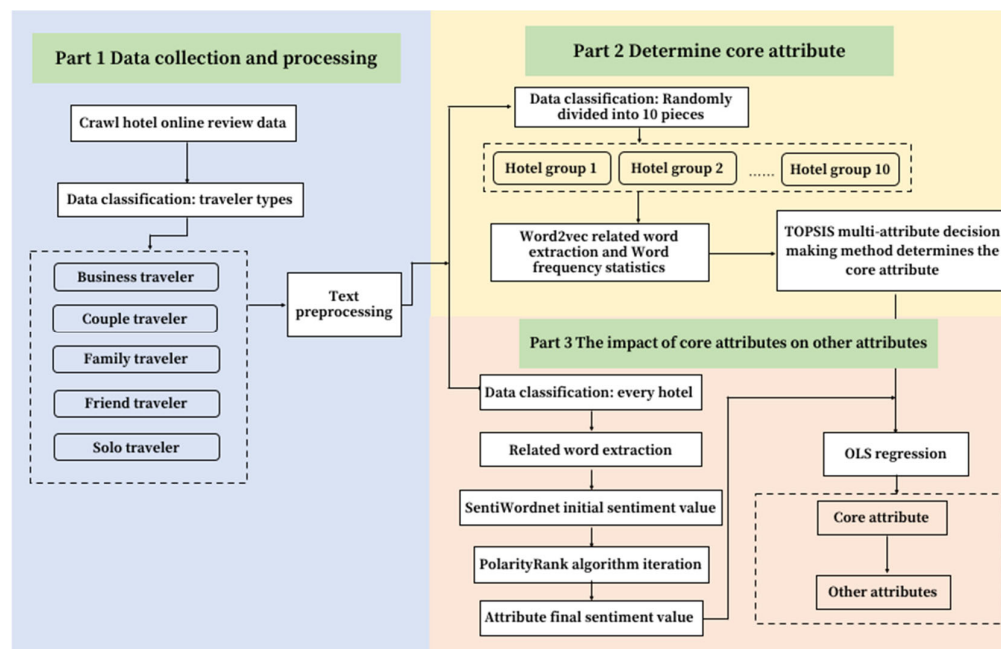


Figure 1. Research design framework (attribute-correlation analysis).

3.1. Data Collection and Processing

Data collection was conducted in May 2023 by crawling online reviews of London hotels from TripAdvisor.com, covering reviews posted by 30 April 2023. We then implemented the following filtering criteria: (1) removing non-English reviews; (2) eliminating duplicate reviews submitted by the same user for the same hotel; (3) discarding reviews with empty textual content; (4) excluding reviews with fewer than 10 characters; and (5) filtering out reviews from hotels with fewer than three stars. After strict filtering, a total of 885,490 valid reviews from 320 three-star and above hotels were retained for subsequent analysis. The collected hotel information includes hotel name, star rating, overall rating, review volume and address. The online review data contains reviewer ID, reviewer location, traveler type, overall rating, textual review, and scores for six core attributes: value, location, service, room quality, cleanliness and sleep quality.

Travelers' hotel preferences and attribute evaluation experiences during accommodation are significantly shaped by their traveler types [6,7]. In the hotel tourism context, five dominant traveler categories are identified: business, couple, family, friend, and solo travelers. Accordingly, this study divided the cleaned dataset into five subsets based on traveler type and conducted standardized text preprocessing. The preprocessing pipeline included text lowercasing, spelling error correction, removal of non-English characters, numbers and punctuation marks, word segmentation, and stop-word elimination.

3.2. Attribute-Related Words Extraction and Frequency Statistics

Following the official attribute classification of TripAdvisor, this study adopts six core hotel evaluation dimensions: value, location, service, room, cleanliness, and sleep quality, which constitute the primary indicators for travelers' hotel assessments. Nevertheless, travelers rarely use these exact dimension terms in their review texts. For example, travelers often employ alternative expressions such as position, scenery and located to describe the location attribute. To address this semantic discrepancy, this study applies the Word2Vec model to calculate word semantic similarity and group semantically synonymous terms under corresponding attribute dimensions.

Word2Vec is a classic word embedding algorithm in natural language processing (NLP) that effectively captures syntactic and semantic information by mapping similar words to adjacent vector spaces [30]. It supports two training architectures: Continuous Bag of Words (CBOW) and Skip-Gram. The CBOW architecture predicts target words from surrounding context words and performs better on small datasets, whereas the Skip-Gram architecture infers contextual words from a given target word and exhibits superior performance on large corpora [31]. Given the large sample size of this study and the predefined six core attribute keywords, the Skip-Gram algorithm was selected for model training.

Before training the Word2Vec model, raw review texts were systematically preprocessed through lowercasing, spelling correction, elimination of non-English characters and digits, punctuation removal, tokenization, and deletion of NLTK English stop words. Model hyperparameters were set as follows: vector dimensionality of 100, window size of 5, minimum word frequency threshold of 5, and 10 training epochs. The negative sampling size was set to the default value of 5, and a fixed random seed was adopted to guarantee experimental reproducibility. For each traveler group, the review dataset was randomly split into 10 subsets for independent Word2Vec training. Each trained model served as an independent evaluation reference for the subsequent TOPSIS analysis. This bootstrapping-based multi-criteria design reduces biases caused by single-split randomness and substantially improves the reliability and generalizability of core attribute identification results.

This study extracted attribute-related terms by retaining words with a semantic similarity score above 0.5 [32]. The overall frequency of each attribute was aggregated from the frequencies of all its matched related terms. Word mention frequency reflects travelers' active attention toward specific hotel attributes. Such spontaneous textual mentions indicate strong psychological salience in consumption experience, which can serve as a valid proxy for attribute importance [33]. This measurement logic is supported by multiple theoretical foundations. First, the availability heuristic suggests that individuals assess attribute importance based on cognitive accessibility; frequent spontaneous mentions in reviews represent high accessibility and perceived importance [34]. Second, usage-based linguistics posits that linguistic frequency determines cognitive salience, whereby frequently encountered terms are strongly entrenched in memory and easily activated [35]. Third, cue utilization theory states that consumers rely on salient and

identifiable cues for product evaluation [36]; frequently mentioned attributes function as such critical cues in online review scenarios. Consistent with these theories, word frequency is widely recognized as a reliable indicator of term importance in text mining research [37]. Therefore, this study interprets highly frequent attribute mentions as indicators of stronger psychological salience and greater practical importance for travelers.

3.3. Determine Core Attributes

TOPSIS is a multi-attribute decision-making method [38]. The m attributes of n candidate networks in TOPSIS can be represented as a matrix D ($n \times m$), where $A_1, A_2, \dots, A_m$ indicates alternatives, $C_1, C_2, \dots, C_n$ represents the criterion for measuring the alternative, x_{mn} represents the score of the standard C_n in the alternative A_m . W is the weight matrix, and w_n is the weight of the standard C_n . In this study, the alternatives are six hotel attributes (value, location, service, room, cleanliness, sleep quality), as we aim to rank these attributes to identify the core attribute for each traveler type. The evaluation criteria are ten randomly divided review subsets; x_{mn} represents the word frequency of the m -th attribute in the n -th subset; equal weights are adopted with $w_n = 1/10$.

Among MCDM methods, TOPSIS selects the option closest to the positive ideal solution and farthest from the negative ideal solution. The positive ideal solution maximizes benefits and minimizes costs, while the negative ideal solution does the opposite. This method is widely used in the supply chain and logistics management, business and marketing management and other fields. TOPSIS makes full use of attribute information, provides a basic ranking of alternatives, and does not require independence between attributes [39,40]. Compared with other multi-attribute decision-making methods (such as analytic hierarchy process and network analysis), TOPSIS technology is more suitable for this study. The TOPSIS decision matrix is presented as follows:

$$D = \begin{matrix} & C_1 & C_2 & \dots & C_n \\ \begin{matrix} A_1 \\ A_2 \\ \vdots \\ A_m \end{matrix} & \begin{bmatrix} x_{11} & x_{12} & \dots & x_{1n} \\ x_{21} & x_{22} & \dots & x_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ x_{m1} & x_{m2} & \dots & x_{mn} \end{bmatrix} \end{matrix}$$

Taking business travelers as an example, Table 2 shows the data of TOPSIS decision word frequency.

Table 2. TOPSIS decision word frequency of business travelers.

Attribute	C_1	C_2	C_3	C_4	C_5	C_6	C_7	C_8	C_9	C_{10}
value	6254	10,851	8891	945	7601	8635	13,738	2708	6928	1368
location	24,035	38,395	31,591	21,455	34,762	34,327	51,318	32,635	35,758	39,592
service	2999	11,430	10,808	11,712	6999	13,995	29,026	11,510	3393	5279
room	23,846	35,149	23,115	23,952	38,237	32,238	27,075	30,167	27,142	21,959
cleanliness	22,953	43,975	28,791	33,367	37,418	38,435	45,140	11,564	31,144	33,905
sleep quality	5841	5555	1460	5695	5892	5334	7738	5187	1754	6069

According to the TOPSIS calculation procedure [41], researchers first standardize word frequency data and then construct the normalized fuzzy decision matrix. Afterwards, the positive and negative ideal solutions are determined. Researchers further calculate the distances to the two ideal solutions and the relative closeness degree of each alternative, and finally select the attribute with the highest relative closeness value as the core attribute. In this study, the core attribute is defined as the attribute that obtains the highest relative closeness value after calculating the frequencies of attribute-related words via the TOPSIS method for a specific traveler group. The specific calculation formulas are presented as follows.

Step 1: Standardize word frequency $r_{ij} = \frac{x_{ij}}{\sqrt{\sum_{i=1}^m x_{ij}^2}}$.

Step 2: Construct the normalized fuzzy decision matrix $V = [v_{ij}]_{m \times n}$, $v_{ij} = w_j \times r_{ij}$.

Step 3: Identify ideal and negative-ideal solutions $A^+ = [v_1^+, \dots, v_n^+]$, $A^- = [v_1^-, \dots, v_n^-]$.

Step 4: Calculate the distance for both ideal and negative-ideal solutions $D_i^+ = \sqrt{\sum_{j=1}^n (v_j^+ - v_{ij})^2}$, $D_i^- = \sqrt{\sum_{j=1}^n (v_j^- - v_{ij})^2}$.

Step 5: Calculate the relative closeness to the ideal solution $C_{TOPSIS} = \frac{(D_i^-)}{(D_i^- + D_i^+)}$.

3.4. Sentiment Analysis

Sentiment analysis determines the emotional tendency of text using supervised or unsupervised methods. This study opted for the unsupervised dictionary-based approach due to its efficiency and maturity, using the SentiWordNet sentiment dictionary to conduct sentiment analysis, which assigns positive (PosScore) and negative sentiment values (NegScore) to words based on their parts of speech. Given the extensive review sample size and varying parts of speech in reviews, this study retained nouns, verbs, adverbs, and adjectives. For each retained word, if multiple synsets were available in SentiWordNet, we calculated the arithmetic mean of the PosScore and NegScore values across all synsets and used these averages as the word's initial positive sentiment value e_i^+ and the initial negative sentiment value e_i^- of the word.

