DOKUZ EYLÜL UNIVERSITY
GRADUATE SCHOOL OF NATURAL AND APPLIED SCIENCES

GENERATING SOLUTIONS FOR DATA-DRIVEN DECISION MAKING PROBLEMS BASED ON BAYESIAN MODELS

by
Sedef ÇALI

September, 2023
İZMİR

GENERATING SOLUTIONS FOR DATA-DRIVEN DECISION MAKING PROBLEMS BASED ON BAYESIAN MODELS

A Thesis Submitted to the
Graduate School of Natural and Applied Sciences of Dokuz Eylül University
In Partial Fulfillment of the Requirements for the Degree of Master of Science
in Industrial Engineering, Industrial Engineering Program

by
Sedef ÇALI

September, 2023
İZMİR

Ph.D. THESIS EXAMINATION RESULT FORM

We have read the thesis entitled **"GENERATING SOLUTIONS FOR DATA-DRIVEN DECISION MAKING PROBLEMS BASED ON BAYESIAN MODELS "** completed by **SEDEF ÇALI** under supervision of **PROF.DR. ADİL BAYKASOĞLU** and we certify that in our opinion it is fully adequate, in scope and in quality, as a thesis for the degree of Doctor of Philosophy.

Prof. Dr. Adil BAYKASOĞLU

Supervisor

Prof. Dr. Latif SALUM	Doç. Dr. Kökten Ulaş BİRANT
Thesis Committee Member	Thesis Committee Member
Doç. Dr. Zeynel Abidin ÇİL	Doç. Dr. Damla KIZILAY
Examining Committee Member	Examining Committee Member

Prof. Dr. Okan FISTIKOĞLU

Director

Graduate School of Natural and Applied Sciences

ACKNOWLEDGMENT

Firstly, I would like to thank to my supervisor, Prof. Dr. Adil BAYKASOĞLU, for his guidance and support. He offers me insightful counsel for both academic and personal matters. I would also like to present my thanks to the members of the thesis committee, Prof. Dr. Latif SALUM and Doç. Dr. Kökten Ulaş BİRANT for their valuable comments. They have excellent human relation skills.

I would like to express that this dissertation has been supported by 2211-C National PhD Scholarship Program in the Priority Fields in Science and Technology of the Scientific and Technological Research Council of Türkiye (TÜBİTAK). I am very proud and honored to be a part of TÜBİTAK, and to have served my country with the valuable works.

I would like to give special thanks to my parents Nevser ÇALI and Kasım ÇALI, to my grandmother Münevver EVGİN, to my brother Mehmet ÇALI, and to my sister Esra TUNÇER ÇALI for their support and love. I consider myself very lucky to have such a great family.

Sedef ÇALI

GENERATING SOLUTIONS FOR DATA-DRIVEN DECISION MAKING PROBLEMS BASED ON BAYESIAN MODELS

ABSTRACT

Multi-criteria decision-making (MCDM) is a sub-field of operations research which assists decision-makers in choosing the most appropriate alternative among the alternatives by considering multiple and conflicting criteria assessments. In the information age we are in, the rapid developments in Internet technologies, the expansion of e-commerce, the rising volume of data, and diversity lead to new data-driven decision-making problems. Traditional MCDM analysis remains weak in solving new decision-making problems brought by the big data age, addressing uncertainties and criteria interactions.

This thesis addresses the problems of inferring to what extent customers give weight to product characteristics and determining product rankings from online customer reviews. Decision-support methodologies have been developed which use machine learning algorithms based on Bayes' theorem, enhancing the shortcomings of MCDM analysis, to solve these data-driven decision-making problems.

In the first part of the thesis, a new criterion weighting method was developed using the Markov chain Monte Carlo simulation. The method has been applied to the dataset, including the product characteristics and overall customer ratings on the mobile phones sold on e-commerce website, to determine which characteristic of the product customers put more weight on. In the second part, the hotel ranking methodology, including the analysis of both numerical and textual customer reviews, is developed using Bayesian networks and the graph theory-matrix approach. A new sentiment analysis algorithm has been introduced to convert textual reviews in Turkish into numerical scores. Therefore, this thesis contains original contributions to the literature of MCDM integrated with big data analysis and machine learning.

Keywords: Bayesian networks, MCDM, sentiment analysis, Bayesian inference, big data.

VERİ GÜDÜMLÜ KARAR VERME PROBLEMLERİNE BAYESÇİ MODELLERLE ÇÖZÜMLER SUNULMASI

ÖZ

Çok kriterli karar verme (ÇKKV), çok sayıda ve birbirleriyle çelişen kriter değerlendirmelerini dikkate alarak alternatifler içerisinden en uygununu seçmeleri konusunda karar vericilere destek olan yöneylem araştırmasının bir alanıdır. İçinde bulunduğumuz bilgi çağında, internet teknolojilerindeki hızlı gelişmeler, e-ticarete olan yönelim, giderek artan veri hacmi ve çeşitliliği yeni veri-güdümlü karar verme problemlerini ortaya çıkarmaktadır. Geleneksel ÇKKV analizi büyük veri çağının getirdiği yeni karar verme problemlerinin çözümlerinde, belirsizlikleri ve kriter etkileşimlerini ele almakta zayıf kalmaktadır.

Bu tezde, online müşteri yorumlarından müşterilerin ürünlerin hangi özelliklerine ne ölçüde önem verdiğine dair çıkarsama yapma ve ürün sıralamalarını belirleme problemleri ele alınmıştır. Bu veri güdümlü karar verme problemlerinin çözümü için ÇKKV analizinin zayıf yönlerini destekleyen Bayes teoremine dayalı makine öğrenmesi algoritmalarına başvuran karar destek methodolojileri geliştirilmiştir.

Tezin ilk kısmında Markov zinciri Monte Carlo simulasyonuna dayalı yeni bir kriter ağırlığı belirleme yöntemi geliştirilmiştir. Geliştirilen yöntem, bir e-ticaret sitesinde satışa sunulan cep telefonlarına ait ürün özellikleri ve genel müşteri puanlaması verisine uygulanarak müşterilerin ürünün hangi özelliğine daha çok önem verdikleri çıkarsanmıştır. İkinci kısımda, hem sayısal hem metinsel müşteri değerlendirmeleri analizi içeren ve Bayes ağları ve çizge teorisi-matris yaklaşımından yararlanan otel sıralaması methodolojisi geliştirilmiştir. Türkçe müşteri yorumlarını sayısal puanlamaya dönüştürmek üzere yeni bir duygu analizi algoritması sunulmuştur. Dolayısıyla bu tez büyük veri analizi ve makine öğrenmesi entegreli ÇKKV analizi literatürüne orijinal katkılar içermektedir.

Anahtar kelimeler: Bayes ağları, ÇKKV, duygu analizi, Bayesçi çıkarsama, büyük veri.

CONTENTS

Page

THESIS EXAMINATION RESULT FORM ii
ACKNOWLEDGMENT iii
ABSTRACT iv
ÖZ v
CONTENTS vi
LIST OF FIGURES ix
LIST OF TABLES x

CHAPTER ONE - INTRODUCTION 1

1.1 Research Problems and Contributions 3
1.2 Structure of the Thesis 5

CHAPTER TWO - A BAYESIAN BASED APPROACH FOR ANALYZING CUSTOMER'S ONLINE SALES DATA TO IDENTIFY WEIGHTS OF PRODUCT ATTRIBUTES 7

2.1 Chapter Introduction 7
2.2 Related Works 10
2.3 Theoretical Background 14
2.3.1 Bayes' Theorem 14
2.3.2 MCMC Methods 15
2.3.3 Dirichlet Distribution 16
2.3.4 Kernel Density Estimation 17
2.4 Problem Description 19
2.5 Methodology 20
2.5.1 Data Preparing Phase 21
2.5.2 Inference Phase 25

2.6 Case Study 27
2.7 Results of Proposed Method 30
2.8 Comparative Study 34
2.8.1 Non Linear Mathematical Model 34
2.8.2 Linear Regression Analysis with Categorical Predictors 36
2.8.3 Multilayer Perceptron Neural Network 38
2.9 Chapter Conclusions 41
2.9.1 Contributions and Practical Implications 41

CHAPTER THREE - RANKING HOTELS BASED ON BAYESIAN NETWORK AND SENTIMENT ANALYSIS WITHIN A MULTI-CRITERIA DECISION-MAKING FRAMEWORK 44

3.1 Chapter Introduction 44
3.2 Related Works 47
3.2.1 Studies on Product/Service Ranking through Online Reviews 47
3.2.2 Studies Using BN in MCDM 49
3.3 Problem Description 54
3.4 Methodology 56
3.4.1 Online Review Analysis 56
3.4.1.1 Data Crawling 56
3.4.1.2 Preprocessing Online Reviews 56
3.4.1.3 Extracting Aspect Expressions and Compiling Sentiment Lexicons 57
3.4.1.4 Computing Sentiment Scores of Criteria in Textual Reviews 59
3.4.1.5 Determining Evaluation Values in the Decision Matrix 62
3.4.2 BN Analysis 63
3.4.2.1 Bayesian Reasoning 66

3.4.2.2 Determining the Relative Importance of Criteria via Mutual Information 68
3.4.3 MCDM Analysis 69
3.4.3.1 GTMA 69
3.5 Case Study 72
3.5.1 Analyzing Online Reviews 72
3.5.2 BN Analysis 74
3.5.2.1 Accuracy of the Model 78
3.5.2.2 Sensitivity Analysis and What-If Analysis 78
3.5.3 MCDM Analysis through BN-GTMA Model 80
3.6 Comparative Analysis and Discussion 83
3.7 Chapter Conclusions 84

CHAPTER FOUR - CONCLUSIONS 87

REFERENCES 89

LIST OF FIGURES

Page

Figure 2.1 Metropolis-Hastings algorithm 16
Figure 2.2 Solution procedure for the problem 21
Figure 2.3 Flowchart of data collection process 22
Figure 2.4 MCMC algorithm for attribute weight inference 27
Figure 2.5 Screenshot of e-commerce webpage on Hepsiburada.com 28
Figure 2.6 Histograms of the variables of a)Screen size (inch), b) Camera Resolution (MP), c) Memory (GB), d) Battery (mAh), e) Average online ratings, and f) Number of reviews. 29
Figure 2.7 Posterior probability distributions of weights with Gaussian kernel: $w = [0.136, 0.497, 0.169, 0.196]$. 31
Figure 2.8 Posterior probability distributions of weights with Triweight kernel: $w = [0.139, 0.490, 0.159, 0.209]$ 31
Figure 2.9 Posterior probability distributions of weights with Epanechnikov kernel: $w = [0.143, 0.492, 0.152, 0.211]$ 32
Figure 2.10 Posterior probability distributions of weights with uniform kernel: $w = [0.141, 0.468, 0.147, 0.242]$. 32
Figure 3.1 Solution procedure for the problem of product/service ranking through online reviews 57
Figure 3.2 Function main of aspect-based sentiment analysis 59
Figure 3.3 Function of SentimentScores (sentiment vectors VRir) 61
Figure 3.4 An example BN 64
Figure 3.5 Types of reasoning: a) diagnostic, b) predictive, c) intercausal, and d) combined (Korb & Nicholson, 2010) 67
Figure 3.6 A review crawled from otelpuan.com in September 2022 72
Figure 3.7 Frequency statistics from the dataset 74
Figure 3.8 The BN estimated from the dataset 77
Figure 3.9 Posterior probabilities of Recommendation, given evidence that the others are equal to (a) VerySat and (b) DisSat. 79
Figure 3.10 Digraph representation of the hotel ranking problem. 80

LIST OF TABLES

Page

Table 2.1 Data matrix after discretization 23
Table 2.2 Data matrix after density estimation 24
Table 2.3 Descriptive summary of the dataset 29
Table 2.4 Result of discretization 30
Table 2.5 Posterior probabilities for the order of attribute weights according to different kernels 33
Table 2.6 Results of regression analysis 37
Table 2.7 MLP network information 39
Table 2.8 Comparative study results 40
Table 3.1 Variables in the BN 75
Table 3.2 Relative importance of the hotel evaluation criteria 81
Table 3.3 Evaluation values of the alternative hotels 81
Table 3.4 Permanent matrices of hotels 82
Table 3.5 Customer satisfaction index values of hotels and ranks 83
Table 3.6 Comparison of the ranking with the results of SAW and TOPSIS 84

CHAPTER ONE
INTRODUCTION

Rational decision-making is of great importance for companies and individuals. MCDM is used to support both individuals and organizations in choosing the optimal alternative among a set of alternatives by taking into account multiple evaluating criteria. MCDM consists of multi-attribute decision-making (MADM) and multi-objective decision-making (MODM). MADM evaluates a finite number of alternatives, whereas MODM deals with endless alternatives. In this dissertation, we deal with problems with a finite number of alternatives and refer to MADM while using the general term MCDM.

Decision-makers with limited numbers have provided evaluation information on the alternatives for the MCDM process via the conventional surveys with a few questions in general. Nowadays, users can express their opinions about products and discuss their service-related experiences online. Online reviews are preferable to surveys in that they are publicly available, simple to gather, enthusiastically posted by customers and inexpensive in terms of both time and money. In this manner, the huge amounts of data on the internet wait to be processed for decision-making analysis. These secondary data have grown to be an important information source as information sharing has increased. However, the massive volumes of data on the internet and the rising trend of information sharing give rise to new problems and necessitate MCDM being evolved to solve these problems. Many researchers combined MCDM methodologies with data mining, machine learning, and predictive analytics techniques to create data-driven decision-making systems that differ from the earlier MCDM models (Yalcin et al., 2022). Despite increased interest in MCDM based on user-generated big data, there are still research gaps for comprehensive MCDM models integrated with data mining and machine learning.

This dissertation makes use of the methods based on Bayes theory in probabilistic machine learning to propose new models for the data-driven decision-making problems. Markov chain Monte Carlo (MCMC) and Bayesian networks (BNs) are (Murphy, 2012) utilized algorithms in this thesis. MCMC methods generate dependent samples by creating Markov chains to inference the desired probability distributions (Andrieu et al., 2003). Bayesian models utilize Bayes theory based on conditional probabilities. In Bayesian models, posterior distributions are inferred by prior probabilities updated with the new information arrives. These models have several advantages such that they enable decision-makers to incorporate the knowledge to the statistical inference, to extract valuable information from small data, and to estimate probability distributions of parameters instead of point estimation.

A Bayesian network (BN) is a probabilistic graphical model extracting the statistical relationships in the data. It is capable of extracting probabilistic relationships between the random variables and illustrating these relationships with directed acyclic graph. It can be used in improving disadvantages aspects of MCDM techniques. In literature, there are studies integrating BNs and MCDM techniques or developing MCDM methodologies only using BNs. However, most of them utilize knowledge-driven BNs, in which the network and the parameters of the BN are learned through expert knowledge. Data-driven BNs employed in MCDM problems are rather scarce in the literature. In addition these BNs are learned based on structured data. In this respect, one of the main purposes of this thesis is to develop data-driven MCDM model using potential patterns extracted from available data by using techniques based on Bayesian theory so as to help decision-makers.

Online customer reviews include not only ratings but also textual reviews. These texts must be transformed into numerical ratings using text mining techniques. An analytical technique called sentiment analysis is utilized to mine texts and assess the sentimental content in these texts (Fan et al., 2020). Studies on product ranking through online reviews typically focus on Chinese or English reviews, but studies examining Turkish reviews are extremely rare (Çalı & Balaman, 2019). We created

an aspect-based sentiment analysis method for mining Turkish reviews in light of this research gap.

1.1 Research Problems and Contributions

This thesis has introduced two novel methodologies aiming to solve two data-driven decision-making problems. These research problems are related to a critical problem in MCDM literature, weighing the evaluation criteria (Zardari et al., 2015)**.** We have looked for solutions to this problem with different characteristics under different circumstances.

The first problem is determining to what extent customers give importance to the characteristics of the product. This problem is encountered in datasets retrieved from e-commerce websites, where customers are only allowed to post the overall performance ratings regarding the products. Although some websites allow for attribute-specific examination, the data may not be sufficient. The thesis aims to present a novel Bayesian methodology that can answer the following questions regarding the problem:

- How can we infer criteria weights while only the overall satisfaction scores of alternatives are known?
- How can deeper knowledge on customer preferences be explored with relatively fewer data including randomness?
- How can we infer these criteria weights without knowing any statistical parameters?

The developed method, detailed and introduced in Chapter 2, can be summarized in this way. The inputs are overall customer satisfaction scores and product attribute characteristics. Kernel density estimation is utilized to estimate distributions of satisfaction scores of product attributes. The impacts of product attributes on customer satisfaction are considered as weights of mixture kernel distributions and posterior distributions of weights are simulated with a MCMC, namely Metropolis-

Hastings algorithm. Although the applicability of the method is demonstrated within a case study that includes online customer sales data related to mobile phones, the method can be applied to any decision-making problem with overall scores.

The contributions of this methodology are mentioned below:

- This study presents a new solution methodology in acquiring weights to criteria for MCDM analysis under uncertain environment via Bayesian inference. To the best of our knowledge, it is the first study extracting criterion weights only considering overall performance scores of alternatives based on kernel density estimation and Metropolis-Hastings algorithm.
- From the perspective of case study, the study fulfills the research gap in customer satisfaction literature with implementing the Bayesian solution in discovering impact of product attributes based on online data.

The second problem is ranking products/services based on online customer reviews. We have proposed a detailed literature review in Chapter 3 and we have inferred that a comprehensive approach is needed in product ranking through online reviews considering statistical relationships between criteria. In this manner, this thesis aims to present a new methodology that can solve the following questions regarding the so-called problem:

- How can we analyze statistical relationships between criteria performances extracted from the online customer reviews?
- How can we incorporate the interrelationships between criteria into MCDM analysis based on online customer data?
- How can we transform the textual reviews in Turkish into numerical scores to incorporate them into MCDM analysis?

The developed methodology, detailed and introduced in Chapter 3, can be summarized in this paragraph. A product/service ranking methodology is proposed based on analyzing online customer reviews, including both numerical ratings and

textual reviews. A BN is estimated to explore statistical dependencies in the dataset. Text reviews in Turkish are transformed into numerical scores via a new aspect-based sentiment analysis algorithm. The relative importance of criteria acquired by BN analysis are passed to graph theory matrix approach (GTMA). In other words, an integrated model of BN-GTMA is developed for MCDM analysis. The proposed methodology is applied to a case study including some of the ski hotels in Turkey.

The contributions of this methodology to the literature are as followings:

- This study presents a methodology that employs BN in a MCDM framework to analyze online reviews to discover relationships between customer satisfaction regarding different criteria of alternative products/services and, therefore, to provide decision-making support for both potential customers and the related firms.
- The proposed methodology uses BN analysis to determine the relative importance of criteria to be used in the product/service ranking system of MCDM analysis. In other words, it presents BN-GTMA integrated model.
- BN is constructed based on online ratings as well as textual reviews, which are transformed into numerical values through a proposed aspect-based sentiment analysis algorithm. The aspect-based sentiment analysis algorithm considers different sentiment degrees in the reviews. It is implemented to Turkish reviews on ski hotels in Turkey.

1.2 Structure of the Thesis

The thesis includes four chapters. Chapter 1 mentions the scope, purpose and motivation of the thesis. Chapter 2 introduces the developed Bayesian methodology extracting the weights of product attributes on customer satisfaction scores expressed via online ratings. Chapter 3 introduces the developed product/service ranking methodology based on BN and sentiment analysis within a MCDM framework. Chapter 4 mentions the conclusions. Chapter 2 has published as a research article in

the journal of “Expert Systems with Applications”. Chapter 3 has been submitted to a journal, and it is under review.

CHAPTER TWO
A BAYESIAN BASED APPROACH FOR ANALYZING CUSTOMER'S ONLINE SALES DATA TO IDENTIFY WEIGHTS OF PRODUCT ATTRIBUTES

2.1 Chapter Introduction

With the global health crises COVID – 19, the trend of online shopping has significantly increased. UN trade and development experts (UNCTAD, 2020) reported that the pandemic has led to sales growth of e-commerce websites by pointing out "dramatic" rise in online marketplace's share of all retail sales, from 16 per cent to 19 per cent in 2020. It is undeniable the fact that these web platforms are important source of information to the non-experienced customers, the market researchers, and the corresponding companies because most of the e-commerce websites enable the customers to express their satisfaction on webpages. In addition, several studies mentioned the significant influence of online evaluations on both product sales and customer's decision making processes (Chevalier & Mayzlin, 2006; X. Li & Hitt, 2008; Zhang et al., 2013). Therefore, analyzing online ratings and reviews become more of an issue along with the change of human behaviors in shopping.

Analyzing customer opinions have been conducted by using several approaches such as interviews, phone surveys and questionnaire surveys. However, these traditional approaches are time-consuming and labor-intensive tools. Speaking of online customer opinion data, they are acquired with relatively less effort and necessitated less time in contrast with surveys. Moreover, customers are willing to express their opinions online, whereas it might be possible that customers reluctantly fill the survey forms (Groves, 2006). Due to these reasons, online customer data is a prevailing information source to understand customer satisfaction.

Customers post their evaluations on webpages with several ways such as numeric ratings and text reviews. In text reviews, customers can write their experiences on

the products in free format, whereas they summarize their opinions by clicking a rate from predetermined numerical scale. Summary statistics of these ratings are displayed on webpages. Besides the text reviews, valuable information contained in online ratings cannot be negligible. Several studies indicate relationship between online ratings and customer satisfaction (Engler et al., 2015; Gu & Ye, 2014; Thirumalai & Sinha, 2011). Numerical ratings have a big role on expressing customer satisfaction and thereby positive or negative impact on purchase decision of potential customers (De Maeyer, 2012). Product ratings enable to gain deeper understanding of pricing and quality decisions in supply chain (X. Wang et al., 2019) and of market outcomes such as price, demand, product sales and profits (Sun, 2012).

Online customer review characteristics have influence on customer choice. Kostyra et al. (2016) revealed that the higher average rating leads to the higher choice probability. They also examined the importance of brand, price, and technical product attributes changes based on the presence of online customer ratings. Therefore, it is crucial to examine online ratings comprehensively.

Besides, some e-commerce web-sites such as Amazon, eBay etc. allow for rating to overall satisfaction of the product rather attribute specific rating. Even though some websites permit for attribute-specific evaluation, there might not sufficient data. Based on above-mentioned arguments, we aim at proposing novel methodology in deriving the impact of product attributes on customer satisfaction by analyzing online customer ratings. The overall product ratings and attribute characteristics are considered as inputs for the model, and the strength of each attribute on total satisfaction is calculated based on Bayesian approach.

The study proposes a methodology for extracting impact of product attributes on customer satisfaction, which also corresponds to a critical problem in MCDM literature. This problem is determination of criterion weights. These weights to criteria have an important role for measuring overall preferences of alternatives. There are several approaches in determination of criterion weights in MCDM literature (Diakoulaki et al., 1995; J. J. Wang et al., 2009; Zardari et al., 2015).

However, the traditional approaches take into consideration performance score of each attribute in computational process, and they have weaknesses in reasoning under uncertainties. There is a research gap in determination of criterion weights for situations in which only overall satisfaction scores of alternatives are known. This study presents a new solution methodology in acquiring weights to criteria for MCDM analysis under uncertain environment via Bayesian inference. To the best of our knowledge, it is the first study extracting criterion weights only considering overall performance scores of alternatives based on kernel density estimation and Metropolis-Hastings algorithm. From the perspective of case study, the study fulfills the research gap in customer satisfaction literature with implementing the Bayesian solution in discovering impact of product attributes based on online data.

