Journal of
Theoretical and Applied
Electronic Commerce Research

MDPI

Article

Unveiling Customer Preferences from Online Tourism Reviews: A Data-Driven Framework for Customer Segmentation and Hotel Recommendation

Zhongmin Pu

College of Management Science, Chengdu University of Technology, Chengdu 610059, China; puzhongmin@cdut.edu.cn

Abstract

Online tourism platforms generate large volumes of customer reviews containing rich information about hotel experiences, yet customers' heterogeneous attribute preferences remain difficult to identify and utilize in customer segmentation and hotel recommendation. To address this gap, this study proposes an online review data-driven framework for customer segmentation and hotel recommendation. Specifically, the framework employs Bert-LDA and RNTN to extract key attributes and their sentiment from online reviews. Several machine learning models then assess the nonlinear relationship between attributes and overall customer satisfaction. Notably, the Local Interpretable Model-Agnostic Explanations algorithm is applied for the first time to evaluate each attribute's contribution to customer satisfaction in individual reviews, forming the basis for customer clustering. Next, a novel distance-based group member contribution measurement method is proposed to evaluate the differentiated attribute weights among segmented customer groups. Finally, the PLTS-VIKOR method is applied to generate differentiated recommendation lists for each group. The framework is evaluated through a case study based on 28,502 online tourism reviews. Results reveal substantial heterogeneity in attribute priorities across segments, while room quality and service remain important across most segments. The framework produces differentiated hotel rankings that reflect each segment's distinct preferences, and provides a practical approach for hotel recommendation.

Keywords: online reviews; customer segmentation; hotel recommendation; customer preferences; interpretable machine learning; multi-attribute decision-making

Check for updates

Academic Editor: Sang-Lin Han

Received: 21 September 2026
Revised: 22 October 2026
Accepted: 30 October 2026
Published: 4 November 2026

Copyright: © 2026 by the author. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license.

1. Introduction

With the advent of the digital age and the rapid development of social media, online reviews on various e-commerce platforms have become a crucial information source for corporate decision-making and market analysis. Particularly in the hospitality industry, online reviews not only reflect customer satisfaction and experiences but also provide rich insights into customer needs and behavior patterns [1]. Extracting preference information from these reviews is crucial for understanding customer needs, facilitating targeted service improvements, and formulating more effective marketing strategies [2]. Currently, significant attention is being given to leveraging natural language processing techniques, such as attribute extraction and sentiment analysis techniques, to mine customer preferences from online reviews, resulting in notable outcomes [3,4]. However, there is limited research on extracting customer preference information from online reviews specifically for customer segmentation and hotel recommendation.

To be specific, there are three research gaps. First, existing review-based customer segmentation provides a limited basis for revealing customers' attribute preferences. Traditional customer segmentation methods primarily rely on demographic or behavioral data, which do not directly reflect customer attribute preferences and often suffer from challenges such as data acquisition difficulties and limited timeliness [5]. Recent studies have used multi-attribute ratings [5], attribute sentiment, mention frequency, or TF-IDF values [6] as segmentation bases, but ratings and sentiments mainly reflect current satisfaction rather than attribute prioritization, and mention frequency does not necessarily reflect an attribute's importance in shaping satisfaction [7]. In addition, although the relationship between attribute performance and satisfaction is nonlinear [8–10], common machine learning methods such as XGBoost, RF and SVM are largely black-box models and, at best, measure overall rather than review-level attribute importance [7]. Second, existing group recommendation research has not fully translated heterogeneous segment preferences into segment-specific hotel recommendations while distinguishing individual member contributions. Prior work relies mainly on online ratings [11], which can introduce biases because identical ratings may stem from different concerns, such as service versus cleanliness; reviews offer more detailed preference information [12], but assigning member weights through geographic or social-network information often faces challenges related to data availability and quantification. Third, there is a lack of an integrated framework that links review mining, interpretable machine learning, customer segmentation, and hotel recommendation for customer segmentation and hotel recommendation. Most studies address isolated components, leaving unclear how review-level explanations can be aggregated into segment-level preferences and further translated into differentiated hotel rankings.

To address these gaps, this study proposes an online review-driven framework for preference-based customer segmentation and hotel recommendation. Specifically, it seeks to answer three research questions: RQ1: How can attribute-level importance in shaping customer satisfaction be identified and interpreted from online hotel reviews using an interpretable machine learning approach? RQ2: How can customers with different hotel attribute preferences be segmented based on review-level attribute-importance profiles? RQ3: How can individual preference profiles be translated into segment-specific attribute weights and incorporated into a multi-attribute decision-making model to generate differentiated hotel recommendations? The framework integrates individual-level attribute contributions, customer preference profiles, segment-specific attribute weights, and hotel rankings, making the recommendation process transparent and traceable. The main contributions can be summarized as follows: (1) We conduct a novel online review-driven framework integrating online review mining, interpretable machine learning, with the PLTS-VIKOR method for customer segmentation and hotel recommendation. (2) The LIME algorithm is employed to quantify the contribution of customer-relevant attributes at the individual review level by explaining nonlinear machine learning models, thereby providing an explainable basis for customer segmentation. (3) A novel distance-based method is proposed to measure the contribution of individual group members, enabling the determination of differentiated attribute weights for hotel recommendations. (4) The effectiveness and robustness of the proposed framework is demonstrated through a comprehensive case study on 28,502 TripAdvisor reviews, showing its ability to generate distinct customer segments and deliver differentiated hotel recommendations.

The structure of the remaining sections is as follows: Section 2 reviews relevant research on customer segmentation and hotel recommendation based on online reviews. Section 3 presents the research methodology of this study, including customer preference extraction, customer segmentation and hotel recommendation. Section 4 validates the effectiveness of the approach through a case study from TripAdvisor. Section 5 further compares and

discusses the proposed method with other approaches. Finally, Section 6 concludes the study, highlighting the implications and avenues for future research.

2. Literature Review

2.1. Theoretical Background

This study draws on benefit segmentation theory to conceptualize customer heterogeneity in attribute preference and multi-attribute decision-making (MADM) theory to support hotel evaluation and recommendation. Benefit segmentation theory, introduced by Haley [13], proposes that consumers can be segmented according to the benefits they seek from products or services, reflecting heterogeneity in their preferences and decision priorities. In the hospitality industry, customers may differ in the importance they attach to hotel attributes such as service, cleanliness, location, and value for money. For example, some customers may prioritize location and convenience, whereas others may place greater emphasis on cleanliness or value for money. Online reviews provide information about these attribute-level evaluations, making it possible to investigate preference heterogeneity beyond demographic characteristics or overall hotel ratings. From a benefit-segmentation perspective, such differences provide a basis for grouping customers according to the relative importance they attach to hotel attributes. Previous studies have applied aspect-based sentiment analysis (ABSA) to online hotel reviews to identify attribute-level evaluations and customer preferences [14]. However, how to further translate the relative contributions of different attributes expressed in reviews into features suitable for customer segmentation remains an important research question. To address this issue, this study uses Local Interpretable Model-Agnostic Explanations (LIME) to interpret the predictions of a customer satisfaction model at the individual-review level, construct review-level attribute-importance profiles, and subsequently identify groups of reviews with similar attribute-importance patterns through clustering.

MADM theory provides a systematic basis for evaluating alternatives across multiple criteria with potentially different weights. In hotel selection, customers may assign different levels of importance to different attributes, such as service and cleanliness, leading to different hotel rankings under different weighting schemes. Accordingly, identifying differences in attribute preferences is an important prerequisite for generating differentiated hotel recommendations. Furthermore, MADM methods have been applied to hotel evaluation and recommendation based on online reviews. For example, Yang, Liao [15] developed a hotel selection model integrating aspect-level sentiment analysis, prospect theory, and fuzzy inference, thereby incorporating attribute evaluations and customer risk preferences into the hotel decision-making process. Based on this, this study links review-based customer segmentation with differentiated attribute weighting and applies PLTS-VIKOR to generate hotel recommendations for different customer segments.

2.2. Customer Segmentation Related Studies

Customer segmentation is essential for achieving market success, as it enables corporations to accurately identify and meet the needs of different customer groups [16]. Research on customer segmentation originated in 1956 [17], and early studies relied on demographic characteristics, including age, gender, income, occupation, ZIP code and education level [18]. While these static indicators provided a preliminary customer profile, they were often too broad to capture subtle differences in customer behavior and attitudes. Subsequent to technological advancements in marketing, researchers began to focus on customer behavior traits, such as purchasing habits, usage frequency and purchasing channels [19]. These dynamic behavioral insights helped corporations gain a better understanding of customers, but challenges like high data privacy and difficulties in capturing sentiment factors re-

mained. In the digital age, needs-based customer segmentation has gained increasing attention [20]. Online reviews and ratings contain rich information about customer needs and preferences, revealing not only satisfaction and dissatisfaction but also specific expectations and focal points regarding products or services [21]. Utilizing natural language processing techniques to extract customer preferences from online reviews, particularly key attributes and attribute sentiment, has become well-developed [22]. However, the integration of these preference-based insights into customer segmentation frameworks remains sparsely investigated.