To improve the dictionary's versatility, sentiment scores were adjusted using the PolarityRank algorithm. Montejo-Ráez et al. proposed a random walk weighting method over SentiWordNet for sentiment polarity detection on Twitter, showing that this unsupervised technique is "promising" for short-text sentiment classification [42]. Bhatnagar et al. proposed an aspect-based opinion mining framework for the tourism sector using SentiWordNet, extracting explicit aspects from TripAdvisor hotel reviews and assigning sentiment scores based on SentiWordNet, further demonstrating the feasibility of this approach for hotel review sentiment analysis [43]. Chen et al. directly employed PolarityRank on TripAdvisor hotel reviews to construct a sentiment lexicon for the hospitality domain, confirming its effectiveness on hotel review texts [44].

This algorithm assigned positive and negative sentiment values to each word, creating a graph where emotions propagate through nodes via PangRank's correlation calculations. Specifically speaking, the SentiWordNet dictionary assigned initial positive and negative values to each word. If a word is not in the dictionary, both values were set to 0. Then a directed graph was built based on a syntactic tree, connecting nodes with edges. The graph $G = (V, E)$ consists of nodes V and edges E , where each node represents a word in the review, edges are constructed based on syntactic dependency relations, and each edge has a non-zero weight w_{ij} representing the connection strength between nodes. In this study, the node similarity was used as the weight, with d as the iterative parameter set to 0.85 to ensure the convergence of PR^+ and PR^- functions. Compared to conventional dictionary-based sentiment analysis, PolarityRank offers two key improvements: (1) contextual propagation through the syntactic graph allows semantically related words to influence each other's sentiment values; (2) out-of-vocabulary words can receive sentiment values via propagation from neighboring nodes, reducing information loss. The algorithm iterated the sentiment scores until they reached a fixed point or a specified threshold. According to most studies and the actual situation of this study, the convergence condition was set as the number of iterations being 200 or the affective value difference between the next iteration and the previous iteration being less than or equal to 10^{-5} . Finally, we obtained the PR^+ and PR^- after each node iteration. The calculation formula is as follows:

$$PR^+(v_i) = (1-d)e_i^+ + d \left(\sum_{j \in E(v_i)} \frac{w_{ij}}{\sum_{k \in E(v_j)} |w_{jk}|} PR^+(v_j) \right)$$

$$PR^-(v_i) = (1-d)e_i^- + d \left(\sum_{j \in E(v_i)} \frac{w_{ij}}{\sum_{k \in E(v_j)} |w_{jk}|} PR^-(v_j) \right)$$

Since the purpose of this study is to convert text reviews into digital data and apply them to the regression model to explore whether travelers' emotional experience of core attributes will affect other attributes, both PR^+ and PR^- values can achieve this purpose. This study selects the PR^+ value to reflect the travelers' emotional experience. In the interval of $[0, 1]$, the larger the value, the stronger the emotion and the better the experience.

3.5. Construction of Regression Model

In this study, OLS regression analysis was performed on standardized sentiment values at the hotel level. The core attributes identified via TOPSIS serve as independent variables, while other attributes are set as dependent variables. Previous studies have confirmed that hotel star rating, ranking, public praise, green practices and surrounding environments can affect travelers' perceptions of hotel attributes [45–47], and review volume can interfere with travelers' rating tendencies [48]. In addition, overall hotel ratings comprehensively reflect travelers' evaluations of various attributes, and controlling for overall ratings helps eliminate confounding effects from other factors.

Therefore, this study adopts hotel star rating, hotel ranking, overall rating, hotel honors, review volume, the number of nearby restaurants, nearby scenic spots, traffic conditions and eco-friendliness as control variables. All control variables are retrieved from TripAdvisor. Hotel honors, traffic conditions and eco-friendliness are defined as dummy variables, assigned a value of 1 if the hotel has the corresponding condition and 0 otherwise. The regression equation is shown below.

$$c_j = \text{Constant} + \beta \text{Core_Attribute} + \delta \text{Control} + \varepsilon$$

where c_j is the other attributes, $j = 1, 2, \dots, 5$.

4. Result of Attribute Correlation

4.1. Result of Attribute-Related Words and Word Frequency

According to the previous description, this study used Word2Vec algorithm to extract the relevant words of six hotel attributes from the online reviews of five different types of travelers and counts the word frequency. Taking business travelers as an example, Table 3 shows the partial results.

Table 3. Attribute-related words of business travelers.

Value	Similarity	Frequent	Location	Similarity	Frequent	Service	Similarity	Frequent
represent	0.638837	32	riverside	0.619363	11	competent	0.609132	26
criterion	0.593156	6	closely	0.615396	14	professionalism	0.595891	35
overpriced	0.587698	25	situate	0.604744	220	attentiveness	0.594155	8
astronomical	0.53448	5	position	0.599979	168	satisfaction	0.587255	24
economical	0.533454	6	distance	0.579145	611	politeness	0.583503	7
equivalent	0.529812	14	central	0.57314	1538	exemplary	0.583214	29
extortionate	0.528309	16	spot	0.563325	208	respectful	0.575655	10
competitive	0.52543	38	locate	0.562888	1058	outstanding	0.571599	204
reasonable	0.514136	525	tube	0.553433	1690	exceptional	0.571268	221
worthwhile	0.513213	15	transport	0.513327	229	food	0.53478	2769
Room	Similarity	Frequent	Cleanliness	Similarity	Frequent	Sleep Quality	Similarity	Frequent
small	0.736122	8	necessity	0.660316	19	undisturbed	0.605741	5

bedroom	0.722844	419	neat	0.65925	57	buzzing	0.590099	5
sufficiently	0.688318	16	comfortable	0.641169	3294	awake	0.587166	89
adequately	0.6723	15	wonderfully	0.629327	27	soundly	0.578092	9
standard	0.644841	16	uncommon	0.621839	6	disturb	0.576136	119
windowless	0.639212	12	renovated	0.615702	7	soundproof	0.572821	44
furnished	0.638651	23	aesthetically	0.607802	7	peaceful	0.543069	67
moderately	0.638142	8	organized	0.594292	6	quilt	0.537765	7
oppressive	0.637878	5	decent	0.5138	681	silence	0.520221	10
dreamy	0.631787	8	refurbish	0.500081	195	mattress	0.508809	137

Value refers to travelers' evaluation of paid costs and acquired benefits, such as descriptions of being overpriced or economical. As for the location attribute, travelers frequently mention convenient transportation (transportation, subway) and surrounding environments (riverside, scenic spots, downtown areas). Service covers hotel staff attitudes (polite, respectful) and supplementary services such as catering services. Room-related descriptions mainly involve room space (spacious, narrow) and in-room facilities (windowless rooms, bedrooms, complete furnishings). Cleanliness reflects the tidiness and sanitary condition of hotel supplies (renovated, tidy). The sleep quality attribute involves quilt comfort, mattress softness and sound insulation effect (restless sleep, sound sleep, quiet surroundings). These elements constitute hotel product and service attributes and jointly influence travelers' accommodation experience.

4.2. Core Attributes of Different Traveler Types

The core attributes of the five traveler groups were identified via the TOPSIS analysis. Tables 4–8 present the corresponding results for business, couples, family, friends, and solo travelers. D represents the positive ideal solution distance, D- represents the negative ideal solution distance, and C indicates the relative proximity to the ideal solution. The attribute with the highest relative proximity is the core attribute for that traveler type. Results indicate that business travelers prioritize location, while couples, friends, and solo travelers prioritize room, and family travelers prioritize cleanliness.

Table 4. TOPSIS results of business travelers.

Item	D	D-	C	Rank
value	1.683	0.226	0.118	5
location	0.273	1.711	0.862	1
service	1.482	0.469	0.24	4
room	0.552	1.413	0.719	3
clean	0.471	1.625	0.775	2
sleep quality	1.744	0.154	0.081	6

Table 5. TOPSIS results of couple travelers.

Item	D	D-	C	Rank
value	1.032	0.654	0.388	4
location	0.473	1.11	0.701	3
service	1.455	0.095	0.061	6
room	0.1	1.465	0.936	1
clean	0.157	1.409	0.899	2
sleep quality	1.388	0.233	0.144	5

Table 6. TOPSIS results of family travelers.

Item	D	D-	C	Rank
value	0.835	0.836	0.5	4
location	0.341	1.271	0.789	3
service	1.533	0.054	0.034	6
room	0.237	1.4	0.855	2
clean	0.176	1.461	0.893	1
sleep quality	1.427	0.169	0.106	5

Table 7. TOPSIS results of friend travelers.

Item	D	D-	C	Rank
value	1.798	0.248	0.121	6
location	0.442	1.599	0.783	2
service	1.713	0.339	0.165	5
room	0.273	1.783	0.867	1
clean	0.519	1.564	0.751	3
sleep quality	1.522	0.46	0.232	4

Table 8. TOPSIS results of solo travelers.

Item	D	D-	C	Rank
value	1.774	0.228	0.114	6
location	0.364	1.607	0.815	2
service	1.684	0.276	0.141	4
room	0.198	1.795	0.901	1
clean	0.868	1.207	0.582	3
sleep quality	1.768	0.232	0.116	5

Business travelers refer to people who take trips for work purposes. They prioritize hotels located in business districts or areas with convenient transportation to save time, boost work efficiency and ease travel fatigue caused by tight schedules. Couples pursue romance and intimacy during trips and cherish private time spent in hotel rooms. They are more easily affected by sensory feelings than practical functions [49], and expect a warm and romantic atmosphere to strengthen mutual affection. Room comfort, lighting, atmosphere and interior decoration greatly influence their intimate experience.

Family travelers attach great importance to cleanliness out of consideration for family members' health and safety. They also pay much attention to the overall hotel impression. Cleanliness is essential to accommodation comfort and forms travelers' first impression of a hotel. Failure to meet family travelers' cleanliness expectations will weaken their overall evaluation and lower satisfaction levels. Groups of friends tend to choose spacious rooms suitable for group accommodation, which can provide enough large beds. They need extra space for luggage and daily activities, and also value in-room entertainment facilities such as televisions and gaming devices to enrich group leisure time. Solo travelers pay close attention to room attributes, hoping to obtain quiet and comfortable private space to stay away from external disturbances. A tranquil and private room environment can effectively improve their overall accommodation experience [50].

4.3. Result of Attribute Sentiment Analysis

Then SentiWordNet dictionary and PolarityRank algorithm were used to obtain the sentiment value of each hotel attribute. The higher the sentiment value, the higher the satisfaction of travelers. Taking one of the hotels (The Royal Horseguards Hotel) as an

example, Table 9 shows the top 10 factors of business travelers' sentiment value for the hotel attributes, where *simi* is the similarity between the factor and attribute, and *senti* is the sentiment value of the factor.

Table 9. Sentiment score of attribute-related words for business travelers.

