Hacettepe Journal of Mathematics & Statistics

Hacet. J. Math. Stat.
Volume 55 (1) (2026), 303 – 333
DOI : 10.15672/hujms.1757978

Research Article

Bayesian network and sentiment analysis for evaluating online hotel reviews under multiple criteria

Sedef Çalı[1], Adil Baykasoğlu*[2]

[1] *The Graduate School of Natural and Applied Sciences, Dokuz Eylül University, Main Campus, Izmir, Turkey*
[2] *Department of Industrial Engineering, Faculty of Engineering, Dokuz Eylül University, Main Campus, Izmir, Turkey*

Abstract

Online travel agency platforms provide extensive hotel reviews that reflect customer perceptions on multiple criteria. A novel multi–criterion decision–making approach is introduced that integrates Bayesian networks and the graph theory matrix approach to rank hotels based on online customer reviews. A sentiment analysis algorithm is developed to extract sentiment orientations from textual reviews. A Bayesian network is trained using both numerical ratings and sentiment scores to capture probabilistic dependencies among criteria and generate relative importance weights. The derived weights are embedded into the graph theory matrix approach ranking process. A case study on ski hotels in Turkey demonstrates that the Bayesian network graph theory matrix approach integration reflects customer preferences more effectively than conventional multi–criteria decision–making approaches that assume criterion independence. The results indicate that price–performance is a dominant factor in recommending ski hotels. Service quality and food quality are also important criteria that directly affect recommendation decisions and indirectly influence them through price-performance.

Mathematics Subject Classification (2020). 62C10, 90B50, 68T37, 05C90

Keywords. Bayesian network, graph theory matrix approach, hotel ranking, multi-criteria decision-making, online hotel reviews, sentiment analysis.

1. Introduction

Advances in information technologies, the widespread use of social media and customers growing preference of customers for online platforms have led to a surge in online reviews. These assessments offer valuable insight into customer perceptions of products and services and support both customers and producers in making informed decisions. Before purchasing a product, customers often consult online evaluations on e-commerce websites. Similarly, travelers review opinions posted on online travel agency (OTA) platforms to choose the most suitable hotel. However, examining a large volume of reviews, comparing

*Corresponding Author.
Email addresses: sedefcali@gmail.com (S. Çalı), adil.baykasoglu@deu.edu.tr (A. Baykasoğlu)
Received: 22.08.2025; Accepted: 02.02.2026

hotels using multiple conflicting criteria, and making a rational decision can be challenging. Therefore, data-driven decision-making models have been developed to extract useful information from online review data. One such model involves mining online reviews and ranking alternatives within a multi-criterion decision-making (MCDM) framework.

Customers express their needs, satisfaction, and concerns through online hotel reviews, which appear as numerical ratings or textual comments [60]. These reviews are generated independently by the customers. Although some travelers provide ratings for certain criteria, others express their opinions through text [35]. For this reason, analyzing both ratings and textual reviews is essential. Sentiment analysis is used to convert text-based opinions into numerical scores. As a core Natural Language Processing (NLP) task [5], sentiment analysis can be performed at different levels, such as document-, sentence-, or aspect-level. This study employs aspect-level sentiment analysis because it identifies sentiments associated with specific hotel attributes [5].

Hotel ranking based on online reviews is a typical MCDM problem consisting of three stages [17], which are commonly adopted in the related literature. The first stage extracts hotel criteria using data mining techniques. The second stage identifies sentiment orientations related to each criterion by analyzing ratings and sentiment-converted text reviews, which are then used to construct decision matrices. The final stage applies MCDM techniques to rank alternative hotels.

Sentiment analysis approaches can generally be categorized into two types: lexicon-based and machine learning-based methods. As lexicon-based sentiment analysis is well suited for sentence-level evaluation [29], this study adopts such an approach to determine sentiment orientations for hotel criteria. The sentiment algorithm used in this study incorporates several key features. It follows distance and semantic rules to match aspect expressions with sentiment words accurately, enabling precise identification of sentiment orientations. In addition, instead of relying on binary classification, the algorithm assigns five sentiment categories to capture different strengths of customer opinions.

Research on decision-support models for analysing online hotel reviews has increased significantly [1,4,7,28,30,35,39,43,53,60,61,63–66]. However, many existing studies overlook dependence relationships between criteria. Assuming independence is unrealistic, and capturing statistical dependencies among criteria is essential. BNs provide a probabilistic graphical framework that models dependencies among variables and enables reasoning based on these relationships [34]. They have been applied in diverse fields such as medical research [37], healthcare analytics [54], supply chain risk management [22], customer satisfaction analysis [8], education [41], and classification algorithm comparison [40]. BN is a suitable machine learning tool for uncovering statistical dependencies among variables and performing scenario analysis. In this study, BN is used to identify probabilistic relationships among hotel criteria and determine their relative importance. Recent studies have also shown that machine learning can automatically detect key service quality criteria in online reviews, helping hotel managers target improvements based on real feedback [58]. Building on this trend, the proposed methodology employs BN to reveal how service criteria influence travellers recommendations. As a result, hotel managers can develop more informed strategies based on these data-driven insights.

The MCDM method is applied in the final stage to obtain the hotel ranking. Conventional MCDM methods such as Simple Additive Weighting (SAW) and Technique for Order Preference by similarity to ideal solution (TOPSIS) assume that the criteria are independent [56] and fail to capture the interrelationships between criteria. To address this limitation, Bayesian network (BN) and graph theory matrix approach (GTMA) are integrated to achieve a more realistic ranking of alternative hotels.

In summary, this study offers several notable contributions to the literature:

An online-review-based hotel ranking methodology is introduced that employs BN

within an MCDM framework. In addition to hotel ranking, a detailed probabilistic analysis is also proposed to identify the most important criteria that affect customers' recommendation decision. Thus, the methodology provides decision-making support for both potential customers and hotel managers. This represents one of the first attempts to employ MCDM integrated with BN for hotel ranking based on online reviews.

The proposed method uses BN analysis to evaluate the relative importance of criteria within the hotel ranking system. In simpler terms, it introduces a new MCDM approach that combines BN with GTMA.

A sentiment analysis algorithm is provided that discovers multi-granularity sentiments in reviews. Sentiment scores are combined with numerical ratings. This enables a more sensitive representation of customer perceptions. The methodology is applied to Turkish hotel reviews and remains relatively underexplored in the literature.

The remainder of the paper is organized as follows: Section 2 outlines the relevant literature. Section 3 offers a description of the problem and introduces notation. The proposed approach is explained in Section 4. In Section 5, a case study focused on ski hotel reviews to illustrate the reliability of the ranking system. Section 6 presents a comparative analysis, and Section 7 provides the conclusions.

2. Related works

2.1. Hotel rankings through online reviews

Many researchers have conducted studies on hotel ranking methods derived from online customer reviews. These studies start with extracting customer perception and transforming them into quantitative evaluation measures. Some researchers focused solely on online ratings. For example, Peng et al. [39] introduced a probabilistic linguistic integrated cloud model to analyze online ratings in the context of the hotel selection problem. Online evaluations across various criteria were integrated via the proposed information fusion tools based on the Heronian mean operator, considering their interrelationships. Yu et al. [63] introduced a stochastic dominance approach for hotel ranking. In their study, online ratings on hotel criteria were represented by probabilistic sets of linguistic terms in which the frequency of reviews for a given criterion is interpreted as the probability of its linguistic expression. However, customers can articulate their opinions on services through text as well, and valuable information can be gleaned from these reviews. Therefore, other research streams emphasized the role of text comments in the analysis of online reviews. Aggarwal et al [1] analysed textual reviews using topic modelling only to identify evaluation criteria. Liang et al. [28] used sentiment analysis to derive sentiment tendencies in textual reviews. Afterwards, they converted sentiment scores into linguistic distribution assessments. Nie et al. [35] developed more sensitive sentiment analysis algorithm. A sentiment dictionary was semantically divided to classify words into seven sentiment levels, guided by semantic principles. The general evaluation values for the hotel criteria were represented using assessments based on linguistic distribution. Nevertheless, textual reviews alone do not fully capture all dimensions of customer evaluation. To address this issue, Bi et al. [4] ranked hotels using prospect theory, incorporating not only ratings and text reviews, but also price and location information.

The weighting of criteria plays a crucial role in this problem since it affects the final ranking of alternatives. Some researchers utilised subjective weightings. Bi et al. [4] carried out a hotel ranking study in which the weights of hotel criteria were provided by travellers. Some researchers extracted criteria weights from data to ensure objectivity. For instance, Nie et al. [35] and Peng et al. [39] developed optimization models. Zheng et al. [65] determined the weights using the entropy method that evaluates the amount of information conveyed by each criterion. Zhao et al. [66] used maximum deviation method that measures the dispersion of criterion values and assigns higher weights to criteria that

exhibit greater variability. On the other hand, there are some efforts to consider criteria interactions. Tao and You [53] proposed an online review-based hotel selection study that handles interdependencies among criteria using the Choquet integral. However, the quantity of parameters in the Choquet integral experiences exponential growth as the number of criteria increases. Wang et al. [60] derived the weights of the criteria by combining the term frequency (TF)-inverse document frequency (IDF) and Word2Vec algorithms. TF-IDF used to weight the keywords that appear frequently in the comments. Word2Vec then measured how strongly each keyword is semantically related to hotel criteria. The correlations among criteria were captured indirectly through semantic similarity between keywords and criteria. Despite these contributions, there is a lack of comprehensive analyses that explicitly model probabilistic relationships between criteria.

Several MCDM methods and their extensions have been performed in ranking hotels. For instance, Ren et al. [43] developed a data-driven analytic hierarchy process (AHP) and drawn it on hotel selection. Aggarwal et al. [1] used picture fuzzy TOPSIS approach to rank hotels. Wang et al. [60] used the fuzzy TODIM method of a picture to recommend the most suitable hotel for each type of customer. Çal and Balaman [7] presented an approach that examines text reviews and performs hotel ranking using intuitionistic fuzzy (IF) ELECTRE integrated with VIKOR. However, these methods tend to simplify the decision-making process by ignoring interdependencies among criteria. They generally do not account for the joint influence of attributes on customer preferences. These limitations highlight the need for a ranking mechanism that can effectively exploit these dependencies.

Based on the literature review, we summarize the comparison our study with the common hotel ranking studies in Table 1.

2.2. GTMA and BNs for dependency-aware decision-making

GTMA is a convenient ranking mechanism in decision-making problems that consider dependencies. In contrast to other MCDM methods, GTMA boasts more favorable characteristics [2, 42]. The computational process of GTMA is quite clear, allowing decision-makers to manage the interrelationships between criteria and depict them through graphical representations. Incorporating the permanent matrix within this method improves comprehension of the criteria and improves the modeling of the decision-making structure by encompassing all potential structural components and their relative importance. A small change in performance scores can significantly influence the results, thereby simplifying the process of grading the alternatives from highest to lowest [2].

GTMA is not only employed as a stand-alone method, but is also integrated with other methodologies. Table 2 presents studies that combine several methods and models with GTMA. Although some studies applied additional models such as the Best-Worst Method (BWM) [20], DEMATEL [21], fuzzy DEMATEL [59], or Delphi [18] sequentially along with GTMA, other studies used methods such as AHP [55], ANP [19], and interpretive structural modeling (ISM) [25] to directly feed the GTMA framework. Despite these integrations, the existing studies have some drawbacks. These hybrid methods cannot extract information directly from the dataset. Instead, they require decision-makers to assess dependencies and performance scores. Moreover, they are not capable of incorporating probabilistic dependencies into the computation. BN can serve as a valuable tool for extracting information from datasets, including online customer reviews, and feeding the extracted information to GTMA. For example, Maiyar et al. [31] used BN to extract structural relationships among criteria in the context of the design of fashion products, focusing on customer preference orders. However, the study did not address the determination of weights and the ranking of products.

Table 1. Comparisons with several existing studies on hotel selection.