The significance of the research in terms of case study is twofold. The results can shed light on which attributes product designers on mobile phone market should concentrate to enhance product sale volumes as well as customer satisfaction. Analyzing online reviews are subject of computer science and engineering design literature as well as subject of business and marketing literature. For research perspective, this approach points out how deeper knowledge on customer preferences can be explored with relatively less data and computational effort. Although the proposed methodology is implemented in online data, it can also be applied to other decision-making problems in which only overall satisfaction scores of alternatives are known.

The rest of this chapter is organized as follows. Related works are mentioned in Section 2.2. Section 2.3 introduces methods and basic concepts needed afterwards. Section 2.4 describes the problem and gives notations. Section 2.5 explains details of the proposed approach. Section 2.6 proposes case study of analyzing the impact of mobile phone attributes on overall customer satisfaction expressed with online ratings. Section 2.7 discusses the results. Section 2.8 presents comparison analysis. Finally, Section 2.9 provides conclusions and practical implications.

2.2 Related Works

Customer satisfaction is crucial determinant of customer loyalty (Z. Deng et al., 2010) and of purchase intention (Dash et al., 2021) such that increase in customer satisfaction brings about increase in customer loyalty and product sales. Customer satisfaction is referred as one of the results of relational marketing and is a long-term phenomenon, not an abstract and event-oriented result (Lin, 2013). It leads to repurchasing and making recommendation to potential customers. Maintaining an old customer is more cost-effective than developing a new customer. Therefore, it is of significant importance for companies to systematically measure customer satisfaction and thereby get rational insight on customer opinions in both enhancing customer royalty and improving product sales.

Customer satisfaction can be measured through online data or customer surveys leading to the companies in product design and marketing (Chiu et al., 2011; Hou et al., 2019). It is crucial to explore to what extent each product attribute has impact on customer satisfaction (Y. Wang et al., 2018). There are several studies analyzing the impact of product attributes on overall customer satisfaction based on both conventional techniques such as surveys and relatively new techniques on analyzing online reviews. For instance, Farooq et al. (2018) revealed service quality dimensions such as airline tangibles, terminal tangibles, personnel services, empathy and image to have a significant impact on customer satisfaction of Malaysia Airlines by conducting a questionnaire survey. Ahmad et al. (2012) proposed rule-based method to analyze data features regarding various characteristics of customer satisfaction, in which rules are derived by using a probabilistic feature-selection technique. Imtiaz & Islam (2020) analyzed product reviews on mobile phones based on sentiment analyses to identify how product attributes effect on customer satisfaction. Chen et al. (2022) proposed a hierarchical Bayesian model to investigate how consumers differ geographically in terms of their preferences in purchasing mobile phones analyzing online reviews. R. Chen & Xu (2017) proposed a methodology in discovering most influential aspects of overall ratings based on product ontology, topic modeling and regression analysis. In that study, product

ontology and Latent Dirichlet Allocation methods are made use of specifying product aspects. Polarity scores of these aspects are determined by conducting sentiment analysis. Finally, each aspect's impact on overall ratings is obtained based on multiple regression analysis. The proposed method is implemented into online data regarding single lens reflex camera, and most dominant aspects of product on customer satisfaction are revealed as cost performance, image quality and product integrity. Peláez et al. (2018) developed an approach to extract the importance of consumer purchasing criteria based on sentiment analysis and Choquet integral.

Some studies focus on product design processes based on online data. Jin et al. (2014) prioritized engineering characteristics in quality function deployment based on online reviews by proposing integer mathematical model. Y. Wang et al. (2018) proposed logistic regression model to estimate the impact of product attributes on customer satisfaction scores. They analyzed the online reviews of washing machines by employing Latent Dirichlet Allocation to explore significant product attributes, and investigated the differences of customer satisfaction to the price level of washing machines. Kim & Noh (2019) made use of factor analysis and text analysis in finding the major factors pertaining to washing machine design, and they performed linear regression to estimate the influence of each factor to customer satisfaction.

Apart from evaluations of products, some studies evaluate online merchants' performances. Qu et al. (2008) implemented content analysis and regression analysis to determine the most important factors that affect the overall rating of online merchants. They point out that consumers give more importance to post-transaction services than online transaction services when they rate online merchants' performance in Yahoo's merchant review system. This study can assist online merchants to enhance their services and gain competitive advantages in electronic markets by revealing the weights customers put on criteria. Denguir-Rekik et al. (2009) proposed a system of multi-criteria evaluation of e-commerces based on a possibilistic framework to obtain the impact of variability and the divergence of customers' evaluations labeled by star scores regarding each criterion to the global evaluation.

In the present study, we focus on online product ratings. The proposed method puts forward an efficient and practical way to discover influence of product attribute on overall customer satisfaction. It differentiates from previous researches in terms of utilized methods. The satisfaction scores of product attributes are regarded as random variables whose probability distributions are estimated by utilizing non-parametric methods, namely kernel density estimation and impacts of product attributes are regarded as mixing proportions (weights) of mixture distributions whose posterior distributions are simulated via Metropolis-Hastings algorithm. Thereby, both strengths of product attributes on customer satisfaction and satisfaction scores of product attributes are represented with random variables instead of scalar values, and they include more information according to a scalar value. In addition, Bayesian inference algorithm enables to update posterior distributions of weights together with gaining new information on weights. Furthermore, this method contains a clear logic and a simple computational process.

In the literature, Huang et al. (2012) made use of kernel density estimation and MCMC for proposing yield based process capability indices. The distribution function of quality characteristic is obtained via Kernel density estimation. Metropolis-Hastings algorithm is used for sampling from the estimated density. Process capability indices are calculated by counting the nonconforming Metropolis-Hastings samples. We utilize kernel density estimation and MCMC like Huang et al. (2012). However, there are some differences both in focus of problem and general methodology. For instance, mixture proportions of mixture kernel functions are estimated via MCMC sampling in our study. The estimated mixture proportions are random variables displayed with histograms. The process capability indices in Huang et al. (2012) are real numbers obtained with an equation whose inputs are retrieved from Metropolis-Hastings samples. Our study generates a new and effective solution methodology for the problem, which is the estimation of product feature weights (impact) on customer satisfaction in uncertain environment; the present approach can also be related to the weight determination problem in the MCDM. Attribute weights are important component in MCDM analysis, because weights play an important role in result of analysis and weights can change ranking of alternatives. Several methods

have been proposed for attribute weight determination, which can be grouped into subjective, objective and hybrid methods. Pena et al. (2020) and Pena et al. (2021) reviewed the weighting methods and proposed taxonomy. The preference of the weighting method in MCDM analysis is related to circumstances such as the availability of data, specifications of experts and the considered decision-making problem. According to the traditional classification, the relative importance of each criterion is acquired based on the analyst's subjective evaluation in subjective methods. For instance, Analytical Hierarchy Process (AHP) is one of the most common subjective weighting methods based on pairwise comparisons between the criteria (Saaty, 2003). However, objective weighting methods extract the weights from data based on mathematical models without querying the decision-makers or analysts. Entropy method (H. Deng et al., 2000), the Correlation Coefficient and Standard Deviation methods (Y. M. Wang & Luo, 2010) and the Criteria Importance through Inter-criteria Correlation, namely, CRITIC (Diakoulaki et al., 1995) are the most widely used objective weighting methods. The performance scores of attributes are given in decision matrices in all of the above-mentioned methods. These methods are incapable of extracting the weights from insufficient data, and they can only deliver deterministic weights. However, the proposed method enables researchers to extract weights based only on overall performance scores and can cope with uncertainties by presenting stochastic weights.

In addition, there are few studies that integrate some famous theories regarding customer satisfaction and MCDM methods in the literature. Kano model is one of the famous theories related to customer satisfaction, which categorizes the attributes of a product or service based on how well they can satisfy customer needs. Ghorbani et al. (2013) deal with the supplier selection problem characterized as the MCDM problem, in which the importance weights of criteria are calculated based on fuzzy Kano questionnaire and fuzzy analytical hierarchy process. Kano model categories the criteria in five types of quality attribute. AHP plays the main role in criteria weighting. Our study purposes to estimate the weights of product attributes on aggregated customer satisfaction scores having been presented on online platforms. Customer satisfaction modeling or measurement is not within the scope of this work.

Customer satisfactions are already available as online product ratings. The research question is what the weight of each product attribute in these aggregated customer satisfaction scores is. Estimating the impact of product attributes on customer satisfaction points out this question in this study.

2.3 Theoretical Background

In this section, we explain the methods and concepts employed in the proposed methodology. Firstly, Bayes' theorem is given. MCMC, a fundamental method used in Bayesian inference, is illustrated. Dirichlet distribution is addressed, which is one of the most common prior distributions used in this field. Finally, the Kernel density estimation method is explained.

2.3.1 Bayes' Theorem

Consider E and F as two events such that $P(A) \neq 0$ and $P(B) \neq 0$, the conditional probability of A given B is computed by Equation (2.1) which is the Bayes' theorem in probability calculus.

$$P(A|B) = \frac{P(B|A)P(A)}{P(B)} \tag{2.1}$$

In this equation, the term $P(A|B)$ is called *posterior probability*, the term $P(B|A)$ is interpreted as *likelihood*, the term $P(A)$ is *marginal probability* or *prior probability*, and the term $P(B)$ is called *normalizing term*.

Bayesian inference is defined as computing a conditional probability by utilizing the above-mentioned formula (Neapolitan, 2004).

2.3.2 MCMC Methods

Monte Carlo integration generates samples from the desired distribution, and then computes sample averages that roughly reflect expectations. MCMC generates these samples by repeatedly executing a cleverly designed Markov chain (Gilks et al., 1995, p. 1).

MCMC applications are generated mostly for Bayesian inference. In Bayesian inference, we can encounter computations of posterior distributions, which often require solutions of complex integrals. In such circumstances, MCMC is used as an approximation approach to estimate posterior distributions and complex integrals. There are various types of MCMC algorithms, one of which is Metropolis-Hastings algorithm that was developed by Metropolis et al. (1953). This algorithm constructs a Markov chain of which stationary distribution is $\pi(\mathrm{x})$.

Figure 2.1 describes Metropolis-Hastings algorithm providing steps to reach of the desired distribution. According to the algorithm, the first step is generating sample point x^* from proposal distribution $q(.|x^{(t)})$, then it must be checked whether the sample point generated is accepted for the next state of $x^{(t+1)}$. Therefore, the acceptance probability is calculated. If acceptance probability is higher than a random number drawn from uniform distribution, the next step becomes $x^{(t+1)} = x^*$; otherwise the next step in the chain is as the previous one.

1: Initialize $x^{(0)}$; set $t = 0$

2: For $t = 0$ to $N - 1$

3: Sample a point x^* from $q(.|x^{(t)})$

4: Sample a random variable $u \sim Uniform\,(0,1)$

5: Calculate the acceptance probability:

6: $$R\left(x^*, x^{(t)}\right) = \min\left\{1, \quad \frac{\pi(x^*)q(x^{(t)}|x^*)}{\pi(x^{(t)})q(x^*|x^{(t)})}\right\}$$

7: If $u < \alpha$

8: $x^{(t+1)} = x^*$

9: else

10: $x^{(t+1)} = x^{(t)}$

11: end

12: end

Figure 2.1 Metropolis-Hastings algorithm

2.3.3 Dirichlet Distribution

Dirichlet distribution is a multivariate generalization of the beta distribution, whose support includes k-dimensional vectors of which each member is a real number interval zero and one, and sum of the members equals to one. The joint distribution function of a random vector $w_j = (w_1, w_2, \dots, w_k)$ following Dirichlet distribution is in Equation (2.2).

$$f(w_1, w_2, \dots, w_k; \alpha_1, \alpha_2, \dots, \alpha_k) = \frac{\Gamma(\alpha_1 + \cdots + \alpha_k)}{\Gamma(\alpha_1) \dots \Gamma(\alpha_k)} \prod_{j=1}^{k} w_j^{\alpha_j - 1}, \tag{2.2}$$

where $\sum_{j=1}^{k} w_j = 1$, $w_j \geq 0$ for $\forall j$, and $\Gamma(.)$ is gamma function (Giudici et al., 2013, p. 6).

It is not possible to draw sample from this distribution directly. There are two standard approaches for sampling. In order to sample a random weight vector $w_j = (w_1, w_2, \dots, w_k)$ from the k-dimensional Dirichlet distribution with parameters $(\alpha_1, \alpha_2, \dots, \alpha_k)$, we utilize the fastest approach (Gelman et al., 2013, p. 583). That is, we draw k independent random samples $(x_1, x_2, \dots, x_k)$ from gamma distributions each with probability density in Equation (2.3).

$$x_j \sim Gamma(\alpha_i, 1) = \frac{x_j^{\alpha_i - 1} e^{-x_j}}{\Gamma(\alpha_j)}, \tag{2.3}$$

where $\Gamma(.)$ is gamma function, and then normalize the samples by $w_j = x_j / \sum_{j=1}^{k} x_j$.

2.3.4 Kernel Density Estimation

There are two approaches to estimate probability distribution function of the observed data which are parametric and non-parametric (Silverman, 2018). Parametric approach assumes that the observed data are drawn from a known distribution and estimates the parameters of the distribution from data. Otherwise, non-parametric approach estimates density function based on data. Kernel density estimation is one of the non-parametric approaches. Consider n independent and identically distributed samples $X_1, \dots, X_n$ drawn from a univariate distribution with unknown density f. This non-parametric approach estimates unknown density function using kernel density estimator denoted by $\hat{f}$ in Equation (2.4).

$$\hat{f}(x) = \frac{1}{nh} \sum_{i=1}^{n} K\left(\frac{x - X_i}{h}\right) \tag{2.4}$$

where h is smoothing parameter (bandwidth) and K is non-negative kernel function.

There are various kernels in literature (Härdle, 1991). Different kernels generate different shapes of densities. The most common kernel functions are Uniform, Epanechnikov quadratic, Triweight and Gaussian kernels proposed with Equations (2.5)-(2.8). The main difference of Gaussian kernel is that its support is real numbers, whereas the supports of the others are restricted. However, kernel choice is less important than bandwidth (Silverman, 2018).

Uniform kernel:

$$K\left(\frac{x-X_i}{h}\right)=\begin{cases}0.5 & \text{if } \left|\frac{x-X_i}{h}\right|\leq 1\\ 0 & \text{otherwise}\end{cases} \tag{2.5}$$

Epanechnikov quadratic kernel:

$$K\left(\frac{x-X_i}{h}\right)=\begin{cases}0.75*\left(1-\left(\frac{x-X_i}{h}\right)^2\right) & \text{if } \left|\frac{x-X_i}{h}\right|\leq 1\\ 0 & \text{otherwise}\end{cases} \tag{2.6}$$

Triweight kernel:

$$K\left(\frac{x-X_i}{h}\right)=\begin{cases}\frac{35}{32}*\left(1-\left(\frac{x-X_i}{h}\right)^2\right)^3 & \text{if } \left|\frac{x-X_i}{h}\right|\leq 1\\ 0 & \text{otherwise}\end{cases} \tag{2.7}$$

Gaussian kernel:

$$K\left(\frac{x-X_i}{h}\right)=\frac{1}{\sqrt{2\pi}h}e^{\left(-\frac{(x-X_i)^2}{2h^2}\right)} \tag{2.8}$$

Bandwidth has influence on estimator such that it controls degree of smoothing. Bandwidth can be determined with subjectively as well as using data based selection procedures (Silverman, 2018).

2.4 Problem Description

There are vast numbers of customer ratings regarding products sold in e-commerce websites. This thesis aims to explore how much importance the customers give to each attribute of the sold product by mining these customer ratings. The concerned problem can be defined as determining the impact of product attributes on customer satisfaction based on online product ratings. Consider that a researcher is responsible for examining a product such as mobile phone, laptop, camera etc. sold in e-commerce website. Information regarding product attributes on all of brands and models of this product can be found in the website. In addition, this website allows for the customers to declare their satisfaction with the product by rating. To assist this analysis, the proposed methodology can give a support in exploring on which characteristic of the product customers put more weight through overall satisfaction scores. Based on the result of the analysis, companies manufacturing this product give more importance to this attribute in designing phase. According to marketing perspective, companies highlight this attribute of the product in advertising to attract new customers.

The notations with their implications are introduced as below, which represent the sets and variables used in the proposed methodology.

- $P = \{P_1, P_2, \dots, P_m\}$: a set of products with customer ratings sold in e-commerce website, where P_i indicates ith product. The expression 'product' indicates any brand or model hereafter.
- $A = \{A_1, A_2, \dots, A_k\}$: a set of attributes of the product, where A_j indicates the jth attribute.
- $C_j = \{C_{j1}, C_{j2}, \dots, C_{jn_j}\}$: a collection of categories of jth attribute created after discretization, where n_j denotes the number of categories of the jth attribute.
- A_{ij}: indicates category of attribute A_j belonging to product P_i and $A_{ij} \in C_j$.
- $W = \{w_1, w_2, \dots, w_k\}$: vector of weights of attributes, where w_j indicates the weight of jth attribute with the constraint that $w_j > 0$ and $\sum_{j=1}^{k} w_j = 1$.

- z_i : overall score of product P_i, $i = 1,2, \dots, m$. Each value z_i is a member of the set Z.
- t_j: indicates a value of product attribute A_j taken by any product P_i in database (i.e., 6.4 inches for screen size).
- Y_{ij} : random variable denotes satisfaction score of product P_i with reference to attribute A_j, and y_{ij} indicates the realization of random variable Y_{ij}.
- $f(Y_{ij} = y_{ij})$: distribution function of random variable Y_{ij}.

2.5 Methodology

In this section, we explain the proposed methodology.

Figure 2.2 gives solution procedure for the stated problem of determining the impact of product attributes on customer satisfaction based on online product ratings. The method mainly consists of two phases, namely data preparing and inference. Data preparing phase incudes three steps, which are crawling online product ratings, discretization and estimating kernel densities of satisfaction scores of products regarding attributes, respectively. Then, inference phase follows three steps which are obtaining mixtures of kernel densities, acquiring likelihood function via mixture densities and utilizing Bayesian inference algorithm to compute attribute weights, respectively.

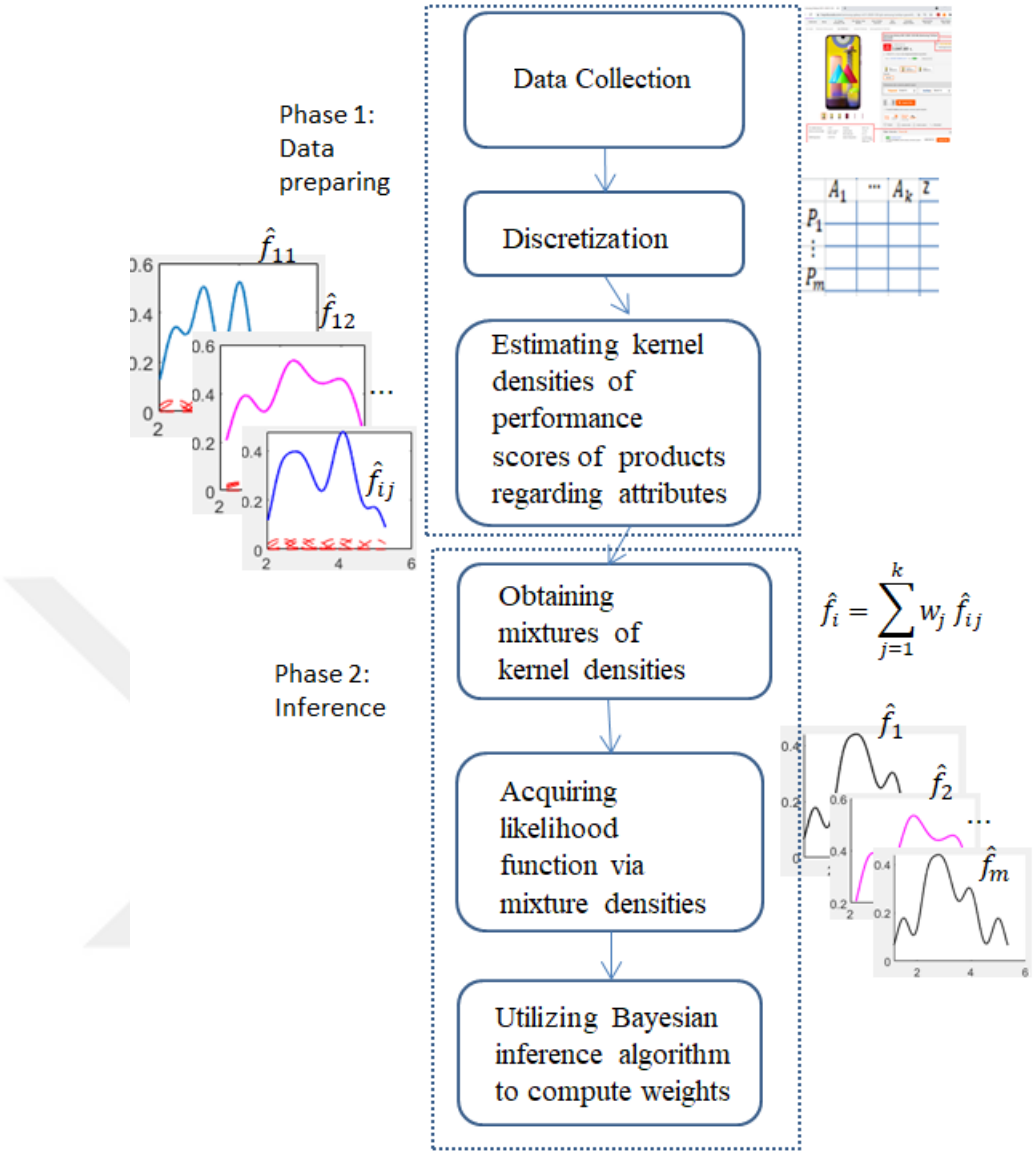


Figure 2.2 Solution procedure for the problem

2.5.1 Data Preparing Phase

The first step of data preparing is gathering the data consisting of characteristics of product attributes and overall numeric ratings. Data collection consists of three steps as shown in Figure 2.3. Data collection process starts with making decisions on the product type and attributes to be analyzed, and on the web pages from which the data are collected. For instance, mobile phone is the product type addressed in this study. At the next step, a web crawling tool is used for retrieving attribute

information and overall product ratings on products from the web pages. The collected data is saved in a comma-separated values (CSV) file. At the last step, the appropriate products are filtered for the analysis. In this study, the mobile phones with at least 50 reviews are used for the analysis.