Value	Simi	Senti	Location	Simi	Senti	Service	Simi	Senti
quality	0.577287	0.559852	central	0.629027	0.415307	helpful	0.510222	0.765962
cheap	0.556258	0.527666	station	0.555726	0.398024	food	0.537215	0.625400
worthwhile	0.517218	0.494978	modern	0.505000	0.391869	restaurant	0.518134	0.612444
choice	0.546376	0.383385	tube	0.541433	0.375761	superfriendly	0.580817	0.602633
overall	0.570542	0.348483	locate	0.650202	0.363025	professional	0.532019	0.430838
reasonable	0.616141	0.199347	convenient	0.584900	0.329633	courteous	0.509895	0.416870
average	0.502694	0.155294	distance	0.536350	0.261762	outstanding	0.605534	0.355646
expensive	0.541533	0.080329	city	0.500295	0.261603	hospitality	0.535744	0.297809
overprice	0.511874	0.062635	attraction	0.506980	0.249149	impeccable	0.535442	0.256733
budget	0.564085	0.053268	spot	0.603478	0.188026	lunch	0.517309	0.135958
Room	Simi	Senti	Cleanliness	Simi	Senti	Sleep Quality	Simi	Senti
comfortable	0.543955	0.918822	facility	0.509092	0.314405	quiet	0.553433	0.491846
bathroom	0.612665	0.906870	decent	0.519788	0.239270	sound	0.509699	0.255444
bed	0.628318	0.881726	beautifully	0.519775	0.191743	hear	0.525341	0.247885
view	0.506294	0.864418	decorate	0.551169	0.190695	peaceful	0.577938	0.193542
floor	0.583717	0.531510	stylish	0.621904	0.149039	silent	0.640427	0.191266
spacious	0.658690	0.421008	maintain	0.633776	0.139624	noisy	0.613058	0.188674
modern	0.553493	0.391869	equip	0.680631	0.139118	bother	0.557164	0.112116
size	0.547282	0.342521	neat	0.730434	0.106171	uncomfortable	0.506855	0.106264
amenity	0.545319	0.316956	refurbish	0.563559	0.119918	awake	0.610088	0.070748
upgrade	0.508889	0.259786	classy	0.503018	0.112448	neighbour	0.534201	0.044320

It can be seen that the high cost performance is conducive to improving the experience of travelers, convenient transportation and the location near the scenic spot can bring better experience to travelers, professional and hospitable staff, catering and other additional services are preferred by travelers. Broad view, exquisite room layout and good hygiene, undisturbed sleep environment can enhance the goodwill of travelers. Through the emotional analysis of text reviews of each hotel, this study obtains the emotional experience of different types of travelers on 6 attributes of each hotel. Attribute sentiment value is the sum of the similarity between factor and attribute times the sentiment of factors divided by the sum of the similarity. The formula can be expressed as:

$$total_senti = \frac{\sum(sim_i \times senti_i)}{\sum sim_i}$$

where *total_senti* is the attribute sentiment value, *simi* is the similarity between attribute-related words and attributes, and *senti* is the sentiment value of attribute-related words. Using business travelers as an example, Table 10 shows the sentiment value of business travelers for each attribute in the top 10 hotels.

Table 10. Attribute sentiment scores of business travelers.

Hotel	Value	Location	Service	Room	Cleanliness	Sleep Quality
1	0.084889	0.105930	0.115287	0.102952	0.102748	0.067075
2	0.117199	0.127451	0.124550	0.098693	0.120944	0.081748
3	0.074844	0.101117	0.097460	0.078600	0.108177	0.054767
4	0.123690	0.110749	0.097665	0.072813	0.106463	0.086594

5	0.115644	0.108356	0.101733	0.080344	0.113562	0.072890
6	0.116513	0.104222	0.104323	0.074607	0.105547	0.068819
7	0.086799	0.098811	0.077393	0.103105	0.082552	0.076477
8	0.067941	0.077997	0.110003	0.108325	0.112054	0.088328
9	0.091794	0.093873	0.095529	0.074493	0.102919	0.055144
10	0.097865	0.107993	0.110651	0.110934	0.102057	0.104610

4.4. Result of Regression Analysis

Prior to regression analysis, we conducted a variance inflation factor (VIF) test to check for multicollinearity. The results showed that all variables had tolerances greater than 0.1 and VIF values below 10, indicating no significant multicollinearity. Subsequently, the Durbin-Wu-Hausman (DWH) test was performed, and the results (presented in Tables 11–15) confirmed the absence of endogeneity issues in the model. Following these tests, OLS regression analysis was conducted, with the results presented in Tables 16–20.

Table 11. DWH test for business travelers.

Test	Null Hypothesis	Test Result	Conclusion
Durbin test	Explanatory variables are exogenous (no endogeneity)	$\chi^2(1) = 0.277, p = 0.599$	Fail to reject the null hypothesis
Wu-Hausman test	Explanatory variables are exogenous (no endogeneity)	$F(1305) = 0.267, p = 0.606$	Fail to reject the null hypothesis

Table 12. DWH test for couple travelers.

Test	Null Hypothesis	Test Result	Conclusion
Durbin test	Explanatory variables are exogenous (no endogeneity)	$\chi^2(1) = 2.095, p = 0.148$	Fail to reject the null hypothesis
Wu-Hausman test	Explanatory variables are exogenous (no endogeneity)	$F(1304) = 2.016, p = 0.157$	Fail to reject the null hypothesis

Table 13. DWH test for family travelers.

Test	Null Hypothesis	Test Result	Conclusion
Durbin test	Explanatory variables are exogenous (no endogeneity)	$\chi^2(1) = 2.907, p = 0.088$	Fail to reject the null hypothesis
Wu-Hausman test	Explanatory variables are exogenous (no endogeneity)	$F(1303) = 2.804, p = 0.095$	Fail to reject the null hypothesis

Table 14. DWH test for friend travelers.

Test	Null Hypothesis	Test Result	Conclusion
Durbin test	Explanatory variables are exogenous (no endogeneity)	$\chi^2(1) = 0.372, p = 0.542$	Fail to reject the null hypothesis
Wu-Hausman test	Explanatory variables are exogenous (no endogeneity)	$F(1303) = 0.356, p = 0.551$	Fail to reject the null hypothesis

Table 15. DWH test for solo travelers.

Test	Null Hypothesis	Test Result	Conclusion
Durbin test	Explanatory variables are exogenous (no endogeneity)	$\chi^2(1) = 1.394, p = 0.238$	Fail to reject the null hypothesis
Wu-Hausman test	Explanatory variables are exogenous (no endogeneity)	$F(1300) = 1.338, p = 0.248$	Fail to reject the null hypothesis

Table 16. OLS regression results for business travelers.

	(1)	(2)	(3)	(4)	(5)
	Value	Service	Room	Sleep	Cleanliness
location	0.227 *** (0.055)	0.27 *** (0.052)	−0.158 *** (0.048)	0.012 (0.073)	0.292 *** (0.052)
value		0.169 *** (0.055)	−0.262 *** (0.047)	0.236 *** (0.074)	0.343 *** (0.053)
service	0.181 *** (0.059)		0.183 *** (0.05)	0.052 (0.077)	0.056 (0.058)
room	−0.356 *** (0.064)	0.232 *** (0.063)		0.728 *** (0.077)	0.483 *** (0.059)
sleep quality	0.14 *** (0.044)	0.029 (0.043)	0.317 *** (0.033)		−0.006 (0.043)
cleanliness	0.361 *** (0.055)	0.055 (0.057)	0.374 *** (0.046)	−0.01 (0.077)	
star	−0.303 *** (0.075)	0.547 *** (0.067)	−0.007 (0.066)	−0.456 *** (0.096)	−0.059 (0.075)
rank	0.051 (0.061)	−0.238 *** (0.058)	−0.007 (0.053)	−0.003 (0.08)	0.107 * (0.059)
overall rating	−0.209 (0.13)	−0.109 (0.126)	−0.057 (0.112)	0.25 (0.17)	0.506 *** (0.124)
review number	0.125 ** (0.053)	0.045 (0.051)	−0.328 *** (0.042)	0.065 (0.069)	0.1 * (0.052)
hotel honor	−0.191 * (0.113)	0.086 (0.11)	0.082 (0.097)	−0.037 (0.148)	−0.127 (0.111)
nearby res	−0.016 (0.079)	0.039 (0.077)	−0.167 ** (0.067)	−0.02 (0.103)	0.043 (0.077)
nearby spot	−0.029 (0.068)	0.025 (0.066)	0.12 ** (0.058)	−0.057 (0.089)	−0.114 * (0.066)
traffic	0.069 (0.159)	−0.427 *** (0.152)	0.207 (0.136)	0.275 (0.206)	−0.1 (0.155)
Eco friendly	−0.394 *** (0.132)	0.049 (0.129)	−0.239 ** (0.114)	0.063 (0.174)	0.093 (0.13)
_cons	2.088 *** (0.668)	−0.787 (0.655)	1.55 *** (0.575)	−1.361 (0.879)	−2.212 *** (0.65)
Observations	317	317	317	317	317
R-squared	0.492	0.525	0.626	0.432	0.517

Standard errors are in parentheses; *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.**Table 17.** OLS regression results for couple travelers.

	(1)	(2)	(3)	(4)	(5)
	Value	Location	Service	Cleanliness	Sleep Quality
room	−0.521 *** (0.069)	−0.012 (0.069)	0.175 ** (0.079)	0.499 *** (0.051)	0.274 *** (0.083)
value		0.101 * (0.053)	0.193 *** (0.06)	0.274 *** (0.042)	0.126 * (0.064)
location	0.119 * (0.062)		0.388 *** (0.062)	0.259 *** (0.046)	0.019 (0.07)
service	0.171 *** (0.053)	0.294 *** (0.047)		0.087 ** (0.042)	−0.072 (0.061)
cleanliness	0.456 *** (0.069)	0.367 *** (0.065)	0.163 ** (0.078)		0.181 ** (0.083)

sleep quality	0.099 *	0.013	−0.064	0.085 **	
	(0.051)	(0.047)	(0.054)	(0.039)	
star	−0.587 ***	−0.074	0.505 ***	0.139 **	−0.158 *
	(0.072)	(0.074)	(0.08)	(0.061)	(0.09)
rank	0.152 **	−0.024	−0.083	0.077 *	0.103
	(0.06)	(0.056)	(0.064)	(0.046)	(0.068)
overall rating	−0.127	−0.197 *	0.169	0.28 ***	0.24 *
	(0.124)	(0.114)	(0.131)	(0.094)	(0.139)
review number	−0.116 *	−0.26 ***	0.069	0.026	0.03
	(0.061)	(0.054)	(0.065)	(0.047)	(0.069)
hotel honor	0.028	−0.01	0.088	−0.123	0.044
	(0.109)	(0.101)	(0.116)	(0.084)	(0.123)
nearby res	0.04	0.022	0.059	−0.153 ***	0.122
	(0.076)	(0.07)	(0.081)	(0.058)	(0.086)
nearby spot	−0.07	0.13 **	0.012	0.007	−0.022
	(0.065)	(0.059)	(0.069)	(0.05)	(0.073)
traffic	−0.228	−0.084	−0.069	0.323 ***	0.011
	(0.151)	(0.14)	(0.161)	(0.116)	(0.172)
Eco friendly	−0.045	0.108	−0.175	0.001	0.16
	(0.127)	(0.117)	(0.135)	(0.099)	(0.143)
_cons	2.924 ***	1.479 **	−2.778 ***	−1.678 ***	−2.797 ***
	(0.688)	(0.648)	(0.733)	(0.54)	(0.782)
Observations	318	318	318	318	318
R-squared	0.528	0.598	0.47	0.717	0.299

Standard errors are in parentheses; *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table 18. OLS regression results for family travelers.