Recently, a few studies have begun exploring customer segmentation by extracting need and preference information from online reviews [23]. For example, Joung and Kim [7] implemented customer segment analysis and product improvements using attribute importance extracted from online reviews. In addition, research on hotel customer segmentation has also explored approaches based on attribute preferences and online review data. For example, Nessel, Kosciólek [24] developed a benefit-segmentation approach using electronic word-of-mouth attribute ratings to estimate tourists' preferences for accommodation attributes and identify groups with similar preference patterns. Nilashi, Minaei-Bidgoli [25] have segmented eco-friendly hotel tourists based on online ratings and ranked the key attributes of each segment. These studies often treat each review as a source of customer preference information, with sentiment values or attribute importance serving as segmentation bases. In particular, sentiment values are typically represented based on sentiment analysis results or online ratings [26,27], while attribute importance is often measured by the frequency of the attribute being mentioned [28]. Importantly, sentiment-based and rating-based approaches can provide valuable information about customers' expressed evaluations, and aspect-based sentiment analysis (ABSA) has been widely used to extract attribute-level opinions and support customer profiling. However, attribute sentiment and overall ratings primarily reflect evaluations of particular attributes or overall experiences, and do not necessarily indicate the relative contribution of each attribute to overall customer satisfaction. Similarly, although attribute frequency captures the prominence of topics in reviews, mention frequency alone does not establish an attribute's importance in shaping satisfaction. Therefore, the key to effective customer segmentation based on online reviews lies in developing robust metrics for measuring attribute importance.

Previous studies have shown that machine learning models can capture nonlinear relationships between hotel attributes and overall customer satisfaction. However, conventional predictive models, including SVM, RF and XGBoost, are black-box models, meaning that the underlying factors driving customer satisfaction and dissatisfaction are not interpretable [29]. Although interpretable models such as decision trees can reveal general attribute effects, they may not adequately capture review-specific variations in attribute contributions. Joung and Kim [7] attempted to use SHAP algorithms to interpret the impact of attributes on customer satisfaction, making it possible to analyze attribute importance in each review. Drawing from previous studies, this study employs LIME to explain individual satisfaction predictions and obtain review-level attribute-importance profiles. These profiles provide the input features for the subsequent clustering analysis.

2.3. Hotel Recommendation Based on Online Reviews

Online reviews provide an important information source for hotel recommendation because they contain not only overall evaluations but also detailed opinions concerning hotel attributes, such as rooms, service, cleanliness, location, and facilities. Earlier hotel recommendation studies relied predominantly on numerical ratings and collaborative-filtering techniques to infer customer preferences and identify suitable alternatives. However, overall ratings provide only an aggregated evaluation and cannot fully explain why customers

prefer or reject a particular hotel. To overcome this limitation, researchers have increasingly incorporated textual reviews and multi-attribute evaluations into hotel recommendation. For example, Rianthong and Dumrongsiri [30] developed a multidimensional hotel-room ranking approach for online travel agencies, while Zhao and Li [31] integrated online ratings and textual reviews from multiple websites to support hotel selection. Ray and Garain [32] combined sentiment analysis and aspect categorization in an ensemble-based hotel recommender system. Xia, An [33] subsequently incorporated multidimensional information into a deep neural network model for smart hotel recommendation. These studies demonstrate that extracting attribute-level information from online reviews can provide a richer and more interpretable basis for hotel recommendation than relying exclusively on overall ratings.

More recent research has shifted from general hotel rankings toward personalized, context-aware, and multi-attribute recommendation. Wang and Chen [34] combined collaborative filtering with probabilistic linguistic term sets to represent uncertain customer evaluations. Zhang and Cheng [35] developed a review-driven hotel recommendation mechanism that considers cooperative and interactive relationships among multiple hotel attributes. Pu and Xu [12] proposed a two-stage hotel recommendation model that extracts both public and personalized preferences from online reviews and incorporates customers' risk attitudes into hotel filtering and ranking. Liu and Dang [36] further considered attribute importance and group consensus in an adaptive online–offline hotel recommendation system. Collectively, these studies indicate that hotel recommendation is evolving from rating-based matching toward review-driven approaches that incorporate attribute-level preferences, contextual information, preference uncertainty, and individual heterogeneity. Nevertheless, most studies generate personalized recommendations for individual customers or provide general rankings for the entire customer population. The potential connection between review-driven customer segmentation and differentiated hotel recommendation for distinct customer segments remains insufficiently explored.

Several gaps remain in the integration of review-derived customer preferences and hotel recommendation. First, existing approaches often rely on overall ratings, attribute sentiment, review similarity, or historical behavior to infer preferences, while the relative contributions of individual attributes to predicted customer satisfaction receive less attention. Second, customer segmentation and hotel recommendation are frequently investigated as separate tasks, leaving limited understanding of how segment-level differences in attribute importance can be translated into differentiated hotel rankings. Third, aggregating preferences at the segment level may obscure differences among individual members, potentially limiting the extent to which recommendations accommodate both group-level utility and individual needs. These gaps motivate an integrated framework that connects review-level attribute-importance analysis, customer segmentation, and hotel recommendation while accounting for differences in attribute priorities and individual regret. The distinctions between this study and previous research are summarized in Table 1.

Table 1. The distinctions between this study and previous research.

Reference	Data Sources	Attribute Extraction Method	Sentiment Analysis Method	Customer Segmentation Method	Recommendation Method
Zhao and Li [31]	3773 hotel reviews	Leximancer	Stanford CoreNLP	None	Extended TOPSIS
Ray and Garain [32]	58,620 hotel reviews	Word2Vec	BERT and Random Forest	None	cosine similarity
Ahani and Nilashi [5]	6342 restaurant reviews	LDA	None	SOM	None

Table 1. *Cont.*

Reference	Data Sources	Attribute Extraction Method	Sentiment Analysis Method	Customer Segmentation Method	Recommendation Method
Sun and Mao [20]	29,216 car reviews	LDA	TF-IDF and SVM	K-means	None
Joung and Kim [7]	5045 product reviews	Word2Vec	VADER	SHAP and AP clustering	None
Pu and Xu [12]	18,721 hotel reviews	LDA	Stanford CoreNLP	None	Hotel filtering and cosine similarity
This paper	28,502 hotel reviews	BERT-LDA	RNTN	XGBoost-LIME and K-means	PLTS-VIKOR

3. The Proposed Framework

The overall structure of the proposed framework is shown in Figure 1, which includes three parts: preference mining, customer segmentation, and hotel recommendation. As shown in Figure 1, the preference extraction part aims to mine customers' attribute preferences from online reviews, which consists of attribute extraction based on the BERT-LDA model and sentiment analysis using the RNTN algorithm. Subsequently, the customer segmentation part focuses on establishing segmentation criteria and dividing customer groups, comprising three components: (1) constructing review classifiers through machine learning algorithms; (2) calculating attribute importance for individual reviews using LIME interpretable machine learning; (3) performing customer clustering via the K-Means algorithm. Finally, the hotel recommendation part is designed to provide differentiated recommendation results for distinct customer clusters, incorporating both attribute weight optimization through distance measurement and Top-N recommendation generation based on the PLTS-VIKOR method.

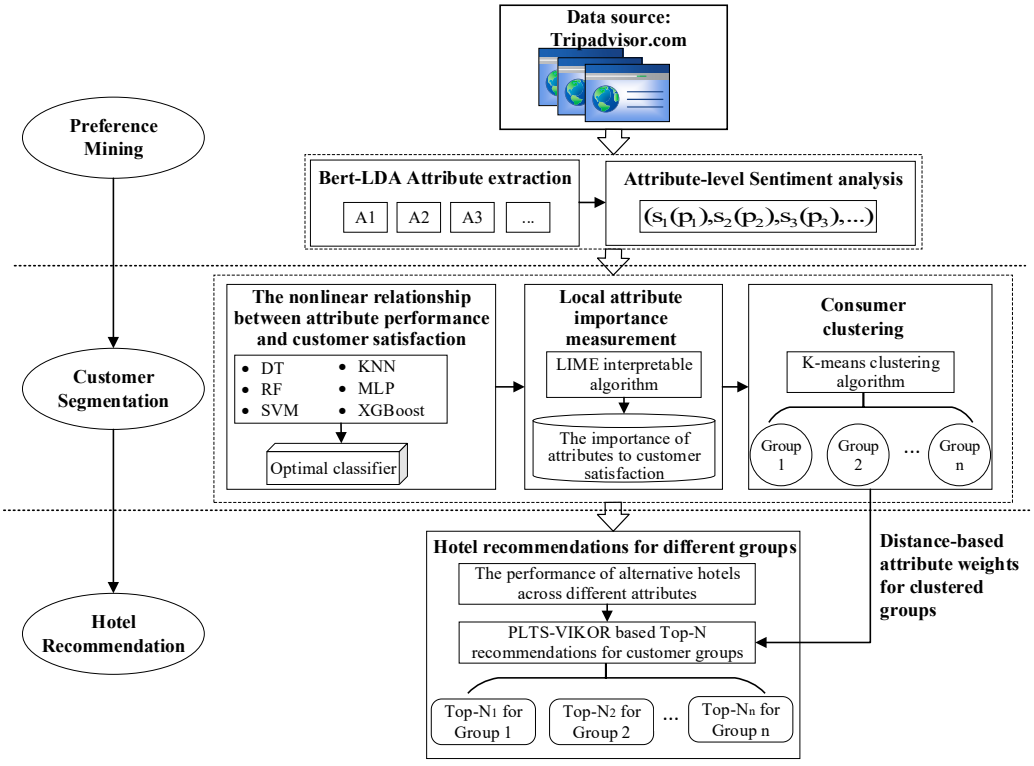


Figure 1. The overall structure of the proposed framework.