Study	Comment Type		Sentiment Scales	Criteria Interaction	Ranking-Inference	Method/ Model	Language
	Rating	Text					
Yu et al. [64]	✓					Extended VIKOR method	
Peng et al. [39]	✓			✓		Cloud-based method	
Yu et al. [63]	✓					Stochastic dominance-based approach	
Ma et al. [30]	✓					X-means clustering and TOPSIS	
Aggarwal et al. [1]		✓	3			PF-TOPSIS method	English
Liang et al. [28]		✓	5			Distribution linguistic VIKOR	English
Çal and Balaman [7]		✓	3			IF-ELECTRE integrated with VIKOR	Turkish
Zheng et al. [65]		✓	5			TOPSIS	Chinese
Zhao et al. [66]		✓	C			Prospect theory	Chinese
Nie et al. [35]		✓	7			Evidence theory	English
Tao and You [53]		✓	4	✓		Choquet integral based method	English
Wang et al. [60]	✓	✓	3	✓		PF-TODIM method	English
Bi et al. [4]	✓	✓	3			Prospect theory	English
Ren et al. [43]	✓	✓	5			AHP	Chinese
Wu et al. [61]	✓	✓	5			PF-GLDS based on prospect theory	English
This study	✓	✓	5	✓	✓	BN-GTMA method	Turkish

Notes: Ranking-Inference: Ranking and Probabilistic Inference. PF-GLDS: picture fuzzy gained and lost dominance score. C: Continuous (no discrete sentiment scale).

Table 2. Comparison with studies using hybrid GTMA.

Reference	Data Source	Research Objective	Method/ Model
Gupta et al. [20]	Expert	Assessing sustainability for big data analytics usage in logistics	BWM
Gupta et al. [21]	Expert	Analysing the barriers in blockchain adoption for carbon reduction	DEMATEL
Virmani et al. [59]	Expert	Assessing industrys preparedness for artificial intelligence adoption in healthcare supply chains	fuzzy DEMATEL
Tuljak-Suban and Bajec [55]	Expert	Evaluating logistics service providers	AHP
Goyal and Grover [19]	Expert	Evaluating manufacturing system effectiveness	ANP
Kavilal et al. [25]	Expert	Measuring supply chain complexity in automobile industry	ISM
Giri et al. [18]	Expert	Analysing challenges for a sustainable lithium-ion battery supply chain in electric vehicles	Delphi
This study	Online data	Evaluating ski hotels based on customer reviews	BN

3. Problem description

This section presents a clear definition of the problem and introduces the key terminology used. Customers (travellers) evaluate alternative hotels considering several criteria, and they make a final decision about whether they are satisfied. It is important to discover the relationships between the degree of satisfaction of the criteria and the final satisfaction decisions to infer significant insights for businesses. The present study aims to offer a hotel ranking system that considers statistical relationships obtained through BN analysis. The system enables hotel managers to infer which hotel criteria customers give more importance to when recommending it. The system aids potential travellers in their reservation decisions by proposing a hotel ranking derived from previous travellers experiences. The variables and sets used in the methodology are provided below, along with their notations and explanations.

- $H = \{H_1, H_2, \ldots, H_m\}$: a set of hotels which customers evaluate with online reviews in OTA website, where H_i indicates the ith hotel.
- $Cr = \{Cr_1, Cr_2, \ldots, Cr_n\}$: a set of hotel criteria, where Cr_j indicates the jth criterion.
- $W = \{w_1, w_2, \ldots, w_n\}$: vector of weights of criteria, where w_j indicates the weight of jth criterion with the constraint that $w_j > 0$ and $\sum_{j=1}^{n} w_j = 1$.
- $AE_j = \{AE_j^1,\ AE_j^2,\ \cdots,\ AE_j^p\}$: a word list including several aspect expressions AE_j^p (noun or noun phrase) describing the criterion C_j.
- $D = \{D^{++}, D^{+}, D^{-}, D^{--}, D_{\text{hes}}\}$: represents the sentiment lexicon set, comprising strongly positive (D^{++}), positive (D^{+}), negative (D^{-}), strongly negative (D^{--}), and hesitant (D_{hes}) sentiment lexicons.
- $Adv = \{Adv^{+}, Adv^{-}\}$: a set of adverb lists. Adv^{-} is a list of adverbs that intensify only negative sentiments.

- $D^{++} = D_j^{++} \cup D_{\text{gen}}^{++}$: refers to the union of criterion-specific sentiment lexicons and general lexicon with strongly positive sentiment expressions.
- $D^{+} = D_j^{+} \cup D_{\text{gen}}^{+}$: refers to the union of criterion-specific sentiment lexicons and general lexicon with positive sentiment expressions.
- $D^{--} = D_j^{--} \cup D_{\text{gen}}^{--}$: gathers the lexicons involving strongly negative expressions linked to each criterion and general strongly negative sentiment expressions.
- $D^{-} = D_j^{-} \cup D_{\text{gen}}^{-}$: gathers the lexicons involving negative expressions linked to each criterion and general negative sentiment expressions.
- Let $R_i = \{R_{i1}, R_{i2}, \cdots, R_{ir}, \ldots, R_{iR_i}\}$ be a set of reviews regarding to ith alternative hotel.
- Let $S_{jr} = \{S_{jr1}, S_{jr2}, \cdots, S_{jrs}, \ldots, S_{jrS_{jr}}\}$ be a set of (sub) sentences in rth review regarding to ith alternative hotel.
- $\mathbf{V} = [ae, sw, adv, neg, lb]$ is a sentiment vector discovered in a review. It includes aspect expression, sentiment word, adverb, negation and labels, respectively.
- Let Labels = {SPos, Pos, Neu, Neg, SNeg} be a set of labels pointing out polarity categories, where SPos = Strongly Positive, Pos = Positive, Neu = Neutral, Neg = Negative, and SNeg = Strongly Negative. A sentiment vector is classified into one of the five polarities.
- score(Labels): function assigning a sentiment degree to each polarity category in Labels within $[0, 1]$.
- $\text{SentimentDegree}_{V_p}$ is sentiment degree of sentiment vector V_p.
- $\text{TSentimentDegree}_{Cr_j}$ is total sentiment degrees acquired by summing of sentiment degrees of sentiment vectors regarding criterion Cr_j found.
- $\text{SentimentScore}_{Cr_j}$ is sentiment score of review according to criterion Cr_j, where N_{Cr_j} be the number of extracted sentiment vectors related to criterion Cr_j in review R_i.

4. Methodology

The proposed methodology is explained in a systematic manner in this section. Figure 1 outlines the solution procedure for the problem of hotel ranking through online reviews. The method consists primarily of three stages: online review analysis, BN analysis, and MCDM analysis. Online review analysis produces sentiment- and rating-based evaluation scores extracted from individual customer reviews.

Sentiment scores are computed using the developed sentiment analysis algorithm. In the second stage, all hotel evaluation scores are used as input data for BN training. The BN captures probabilistic dependencies among hotel criteria and their influence on the recommendation decision, based on individual observations. It helps reveal which criteria customers give more importance when recommending hotels.

The mutual information is used to calculate the relative importance of the criteria. The outputs of both the first and second stages are employed in the MCDM analysis. The review-level scores of each hotel alternative are averaged and used as input to construct the diagonal assessment matrix for each hotel. The weights of the criteria (BN output) are used to create the relative importance matrix. The final evaluation matrix is constructed for each hotel. After computing the permanent scores of the evaluation matrices, the final hotel ranking is revealed.

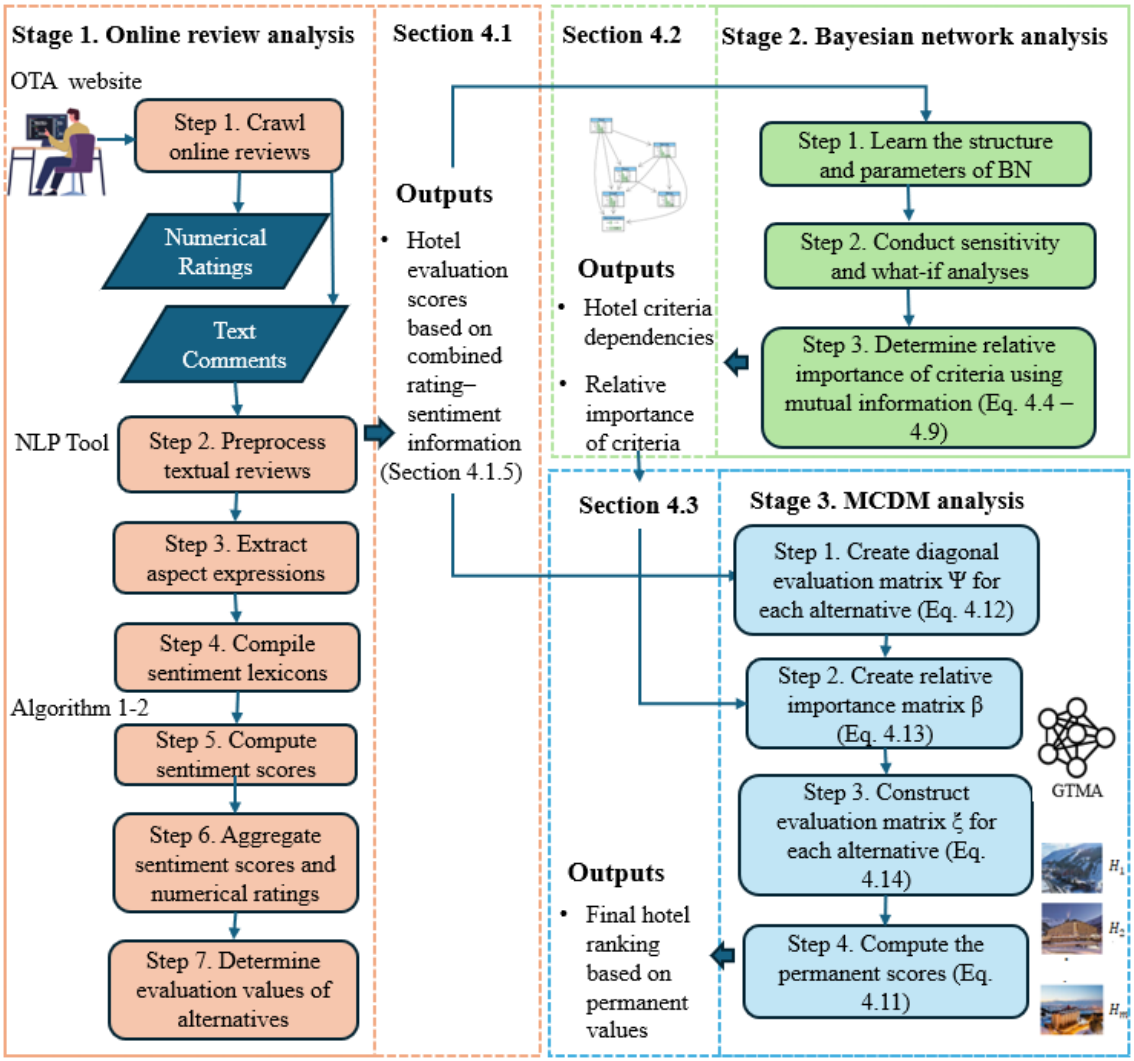


Figure 1. Solution procedure for the problem of hotel ranking through online reviews.

4.1. Online review analysis

4.1.1. Data crawling. A web crawler is used to gather online reviews for data analysis. It is designed to retrieve reviews of relevant hotels from websites. Certain websites allow customers to submit both written reviews and numerical ratings. Textual reviews and numerical ratings are extracted from the web pages for analysis.

4.1.2. Preprocessing online reviews. Textual reviews require data preprocessing before conducting sentiment analysis. In this study, the Turkish textual reviews undergo preprocessing using ITU Turkish Natural Language Processing Web Service [15, 16]. This NLP service provides various tools, including normalization, word analysis, morphological analysis, named entity recognition, and sentence analysis. The sentences are segmented into tokens, which are the smallest units of text such as words and punctuation marks. Words are dissected into their affixes and roots, and part-of-speech (POS) tags (adjective, adverb, noun, verb, etc.) are assigned to the tokens.