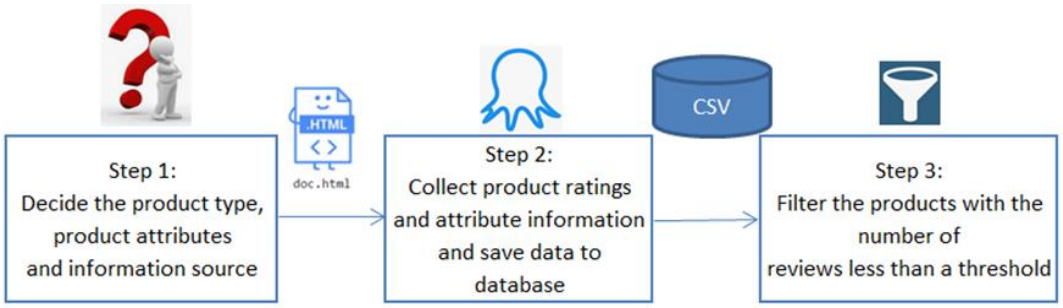


Figure 2.3 Flowchart of data collection process

Two types of product attributes are concerned, namely categorical attributes and continuous attributes. For instance, memory is a categorical attribute for product of mobile phone, which takes a value from a set with relatively less members such as C_3 = {8, 16, 32, 64, 128, 256, 512}. On the other hand, Battery power is a continuous attribute, which takes a value from an interval such as between 1560 and 7000 mAh. In order to discretize continuous attributes, we use equal width discretization, in which the number line between $\min(t_j)$ and $\max(t_j)$ are divided into predetermined number of equal widths (Dougherty et al., 1995). For a detailed explanation, consider the Equation (2.9) where each function $f_j: \mathbb{R} \to \{C_{j1}, C_{j2}, \dots, C_{jn_j}\}$ transforms the value of t_j into categorical symbol.

$$f_j(t) = \begin{cases} C_{j1}, & \min(t_j) \le t < \min(t_j) + wd_j \\ C_{j2}, & \min(t_j) + wd_j \le t < \min(t_j) + 2 * wd_j \\ & \vdots \\ C_{jn_j}, & \min(t_j) + (n_j - 1) * wd_j \le t \le \max(t_j) \end{cases} \quad (2.9)$$

where $t \in t_j$, $wd_j = (\max(t_j) - \min(t_j)) / n_j$, and $j = 1,2 \dots, k$.

After the discretization step, we illustrate data in Table 2.1. It displays overall score z_i of each product available in the website having been obtained by averaging all ratings for the product, as well as displays the categorical values denoted by A_{ij} ($A_{ij} \in C_j$).

Table 2.1 Data matrix after discretization

	Attribute (j)				Overall score
Product (i)	1	2	...	k	Z
1	A_{11}	A_{12}	...	A_{1k}	z_1
2	A_{21}	A_{22}	...	A_{2k}	z_2
⋮	⋮	⋮	⋮	⋮	⋮
m	A_{m1}	A_{m2}	...	A_{mk}	z_m

We consider the satisfaction score of product P_i with reference to attribute A_j as a continuous random variable Y_{ij}. In order to acquire values of random variables Y_{ij} , the procedure considering overall scores z_i and categorical values A_{ij} are followed by using Equation (2.10).

$$Y_{ij} = y_{ij}, \qquad y_{ij} \subseteq Z, \qquad y_{ij} \ni z_r \quad r \,|A_{ij} = A_{rj} \tag{2.10}$$

where $y_{ij} = \{y_{ij}^1, y_{ij}^2, \dots y_{ij}^t, \dots, y_{ij}^s\}$ denotes the set containing values of random variable Y_{ij} with s number of values and a value of random variable provides assignment of $y_{ij}^t = z_i$.

Another way to interpret the procedure is that, satisfaction scores of products with respect to attributes are represented by random variables whose values are extracted from overall scores such that the scores of products belonging to same category concerning an attribute are considered as observations of these random variables.

The distribution of Y_{ij} is computed by kernel density estimation approach (Giudici et al., 2013, p. 327) which employs Equation (2.11) to acquire the function fitted these values.

$$\hat{f}(Y_{ij}) = \frac{1}{sh}\sum_{t=1}^{s} K\left(\frac{Y_{ij} - y_{ij}^{t}}{h}\right) \tag{2.11}$$

where *K* is a kernel function and *h* is bandwidth. Kernel density estimation is performed with Gaussian, Uniform, Epanechnikov quadratic and Triweight kernel functions separately. For instance, kernel density estimation approach generates density distribution by averaging of s Gaussian functions with each value considered as a center with Equation (2.12) when Gaussian kernel is used.

$$\hat{f}(Y_{ij}) = \frac{1}{s}\sum_{t=1}^{s} \frac{1}{\sqrt{2\pi}h} e^{\left(-\frac{\left(Y_{ij}-y_{ij}^{t}\right)^{2}}{2h^{2}}\right)} \tag{2.12}$$

Bandwidth is determined using rule of thumb, which gives optimum bandwidth $\hat{h}_o$ as $1.06\hat{\sigma}n^{-1/5}$ where n is number of observations and $\hat{\sigma}$ is estimator of standard deviation. For other non-Gaussian kernels, the equivalent bandwidth $\hat{h}_1$ is acquired by referencing normal density. The details for the calculus of bandwidth are given in (Härdle, 1991; Silverman, 2018). The result of above-mentioned computation can be illustrated in Table 2.2.

Table 2.2 Data matrix after density estimation

	Attributes				Overall
P_i	A_1	A_2	...	A_k	score (z)
1	$\hat{f}(Y_{11} = y_{11})$	$\hat{f}(Y_{12} = y_{12})$	...	$\hat{f}(Y_{1k} = y_{1k})$	4.6
2	$\hat{f}(Y_{21} = y_{21})$	$\hat{f}(Y_{22} = y_{22})$	...	$\hat{f}(Y_{2k} = y_{2k})$	4.8
⋮	⋮	⋮	⋮	⋮	⋮
m	$\hat{f}(Y_{m1} = y_{m1})$	$\hat{f}(Y_{m2} = y_{m2})$	...	$\hat{f}(Y_{mk} = y_{mk})$	4.5

2.5.2 Inference Phase

After the marginal satisfaction scores are acquired, we suggest a relationship between these scores and corresponding overall scores such that it enables us to derive the importance weights of product attributes. The suggested relationship is to define overall satisfaction score of each product as an observation of random variable Z_i distributed with mixture density in Equation (2.13):

$$Z_i \sim \sum_{j=1}^{k} w_j \hat{f}(Y_{ij} = z_i \,|i, j) \tag{2.13}$$

where w_j indicates the corresponding attribute weights with $w_j > 0$,and $\sum_{j=1}^{k} w_j = 1$.

The above-mentioned relationship is mixture of kernel distributions. Weighted sum of the marginal densities of attribute scores forms the mixture distribution of overall customer satisfaction score. The weight w_j indicates impact of product attribute A_j on overall rating score.

MCMC is utilized to obtain these weights, in other words, to sample from posterior distributions of weights. The attribute weights are assumed to lie on k-dimensional simplex, which is a space in R^k and any data point $w = (w_1, w_2, \dots, w_k)$ satisfies the conditions that $w_j > 0$ and $\sum_{j=1}^{k} w_j = 1$. Dirichlet distribution is an appropriate distribution to express the attribute weights since it provides a sample vector whose elements sum to 1. Prior distribution of weight vector is derived from a Dirichlet distribution, and this distribution is also used as proposal to construct independence Markov chain. Consequently, Metropolis-Hastings acceptance ratio equals the likelihood ratio in this MCMC algorithm (Giudici et al., 2013, p. 204). The acceptance ratio is viewed in Equation (2.14).

$$R\left(w^{(t)}, w^{*}\right) = \frac{L(w^{*}|z_1, \ldots, z_i, \ldots, z_m)}{L(w^{(t)}|z_1, \ldots, z_i, \ldots, z_m)} \tag{2.14}$$

where w is weight vector, z are observed data with likelihood function L(w|z) for parameters w which have prior distribution $p(w)$. Meanwhile, likelihood of attribute weights is expressed with Equation (2.15).

$$L(w|z_1, \ldots, z_i, \ldots, z_m) = \prod_{i=1}^{m} \sum_{j=1}^{k} w_j \hat{f}(Y_{ij} = z_i). \tag{2.15}$$

Logarithmic transformation is utilized for more rational computation, thereby product operator is converted to sum operator in Equation (2.15). Acceptance probability is calculated according to logarithmic likelihoods.

Figure 2.4 presents the steps followed to obtain posterior distributions of attribute weights. In each iteration *t*, random value of vector w is proposed based on $Dirichlet(\alpha_1, \alpha_2, \ldots, \alpha_k)$, and the proposal is accepted considering acceptance probability and a uniformly distributed random number. Finally, they converge to a sample from $p(w|z)$ and weight vector $w = (w_1, w_2, \ldots, w_k)$ is explored.

1: Initialize $w^{(0)}$; set $t = 0$

2: For $t = 0$ to $N - 1$

3: Sample a point w^* from $Dirichlet(\alpha_1, \alpha_2, \dots, \alpha_k)$

4: Sample a random variable $u \sim Uniform\,(0,1)$

5: Calculate the acceptance probability:

6: $$R\left(w^{(t)}, w^*\right) = \min\{1, \ln\left(L(w^*|z_1, \dots, z_i, \dots, z_m)\right) - \ln\left(L\left(w^{(t)}|z_1, \dots, z_i, \dots, z_m\right)\right)\}$$

7: If $\ln(u) < R\left(w^{(t)}, w^*\right)$

8: $w^{(t+1)} = w^*$

9: else

10: $w^{(t+1)} = w^{(t)}$

11: end

12: end

Figure 2.4 MCMC algorithm for attribute weight inference

2.6 Case Study

In this section, we demonstrate how the proposed method is applied in exploring to what extent customers give importance to each technical attribute of the mobile phones sold in an e-commerce website of Hepsiburada (https://www.hepsiburada.com/). This website is one of the popular e-commerce website in Turkey, which allows their customers for posting their comments and ratings in the form of 1-5 star scale. We demonstrate the retrieved data by rectangle frames in Figure 2.5. The product names, information regarding to product attributes, overall ratings and the number of reviews are gathered into a database via web scraping software, namely, Octoparse (https://www.octoparse.com/).

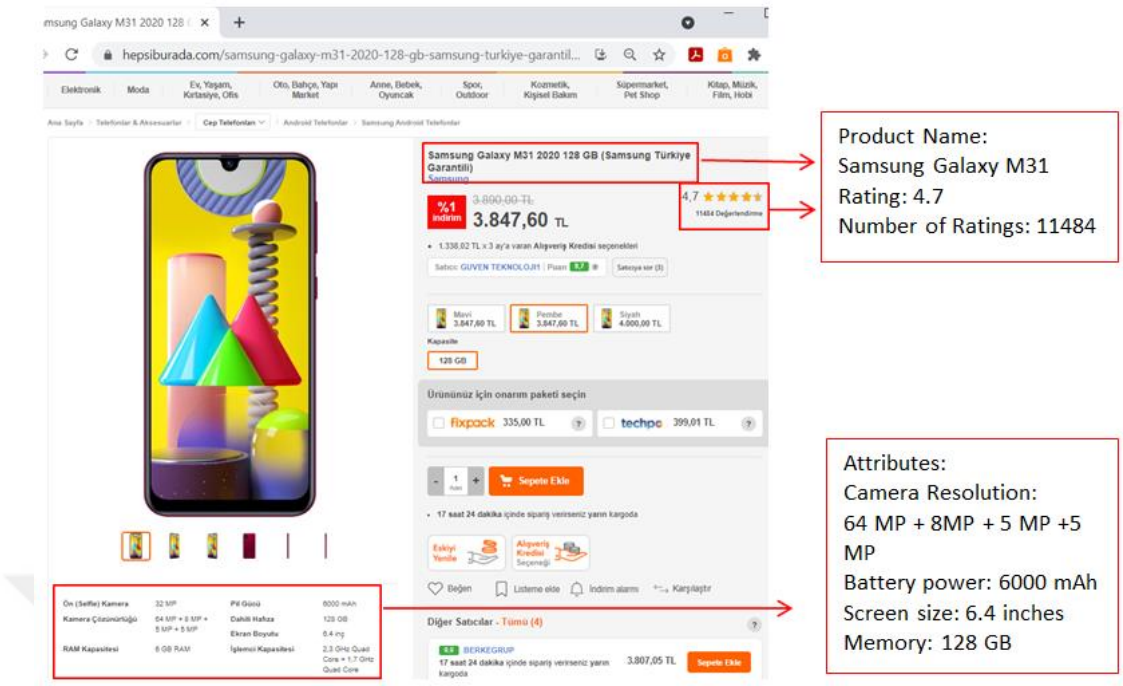


Figure 2.5 Screenshot of e-commerce webpage on Hepsiburada.com

After discarding the products with less than 50 ratings, the database includes 163 products. The considered product attributes are A_1 (Screen size), A_2 (Camera resolution), A_3 (Memory), A_4 (Battery power). These attributes are chosen since they are among commonly mentioned attributes in the website. The details of the collected data regarding to mobile phones are shown in Table 2.3. This table includes some descriptive statistics of product attributes, overall ratings and the number of reviews. Graphical representations of dataset are displayed in Figure 2.6. The variables except Memory are continuous. They are displayed via histograms with seven bins. The y-axes of the histograms give relative probabilities that are the ratio of the number of observations in the bins to the total number of elements in the input data. Some explanations from these figures are as follows. For instance, 32% of the mobile phones have 64 GB memory. 47.2% of the mobile phones have a screen size between 6.1 and 6.52 inches. The camera resolutions of 52.7% of the mobile phones are smaller than 24 MP. The percentage of the mobile phones with a battery size in the interval [3000 3800] mAh is 33. The higher percentages of the average online ratings are in the interval [4.45 4.7]. 75% of the mobile phones have reviews between 50 and 460.

Table 2.3 Descriptive summary of the dataset

Descriptive statistics		Mean	Standard deviation	Min	Max
Product attributes	Screen size (inch)	6.045	0.572	4	6.9
	Camera resolution (MP)	34.815	28.996	2	168
	Memory (GB)	81.128	63.671	8	512
	Battery (mAh)	3732.134	962.447	1560	7000
	Number of reviews	406.539	518.227	50	3193
Overall ratings	Average satisfaction scores	4.560	0.238	3.3	4.9

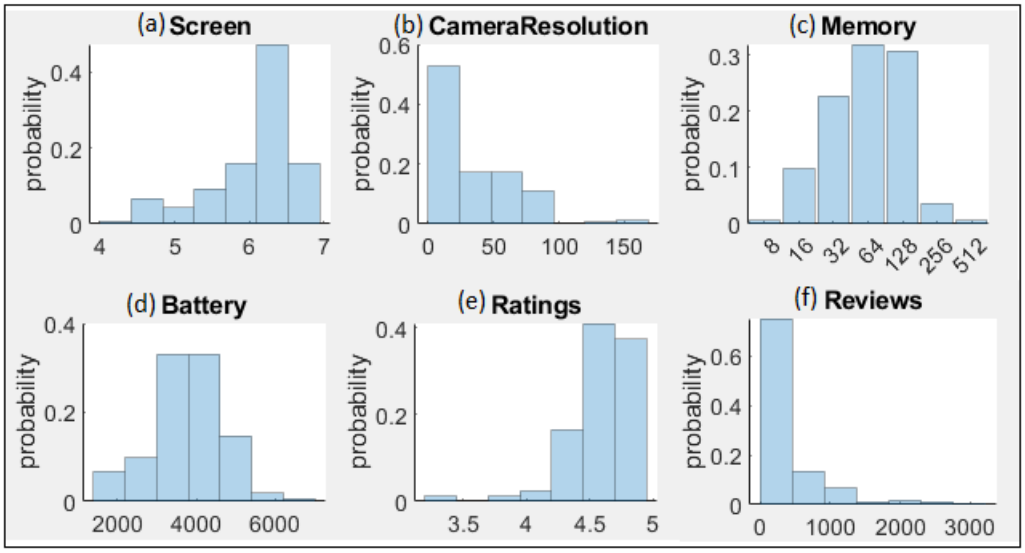


Figure 2.6 Histograms of the variables of a)Screen size (inch), b) Camera Resolution (MP), c) Memory (GB), d) Battery (mAh), e) Average online ratings, and f) Number of reviews.

Since the attributes excluding Memory are continuous attributes, they are discretized by using Equation (2.9). All attributes have seven categories of which logic is analogous with linguistic variables extremely low, very low, low, medium,

high, very high, and extremely high, but indications are displayed by symbols in Table 2.4.

Table 2.4 Result of discretization

	Attributes				
P_i	A_1	A_2	A_3	A_4	Overall score (z)
1	C_{17}	C_{22}	256 GB	C_{44}	4.6
2	C_{17}	C_{24}	128 GB	C_{47}	4.8
3	C_{16}	C_{24}	128 GB	C_{46}	4.7
⋮	⋮	⋮	⋮	⋮	⋮
163	C_{16}	C_{22}	128 GB	C_{44}	4.5

After estimating random variables and densities, we conduct MCMC algorithm for 10000 iterations with the first 1000 iterations discarded as burn-in samples. We check the convergence by acquiring convergence parameter $\hat{R} = 1.00$ obtained for 10 chains. Meanwhile, we control efficient sample sizes and autocorrelation plots. Symmetric Dirichlect distribution whose parameters are taken the same value is determined as prior distribution of weight vector and value of 1 is assigned to concentration parameter α. This means that Dirichlet distribution becomes the uniform distribution over the simplex. Concentration parameter is determined as 1 when there is no prior knowledge on weight vector. According to presence of prior knowledge on weights, concentration parameter can be changed, and thereby different posterior distributions can be revealed.

2.7 Results of Proposed Method

The proposed method is coded in MATLAB by using a computer with i7-106567 processor number and 1.30 GHz CPU. Gaussian, Uniform, Epanechnikov quadratic and Triweight kernel functions are separately utilized while distributions of random variables are estimated. As the result of simulations, the posterior distributions of attribute weights are displayed in the following Figure 2.7 - Figure 2.10.

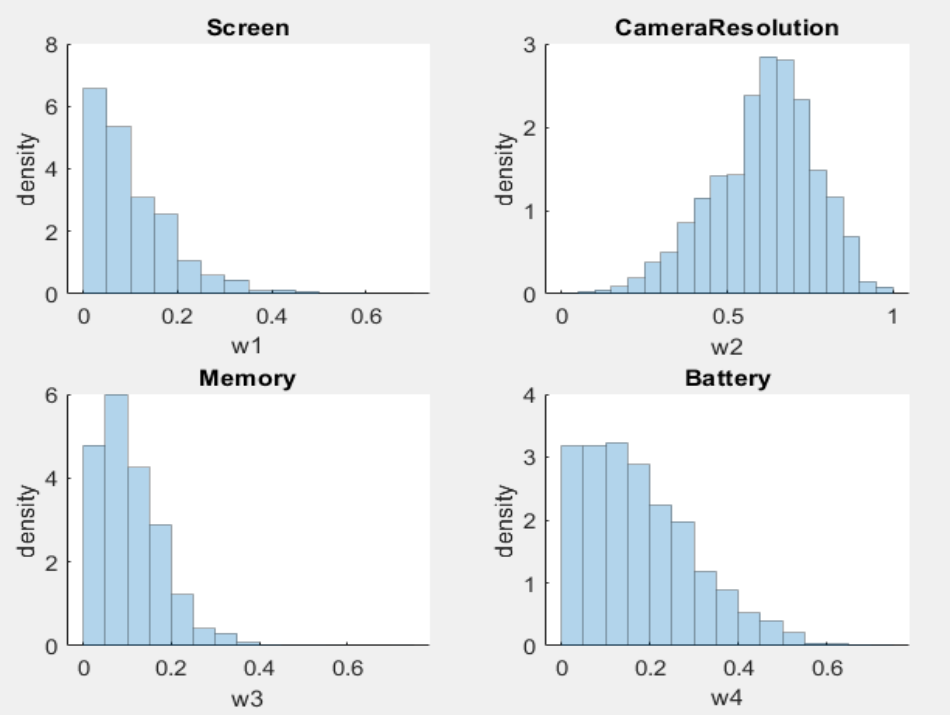


Figure 2.7 Posterior probability distributions of weights with Gaussian kernel: $\overline{\mathrm{w}} = [0.136, 0.497, 0.169, 0.196]$

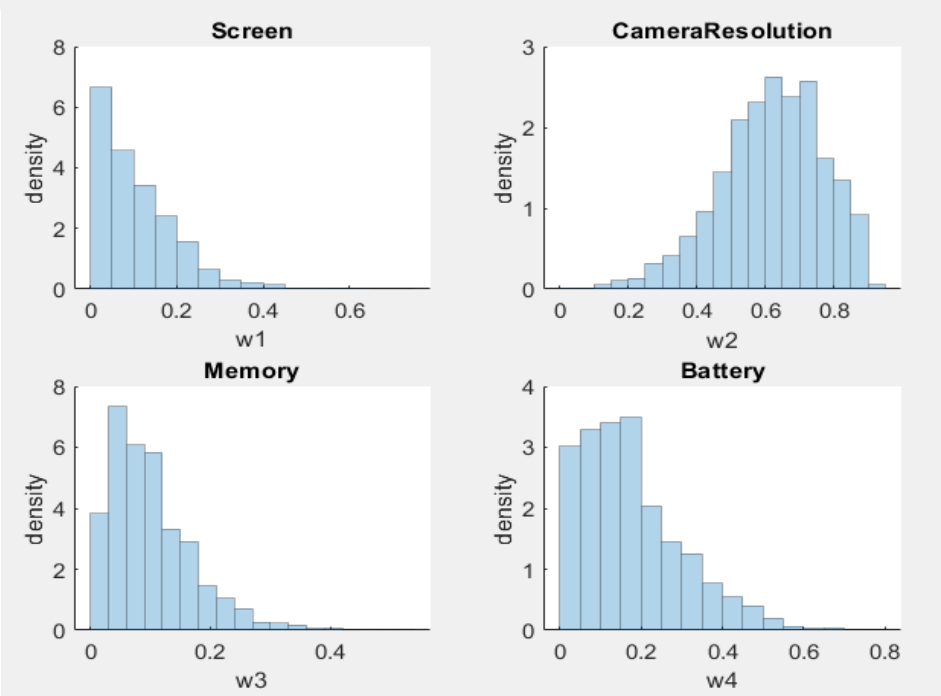


Figure 2.8 Posterior probability distributions of weights with Triweight kernel: $\overline{\mathrm{w}} = [0.139, 0.490, 0.159, 0.209]$

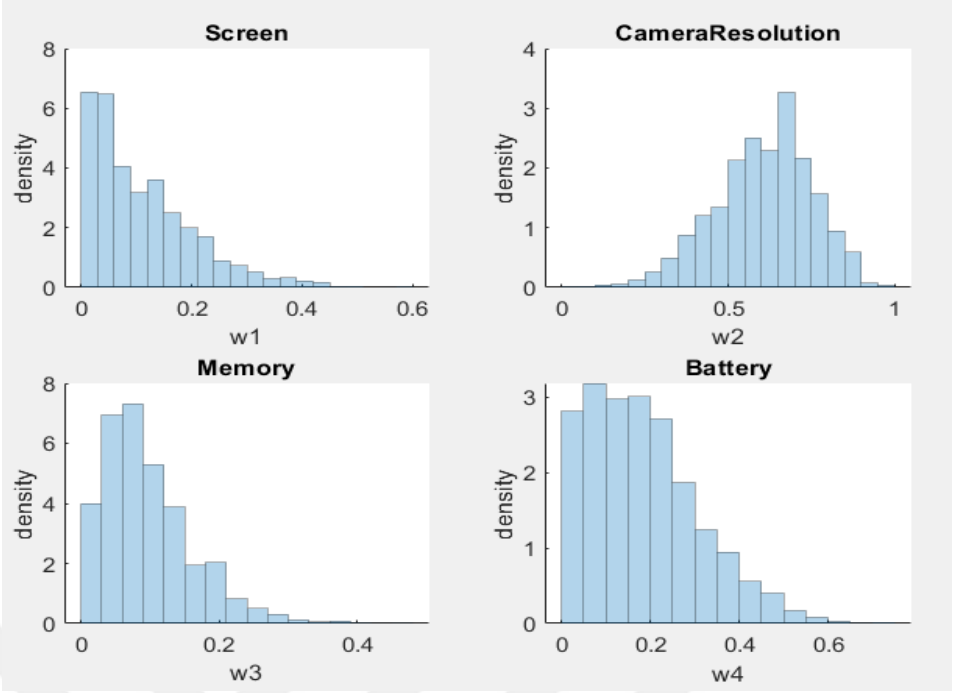


Figure 2.9 Posterior probability distributions of weights with Epanechnikov kernel: $\overline{w} = [0.143, 0.492, 0.152, 0.211]$