3.1. Customer Preference Extraction from Online Reviews

Preference mining aims to extract customer-focused attributes and their sentiment attitudes from online reviews, providing a basis for customer segmentation and hotel recommendations. The BERT-LDA model is used for attribute extraction, combining the powerful text representation capabilities of Bidirectional Encoder Representations from Transformers (BERT) with the topic discovery capabilities of Latent Dirichlet Allocation (LDA). Compared to traditional topic extraction models, BERT-LDA can capture contextual semantic information in complex review texts while identifying latent topics, thereby supporting more context-sensitive attribute extraction [37]. Unlike clustering-based topic models such as BERTopic, which may produce fragmented topics in some settings, BERT-LDA uses a probabilistic generative mechanism to derive interpretable attribute dimensions, which is beneficial for the subsequent customer segmentation task, where consistent and interpretable attributes are required. The process of the BERT-LDA model involves the following four steps: (1) text preprocessing, encompassing text cleaning, tokenization, stop-word removal and stemming; (2) text representation, where the pre-trained BERT model encodes texts, generating contextually relevant word embeddings; (3) dimensionality reduction, transforming high-dimensional BERT embeddings into lower dimensions; (4) LDA topic extraction, applying LDA model to the reduced embeddings to extract latent topics.

Online reviews shed light on customer preferences for specific attributes and provide sentiment analysis for these attributes. To thoroughly mine fine-grained sentiment information across different attributes, this study employs the Recursive Neural Tensor Network (RNTN) model for sentiment analysis. Although more recent pretrained language models exist, such as BERT and RoBERTa, RNTN was selected because its compositional architecture explicitly models how the meanings and sentiment polarities of words and phrases are combined according to syntactic structure [38]. This property is particularly relevant to online reviews containing negation, intensification, and sentiment contrast. Compared with approaches that primarily rely on pretrained contextual representations, RNTN provides an explicit syntactic-compositional mechanism for modeling sentiment. Moreover, the RNTN used in this study generates a standardized probability distribution over five sentiment categories (very negative, negative, neutral, positive, very positive), which can be directly incorporated into the subsequent quantitative decision-making analysis. For instance, for a review that mentions the hotel room attribute "The room is comfortable", the RNTN sentiment analysis yields [0.0011, 0.01396, 0.1742, 0.6762, 0.1345], indicating a positive sentiment orientation. When multiple sentences within the same review are associated with the same attribute, their five-category sentiment probability distributions are averaged to obtain an attribute-level sentiment distribution for that review.

To effectively utilize the sentiment orientation and corresponding probability information obtained from sentiment analysis, this study introduces the Probabilistic Linguistic Term Set (PLTS) to measure attribute performance, facilitating subsequent information integration and computation. PLTS is a tool for handling uncertain and ambiguous information, combining linguistic terms with associated probabilities to express uncertain information more flexibly and intuitive [39]. For example, the sentiment analysis result of the review "The room is comfortable" is represented in PLTS as $\{s_1(0.0011), s_2(0.01396), s_3(0.1742), s_4(0.6762), s_5(0.1345)\}$, indicating an overall positive evaluation of the room by the customer. Additionally, to facilitate the quantification of sentiment scores from reviews, this study introduces the PLTS score function [40], defined by the following formula:

$$E(L(p)) = \sum_{k=1}^{\#L(p)} r^{(k)} p^{(k)} / \#L(p) \quad (1)$$

where $L(p) = \{L^{(k)}(p^{(k)}) | k = 1,2,\ldots,\#L(p))\}$ is a PLTS, $r^{(k)}$ is the subscript of linguistic term $L^{(k)}$ and the expectation value $E(L(p)) \in [0,1]$.

3.2. Customer Segmentation Based on LIME Interpretable Algorithm

Given the nonlinear relationship between attribute performance and customer satisfaction, this study first employs machine learning algorithms to assess this relationship. Subsequently, the LIME algorithm is utilized to interpret the inputs (attribute performance) and outputs (customer satisfaction) of the machine learning model, determining the contribution of input to output for each review. Based on the LIME algorithm, the contribution of each attribute to the predicted satisfaction can be determined for each customer review.

3.2.1. Capturing Nonlinear Relationship Between Attribute Performance and Customer Satisfaction

Existing studies indicate that the relationship between attribute performance and customer satisfaction is nonlinear, and some machine learning algorithms have been used to measure this nonlinear relationship. As a result, this study first compares the performance of several common machine learning classification algorithms to identify the optimal classifier for predicting satisfaction. The detailed process is as follows.

Step 1: Model selection. This study employs six common machine learning algorithms, including DT, RF, SVM, KNN, MLP and XGBoost. Each of these is suitable for the classification problem within the context of this study and has its own advantages. Notably, DT, RF and XGBoost are tree-based models, KNN is a distance-based model, SVM is a model based on linear boundaries, and MLP is a model based on neural networks.

Step 2: Determining inputs and outputs. The inputs for this study are the sentiment scores for each review across various attributes, while the outputs are classifications of the reviews as either positive or negative. To ensure the accuracy of the classifier, customer satisfaction derived from online ratings is treated as a binary classification problem. Following Kirilenko and Stepchenkova [41], ratings of 10, 20 and 30 are considered as negative reviews, whereas ratings of 40 and 50 are considered as positive reviews during the classifier training process. Although the classification of 30-rating reviews remains controversial, previous studies have suggested that such ratings may indicate experiences approaching service failure [42]. Accordingly, this study remains consistent with previous hotel review research to identify dissatisfied customers and support alternative hotel recommendations [12].

Step 3: Model training and evaluation. During the classifier training, the ratio of positive to negative reviews is maintained at 1:1. A five-fold cross-validation method is utilized to evaluate the model, with performance metrics including Accuracy, F1 Score, Precision and Recall. Accuracy represents the proportion of correctly classified reviews among the total number of reviews. Precision indicates the proportion of reviews predicted as positive that are actually positive. Recall measures the proportion of all actual positive reviews that are correctly classified as positive by the model. The F1 Score balances both Precision and Recall, providing a comprehensive measure of the model's performance. The calculation formulas for these metrics are as follows:

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \tag{2}$$

where TP refers to the number of reviews correctly classified as positive, while TN denotes the number of reviews accurately identified as negative. FP is the number of negative reviews mistakenly labeled as positive, and FN is the number of positive reviews incorrectly categorized as negative.

$$Precision = \frac{TP}{TP + FP} \tag{3}$$

$$Recall = \frac{TP}{TP + FN} \tag{4}$$

$$F1\ Score = 2 \times \frac{Precision \times Recall}{Precision + Recall} \quad (5)$$

Step 4: Determining the optimal classifier. The performance metrics obtained from 5-fold cross-validation for the six classification algorithms are averaged and compared. The model exhibiting the best performance across these metrics is selected as the optimal classifier for review satisfaction.

3.2.2. Measuring Relative Attribute Importance with LIME Interpretable Algorithm

Since this study focuses on the impact of variations in attribute performance within individual customer reviews on customer satisfaction (i.e., local attribute importance), traditional machine learning algorithms such as DT and XGBoost, which assess global attribute importance, cannot specifically measure the importance of attributes within each review. Consequently, this study adopts the LIME algorithm for evaluating local attribute importance. Both LIME and SHAP can provide local explanations for individual predictions. SHAP has the advantage of a theoretically grounded additive attribution framework and generally provides more stable explanations, whereas LIME offers greater model independence and can be applied consistently to heterogeneous classifiers without relying on model-specific explanation procedures. Because the optimal classifier is selected empirically from six structurally different algorithms, LIME enables the study to maintain a unified local explanation procedure regardless of the selected model. In addition, its review-specific explanations are consistent with the objective of capturing heterogeneous attribute preferences at the individual customer level. The stability of the resulting explanations is further evaluated through repeated estimation and parameter-sensitivity analysis. The LIME algorithm approximates the predictions of complex models by training a simple, interpretable model in the local neighborhood of each sample to explain its prediction results. Its objective function is given as follows:

$$min_g \sum_i \pi_x(x_i')(f(x_i') - g(x_i'))^2 \quad (6)$$

where x_i' is the *ith* domain data generated around the original sample x, $f(x_i')$ is the complex model, $g(x_i')$ is the local simple model, and $\pi_x\ (x_i')$ represents the weight of the domain data. The calculation formula is shown as follows:

$$\pi_x(x_i') = exp(-\frac{D(x, x_i')^2}{\sigma^2}) \quad (7)$$

where $D(x, x_i')^2$ is the Euclidean distance between the neighborhood data x_i' and the original sample x, and σ is a parameter used to control the decay of weights.

The specific procedure for using the LIME model in this study is shown as follows: First, the reviews are transformed into attribute-based vector representations through data preprocessing. Next, the nonlinear relationship between input (multi-attribute sentiment values) and output (overall ratings) for the reviews is measured using the machine learning classification model with the best performance. Finally, LIME is employed to obtain the local surrogate coefficients for each review. The absolute values of these coefficients are used to represent the relative local influence strengths of the corresponding attributes. To improve the consistency of the attribute vectors across reviews, the absolute coefficients are normalized within each review so that their sum equals one. The resulting vectors represent review-specific relative attribute influence profiles and are subsequently used for customer clustering. These normalized values are interpreted as relative local contributions of attribute importance.

3.2.3. Customer Clustering Based on the K-Means Algorithm

This study employs the K-means algorithm to cluster the normalized, review-specific attribute influence profiles derived from LIME in Section 3.2.2. The resulting clusters are then treated as customer segments, representing similar customer preferences. The K-Means algorithm is particularly popular due to its simplicity, computational efficiency, adaptability to different datasets, and the interpretability of its results [43]. The K-Means algorithm partitions the data into K clusters by iteratively optimizing to minimize the distance between data points and their respective cluster centers. The selection of the K value is important in this algorithm, as it significantly influences the cluster performance. In this study, the optimal number of clusters K is determined using the elbow method. Specifically, we set K values ranging from 2 to 30, compute the clustering cost (Sum of Squared Errors, SSE) for each K, and plot these costs. The optimal K was determined jointly by the elbow method, where the elbow point is identified as the K at which the rate of decrease in SSE slows significantly, and the silhouette coefficient, which was used to validate clustering quality.

$$SSE = \sum_{i=1}^{K} \sum_{x \in C_i} \| x - \mu_i \|^2 \tag{8}$$

where μ_i is the center of the *ith* cluster, and C_i represents all data points in the *ith* cluster.