4.1.3. Extracting aspect expressions and compiling sentiment lexicons. Customers evaluate hotels by expressing their opinions about their aspects. The criteria consist of related aspects. In a general sense, aspect expressions are nouns or noun phrases that express a criterion. For example, the expression of the aspect "hizmet kalitesi (quality of service) indicates the aspect hizmet (service), which is a member of the Service criterion. The aspect expression "telesiyej" (chair lift) indicates the criterion Skiing. Let AE denote the set covering the subsets of AE_j $(j = 1, 2, \cdots, n)$, where AE_j consists of aspect expressions related to the criterion Cr_j in all reviews. The frequencies of tokens based on POS tags are obtained, and filtered tokens are manually assigned to related subsets of aspect expressions.

Sentiment lexicons are compilations of words and phrases used to classify whether a given text conveys a positive or negative sentiment. A dictionary-based sentiment analysis is employed, where the sentiment word sets are constructed using a dictionary. Seed words are acquired by considering the frequencies of tokens with POS tags of adjectives, adverbs, and verbs. Subsequently, sentiment lexicons are expanded using the Turkish sentiment dictionary SentiTurkNet [13]. These lexicons encompass general sentiment lexicons and specific sentiment lexicons, similar to the construction of the sentiment lexicons explained in [7]. For example, customers might use "yardmsever" (helpful) to describe staff in hotels, and this sentiment word is added to a sentiment lexicon related to the criterion Service. Sentiment words such as "berbat" (terrible) and "muhteem" (excellent) can be applied to all criteria, thus being included in general sentiment lexicons. However, five types of sentiment lexicons are compiled by considering the sentiment degrees of sentiment words and phrases in a different manner from [7]. These lexicons include strongly positive, positive, hesitant, negative, and strongly negative sentiment lexicons. Sentiment words are added to lexicons based on their sentiment degrees. For example, the words "perfect" and "good" do not convey the same sentiment intensity when describing a criterion. Additionally, lists of adverbs are compiled, incorporating adverbs that influence sentiment words to varying degrees. For instance, the adverbs "çok" (very), "epey" (quite), "oldukça" (pretty much), "ar" (too much), "gayet" (extremely), "asla" (never), "hiç" (at all) enhance the sentiment degrees of sentiment words. Negation expressions such as "deil" (not), "olmayan" (not to be), "olmal" (must be), etc., which transform sentiment polarity, are considered as well.

4.1.4. Computing sentiment scores of criteria in textual reviews. This section presents the sentiment analysis algorithm used to compute these sentiment scores. The main function of the sentiment analysis is proposed in Algorithm 1.

Algorithm 1 Function main of aspect-based sentiment analysis

for each review R_{ir} in document R_i **do**
  Prepare a null list for sentiment vectors $V = [ae, sw, adv, neg, lb]$
  Tokenize (R_i)
  Extract (sub)sentences via **splitting rules**
  for each token in (sub)sentence S_{irs} **do**
    Examine tokens according to **distance rule a**
    if token $t_a \in$ set AE_j **then**
      Assign t_a to position of "ae" in sentiment vector
      Assign label of t_a to position of "lb" in sentiment vector
    end if
    Examine tokens according to **distance rule b**
    if token $t_b \in$ set D_j or the set D_{gen} **then**
      Assign t_b to position of "sw" in sentiment vector
      Assign label of t_b to position of "lb" **provided that it is null**
    end if
    Examine tokens according to **distance rule c**
    if token $t_c \in$ the set Adv **then**
      Assign t_c to position of "adv" in sentiment vector
    end if
    Examine tokens according to **distance rule d**
    if token $t_d \in$ the set **Negations then**
      Assign t_d to position of "neg" in sentiment vector
    end if
  end for
end for
Function *SentimentScores* (sentiment vectors V_{R_i})

The algorithm seeks sentiment vectors denoted by $V = [ae, sw, adv, neg, lb]$ in the reviews. The elements of a vector include aspect expression, sentiment word, adverb, negation, and labels. The algorithm searches for the aspect expression and its corresponding sentiment word in the review and records them in the sentiment vector. Labels are used to indicate which criterion is described in the sentiment vector. Since a sentiment vector may contain more than one label, "lb" could encompass multiple labels. For example, "geni" (large/big) might describe either a room or a ski run. The sentiment vector with the sentiment word "geni" includes the labels "skiingPos" and "roomPos." Subsequently, if any adverb or negation influencing the sentiment score of the vector in a positive or negative direction is identified, it is included in the sentiment vector.

The members of the sentiment vector are detected by considering the positions of tokens and some semantic rules as in [7].

Sentence splitting rules: If punctuation such as a full stop, exclamation point, question mark, or ellipsis mark is detected, then the review is split into sentences. If the extracted token has a suffix with tense forms ("-di/-d," etc.) or modal forms ("-meli/-mal," etc.), the sentence is split into sub-sentences from that token. Once sentence connectors (such as although, despite, except, and without) are detected, the sentence is split into two parts. Aspect expressions and sentiment words are then looked for in the next fourteen tokens after the latest sentence vector is discovered.

Distance rules: (a) The subsequent five tokens are searched when sentiment words are just found; (b) if an aspect expression and a sentiment word are found, then continue looking for further sentiment words inside the next five tokens; (c) negation expressions are searched in the next three tokens once the last word of the couple of aspect expression

and sentiment phrase has been saved; (d) if a sentiment word is detected, then search the two tokens before the saved word to find any adverbs.

After obtaining the sentiment vectors in a review, the steps of Algorithm 2 are followed. Each sentiment vector with respect to each criterion is checked, respectively. If a sentiment vector has a member of the strongly positive sentiment lexicon D^{++} or if it has a member of the positive sentiment lexicon D^{+} and an adverb strengthening the sentiment (Adv^{+}), the sentiment vector is assigned to the polarity category of strongly positive (SPos). Afterwards, its polarity category is updated to be positive (Pos) provided that a negation appears in the sentiment vector. Hence, the sentiment degree of the vector is decreased to the score of (Pos). If a sentiment vector has a member of the strongly negative sentiment lexicon D^{--} or if it has a member of the negative sentiment lexicon D^{-} and an adverb intensifying the negative sentiment (Adv), the sentiment vector is assigned to the polarity category of strongly negative (SNeg). Afterwards, its polarity category is updated to be negative (Neg) provided that a negation appears in the sentiment vector. Hence, the sentiment degree of the vector is increased to the score of (Neg). The other steps are simply inferred by readers by following the pseudocode in Algorithm 2.

Algorithm 2 Function of *SentimentScores* (sentiment vectors V_{R_i})

Initialize: $SentimentDegree_{V_v} \leftarrow 0$, $TSentimentDegree_{Cr_j} \leftarrow 0$, $SentimentScore_{Cr_j} \leftarrow 0$
for each sentiment vector $V_v = [ae_v, sw_v, adv_v, neg_v, lb_v]$ in review R_i **do**
  for each Criterion Cr_j in the set Cr **do**
    for each V_v with label of criterion Cr_j **do**
      if ($sw_v \in D^{++}$) or ($adv_v \in Adv^{+}$ and $sw_v \in D^{+}$) **then**
        Assign $SentimentDegree_{V_v} \leftarrow$ score(𝕊**Pos**)
        if not.null(neg_v) **then**
          Decrease $SentimentDegree_{V_v} \leftarrow$ score(Pos)
        end if
      else if ($sw_v \in D^{+}$) or ($sw_v \in D^{-}$ and not.null(neg_v)) **then**
        Assign $SentimentDegree_{V_v} \leftarrow$ score(Pos)
      else if ($sw_v \in D_{hes}$) **then**
        Assign $SentimentDegree_{V_v} \leftarrow$ score(Neu)
      else if ($sw_v \in D^{-}$) or ($sw_v \in D^{+}$ and not.null(neg_v)) **then**
        Assign $SentimentDegree_{V_v} \leftarrow$ score(Neg)
        if $adv_v \in Adv^{-}$ **then**
          Decrease $SentimentDegree_{V_v} \leftarrow$ score(𝕊**Neg**)
        end if
      else if ($sw_v \in D^{--}$) or ($adv_v \in Adv$ and $sw_v \in D^{-}$) **then**
        Assign $SentimentDegree_{V_v} \leftarrow$ score(𝕊**Neg**)
        if not.null(neg_v) **then**
          Increase $SentimentDegree_{V_v} \leftarrow$ score(Neg)
        end if
      end if
      $TSentimentDegree_{Cr_j} \leftarrow TSentimentDegree_{Cr_j} + SentimentDegree_{V_v}$
      $N_{V_{Cr_j}} \leftarrow N_{V_{Cr_j}} + 1$
    end for
    $SentimentScore_{Cr_j} \leftarrow TSentimentDegree_{Cr_j} / N_{V_{Cr_j}}$
  end for
end for
Output: $[SentimentScore_{Cr_1}, SentimentScore_{Cr_2}, \ldots, SentimentScore_{Cr_j}]$

Each sentiment vector has been assigned to one of the polarity categories, namely strongly positive (SPos), positive (Pos), neutral (Neu), negative (Neg), and strongly negative (SNeg). These polarity categories have different sentiment degrees characterized by a score function mapping to the [0, 1] interval. The sentiment scores of the review are finally calculated using the average sentiment degrees of the sentiment vectors.

Examples: (1) Strongly positive opinions are expressed in sentences like "oda çok güzel" (the room is very good), "oda mükemmel" (the room is perfect). (2) Positive opinions are reflected in sentences such as "oda güzel" (the room is good), "oda çok güzel deil" (the room is not very good). (3) Neutral opinions can be seen in sentences like "oda idare eder" (the room is so-so). (4) Negative opinions are reflected in sentences such as "oda çok berbat deil" (the room is not very awful), "oda kötü" (the room is bad). (5) Strongly negative opinions are conveyed through sentences like "oda berbat" (the room is awful), "oday hiç beenmedim" (I didn't like the room at all), "oda hiç iyi deil" (the room is not good at all).

4.1.5. Determining evaluation values of alternatives. Both ratings and reviews are utilised to obtain evaluation values. The evaluation value of alternative hotel H_i with respect to criterion Cr_j is denoted by x_{ij} and is computed according to the following rules:

Rule 1: If there is only a numerical rating for the alternative hotel H_is performance on criterion Cr_j in review r, the corresponding numerical rating is assigned to x_{ij}^r.

Rule 2: If there is only a sentiment score for the alternative hotel H_is performance on criterion Cr_j in review r, the corresponding sentiment score is assigned to x_{ij}^r.

Rule 3: If there is both a sentiment score and a numerical rating for the alternative hotel H_is performance on criterion Cr_j in review r, the average of the corresponding scores is assigned to x_{ij}^r.

Finally, the average of x_{ij}^r is assigned to x_{ij}. The review-level values x_{ij}^r are used as individual observations in the BN analysis, whereas the aggregated hotel-level values x_{ij} are employed in the MCDM analysis.

4.2. BN analysis

BN is a probabilistic model composed of a network structure and a global probability distribution. The network structure is a directed acyclic graph (DAG) $G = (N, A)$ in which the nodes N represent a set of variables $V = \{V_1, V_2, \ldots, V_n\}$ and arcs A point probabilistic dependencies between these variables. The graph structure includes two categories of nodes: parents and children, with directional links pointing from the parent to the child. Assume that node Λ functions as a parent, while node Γ is its corresponding child. The probabilistic dependency between Λ and Γ is obtained according to Bayes theorem by calculating $P(\Gamma \mid \Lambda)$, the probability of Γ given Λ:

$$P(\Gamma \mid \Lambda) = \frac{P(\Lambda \mid \Gamma)P(\Gamma)}{P(\Lambda)} \tag{4.1}$$

A global probability distribution V is factored into local probability distributions according to conditional independence assumptions and the chain rule, as shown in Eq. (4.2).

$$P(V) = P(V_1, V_2, \ldots, V_n) = \prod_i P(V_i \mid \Omega_{V_i}) \tag{4.2}$$

where Ω_{V_i} denotes the parents of V_i. The conditional probabilities in this equation belong to distinct parametric families based on whether the variables are discrete or continuous.