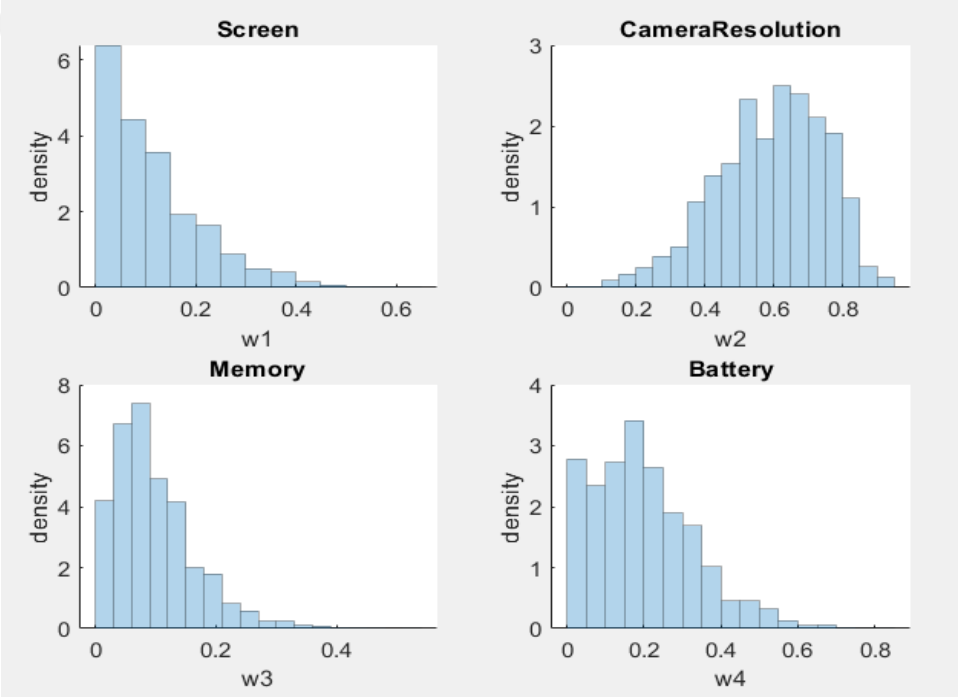


Figure 2.10 Posterior probability distributions of weights with uniform kernel: $\overline{w} = [0.141, 0.468, 0.147, 0.242]$

Posterior probability of order of attribute weights can be revealed by evaluating posterior sample values obtained from MCMC method (Musal et al., 2012). Table 2.5 shows some posterior probabilities regarding the order of weights with reference to different kernel functions. For instance, weight of A_4 (Battery) is higher than weight of Screen (A_1) at 67.6 percentages of simulated values with reference to Gaussian kernel. The order $w_2 \succ w_4 \succ w_3 \succ w_1$ is found at 26 percentages of simulated values, which is the highest proportion. The proportion rates of this ranking are the highest scores for the results with the other kernels as well. This ranking is found at 29.1 percentages of simulated values with Triweight kernel, at 28.3 percentages of simulated values with Epanechnikov quadratic kernel, and 28.8 percentages of simulated values with Uniform kernel functions. In evaluating these probabilities, it should be taken into account the fact that there are 24 combinations for order of four weights. Consequently, we deduce that the customers give more importance to Camera Resolution (A_2), whereas give less importance to Screen size (A_1) in evaluating a mobile phone.

Table 2.5 Posterior probabilities for the order of attribute weights according to different kernels

	Gaussian kernel	Triweight kernel	Epanechnikov Quadratic kernel	Uniform kernel
$P(w_2 \succ w_4)$	0.930	0.945	0.938	0.911
$P(w_3 \succ w_1)$	0.545	0.513	0.498	0.493
$P(w_2 \succ w_1)$	0.976	0.982	0.976	0.970
$P(w_2 \succ w_3)$	0.986	0.988	0.993	0.989
$P(w_4 \succ w_3)$	0.656	0.686	0.697	0.715
$P(w_4 \succ w_1)$	0.676	0.662	0.661	0.694
$P(w_2 \succ w_4 \succ w_3 \succ w_1)$	0.260	0.291	0.283	0.288
$P(w_4 \succ w_2 \succ w_3 \succ w_1)$	0.031	0.025	0.022	0.033
$P(w_2 \succ w_3 \succ w_4 \succ w_1)$	0.134	0.113	0.116	0.109
$P(w_2 \succ w_4 \succ w_1 \succ w_3)$	0.206	0.203	0.202	0.201

2.8 Comparative Study

In this section, comparisons regarding rankings are proposed and strengths of the method are explained. Nonlinear mathematical model, ordinary least squares (OLS) regression model and multilayer perceptron models are created. Five-fold cross validation is implemented for comparative study. This method splits data into five equal width folds. Four folds are trained to construct the model and the remained fold is used to test it. This process is repeated five times; thereby the average performance metrics of five models is taken. The performance metrics are mean squared error (MSE), root mean squared error (RMSE) and mean absolute percentage error (MAPE), which are presented in Equations (2.16)-(2.18).

$$MSE = \frac{\sum_{i=1}^{m}(z_i - \hat{z}_i)^2}{m} \tag{2.16}$$

$$RMSE = \sqrt{\frac{\sum_{i=1}^{m}(z_i - \hat{z}_i)^2}{m}} \tag{2.17}$$

$$MAPE = 100 * \frac{1}{m}\sum_{i=1}^{m}\frac{|z_i - \hat{z}_i|}{z_i} \tag{2.18}$$

where $\hat{z}_i$ is predicted overall satisfaction score of product, z_i is actual overall satisfaction score and *m* is the number of products.

2.8.1 Non Linear Mathematical Model

We express our problem as an optimization problem and write a mathematical model to solve it. Then, we compare the result of this model and the Bayesian approach with different kernels. Mathematical model considers mean of random variables as satisfaction scores of product attributes. Decision variables are scalar instead of random variables. It aims at minimizing MSE expressed in Equation (2.16).

Minimize

$$MSE = \frac{\sum_{i=1}^{m}(z_i - \hat{z}_i)^2}{m}$$

Subject to:

$$\sum_{j=1}^{k} w_j * \bar{Y}_{ij} = \hat{z}_i, \qquad \forall i \tag{2.19}$$

$$\sum_{j=1}^{k} w_j = 1 \tag{2.20}$$

$$w_j \in (0,1] \tag{2.21}$$

In objective function, $\hat{z}_i$ is predicted overall satisfaction score of product and z_i is actual overall satisfaction score. The predicted overall score $\hat{z}_i$ is computed with constraint (2.19), where $\bar{Y}_{ij}$ denotes mean of random variable Y_{ij}, satisfaction score of P_i with reference to A_j acquired by Equation (2.10). Constraints (2.20) and (2.21) express decision variables of weight vector w.

According to five-fold cross validation conducted in MATLAB, average MSE of non-linear optimization method solved by interior-point-convex algorithm is obtained as 0.0483. Decision variable vector with lowest MSE is obtained as (4.94E-08 0.606 0.393 2.80E-08). Camera Resolution has the highest impact on overall score according to mathematical model. The ranking of weights is $w_2 \succ w_3 \succ w_1 \succ w_4$. The average MSEs of Bayesian method with different kernels are similar. The obtained scores are 0.0386, 0.0387, 0.0387 and 0.0388 with Uniform, Triweight, Epanechnikov quadratic and Gaussian kernels, respectively. The final ranking regarding impact of product attributes is $w_2 \succ w_4 \succ w_3 \succ w_1$.

Bayesian methodology is superior to optimization model in terms of computational efficiency as well. It takes about 41 sec for 326000 kernel function assessment in Metropolis-Hastings sampling. However, it takes about 86 sec to obtain feasible solution for optimization model with one objective function.

Moreover, Bayesian methodology deals with performance scores represented as random variables. The weights are acquired as random variables. On the other hand, optimization model only considers expected values of performance scores regarding attributes. It does not give importance weights of attributes as random variable. Metropolis-Hastings is a bootstrap sampling from the nonparametric kernel density model. A simulation can be conducted even though having incomplete dataset and the impact of product attributes are estimated as random variables according to simulation result. Therefore, we have realized more comprehensive analysis with Bayesian approach considering uncertainties.

2.8.2 Linear Regression Analysis with Categorical Predictors

The attributes Screen size (A_1), Camera Resolution (A_2), Memory (A_3), Battery (A_4) are considered as categorical predictors. Each attribute has seven categories as mentioned in section of Case Study. The first categories of categorical predictors (C11, C21, C31, and C41) are determined as reference levels in the model. Linear regression model of overall satisfaction scores is created as a function of attributes (Alsaffar, 2017). The formal model is given by Equation (2.22).

$$Z = \beta_0 + \beta_1 * I_{(A_1=C12)} + \cdots + \beta_6 * I_{(A_1=C17)} + \cdots + \beta_{19} * I_{(A_4=C42)} + \cdots + \beta_{24} * I_{(A_4=C47).} \quad (2.22)$$

where $I_{(A_i=C_{jk})}$ are indicator variables. For instance, the value of $I_{(A_1=C12)}$ is one if the value of Screen size (A_1) attribute is in category of C12. The overall satisfaction scores are indicated with vector of Z.

This regression model is created and solved in MATLAB. Table 2.6 displays regression coefficients (β), in other words, estimated contributions of attribute categories to overall satisfaction scores. The average MSEs of this regression model is 0.0581. Consequently, the proposed Bayesian method gives better result in cross validation comparison. It is important to highlight that regression analysis is not for weight determination. We have just found the impact of each attribute category to

overall satisfaction score. The range of coefficients for each attribute is determined according to maximum and minimum values of coefficients. A high range value indicates that the related attribute has higher impact on change of the customer satisfaction score. Memory (A_3) has the highest impact on change of overall customer satisfaction scores.

Table 2.6 Results of regression analysis

Attributes	Categories	Coefficients	Std. Error	Range	Importance values
A1	C12	0.118	0.135	[-0.088 0.126]	0.159
	C13	0.166	0.170		
	C14	0.089	0.178		
	C15	0.116	0.192		
	C16	0.126	0.192		
	C17	-0.088	0.202		
A2	C22	0.383	0.179	[0 0.382]	0.285
	C23	0.270	0.157		
	C24	0.066	0.160		
	C25	0.345	0.162		
	C26	0.119	0.165		
	C27	0.231	0.171		
A3	C32	0.366	0.290	[0 0.519]	0.387
	C33	0.462	0.291		
	C34	0.511	0.300		
	C35	0.520	0.303		
	C36	0.436	0.320		
	C37	0.343	0.391		
A4	C42	-0.127	0.140	[-0.186 0.037]	0.167
	C43	-0.186	0.160		
	C44	-0.122	0.149		
	C45	-0.015	0.154		
	C46	-0.014	0.180		
	C47	0.038	0.180		
	(Intercept)	3.8	0.2142		

2.8.3 Multilayer Perceptron Neural Network

Multilayer Perceptron (MLP) Neural Network is one of the most widely utilized classes of neutral networks that can be effectively implemented to different complex problems (Baykasoğlu & Baykasoğlu, 2017; Kulluk et al., 2013; Shepherd, 2012). MLP is a supervised method which enables comparison of model-predicted results with known values of target variable. MLP includes three main parts: one input layer, one or more hidden layers and one output layer. The input layer contains the predictors, the hidden layers contain unobservable nodes, and output layer contain dependent variables. MLP uses back propagation learning algorithm for training process. In a MLP, an activation function is applied to each unit so as to deal with the potential nonlinearities between variables. Hyperbolic Tangent (Tanh) is one of the most well-known transfer functions. MLP is briefly introduced since it is out of scope of this research. More detailed information on MLP can be found in several studies (Shepherd, 2012; Xie et al., 2022).

MLP is implemented in IBM SPSS Neural Networks 28. Sum of squares error is utilized as the performance metric. Table 2.7 shows specifications of the best architecture of the MLP.

Table 2.7 MLP network information

Input layer	Factors	Screen
		Camera
		Memory
		Battery
	Number of units[a]	28
Hidden Layer(s)	Number of Hidden Layers	1
	Number of Units in Hidden Layer 1[a]	1
	Activation Function	Hyperbolic tangent
Output Layer	Dependent Variables 1	Overall satisfaction scores
	Number of Units	1
	Rescaling Method for Scale Dependents	Standardized
	Activation Function	Identity
	Error Function	Sum of squares

[a]Excluding the bias unit

According to five-fold cross validation, the average MSEs of MLP model is calculated as 0.0421. In addition, IBM SPSS Neural Networks 28 allows to perform sensitivity analysis, which computes the importance of each predictor in determining the neural network using variance-based technique (IBM Corp., 2022b; Saltelli et al., 2004). Predictors are ranked according to the sensitivity measure that is defined as follows:

$$S_i = \frac{V_i}{V(Y)} = \frac{V(E(Y|X_i))}{V(Y)} \tag{2.23}$$

where V(Y) is the unconditional output variance. In the numerator, the expectation operator E calls for an integral over X_{-i} that is, over all factors but X_i, then the variance operator V implies a further integral over X_i. Predictor importance

is then computed as the normalized sensitivity. The total sum of the overall variable importance of the ANNs is 1 (IBM Corp., 2022a).

The importance values for attributes are 0.156, 0.322, 0.294, 0.229, respectively. Sensitivity analysis showed that Camera Resolution (A_2) is the most significant determinant of overall customer satisfaction scores.

In summary, Bayesian approach gives better performance metrics than other models. Table 2.8 presents the performance metrics and rankings regarding to the attributes. In addition, it is important to highlight the main purposes of OLS regression analysis and MLP are not weight determination. The obtained importance values give the relative contribution of each input variable to prediction. Moreover, Bayesian approach gives importance weights of attributes as random variable. The weights acquired by Bayesian approach have direct contribution in calculating overall performance scores.

Table 2.8 Comparative study results

	MSEs	RMSEs	MAPEs	Rankings
The proposed method with Gaussian Kernel	0.0386	0.1867	2.81	$w_2 \succ w_4 \succ w_3 \succ w_1$
The proposed method with Triweight Kernel	0.0387	0.1869	2.81	$w_2 \succ w_4 \succ w_3 \succ w_1$
The proposed method with Epanechnikov Kernel	0.0387	0.1871	2.81	$w_2 \succ w_4 \succ w_3 \succ w_1$
The proposed method with Uniform Kernel	0.0388	0.1873	2.81	$w_2 \succ w_4 \succ w_3 \succ w_1$
OLS Regression	0.0581	0.2349	3.85	$A_3 \succ A_2 \succ A_4 \succ A_1$
MLP Neutral Network	0.0421	0.1948	3.10	$A_2 \succ A_3 \succ A_4 \succ A_1$
Non-Linear Mathematical Model	0.0483	-	-	$w_2 \succ w_3 \succ w_1 \succ w_4$

2.9 Chapter Conclusions

This thesis proposes a novel method for providing the impact of product attributes on customer satisfaction by analyzing online customer ratings. In the method, satisfaction scores of products with respect to attributes are represented by random variables whose values are extracted from overall scores such that the scores of products belonging to same category concerning an attribute are considered as observations of these random variables. The probability distribution functions of the random variables are calculated via kernel density estimation method. Overall scores are considered as values from mixture distributions made up of random variables pertaining to product attributes. Impacts of product attributes on overall scores are mixing proportions of random variables regarding product attributes, in other words, weight vector. We inference this weight vector by Bayesian inference algorithm, namely MCMC.

The proposed method is applied to analyses of online customer ratings on mobile phones. Based on our dataset, we acquire that camera resolution has highest impact on customer satisfaction. In addition, we conduct cross validation between Bayesian method, non-linear mathematical programming model, OLS regression model and MLP neutral network model. As a result, we get the better performance scores by employing the proposed Bayesian model. Moreover, different kernels are utilized in kernel density estimation such as Epanechnikov, Uniform, Triweight and Gaussian kernels. The results show that the kernel choice has not effect on rankings of weights. We obtain similar rankings and MSE scores with all of them. Section 2.9.1 summaries the major contributions of this study together with its practical implications.

2.9.1 Contributions and Practical Implications

Our study explores the weights of product attributes hidden in the customer satisfaction scores. This process is like deductive reasoning in reverse engineering, which gains knowledge from a finished product. We have attempted to find out how

the ratings of products are realized in minds of customers through an innovative analytical method based on Bayesian theory. This method uses the overall ratings such a finished object in reverse engineering, and attribute information of the rated products, and then extracts the weights of product attributes in the composition of the overall rating scores of these products. The overall rating scores, in other words, customer satisfaction scores are composed of the satisfaction level on each product attribute. What is the weight of each product attribute in this composition? This is the research question of the present study.

The weights, in other words, the output of the proposed method, can be used in several applications. Product ranking studies based on MCDM can use these weights as criterion weights. Therefore, this method can be regarded as a weighting method for MCDM analysis. It is worth mentioning that the proposed method is not only applicable to the case studies on customer satisfaction if it will be used in MCDM analysis. Although the analysis of online customer ratings is addressed in this thesis, it can also be applied to other decision-making problems, in which only data on overall satisfaction scores of alternatives are available. In other words, this study also proposes a new solution methodology in acquiring weights to criteria for MCDM analysis in an uncertain environment via Bayesian inference. To the best of our knowledge, it is the first study that extracts criterion weights by considering only overall performance scores of alternatives based on kernel density estimation and Metropolis-Hastings algorithm.

From the perspective of the case study, the present study fulfills the research gap in customer satisfaction literature by implementing the Bayesian solution in discovering the weights of product attributes on customer satisfaction scores expressed via online ratings. Product recommendation studies can make use of these weights to find appropriate products to be proposed to customers. The products with superior in the characteristic, to which our method assigns the highest weight, can be recommended on the web-sites. Adding such specialized product recommendation systems to the websites can increase the purchasing of the customers in these e-commerce websites. The proposed method has a clear logic and a simple

computational process to support the companies in making decisions in product design and to gain insight into customer preferences for marketing. Advertising messages of the products can be defined according to these weights. Advertising companies might seek to call customer attention to the product by highlighting the superiority of the product attribute with the highest weight.

In terms of future research, the proposed method can be applied to online product ratings from other markets such as laptops, camera etc. The extended methodology might be developed based on online reviews as well as online ratings to obtain wider knowledge from customer opinions.

CHAPTER THREE
RANKING HOTELS BASED ON BAYESIAN NETWORK AND SENTIMENT ANALYSIS WITHIN A MULTI-CRITERIA DECISION-MAKING FRAMEWORK

3.1 Chapter Introduction

Developments in information technologies, usage of social media, and changes in customer shopping choices tending to e-commerce websites rather than shopping centers pave the way for an explosion of online evaluations on various platforms. These online evaluations are helpful sources of information about customer opinions regarding a product or service, and they support customers and producers in effective decision-making. Customers search for online evaluations on e-commerce websites regarding alternative products such as mobile phones before making purchasing choices; similarly, travelers intend to read opinions related to alternative hotels posted to tourism websites by experienced tourists in order to optimize their booking decisions and choose the most appropriate alternative hotels. However, it is impossible for a customer to read a tremendous amount of reviews, compare alternatives by taking into account their several conflicting features, and make a rational decision. The researchers propose various data-driven decision-making models to aid customers in extracting valuable information from online data in a systematic and rapid way, one of which is based on mining online reviews and ranking the products within a MCDM procedure.

Consumers convey their requirements, satisfaction, and problems through online reviews, which exist in two forms: ratings and textual reviews (L. Wang et al., 2020). Online reviews regarding products are generated without the intervention of any agents aside from consumers. Consumers might choose to assign a rating to any criterion while posting textual reviews regarding the other criteria (Nie et al., 2020). Therefore, it is more valuable to analyze textual reviews as well as ratings. In order to analyze textual reviews and transform them into numerical scores like ratings, sentiment analysis is performed. Sentiment analysis, a task of Natural Language

Processing (NLP), is a computational study of extracting sentiment and opinions from texts (Birjali et al., 2021). The task of sentiment analysis is mainly performed at the document level, sentence level, or aspect level. Aspect-based sentiment analysis is within the scope of the study since it focuses on the identification of sentiment orientations regarding specific aspects of entities. Numerous review articles about sentiment analysis have been proposed. Interested readers can learn more details regarding this technique by examining one of the latest research papers (Birjali et al., 2021).

Product/service ranking based on online reviews can be regarded as an MCDM problem, which includes three main stages (Fan et al., 2020). The first stage begins with extracting product attributes to be evaluated in comparison to alternative products. Data mining techniques are employed in exploring product attributes. This stage refers to the determination of criteria in the MCDM procedure. The second stage is identifying sentiment orientations of product attributes through online review analysis. The online ratings and/or text reviews transformed into scores via sentiment analysis are utilized to express the performance values of alternatives regarding criteria and criteria weights; thereby, decision matrices are created. The last stage concludes with the ranking of alternatives by application of MCDM techniques.

The sentiment analysis used for text reviews is primarily divided into two types: lexicon-based sentiment analysis and machine learning-based sentiment analysis. Since lexicon-based sentiment analysis is suitable for sentence-level sentiment analysis (Y. Liu et al., 2017), we have obtained the sentiment orientations of criteria mentioned in text reviews using the lexicon-based sentiment analysis approach in this study. The proposed sentiment analysis algorithm has several important features. It follows distance rules and semantic rules for correctly matching aspect words and sentiment words, thereby accurately identifying the sentiment orientations of criteria in the reviews. Rather than utilizing bilateral classification, the sentiment analysis algorithm makes it possible to assign five-scale labels to the sentiment orientations of criteria; hence, the sentiment strengths in the reviews can be captured.

There is a growing interest in developing decision-support models that analyze online reviews (Çalı & Balaman, 2019; Çalı & Baykasoğlu, 2022; Fan et al., 2018; Liang et al., 2019; Luo et al., 2021; Nie et al., 2020; H.-G. Peng et al., 2018; Yu et al., 2018, 2020). However, existing studies have different shortcomings, such that the dependence relationships between criteria are neglected. The assumption of independence between criteria is not realistic, and it is necessary to capture statistical dependencies between criteria. A BN is a probabilistic graphical model which displays the probabilistic relationships between a large number of variables in a system and performs probabilistic inference with those variables (Neapolitan, 2004). BNs have been used for various projects to date, including medical research (Park et al., 2018), healthcare analytics (Topuz et al., 2018), supply chain risk management (Hosseini & Ivanov, 2020), customer satisfaction analysis (Chakraborty et al., 2016), education (Pietro et al., 2015), and comparison of classification algorithms (Y. Peng et al., 2011). BN is an appropriate machine learning tool that discovers statistical dependencies between random variables and allows decision-makers to conduct scenario analysis with an inference mechanism. This study makes use of BN to obtain the probabilistic relationship between the criteria of the considered MCDM problem and capture the relative importance of the criteria.