3.3. Hotel Recommendation with Differentiated Group Weights

This section first introduces a systematic approach for assigning differentiated attribute weights for various groups by considering group members' contributions. Subsequently, the PLTS-VIKOR method is developed for generating hotel recommendation results for different customer groups.

3.3.1. Determining Attribute Weights by Considering Group Members' Contributions

Considering that the attribute importance of customer reviews at the cluster center is more representative than those at the cluster edges within each cluster group, this study proposes an attribute weighting method by assigning customer review weights based on their distance from the cluster center, with a gradual decrease in importance as the distance increases. The calculation process is presented as follows.

Step 1: Calculating the distance of each customer review to its cluster center.

$$d_{ik} = \sqrt{\sum_{j=1}^{m} (x_{ij} - c_{kj})^2} \tag{9}$$

where x_{ij} is the importance value of the customer review i on the attribute j, c_{kj} represents the importance value of the cluster center k on the attribute j, and m is the number of attributes.

Step 2: Computing the customer review weights using the distance d_{ik}.

$$w_i = \exp(-d_{ik}) \tag{10}$$

where the exponential decay function is adopted to assign review weights according to their distances from the cluster center. The underlying rationale is that reviews closer to the center are more representative of the segment-level attribute influence profile and therefore receive greater weights, whereas the contributions of reviews decrease smoothly as their distances increase.

Step 3: Calculating the weighted average value of all customer reviews on each attribute.

$$\overline{x}_{kj} = \frac{\sum_{i \in C_k} w_i x_{ij}}{\sum_{i \in C_k} w_i} \tag{11}$$

where C_k is the set of all customer reviews belonging to the cluster k.

Step 4: Normalizing the weighted average values.

$$\tilde{x}_{kj} = \frac{\overline{x}_{kj}}{\sum_{j=1}^{m} \overline{x}_{kj}} \tag{12}$$

where $\sum_{j=1}^{m} \tilde{x}_{kj} = 1$.

3.3.2. Hotel Recommendation Based on the PLTS-VIKOR Method

This study employs the VIKOR method with the PLTS environment for hotel recommendations. The PLTS-VIKOR method effectively balances maximizing the group utility for the majority while minimizing individual regret for the opposition, thereby making the results of multi-attribute decision-making problems more reasonable. The PLTS-VIKOR method consists of the following three key steps.

Step 1: Constructing the decision-making matrix M.

$$M = [y_{ij}]_{m \times n} \tag{13}$$

where y_{ij} is the performance of the *ith* hotel under the *jth* attribute, and y_{ij} can be calculated using the PLTS score function in the Formula (1).

Step 2: Determining the positive ideal solution A^+ and the negative ideal solution A^-.

$$A^+ = \{y_1^+, y_2^+, \ldots, y_n^+\} \tag{14}$$

$$A^- = \{y_1^-, y_2^-, \ldots, y_n^-\} \tag{15}$$

where $y^+{}_j = \max_i y_{ij}$ and $y^-{}_j = \min_i y_{ij}$.

Step 3: Calculating the distance of each hotel from the positive ideal solution (S_{ik}) and the negative ideal solution (R_{ik}).

$$S_{ik} = \sum_{j=1}^{n} \tilde{x}_{kj} \frac{y_j^+ - y_{ij}}{y_j^+ - y_{ij}^-} \tag{16}$$

$$R_{ik} = \max_j \left(\tilde{x}_{kj} \frac{y_j^+ - y_{ij}}{y_j^+ - y_{ij}^-}\right) \tag{17}$$

where $\tilde{x}_{kj}$ is the weight of the *jth* attribute under the *kth* group.

Step 4: Calculating the comprehensive index Q_{ik}.

$$Q_{ik} = v\left(\frac{S_{ik} - S_{ik}^-}{S_{ik}^+ - S_{ik}^-}\right) + (1 - v)\left(\frac{R_{ik} - R_{ik}^-}{R_{ik}^+ - R_{ik}^-}\right) \tag{18}$$

where $S_{ik}^+ = \max_i S_{ik}$, $S_{ik}^- = \min_i S_{ik}$, $R_{ik}^+ = \max_i R_{ik}$, $R_{ik}^- = \min_i R_{ik}$ and v is a weight coefficient, typically set to 0.5. The smaller the Q_{ik} value, the better the performance of the hotel i under the *kth* group weight.

4. Case Study

4.1. Data Collection and Preprocessing

The data used in this study were obtained from TripAdvisor, a globally recognized travel platform that provides customer feedback through online reviews. A Python 3.13 based web scraping program was used to collect online hotel reviews in Manchester, UK. Manchester was selected because it offers a diverse hotel market with a substantial volume of online customer feedback. The data collection was conducted between 12 April 2025 and 27 April 2025. The collected data comprised hotel information (name, location, overall rating, number of reviews, and multi-attribute ratings), reviewer information (nickname, address, contributions, and helpful votes), and review details (title, comment, date of stay,

trip type, overall score, multi-attribute ratings, and helpful votes). Figure 2 presents an example of a customer review on TripAdvisor.

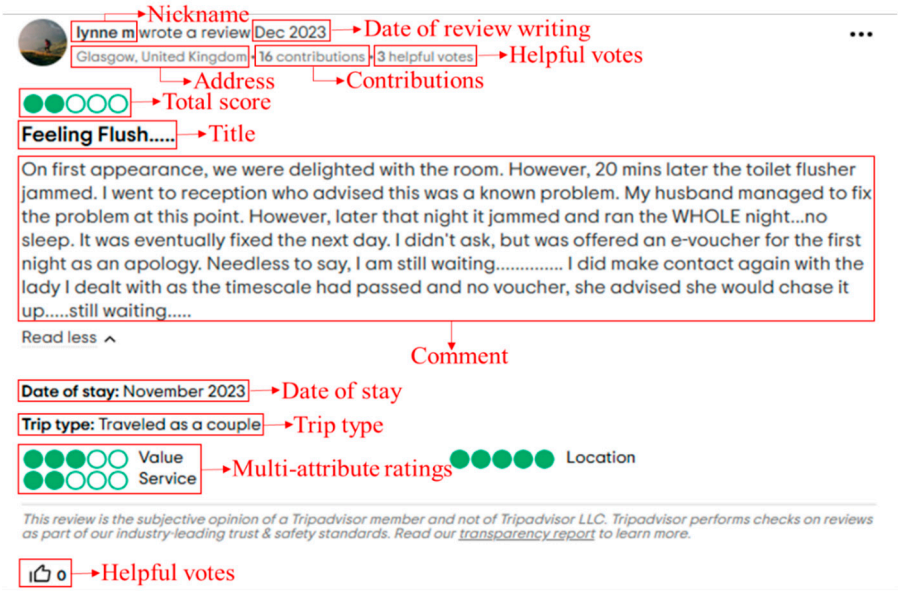


Figure 2. An example of a customer review on TripAdvisor.

The data preprocessing procedure consisted of two main stages: hotel selection and review screening. In the first stage, eligible hotels were identified according to predefined selection criteria. In the second stage, reviews associated with the selected hotels were screened to construct the final dataset for analysis. Hotels were selected based on three criteria: (1) they were classified as 4- or 5-star hotels in Manchester, to focus on hotels within relatively comparable quality categories; (2) their overall ratings on TripAdvisor were greater than 4.0, indicating relatively high customer evaluations; (3) they had more than 100 English-language online reviews, ensuring a sufficient volume of textual data for analysis. Applying these criteria resulted in a final selection of 52 hotels for the hotel recommendation analysis. In the review screening stage, the 32,267 online reviews associated with the 52 selected hotels, covering the period from April 2003 to March 2025. Data preprocessing was conducted in three stages. First, reviews containing fewer than 10 words were removed [23]. Second, non-English reviews were identified and excluded. Third, duplicate reviews were removed. After these steps removed 3765 reviews, the final dataset comprised 28,502 reviews for further analysis, as shown in Table 2.

Table 2. Descriptive statistics of the case-study dataset.

Metric	Hotels	Metric	Online Reviews	Metric	Total Scores
Total Number	52	Total Number	28,502	Total Number	28,502
Maximum Number of Overall Rating	5	Average Number of Words	127.03	Average Score	3.24
Minimum Number of Overall Ratings	4	Median Number of Words	134	Median Score	4
Maximum Number of Star	5	Maximum Number of Words	2198	Maximum Score	5
Minimum Number of Star	4	Minimum Number of Words	10	Minimum Score	1

4.2. Customer Preference Extraction

Based on the BERT-LDA method described in Section 3.1, with the number of topics set to 9, this study extracted 9 attributes and corresponding keywords from the hotel reviews, as detailed in Table 3. These attributes are Value (A1), Room (A2), Location (A3), Service (A4), Cleanliness (A5), Facility (A6), Food (A7), Transportation (A8) and Sleep (A9).

Table 3. Hotel attributes along with corresponding keywords identified through Bert-LDA.