There are two BN approaches in terms of network construction. Expert-driven BNs require domain knowledge and expert experience in forming network structures and eliciting probabilities. Data-driven BNs explore conditional probabilities and/or structure of

BN from data using algorithms. Learning a BN is accomplished in two stages. Structure learning is the first, in which the network structure is learned from the data, and parameter learning is the second, in which local distributions are learned in accordance with the network built in the first phase. To learn the BN structure, two different types of algorithms are used: constraint-based algorithms and score-based algorithms. There are also hybrid approaches, which combine both types.

Score-based learning algorithms refer to the use of various optimization algorithms for the problem of learning the structure of the BN [48]. Score-based learning algorithms aim to identify a network structure that optimizes a predefined scoring function. There are various algorithms such as Hill Climbing (HC) and Tabu Search. When these algorithms are applied, the search space is scanned until the improvements in the calculated score value are removed. On the other hand, constraint-based algorithms detect the presence and direction of arrows between nodes by applying conditional independence tests.

Hybrid algorithms perform conditional independence tests to restrict the search space before score-based learning. Although there is no clear distinction as to which method should be used, Daly et al. [11] point out that studies, in which parameters are learned and models are produced for probabilistic inference following the learning of the network structure, tend to resort to score-based learning. In addition, prior information about the data can be incorporated into all learning algorithms using blacklist and whitelist arguments [47]. A whitelist guaranties that a set of arcs being in the network, whereas a blacklist prevents a set of arcs from being in the network structure.

BN is constructed using mathematical methods in which the structure of the network and the parameters are estimated from the data in this study. First, the HC algorithm is used to determine the network structure. The algorithm starts with an empty DAG. The DAG score is improved iteratively by applying modifications, such as adding, reversing, and deleting arcs, at each iteration. This process continues until no further improvement is possible. The Akaike Information Criterion (AIC) is employed as the score function to select the most suitable network structure for the data. The AIC score is defined in Eq. (4.3).

$$AIC = logL(V) - k * nparams(V) \tag{4.3}$$

where $logL$ calculates the likelihood of the observed data given the model parameters, $nparams$ gives the number of free parameters in BN, and k is the penalty coefficient. These terms jointly balance model fit and model complexity. The network structure that maximises this metric value is determined as the most suitable model for the data [47]. After the BN structure is learned, the parameters are estimated. Bayesian estimation combines prior distributions with observed data through Bayes theorem to obtain conditional probability table values [48]. Since this study uses discrete values in the BN training process, the parameters are estimated with a multinomial likelihood with a uniform Dirichlet prior. The learnt conditional probability values reflect the strengths of the relationships between the variables in the network.

BNs allow decision-makers to conduct probabilistic reasoning and thereby obtain posterior distributions. The two common types of reasoning are predictive and diagnostic reasoning. Predictive reasoning proceeds from parent nodes to child nodes. The beliefs regarding child nodes are updated with evidence regarding the state of a parent node being instantiated. On the other hand, the direction of reasoning is from child nodes to parent nodes in diagnostic reasoning. The beliefs related to parent nodes are updated once evidence related to a child node is inserted. The belief updating is performed with exact algorithm [26].

4.2.1. Determining the relative importance of criteria using mutual information. Criteria weighting methods are used to determine the relative importance of evaluation criteria in MCDM. Subjective methods require the subjective evaluation of decision

makers. Among them, AHP [45] and its fuzzy extensions [24] are the most widely used techniques and are based on pairwise comparisons of criteria. However, pairwise comparisons become time-consuming as the number of criteria increases. Moreover, our study focuses on extracting the opinions of many customers. It is difficult to obtain subjective importance assessments from all customers. Objective weighting methods derive the criteria weights from data. For example, the entropy method [50] measures the uncertainty in information about the values of the criteria. If alternatives get similar scores for a criterion, that criterion exhibits higher entropy and is assigned a lower weight. Online-review-based weighting methods rely on the frequency of terms associated with specific criteria. The underlying assumption is that the criteria mentioned more frequently in reviews are considered more important by customers [28]. These methods typically assume that criteria are independent, and they do not allow the analyst to explore statistical relationships between them. They cannot consider how each criterion influences the final decision. Conversely, in real-world decision-making, criteria interact with one another [53]. A change in one criterion can affect the evaluation of another. Therefore, BN is employed as a weighting mechanism using the concept of mutual information. BNs can model these dependencies through conditional probability distributions. BN quantifies the contribution of information of each criterion to the recommendation decision using posterior probabilities and mutual information.

Mutual information [50] is one of the most popular metrics utilised to rank information sources. It is based on the idea that the uncertainty related to any random variable can be described by the entropy function [38].

$$H(Y) = -\sum_{y} P(y) \log P(y) \tag{4.4}$$

The uncertainty related to Y, given that random variable T is instantiated as t, is obtained via Eq. (4.5).

$$H(Y \mid t) = -\sum_{y} P(y \mid t) \log P(y \mid t) \tag{4.5}$$

Eq. (4.6) computes the average residual uncertainty in Y given T.

$$H(Y \mid T) = \sum_{t} H(Y \mid t) P(t) \tag{4.6}$$

Then, the mutual information expressing the total uncertainty-reducing potential of T is calculated by Eq. (4.7).

$$I(Y,T) = H(Y) - H(Y \mid T) \tag{4.7}$$

Mutual information is symmetric ($I(Y,T) = I(T,Y)$) and a non-negative quantity. It shows the dependency between two random variables. If mutual information is zero, it means that the random variables are mutually independent. Let $Cr = \{Cr_1, Cr_2, \ldots, Cr_n\}$ be the set of random variables which describe the criteria and the random variable Y represents the customers' recommendation decision. The impact of each criterion on recommendation decision is obtained using mutual information. These measures can be reflected in decision-making methodology as criteria weights with Eq. (4.8) or relative importance values with Eq. (4.9).

$$w_j = \frac{I(Cr_j, Y)}{\sum_{j=1}^{n} I(Cr_j, Y)}, \tag{4.8}$$

where w_j indicates the weight of criterion Cr_j and $\sum_{j=1}^{n} w_j = 1$. The relative importance of criterion Cr_i over criterion Cr_j is obtained by drawing inspiration from [27]:

$$r_{ij} = \frac{I(Cr_i, Y)}{I(Cr_i, Y) + I(Cr_j, Y)}. \tag{4.9}$$

To sum up, higher mutual information values indicate that knowing the satisfaction level of a criterion provides more information about the recommendation decision by reducing its uncertainty. Accordingly, a criterion is considered more important if customer satisfaction with that criterion exhibits a stronger probabilistic dependency with the recommendation outcome. The criteria are weighted based on their contribution to explaining customers recommendation behaviour. Although criterion importance is quantified with respect to the recommendation decision, the influence of inter-criterion interactions is inherently embedded in the probabilistic structure of the BN.

4.3. MCDM analysis

4.3.1. GTMA. GTMA is a method employed in MCDM analysis that is based on the principles of matrix operations and graph theory. The computation of permanent introduced by Binet and Cauchy in 1812 [33] is a key component of the graph theory matrix approach. In combinatorial mathematics, the permanent is used to represent the total number of perfect matchings within a bipartite graph. Consider an $m \times n$ matrix, namely $A = (a_{ij})$ where $m \leq n$, Eq. (4.10) [33] calculates the permanent of this matrix.

$$\text{Per}(A) = \sum_{\sigma} a_{1\sigma(1)} a_{2\sigma(2)} \cdots a_{m\sigma(m)} \tag{4.10}$$

where the summation is taken over all bijective mappings from $\{1, \ldots, m\}$ to $\{1, \ldots, n\}$. The sum of all diagonal product of A, one of which is expressed like $a_{1\sigma(1)} a_{2\sigma(2)} \cdots a_{m\sigma(m)}$, gives the permanent of A [2]. It is more complicated to calculate the permanent rather than determinant since Gaussian elimination cannot be used in the calculation process [2]. Valiant [57] proved that determining a matrix's permanent falls into the category of NP-Hard. Thus, Ryser [44] had developed an algorithm to approximate the permanent of an $n \times n$ matrix, which is given in Eq. (4.11).

$$\text{Per}(A) = (-1)^n \sum_{S \subseteq \{1,\ldots,n\}} (-1)^{|S|} \prod_{i=1}^{n} \sum_{j \in S} a_{ij} \tag{4.11}$$

where, the sum is over all subsets of $\{1, \ldots, n\}$ and $|S|$ is the number of elements in S. GTMA has several advantages such as modelling hierarchical problems and considering criteria interaction. The method also has a clear and straightforward computational procedure. In this study, BN analysis is employed to obtain the components of GTMA. The BNGTMA method consists of three steps, which are explained as follows.

Step-1: The elements of the MCDM framework, namely the criteria, sub-criteria, and alternatives are identified.

Step-2: The interrelation between the criteria is specified in this step. The normalization process is performed. A diagonal matrix (Ψ_i) of size $n \times n$ shown in Eq. (4.12) is created for each alternative to represent criteria-based evaluations. The evaluation scores of the alternatives with respect to the criteria are derived from the outputs of the online review analysis (explained in Section 4.1.5). Each x_{ij} indicates the assessment value of alternative i regarding the criterion j.

$$\Psi_i = \begin{bmatrix} x_{i1} & 0 & \cdots & 0 \\ 0 & x_{i2} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & x_{in} \end{bmatrix} \tag{4.12}$$

Eq. (4.13) shows the matrix of relative importance (or interrelation) β, where each r_{ij} indicates how important the criterion i is in proportion to the combined importance of

the criteria i and j [27]. The values of r_{ij} are obtained via Eq. (4.9) considering mutual information (explained in Section 4.2.1).

$$\beta = \begin{bmatrix} 0 & r_{12} & \cdots & r_{1n} \\ r_{21} & 0 & \cdots & r_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ r_{n1} & r_{n2} & \cdots & 0 \end{bmatrix} \tag{4.13}$$

Step-3: The evaluation matrix ξ_{H_i} for each alternative H_i is constructed by adding the criteria evaluation matrix for the alternative to the relative importance (interrelation) matrix, which is given by Eq. (4.14).

$$\xi_{H_i} = \Psi_i + \beta = \begin{bmatrix} x_1 & r_{12} & \cdots & r_{1n} \\ r_{21} & x_2 & \cdots & r_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ r_{n1} & r_{n2} & \cdots & x_n \end{bmatrix} \tag{4.14}$$

The evaluation score for each alternative is computed through the permanent of the matrix ξ_{H_i} as defined in Eq. (4.11). Alternatives are ordered from highest to lowest based on their permanent values. The top-ranked alternative, according to the maximum permanent value, is considered the most favorable.

5. Case study

The proposed methodology is employed in a case study examining reviews of five-star ski hotels in Turkey.

5.1. Analysing online reviews

Customer reviews of seven alternative ski hotels labelled as H_1 through H_7 were crawled from the otelpuan.com website (https://www.otelpuan.com/tr) in September 2022. This OTA website is widely used in Turkey. Customers evaluate hotels according to criteria, namely *Price-Performance* (Cr_1), *Food* (Cr_2), *Room* (Cr_3), *Skiing* (Cr_4), and *Service* (Cr_5). They also indicate their decision on whether they recommend the hotel by posting *recommended*, *not recommended*, or *hesitant*. The dataset comprises textual reviews, ratings, and recommendations for alternative hotels. An illustrative review is presented in Figure 2.



Figure 2. A review crawled from otelpuan.com in September 2022.

The criteria variables are highlighted with red rectangles, while the target variable (*Recommendation*) is outlined with a green rectangle. Customers provide continuous numerical ratings within the range [1, 10] for *Food*, *Room*, *Skiing*, and *Service*. *Price-Performance* is assessed using linguistic variables from the term set (Very Bad, Poor, Medium, Good, and Excellent).