	(1)	(2)	(3)	(4)	(5)
	Value	Location	Service	Room	Sleep Quality
cleanliness	0.296 ***	0.301 ***	0.221 ***	0.317 ***	0.081
	(0.054)	(0.061)	(0.058)	(0.046)	(0.062)
value		0.311 ***	0.216 ***	−0.312 ***	0.277 ***
		(0.062)	(0.059)	(0.047)	(0.061)
location	0.248 ***		0.173 ***	−0.174 ***	−0.02
	(0.049)		(0.053)	(0.044)	(0.056)
service	0.196 ***	0.197 ***		0.193 ***	−0.08
	(0.054)	(0.06)		(0.047)	(0.06)
room	−0.402 ***	−0.281 ***	0.275 ***		0.632 ***
	(0.061)	(0.071)	(0.067)		(0.062)
sleep quality	0.231 ***	−0.02	−0.074	0.41 ***	
	(0.051)	(0.059)	(0.055)	(0.04)	
star	−0.298 ***	0.236 ***	0.486 ***	0.062	−0.271 ***
	(0.069)	(0.078)	(0.069)	(0.062)	(0.076)
rank	0.007	0.032	−0.135 **	0.021	0.006
	(0.056)	(0.063)	(0.058)	(0.049)	(0.061)
overall rating	−0.139	−0.209	0.095	−0.356 ***	0.053
	(0.12)	(0.134)	(0.126)	(0.104)	(0.131)
review number	0.174 ***	−0.114 **	−0.006	−0.236 ***	−0.117 **
	(0.05)	(0.057)	(0.054)	(0.043)	(0.056)
hotel honor	−0.225 **	0.043	0.175	0.011	0.073
	(0.109)	(0.123)	(0.115)	(0.097)	(0.12)
nearby res	−0.081	0.126 *	−0.095	−0.041	0.031

	(0.067)	(0.075)	(0.071)	(0.06)	(0.074)
nearby spot	−0.025	0.11	0.141 **	0.046	0.035
	(0.064)	(0.072)	(0.067)	(0.057)	(0.071)
traffic	−0.077	−0.052	−0.086	−0.089	0.108
	(0.087)	(0.098)	(0.092)	(0.077)	(0.095)
Eco friendly	−0.153	0.007	0.065	−0.127	0.221
	(0.125)	(0.141)	(0.132)	(0.11)	(0.137)
_cons	2.488 ***	−0.22	−1.607 **	2.746 ***	−1.283
	(0.708)	(0.808)	(0.753)	(0.616)	(0.787)
Observations	317	317	317	317	317
R-squared	0.541	0.425	0.494	0.643	0.501

Standard errors are in parentheses; *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table 19. OLS regression results for friend travelers.

	(1)	(2)	(3)	(4)	(5)
	Value	Location	Service	Cleanliness	Sleep Quality
room	−0.353 ***	−0.334 ***	0.248 ***	0.64 ***	0.636 ***
	(0.07)	(0.078)	(0.069)	(0.053)	(0.068)
value		0.224 ***	0.226 ***	0.252 ***	0.411 ***
		(0.062)	(0.055)	(0.049)	(0.056)
location	0.184 ***		0.19 ***	0.174 ***	0.142 **
	(0.051)		(0.05)	(0.045)	(0.055)
service	0.239 ***	0.243 ***		0.001	−0.019
	(0.057)	(0.064)		(0.052)	(0.063)
cleanliness	0.321 ***	0.27 ***	0.001		−0.012
	(0.062)	(0.07)	(0.063)		(0.069)
sleep quality	0.364 ***	0.153 **	−0.016	−0.008	
	(0.05)	(0.059)	(0.053)	(0.048)	
star	−0.366 ***	0.385 ***	0.578 ***	0.019	−0.239 ***
	(0.077)	(0.085)	(0.07)	(0.07)	(0.083)
rank	0.032	−0.082	−0.06	0.155 ***	−0.02
	(0.058)	(0.064)	(0.057)	(0.051)	(0.062)
overall rating	−0.152	−0.323 **	0.311 **	0.231 **	0.285 **
	(0.126)	(0.138)	(0.122)	(0.111)	(0.133)
review number	0.011	−0.016	0.018	0.27 ***	−0.248 ***
	(0.054)	(0.059)	(0.053)	(0.045)	(0.056)
hotel honor	−0.103	0.058	−0.001	0.064	−0.177
	(0.112)	(0.124)	(0.11)	(0.1)	(0.119)
nearby res	−0.019	−0.058	−0.03	−0.114 *	0.115
	(0.071)	(0.078)	(0.069)	(0.062)	(0.075)
nearby spot	0.052	0.043	0.112 *	−0.066	−0.033
	(0.068)	(0.074)	(0.066)	(0.06)	(0.072)
traffic	−0.044	0.115	−0.07	−0.024	0.008
	(0.091)	(0.1)	(0.088)	(0.08)	(0.096)
Eco friendly	−0.196	0.044	−0.174	0.07	0.114
	(0.132)	(0.146)	(0.128)	(0.117)	(0.14)
_cons	2.591 ***	−0.196	−3.131 ***	−1.524 **	−1.101
	(0.732)	(0.824)	(0.705)	(0.656)	(0.791)
Observations	317	317	317	317	317
R-squared	0.496	0.387	0.521	0.604	0.615

Standard errors are in parentheses; *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table 20. OLS regression results for solo travelers.

	(1)	(2)	(3)	(5)	(6)
	Value	Location	Service	Cleanliness	Sleep Quality
room	−0.325 *** (0.056)	−0.274 *** (0.071)	0.467 *** (0.07)	0.652 *** (0.05)	0.822 *** (0.07)
value		0.584 *** (0.063)	0.389 *** (0.07)	0.494 *** (0.055)	0.314 *** (0.082)
location	0.381 *** (0.041)		0.133 ** (0.059)	0.08 (0.05)	0.291 *** (0.065)
service	0.241 *** (0.043)	0.126 ** (0.056)		−0.044 (0.049)	−0.047 (0.066)
cleanliness	0.433 *** (0.048)	0.107 (0.067)	−0.062 (0.069)		−0.19 ** (0.078)
sleep quality	0.15 *** (0.039)	0.213 *** (0.048)	−0.037 (0.051)	−0.103 ** (0.042)	
star	−0.156 *** (0.049)	0.179 *** (0.061)	0.318 *** (0.061)	−0.121 ** (0.053)	−0.06 (0.072)
rank	−0.017 (0.041)	−0.075 (0.05)	0.034 (0.052)	0.002 (0.044)	0.133 ** (0.059)
overall rating	−0.081 (0.088)	−0.222 ** (0.108)	0.206 * (0.111)	0.196 ** (0.093)	0.254 ** (0.127)
review number	−0.086 ** (0.036)	0.118 *** (0.044)	0.056 (0.046)	0.042 (0.038)	−0.349 *** (0.048)
hotel honor	−0.026 (0.08)	0.193 * (0.099)	0.034 (0.102)	−0.189 ** (0.085)	−0.235 ** (0.115)
nearby res	−0.046 (0.05)	0.019 (0.062)	−0.027 (0.063)	0.035 (0.053)	−0.003 (0.072)
nearby spot	0.052 (0.048)	−0.03 (0.059)	0.015 (0.061)	−0.05 (0.051)	0.014 (0.069)
traffic	−0.094 (0.064)	0.063 (0.08)	0.004 (0.082)	0.019 (0.069)	−0.012 (0.093)
Eco friendly	−0.084 (0.093)	0.069 (0.115)	0.09 (0.118)	−0.066 (0.099)	−0.231 * (0.133)
_cons	1.761 *** (0.511)	−0.004 (0.644)	−2.312 *** (0.648)	−0.522 (0.555)	−1.072 (0.75)
Observations	314	314	314	314	314
R-squared	0.755	0.625	0.605	0.721	0.701

Standard errors are in parentheses; *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Business travelers are mainly occupied with official matters and conferences, and tend to choose hotels close to workplaces or conference venues [6]. The regression results reveal that location, the core attribute of business travelers, exerts significant effects on other attributes: value (0.227 ***), service (0.270 ***), room (−0.158 ***), and cleanliness (0.292 ***). Since most business travel costs are covered by employers, such travelers are less sensitive to value-related factors [16]. A well-located hotel enables them to return comfortably after busy workdays and greatly improves their overall accommodation experience. Meanwhile, location shows no significant influence on sleep quality, which is in line with the behavioral traits of business travelers. Existing studies have confirmed that business travelers attach great importance to sleep quality [51,52]. Sufficient sound sleep helps them restore energy and stay energetic for subsequent business activities [50], so they will make independent evaluations of sleep quality separately.

Couples pursue romance and a pleasant atmosphere when travelling and value privacy, regarding hotel rooms as vital spaces for intimate interaction [49]. The regression results show that except for location (−0.012), their core attribute room exerts significant effects on other attributes, including value (−0.521 ***), service (0.175 **), cleanliness (0.499 ***) and sleep quality (0.274 ***). Hence, the warm and romantic atmosphere of rooms can shape couples' perceptions of other hotel attributes. Furthermore, room conditions have no significant influence on location. Since couples tend to pick hotels at distinctive and meaningful locations related to their emotional experiences to leave precious memories, they are inclined to assess hotel location independently.

Family travelers usually travel with children or elderly family members. They attach great importance to cleanliness and prefer hotels with tidy surroundings and cozy beds to gain a homely sense of comfort during trips. Their core attribute, cleanliness, exerts significant effects on value (0.296 ***), location (0.301 ***), service (0.221 ***) and room (0.317 ***), yet shows no obvious impact on sleep quality. In reality, family travelers often spend plenty of energy taking care of family members in the daytime. They crave quiet surroundings for sound rest at night [53], so they will separately assess sleep quality.

Groups of friends focus more on social interaction experiences and favor spacious, comfortable rooms suitable for group accommodation and recreational activities. Their core attribute room poses significant influences on value (−0.353 ***), location (−0.334 ***), service (0.248 ***), cleanliness (0.64 ***) and sleep quality (0.636 ***). Obviously, a pleasant in-room atmosphere brings them relaxation and joy, and further affects their judgments on other service attributes.

Solo travelers place the highest emphasis on room attributes. This core attribute exerts significant effects on value (−0.325 ***), location (−0.274 ***), service (0.467 ***), cleanliness (0.652 ***) and sleep quality (0.822 ***). This is mainly because solo travelers desire exclusive private space and thus focus closely on room quality and comfort [6]. Other auxiliary service perceptions are easily shaped by the direct experience brought by room conditions.

We also identify an interesting finding that room attributes are negatively correlated with location and value respectively. In line with rent competition theory, hotels situated in prime bustling areas bear higher land costs [54]. To secure superior geographical positions, some hotels tend to compromise on room conditions, such as downsizing room space or cutting indoor facilities. Travelers therefore have to make trade-offs between location advantages and room quality. They may accept simpler rooms for a better location, or choose well-furnished rooms at the cost of a less ideal location.

As for the negative correlation between room quality and perceived value, it can be explained by the psychological contrast effect [55]. When assessing multiple dimensions, people's judgment on one aspect will interfere with that on another. Travelers satisfied with superior room conditions are more likely to feel the charge is overpriced, thus forming negative evaluations of overall value. Based on transaction utility theory [56,57], consumers' value perception relies on both acquisition utility (whether room quality matches the price) and transaction utility (whether the actual price is lower than their internal reference price). When travelers enjoy high-quality rooms, their inner expectation of cost-effective consumption will lower their psychological reference price. Consequently, the real price exceeds this adjusted standard, producing negative transaction utility and resulting in unsatisfactory value perception even with a satisfactory room experience.