Attribute	Representative Keywords	Numbers
Value (A1)	Value, price, expense, spend, bonus, charge, cost, credit, pay, money	16
Room (A2)	Room, bath, suite, space, bedroom, decoration, toilet, furniture, door	25
Location (A3)	Location, nearby, shop, site, market, walk, situation, close, central	22
Service (A4)	Service, book, manage, luggage, treat, staff, attitude, front desk, reception	26
Cleanliness (A5)	Cleanliness, housekeeping, wash, sweep, carpet, towel, smell, vacuum	21
Facility (A6)	Facility, swimming, gym, laundry, fitness, equipment, club, bar, wifi	23
Food (A7)	Food, meal, breakfast, water, taste, buffet, dish, drink, coffee, snack...	16
Transportation (A8)	Transportation, subway, station, airport, flight, taxi, car, traffic, drive, shuttle	15
Sleep (A9)	Sleep, wake, asleep, bed, sheet, quilt, night, blanket, bedding, pillow	12

Using the RNTN algorithm mentioned in Section 3.1, sentiment analysis results for each review on different attributes can be obtained. For instance, as shown in Table 4, the review mentions three attributes: Room (A2), Location (A3), and Service (A4). The sentiment analysis results for these attributes are Negative (A2), Positive (A3), and Negative (A4), respectively.

Table 4. Sentiment analysis result of an online review.

Online Review	Good location. However, staff were more keen to chat with each other and ignore multiple guest requests as our room was missing a TV remote.
Attributes	Room (A2), Location (A3), Service (A4)
Sentiment Distribution of Room (A2)	[0.1283, 0.3787, 0.3051, 0.1718, 0.01591]
PLTS Presentation of Room (A2)	$\{s_1(0.1283), s_2(0.3787), s_3(0.3051), s_4(0.1718), s_5(0.0159)\}$
Sentiment Orientation	Negative

4.3. Customer Segmentation

The customer segmentation process involves three key steps: developing the relationship between attribute performance and customer satisfaction to identify the optimal classifier, evaluating the importance of each attribute in individual reviews using the LIME algorithm, and classifying customers using the K-means algorithm. Firstly, a five-fold cross-validation method was used to evaluate the performance of the DT, RF, SVM, KNN, MLP and XGBoost using accuracy, F1 score, precision, and recall. Based on the results reported in Table 5, XGBoost demonstrated the best overall performance and was selected as the optimal classifier.

Secondly, the trained XGBoost algorithm is used to predict the nonlinear relationship between attribute performance and customer satisfaction. The LIME algorithm is then employed to interpret the model predictions and estimate the local importance of each attribute for individual reviews. For subsequent customer segmentation, a subset of 4000 reviews from dissatisfied customers (i.e., those with ratings of 30 or below) was randomly selected from the full dataset. The resulting normalized attribute-importance profiles were used as inputs to the K-means algorithm for customer segmentation and to support alternative hotel

recommendations. Figure 3 illustrates the LIME results for a randomly selected review. For this customer, Service (A4) is the most important attribute, followed by Cleanliness (A5), whereas Room (A2) has the lowest importance.

Table 5. The performance of different machine learning classification algorithms.

Algorithms	Accuracy	F1 Score	Precision	Recall
DT	0.8264	0.8295	0.8254	0.8231
RF	0.8740	0.8737	0.8753	0.8741
SVM	0.8718	0.8717	0.8728	0.8718
KNN	0.8234	0.8229	0.8270	0.8234
MLP	0.8686	0.8730	0.8751	0.8736
XGBoost	0.8781	0.8780	0.8795	0.8781

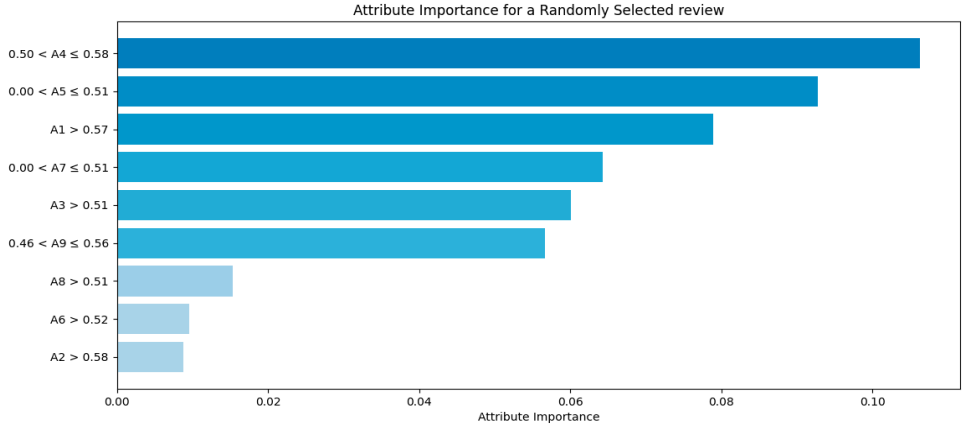


Figure 3. The results for a randomly selected review based on the LIME algorithm.

After extracting attribute importance from customer reviews, the K-means clustering algorithm is employed to group customers with similar attribute preferences. To determine the appropriate number of clusters, the elbow method is first applied, followed by an examination of the silhouette coefficient as a complementary clustering validity measure. As shown in Figure 4, the elbow plot exhibits a relatively noticeable bend at K = 6. Moreover, the silhouette coefficient at K = 6 is 0.4381, which is higher than those at the neighboring values, indicating a reasonable balance between within-cluster cohesion and between-cluster separation. Based on these two criteria, 6 clusters are selected for subsequent analysis.

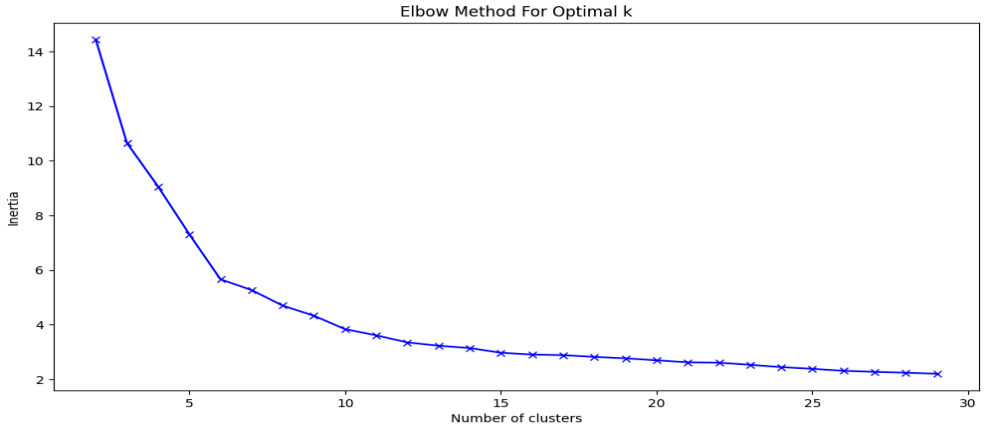


Figure 4. Optimal number of clusters using the elbow method.

When K = 6, the K-means clustering results are illustrated in Figure 5. This figure displays a scatter plot of the data reduced to two dimensions using PCA. The clusters are represented by 6 different colors: blue, orange, green, purple, cyan and magenta, where each corresponding to a different cluster category. The red markers indicate the cluster centers. The number of customer reviews in clusters 0, 1, 2, 3, 4 and 5 are 367, 861, 927, 917, 480 and 448, respectively.

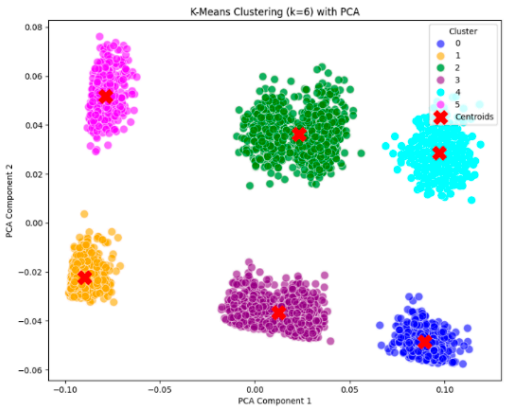


Figure 5. K-means clustering of attribute importance with PCA.

4.4. Hotel Recommendation

Each cluster is treated as a customer segment for calculating the attribute importance. Using the formula from Section 3.3.1, attribute weights for each customer segment are derived, as shown in Table 6. The results indicate a differentiated distribution of attribute importance across different customer segments. Specifically, Group 1 considers the importance of Service (A4) to be significantly higher than other attributes, while Group 2 and Group 6 prioritize Room (A2) and Service (A4). Group 3 regards Cleanliness (A5) and Room (A2) as relatively important, and Groups 4 and 5 view Room (A2) as a key attribute affecting overall satisfaction. Overall, Room (A2) and Service (A4) are relatively important across all customer segments, indicating that the core customer needs are the comfort of the room and the convenience of service.

Table 6. Differentiated attribute weights across customer segments.

Group Number	A1	A2	A3	A4	A5	A6	A7	A8	A9
Group 1	0.1363	0.0679	0.0855	0.2845	0.1805	0.0388	0.0973	0.0225	0.0867
Group 2	0.1139	0.2272	0.0737	0.2331	0.1866	0.0418	0.0128	0.0195	0.0914
Group 3	0.1290	0.2341	0.0984	0.0164	0.2330	0.0562	0.0910	0.0261	0.1157
Group 4	0.0861	0.3125	0.0608	0.1966	0.1653	0.0338	0.0583	0.0158	0.0707
Group 5	0.0684	0.4102	0.0791	0.0122	0.2215	0.0479	0.0469	0.0204	0.0935
Group 6	0.1050	0.1967	0.0673	0.2087	0.1565	0.0371	0.1295	0.0168	0.0824

After obtaining the attribute weights for different clusters, the VIKOR method is employed to provide differentiated hotel recommendation lists for each customer segment. Based on the sentiment analysis algorithm outlined in Section 3.1, sentiment analysis results for the hotel reviews are obtained. And then, using formula (1), the aggregated PLTS results for hotel attribute performance are calculated. For example, the attribute performance for hotel H1 is detailed in Table 7. Table 7 indicates that the attributes Service (A4) and Value (A1) exhibit the best performance for this hotel, while the performance of Transportation (A8) is relatively inferior.