Customers assign scores to *Price-Performance* to convey their satisfaction with the overall costs of hotel facilities, considering the performance provided by the hotel. *Service* is evaluated by rating employees' attitudes and service quality. *Skiing* is assessed based on the hotel's ski facilities. The scores for *Food* reflect customer satisfaction with various aspects, including quality, taste, and presentation. The *Room* criterion is rated to express satisfaction with various aspects such as cleanliness, design, and landscape.

Customers often express their opinions through textual reviews rather than providing numerical ratings for some evaluations. To address this, the aspect-based sentiment analysis algorithm (explained in Section 4.1) is used to analyze textual reviews and convert them into numerical scores, similar to ratings, to avoid information loss. If a customer review lacks a numerical rating for a criterion, the score derived from sentiment analysis is utilised for that specific criterion. In cases where a customer review includes both a numerical rating and a sentiment score for a criterion, the average of these scores is considered the final score for that criterion. In the data set, customers frequently assign points such as 2, 4, 6, 8 and 10 to criteria. Hence, these scores are used as linguistic variables for *Price-Performance*.

5.2. BN analysis

After the online review analysis, the next stage is the BN analysis. The frequency statistics of the dataset are illustrated in Figure 3.

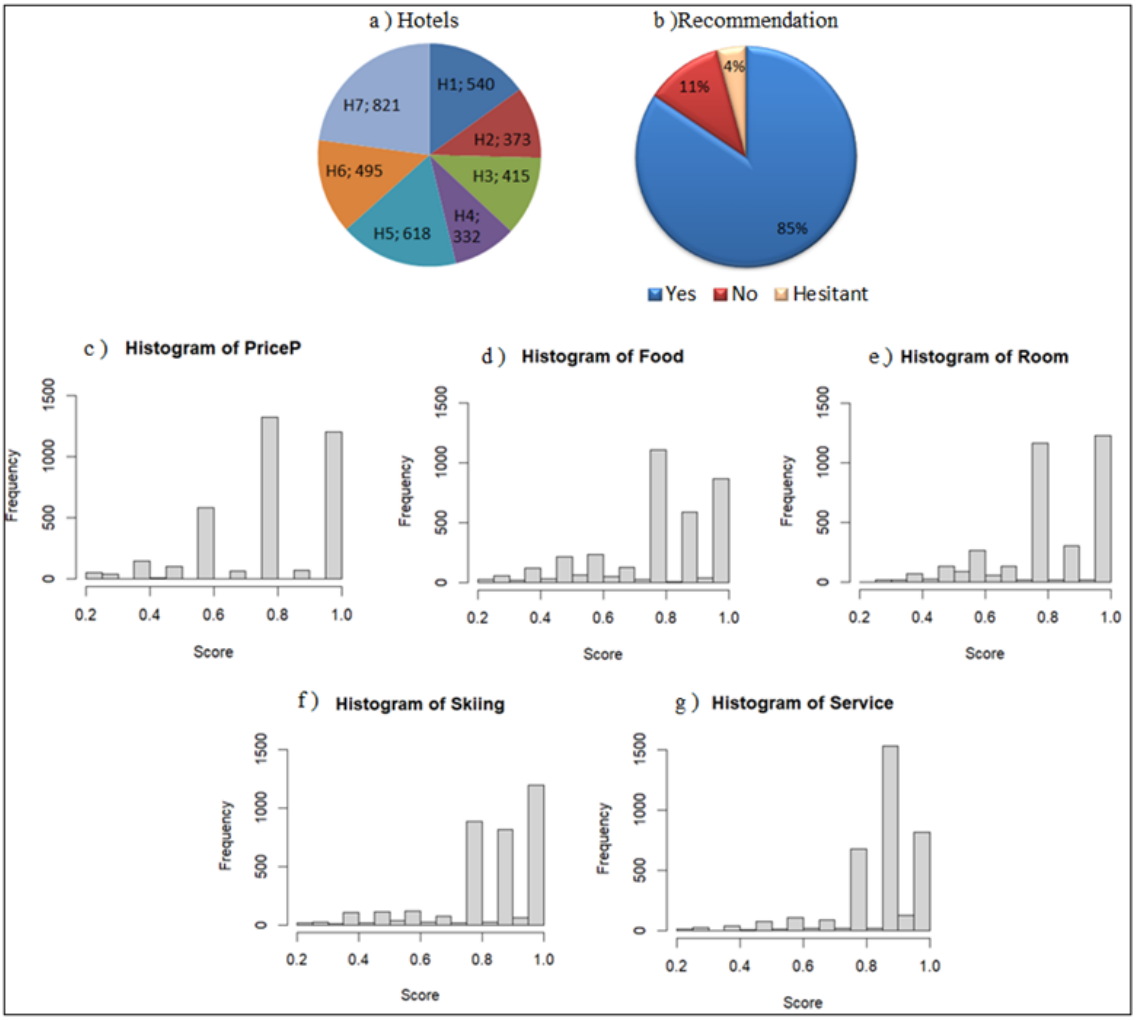


Figure 3. Frequency statistics from the dataset.

The pie chart at the top left of the figure shows the distribution of reviews of alternative hotels. The total number of reviews is 3594. The pie chart at the top right of the figure depicts the distribution of the recommendation choices of customers. The final scores for the criteria are illustrated with histograms. Criteria variables are continuous variables that have values from 0 to 1. They are discretised by dividing them into groups of equal length for BN analysis. Hence, criteria variables are categorised into five states. Since the first state that includes the score interval [0, 0.20] has scarce data, we merge the first and second states. The target variable Recommendation is categorised into two states, namely *Yes* and *No*. Customers answering *Hesitant* to the recommendation choice are included in No due to their scarcity. Finally, the variables and their states are presented in Table 3. This table also gives the intervals representing the states.

Table 3. Variables in the BN.

Variable Type	**Name (Code)**	**State Name (Code):**	**Interval**
Criteria variables	Price-Performance (Cr_1),	Dissatisfied (DisSat):	[0, 0.40]
	Food (Cr_2),	Neutral (Neut):	[0.41, 0.60]
	Room (Cr_3),	Satisfied (Sat):	[0.61, 0.80]
	Skiing (Cr_4),	VerySatisfied (VerySat):	[0.81,1]
	Service (Cr_5).		
Target variable	Recommendation (Y)	Recommended (Yes)	-
		Not Recommended (No)	-

BN is learned from the dataset, which is modelled using the `bnlearn` package in R [47]. Before learning the structures of BNs, some logical constraints are integrated into BN modelling by creating blacklists that prevent two types of arcs: (i) arcs directed from the target variable *Recommendation* to any other variable, and (ii) arcs directed from *Price-Performance* to any criterion variable. BN is constructed to determine the dependence structure between the variables; hence it identifies which criteria variables have a direct impact on *Recommendation*. If there is an association between the target variable and any criterion variable, the arc must be directed from the criterion variable to *Recommendation*. Likewise, if there is an association between *Price-Performance* and any other criterion variable, the arc must be directed from the criterion variable to *Price-Performance*. Logically, customers might assign the satisfaction score for *Price-Performance* according to the performances of the other criteria offered by the hotel. They make their decision on the recommendation by considering the perceived performances of the criteria.

The structure of the BN is learned by using the HC algorithm with the AIC score function, and the local distributions are estimated with Bayesian methods. The learned BN, along with the marginal probabilities of the variables, is shown in Figure 4.

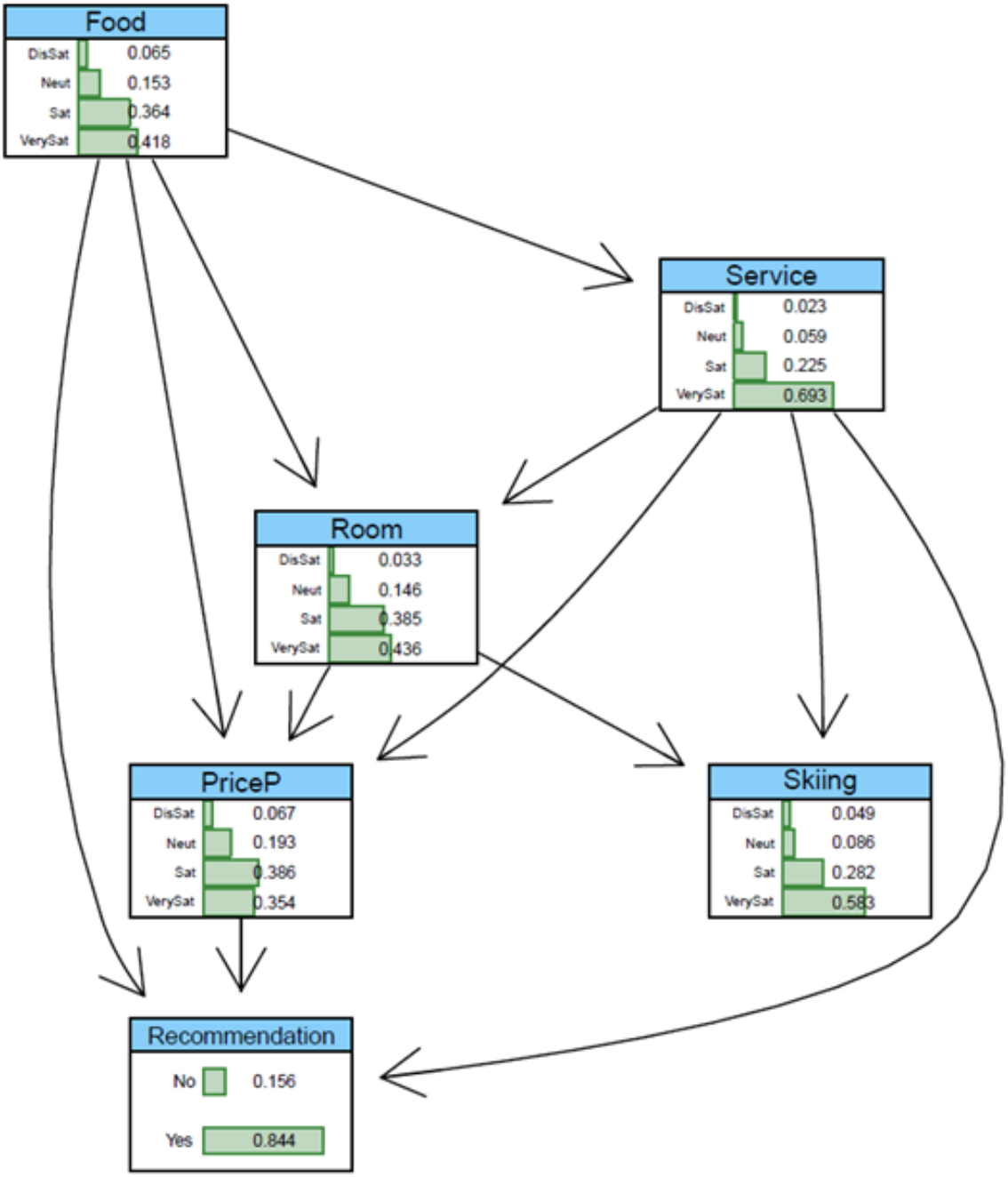


Figure 4. The BN estimated from the dataset.

As mentioned earlier, each criterion variable has four possible states, and the target variable has two possible states. The corresponding value written beside a state is the probability of observing the state of the related variable. In the BN graph, *Price-Performance* is abbreviated as *PriceP*. The BN graph allows the deduction of direct and indirect relationships between variables. It indicates that criterion variables *Price-Performance*, *Service*, and *Food* have a direct influence on *Recommendation*. *Skiing* does not have a direct relationship with *Price-Performance* and *Food* but exhibits indirect connections via certain paths. The highest probability for variables, except *Price-Performance*, is observed at the state *VerySat*. Conversely, *Price-Performance* has its highest probability for the state *Sat*, with a probability of 38.6%.