Different types of travelers hold distinct satisfaction degrees toward various hotel attributes according to their personalized demands. Perceptions of core attributes exert remarkable influences on other attributes, which verifies their interdependent relationships. Specifically, satisfaction with room attributes is negatively correlated with that of location and value. Room, location and value belong to tangible evaluation

dimensions, whose actual conditions are easily affected by external constraints and hard to adjust flexibly. Hotels located in prime areas may reduce room standards to offset high operating costs, forcing travelers to make trade-offs between geographical advantages and room quality during selection.

By contrast, service, cleanliness and sleep quality mainly reflect travelers' emotional feelings, and core attributes serve as key factors dominating overall hotel evaluations. To be specific, high satisfaction with core attributes can compensate for unsatisfactory experience of secondary attributes, whereas negative feelings toward core attributes will weaken travelers' favorable impressions of other attributes and further lower overall hotel evaluation. This evaluation mechanism is fully consistent with the halo effect.

5. Hotel Recommendation Model Based on Attribute Correlation

Hotel attributes are essential for building hotel recommendation models, as they directly reflect traveler preferences and improve model performance. This study verifies the correlations among hotel attributes, confirming that travelers' experience of core attributes affects their evaluations of other attributes. Recommendation models incorporating such attribute correlations can better support travelers' decision-making. Figure 2 presents the overall research framework.

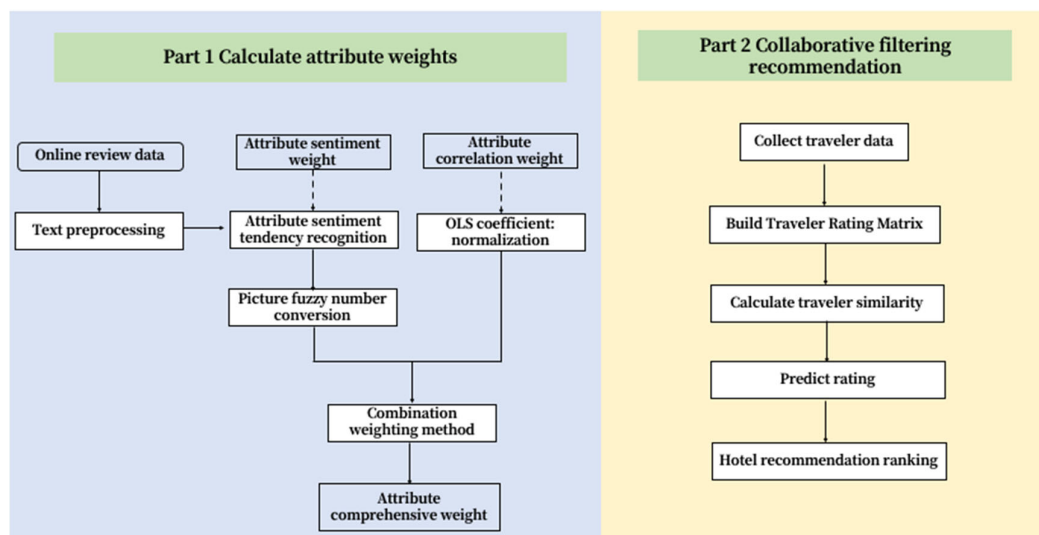


Figure 2. Research design framework (recommendation model).

5.1. Build the Hotel Recommendation Model

This study first assigned scores to six attributes across different traveler groups based on verified attribute correlations. Attribute weighting integrated travelers' emotional tendencies extracted from textual reviews and regression coefficients that reflect the influence intensity of core attributes on other attributes. The calculated weights were further embedded into the collaborative filtering algorithm to explore how incorporating attribute correlations affects hotel recommendation results.

5.1.1. Attribute Sentiment Recognition

For the first step, we collected 885,490 online reviews of 320 hotels in London with three stars and above on TripAdvisor. The second step involved data cleaning and text preprocessing, including converting text reviews to lowercase, correcting spelling errors, and removing stop words. The third step was to identify the attribute sentiment value in each online review.

In each review, attribute sentiment values above 0.5 are considered positive, below −0.5 as negative, and between −0.5 and 0.5 as neutral. If an attribute is not mentioned, the sentiment for that attribute is marked as missing. The results categorize sentiments as positive, neutral, negative, or missing. Referring to Tao et al. (2022) [29], let $S = s_1, s_2, s_3, s_4$ denote the set of sentiment orientations, where s_1, s_2, s_3, s_4 represent the sentiment orientations of positive, neutral, negative and refusal, respectively. s_e is the e th sentiment orientation, $e = 1, 2, 3, 4$. Let ξ_{ijk} denote the sentiment orientation of the j th attribute in the k th online review of the i th alternative hotel, $\xi_{ijk} \in S, i = 1, 2, \dots, m, j = 1, 2, \dots, n, k = 1, 2, \dots, N_i$. Let h_{ij}^e represent the number of online reviews of the i th hotel with a sentiment orientation for the j th attribute as $s_e, i = 1, 2, \dots, m, j = 1, 2, \dots, n, e = 1, 2, 3, 4$. Then h_{ij}^e can be determined based on ξ_{ijk} , and the calculation formula is:

$$h_{ij}^e = \sum_{k=1}^{N_i} \varphi^e(\xi_{ijk}), i = 1, 2, \dots, m, j = 1, 2, \dots, n, e = 1, 2, 3, 4 \quad (1)$$

where $\varphi^e(\xi_{ijk})$ is the indicator function defined as:

$$\varphi^e(\xi_{ijk}) = \begin{cases} 1, & \xi_{ijk} = s_e \\ 0, & \xi_{ijk} \neq s_e \end{cases}, i = 1, 2, \dots, m, j = 1, 2, \dots, n, e = 1, 2, 3, 4 \quad (2)$$

5.1.2. Transforming Attributes into Picture Fuzzy Number

A fuzzy number can represent support, hesitation, opposition, and abstention for events using fuzzy sets and membership functions. This method is useful in dealing with uncertainty and fuzziness in various fields. Travelers' positive emotions in reviews can be considered as support, negative emotions as opposition, neutral emotions as hesitation, and missing sentiments as abstention. Therefore, the emotional orientations of hotel attributes in reviews can be transformed into the form of fuzzy numbers. The specific transformation process is as follows:

Let p_{ij}^e represents the probability that hotel A_i 's sentiment orientation towards attribute c_j is s_e . The calculation formula is:

$$p_{ij}^e = \frac{h_{ij}^e}{N_i}, i = 1, 2, \dots, m, j = 1, 2, \dots, n, e = 1, 2, 3, 4 \quad (3)$$

Based on the definition of fuzzy numbers, the probabilities $p_{ij}^1, p_{ij}^2, p_{ij}^3$ and p_{ij}^4 determined by Equation (3) can be respectively regarded as the support membership, hesitation membership, opposition membership, and abstention membership of alternative hotel A_i with respect to attribute c_j . Therefore, let $r_{ij} = \langle \mu_{ij}, \eta_{ij}, \gamma_{ij}, \pi_{ij} \rangle$ represent the fuzzy number form of the sentiment orientation identified from online reviews for alternative hotel A_i with respect to attribute c_j , where $\mu_{ij} = p_{ij}^1, \eta_{ij} = p_{ij}^2, \gamma_{ij} = p_{ij}^3$ and $\pi_{ij} = p_{ij}^4$ represent the support membership, hesitation membership, opposition membership, and abstention membership, $i = 1, 2, \dots, m, j = 1, 2, \dots, n$.

5.1.3. Determine Attribute Weight

Travelers' emotional expressions in text reviews indicate their satisfaction with different attributes. By converting these emotional orientations into fuzzy numbers, we calculated the emotional weights of the attributes. In addition, as seen from the previous regression analysis results, there are correlations between attributes. Core attributes have different degrees of influence on other attributes, and the degree of influence is reflected in the regression coefficients. Based on this, we comprehensively consider the results of textual sentiment measurement and the correlation between attributes to assign comprehensive weights to the attributes. The specific process is as follows:

First, based on the definition of the entropy of fuzzy sets and the fuzzy numbers r_{ij} of alternative hotels with respect to attribute c_j , the entropy $E_j(R)$ of attribute c_j can be calculated using the following formula:

$$E_j(R) = \frac{1}{m} \sum_{i=1}^m \left(\frac{1}{\ln 4} (-\mu_{ij} \ln \mu_{ij} - \eta_{ij} \ln \eta_{ij} - \gamma_{ij} \ln \gamma_{ij} - \pi_{ij} \ln \pi_{ij}) \right) \quad (4)$$

Furthermore, the weights $\omega_{j(o)}$ of each attribute can be determined using the entropy method, with the calculation formula as follows:

$$\omega_{j(o)} = \frac{1 - E_j(R)}{n - \sum_{j=1}^n E_j(R)} \quad (5)$$

Next, based on the OLS regression analysis results, the equation coefficients represent the degree of influence of core attributes on other attributes. Using the equation coefficients to assign weights to the corresponding attributes, where the core attributes are assigned a value of 1, and the attributes with no significant relationship to the core attributes are assigned a value of 0. Then, the coefficients β_j are normalized to obtain the weights of attribute correlations.

$$\omega_{j(s)} = \frac{|\beta_j|}{\sum_{j=1}^n |\beta_j|} \quad (6)$$

Finally, based on the weights $\omega_{j(o)}$ of different types of travelers for attribute c_j calculated from Equations (4) and (5), as well as the attribute correlation weights $\omega_{j(s)}$ calculated from Equation (6), the comprehensive weight ω_j of attribute c_j can be determined using the combination weighting method, with the calculation formula as follows:

$$\omega_j = \theta_1 \omega_{j(s)} + \theta_2 \omega_{j(o)}, j = 1, 2, \dots, n \quad (7)$$

where $0 \leq \omega_j \leq 1$, θ_1, θ_2 are parameters and satisfy $0 \leq \theta_1 \leq 1$, $0 \leq \theta_2 \leq 1$, $\theta_1 + \theta_2 = 1$.

5.1.4. Hotel Recommendation Based on Collaborative Filtering Algorithm

Collaborative filtering is a widely adopted recommendation algorithm. It requires no predefined item features or user attributes, and generates recommendations according to user behaviors and user similarity. Its core principle is to calculate user similarity via the Pearson correlation coefficient or cosine similarity, select k nearest neighbors for each user, and further predict users' ratings on unevaluated items [25]. Nevertheless, when adopted for multi-attribute rating data, traditional collaborative filtering has obvious drawbacks, namely user rating scale heterogeneity [58]. Users hold distinct rating habits and evaluation standards. For instance, some users regard 6 points out of 10 as a highly satisfactory score, while others treat the same score as unsatisfactory. Such divergence wherein identical numerical ratings convey different subjective attitudes will cause systematic bias in user similarity computation and subsequent rating prediction. Traditional cosine similarity merely compares original rating vectors and ignores individual rating bias, which results in improper neighbor selection and poor recommendation performance.

To address this limitation, the present study incorporates attribute correlations into the construction of the user-hotel rating matrix. Specifically, instead of directly using users' raw attribute-level ratings, we construct the rating matrix based on users' sentiment scores for each attribute and the observed inter-attribute correlations, which may reflect the influence of core attributes on travelers' overall evaluation patterns. By doing so, the rating matrix more closely reflects users' true preferences, thereby mitigating the negative impact of user rating scale heterogeneity and improving the accuracy of user similarity calculation and hotel recommendation performance.