The MCDM method is performed in the last stage to obtain product ranking. However, the traditional MCDM methods such as SAW, TOPSIS, ELECTRE, VIKOR, and PROMETHEE assume that the criteria are independent (Tzeng & Huang, 2011). They are not capable of capturing interrelationships between the criteria. Therefore, this study integrates BN with GTMA to rank the alternatives more conveniently.

To sum up, the main contributions of this thesis to the literature are as follows:

- The methodology described in this thesis uses BN in an MCDM framework to analyze online reviews in order to identify relationships between customer satisfaction regarding various product criteria and, as a result, to support decision-making for both potential customers and the associated businesses.

- The suggested methodology uses BN analysis to evaluate the relative weights of the criteria that will be applied to the MCDM analysis's product ranking system. In other words, it presents BN-GTMA model.
- BN is built from both numerical ratings and textual reviews that have been converted into numbers using a proposed aspect-based sentiment analysis method. The aspect-based sentiment analysis method takes into account the various sentiment degrees in the reviews.

The remainder of this chapter is structured as follows: Section 3.2 mentions related works. Section 3.3 provides a description of the problem and notations. The proposed approach is described in detail in Section 3.4. In Section 3.5, a case study dealing with hotel reviews is provided to demonstrate the reliability of the ranking system. Section 3.6 gives the comparative analysis and Section 3.7 gives the conclusions.

3.2 Related Works

3.2.1 Studies on Product/Service Ranking through Online Reviews

Numerous researchers have been interested in studies on product/service ranking techniques based on online customer reviews; some of them consider either online ratings or text reviews, while others have analyzed both of them.

Yu et al. (2018) associated online ratings with linguistic terms and proposed a MCDM process based on linguistic distribution assessments. Alternative hotels are ranked via the VIKOR method, and the weights of the criteria are acquired by considering the frequencies of the evaluation values on the criteria. H.-G. Peng et al. (2018) proposed a probabilistic linguistic integrated cloud model to analyze online ratings for the hotel selection problem. They integrated online evaluations on different criteria via the proposed information fusion tools based on the Heronian mean operator, considering their interrelationships, and obtained the criteria weights by optimization models maximizing deviations. Yu et al., (2020) proposed a

stochastic dominance approach for hotel ranking, expressing online ratings on hotel criteria by probabilistic linguistic term sets in which the frequency of each common review is defined as the probability value of the related linguistic term.

Besides online ratings, customers can express their opinions on the products or services with texts and valuable information might be discovered from these reviews. Liang et al. (2019) transformed the sentiment orientations on hotel criteria extracted from text reviews into the format of distribution linguistic assessments. They combined the frequency of words and the deviation degree of criterion evaluations to calculate criteria weights and ranked the hotels through DL-VIKOR. Çalı & Balaman (2019) proposed a method that analyzes text reviews and realizes the ranking via intuitionistic fuzzy ELECTRE integrated with VIKOR. They deal with hotel ranking as a case study, and the weights of hotel criteria are acquired by the entropy method. Nie et al. (2020) proposed a hotel selection methodology based on text reviews. They handled criteria value, location, sleep quality, room, cleanness, and service as provided by TripAdvisor and used word2vec and the concept of similarity degree to extract words expressing these criteria and sentiment words. They build a semantic partitioned sentiment dictionary in which the sentiment words are mapped into the seven sentiment scales by considering semantic rules. Linguistic distribution assessments are used to express the overall evaluation value regarding each criterion for each hotel, and the ranking of hotels is obtained by an evidence-based approach. The criteria weights are computed by an optimization model that maximizes deviations. L. Wang et al. (2020) analyzed the preferences of different types of hotel customers in choosing hotels based on text comments. Key factors regarding criteria are determined by word2vector, and criteria weights are calculated with the term frequency-inverse document frequency algorithm. Finally, the picture fuzzy TODIM method proposed the most convenient hotel for each type of hotel customer. Luo et al. (2021) proposed an online evaluation ranking method for tourism attractions based on probabilistic linguistic term sets and the integrated determination of objective criteria weights (IDOCRIW)-combined compromise solution (COCOSO) model.

Aside from the studies on the assessment and choice of hotels, there are valuable studies addressing product ranking and weight determination through online reviews. Fan et al. (2018) transformed the online ratings of the product on several criteria into discrete percentage distributions, computed the stochastic dominance degrees for comparing each product pair regarding each criterion, and made use of PROMETHEE-II for ranking the alternative products. They applied the method to the case study of automobile ranking, in which the criteria weights are directly provided by the consumer. Çalı & Baykasoğlu (2022) developed a Bayesian-based approach in which online ratings and product characteristics are analyzed to discover the impacts of product attributes on overall customer satisfaction. F. Liu et al. (2023) proposed a physician selection method based on online reviews handling interactive criteria through the integration of MCDM and a machine learning method called principal component analysis. They considered the recommendation degree as a criterion, whereas we have considered the recommendation choice (the target variable in the BN) as a tool to reach the relative importance between criteria from a different perspective.

The above-mentioned studies have proposed several important contributions to the solution of product/service ranking based on online reviews. However, the comprehensive analysis that discovered the probabilistic relationships between the criteria has been lacking in the literature. Therefore, this is the main motivation of our study, in which BN analysis is performed.

3.2.2 Studies Using BN in MCDM

A literature review is conducted for exploring studies dealing with BNs in the MCDM framework. Expert-driven BN is more common than data-driven BN in the studies integrating MCDM with BN.

Fenton & Neil (2001) proposed a generic decision-making approach in which BNs and MCDM are used in complementary ways. BNs enable to model of uncertain

criteria depending on various factors, and hence to predict performance scores of these criteria for given alternatives to be analyzed with any MCDM method.

Kaya & Yet (2019) employed the output of DEMATEL to construct an expert-based BN, and applied their methods to a supplier selection problem. They used the ranked node method to obtain conditional probabilities. To eliminate cycles between criteria of product quality and quality system certification, multiple time frames are modeled.

Yang et al. (2019) proposed a BN network model constructed based on DEMATEL's causality diagram to determine the key factors in ensuring the development of metropolitan economy and logistics coordinately. The conditional probabilities are determined via questionnaires to which the experts respond with linguistic variables characterized by triangular fuzzy numbers.

Meng et al. (2019) proposed an integrated method utilizing DEMATEL-BN to foresee the probable accidents triggered by leakage on offshore platforms. DEMATEL has no role in the construction of BN; it just specifies critical risk indicators. BN is structured based on the Fault Tree and Event tree model and used to calculate posterior probabilities of indicators.

F. Li et al. (2019) combined DEMATEL, Interpretive Structural Modeling, and BN to analyze accident causations of urban buried gas pipeline networks. Integration of ISM to DEMATEL enables to visualization of the dependencies of risk factors in a hierarchical network model to be employed in construction of BN later.

Zhang et al. (2020) presented an approach for risk analysis of foundation pit collapse during construction based on fuzzy BN, of which probability parameters are revealed by using fuzzy AHP and considering the confidence levels of experts conducting pairwise comparisons. Expert evaluations in eliciting probabilities are expressed with triangular fuzzy numbers, but their crisp values are used in the inference process.

Zubair et al. (2015) proposed a paper aiming to prioritize nuclear accident precursors and select the most important ones based on a combination of BBN and AHP methods and handling the Fukushima accident as a case study. Expert-based BBN is employed to reduce the number of precursors to be evaluated with AHP to obtain the importance weights, and then another BBN is constructed to analyse posterior probabilities of the precursors, in which prior probabilities are estimated by experts and conditional node tables are formed based on AHP weights.

Ghavami et al. (2020) proposed a methodology assessing risk of failure in urban sewer pipelines, and prioritizing the inspection of the sewer pipes to prevent any environmental, economic, or social threat. The risk of failure for a sewer pipe is computed by the product of Probability of Failure (PoF) and Consequence of Failure (CoF) regarded as quantification of the failure impact of sewer pipeline. The probabilities of failures are obtained using BBN of which the learning process is performed with data acquired from reports, interviews, and GIS maps. On the other hand, the consequences of failures are computed with DEA where the criteria weights are obtained by AHP through experts' assessments, and the criteria scores are calculated with the help of GIS and its spatial analyses.

Yang et al. (2009) developed MCDM methodology for safety management expressing decision attributes such as alternatives and criteria with nodes in BN. They define the states of attribute nodes with linguistic variables characterized by fuzzy numbers. After the calculation of prior and posterior expected utilities based on fuzzy and probability measures, the differences of them are put into the decision matrix. To obtain criteria weights, they combine two types of weight measures while one is obtained via AHP; the other is obtained by entropy method based on differences in expected utilities. Fuzzy TOPSIS is employed for ranking. The proposed method is applied to a container transportation delay problem.

Roozbahani et al. (2018) predicted monthly groundwater levels using Dynamic BNs. Then, they created different scenarios and updated the network to predict the future condition of groundwater according to these scenarios. Finally, they used

PROMETHEE-II to rank these scenarios according to criteria as economic, social, and technical.

Zhao et al. (2020) analyzed several factors and sub-factors leading to the escalation of domino effects in the coal gasification process by using the fuzzy Analytical Hierarchy Method and estimating the probabilities of domino effects via BNs in which each device represents a node. The conditional probabilities, in other words, escalation probabilities of domino effects are determined with the result of FAHP, in which the experts respond to the questions on the importance degrees of the factors by using linguistic variables characterized by fuzzy numbers.

Li et al. (2020) presented a risk assessment model for gas explosion in a coal mine including FAHP and expert-based BN. FAHP is employed to obtain the importance weights of experts by considering subjective reliability related to experts for each node and objective factors such as job fields, education level, and experience. After obtaining fuzzy probability scores given to nodes by experts, these scores are multiplied by expert weights. The prior and conditional probabilities are computed after defuzzification and normalization processes. Causal inference gives the risk level of a gas explosion in the handled coal mine. The diagnostic inference gives posterior probabilities of risk factors when observing a gas explosion and sensitivity analyses give the most important risk factors. Decision-makers can strengthen these risk factors to prevent gas explosion accidents.

Nepal et al. (2013) proposed a framework to prioritize customer satisfaction attributes in the product development stage. They construct a network representing relationships among attributes and other criteria that affect the decision-making process of the company. Firstly, they obtain weighting via ANP based on pairwise comparisons. Then, BN enables experts to express their judgments on prior probabilities of attributes and criteria, as well as conditional probabilities which indicate the belief probabilities of criteria for each state when each state of attributes is known. These conditional probabilities are reduced to belief probabilities aggregating all states by using the utility function. Afterward, the obtained marginal

probabilities of attributes, in other words, BBN's weights are combined with ANP's result to present final weightings. Finally, ANP-BBN results are compared with the AHP method, and sensitivity analyses are conducted which include the changes in causal relationships.

Y. Huang & Bian (2009) proposed an intelligent tourist attraction recommendation system integrating BN, ontology theory, AHP, and spatial web service technology. The system offers a convenient place to a user based on MCDM analysis handled both BN responses considering online travel information on traveler behaviors and spatial information accessed through GIS.

Maiyar et al. (2019) proposed a study aiming to identify the structural relationship between product attributes and prioritize customer preferences regarding these attributes while exploiting textual social media data derived from fashion blogs in Germany. The problem is formulated as a BN structure learning problem with multiple products attributes as nodes and their corresponding parent–child relationships as directed arcs connecting the nodes. The proposed NP-hard problem is solved by using four hybrid metaheuristics self-learning particle swarm optimization (SLPSO) with the K2 algorithm (SLPSOK2), particle swarm optimization/K2score, Genetic Algorithm/K2 score, and ant colony optimisation/K2 score.

In sum, DEMATEL is the most popular MCDM method used in the integration of BN. The others are AHP, ANP, TOPSIS, and PROMETHEE. The other motivation of our study is to make use of a different MCDM method, which has been addressed before neither in product/service ranking through online reviews nor in the integration of data-driven BN with MCDM. Compared to the other MCDM methods, GTMA has more favorable characteristics (Baykasoglu, 2014; Rao, 2006). GTMA's computational process is rather straightforward. It allows the decision analysts to deal with the interrelationships between the criteria and illustrate them with graphical representations. The matrix permanent concept used in this method aids in better understanding the criteria and describing the decision-making problem since it

includes all potential structural components of the criteria and their relative importance. Even a slight change in the performance values of the criteria might significantly alter the outcome, so it is simpler to rank the alternatives with obvious differences in decreasing order.

3.3 Problem Description

Customers evaluate alternatives by considering several criteria, and they make a final decision about whether or not they are satisfied. It is considerable to discover the relationships between the satisfaction degrees for criteria and final satisfaction decisions to infer significant insights for businesses. The present study aims to offer a product ranking system that considers statistical relationships obtained through BN analysis. The system enables companies to infer information on which criteria of the product customers give more importance when recommending it. The system aids potential customers in their purchase decisions by proposing product ranking.

The variables and sets used in the methodology are provided below, along with their notations and explanations.

- $H = \{H_1, H_2, \dots, H_m\}$: a set of products/services which customers evaluate with online reviews in e-commerce website, where H_i indicates ith product. The expression 'alternatives' point out hotels for our case study.
- $C = \{C_1, C_2, \dots, C_n\}$: a set of criteria of the product, where C_j indicates the jth criterion.
- $W = \{w_1, w_2, \dots, w_k\}$: vector of weights of criteria, where w_j indicates the weight of jth criterion with the constraint that $w_j > 0$ and $\sum_{j=1}^{k} w_j = 1$.
- $AE_j = \{AE_j^1,\ AE_j^2, \cdots, AE_j^p\}$: a word list including several aspect expressions AE_j^p (noun or noun phrase) describing the criterion C_j.
- $D = \{D^{++}, D^{+}, D^{-}, D^{--}, D_{hes}\}$ describes the set of sentiment lexicons which includes strongly positive sentiment lexicons denoted by D^{++}, positive sentiment lexicons denoted by D^{+}, negative sentiment lexicons denoted by D^{-},

strongly negative sentiment lexicons denoted by D^{--} and hesitant sentiment lexicon denoted by D_{hes}.

- $Adv = \{Adv^{+}, Adv^{-}\}$: a set of adverb lists. Adv^{-} is list of adverbs that intensifying only negative sentiments.
- $D^{++} = D_j^{++} \cup D_{gen}^{++}$ includes strongly positive sentiment lexicons created for each criterion and strongly positive general sentiment lexicon.
- $D^{+} = D_j^{+} \cup D_{gen}^{+}$ includes positive sentiment lexicons created for each criterion and positive general sentiment lexicon.
- $D^{--} = D_j^{--} \cup D_{gen}^{--}$ includes strongly negative sentiment lexicons created for each criterion and strongly negative general lexicon.
- $D^{-} = D_j^{-} \cup D_{gen}^{-}$ includes negative sentiment lexicons created for each criterion and negative general sentiment lexicon.
- Let $R_i = \{R_{i1}, R_{i2}, \cdots, R_{ir}, \ldots, R_{iR}\}$ be a set of reviews (document) including reviews (R_{ir}) on ith alternative.
- Let $S_{ir} = \{S_{ir1}, S_{ir2}, \cdots, S_{irs}, \ldots, S_{irS}\}$ be a set of (sub) sentences in rth review (R_{ir}) regarding to ith alternative.
- V =[ae, sw, adv, neg, cr] 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].
- SentimentDegree_V_v is sentiment degree of sentiment vector V_v. TSentimentDegree_C_j is total sentiment degrees acquired by summing of sentiment degrees of sentiment vectors regarding criterion C_j found. SentimentScore_C_j is sentiment score of review according to criterion C_j, $N_V_C_j$ be the number of extracted sentiment vectors related to criterion C_j in review R_{ir}.

3.4 Methodology

The proposed methodology is explained step by step in this section. Figure 3.1 gives the solution procedure for the problem of product ranking through online reviews. The method mainly consists of three phases, namely online review analysis, BN analysis, and MCDM analysis. Online review analysis stands for preparation for BN analysis. It consists of four steps, which are data crawling, preprocessing online reviews, extracting aspect expressions and compiling sentiment lexicons, and computing sentiment scores of criteria in textual reviews, respectively. Then, BN's structures and parameters are learned in the BN analysis phase. The output of BN analysis is used for product ranking obtained through MCDM analysis.

3.4.1 Online Review Analysis

3.4.1.1 Data Crawling

A web crawler is used for collecting online data utilized for big data analysis. We developed a Python-based web crawler to collect reviews of related products from the website. Some websites allow customers to post both textual reviews and numerical ratings. Both textual reviews and numerical ratings are retrieved from the web pages for this study.

3.4.1.2 Preprocessing Online Reviews

The textual reviews are required for the data preprocessing process before sentiment analysis. In this thesis, the textual reviews in Turkish are preprocessed through ITU Turkish Natural Language Processing Web Service (Eryiğit, 2014; Eryiğit et al., 2008). This NLP service offers many tools, such as normalization, word analysis, morphological analysis, entity name recognition, and sentence analysis. The sentences are divided into their tokens, which are the smallest parts of the text such as words and punctuation marks. The words are separated into their

affixes and roots. Part of speech (POS) tags (adjective, adverb, noun, verb, etc.) are assigned to tokens.

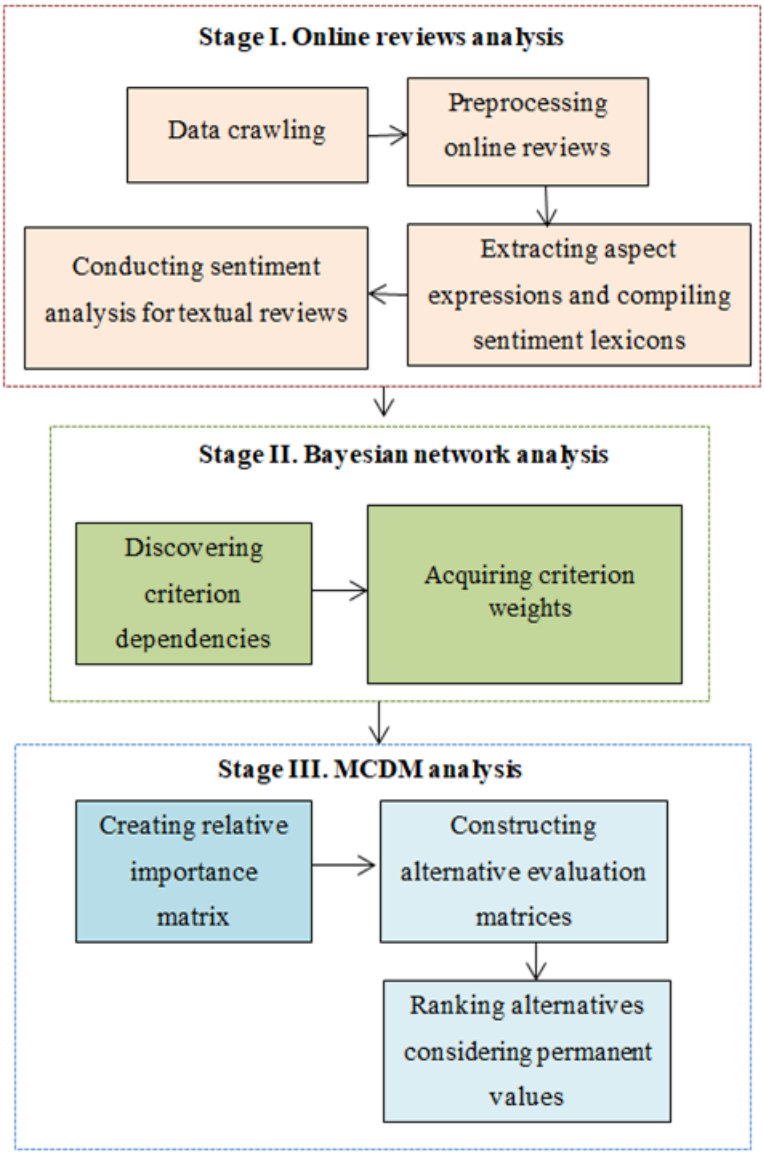


Figure 3.1 Solution procedure for the problem of product/service ranking through online reviews

3.4.1.3 Extracting Aspect Expressions and Compiling Sentiment Lexicons

Customers evaluate alternatives by expressing opinions on their aspects. The criteria consist of related aspects. In a general sense, aspect expressions are nouns or noun phrases expressing a criterion. For example, the aspect expression “hizmet

kalitesi" (quality of service) indicates the aspect "hizmet" (service), which is a member of the criterion *Service*. The aspect expression telesiyej (chair lift) indicates the criterion *Skiing*. Let AE denote the set covering the subsets of AE_j $(j = 1,2,\dots,n)$, and AE_j consists of aspect expressions concerning criterion C_j in all the reviews. The frequencies of tokens considering POS tags are obtained and the filtered tokens are assigned to related subsets of aspect expressions manually.

Sentiment lexicons are lists of words and phrases used in classifying whether a given text expresses a positive or negative sentiment. This study conducts dictionary-based sentiment analysis, in which the process of sentiment word set construction is based on a dictionary. The seed words are obtained by considering the frequencies of tokens with POS tags of adjectives, adverbs, and verbs. Then, sentiment lexicons are extended with the help of the Turkish sentiment dictionary SentiTurkNet (Dehkharghani et al., 2016). Sentiment lexicons comprise general sentiment lexicons and specific sentiment lexicons as in Çalı & Balaman (2019). For instance, customers might use "yardımsever" (helpful) to describe staff in hotels. This sentiment word is added to a sentiment lexicon concerning the criterion *Service*. Sentiment words such as "berbat" (terrible) and "muhteşem" (excellent) can be used for all criteria; hence, they are added to general sentiment lexicons. However, we compile five types of sentiment lexicons by considering the sentiment degrees of sentiment words and phrases differently from Çalı & Balaman (2019). These lexicons are strongly positive, positive, hesitant, negative, and strongly negative sentiment lexicons. Sentiment words are added to lexicons by considering their sentiment degrees. For example, the word "perfect" and the word "good" do not express the same sentiment intensity when describing a criterion. Moreover, adverb lists are compiled, which include adverbs that affect sentiment words to different degrees. For example, the adverbs "çok" (very), epey (quite), oldukça (pretty much), aşırı (too much), gayet (extremely), asla (never), hiç (at all) strengthen the sentiment degrees of sentiment words. Negation expressions ("değil, olmayan, olmalı"/be not, not to be, must be, etc.) transforming sentiment polarity are put into a list as well.

3.4.1.4 Computing Sentiment Scores of Criteria in Textual Reviews

In this part of our work, the aspect-based sentiment analysis algorithm is explained, which is used to compute sentiment scores of criteria in textual reviews. The main function of the aspect-based sentiment analysis is proposed in Figure 3.2.

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

Figure 3.2 Function main of aspect-based sentiment analysis

The algorithm searches for sentiment vectors denoted by V =[ae, sw, adv, neg, cr] in the reviews. The members of a vector are aspect expression, sentiment word,

adverb, negation, and labels. The aspect expression and corresponding sentiment word are searched in the review and saved in the sentiment vector. Labels are used to keep track of which criterion is described in the sentiment vector. Since a sentiment vector might include more than one label, "cr" might include more than one label. For instance, "geniş" (large/big) might be used to describe either a room or a ski-run. The sentiment vector with the sentiment word "geniş" includes the labels of "skiingPos" and "roomPos". Then, if any adverb or negation that has an influence on the sentiment score of the vector in a positive or negative direction is detected, it is kept in the sentiment vector.