The primary objective of the BN in this study is to support probabilistic inference and sensitivity analysis rather than out-of-sample prediction. For this reason, a conventional trainingtesting split is not adopted. Instead, model validation follows an approach commonly used in probabilistic inference studies [41]. According to this approach, the quality of the learned BN is assessed using the receiver operating characteristic (ROC) curve [32] and the area under the curve (AUC) [52]. After estimating the network, 3,600 synthetic observations are generated from the fitted model. The ROC curve is then plotted based on simulated data to evaluate the models goodness of fit. The ROC curve illustrates the relationship between the true positive rate (sensitivity) and the true negative rate (1-specificity). It is constructed by comparing the predicted probabilities of the model for the target variable with the simulated results. The AUC metric summarizes the ROC curve into a single index and quantifies the discriminant ability of the BN. It ranges from 0.5 (random performance) to 1.0 (perfect discriminant performance). According to Swets [52], an AUC value above 0.80 indicates good discriminant performance. In this study, the AUC value for the target variable *Recommendation* is 0.92, demonstrating a high level of discriminative performance. This result confirms that the learned BN successfully captures the underlying dependencies in the data and provides reliable probabilistic predictions for subsequent sensitivity and what-if analyses.

5.2.1. Sensitivity analysis and what-if analysis. The BN graph visually represents variables directly linked to *Recommendation* through directed arcs. However, to integrate these relationships into the decision-making process, a measure is necessary. Hence, the degree of impact of each criterion on the recommendation choice is calculated using mutual information in sensitivity analysis. Mutual information values between the criteria variables and *Recommendation* are computed, respectively, as 0.204, 0.125, 0.062, 0.012, and 0.134 bits by Eqs. (4.4)–(4.7). According to the analysis, the criteria can be ordered from most important to least important in determining the recommendation choices of the customers as follows: *Price-Performance*, *Service*, *Food*, *Room*, and *Skiing*. BN is a particularly suitable machine learning tool due to its inference mechanism. Hotel managers can leverage this aspect of BN to analyse the effects of variable states on the recommendation choice. The results of the what-if analysis are illustrated in Figure 5.

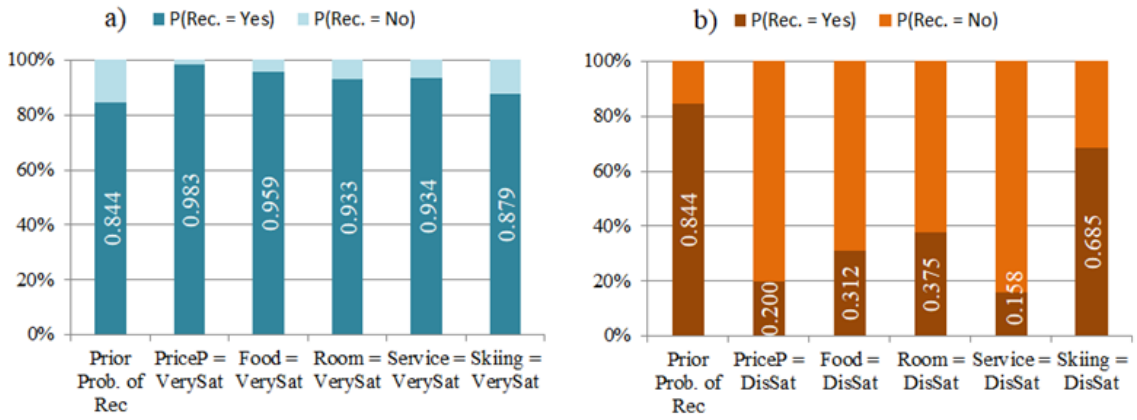


Figure 5. Posterior probabilities of Recommendation, given evidence that the others are equal to (a) *VerySat* and (b) *DisSat*.

The initial bar in each graph represents the probability of *Recommendation* without any evidence (baseline). The left segment (a) of the figure displays each posterior probability of *Recommendation* given evidence that each variable is at the state of *VerySat*, while the right segment (b) of the figure shows each posterior probability of *Recommendation*

given evidence that each variable is at the state of *DisSat*. The probability of positive recommendation choice (*Recommendation* = *Yes*) for hotels increases from 0.844 to 0.983 when there is evidence that customers are very satisfied with *Price-Performance*. Conversely, the probability of positive recommendation choice (*Recommendation* = *Yes*) for hotels decreases from 0.844 to 0.200 when there is evidence that customers are dissatisfied with *Price-Performance*. The figure also illustrates significant changes in the probability of *Recommendation* when setting evidence for *Food* and *Service*. The importance of these variables is highlighted in sensitivity analysis through mutual information as well. In addition, the changes in the probability distribution of *Recommendation* when there is evidence that *Skiing* is at *VerySat* or at *DisSat* are not significant.

5.3. MCDM analysis through GTMA integrated with BN

A digraph is used to describe the problem, in which nodes indicate the criteria and edges indicate interrelationships between the criteria. A directed edge from node i to node j shows the relative importance of criterion i over criterion j. The digraph of the hotel ranking problem depicted in this study is illustrated in Figure 6.

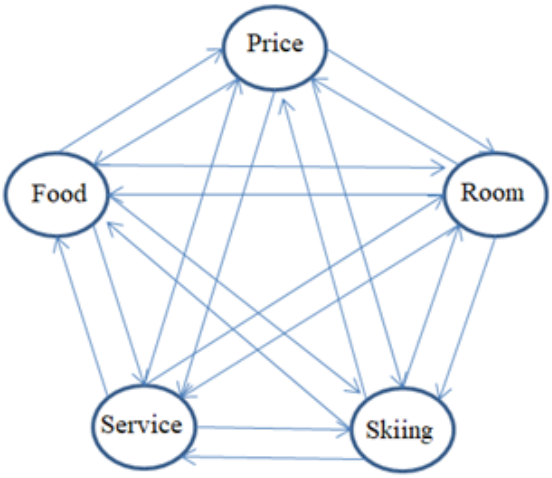


Figure 6. Digraph representation of the hotel ranking problem.

The GTMA integrated with BN is implemented step by step. Firstly, the relative importance values are calculated with Eq. (4.9), and the relative importance matrix is presented in Table 4.

Table 4. Relative importance of the hotel evaluation criteria.

Criteria	Price-Performance	Food	Room	Skiing	Service
Price-Performance	-	0.620	0.767	0.944	0.604
Food	0.380	-	0.668	0.912	0.483
Room	0.233	0.332	-	0.838	0.316
Skiing	0.056	0.088	0.162	-	0.082
Service	0.396	0.517	0.684	0.918	-

In the BN stage, mutual information values between the criteria variables and *Recommendation* are computed as 0.204, 0.125, 0.062, 0.012, and 0.134 bits, respectively. That is, the total uncertainty-reducing potential of *Price-Performance* is 0.204 bits, while the total uncertainty reducing potential of *Food* is 0.125 bit. *Price-Performance* accounts for

62% of the sum of these two values (0.204+0.125). The relative importance of *Price-Performance* is higher than that of Food in terms of hotel recommendation. Secondly, the performance scores of alternative hotels with respect to the criteria are calculated based on online review analysis outputs. Performance values of hotels are found in Table 5. These values are normalised by dividing by the maximum values.

Table 5. Evaluation values of the alternative hotels.

Hotel	**Price-Performance**	**Food**	**Room**	**Skiing**	**Service**
H_1	0.726	0.700	0.801	0.841	0.837
H_2	0.793	0.822	0.836	0.855	0.852
H_3	0.796	0.854	0.785	0.847	0.880
H_4	0.756	0.769	0.804	0.811	0.843
H_5	0.792	0.771	0.812	0.815	0.863
H_6	0.797	0.799	0.764	0.854	0.859
H_7	0.855	0.823	0.887	0.867	0.888

Lastly, permanent matrices of hotels are formed by considering the relative importance of criteria and normalised evaluation values given in Table 6.

Table 6. Permanent matrices of hotels.

(a) H_1

	Cr_1	Cr_2	Cr_3	Cr_4	Cr_5
Cr_1	0.849	0.620	0.767	0.944	0.604
Cr_2	0.380	0.819	0.668	0.912	0.483
Cr_3	0.233	0.332	0.903	0.838	0.316
Cr_4	0.056	0.088	0.162	0.970	0.082
Cr_5	0.396	0.517	0.684	0.918	0.942

(b) H_2

	Cr_1	Cr_2	Cr_3	Cr_4	Cr_5
Cr_1	0.927	0.620	0.767	0.944	0.604
Cr_2	0.380	0.962	0.668	0.912	0.483
Cr_3	0.233	0.332	0.943	0.838	0.316
Cr_4	0.056	0.088	0.162	0.986	0.082
Cr_5	0.396	0.517	0.684	0.918	0.960

(c) H_3

	Cr_1	Cr_2	Cr_3	Cr_4	Cr_5
Cr_1	0.930	0.620	0.767	0.944	0.604
Cr_2	0.380	1.000	0.668	0.912	0.483
Cr_3	0.233	0.332	0.886	0.838	0.316
Cr_4	0.056	0.088	0.162	0.977	0.082
Cr_5	0.396	0.517	0.684	0.918	0.991

(d) H_4

	Cr_1	Cr_2	Cr_3	Cr_4	Cr_5
Cr_1	0.884	0.620	0.767	0.944	0.604
Cr_2	0.380	0.900	0.668	0.912	0.483
Cr_3	0.233	0.332	0.906	0.838	0.316
Cr_4	0.056	0.088	0.162	0.936	0.082
Cr_5	0.396	0.517	0.684	0.918	0.949

(e) H_5

	Cr_1	Cr_2	Cr_3	Cr_4	Cr_5
Cr_1	0.926	0.620	0.767	0.944	0.604
Cr_2	0.380	0.902	0.668	0.912	0.483
Cr_3	0.233	0.332	0.916	0.838	0.316
Cr_4	0.056	0.088	0.162	0.940	0.082
Cr_5	0.396	0.517	0.684	0.918	0.972

(f) H_6

	Cr_1	Cr_2	Cr_3	Cr_4	Cr_5
Cr_1	0.932	0.620	0.767	0.944	0.604
Cr_2	0.380	0.935	0.668	0.912	0.483
Cr_3	0.233	0.332	0.861	0.838	0.316
Cr_4	0.056	0.088	0.162	0.985	0.082
Cr_5	0.396	0.517	0.684	0.918	0.967

(g) H_7

	Cr_1	Cr_2	Cr_3	Cr_4	Cr_5
Cr_1	1.000	0.620	0.767	0.944	0.604
Cr_2	0.380	0.963	0.668	0.912	0.483
Cr_3	0.233	0.332	1.000	0.838	0.316
Cr_4	0.056	0.088	0.162	1.000	0.082
Cr_5	0.396	0.517	0.684	0.918	1.000

Hotel customer satisfaction index values of the alternative hotels are obtained using the matrix permanent formula and given with their ranks in Table 7.

Table 7. Customer satisfaction index values of hotels and ranks.

Hotel	**Hotel Customer Satisfaction Index**	**Rank**
H_1	4.294	7
H_2	4.944	2
H_3	4.942	3
H_4	4.453	6
H_5	4.630	5
H_6	4.724	4
H_7	5.407	1

The methodology suggests the following ranking: $H_7 \succ H_2 \succ H_3 \succ H_6 \succ H_5 \succ H_4 \succ H_1$, where "$\succ$" denotes "superior to". It is evident that H_7 has the highest-ranking score. H_2 and H_3 have almost the same value which follows the score of H_7. This ranking can be presented to potential customers as advice derived from customer reviews through a scientific decision-support tool.

6. Comparative analysis and results

This section compares the results with those obtained using two well-known MCDM methods: SAW [9] and TOPSIS [23]. The criteria weights are derived by Eq. (4.8) while performing SAW and TOPSIS.

SAW calculates the overall score of an alternative by multiplying each criterions value with its weight and summing the resulting products. However, this method may mislead decision-makers since it ignores interactions between criteria [12]. TOPSIS identifies the best option by increasing its closeness to the ideal positive solution and reducing its proximity to the negative ideal alternative [36]. The process starts with building a weighted normalised matrix, after which both the positive ideal (best values) and negative ideal alternatives are defined. The distance values of alternatives from the positive ideal alternative and the negative ideal alternative are calculated. A ratio for each alternative, the division of the distance to the negative ideal alternative by the sum of the distances to the positive and negative ideal alternatives, is computed. However, this method is not capable of dealing with interrelationships between the criteria, similar to SAW.