Table 7. The performance of Hotel 1 in the 9 attributes.

Attribute	PLTS Results	Performance
Value (A1)	$\{s_1(0), s_2(0.2214), s_3(0.2214), s_4(0.55), s_5(0.0071)\}$	0.6685
Room (A2)	$\{s_1(0), s_2(0.314), s_3(0.2558), s_4(0.4264), s_5(0.0039)\}$	0.6241
Location (A3)	$\{s_1(0), s_2(0.2432), s_3(0.2432), s_4(0.4955), s_5(0.018)\}$	0.6576
Service (A4)	$\{s_1(0), s_2(0.2216), s_3(0.1796), s_4(0.5808), s_5(0.018)\}$	0.6790
Cleanliness (A5)	$\{s_1(0), s_2(0.3065), s_3(0.2097), s_4(0.4516), s_5(0.0323)\}$	0.6420
Facility (A6)	$\{s_1(0), s_2(0.275), s_3(0.225), s_4(0.475), s_5(0.025)\}$	0.6500
Food (A7)	$\{s_1(0), s_2(0.2676), s_3(0.2676), s_4(0.4577), s_5(0.007)\}$	0.6408
Transportation (A8)	$\{s_1(0), s_2(0.4074), s_3(0.3704), s_4(0.2222), s_5(0)\}$	0.5630
Sleep (A9)	$\{s_1(0), s_2(0.243), s_3(0.2804), s_4(0.4579), s_5(0.0187)\}$	0.6505

Finally, the VIKOR method discussed in Section 3.3.2 is utilized for hotel recommendations. Based on the performance of each alternative hotel across various attributes and the differentiated attribute weights for each customer segment, the Top-10 hotel recommendations for each segment are derived in Table 8.

Table 8. Differentiated hotel recommendation results for each customer segment.

Group Name	Hotel Ranking
Group 1	H39 > H50 > H40 > H15 > H18 > H25 > H1 > H33 > H47 > H37
Group 2	H15 > H18 > H39 > H1 > H40 > H25 > H33 > H50 > H19 > H30
Group 3	H15 > H18 > H1 > H19 > H20 > H40 > H39 > H25 > H50 > H30
Group 4	H18 > H15 > H40 > H39 > H1 > H19 > H33 > H25 > H20 > H50
Group 5	H18 > H19 > H20 > H15 > H1 > H40 > H32 > H39 > H25 > H33
Group 6	H18 > H39 > H40 > H15 > H25 > H1 > H50 > H33 > H47 > H19

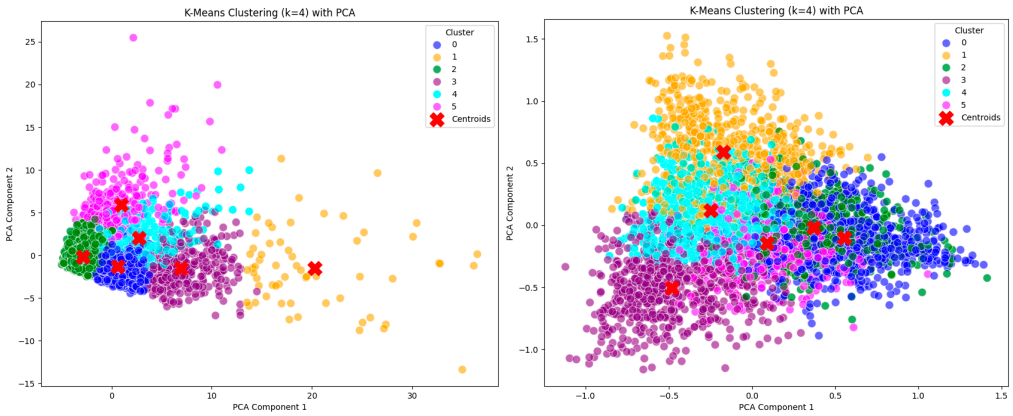
Table 8 indicates that the hotel recommendations differ based on the varying preferences for hotel attributes among the customer segments. For Group 1, which places a significant emphasis on Service (A4), the top three recommended hotels are H39 > H50 > H40. For Groups 2 and 3, which prioritize Room (A2), the recommendations show some similarities, with the top three hotels being H15 > H18 > H39 and H15 > H18 > H1, respectively. For Groups 4, 5 and 6, the top recommended hotel is H18. Overall, hotel H18 demonstrates superior performance and is thus more frequently prioritized in the recommendations compared to other hotels.

5. Result Comparison and Discussion

5.1. Effectiveness of Different Customer Segmentation Methods

In the process of customer segmentation based on online reviews, establishing clustering bases is a significant factor contributing to the variability in segmentation results. Previous studies have primarily relied on three metrics as the basis for clustering: multi-attribute ratings directly provided by customers (M1) [5], attribute frequencies based on reviews (M2) [44], and attribute TF-IDF values based on reviews (M3) [6]. The M1 metric suffers from a pronounced data sparsity issue, as few customers rate each attribute's performance comprehensively. For instance, as shown in Figure 2, the customer only rated the attributes of Value, Location and Service, making it challenging to determine sentiment scores for unmentioned attributes. Additionally, M1 only accounts for customers' sentiments without considering the importance of attributes. The M2 metric considers attribute importance based on frequency, whereas the M3 metric evaluates attribute importance by incorporating both attribute frequency and inverse document frequency. Given that M1

suffers from data missing issues in the dataset used in this study, M2 and M3 were selected for a comparative analysis with the customer segmentation in this case. The results are illustrated in Figure 6.



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

Figure 6. (**a**) Customer segmentation results using Method M2. (**b**) Customer segmentation results using Method M3.

From Figures 5 and 6, it can be observed that the application of the LIME explainable machine learning algorithm yields the more effective customer segmentation than the frequency-based and TF-IDF-based methods. Specifically, Figure 6a, which portrays customer segmentation based on attribute frequency, exhibits a lack of discernibility among clusters, particularly with Clusters 1 and 5 displaying marked dispersion and poor clustering results. Conversely, Figure 6b, depicting customer segmentation based on attribute TF-IDF values, reveals instances of overlap between Clusters 0, 2, and 5, as well as indistinct distinctions between clusters and a notable dispersion of data points within some clusters, leading to suboptimal segmentation results. In contrast, the customer segmentation results presented in Figure 6 demonstrate clearer distinctions between groups and more concentrated data points within each cluster. To quantitatively evaluate these differences, the Silhouette Coefficient was calculated for all three methods under the same setting of clusters. The LIME-based method achieved a Silhouette Coefficient of 0.4381, substantially higher than those of the frequency-based method (0.2104) and the TF-IDF-based method (0.1678). These results indicate that the LIME-based method achieves better cluster cohesion and separation in this case study.

From a theoretical perspective, these findings suggest that review-based customer segmentation can be informed not only by the frequency with which customers mention particular attributes but also by the relative contributions of these attributes to model-predicted satisfaction. Compared with frequency- and TF-IDF-based representations, LIME-derived attribute-importance profiles provide an alternative basis for identifying differences in attribute-importance patterns across reviews. This finding extends existing review-based segmentation approaches by introducing an interpretable, prediction-oriented perspective on customer heterogeneity.

5.2. Impact of Attribute Weights on Hotel Recommendations Across Customer Segments

In this study, the attribute preferences of different customer groups are used as the basis for differentiated hotel recommendations. Consequently, it is essential to investigate how changes in attribute weights affect the recommendation results. In the context of this case study, the simulation experiment was conducted by randomly generating 1000 sets of attribute weights using Python. The variation in hotel rankings under different weight

settings is illustrated in Figure 7. In the figure, each row represents an alternative hotel, with colors ranging from dark blue to yellow indicating higher hotel rankings. Greater color fluctuation within a row means a more significant impact of different weight settings on hotel rankings.

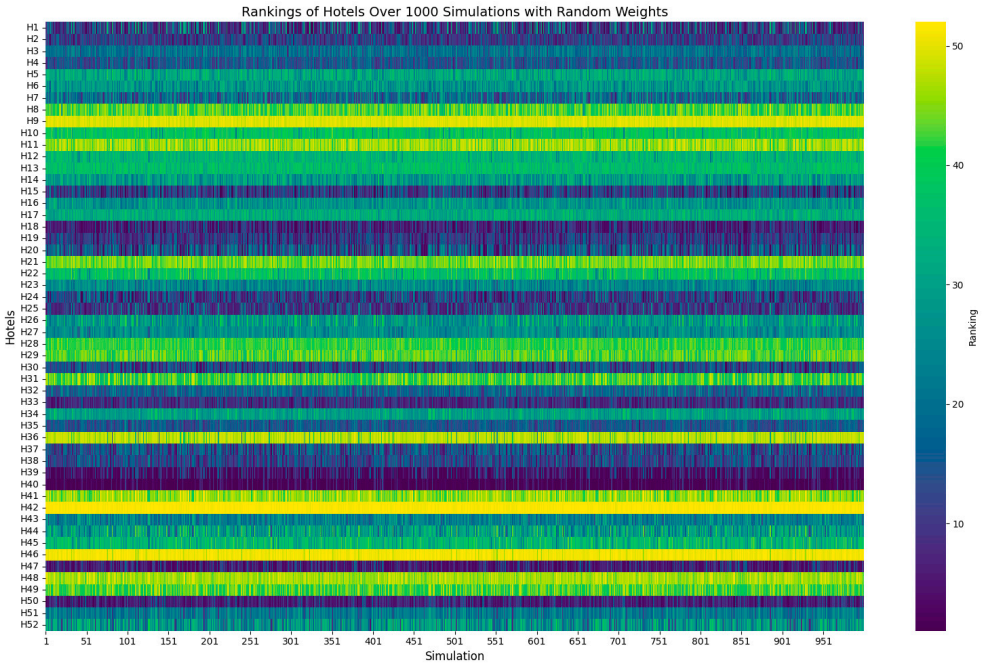


Figure 7. Variation in hotel ranking under different attribute weight settings.