The members of the sentiment vector are detected by considering the positions of tokens and some semantic rules by referencing Çalı & Balaman (2019).

Sentence Splitting Rules: If punctuation such as a full stop, exclamation point, question mark, or triple dot is detected, then the review is split into sentences. If the extracted token has the suffix of verb and verbal in tense forms (-di/-dı etc.) or in modal forms (-meli/-malı etc.) in a sentence, then it is spit into sub sentences from the token. Once sentence connectors (such as although, despite, except, and without) are detected, the sentence is split into two parts. Looking for aspect expressions and sentiment words in the next fourteen tokens, after the latest sentence vector is discovered (Çalı & Balaman, 2019).

Distance Rules: (a) The subsequent five tokens must be 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; (d) 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 (Çalı & Balaman, 2019).

Distance Rule (c): 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 in Figure 3.3 are followed.

Initialize : SentimentDegree_V_v ← 0, TSentimentDegree_C_j ← 0, SentimentScores_C_j ← 0
For each sentiment vectors $V_v = [\text{ae}_v, \text{sw}_v, \text{adv}_v, \text{neg}_v, cr_v]$ in review R_{ir} **do**:
  For each Criterion C_j in the set C **do**:
    For each V_v with label of criterion C_j **do**:
      If (sw_v ∈ the set D^{++}) **or** (adv_v ∈ the set Adv^+ **and** sw_v ∈ the set D^+) **then**:
        Assign SentimentDegree_V_v ← **score** (SPos)
        If not.null(neg_v) **then**:
          Decrease SentimentDegree_V_v ← **score** (Pos)
      Else if (sw_v ∈ the set D^+) **or** (sw_v ∈ the set D^- **and** not.null(neg_v)) **then**:
        Assign SentimentDegree_V_v ← **score** (Pos)
      Else if (sw_v ∈ the set D_{hes}) **then**:
        Assign SentimentDegree_V_v ← **score** (Neu)
      Else if (sw_v ∈ the set D^-) **or** (sw_v ∈ the set D^+ **and** not.null(neg_v)) **then**:
        Assign SentimentDegree_V_v ← **score** (Neg)
        If adv_v ∈ the set Adv^- **then**:
          Decrease SentimentDegree_V_v ← **score** (SNeg)
      Else if (sw_v ∈ the set D^{--}) **or** (adv_v ∈ the set Adv **and** sw_v ∈ the set D^-) **then**:
        Assign SentimentDegree_V_v ← **score** (SNeg)
        If not.null(neg_v) **then**:
          Increase SentimentDegree_V_v ← **score** (Neg)
      TSentimentDegree_C_j ← TSentimentDegree_C_j + SentimentDegree_V_v
      $N_V_C_j \leftarrow N_V_C_j + 1$
    SentimentScore_C_j ← TSentimentDegree_C_j / $N_V_C_j$
Output: [SentimentScore _C_1, SentimentScore _C_2, …, SentimentScore_C_j]

Figure 3.3 Function of SentimentScores (sentiment vectors $V_{R_{ir}}$)

Each sentiment vector regarding 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 pseudo-code in Figure 3.3.

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) with different sentiment degrees characterized by a score function mapping to [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 such as oda çok güzel (the room is very good), oda mükemmel (the room is perfect). (2) Positive opinions are such as oda güzel (the room is good), oda çok güzel değil (the room is not very good), (3) Neutral opinions are such as oda idare eder (the room is so so). (4) Negative opinions are such as oda çok berbat değil (the room is not very awful), oda kötü (the room is bad). (5) Strongly negative opinions are such as oda berbat (the room is awful), odayı hiç beğenmedim (I didn't like the room at all), oda hiç iyi değil (the room is not good at all).

3.4.1.5 Determining Evaluation Values in the Decision Matrix

We have used both ratings and reviews in obtaining evaluation values. The evaluation value of the alternative H_i with concerning the criterion C_j is represented by x_{ij} and computed by the following rules:

Rule 1: If there is only a numerical rating for the alternative hotel H_i's performance on the criterion C_j in the 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_i's performance on the criterion C_j in the 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_i's performance on the criterion C_j in the review r, the average of the corresponding scores is assigned to x_{ij}^r.

Finally, x_{ij} is assigned to the average of x_{ij}^r.

3.4.2 BN Analysis

BN is a probabilistic model composed of a network structure G and a joint probability distribution P. Network structure is a directed acyclic graph (DAG) $G = (N, A)$ in which the nodes N denote a set of variables $X = \{X_1, X_2, \dots, X_n\}$ and arcs A point out probabilistic dependencies between these variables. The graph includes two types of nodes, namely *parent nodes* and *child nodes*. The arrows are directed from parent nodes to child nodes. The nodes with no incoming arrows are *root nodes*. The nodes with no outgoing arrows are *leaf nodes*. The probabilistic dependency between a parent and a child is obtained according to Bayes' theorem. The *descendants* of a node are all the nodes that can be reached from the node. The *non-descendants* of a node are all the nodes to which there is no path from the node.

BN satisfies the *Markov condition,* which means each variable is conditionally independent of its *non-descendants* given its parents. The joint probability distribution P is factorized into local probability distributions according to

conditional independence assumptions implied by the Markov condition and chain rules with the following Equation (3.1).

$$P(X) = P(X_1, X_2, \dots, X_n) = \prod_i P(X_i|\Omega_{X_i}) \tag{3.1}$$

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

Here, we gave an example BN whose DAG is shown in Figure 3.4 to explain the italicized terms that were previously mentioned.

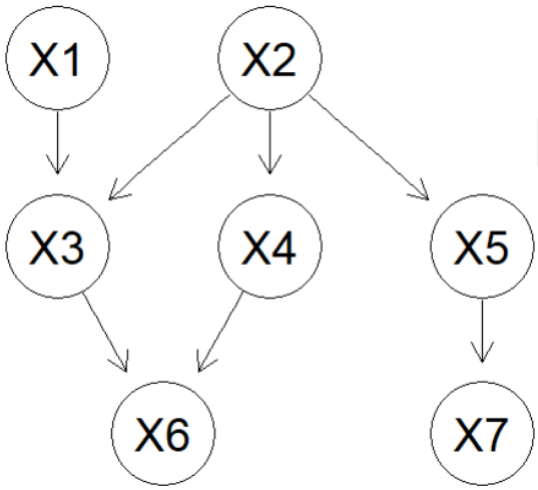


Figure 3.4 An example BN

The BN consists of seven nodes. X1 and X2 are only defined as parent nodes. They are root nodes as well. X3 is child of X1, but it is parent of X6. The parents of node X3 are X1 and X2. X3 is conditionally independent with X5 and X7 given its parents considering the Markov condition. This probabilistic independency is denoted as X3 $\perp\!\!\!\perp_P$ {X5, X7} | {X1, X2}. The joint distribution of BN is obtained as below:

$$P(X1, X2, X3, X4, X5, X6, X7) = P(X1)P(X2)P(X3|X1, X2)P(X4|X2)P(X5|X2)P(X6|X3, X4)P(X7|X5) \quad (3.2)$$

There are two BN approaches in terms of the construction of a network. Expert-driven BNs require domain knowledge and the experience of experts in forming network structures and in eliciting probabilities. Data-driven BNs explore conditional probabilities and/or structure of BN from data by 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 the local distributions are learned in accordance with the network built in the first phase. In order to learn BN structure, two different types of algorithms are used: constraint-based algorithms and score-based algorithms. There are also hybrid approaches, which are a mixture of the two types.

Score-based learning algorithms refer to the use of various optimization algorithms for the problem of learning the structure of the BN (Scutari & Denis, 2015). Score-based learning algorithms search a network structure with the highest goodness-of-fit score. There are various algorithms such as Hill Climbing and Tabu Search. When these algorithms are applied, the search space is scanned until the improvement in the calculated score value ends, usually starting from an empty network structure without arcs between the variables, adding, removing, or changing the direction of arcs at each step. 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 searching. Although there is no clear distinction as to which method should be used, Daly et al., (2011) 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 by using the blacklist and whitelist arguments (Scutari, 2010). A whitelist guarantees a

set of arcs being in the network, whereas a blacklist prevents a set of arcs from being in the network structure.

BN is created by mathematical methods in which the network structure and the parameters are estimated from the data in this study. While determining the most suitable network structure for the data, some metrics are considered, one of the most used is Akaiki Information Criterion (AIC) (Pietro et al., 2017). The network structure that maximizes this metric value is determined as the most suitable model for the data. At first, a score-based algorithm is used to determine the network structure. After the BN structure, the parameters are estimated with either maximum likelihood estimation or Bayesian estimation (Scutari & Denis, 2015, p.111). Since this study uses discrete values in the training process of the BN, the estimated parameters are from the multinomial distribution specified by conditional probability table values, which also reflect the strengths of the relationships between the variables in the network.

3.4.2.1 Bayesian Reasoning

BNs allow decision-makers to conduct probabilistic reasoning and thereby obtain posterior distributions. The types of reasoning are predictive, diagnostic, inter causal (explaining away), and combined reasoning. We have explained these types of reasoning with the example of BN for a lung cancer problem borrowed from the literature (Korb & Nicholson, 2010, p.31). Figure 3.5 illustrates all of them with BNs, in which the nodes indicate the following variables: X1: Pollution, X2: Smoker, X3: Cancer, X4: XRay, X5: Dyspnoea.

Predictive reasoning is from parent nodes to child nodes. In other words, it queries the effects given the evidence about causes. Besides, the effects of parent nodes on child nodes can be analyzed by setting evidence on different levels of parent nodes. The beliefs (posterior probability) regarding child nodes are updated with evidence regarding the state of a parent node being instantiated. For instance, the belief of the patient developing cancer and the beliefs of the symptoms such as observing positive

X-ray results and Dyspnoea will increase given the information that the patient is a smoker. 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. Intercausal reasoning deals with the mutual causes of a common result. Given the information that the patient is cancer, the posterior probabilities of the causes (smoke and pollution) will increase, however, the posterior probability of the cause of exposure to pollution will decrease together with the evidence that the patient is also a smoker. That is to say, the cause of pollution has been explained away. In addition, combined reasoning can be realized such as the combination of predictive and diagnostic reasoning demonstrated by Figure 3.5(d).

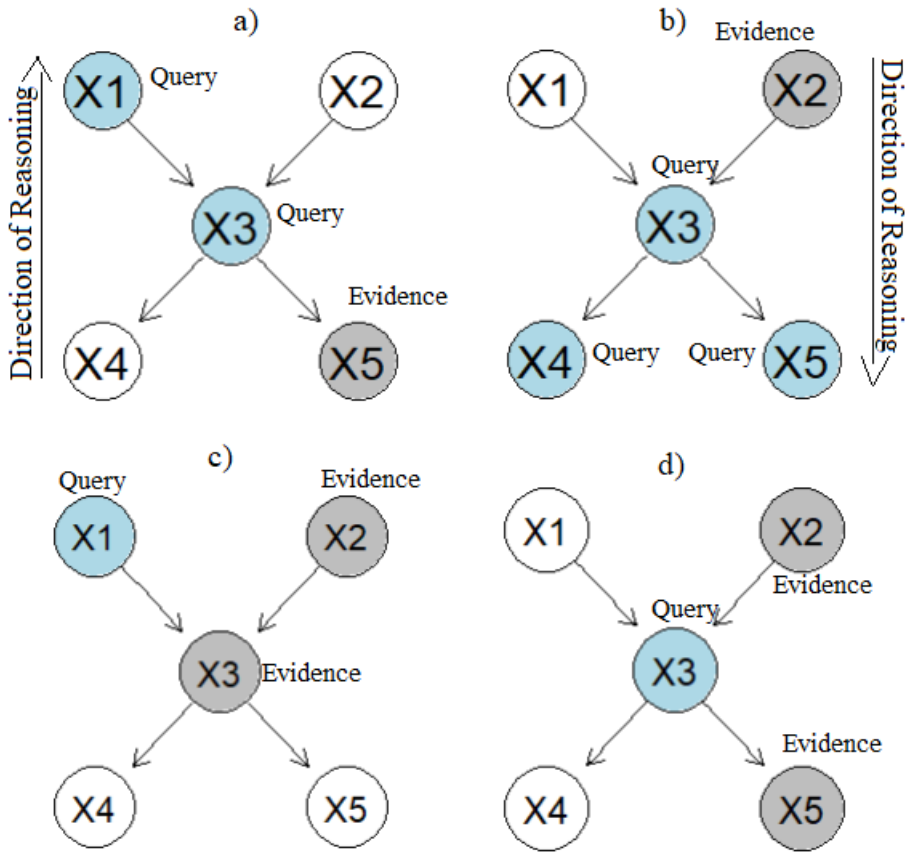


Figure 3.5 Types of reasoning: a) diagnostic, b) predictive, c) intercausal, and d) combined (Korb & Nicholson, 2010)

The belief updating is performed with either exact or approximate algorithms (Koller & Friedman, 2009).

3.4.2.2 Determining the Relative Importance of Criteria via Mutual Information

Mutual information (Shannon & Weaver, 1949) is one of the most popular metrics utilized for ranking information sources. It is based on the idea that the uncertainty related to any random variable can be described by the entropy function (Pearl, 1988) in Equation (3.3).

$$H(Y) = -\sum_{y} P(y) \log P(y) \tag{3.3}$$

The uncertainty related to Y, given that random variable Z is instantiated to z, is obtained via the following Equation (3.4)

$$H(Y|z) = -\sum_{y} P(y|z)\, log\, P(y|z), \tag{3.4}$$

and Equation (3.5) computes the average residual uncertainty in Y given Z.

$$H(Y|Z) = \sum_{z} H(Y|z)P(z) \tag{3.5}$$

Then, mutual information expressing the total uncertainty-reducing potential of Z is calculated by Equation (3.6).

$$I(Y,Z) = H(Y|Z) - H(Y) \tag{3.6}$$

Mutual information is symmetric ($I(Y,Z) = I(Z,Y)$) and non-negative quantity. It shows the dependency between two random variables. If mutual information is obtained as zero, it means that the random variables are mutually independent.

Let consider $C = \{C_1, C_2, \dots, C_n\}$ be the set of random variables which describe the criteria and the random variable Y represents the recommendation choice. We have mentioned the decision on whether or not a customer recommends a hotel as "recommendation choice" in this thesis. The impact degree of each criterion on recommendation choice is obtained by mutual information. These measures can be reflected in decision-making methodology as criteria weights with Equation (3.7) or relative importance values with Equation (3.8).

$$w_j = \frac{I(C_j, Y)}{\sum_{j=1}^{n} I(C_j, Y)}, \tag{3.7}$$

where w_j indicates weight of criterion C_j and $\sum_{j=1}^{n} w_j = 1$.

The relative importance of criterion C_i over criterion C_j is obtained by drawing inspiration from (Koulouriotis & Ketipi, 2011):

$$r_{ij} = \frac{I(C_i, Y)}{I(C_i, Y) + I(C_j, Y)}. \tag{3.8}$$

3.4.3 MCDM Analysis

3.4.3.1 GTMA

GTMA is a technique used in MCDM analysis and it is based on graph theory and matrix algebra. The computation of permanent introduced by Binet and Cauchy in 1812 (Minc, 1978) is a key component of the GTMA. Permanent is utilized in combinatorial mathematics, which describes the number of perfect matching in a bipartite graph. Let $A = (a_{ij})$ be an m x n matrix, where $m \leq n$. The permanent of a matrix A is obtained by using Equation (3.9).

$$Per(A) = \sum_{\sigma} a_{1\sigma(1)} a_{2\sigma(2)}, \dots, a_{m\sigma(m)} \tag{3.9}$$

where, the summation extends over all one-to-one functions from $\{1, \dots, m\}$ to $\{1, \dots, m\}$ (Minc, 1978). The sequence $a_{1\sigma(1)}a_{2\sigma(2)}, \dots, a_{m\sigma(m)}$ indicates the diagonal of matrix A and the product $a_{1\sigma(1)}a_{2\sigma(2)}, \dots, a_{m\sigma(m)}$is a diagonal product of A. Thereby, the permanent of A is the sum of all diagonal products of A (Baykasoglu, 2014).

The computation of permanent is more challenging than the determinant since permanent cannot be computed by using Gaussian elimination, unlike the determinant (Baykasoglu, 2014). Valiant (1979) demonstrated that computing the permanent of a matrix is #P-hard and thus unlikely to have a polynomial time solution. Ryser (1963) had developed an $O(n^2 2^n)$ time algorithm to compute the permanent of an n × n matrix, which is given in Equation (3.10).

$$Per(A) = (-1)^n \sum_{S \subseteq \{1,\dots,n\}} (-1)^{|S|} \prod_{i=1}^{n} \sum_{j \in S} a_{ij} \tag{3.10}$$

where, the sum is over all subsets of {1, . . ., n} and |S| is the number of elements in S.

GTMA has several advantages such as modeling hierarchical problems and considering criteria interaction. This approach has a simple computational procedure as well. We have used BN analysis to obtain components of GTMA. The methodology of GTMA integrated with BN results consists of three steps, which are explained as follows.

Step 1: The components of MCDM which are criteria, sub-criteria, and alternatives are specified.

Step 2: The relative importance (or interrelation) between the criteria are determined. The performance scores of alternatives concerning criteria are obtained based on BN analysis outputs. The normalization process is performed. A criteria rating matrix (Ψ_i) for each alternative is constructed which is an nxn square diagonal

matrix shown in Equation (3.11). Each x_{ij} is computed in the online review analysis discussed in the previous section.

$$\Psi_i = \begin{bmatrix} x_{i1} & 0 & \cdots & 0 \\ 0 & x_{i2} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & x_{in} \end{bmatrix} \tag{3.11}$$

The relative importance (or interrelation) matrix β is formed in which the values of r_{ij} represent the relative importance of criterion i relatively to the cumulative importance of criteria i and j (Koulouriotis & Ketipi, 2011). The values of r_{ji} is obtained via Equation (3.8) considering mutual information in this study. The relative importance (interrelation) matrix β is formed by Equation (3.12).

$$\beta = \begin{bmatrix} 0 & r_{12} & \cdots & r_{1n} \\ r_{21} & 0 & \cdots & r_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ r_{n1} & r_{n2} & \cdots & 0 \end{bmatrix} \tag{3.12}$$

Step 3: The alternative evaluation matrix ξ_{H_i} for each alternative H_i is constructed by summing the criteria rating matrix for the alternative and relative importance (interrelation) matrix, which is given by Equation (3.13).

$$\xi_{H_i} = \Psi_i + \beta = \begin{bmatrix} x_1 & r_{12} & \cdots & r_{1n} \\ r_{21} & x_2 & \cdots & r_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ r_{n1} & r_{n2} & \cdots & x_n \end{bmatrix} \tag{3.13}$$

The rating of each alternative is obtained by calculating the permanent of matrix ξ_{H_i} by using Equation (3.10). The permanent values of alternatives are ranked in descending order. The alternative with the highest permanent value is the best.

3.5 Case Study

The proposed methodology is applied to a case study that analyzes hotel reviews regarding five-star ski hotels in Turkey.

3.5.1 Analyzing Online Reviews

We have crawled reviews on 7 alternative ski hotels from otelpuan.com (https://www.otelpuan.com/tr) in September 2022. It is a popular website in Turkey, including hotel reviews. The alternative hotels are Ağaoglu My Mountain Uludağ (H_1), Bof Hotels Uludağ Ski & Luxury Resort (H_2), Grand Kartal Hotel (H_3), Grand Yazıcı Uludağ (H_4), Karinna Hotel Uludağ (H_5), Polat Erzurum Resort Hotel (H_6), and Sway Hotels Palandöken (H_7). Customers evaluate hotels with reference to criteria, namely, *Price-Performance* (C_1), *Food* (C_2), *Room* (C_3), *Skiing* (C_4), and *Service* (C_5). They also propose their decision on whether or not they recommend the hotel by posting *recommended*, *not recommended*, or *hesitant*. The dataset includes textual reviews, ratings, and recommendation choices regarding alternative hotels. An example review is shown in Figure 3.6.



Figure 3.6 A review crawled from otelpuan.com in September 2022

The criteria variables are framed with red rectangles. The target variable (*Recommendation*) is framed with a green rectangle. Customers assign continuous

numerical ratings in the interval [1, 10] to *Food*, *Room*, *Skiing*, and *Service*. *Price-Performance* is evaluated with linguistic variables from the term set (Very Bad, Poor, Medium, Good, and Excellent).

Customers assign scores to *Price-Performance* for expressing their satisfaction with the costs of all hotel facilities while also considering how well the performance is provided by the hotel. They evaluate employees' attitudes and service quality in hotels by rating to *Service*. They assess the ski facilities of the hotel by rating *Skiing*. The scores given to *Food* represent how customers are satisfied with the quality, taste, presentation, etc. Customers assign rates to *Room* for expressing their pleasure in terms of several aspects such as cleanness, design, landscape, etc.

Customers have preferred to write textual reviews regarding criteria instead of posting numerical ratings in some evaluations. In this manner, aspect-based sentiment analysis algorithm (explained in Section 3.4.1) is used to analyze textual reviews and transform them to numerical scores like ratings to prevent loss of information. If a customer review does not include a numerical rating on any criterion, the score obtained by sentiment analysis is considered for the corresponding criterion. If a customer review includes both a numerical rating and sentiment score for any criterion, the average of these scores is considered for the final score of the corresponding criterion. It is observed in the dataset that customers have mostly assigned 2, 4, 6, 8, and 10 points to criteria. Therefore, we have used these scores for linguistic variables of *Price-Performance* and labels of sentiment analysis. Normalized values are used for labels of sentiment analysis, and the numerical ratings are transformed to normalized form as well. In other words, the function of *Score* assigns 0.2, 0.4, 0.6, 0.8, and 1 to Strongly Negative (SNeg), Negative (Neg), Neutral (Neu), Positive (Pos), and Strongly Positive (SPos) sentiment vectors in Algorithm 2. These scores are matched with linguistic variables of *Price-Performance*, as well.

3.5.2 BN Analysis

After the online review analysis, the next stage is the BN analysis. The frequency statistics of the dataset are illustrated in Figure 3.7.