The ranking of hotels based on SAW and TOPSIS is presented in Table 8. H_7 is in the best position in the ranking produced by each of the three methods. However, H_2 and H_3 compete for the second position in the ranking. BN-GTMA introduces H_2 as the second-ranked hotel, whereas SAW and TOPSIS propose H_3 as the second-ranked hotel. The differences arises because SAW and TOPSIS are based on linear aggregation. They tend to favour alternatives with higher scores on isolated criteria. Although H_3 slightly outperforms H_2 on some individual criteria such as Food and Service, H_2 delivers a more balanced joint performance on Price-Performance, Service and Food. BN-GTMA assigns a higher rank to H_2 since it evaluates hotels based on joint structural dominance rather than isolated criterion values. The matrix permanent concept in BN-GTMA considers all possible structural components of the criteria together with their relative importance. A small variation in the criteria brings about a high difference in permanent values. In other words, even minor performance differences can be captured by the system and result in noticeable changes in the positions of alternatives. This property enhances the discriminative power of the ranking system [2, 3, 42].

Table 8. Comparison of rankings with the results of SAW and TOPSIS.

Methods	**Criteria weights**	H_1	H_2	H_3	H_4	H_5	H_6	H_7
SAW	$w = [0.380,\ 0.233,\ 0.115,\ 0.022,\ 0.250]$	7	3	2	6	5	4	1
TOPSIS	$w = [0.380,\ 0.233,\ 0.115,\ 0.022,\ 0.250]$	7	3	2	6	5	4	1
BN-GTMA	Relative importance matrix	7	2	3	6	5	4	1

Moreover, the proposed methodology is compared with that developed by [7]. They propose a hotel ranking methodology based on online reviews and IF-ELECTRE integrated with VIKOR. As shown in Table 9, different ranking results are obtained across the methods. Both methods rank H_7 as the best hotel and H_1 as the worst hotel. However, the ranking positions of some hotels (H_3, H_5 and H_2) do not remain the same. The differences stem from methodological variations. Çal and Balaman [7] make significant contributions to the literature on hotel ranking through online reviews; however, there are some limitations. In their methodology, they identify the sentiment tendency of each sentence in the reviews by considering positive, negative, and hesitant degrees. However, our sentiment analysis algorithm extracts sentiment orientations from reviews by considering a sentiment scale ranging from strongly positive to strongly negative, with the help of adverbs and additional lexicons. Their study considers only textual reviews, while numerical ratings are excluded. They employ the entropy method to derive the

criteria weights, yielding the weight vector: $w = [0.338,\ 0.018,\ 0.063,\ 0.366,\ 0.214]$. That is, the importance ranking of the criteria is price-performance, skiing, service, food and room, respectively. However, the entropy method derives criteria weights solely from the variability of criteria evaluations, without considering interactions or dependencies.

Table 9. Comparative analysis with the study by [7].

Methods	**Ranking**
The proposed method	$H_7 \succ H_2 \succ H_3 \succ H_6 \succ H_5 \succ H_4 \succ H_1$
The method developed by [7]	$H_7 \succ H_6 \succ H_3 \succ H_5 \succ H_2 \succ H_4 \succ H_1$

The advantage of the BN-GTMA framework is directly linked to the use of mutual information as a weighting mechanism. Unlike entropy, mutual information measures how strongly each criterion is related to the recommendation decision, which leads to different criteria weights in the ranking. A criterion is more weighted if customer satisfaction with that criterion exhibits a stronger probabilistic dependency with the recommendation. Price-Performance is found to be the most important criterion that influences recommendation decision of customers. Service and Food follow as other significant criteria. These findings are consistent with prior studies in tourism and hospitality. Singgalen [51] reports that price-performance (expressed as perceived value in the study) and service quality are among the strongest predictors of customers recommendation intentions. According to Singgalen [51], 90% of the guests who assign high rates to cleanliness, service quality, and perceived value are more likely to recommend the property and revisit it in the future. Ski hotels marketing studies also support the view that price-performance significantly affects overall satisfaction and customer attitudinal loyalty [6, 10]. Hotel managers should aim to raise perceived value by balancing benefit and cost [6]. They should follow effective pricing policies not only to improve customer retention but also to attract new customers. More personalised services may be offered at the same price level. They should primarily focus on training staff to provide kinder, faster, and more attentive service.

Food emerges as a key criterion for customers staying in ski hotels, consistent with Shang et al. [49]. They associate this result with the fact that customers in ski hotels value food variety from a sports-oriented perspective. Hotel managers should offer a wide variety of high-quality and delicious food to enhance customer satisfaction.

In addition, Dogan et al. [14] examine the relationships between hotel attributes and priceperformance using a fuzzy rule-based system and show that these relationships should not be treated as linear. Their observations are consistent with the need for modelling dependencies among hotel criteria, which our study addresses through the BNGTMA framework. The BN structure reveals the impact of the other hotel attributes on price-performance. PricePerformance is directly influenced by Food, Room, and Service, while other criteria such as Skiing affect PricePerformance indirectly through Service. PricePerformance emerges from joint impacts rather than isolated criterion evaluations.

7. Conclusion and future work

This paper proposes an online review ranking methodology integrating BN, sentiment analysis, and GTMA. From a scientific perspective, there are three main contributions. First, it introduces a probabilistic weighting mechanism based on mutual information derived from a BN. Second, it proposes a new MCDM approach, which involves the integration of BN and GTMA to determine the ranking of hotels. This approach can be used for other data-driven decision-making problems beyond hotel ranking based on online reviews. Third, an aspect-based sentiment algorithm is developed to analyze hotel reviews in Turkish; however, it can be adapted to other languages.

From a managerial perspective, the proposed framework provides actionable insights for hotel managers. The BN-based sensitivity and what-if analyses reveal how changes in customer satisfaction of hotel criteria affect the probability of customer recommendation. The model allows managers to evaluate potential improvement scenarios before implementing costly operational changes. Hotel managers can deduce which criteria customers value more when recommending hotels to others. Marketing strategies can be determined according to the criteria that customers actually value. In addition, the proposed ranking allows hotel managers to benchmark their performance against competitors and identify their relative position in the market. Beyond managerial use, the proposed ranking also helps potential customers make informed reservations without reviewing large volumes of comments.

Despite its various contributions, the paper has some limitations. The empirical analysis is based on reviews collected from a single online travel agency platform, which may introduce platform-specific effects. In addition, the case study focuses on seven popular ski hotels in Turkey. Customer preferences can differ across hotel categories, destinations, or cultural contexts. The number of customers who selected the hesitant option for recommendation is relatively small in the dataset. Therefore, these observations are grouped into the not-recommended category in the analysis. These limitations do not affect the validity of the proposed methodology but indicate that the empirical findings should be interpreted within the context of the dataset.

Several extensions may be considered in future research. Customer reviews could be crawled from multiple websites to enrich the data source [62]. The datasets containing multilingual reviews could be analysed, and deep learningbased sentiment analysis models could be integrated to better cope with multilingual datasets [46]. The MCDM component of the study could be enhanced through the integration of fuzzy set theory. Moreover, the proposed methodology can be applied to other service and product domains, such as restaurants, healthcare services, and car selection problems. Finally, future studies could employ dynamic BNs to capture changes in customer preferences over time.

Acknowledgements

The author would like to thank the anonymous referees for their careful reading and valuable comments, which helped improve the quality of this paper.

Author contributions. S. Çal: Conceptualization, Methodology, Software, Validation, Formal analysis, Investigation, Data curation, Writing - Original Draft, Writing - Review & Editing. **A. Baykasolu:** Conceptualization, Methodology, Writing - Review & Editing, Supervision.

Conflict of interest statement. The authors confirm that there is no conflict of interest with any individual or organization for this study.

Funding. No public, private, or nonprofit fund organisation offered specific grant for this study.

Data availability. Any additional information related to the study can be requested from the corresponding author.

References

[1] A.G. Aggarwal, S. Aggarwal and V. Jindal, *Ranking of hotels using customer reviews: an LDA-picture fuzzy TOPSIS approach*, Int. J. Syst. Assur. Eng. Manag. **16** (5), 1885-1898, 2025, doi:10.1007/s13198-025-02742-3.

[2] A. Baykasoglu, *A review and analysis of graph theoretical-matrix permanent approach to decision making with example applications*, Artif. Intell. Rev. **42**, 573-605, 2014, doi:10.1007/s10462-012-9354-y.

[3] A. Baykasoglu, *"Graph theory" and "matrix method" based approach for business process modeling/simulation software selection*, J. Fac. Eng. Archit. Gazi Univ. **28** (3), 555-566, 2013.

[4] J.-W. Bi, T.-Y. Han, Y. Yao and H. Li, *Ranking hotels through multi-dimensional hotel information: a method considering travelers preferences and expectations*, Inf. Technol. Tour. **24**, 127-155, 2022, doi:10.1007/s40558-022-00223-y.

[5] M. Birjali, M. Kasri and A. Beni-Hssane, *A comprehensive survey on sentiment analysis: approaches, challenges and trends*, Knowl.-Based Syst. **226**, 107134, 2021, doi:10.1016/j.knosys.2021.107134.

[6] L. Bonnefoy-Claudet and N. Ghantous, *Emotions' impact on tourists' satisfaction with ski resorts: the mediating role of perceived value*, J. Travel Tour. Mark. **30** (6), 624-637, 2013, doi:10.1080/10548408.2013.810999.

[7] S. Çal and .Y. Balaman, *Improved decisions for marketing, supply and purchasing: mining big data through an integration of sentiment analysis and intuitionistic fuzzy multi criteria assessment*, Comput. Ind. Eng. **129**, 315-332, 2019, doi:10.1016/j.cie.2019.01.051.

[8] S. Chakraborty, K. Mengersen, C. Fidge, L. Ma and D. Lassen, *A Bayesian Network-based customer satisfaction model: a tool for management decisions in railway transport*, Decis. Anal. **3**, 4, 2016, doi:10.1186/s40165-016-0021-2.

[9] C.W. Churchman and R.L. Ackoff, *An approximate measure of value*, J. Oper. Res. Soc. Am. **2**, 172-187, 1954, doi:10.1287/opre.2.2.172.

[10] J.S. Clark and J.K. Maher, *If you have their minds, will their bodies follow? factors effecting customer loyalty in a ski resort setting*, J. Vacat. Mark. **13** (1), 59-71, 2007, doi:10.1177/1356766706071209.

[11] R. Daly, Q. Shen and S. Aitken, *Learning Bayesian networks: approaches and issues*, Knowl. Eng. Rev. **26**, 99-157, 2011, doi:10.1017/S0269888910000251.

[12] M. Darvish, M. Yasaei and A. Saeedi, *Application of the graph theory and matrix methods to contractor ranking*, Int. J. Proj. Manag. **27** (6), 610-619, 2009, doi:10.1016/j.ijproman.2008.10.004.

[13] R. Dehkharghani, Y. Saygin, B. Yanikoglu and K. Oflazer, *SentiTurkNet: a Turkish polarity lexicon for sentiment analysis*, Lang. Resour. Eval. **50**, 667-685, 2016, doi:10.1007/s10579-015-9307-6.

[14] S. Dogan, M.A. Basaran and K. Kantarci, *Determination of attributes affecting price-performance using fuzzy rule-based systems: online ratings of hotels by travel 2.0 users*, J. Hosp. Tour. Technol. **11** (2), 291-311, 2020, doi:10.1108/JHTT-07-2018-0067.

[15] G. Eryiit, *ITU Turkish NLP web service*, in: Proceedings of the Demonstrations at the 14th Conference of the European Chapter of the Association for Computational Linguistics, 1-4, Gothenburg, Sweden, 2014.

[16] G. Eryiit, J. Nivre and K. Oflazer, *Dependency parsing of Turkish*, Comput. Linguist. **34** (3), 357-389, 2008, doi:10.1162/coli.2008.07-017-R1-06-83.