Based on the color variations in the heatmap shown in Figure 7, it can be observed that changes in attribute weights have a significant impact on hotel rankings, indicating the necessity of tailoring hotel recommendations according to the weight preferences of different customer segments. Specifically, hotels H1, H32 and H41 exhibit substantial color variations across different experiments, suggesting that their rankings are more sensitive to changes in attribute weights. Conversely, hotels H40 and H42 display minimal color variation, indicating that their rankings are relatively stable. Additionally, the overall color for hotels H42, H46 and H9 tends towards yellow, reflecting their relatively poorer overall performance across different simulations. In contrast, hotels H15, H18 and H40 exhibit a color trend towards blue, indicating consistently better overall performance in various experiments. Overall, changes in attribute weights directly influence hotel rankings and recommendation outcomes. By incorporating the attribute-importance profiles derived from the proposed LIME-based customer segmentation method into the recommendation process, the framework accounts for differences in attribute priorities across customer segments. The observed ranking variations highlight the importance of considering segment-specific attribute preferences when generating hotel recommendations, providing further support for the value of integrating interpretable customer segmentation with multi-attribute hotel recommendation to better accommodate heterogeneous customer needs.

5.3. Comparative Performance of Hotel Recommendation Methods

To assess the effectiveness of the PLTS-VIKOR method, this study compares it with several other commonly used multi-attribute decision-making methods in PLTS environment, namely PLTS-TOPSIS (W1) [40], PLTS-PROMETHEE II (W2) [45] and PLTS-COCOSO (W3) [46]. These methods were selected because they represent different decision principles, including distance-based evaluation, outranking relations, and compromise aggregation. The methods are compared from two perspectives. First, Spearman's rank correlation coeffi-

cient and the Top-10 overlap rate are used to examine the consistency between the rankings generated by the proposed method and those produced by the comparison methods. Spearman's coefficient evaluates the consistency of the complete rankings of the 52 hotels, whereas the Top-10 overlap rate measures the proportion of highly ranked hotels identified by both methods. Second, group utility loss, maximum individual regret, and worst individual regret are employed to evaluate the recommendation outcomes. Group utility loss represents the aggregate performance gap between the selected hotels and the ideal solution across all attributes. Maximum individual regret captures the average worst attribute-level performance gap among the Top-10 hotels, while worst individual regret represents the largest regret observed within the Top-10 recommendation list. Lower values for these three indicators indicate more desirable decision outcomes. Table 9 reports the average results across the 6 customer groups.

Table 9. Comparison of hotel-ranking methods across 6 customer groups.

Method	Spearman's Coefficients	Top-10 Overlap	Group Utility Loss	Maximum Individual Regret	Worst Individual Regret
This study	—	—	0.1706	0.0563	0.0758
W1	0.9937	90.00%	0.1663	0.0584	0.0823
W2	0.9840	88.33%	0.1652	0.0593	0.0822
W3	0.9860	88.33%	0.1652	0.0593	0.0822

Notes: Spearman's ρ and the Top-10 overlap rate are calculated relative to the proposed PLTS-VIKOR method. The reported values are averages across the six customer segments. Lower group utility loss, maximum individual regret, and worst individual regret indicate better performance.

The results show a high degree of consistency between PLTS-VIKOR and the comparison methods. Spearman's coefficients range from 0.9840 to 0.9937, while the Top-10 overlap rates range from 88.33% to 90.00%, indicating that the overall rankings and recommended hotel lists are relatively stable across the four MCDM methods in this case study. Although PLTS-VIKOR produces a slightly higher group utility loss, it achieves the lowest maximum individual regret and worst individual regret. This indicates that it accepts a limited reduction in overall utility to reduce the risk of poor performance on attributes important to each customer group. Therefore, PLTS-VIKOR provides a more balanced compromise between overall group utility and unfavorable attribute-level outcomes, supporting its suitability for group recommendation.

6. Conclusions

6.1. Theory and Practice Implications

Online reviews contain rich information on customer preferences. This study uses the LIME interpretable machine learning algorithm to extract customer attribute preferences from each review and constructs an integrated framework for customer segmentation and hotel recommendation. The case study and comparative analysis results indicate that, among various machine learning algorithms, trained XGBoost excels in capturing the nonlinear relationships between attributes and customer satisfaction in online reviews. In addition, the attribute importance derived from the LIME interpretable machine learning algorithm for customer segment outperforms previous methods solely relying on attribute frequency or TF-IDF values. Finally, the proposed hotel recommendation method based on customer segmentation demonstrates a different trade-off between group utility and individual regret compared with the alternative methods.

This study makes several theoretical contributions. Firstly, it contributes an integrated framework linking online review-based customer segmentation with hotel recommendation, which provides a framework for need-based customer classification and recommendation strategies. Secondly, unlike previous customer segmentation approaches that rely on online ratings or attribute frequency, this study quantifies the nonlinear relationships between attribute performance and customer satisfaction using machine learning models, and assesses the importance of attributes in each review through the LIME algorithm, which forms the foundation for a more accurate and interpretable customer segmentation. Thirdly, in the hotel recommendation stage, a distance-based method is proposed to measure the influence of group members, which determines differentiated attribute weights for each segment. Fourthly, this study contributes to group recommendation research by applying PLTS-VIKOR to evaluate hotel alternatives while considering both group utility and individual regret.

The findings of this study have significant practical implications. Firstly, this study introduces a novel framework by integrating customer segmentation with hotel recommendations, offering a new perspective for hospitality industries to identify customer groups based on customer needs. This proposed approach provides valuable insights for offering differentiated services or product recommendations to various segments. Secondly, by quantifying attribute importance in online reviews using the LIME algorithm, this study improves traditional segmentation methods based on ratings or attribute frequency. This improvement enables hotel managers to identify key factors for different customer segments, fostering targeted service improvements and innovations that cater to these segments' unique preferences. Finally, the applicability of the proposed framework extends beyond the hospitality sector, encompassing domains such as e-commerce, online education and healthcare. It offers practical guidance for optimizing group activities and formulating precise marketing strategies in these fields, thereby elevating customer service levels across related industries and fostering data-driven decision-making and resource allocation efficiencies.

6.2. Limitations and Future Directions

This study has several limitations that provide opportunities for future research. First, the proposed framework focuses on needs-based customer segmentation using online reviews while overlooking other potentially informative customer characteristics, such as demographic information (e.g., age, gender, and location) and behavioral data. Future research could integrate multiple data sources to achieve more comprehensive customer profiling and segmentation. Second, although the proposed framework employs the LIME algorithm to provide interpretable attribute importance for customer segmentation, future studies could investigate more advanced explainable models, such as attention-based deep learning methods or graph neural networks, to capture more complex relationships among customer needs and service attributes. Third, this study adopts an exponential decay function to determine review weights based on their distances from cluster centers. Future research could compare it with alternative weighting functions, such as linear and Gaussian functions, to examine their effects on group-level attribute weights and recommendation results. Finally, although this study employs BERT-LDA and RNTN for attribute extraction and fine-grained sentiment analysis, respectively, it does not compare their performance with recent instruction-tuned large language models (LLMs) or zero-shot aspect extraction methods. Future research could benchmark these approaches to evaluate their relative effectiveness and explore their potential to improve the accuracy and robustness of review-based customer segmentation and hotel recommendation.

Funding: The work was supported by the Sichuan Office of Philosophy and Social Science (Grant No. SCJJ26ND388), the Chengdu University of Technology (Grant No. YJ2026-QN007), and Yibin Municipal Federation of Social Sciences (Grant No. YB26Y19).

Informed Consent Statement: Not applicable.

Data Availability Statement: The data availability is restricted. The data were derived from TripAdvisor, and can be obtained from https://www.tripadvisor.com/ (accessed on 21 September 2026) with permission from TripAdvisor.

Conflicts of Interest: The author declares no conflicts of interest.