The first step is to collect traveler behavior data. This study uses hotel names, traveler ID, traveler types, traveler addresses, and traveler sentiment scores for six hotel attributes (value, location, service, room, cleanliness, sleep quality) as the basis for hotel recommendations. Here, the sentiment score is denoted as S_{uj} , representing the sentiment value of traveler u for attribute j , where $u = 1, 2, \dots, U_i$, and $j = 1, 2, \dots, n$.

The second step involves creating a traveler-hotel rating matrix where rows represent travelers and columns represent hotels. Each element reflects the traveler's overall rating for the hotel. Considering that different traveler types prioritize hotel attributes differently, we factor in the weighted correlation between attributes to calculate the traveler's comprehensive rating for the hotel using a straightforward weighting approach. The specific calculation formula is:

$$r_{ui} = \sum_{j=1}^n (S_{uj} \times w_j). \quad (8)$$

where S_{uj} represents the traveler u 's sentiment for attribute j , w_j represents the weight of attribute j , and different types of travelers have different attribute weights. r_{ui} is the comprehensive rating of traveler u for hotel i .

The third step is to calculate user similarity. This study uses cosine similarity to calculate the similarity between users. Cosine similarity is widely used in fields such as image processing and document retrieval. In collaborative filtering, the row vectors in the score matrix represent users. Let the vectors be represented as A and B , where the element values are the user's comprehensive ratings for the hotels. The cosine value of the angle between the two vectors A and B represents the similarity between the two users. The calculation formula (9) is as follows:

$$\text{sim}_{ab} = \cos(A, B) = \frac{A \times B}{\|A\| \times \|B\|} = \frac{\sum_{u=1}^{U_i} (A_u \times B_u)}{\sqrt{\sum_{u=1}^{U_i} (A_u)^2} \times \sqrt{\sum_{u=1}^{U_i} (B_u)^2}} \quad (9)$$

The fourth step is to predict ratings. For the target user, based on the ratings of the k most similar users, the prediction for the target user's unrated items is calculated using the formula (10):

$$p(a, i) = \sum_{a \in S(a, K) \cap R(i)} \text{sim}_{ab} r_{bi} \quad (10)$$

where $S(a, K)$ represents the K users with interests close to user a , $R(i)$ is the set of users who have rated hotel i , sim_{ab} is the interest similarity between user a and user b , and r_{bi} represents user b 's rating for hotel i .

The fifth step is to generate a recommendation list based on the predicted ratings. The list is sorted in descending order of predicted ratings, and the top N hotels are selected as the recommendation results. To measure the effectiveness of the recommendations, this study calculates the precision, MAE, and RMSE. Precision measures the overlap between the recommended list and the items the user actually likes, MAE measures the average absolute error between the predicted ratings and the actual ratings, and RMSE measures the root mean square error between the predicted ratings and the actual ratings. The calculation formulas (11)–(13) are as follows:

$$\text{Precision} = \frac{R(U) \cap F(U)}{U_i} \quad (11)$$

where $R(U) \cap F(U)$ represents the intersection quantity of items in the recommendation list and the items the user actually likes, and U_i represents the total recommendations.

$$\text{MAE} = \frac{1}{t} \sum_{i=1}^t |p_i - f_i| \quad (12)$$

$$\text{RMSE} = \sqrt{\frac{1}{t} \sum_{i=1}^t (p_i - f_i)^2} \quad (13)$$

where p_i represents the predicted rating for the hotel, and f_i represents the actual rating for the hotel, and t represents the number of ratings.

5.2. Case Study

In this study, five hotels in London were selected, namely The Resident Kensington Hotel (H_1), Radisson Blu Edwardian Vanderbilt Hotel (H_2), Radisson Blu Edwardian Grafton Hotel (H_3), Point A London Kings Cross St Pancras Hotel (H_4), and Park Plaza Westminster Bridge London (H_5). These hotels are evaluated based on six attributes: value,

location, room, service, cleanliness, and sleep quality. Additionally, as the purpose of this study is to assess the recommendation effectiveness of a hotel recommendation model that considers attribute correlations, the proportion of attribute correlation weights and text sentiment measurement weights was set at $\theta_1 : \theta_2 = 7:3$, indicating a greater importance of attribute correlation weights. At this point, the core attribute weights are also the highest. The following are the main calculation processes and results.

Firstly, online reviews for the above five alternative hotels were obtained from TripAdvisor. Then, the online reviews for each of the five alternative hotels were categorized based on traveler types, and the sentiment orientations of the six attributes in each review for different traveler groups were identified. The results of the sentiment tendency identification were denoted as ξ_{ijk} , $\xi_{ijk} \in S = \{s_1, s_2, s_3, s_4\}$. Based on ξ_{ijk} , the quantities of the four types of orientations for the six attributes of the five hotels were obtained using Equations (1) and (2). Taking business travelers as an example, the results are shown in Table 21.

Table 21. The numbers of reviews of business travelers' attribute sentiment orientation.

Hotel	Value				Location				Service			
	Pos	Neu	Neg	Null	Pos	Neu	Neg	Null	Pos	Neu	Neg	Null
H_1	356	49	7	99	446	51	5	9	324	60	12	115
H_2	804	147	39	245	946	196	47	46	708	186	48	293
H_3	1084	232	72	384	1263	324	94	91	885	293	99	495
H_4	670	128	30	235	829	154	33	47	502	126	21	414
H_5	2407	452	104	720	2911	502	131	139	1856	610	235	982
Hotel	Room				Cleanliness				Sleep Quality			
	Pos	Neu	Neg	Null	Pos	Neu	Neg	Null	Pos	Neu	Neg	Null
H_1	427	62	10	12	482	23	2	4	136	99	13	263
H_2	788	324	71	52	1037	140	44	14	155	219	56	805
H_3	1097	476	105	94	1406	232	90	44	284	383	117	988
H_4	784	182	56	41	938	72	38	15	216	192	66	589
H_5	2552	727	156	248	3055	357	155	116	493	410	115	2665

Next, according to Equation (3), the probabilities of positive sentiment orientation, neutral sentiment orientation, negative sentiment orientation, and missing items for the six attributes of the five hotels were obtained. Then, based on the definition of picture fuzzy numbers, the probabilities of the four types of sentiment orientations for the six attributes of the five hotels were converted into evaluation results in the form of picture fuzzy numbers. Taking business travelers as an example, the results are shown in Table 22.

Table 22. Fuzzy number results of business travelers about each attribute.

Hotel	Value				Location				Service			
	Pos	Neu	Neg	Null	Pos	Neu	Neg	Null	Pos	Neu	Neg	Null
H_1	0.696673	0.095890	0.013699	0.193738	0.872798	0.99804	0.009785	0.017613	0.634051	0.117417	0.23483	0.225049
H_2	0.651012	0.119028	0.031579	0.198381	0.765992	0.158704	0.038057	0.037247	0.573279	0.150607	0.038866	0.237247
H_3	0.611738	0.130926	0.040632	0.216704	0.712754	0.182844	0.053047	0.051354	0.499436	0.165350	0.055869	0.279345
H_4	0.630292	0.120414	0.028222	0.221072	0.779868	0.144873	0.031044	0.044214	0.472248	0.118532	0.019755	0.389464
H_5	0.653543	0.122726	0.028238	0.195493	0.790388	0.136302	0.035569	0.037741	0.503937	0.165626	0.063807	0.266630
Hotel	Room				Cleanliness				Sleep Quality			
	Pos	Neu	Neg	Null	Pos	Neu	Neg	Null	Pos	Neu	Neg	Null
H_1	0.835616	0.121331	0.019569	0.023483	0.943249	0.045010	0.003914	0.007828	0.266145	0.193738	0.025440	0.514677
H_2	0.638057	0.262348	0.057490	0.042105	0.839676	0.113360	0.035628	0.011336	0.125506	0.177328	0.045344	0.651822
H_3	0.619074	0.268623	0.059255	0.053047	0.793454	0.130926	0.050790	0.024831	0.160271	0.216140	0.066027	0.557562
H_4	0.737535	0.171214	0.052681	0.038570	0.882408	0.067733	0.035748	0.014111	0.203198	0.180621	0.062088	0.554092
H_5	0.692913	0.197393	0.042357	0.067336	0.829487	0.096932	0.042085	0.031496	0.133858	0.111322	0.031225	0.723595

Then, based on Equations (4) and (5), the attribute weights $\omega_{j(o)}$ of the alternative hotels were calculated. Using Equation (6), the attribute correlation weights $\omega_{j(s)}$ were calculated, and then the comprehensive weights ω_j of the attributes were calculated using Equation (7). The results of the comprehensive weights are shown in Table 23.

Table 23. Weight of the 6 attributes for different traveler types.

Traveler Type	Attribute					
	Value	Location	Service	Room	Cleanliness	Sleep Quality
business	0.1222	0.4245	0.1256	0.1087	0.1866	0.0324
couples	0.1858	0.0696	0.0796	0.3379	0.2249	0.1022
family	0.1335	0.1700	0.1017	0.1568	0.4082	0.0298
friend	0.1129	0.1413	0.0820	0.2724	0.2225	0.1689
solo	0.1021	0.1204	0.1219	0.2528	0.2137	0.1891

We then generate hotel recommendation results based on the collaborative filtering algorithm. First, traveler behavior data were collected from TripAdvisor. Taking Park Plaza Westminster Bridge London as an example, Table 24 displays partial sample data, in which the indicators including value, location, service, room, cleanliness and sleep quality refer to sentiment values extracted from online reviews. It can be preliminarily observed from the data that a high sentiment value of core attributes usually corresponds to a higher overall rating. On the contrary, low sentiment scores of core attributes will interfere with users' perception of other attributes, thus leading to lower overall ratings.

Table 24. Traveler behavior data.

User ID	User Address	User Type	Overall Rating	Value	Location	Service	Room	Clean	Sleep
ID001	Morristown	Business	4.00	0.49	0.49	0.19	0.44	0.46	0.44
ID002	None	Business	5.00	0.57	0.50	0.28	0.10	0.03	−0.42
ID003	Athens	Business	1.00	0.46	−0.12	0.14	−0.14	0.00	0.00
ID004	None	Couples	5.00	0.62	0.62	0.81	0.70	0.69	0.56
ID005	Tami R	Couples	5.00	0.31	0.26	0.26	0.56	0.49	0.42
ID006	Sleaford	Couples	1.00	0.10	0.04	0.15	−0.01	−0.02	0.00
ID007	Verwood	Family	4.00	0.34	0.41	0.54	0.44	0.45	0.00
ID008	Oakham	Family	5.00	0.66	0.58	0.51	0.32	0.66	0.42
ID009	Darlington	Family	2.00	0.40	−0.25	−0.30	−0.20	−0.13	−0.16
ID010	Norwich	Friends	5.00	0.45	0.46	0.30	0.53	0.24	0.00
ID011	Somerset	Friends	4.00	0.53	0.35	0.00	0.68	0.51	−0.42
ID012	Nottingham	Friends	1.00	0.11	0.39	0.00	−0.08	0.00	−0.07
ID013	Atlanta	Solo	5.00	0.33	0.09	0.61	0.73	0.58	0.73
ID014	Leicestershire	Solo	4.00	0.53	0.28	0.53	0.79	0.56	0.00
ID015	Stockholm	Solo	1.00	0.31	0.04	−0.17	−0.05	0.04	0.00

Note: To protect the reviewers' privacy, all reviewer IDs have been anonymized using sequential pseudonyms (e.g., ID001, ID002, ID003, etc.) instead of their real usernames or identifiers.