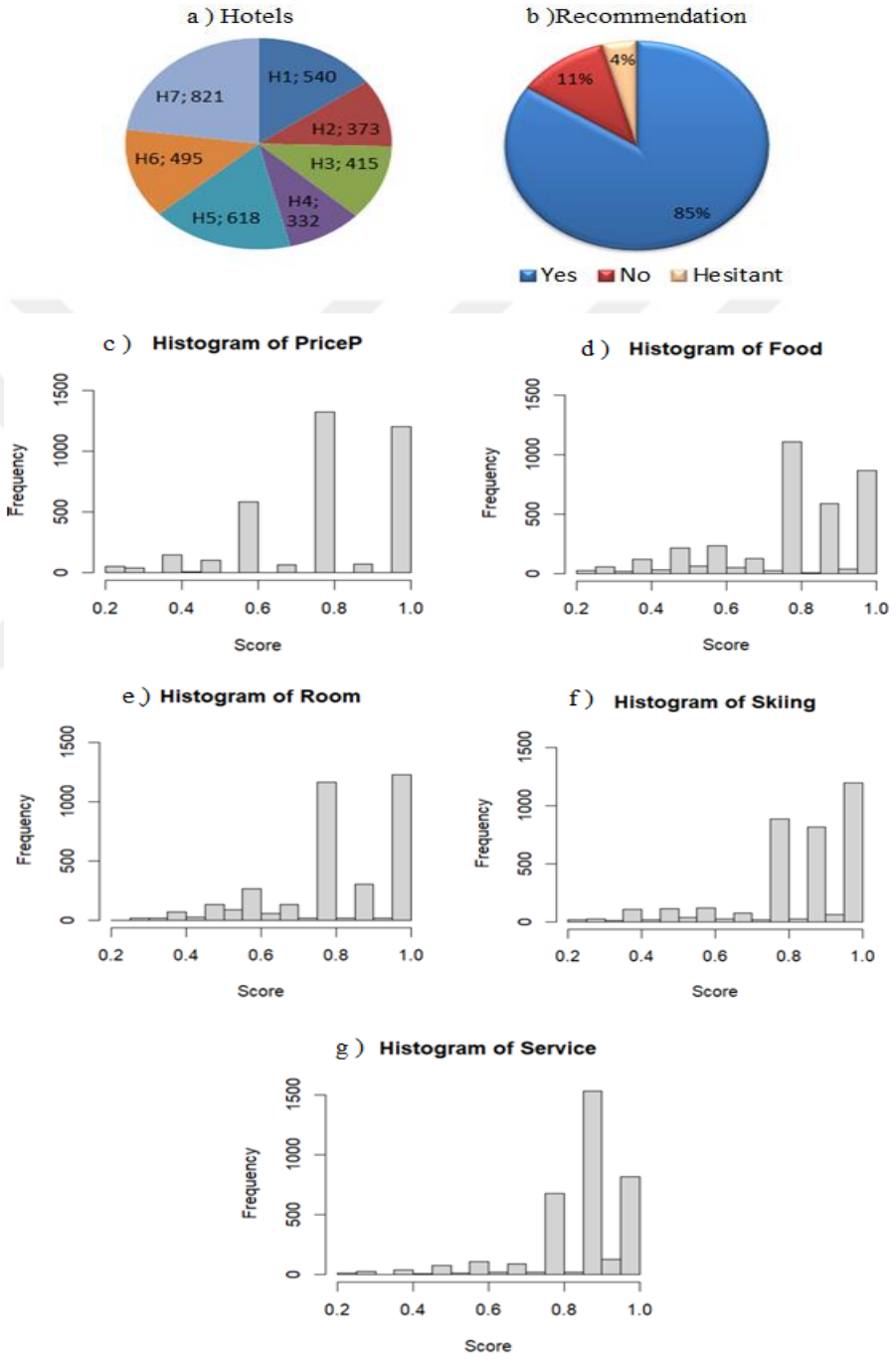


Figure 3.7 Frequency statistics from the dataset

The pie chart at the top left of the figure shows the distribution of reviews of alternative hotels. The number of total reviews is 3594, with no missing values. The pie chart at the top right of the figure depicts the distribution of the recommendation choices of the 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 discretized by dividing them into groups of equal length for BN analysis. Hence, criteria variables are categorized 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. Target variable *Recommendation* is categorized 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 3.1. This tableau also gives the intervals representing the states.

Table 3.1 Variables in the BN

Type	Name (Code)	State Name (Code):	Interval
Criteria variables	Price-Performance (C_1), Food (C_2), Room (C_3), Skiing (C_4), Service (C_5)	Dissatisfied (DisSat):	[0, 0.40]
		Neutral (Neut):	[0.41, 0.60]
		Satisfied (Sat):	[0.61, 0.80]
		Very Satisfied (VerySat):	[0.81, 1]
Target variable	Recommendation (T)	Recommended (Yes)	-
		Not recommended (No)	

BN is learned from the dataset, which is modeled using the bnlearn package in R (Scutari, 2010). Before learning the structures of BNs, we integrated some logical constraints into BN modeling by creating blacklists that prevent the presence of arcs directed from the target variable *Recommendation* to any other variable in the network structures and the presence of arcs directed from any criterion variable to *Price-Performance*. BN is constructed to determine the dependence structure between the variables; hence it extracts 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, the customers might assign the satisfaction score for *Price-Performance* according to the performances of the other criteria presented in the hotel. They make their decision on the recommendation by considering the perceived performances of the criteria.

The structure of BN is learned by using the Hill-Climbing algorithm with the AIC score function (Scutari, 2010), which is defined with

$$AIC = \log P(C_1, C_2, C_3, C_4, C_5, T) - d \tag{3.14}$$

where d is the number of parameters and the local distributions are estimated with Bayesian estimation. The BN learned from the dataset and the marginal probabilities of the variables are given in Figure 3.8.

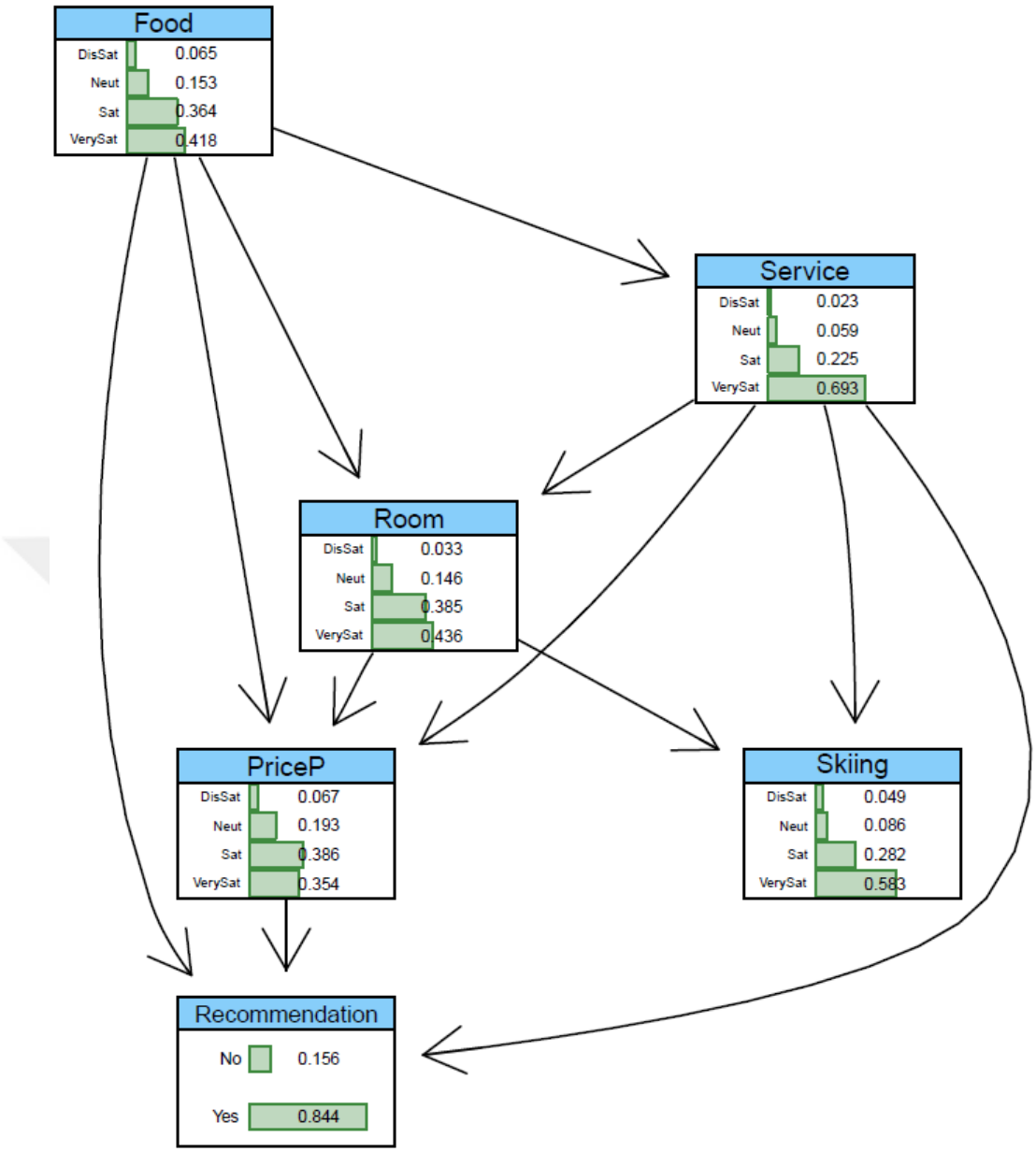


Figure 3.8 The BN estimated from the dataset

As stated before, 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. *Price-Performance* is abbreviated as PriceP in the network.

BN graph enables to deduce of direct and indirect relationships between the variables. BN shows that criteria variables *Price-Performance*, *Service*, and *Food*

have a direct influence on *Recommendation*. *Skiing* has no direct relationship with *Price-Performance* and *Food*. However, they have indirect relationships via the paths. The highest probability of variables except *Price-Performance* is observed at the state *VerySat*. Otherwise, we see that *Price-Performance* has its highest probability for state *Sat* with 38.6 %.

3.5.2.1 Accuracy of the Model

The accuracy of the model is verified, as discussed in (Pietro et al., 2015). A random sample of 3600 observations is simulated from the created BN model. The receiver operating characteristic (ROC) curve (Metz, 1978) is drawn based on these data to demonstrate the goodness of fit. The ROC curve plots the true positive rate (sensitivity) versus the true negative rate (1-spesificity), and the area under the curve is the AUC metric. The AUC value for the target variable *Recommendation* is 0.92, which points out a good accuracy of the model (Swets, 1988).

3.5.2.2 Sensitivity Analysis and What-If Analysis

The BN graph shows the variables directly dependent with *Recommendation* via the directed arcs. However, a measurement is required to incorporate these relations to decision-making procedure. Therefore, the impact degree of each criterion on recommendation choice is calculated based on mutual information measurement in the sensitivity analysis. Mutual information values between the criteria variables and *Recommendation* are computed, respectively, as 0.204, 0.125, 0.06, 0.01, and 0.134 bits by Equations (3.3)-(3.6). According to the analysis, the criteria can be ordered from most important to least important as *Price-Performance*, *Service*, *Food*, *Room*, and *Skiing*, respectively, in determining the recommendation choices of the customers.

BN is a very useful machine learning tool since it has an inference mechanism as well. Hotel managers can use this aspect of BN to observe the impacts of the states of the variables on the recommendation choice. We have performed Bayesian inference

(reasoning) with exact inference algorithm, which is belief updates based on junction trees (Scutari & Denis, 2015, p.115). Figure 3.9 shows the what-if analysis results.



Figure 3.9 Posterior probabilities of Recommendation, given evidence that the others are equal to (a) VerySat and (b) DisSat

The first bar in either graph depicts the probability of *Recommendation* without evidence (baseline). The left part (a) of the figure gives each posterior probability of *Recommendation* given evidence that each variable is at the state of *VerySat*, whereas the right part (a) of the figure gives 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 rises from 0.844 to 0.983, given the evidence that customers are very satisfied with *Price-Performance*. On the other hand, the probability of positive recommendation choice (*Recommendation* = *Yes*) for hotels declines from 0.844 to 0.200, given the evidence that customers are dissatisfied with *Price-Performance*. The figure also depicts the drastic changes in probability of *Recommendation* setting evidence for *Food* and *Service*. We have highlighted the importance of these variables by pointing out mutual information in sensitivity analysis as well. Although we have analyzed ski hotels, the changes in the probability distribution of *Recommendation* are not remarkable given the evidence for *Skiing* at *VerySat* or *DisSat* when compared with the other variables.

3.5.3 MCDM Analysis through BN-GTMA Model

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 the node *i* to the node *j* show the relative importance of criterion *i* over criterion *j*. The digraph of the hotel ranking problem death in this study is illustrated in Figure 3.10.

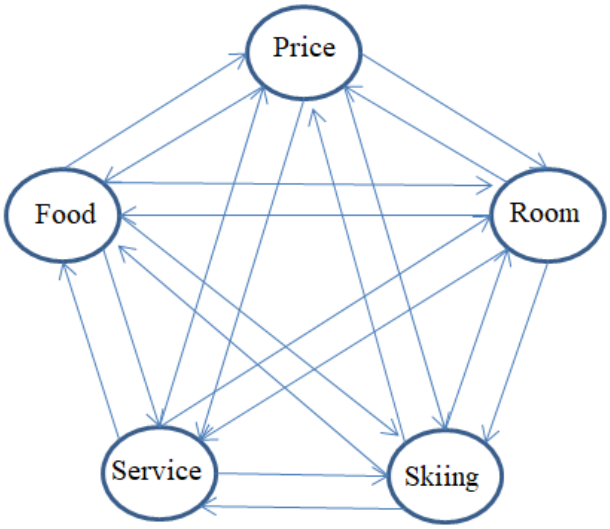


Figure 3.10 Digraph representation of the hotel ranking problem.

The methodology of GTMA integrated with BN is realized step by step. Firstly, the relative importance values are calculated with Equation (3.8), and the relative importance matrix is proposed in Table 3.2.

Table 3.2 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 BN stage, we have mentioned that mutual information between the criteria variables and *Recommendation* are computed as 0.204, 0.125, 0.06, 0.01, and 0.134 bits respectively. That is, total uncertainty reducing potential of *Price-Performance* is 0.204 bit while total uncertainty reducing potential of *Food* is 0.125 bit. *Price-Performance* has 62 % of the sum of these two values (0.204+0.125). The relative importance of *Price-Performance* is higher than *Food* in terms of hotel recommendation. Secondly, the performance scores of alternatives regarding criteria are calculated based on online review analysis outputs. Performance values of hotels are found in Table 3.3. These values are normalized by dividing max values.

Table 3.3 Evaluation values of the alternative hotels

Hotel	Price-Performance	Food	Room	Skiing	Service
H1	0.7264	0.6996	0.8007	0.841	0.8371
H2	0.7931	0.8216	0.8362	0.8549	0.8524
H3	0.7957	0.8542	0.7855	0.8472	0.8798
H4	0.7562	0.7685	0.8036	0.8112	0.843
H5	0.7918	0.7706	0.8123	0.8148	0.8632
H6	0.7975	0.799	0.7636	0.8538	0.8586
H7	0.8555	0.8226	0.8866	0.8669	0.8882

Lastly, permanent matrices of hotels are formed by considering the relative importance of criteria and normalized evaluation values given in Table 3.4.

Table 3.4 Permanent matrices of hotels

	H1				
	C1	C2	C3	C4	C5
C1	0.849	0.62	0.767	0.944	0.604
C2	0.38	0.819	0.668	0.912	0.483
C3	0.233	0.332	0.903	0.838	0.316
C4	0.056	0.088	0.162	0.97	0.082
C5	0.396	0.517	0.684	0.918	0.942

	H2				
	C1	C2	C3	C4	C5
C1	0.927	0.62	0.767	0.944	0.604
C2	0.38	0.962	0.668	0.912	0.483
C3	0.233	0.332	0.943	0.838	0.316
C4	0.056	0.088	0.162	0.986	0.082
C5	0.396	0.517	0.684	0.918	0.96

	H3				
	C1	C2	C3	C4	C5
C1	0.93	0.62	0.767	0.944	0.604
C2	0.38	1	0.668	0.912	0.483
C3	0.233	0.332	0.886	0.838	0.316
C4	0.056	0.088	0.162	0.977	0.082
C5	0.396	0.517	0.684	0.918	0.991

	H4				
	C1	C2	C3	C4	C5
C1	0.884	0.62	0.767	0.944	0.604
C2	0.38	0.9	0.668	0.912	0.483
C3	0.233	0.332	0.906	0.838	0.316
C4	0.056	0.088	0.162	0.936	0.082
C5	0.396	0.517	0.684	0.918	0.949

	H5				
	C1	C2	C3	C4	C5
C1	0.926	0.62	0.767	0.944	0.604
C2	0.38	0.902	0.668	0.912	0.483
C3	0.233	0.332	0.916	0.838	0.316
C4	0.056	0.088	0.162	0.94	0.082
C5	0.396	0.517	0.684	0.918	0.972

	H6				
	C1	C2	C3	C4	C5
C1	0.932	0.62	0.767	0.944	0.604
C2	0.38	0.935	0.668	0.912	0.483
C3	0.233	0.332	0.861	0.838	0.316
C4	0.056	0.088	0.162	0.985	0.082
C5	0.396	0.517	0.684	0.918	0.967

	H7				
	C1	C2	C3	C4	C5
C1	1	0.62	0.767	0.944	0.604
C2	0.38	0.963	0.668	0.912	0.483
C3	0.233	0.332	1	0.838	0.316
C4	0.056	0.088	0.162	1	0.082
C5	0.396	0.517	0.684	0.918	1

Hotel customer satisfaction index values of the alternative hotels are found using matrix permanent formula, and given with the ranks in Table 3.5.

Table 3.5 Customer satisfaction index values of hotels and ranks

Hotel	Hotel Customer Satisfaction Index	Rank
H1	4.294	7
H2	4.944	2
H3	4.942	3
H4	4.453	6
H5	4.630	5
H6	4.724	4
H7	5.407	1

Our methodology proposes the ranking of H7≻H2≻H3≻H6≻H5≻H4≻H1. It can be seen that H7 is the highest ranking score. H2 and H3 have almost the same value which follows the score of H7. This ranking can be presented to potential customers as the advice of the customers posted the reviews, which are acquired through comprehensive analysis, including contributions of scientific tools.

3.6 Comparative Analysis and Discussion

In this section, we have compared our results with the results of two well-known MCDM methods, namely, the technique for order preference by similarity to ideal solution (TOPSIS)(Hwang & Yoon, 1981) and the simple additive weighting (SAW)(Churchman & Ackoff, 1954). The criteria weights are derived by Equation (3.7) while performing SAW and TOPSIS.

In SAW, each criterion score is multiplied by the related criterion weight, and then these products are summed to obtain the total evaluation score of an alternative. However, this method may mislead decision-makers since interactions between criteria are ignored (Darvish et al., 2009). TOPSIS seeks the convenient alternative by maximizing the distance to the positive ideal alternative while minimizing the negative ideal (nadir) alternative (Olson, 2004). At first, the weighted normalized matrix is constructed, and then the positive ideal alternative with the best criteria values and the negative ideal alternative with the worst criteria values are

determined. 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. The alternative with the highest ratio is proposed as the best alternative. This method is not capable of dealing with interrelationships between the criteria, like SAW.

The ranking of hotels based on SAW and TOPSIS is given in Table 3.6. H7 is in the best position in the ranking produced by each of the three methods. However, H2 and H3 compete for the second position in the ranking. The GTMA introduces H2 as the second convenient hotel, whereas SAW and TOPSIS propose H3 as the second convenient hotel. The ranking proposed by GTMA is more rational than the others in this dataset since it takes into account the interrelationships of criteria. It describes the hotel ranking problem better than the other since the matrix permanent concept involves 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, thus making the ranking of the alternatives easier.

Table 3.6 Comparison of the ranking with the results of SAW and TOPSIS

Methods	Criteria weights	Positions in the rankings						
		H1	H2	H3	H4	H5	H6	H7
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
GTMA	Relative importance matrix	7	2	3	6	5	4	1

3.7 Chapter Conclusions

This thesis proposes an online review ranking methodology integrating BN, sentiment analysis, and GTMA. The process starts with online review analysis, in which both online ratings and text reviews are incorporated. The customer's

sentiments on criteria mentioned in the text reviews are extracted by the aspect-based sentiment analysis algorithm. The main purpose of the proposed algorithm is to classify customer reviews according to different sentiment polarity degrees. The process continues with BN analysis discovering the statistical relationships both between the criteria and between each criterion and the recommendation choice. We have computed the relative importance of each criterion based on mutual information used to identify the relevance of the variables in BN. The process ends with ranking the alternatives via matrix permanents. In other words, the components of the GTMA are obtained based on the outputs of the sentiment analysis and the BN analysis.

The developed methodology is applied to a case study aiming to rank ski hotels through online evaluations. We have analyzed seven popular ski hotels in Turkey. The aspect-based sentiment algorithm in this study was developed for analyzing hotel reviews in Turkish; however, it can be adapted for other domains and other languages. This algorithm does not follow a basic process intending bilateral classification; it aims to get a more informative classification based on five categories by detecting the words strengthening the sentiments. This study proposes a new MCDM approach, which involves the integration of BN and GTMA to determine the ranking of the hotels. The superiorities of this method are highlighted through a comparative analysis. This approach can be used for other data-driven decision-making problems besides product ranking problems based on online reviews.

The developed decision support tool has several advantages in terms of implementation. Especially, the BN part of the methodology gives businesses the ability to deduce on which criterion the customers give more value when recommending alternatives to other customers. The managers can develop rational strategies supported by the results of this scientific tool. According to our dataset, it can be deduced that the price-performance scores of the hotels have a great impact on the recommendation decision for ski hotels. The customers of ski hotels consider whether the hotel is a price-performance hotel or not. In this manner, they make decisions about recommending the hotels to other customers. Hotel managers should

follow effective price policies to not only improve customer retention but also attract new customers. Besides having various contributions, the thesis has some limitations as well. Since the main focus of this thesis is BN analysis, the MCDM analysis part of the study is open to improvement with the integration of fuzzy set theory. This analysis is based on one website. In future research, this limitation might be removed with the incorporation of other websites, including hotel reviews, into the analysis.

CHAPTER FOUR
CONCLUSIONS

Massive amounts of data on the internet and the growing trend of information sharing create new issues that call for improvements in MCDM. Thereby, this dissertation proposes two novel data-driven decision-making models with the help of algorithms based on Bayesian theory.

The first is a Bayesian approach for figuring out how product features affect customer satisfaction which also addresses the weight determination problem in MCDM. In the analysis, this method only uses information on product characteristics and overall product scores. The impacts of product attributes on customer satisfaction are considered as weights of mixture kernel distributions and posterior distributions of weights are simulated with the Metropolis-Hastings algorithm in MCMC methods. The applicability of the proposed methodology is demonstrated by a case study, in which online data of the mobile phone market are analyzed.

The significance of the research in terms of the case study is twofold. The results can shed light on which attributes product designers should concentrate on to enhance product sale volumes as well as customer satisfaction. Analyzing online reviews are subject of computer science and engineering design literature as well as the subject of business and marketing literature. From a research standpoint, this method demonstrates how more in-depth knowledge about customer preferences can be captured with comparatively fewer data and computational work.

The second is a novel Bayesian model for hotel ranking through online customer reviews. In the literature, there is no study using data-driven BN to take into account the stage of determining criteria interactions in product/service ranking through online customer reviews. Thus, this dissertation fulfills this research gap. Moreover, this dissertation extracts knowledge from unstructured data covering online ratings and textual reviews. Aspect-based sentiment analysis is proposed to classify online Turkish reviews of ski hotels considering different sentiment degrees. The developed decision-making model proposes two-sided benefits. BN analysis of online customer

reviews enables the ski hotel managers to gain knowledge about to what extent the customers give importance to each criterion while recommending the ski hotels. The MCDM procedure, the integrated BN-GTMA model, enables potential customers to learn the rankings of the hotels before making a reservation decision.