References

1. Zhang, Y.; Li, X.K. Exploring asymmetric effects of service attributes of robots on customer satisfaction in hotels. *Inf. Technol. Tour.* **2026**, *28*, 23. [CrossRef]
2. Drozd, D.; Söilen, K.S. AI Labels, Perceived Authenticity, and Consumer Trust in User-Generated Reviews. *J. Theor. Appl. Electron. Commer. Res.* **2026**, *21*, 154. [CrossRef]
3. Apostol, E.S.; Pisica, A.G.; Truica, C.O. ATESA-BÆRT: A heterogeneous ensemble learning model for Aspect-Based Sentiment Analysis. *Knowl.-Based Syst.* **2025**, *326*, 113987. [CrossRef]
4. Liu, C.T.; Chen, T.; Pu, Q.; Jin, Y. Text Mining for Consumers' Sentiment Tendency and Strategies for Promoting Cross-Border E-Commerce Marketing Using Consumers' Online Review Data. *J. Theor. Appl. Electron. Commer. Res.* **2025**, *20*, 125. [CrossRef]
5. Ahani, A.; Nilashi, M.; Ibrahim, O.; Sanzogni, L.; Weaven, S. Market segmentation and travel choice prediction in SPA hotels through TripAdvisor's online reviews. *Int. J. Hosp. Manag.* **2019**, *80*, 52–77. [CrossRef]
6. Marandi, A.; Tasavori, M.; Najmi, M. New insights into hotel customer's revisiting intentions, based on big data. *Int. J. Contemp. Hosp. Manag.* **2024**, *36*, 292–311. [CrossRef]
7. Joung, J.; Kim, H. Interpretable machine learning-based approach for customer segmentation for new product development from online product reviews. *Int. J. Inf. Manag.* **2023**, *70*, 102641. [CrossRef]
8. Caber, M.; Albayrak, T.; Loiacono, E.T. The classification of extranet attributes in terms of their asymmetric influences on overall user satisfaction: An introduction to asymmetric impact-performance analysis. *J. Travel Res.* **2013**, *52*, 106–116. [CrossRef]
9. Qin, Y.; Luo, C.; Ngai, E.W.T. Deconstructing customer satisfaction recipes: A dynamic configurational framework leveraging the power of online reviews in tourism contexts. *Tour. Manag.* **2025**, *110*, 105181. [CrossRef]
10. Liu, Y.H.; Li, Y.L. Gaining insights for airline satisfaction improvement using interpretable machine learning: Considering customers' interactive perception of service features. *Curr. Issues Tour.* **2026**. [CrossRef]
11. Rostami, M.; Berahmand, K.; Forouzandeh, S.; Ahmadian, S.; Farrahi, V.; Oussalah, M. A novel healthy food recommendation to user groups based on a deep social community detection approach. *Neurocomputing* **2024**, *576*, 127326. [CrossRef]
12. Pu, Z.M.; Xu, Z.S.; Zhang, C.X.; Zeng, X.-J.; Gan, W.D. An online review-driven two-stage hotel recommendation model considering customers' risk attitudes and personalized preferences. *Omega* **2025**, *131*, 103197. [CrossRef]
13. Haley, R.I. Benefit Segmentation: A Decision-oriented Research Tool. *J. Mark.* **1968**, *32*, 30–35. [CrossRef]
14. Kyriakidis, A.; Tsafarakis, S. MUSAsent: Integrating aspect-based sentiment into Multicriteria Satisfaction Analysis. *Omega* **2026**, *143*, 103552. [CrossRef]
15. Yang, S.S.; Liao, H.C.; Koczy, L.T. Preference mining and fuzzy inference for hotel selection based on aspect-based sentiment analysis from user-generated content. *J. Oper. Res. Soc.* **2025**, *76*, 1414–1431. [CrossRef]
16. Salazar-Santander, C.; Mac Cawley, A.F.; Martinez-Troncoso, C. An optimal effectiveness-driven target segment selection modeling approach for marketing campaign management. *Comput. Ind. Eng.* **2025**, *202*, 110945. [CrossRef]
17. Smith, W.R. Product differentiation and market segmentation as alternative marketing strategies. *J. Mark.* **1956**, *21*, 3–8. [CrossRef]
18. Zuo, M.H.; Angelopoulos, S.; Liang, Z.Y.; Ou, C.X.J. Blazing the Trail: Considering Browsing Path Dependence in Online Service Response Strategy. *Inf. Syst. Front.* **2023**, *25*, 1605–1619. [CrossRef]
19. Su, Y.F.; Cai, X.D.; Li, T. FICLRec: Frequency enhanced intent contrastive learning for sequential recommendation. *Inf. Process. Manag.* **2025**, *62*, 104231. [CrossRef]
20. Sun, B.; Mao, H.Y.; Yin, C.S. How to identify product defects and segment consumer groups on an online auto forum. *Int. J. Consum. Stud.* **2022**, *46*, 2270–2287. [CrossRef]
21. Wu, P.; Tang, T.; Zhou, L.; Martinez, L. A decision-support model through online reviews: Consumer preference analysis and product ranking. *Inf. Process. Manag.* **2024**, *61*, 103728. [CrossRef]
22. Shin, S.; Nicolau, J.L. Identifying attributes of wineries that increase visitor satisfaction and dissatisfaction: Applying an aspect extraction approach to online reviews. *Tour. Manag.* **2022**, *91*, 104528. [CrossRef]

23. Kim, S.; Lee, S.; McCulloch, R. A Topic-Based Segmentation Model for Identifying Segment-Level Drivers of Star Ratings from Unstructured Text Reviews. *J. Mark. Res.* **2024**, *61*, 1132–1151. [CrossRef]
24. Nessel, K.; Kosciólek, S.; Wszendybyl-Skulska, E.; Kopera, S. Benefit segmentation in the tourist accommodation market based on eWOM attribute ratings. *Inf. Technol. Tour.* **2021**, *23*, 265–290. [CrossRef]
25. Nilashi, M.; Minaei-Bidgoli, B.; Alrizq, M.; Alghamdi, A.; Alsulami, A.A.; Samad, S.; Mohd, S. An analytical approach for big social data analysis for customer decision-making in eco-friendly hotels. *Expert Syst. Appl.* **2021**, *186*, 115722. [CrossRef]
26. Nilashi, M.; Ahmadi, H.; Arji, G.; Alsalem, K.O.; Samad, S.; Ghabban, F.; Alzahrani, A.O.; Ahani, A.; Alarood, A.A. Big social data and customer decision making in vegetarian restaurants: A combined machine learning method. *J. Retail. Consum. Serv.* **2021**, *62*, 102630. [CrossRef]
27. Wu, X.L.; Liao, H.C. A Choquet integral-based optimization model for product ranking considering the reliability of evaluations. *Appl. Soft Comput.* **2026**, *192*, 114767. [CrossRef]
28. Wang, Z.; Hu, S.J.; Niu, S.F.; Li, S.Y.; Liu, W.D.; Huang, L.Y. Research on product design improvement method based on online review and improvement importance performance competitor analysis. *Expert Syst. Appl.* **2025**, *279*, 127400. [CrossRef]
29. Meng, Y.; Yang, N.H.; Qian, Z.L.; Zhang, G.Y. What Makes an Online Review More Helpful: An Interpretation Framework Using XGBoost and SHAP Values. *J. Theor. Appl. Electron. Commer. Res.* **2021**, *16*, 466–490. [CrossRef]
30. Rianthong, N.; Dumrongsiri, A.; Kohda, Y. Improving the multidimensional sequencing of hotel rooms on an online travel agency web site. *Electron. Commer. Res. Appl.* **2016**, *17*, 74–86. [CrossRef]
31. Zhao, M.; Li, L.Y.; Xu, Z.H. Study on hotel selection method based on integrating online ratings and reviews from multi-websites. *Inf. Sci.* **2021**, *572*, 460–481. [CrossRef]
32. 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. [CrossRef]
33. Xia, H.S.; An, W.Y.; Liu, G.W.; Hu, R.J.; Zhang, J.Z.; Wang, Y. Smart recommendation for tourist hotels based on multidimensional information: A deep neural network model. *Enterp. Inf. Syst.* **2023**, *17*, 1959651. [CrossRef]
34. Wang, E.; Chen, Y.Y.; Li, Y.M. Research on a Hotel Collaborative Filtering Recommendation Algorithm Based on the Probabilistic Language Term Set. *Mathematics* **2023**, *11*, 4106. [CrossRef]
35. Zhang, C.H.; Cheng, X.R.; Li, K.; Li, B. Hotel recommendation mechanism based on online reviews considering multi-attribute cooperative and interactive characteristics. *Omega* **2025**, *130*, 103173. [CrossRef]
36. Liu, P.D.; Dang, R.; Wang, P.; Xu, Y.C.; Zhang, Y.F. Online-offline combined adaptive hotel recommendation system considering attribute importance and group consensus. *Decis. Support Syst.* **2025**, *196*, 114503. [CrossRef]
37. Venugopalan, M.; Gupta, D. An enhanced guided LDA model augmented with BERT based semantic strength for aspect term extraction in sentiment analysis. *Knowl.-Based Syst.* **2022**, *246*, 108668. [CrossRef]
38. Pu, Z.M.; Zhang, C.; Xu, Z.S.; Wang, X.X. A fuzzy decision support model for online review-driven hotel selection by considering risk attitudes of customers. *J. Oper. Res. Soc.* **2024**, *75*, 1407–1420. [CrossRef]
39. Cairang, Y.; Zhang, H.D.; He, Y.P. A three-way decision method based on jensen-shannon divergence and behavior perception under the probabilistic linguistic term environment. *Appl. Soft Comput.* **2026**, *192*, 114724. [CrossRef]
40. Pang, Q.; Wang, H.; Xu, Z.S. Probabilistic linguistic linguistic term sets in multi-attribute group decision making. *Inf. Sci.* **2016**, *369*, 128–143. [CrossRef]
41. Kirilenko, A.P.; Stepchenkova, S.O.; Dai, X. Automated topic modeling of tourist reviews: Does the Anna Karenina principle apply? *Tour. Manag.* **2021**, *83*, 104241. [CrossRef]
42. 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. [CrossRef]
43. Jain, A.; Darbari, J.D.; Kaul, A.; Ramanathan, U.; Jha, P.C. Facility selection model for BOPS service for an omnichannel retail chain. *IEEE Trans. Eng. Manag.* **2022**, *69*, 2857–2870. [CrossRef]
44. Jiang, S.M.; Cai, S.Q.; Olle, G.O.; Qin, Z.Y. Durable product review mining for customer segmentation. *Kybernetes* **2015**, *44*, 124–138. [CrossRef]
45. Zhang, J.R.; Li, M.J.; Lu, J.C. Asymmetric normalized probabilistic linguistic term set based on prospect theory and its application to multi-attribute decision-making. *Soft Comput.* **2023**, *27*, 10427–10445. [CrossRef]
46. Yazdani, M.; Zarate, P.; Zavadskas, E.K.; Turskis, Z. A combined compromise solution (CoCoSo) method for multi-criteria decision-making problems. *Manag. Decis.* **2019**, *57*, 2501–2519. [CrossRef]

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.