Secondly, we constructed the traveler-hotel rating matrix. In the matrix, rows represent travelers, columns represent hotels, and each element represents the comprehensive rating of the travelers for the hotel. Table 25 presents some of the matrix data.

Table 25. The traveler-hotel rating matrix.

User	Hotel 1	Hotel 2	Hotel 3	Hotel 4	Hotel 5
1	0	0	0.27130564	0	0.21988132
2	0	0.004665	0.29179818	0	0.2807257
3	0	0.20195976	0.1963495	0	0.24008417
4	0	0	0.46864283	0.44075337	0.16149974
5	0	0	0.23304562	0.17679692	0.37140592
6	0	0.1606086	0.05435809	0	0.32611536
...	...	...	...	...	...

Next, the cosine similarity was used to calculate the similarity between the traveler and other travelers. Based on the comprehensive ratings of the five most similar tourists, the model predicts the traveler's ratings for unrated hotels. An example of different types of travelers' hotel recommendation rankings under different preferences is shown in Table 26.

Table 26. Ranking results of alternative hotels.

User ID	User Type	Rank				
		H_1	H_2	H_3	H_4	H_5
ID01	Business	5	1	3	4	2
ID02	Couples	4	2	3	5	1
ID03	Family	3	4	1	5	2
ID04	Friend	3	4	1	5	2
ID05	Solo	1	4	5	3	2

Note: All reviewer IDs have been anonymized using sequential pseudonyms (e.g., ID01, ID02, ID03, etc.) instead of their real usernames or identifiers.

To evaluate the recommendation model, the dataset was randomly split into a training set (80%) and a testing set (20%). The training set was used to construct the user-hotel rating matrix, and the testing set was used to evaluate model performance. The model's accuracy is 88.22%, MAE is 0.0264, and RMSE is 0.0719. It can be seen that the model has high accuracy in predicting user preferences, with small differences between predicted and actual values. When not considering the correlation between attributes, the model's accuracy is 85.06%, MAE is 0.0265, and RMSE is 0.0760. Our model improves prediction accuracy by 3.16 percentage points, demonstrating that incorporating attribute correlations into collaborative filtering significantly enhances recommendation performance. To further validate this conclusion, the study expanded the number of hotels and calculated the model's accuracy, MAE, and RMSE considering attribute correlation (S1) and not considering attribute correlation (S2), as shown in Table 27. It can be observed that, across different hotel sample sizes, the model's predictive accuracy considering attribute correlation is higher than that not considering attribute correlation. This indicates that considering the correlation between attributes can more accurately capture the user's overall experience and preferences, aligning better with the user's actual needs and preferences, and providing better recommendation accuracy and practicality.

Table 27. Recommendation effect of different hotel sample sizes.

Hotel Number	Accuracy Rate		MAE		RMSE	
	S1	S2	S1	S2	S1	S2
5	88.22%	85.06%	0.0264	0.0265	0.0719	0.0760
10	81.63%	80.11%	0.0352	0.0388	0.0893	0.0926

15	72.72%	68.34%	0.0584	0.0610	0.1268	0.1278
20	75.64%	69.12%	0.0492	0.0510	0.1126	0.1110
25	73.99%	71.25%	0.0479	0.0547	0.1111	0.1193

Furthermore, to validate the reasonableness of the proportion setting and the significant role of core attributes, the study adjusted the ratio of attribute correlation weight and text sentiment measurement weight, denoted as θ_1 and θ_2 . The results are shown in Table 28. When the ratio of θ_1 and θ_2 is 5:5 or 6:4, there is a problem of core attributes having less weight than other attributes, leading to a decrease in model recommendation accuracy. It can be seen that core attributes have a significant impact on the traveler experience; assigning greater weight to core attributes helps improve the accuracy of hotel recommendations.

Table 28. Recommendation effect of different ratios.

Ratio/ Hotel Number	$\theta_1:\theta_2$	5:5	6:4	7:3
5	Accuracy rate	84.48%	84.20%	88.22%
	MAE	0.0327	0.0270	0.0264
	RMSE	0.0856	0.0763	0.0719
10	Accuracy rate	79.54%	80.68%	81.63%
	MAE	0.0379	0.0373	0.0352
	RMSE	0.0931	0.0965	0.0893
15	Accuracy rate	72.07%	71.61%	72.72%
	MAE	0.0584	0.0522	0.0584
	RMSE	0.1269	0.1175	0.1268
20	Accuracy rate	73.39%	74.94%	75.64%
	MAE	0.0463	0.0473	0.0492
	RMSE	0.1076	0.1105	0.1126
25	Accuracy rate	70.63%	73.28%	73.99%
	MAE	0.0561	0.0509	0.0479
	RMSE	0.1231	0.1153	0.1111

6. Conclusions

6.1. Theoretical Implications

This study adopts online review text mining to identify core hotel attributes for different traveler groups, explores the influencing mechanisms of core attributes on secondary attributes, and constructs an improved hotel recommendation model that incorporates traveler heterogeneity and attribute correlation effects. The major theoretical contributions of this study are summarized as follows:

First, prior literature largely ignores interdependencies among hotel attributes. Although a few studies have attempted to verify attribute relationships, they fail to consider traveler type heterogeneity. In fact, attribute importance varies substantially across different traveler groups. This study identifies the directional influencing effects of core attributes on other attributes and further verifies that such effects differ significantly by traveler type. This study defines core attributes as highly salient and priority dimensions for specific traveler groups; these attributes not only determine travelers' direct perception of hotel features but also reshape their evaluations of other independent attributes. Core attributes dominate overall hotel assessment: strengths in secondary attributes can be compensated by superior core attribute performance, whereas poor core attribute satisfaction significantly undermines travelers' overall hotel impressions. This

directional asymmetric influence mechanism—rather than simple attribute correlation—has not been systematically investigated in existing research.

Second, most existing hotel recommendation models treat hotel attributes as mutually independent and neglect their inherent psychological interdependencies. This study optimizes the traditional research framework by integrating attribute-level sentiment scores and regression-derived correlation coefficients into the construction of the user-hotel rating matrix. The improved collaborative filtering algorithm fully incorporates attribute interactive relationships, instead of relying merely on isolated attribute scores. Accordingly, the proposed model captures fine-grained traveler preference characteristics and effectively enhances the accuracy and practical applicability of hotel recommendation results.

6.2. Practical Implications

This study offers practical implications for hotel managers and travel platforms. Hotel managers can formulate targeted operational strategies according to their resource endowments. For hotels with sufficient resources, optimizing core attributes corresponding to different traveler groups can effectively improve overall guest experience and brand reputation. Specifically, hotel location selection should prioritize transportation accessibility and proximity to tourist attractions or business districts. Guest rooms require regular renovation and upgrading to create a warm and cozy ambiance, while strict cleanliness and hygiene standards must be consistently maintained.

Resource-constrained hotels can concentrate on their dominant core attributes and target specific customer segments to build differentiated themed hotels. For example, hotels in suburban or remote areas can focus on friend groups, solo travelers, and family travelers by prioritizing room quality and cleanliness in operation and promotion. Marketing campaigns should fully highlight these competitive strengths with targeted labels such as “friend gathering” and “solo travel-friendly” to precisely attract target customers. When inherent resource limitations fail to meet traveler expectations, hotels can provide supplementary supportive services for compensation. For instance, operating shuttle buses can effectively mitigate the disadvantage of inconvenient locations.

Travel platforms can deliver customized hotel recommendations for different traveler groups. For example, hotels with intimate atmospheres and distinctive room designs can be preferentially recommended for travelling couples. To further improve recommendation accuracy, travel websites can launch incentive mechanisms to encourage users to publish high-quality reviews and thereby enrich review data volume. In addition, for new users suffering from the cold-start problem, platforms can adopt information flow advertising and guided preference collection to build personalized user profiles. The improved recommendation model can continuously optimize user browsing experience, boost user engagement and conversion rates, and ultimately increase platform traffic and revenue.

6.3. Limitations and Future Research

This study presents several limitations. First, the sentiment analysis results can be further validated via manual annotation or benchmark model comparison in future studies to improve analytical reliability. Second, this study does not conduct baseline comparisons with conventional collaborative filtering approaches (e.g., user-based and item-based CF), which should be supplemented in future work. Third, the dataset is limited to London hotel reviews collected from TripAdvisor before April 2023, which restricts the generalizability of the findings across different platforms, geographical regions and cultural contexts.

Furthermore, this study ignores the potential influence of hotel star ratings and does not distinguish between positive and negative reviews. According to loss aversion theory [59], negative experiences exert a stronger behavioral impact than positive ones, which implies that this study omits asymmetric emotional effects. Finally, this study does not resolve data sparsity and cold-start problems. Future research can adopt hybrid recommendation frameworks and knowledge graph techniques to alleviate these technical limitations.

Author Contributions: Conceptualization, Y.C.; Methodology, Y.C.; Software, S.C.; Validation, S.C.; Formal analysis, S.C. and Z.W.; Investigation, S.C. and Z.D.; Resources, S.Y.; Data curation, S.C. and Z.D.; Writing—original draft, Y.C., S.C. and Z.W.; Writing—review & editing, Y.C. and S.Y.; Visualization, Z.W. and Z.D.; Supervision, S.Y.; Project administration, S.Y.; Funding acquisition, S.Y. All authors have read and agreed to the published version of the manuscript.

Funding: This research has no funding.

Institutional Review Board Statement: Our study uses publicly available social media review data that is openly accessible without login or special permission. No human participants were recruited, no interventions or interactions were performed, and no private or sensitive information was collected. All data processing was conducted on anonymized and de-identified data, and the study poses minimal risk to individuals. The study involves only non-interventional analysis of publicly available data, which is exempt from formal ethics committee/IRB review under relevant guidelines. We confirm that no additional ethics approval is required for this research.

Informed Consent Statement: Informed consent from individual users was waived for the following reasons: (1) the data is publicly available and users have no reasonable expectation of privacy in these public posts; (2) the study poses no more than minimal risk to participants; (3) it is technically impracticable to obtain individual consent from all users of the large dataset; and (4) all data has been irreversibly anonymized, so no individual can be identified from the published results.

Data Availability Statement: The original contributions presented in the study are included in the article; further inquiries can be directed to the corresponding author.