[17] Z.-P. Fan, G.-M. Li and Y. Liu, *Processes and methods of information fusion for ranking products based on online reviews: an overview*, Inf. Fusion **60**, 8797, 2020, doi:10.1016/j.inffus.2020.02.007.

[18] P. Giri, S. Paul and B.K. Debnath, *A hybrid graph-theoretic solution of Delphi and graph theory matrix approach (GTMA) to the challenges of sustainable lithium-ion battery (LIB) supply chains in electric vehicles in India*, J. Energy Storage **97**, 112702, 2024, doi:10.1016/j.est.2024.112702.

[19] S. Goyal and S. Grover, *Manufacturing systems effectiveness measurement by using combined approach of ANP and GTMA*, Int. J. Syst. Assur. Eng. Manag. **4** (4), 404-423, 2013, doi:10.1007/s13198-012-0129-2.

[20] A. Gupta, A. Dwivedi and S.K. Paul, *Developing a sustainability index for implementing big data analytics in the logistics sector*, Sustain. Dev. **33** (S1), 1271-1291, 2025, doi:10.1002/sd.70024.

[21] A. Gupta, R.K. Singh and M.M. Kamal, *Blockchain technology adoption for secured and carbon neutral logistics operations: barrier intensity index framework*, Ann. Oper. Res., 1-34, 2024, doi:10.1007/s10479-024-05824-w.

[22] S. Hosseini and D. Ivanov, *Bayesian networks for supply chain risk, resilience and ripple effect analysis: a literature review*, Expert Syst. Appl. **161**, 113649, 2020, doi:10.1016/j.eswa.2020.113649.

[23] C. Hwang and K. Yoon, *Multiple Attribute Decision Making: Methods and Applications*, Springer, Berlin Heidelberg, 1981.

[24] C. Kahraman, U. Cebeci and Z. Ulukan, *Multi-criteria supplier selection using fuzzy AHP*, Logist. Inf. Manag. **16** (6), 382-394, 2003, doi:10.1108/09576050310503367.

[25] E.G. Kavilal, S. Prasanna Venkatesan and J. Sanket, *An integrated interpretive structural modeling and a graph-theoretic approach for measuring the supply chain complexity in the Indian automotive industry*, J. Manuf. Technol. Manag. **29** (3), 478-514, 2018, doi:10.1108/JMTM-03-2017-0032.

[26] D. Koller and N. Friedman, *Probabilistic Graphical Models: Principles and Techniques*, MIT Press, Cambridge, 2009.

[27] D.-E. Koulouriotis and M.-K. Ketipi, *A fuzzy digraph method for robot evaluation and selection*, Expert Syst. Appl. **38** (9), 11901-11910, 2011, doi:10.1016/j.eswa.2011.03.082.

[28] X. Liang, P. Liu and Z. Wang, *Hotel selection utilizing online reviews: a novel decision support model based on sentiment analysis and DL-VIKOR method*, Technol. Econ. Dev. Econ. **25** (6), 1139-1161, 2019, doi:10.3846/tede.2019.10766.

[29] Y. Liu, J.-W. Bi and Z.-P. Fan, *Ranking products through online reviews: a method based on sentiment analysis technique and intuitionistic fuzzy set theory*, Inf. Fusion **36**, 149-161, 2017, doi:10.1016/j.inffus.2016.11.012.

[30] M. Ma, B. Yu and Z. Xu, *A scalable framework for ranking integration in large-scale online reviews: integrating clustering and multi-attribute decision-making*, Electron. Commer. Res. Appl. **45**, 101570, 2026, doi:10.1016/j.elerap.2025.101570.

[31] L.M. Maiyar, S.J. Cho, M.K. Tiwari, K.D. Thoben and D. Kiritsis, *Optimising online review inspired product attribute classification using the self-learning particle swarm-based Bayesian learning approach*, Int. J. Prod. Res. **57** (10), 3099-3120, 2019, doi:10.1080/00207543.2018.1535724.

[32] C.E. Metz, *Basic principles of ROC analysis*, Semin. Nucl. Med. **8** (4), 283-298, 1978.

[33] H. Minc, *Permanents*, Addison-Wesley, Reading, 1978.

[34] R.E. Neapolitan, *Learning Bayesian Networks*, Pearson Prentice Hall, Upper Saddle River, 2004.

[35] R.-X. Nie, Z.-P. Tian, J.-Q. Wang and K.S. Chin, *Hotel selection driven by online textual reviews: applying a semantic partitioned sentiment dictionary and evidence theory*, Int. J. Hosp. Manag. **88**, 102495, 2020, doi:10.1016/j.ijhm.2020.102495.

[36] D.L. Olson, *Comparison of weights in TOPSIS models*, Math. Comput. Model. **40** (7-8), 721-727, 2004, doi:10.1016/j.mcm.2004.10.003.

[37] E. Park, H.-J. Chang and H.-S. Nam, *A Bayesian network model for predicting post-stroke outcomes with available risk factors*, Front. Neurol. **9**, 1-11, 2018, doi:10.3389/fneur.2018.00699.

[38] J. Pearl, *Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference*, Morgan Kaufmann, San Francisco, 1988.

[39] H.-G. Peng, H.-Y. Zhang and J.-Q. Wang, *Cloud decision support model for selecting hotels on TripAdvisor.com with probabilistic linguistic information*, Int. J. Hosp. Manag. **68**, 124-138, 2018, doi:10.1016/j.ijhm.2017.10.001.

[40] Y. Peng, G. Kou, G. Wang and Y. Shi, *FAMCDM: A fusion approach of MCDM methods to rank multiclass classification algorithms*, Omega **39** (6), 677-689, 2011, doi:10.1016/j.omega.2011.01.009.

[41] L.D. Pietro, R.G. Mugion, F. Musella, M.F. Renzi and P. Vicard, *Reconciling internal and external performance in a holistic approach: a Bayesian network model in higher education*, Expert Syst. Appl. **42** (5), 2691-2702, 2015, doi:10.1016/j.eswa.2014.11.019.

[42] R.V. Rao, *A decision-making framework model for evaluating flexible manufacturing systems using digraph and matrix methods*, Int. J. Adv. Manuf. Technol. **30** (11), 1101-1110, 2006, doi:10.1007/s00170-005-0150-6.

[43] P. Ren, B. Zhu, L. Ren and N. Ding, *Online choice decision support for consumers: data-driven analytic hierarchy process based on reviews and feedback*, J. Oper. Res. Soc. **74** (10), 2227-2240, 2023, doi:10.1080/01605682.2022.2129491.

[44] H.J. Ryser, *Combinatorial Mathematics*, 14th ed., Math. Assoc. of America, New York, 1963.

[45] T.L. Saaty, *Decision-making with the AHP: why is the principal eigenvector necessary*, Eur. J. Oper. Res. **145** (1), 85-91, 2003, doi:10.1016/S0377-2217(02)00227-8.

[46] S. Sagnika, A. Pattanaik, B.S.P Mishra and S.K. Meher, *A review on multi-lingual sentiment analysis by machine learning methods*, J. Eng. Sci. Technol. Rev. **13** (2), 154-166, 2020, doi:10.25103/jestr.132.19.

[47] M. Scutari, *Learning Bayesian networks with the bnlearn R Package*, J. Stat. Softw. **35**, 122, 2010, doi:10.18637/jss.v035.i03.

[48] M. Scutari and J.-B. Denis, *Bayesian Networks with Examples in R*, CRC Press, Boca Raton, FL, 2015.

[49] Z. Shang, J.M. Luo and A. Kong, *Topic modelling for ski resorts: an analysis of experience attributes and seasonality*, Sustainability **14** (6), 3533, 2022, doi:10.3390/su14063533.

[50] C.E. Shannon and W. Weaver, *The Mathematical Theory of Communication*, Univ. of Illinois Press, Urbana, 1949.

[51] Y.A. Singgalen, *Hotel customer satisfaction: a comprehensive analysis of perceived cleanliness, location, service, and value*, J. Bus. Econ. Res. **5** (3), 352-369, 2024, doi:10.47065/jbe.v5i3.6016.

[52] J.A. Swets, *Measuring the accuracy of diagnostic systems*, Science **240** (4857), 1285-1293, 1988, doi:10.1126/science.3287615.

[53] L.-L. Tao and T.-H. You, *A multi-criteria decision-making model for hotel selection by online reviews: considering the traveller types and the interdependencies among criteria*, Appl. Intell. **52**, 12436-12456, 2022, doi:10.1007/s10489-021-03151-2.

[54] K. Topuz, H. Uner, A. Oztekin and M.B. Yildirim, *Predicting pediatric clinic no-shows: a decision analytic framework using elastic net and Bayesian belief network*, Ann. Oper. Res. **263**, 479-499, 2018, doi:10.1007/s10479-017-2489-0.

[55] D. Tuljak-Suban and P. Bajec, *Integration of AHP and GTMA to make a reliable decision in complex decision-making problems: application of the logistics provider selection problem as a case study*, Symmetry **12** (5), 766, 2020, doi:10.3390/sym12050766.

[56] G.-H. Tzeng and J.-J. Huang, *Multiple Attribute Decision Making: Methods and Applications*, CRC Press, New York, 2011.

[57] L.G. Valiant, *The complexity of computing the permanent*, Theor. Comput. Sci. **8** (2), 189-201, 1979, doi:10.1016/0304-3975(79)90044-6.

[58] V. Vargas-Calderón, A. Moros Ochoa, G.Y. Castro Nieto and J.E. Camargo, *Machine learning for assessing quality of service in the hospitality sector based on customer reviews*, Inf. Technol. Tour. **23**, 351-379, 2021, doi:10.1007/s40558-021-00207-4.

[59] N. Virmani, R.K. Singh, V. Agarwal and E. Aktas, *Artificial intelligence applications for responsive healthcare supply chains: a decision-making framework*, IEEE Trans. Eng. Manag. **71**, 8591-8605, 2024, doi:10.1109/TEM.2024.3370377.

[60] L. Wang, X.-K. Wang, J.-J. Peng and J.-Q. Wang, *The differences in hotel selection among various types of travellers: a comparative analysis with a useful bounded rationality behavioural decision support model*, Tour. Manag. **76**, 103961, 2020, doi:10.1016/j.tourman.2019.103961.

[61] J. Wu, C. Fu, E. Zhao, S. Sun and S. Wang, *What makes a better hotel selection? a hybrid approach considering tourists preference and review helpfulness*, Curr. Issues Tour. **28** (10), 1570-1594, 2025, doi:10.1080/13683500.2024.2343098.

[62] W. Xie, W. Zhao and Z. Xu, *Comparative analysis of multi-platform e-commerce online reviews based LDA and sentiment dictionary*, Int. J. Intell. Comput. Cybern. **18** (2), 298-325, 2025, doi:10.1108/IJICC-08-2024-0384.

[63] S.-M. Yu, Z.-J. Du, X.-D. Lin, H.-J. Luo and J.-Q. Wang, *A stochastic dominance-based approach for hotel selection under probabilistic linguistic environment*, Mathematics **8** (9), 1525, 2020, doi:10.3390/math8091525.

[64] S.-M. Yu, J. Wang, J.-Q. Wang and L. Li, *A multi-criteria decision-making model for hotel selection with linguistic distribution assessments*, Appl. Soft Comput. **67**, 741-755, 2018, doi:10.1016/j.asoc.2017.08.009.

[65] J. Zheng, M.Y. Li, D.N. Yuan and S. Xie, *Product ranking considering differences across online review platforms: a method based on intuitionistic fuzzy soft sets*, J. Oper. Res. Soc. **76** (6), 1252-1275, 2025, doi:10.1080/01605682.2024.2417727.

[66] E. Zhao, S. Sun, C. Fu, J. Wu and S. Wang, *How user-generated content influence different types of travelers to select hotels? a perspective with prospect theory*, Inf. Process. Manag. **62** (3), 104049, 2025, doi:10.1016/j.ipm.2024.104049.