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

References

1. Roy, G. Travelers' online review on hotel performance—Analyzing facts with the Theory of Lodging and sentiment analysis. *Int. J. Hosp. Manag.* **2023**, *111*, 103459.
2. Xu, X. Examining the relevance of online customer textual reviews on hotels' product and service attributes. *J. Hosp. Tour. Res.* **2019**, *43*, 141–163.
3. Zhang, C.; Xu, Z.; Gou, X.; Chen, S. An online reviews-driven method for the prioritization of improvements in hotel services. *Tour. Manag.* **2021**, *87*, 104382.
4. Mellinas, J.P.; Nicolau, J.L.; Park, S. Inconsistent behavior in online consumer reviews: The effects of hotel attribute ratings on location. *Tour. Manag.* **2019**, *71*, 421–427.
5. Nicolau, J.L.; Mellinas, J.P.; Martín-Fuentes, E. The halo effect: A longitudinal approach. *Ann. Tour. Res.* **2020**, *83*, 102938.
6. Rhee, H.T.; Yang, S.B. How does hotel attribute importance vary among different travelers? An exploratory case study based on a conjoint analysis. *Electron. Mark.* **2015**, *25*, 211–226.
7. Wang, L.; Wang, X.K.; Peng, J.J.; Wang, J.Q. The differences in hotel selection among various types of travellers: A comparative analysis with a useful bounded rationality behavioural decision support model. *Tour. Manag.* **2020**, *76*, 103961.
8. Lewis, R.C. Isolating differences in hotel attributes. *Cornell Hotel Restaur. Adm. Q.* **1984**, *25*, 64–77.
9. Aijzen, I. The theory of planned behavior. *Organizational Behavior and Human. Decis. Processes* **1991**, *50*, 179–211.
10. Levitt, T. *Marketing Imagination*; Free Press: New York, NY, USA, 1983.
11. Xu, X. Examining an asymmetric effect between online customer reviews emphasis and overall satisfaction determinants. *J. Bus. Res.* **2020**, *106*, 196–210.

12. Jokelainen, J.; Garrod, B.; Sthapit, E.; Pesonen, J. The role of experiential familiarity in shaping hotel-chain competitiveness. *Int. J. Contemp. Hosp. Manag.* **2023**, *36*, 2999–3015.
13. Francesco, G.; Roberta, G. Cross-country analysis of perception and emphasis of hotel attributes. *Tour. Manag.* **2019**, *74*, 24–42.
14. Sharma, T.; Chen, J.S.; Ramos, W.D.; Sharma, A. Visitors' eco-innovation adoption and green consumption behavior: The case of green hotels. *Int. J. Contemp. Hosp. Manag.* **2023**, *36*, 1005–1024.
15. Li, H.; Liu, Y.; Tan, C.W.; Hu, F. Comprehending customer satisfaction with hotels: Data analysis of consumer-generated reviews. *Int. J. Contemp. Hosp. Manag.* **2020**, *32*, 1713–1735.
16. Bi, J.W.; Liu, Y.; Fan, Z.P.; Zhang, J. Exploring asymmetric effects of attribute performance on customer satisfaction in the hotel industry. *Tour. Manag.* **2020**, *77*, 104006.
17. Kano, N. Attractive quality and must-be quality. *J. Jpn. Soc. Qual. Control* **1984**, *31*, 147–156.
18. Quan, L.; Kim, J.J.; Han, H. Customer views on comprehensive green hotel selection attributes and analysis of importance-performance. *J. Travel Tour. Mark.* **2022**, *39*, 535–554.
19. Athanasopoulou, P.; Giovanis, A.; Papakriakopoulos, D.; Gdonteli, K.; Choustoulakis, E. The Role of Traveler Type and Hotel Star Rating in the Effect of Hotel Service Attributes on Customer Satisfaction. *J. Inf. Syst. Eng. Manag.* **2025**, *10*, 413–421.
20. Ervianty, R.M.; Dina, N.Z.; Prasetyo, S. Clustering family-friendly hotels' guests to develop tourism marketing strategies. *J. Geogr. Inst. Jovan Cvijic SASA* **2024**, *74*, 213–227.
21. Doğan, S.; Basaran, M.A.; Kantarci, K. Determination of attributes affecting price-performance using fuzzy rule-based systems: Online ratings of hotels by travel 2.0 users. *J. Hosp. Tour. Technol.* **2020**, *11*, 291–311.
22. Thorndike, E.L. A constant error in psychological ratings. *J. Appl. Psychol.* **1920**, *4*, 25–29.
23. Nilashi, M.; Bagherifard, K.; Rahmani, M.; Rafe, V. A recommender system for tourism industry using cluster ensemble and prediction machine learning techniques. *Comput. Ind. Eng.* **2017**, *109*, 357–368.
24. Kwok, P.K.; Lau, H.Y.K. Hotel selection using a modified TOPSIS-based decision support algorithm. *Decis. Support Syst.* **2019**, *120*, 95–105.
25. Shambour, Q.Y.; Abu-Shareha, A.A.; Abualhaj, M.M. A Hotel Recommender System Based on Multi-Criteria Collaborative Filtering. *Inf. Technol. Control* **2022**, *51*, 390–402.
26. Le, Q.H.; Mau, T.N.; Tansuchat, R.; Huynh, V.N. A multi-criteria collaborative filtering approach using deep learning and Dempster-Shafer theory for hotel recommendations. *IEEE Access* **2022**, *10*, 37281–37293.
27. Nie, R.X.; Tian, Z.P.; Wang, J.Q.; Chin, K.S. Hotel selection driven by online textual reviews: Applying a semantic partitioned sentiment dictionary and evidence theory. *Int. J. Hosp. Manag.* **2020**, *88*, 495.
28. Wu, J.; Liu, C.; Wu, Y.; Cao, M.; Liu, Y. A novel hotel selection decision support model based on the online reviews from opinion leaders by best worst method. *Int. J. Comput. Intell. Syst.* **2022**, *15*, 19.
29. Tao, L.L.; You, T.H. A multi-criteria decision-making model for hotel selection by online reviews: Considering the traveller types and the interdependencies among criteria. *Appl. Intell.* **2022**, *52*, 12436–12456.
30. Ray, B.; Garain, A.; Sarkar, R. An ensemble-based hotel recommender system using sentiment analysis and aspect categorization of hotel reviews. *Appl. Soft Comput.* **2021**, *98*, 106935.
31. Zhang, D.; Xu, H.; Su, Z.; Xu, Y. Chinese comments sentiment classification based on word2vec and SVMperf. *Expert Syst. Appl.* **2015**, *42*, 1857–1863.
32. Mustafa, G.; Usman, M.; Yu, L.; Afzal, M.T.; Sulaiman, M.; Shahid, A. Multi-label classification of research articles using Word2Vec and identification of similarity threshold. *Sci. Rep.* **2021**, *11*, 21900.
33. Hu, N.; Zhang, T.; Gao, B.; Bose, I. What do hotel customers complain about? Text analysis using structural topic model. *Tour. Manag.* **2019**, *72*, 417–426.
34. Tversky, A.; Kahneman, D. Availability: A heuristic for judging frequency and probability. *Cogn. Psychol.* **1973**, *5*, 207–232.
35. Schmid, H.J.; Günther, F. Toward a unified socio-cognitive framework for salience in language. *Front. Psychol.* **2016**, *7*, 1110.
36. Cox, D.F. Cue utilization in the quality perception process. In *Proceedings of the 1967 American Marketing Association Conference*; American Marketing Association: Chicago, IL, USA, 1967.
37. Carroll, J.M.; Roeloffs, R. Computer selection of keywords using word-frequency analysis. *Am. Doc.* **1969**, *20*, 227–233.
38. Yoon, K.P.; Hwang, C.L. *Multiple Attribute Decision Making: An Introduction*; Sage Publications: Thousand Oaks, CA, USA, 1995.
39. de Lima Silva, D.F.; de Almeida Filho, A.T. Sorting with TOPSIS through boundary and characteristic profiles. *Comput. Ind. Eng.* **2020**, *141*, 106328.
40. Nilashi, M.; Mardani, A.; Liao, H.; Ahmadi, H.; Manaf, A.A.; Almukadi, W. A hybrid method with TOPSIS and machine learning techniques for sustainable development of green hotels considering online reviews. *Sustainability* **2019**, *11*, 6013.

41. Page, L.; Brin, S.; Motwani, R.; Winograd, T. *The Pagerank Citation Ranking: Bring Order to the Web*; Technical report; Stanford University: Stanford, CA, USA, 1998.
42. Montejó-Ráez, A.; Martínez-Cámara, E.; Martín-Valdivia, M.T.; López, L.A.U. Random walk weighting over sentiwordnet for sentiment polarity detection on twitter. In *Proceedings of the 3rd Workshop in Computational Approaches to Subjectivity and Sentiment Analysis*; Association for Computational Linguistics: Stroudsburg, PA, USA, 2012; pp. 3–10.
43. Bhatnagar, V.; Goyal, M.; Hussain, M.A. A novel aspect based framework for tourism sector with improvised aspect and opinion mining algorithm. In *Research Anthology on Implementing Sentiment Analysis Across Multiple Disciplines*; IGI Global Scientific Publishing: Hershey, PA, USA, 2022; pp. 314–328.
44. Chen, Y.; Zhong, Y.; Yu, S.; Xiao, Y.; Chen, S. Exploring bidirectional performance of hotel attributes through online reviews based on sentiment analysis and Kano-IPA model. *Appl. Sci.* **2022**, *12*, 692.
45. Rajaguru, R.; Hassanli, N. The role of trip purpose and hotel star rating on guests' satisfaction and WOM. *Int. J. Contemp. Hosp. Manag.* **2018**, *30*, 2268–2286.
46. Yu, Y.; Li, X.; Jai, T.M. The impact of green experience on customer satisfaction: Evidence from TripAdvisor. *Int. J. Contemp. Hosp. Manag.* **2017**, *29*, 1340–1361.
47. Yang, Y.; Mao, Z.; Tang, J. Understanding guest satisfaction with urban hotel location. *J. Travel Res.* **2018**, *57*, 243–259.
48. Zhang, Z.; Zhang, Z.; Yang, Y. The power of expert identity: How website-recognized expert reviews influence travelers' online rating behavior. *Tour. Manag.* **2016**, *55*, 15–24.
49. Kim, D.; Perdue, R.R. The effects of cognitive, affective, and sensory attributes on hotel Choice. *Int. J. Hosp. Manag.* **2013**, *35*, 246–257.
50. Kim, J.; Lee, S. Understanding the Preferences of Solo Travelers: A Study of Hotel Attributes. *J. Travel Res.* **2018**, *57*, 342–356.
51. Xiong, W.; Fan, F.; Qi, H. Effects of Environmental Change on Travelers' Sleep Health: Identifying Risk and Protective Factors. *Front. Psychol.* **2020**, *11*, 724.
52. Chen, H.S.; Severt, K.; Shin, Y.H.; Knowlden, A.; Hilliard, T.W. "How'd you sleep?" measuring business travelers' sleep quality and satisfaction in hotels. *J. Hosp. Tour. Insights* **2018**, *1*, 188–202.
53. Mao, Z.; Yang, Y.; Wang, M. Sleepless nights in hotels? Understanding factors that influence hotel sleep quality. *Int. J. Hosp. Manag.* **2018**, *74*, 189–201.
54. Alonso, W. A theory of the urban land market. *Pap. Reg. Sci.* **1960**, *6*, 149–157.
55. Festinger, L. A theory of social comparison processes. *Hum. Relat.* **1954**, *7*, 117–140.
56. Thaler, R. Mental accounting and consumer choice. *Mark. Sci.* **1985**, *4*, 199–214.
57. Oh, H. Price fairness and its asymmetric effects on overall price, quality, and value judgments: The case of an upscale hotel. *Tour. Manag.* **2003**, *24*, 387–399.
58. Leung, X.Y.; Yang, Y. Are all five points equal? Scaling heterogeneity in hotel online ratings. *Int. J. Hosp. Manag.* **2020**, *88*, 102539.
59. Lin, G.; Lin, M.S.; Song, H. An assessment of prospect theory in tourism decision-making research. *J. Travel Res.* **2024**, *63*, 275–297.

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.