As a result, this thesis adds new Bayesian models to the literature on MCDM using big data analysis and probabilistic machine learning. The analyses of both models are based on data crawled from one website. In future studies, the other websites may be incorporated into the analysis.

REFERENCES

Ahmad, A., Dey, L., & Halawani, S. M. (2012). A rule-based method for identifying the factor structure in customer satisfaction. *Information Sciences, 198*, 118–129. https://doi.org/https://doi.org/10.1016/j.ins.2012.02.056

Alsaffar, A. H. (2017). Empirical study on the effect of using synthetic attributes on classification algorithms. *International Journal of Intelligent Computing and Cybernetics, 10*(2), 111–129. https://doi.org/10.1108/IJICC-08-2016-0029

Andrieu, C., De Freitas, N., Doucet, A., & Jordan, M. I. (2003). An introduction to MCMC for machine learning. *Machine Learning, 50*, 5–43. https://doi.org/https://doi.org/10.1023/A:1020281327116

Baykasoglu, A. (2014). A review and analysis of "graph theoretical-matrix permanent" approach to decision making with example applications. *Artificial Intelligence Review*, *42*(4), 573–605. https://doi.org/10.1007/s10462-012-9354-y

Baykasoğlu, A., & Baykasoğlu, C. (2017). Multiple objective crashworthiness optimization of circular tubes with functionally graded thickness via artificial neural networks and genetic algorithms. *Proceedings of the Institution of Mechanical Engineers, Part C: Journal of Mechanical Engineering Science*, *231*(11), 2005–2016. https://doi.org/10.1177/0954406215627181

Birjali, M., Kasri, M., & Beni-Hssane, A. (2021). A comprehensive survey on sentiment analysis: Approaches, challenges and trends. *Knowledge-Based Systems*, *226*, 107134. https://doi.org/10.1016/j.knosys.2021.107134

Çalı, S., & Balaman, Ş. Y. (2019). Improved decisions for marketing, supply and purchasing: Mining big data through an integration of sentiment analysis and intuitionistic fuzzy multi criteria assessment. *Computers and Industrial Engineering*, *129*, 315–332. https://doi.org/10.1016/j.cie.2019.01.051

Çalı, S., & Baykasoğlu, A. (2022). A Bayesian based approach for analyzing customer's online sales data to identify weights of product attributes. *Expert Systems with Applications*, *210*. https://doi.org/10.1016/j.eswa.2022.118440

Chakraborty, S., Mengersen, K., Fidge, C., Ma, L., & Lassen, D. (2016). A Bayesian Network-based customer satisfaction model: a tool for management decisions in railway transport. *Decision Analytics*, *3*(1). https://doi.org/10.1186/s40165-016-0021-2

Chen, K., Jin, J., Zhao, Z., & Ji, P. (2022). Understanding customer regional differences from online opinions: a hierarchical Bayesian approach. *Electronic Commerce Research*, *22*(2), 377–403. https://doi.org/10.1007/s10660-020-09420-5

Chen, R., & Xu, W. (2017). The determinants of online customer ratings: a combined domain ontology and topic text analytics approach. *Electronic Commerce Research*, *17*(1), 31–50. https://doi.org/10.1007/s10660-016-9243-6

Chevalier, J. A., & Mayzlin, D. (2006). The effect of word of mouth on sales: Online book reviews. *Journal of Marketing Research*, *43*(3), 345–354. https://doi.org/10.1509/jmkr.43.3.345

Chiu, S. I., Cheng, C. C., Yen, T. M., & Hu, H. Y. (2011). Preliminary research on customer satisfaction models in Taiwan: A case study from the automobile industry. *Expert Systems with Applications*, *38*(8), 9780–9787. https://doi.org/10.1016/j.eswa.2011.01.172

Churchman, C. W., & Ackoff, R. L. (1954). An approximate measure of value. *Journal of Operations Research Society of America*, *2*(1), 172–187. https://doi.org/https://doi.org/10.1287/opre.2.2.172

Daly, R., Shen, Q., & Aitken, S. (2011). Learning Bayesian networks: approaches and issues. *Knowledge Engineering Review*, *26*(2), 99–157. https://doi.org/10.1017/S0269888910000251

Darvish, M., Yasaei, M., & Saeedi, A. (2009). Application of the graph theory and matrix methods to contractor ranking. *International Journal of Project Management*, *27*(6), 610–619. https://doi.org/10.1016/j.ijproman.2008.10.004

Dash, G., Kiefer, K., & Paul, J. (2021). Marketing-to-Millennials: Marketing 4.0, customer satisfaction and purchase intention. *Journal of Business Research*, *122*, 608–620. https://doi.org/10.1016/j.jbusres.2020.10.016

De Maeyer, P. (2012). Impact of online consumer reviews on sales and price strategies: A review and directions for future research. *Journal of Product and Brand Management*, *21*(2), 132–139. https://doi.org/10.1108/10610421211215599

Dehkharghani, R., Saygin, Y., Yanikoglu, B., & Oflazer, K. (2016). SentiTurkNet: a Turkish polarity lexicon for sentiment analysis. *Language Resources and Evaluation*, *50*(3), 667–685. https://doi.org/10.1007/s10579-015-9307-6

Deng, H., Yeh, C. H., & Willis, R. J. (2000). Inter-company comparison using modified TOPSIS with objective weights. *Computers and Operations Research*, *27*(10), 963–973. https://doi.org/10.1016/S0305-0548(99)00069-6

Deng, Z., Lu, Y., Wei, K. K., & Zhang, J. (2010). Understanding customer satisfaction and loyalty: An empirical study of mobile instant messages in China. *International Journal of Information Management*, *30*(4), 289–300. https://doi.org/10.1016/j.ijinfomgt.2009.10.001

Denguir-Rekik, A., Montmain, J., & Mauris, G. (2009). A possibilistic-valued multi-criteria decision-making support for marketing activities in e-commerce: Feedback Based Diagnosis System. *European Journal of Operational Research*, *195*(3), 876–888. https://doi.org/10.1016/j.ejor.2007.11.020

Diakoulaki, D., Mavrotas, G., & Papayannakis, L. (1995). Determining objective weights in multiple criteria problems: The critic method. *Computers and Operations Research*, *22*(7), 763–770. https://doi.org/10.1016/0305-0548(94)00059-H

Dougherty, J., Kohavi, R., & Sahami, M. (1995). Supervised and Unsupervised Discretization of Continuous Features. *Machine Learning Proceedings 1995*, *1995*, 194–202. https://doi.org/10.1016/b978-1-55860-377-6.50032-3

Engler, T. H., Winter, P., & Schulz, M. (2015). Understanding online product ratings: A customer satisfaction model. *Journal of Retailing and Consumer Services*, *27*, 113–120. https://doi.org/10.1016/j.jretconser.2015.07.010

Eryiğit, G. (2014). ITU Turkish NLP web service. *Proceedings of the Demonstrations at the 14th Conference of the European Chapter of the Association for Computational Linguistics*, 1–4.

Eryiğit, G., Nivre, J., & Oflazer, K. (2008). Dependency parsing of Turkish. *Computational Linguistics*, *34*(3), 357–389. https://doi.org/10.1162/coli.2008.07-017-R1-06-83

Fan, Z.-P., Li, G.-M., & Liu, Y. (2020). Processes and methods of information fusion for ranking products based on online reviews: An overview. *Information Fusion*, *60*, 87–97. https://doi.org/10.1016/j.inffus.2020.02.007

Fan, Z.-P., Xi, Y., & Liu, Y. (2018). Supporting consumer's purchase decision: a method for ranking products based on online multi-attribute product ratings. *Soft Computing*, *22*(16), 5247–5261. https://doi.org/10.1007/s00500-017-2961-4

Farooq, M. S., Salam, M., Fayolle, A., Jaafar, N., & Ayupp, K. (2018). Impact of service quality on customer satisfaction in Malaysia airlines: A PLS-SEM approach. *Journal of Air Transport Management*, *67*, 169–180. https://doi.org/10.1016/j.jairtraman.2017.12.008

Fenton, N., & Neil, M. (2001). Making decisions: Using Bayesian nets and MCDA. *Knowledge-Based Systems*, *14*(7), 307–325. https://doi.org/10.1016/S0950-7051(00)00071-X

Gelman, A., Carlin, J. B., Stern, H. S., Dunson, D. B., Vehtari, A., & Rubin, D. B. (2013). *Bayesian data analysis* (3rd ed.). Chapman & Hall/CRC.

Ghorbani, M., Mohammad Arabzad, S., & Shahin, A. (2013). A novel approach for supplier selection based on the kano model and fuzzy MCDM. *International Journal of Production Research*, *51*(18), 5469–5484. https://doi.org/10.1080/00207543.2013.784403

Gilks, W. R., Richardson, S., & Spiegelhalter, D. J. (1995). Introducing Markov chain Monte Carlo. In W. R. Gilks, S. Richardson, & D. J. Spiegelhalter (Eds.), *Markov chain Monte Carlo in practice* (pp. 1–19). Chapman and Hall /CRC Press. https://doi.org/https://doi.org/10.1201/b14835

Giudici, P., Givens, G. H., & Mallick, B. K. (Eds.). (2013). *Computational statistics* (2nd ed.). Wiley.

Groves, R. M. (2006). Nonresponse rates and nonresponse bias in household surveys. *Public Opinion Quarterly*, *70*(5), 646–675. https://doi.org/10.1093/poq/nfl033

Gu, B., & Ye, Q. (2014). First step in social media: Measuring the influence of online management responses on customer satisfaction. *Production and Operations Management*, *23*(4), 570–582. https://doi.org/10.1111/poms.12043

Härdle, W. (1991). *Smoothing Techniques: With Implementation in S*. Springer.

Hosseini, S., & Ivanov, D. (2020). Bayesian networks for supply chain risk, resilience and ripple effect analysis: A literature review. *Expert Systems with Applications*, *161*, 113649. https://doi.org/10.1016/j.eswa.2020.113649

Hou, T., Yannou, B., Leroy, Y., & Poirson, E. (2019). Mining customer product reviews for product development: A summarization process. *Expert Systems with Applications*, *132*, 141–150. https://doi.org/10.1016/j.eswa.2019.04.069

Huang, W., Pahwa, A., & Kong, Z. (2012). Kernel density estimation and metropolis-hastings sampling in process capability analysis of unknown distributions. *In ASME 2012 International Manufacturing Science and Engineering Conference Collocated with the 40th North American Manufacturing Research Conference and in Participation with the Int. Conf., MSEC 2012*, 651–660. https://doi.org/https://doi.org/10.1115/MSEC2012-7299

Huang, Y., & Bian, L. (2009). A Bayesian network and analytic hierarchy process based personalized recommendations for tourist attractions over the Internet. *Expert Systems with Applications*, *36*(1), 933–943. https://doi.org/10.1016/j.eswa.2007.10.019

Hwang, C., & Yoon, K. (1981). *Multiple attribute decision making methods and applications*. Springer.

IBM Corp. (2022a). *IBM SPSS Modeler 18.3 Algorithms Guide*. https://www.ibm.com/docs/en/SS3RA7_18.3.0/pdf/AlgorithmsGuide.pdf/

IBM Corp. (2022b). *IBM SPSS Neural Networks*. https://www.ibm.com/docs/en/SSLVMB_28.0.0/pdf/IBM_SPSS_Neural_Network.pdf/

Imtiaz, M. N., & Islam, M. K. Ben. (2020). Identifying Significance of Product Features on Customer Satisfaction Recognizing Public Sentiment Polarity: Analysis of Smart Phone Industry Using Machine-Learning Approaches. *Applied Artificial Intelligence*, *34*(11), 832–848. https://doi.org/10.1080/08839514.2020.1787676

Jin, J., Ji, P., & Liu, Y. (2014). Prioritising engineering characteristics based on customer online reviews for quality function deployment. *Journal of Engineering Design*, *25*(7–9), 303–324. https://doi.org/10.1080/09544828.2014.984665

Kaya, R., & Yet, B. (2019). Building Bayesian networks based on DEMATEL for multiple criteria decision problems: A supplier selection case study. *Expert Systems with Applications*, *134*, 234–248. https://doi.org/10.1016/j.eswa.2019.05.053

Kim, H. S., & Noh, Y. (2019). Elicitation of design factors through big data analysis of online customer reviews for washing machines. *Journal of Mechanical Science and Technology*, *33*(6), 2785–2795. https://doi.org/10.1007/s12206-019-0525-5

Koller, D., & Friedman, N. (2009). *Probabilistic graphical models: principles and techniques*. The MIT Press.

Korb, K. B., & Nicholson, A. E. (2010). *Bayesian artificial intelligence* (2nd ed.). CRC Press.

Kostyra, D. S., Reiner, J., Natter, M., & Klapper, D. (2016). Decomposing the effects of online customer reviews on brand, price, and product attributes. *International Journal of Research in Marketing*, *33*(1), 11–26. https://doi.org/10.1016/j.ijresmar.2014.12.004

Koulouriotis, D. E., & Ketipi, M. K. (2011). A fuzzy digraph method for robot evaluation and selection. *Expert Systems with Applications*, *38*(9), 11901–11910. https://doi.org/10.1016/j.eswa.2011.03.082

Kulluk, S., Özbakir, L., & Baykasoğlu, A. (2013). Fuzzy DIFACONN-miner: A novel approach for fuzzy rule extraction from neural networks. *Expert Systems with Applications*, *40*(3), 938–946. https://doi.org/10.1016/j.eswa.2012.05.050

Li, F., Wang, W., Dubljevic, S., Khan, F., Xu, J., & Yi, J. (2019). Analysis on accident-causing factors of urban buried gas pipeline network by combining DEMATEL, ISM and BN methods. *Journal of Loss Prevention in the Process Industries*, *61*(June), 49–57. https://doi.org/10.1016/j.jlp.2019.06.001

Li, X., & Hitt, L. M. (2008). Self-selection and information role of online product reviews. *Information Systems Research*, *19*(4), 456–474. https://doi.org/10.1287/isre.1070.0154

Liang, X., Liu, P., & Wang, Z. (2019). Hotel selection utilizing online reviews: A novel decision support model based on sentiment analysis and DL-VIKOR method. *Technological and Economic Development of Economy*, *25*(6), 1139–1161. https://doi.org/10.3846/tede.2019.10766

Lin, S. Y. (2013). The Influence of Relational Selling Behavior on Relationship Quality: The Moderating Effect of Perceived Price and Customers' Relationship Proneness. *Journal of Relationship Marketing*, *12*(3), 204–222. https://doi.org/10.1080/15332667.2013.816931

Liu, F., Liao, H., & Al-Barakati, A. (2023). Physician selection based on user-generated content considering interactive criteria and risk preferences of patients. *Omega*, *115*, 102784. https://doi.org/10.1016/j.omega.2022.102784

Liu, Y., Bi, J.-W., & Fan, Z.-P. (2017). Ranking products through online reviews: A method based on sentiment analysis technique and intuitionistic fuzzy set theory. *Information Fusion*, *36*, 149–161. https://doi.org/10.1016/j.inffus.2016.11.012

Luo, Y., Zhang, X., Qin, Y., Yang, Z., & Liang, Y. (2021). Tourism Attraction Selection with Sentiment Analysis of Online Reviews Based on Probabilistic Linguistic Term Sets and the IDOCRIW-COCOSO Model. *International Journal of Fuzzy Systems*, *23*(1), 295–308. https://doi.org/10.1007/s40815-020-00969-9

Meng, X., Chen, G., Zhu, G., & Zhu, Y. (2019). Dynamic quantitative risk assessment of accidents induced by leakage on offshore platforms using DEMATEL-BN. *International Journal of Naval Architecture and Ocean Engineering*, *11*(1), 22–32. https://doi.org/10.1016/j.ijnaoe.2017.12.001

Metropolis, N., Rosenbluth, A. W., Rosenbluth, M. N., Teller, A. H., & Teller, E. (1953). Equation of state calculations by fast computing machines. *The Journal of Chemical Physics*, *21*(6), 1087–1092. https://doi.org/https://doi.org/10.1063/1.1699114

Metz, C. E. (1978). Basic principles of ROC analysis. *Seminars in Nuclear Medicine*, *8*(4), 283–298.

Minc, H. (1978). *Permanents*. Addison-Wesley.

Murphy, K. P. (2012). Machine learning: a probabilistic perspective. In *Chance Encounters: Probability in Education*. The MIT Press.

Musal, R. M., Soyer, R., McCabe, C., & Kharroubi, S. A. (2012). Estimating the population utility function: A parametric Bayesian approach. *European Journal of Operational Research*, *218*(2), 538–547. https://doi.org/10.1016/j.ejor.2011.11.017

Neapolitan, R. E. (2004). *Learning Bayesian networks*. Pearson Prentice Hall.

Nie, R.-X., Tian, Z.-P., Wang, J.-Q., & Chin, K. S. (2020). Hotel selection driven by online textual reviews: Applying a semantic partitioned sentiment dictionary and evidence theory. *International Journal of Hospitality Management*, *88*. https://doi.org/10.1016/j.ijhm.2020.102495

Olson, D. L. (2004). Comparison of weights in TOPSIS models. *Mathematical and Computer Modelling*, *40*(7–8), 721–727. https://doi.org/10.1016/j.mcm.2004.10.003

Park, E., Chang, H.-J., & Nam, H. S. (2018). A Bayesian network model for predicting post-stroke outcomes with available risk factors. *Frontiers in Neurology*, *9*, 1–11. https://doi.org/10.3389/fneur.2018.00699

Pearl, J. (1988). *Probabilistic reasoning in intelligent systems: networks of plausible inference*. Morgan Kaufmann Publishers, Inc.

Peláez, J. I., Cabrera, F. E., & Vargas, L. G. (2018). Estimating the importance of consumer purchasing criteria in digital ecosystems. *Knowledge-Based Systems*, *162*, 252–264. https://doi.org/10.1016/j.knosys.2018.07.023

Pena, J., Nápoles, G., & Salgueiro, Y. (2020). Explicit methods for attribute weighting in multi-attribute decision-making: a review study. *Artificial Intelligence Review*, *53*(5), 3127–3152. https://doi.org/10.1007/s10462-019-09757-w

Pena, J., Nápoles, G., & Salgueiro, Y. (2021). Implicit and hybrid methods for attribute weighting in multi-attribute decision-making: a review study. *Artificial Intelligence Review*, *54*(5), 3817–3847. https://doi.org/10.1007/s10462-020-09941-3

Peng, H.-G., Zhang, H.-Y., & Wang, J.-Q. (2018). Cloud decision support model for selecting hotels on TripAdvisor.com with probabilistic linguistic information. *International Journal of Hospitality Management*, *68*, 124–138. https://doi.org/10.1016/j.ijhm.2017.10.001

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

Pietro, L. D., Mugion, R. G., Musella, F., Renzi, M. F., & Vicard, P. (2015). Reconciling internal and external performance in a holistic approach: A Bayesian network model in higher education. *Expert Systems with Applications*, *42*(5), 2691–2702. https://doi.org/10.1016/j.eswa.2014.11.019

Pietro, L. D., Mugion, R. G., Musella, F., Renzi, M. F., & Vicard, P. (2017). Monitoring an airport check-in process by using Bayesian networks. *Transportation Research Part A: Policy and Practice*, *106*, 235–247. https://doi.org/10.1016/j.tra.2017.09.020

Qu, Z., Zhang, H., & Li, H. (2008). Determinants of online merchant rating: Content analysis of consumer comments about Yahoo merchants. *Decision Support Systems*, *46*(1), 440–449. https://doi.org/10.1016/j.dss.2008.08.004

Rao, R. V. (2006). A decision-making framework model for evaluating flexible manufacturing systems using digraph and matrix methods. *The International Journal of Advanced Manufacturing Technology*, *30*, 1101–1110. https://doi.org/https://doi.org/10.1007/s00170-005-0150-6

Ryser, H. J. (1963). *Combinatorial mathematics* (14th ed.). The Mathematical Association of America.

Saaty, T. L. (2003). Decision-making with the AHP: Why is the principal eigenvector necessary. *European Journal of Operational Research, 145*(1), 85–91. https://doi.org/10.1016/S0377-2217(02)00227-8

Saltelli, A., Tarantola, S., Campolongo, F., & Ratto, M. (2004). *Sensitivity analysis in practice: A guide to assessing scientific models*. Wiley.

Scutari, M. (2010). Learning Bayesian networks with the bnlearn R Package. *Journal of Statistical Software, 35*(3), 1–22. https://doi.org/10.18637/jss.v035.i03

Scutari, M., & Denis, J.-B. (2015). *Bayesian networks with examples in R*. CRC Press.

Shannon, C. E., & Weaver, W. (1949). *The mathematical theory of communication*. University of Illinois Press.

Shepherd, A. J. (2012). *Second-order methods for neural networks: Fast and reliable training methods for multi-layer perceptrons*. Springer Science & Business Media.

Silverman, B. W. (2018). *Density estimation: For statistics and data analysis*. Routledge.

Sun, M. (2012). How does the variance of product ratings matter? *Management Science, 58*(4), 696–707. https://doi.org/10.1287/mnsc.1110.1458

Swets, J. A. (1988). Measuring the accuracy of diagnostic systems. *Science, 240*, 1285–1293. https://doi.org/10.1126/science.3287615

Thirumalai, S., & Sinha, K. K. (2011). Customization of the online purchase process in electronic retailing and customer satisfaction: An online field study. *Journal of Operations Management*, *29*(5), 477–487. https://doi.org/10.1016/j.jom.2010.11.009

Topuz, K., Uner, H., Oztekin, A., & Yildirim, M. B. (2018). Predicting pediatric clinic no-shows: a decision analytic framework using elastic net and Bayesian belief network. *Annals of Operations Research*, *263*(1–2), 479–499. https://doi.org/10.1007/s10479-017-2489-0

Tzeng, G. H., & Huang, J. J. (2011). *Multiple attribute decision making: methods and applications*. CRC Press.

UNCTAD (United Nations Conference on Trade and Development). (2020). *COVID-19 and E-Commerce: Impact on Businesses and Policy Responses.* https://unctad.org/system/files/official-document/dtlstictinf2020d2_en.pdf

Valiant, L. G. (1979). The complexity of computing the permanent. *Theoretical Computer Science*, *8*(2), 189–201. https://doi.org/10.1016/0304-3975(79)90044-6

Wang, J. J., Jing, Y. Y., Zhang, C. F., & Zhao, J. H. (2009). Review on multi-criteria decision analysis aid in sustainable energy decision-making. *Renewable and Sustainable Energy Reviews*, *13*(9), 2263–2278. https://doi.org/10.1016/j.rser.2009.06.021

Wang, L., Wang, X.-K., Peng, J.-J., & Wang, J.-Q. (2020). The differences in hotel selection among various types of travellers: A comparative analysis with a useful bounded rationality behavioural decision support model. *Tourism Management*, *76*(October 2018). https://doi.org/10.1016/j.tourman.2019.103961

Wang, X., Leng, M., Song, J., Luo, C., & Hui, S. (2019). Managing a supply chain under the impact of customer reviews: A two-period game analysis. *European Journal of Operational Research*, *277*(2), 454–468. https://doi.org/10.1016/j.ejor.2019.02.033

Wang, Y., Lu, X., & Tan, Y. (2018). Impact of product attributes on customer satisfaction: An analysis of online reviews for washing machines. *Electronic Commerce Research and Applications*, *29*, 1–11. https://doi.org/10.1016/j.elerap.2018.03.003

Wang, Y. M., & Luo, Y. (2010). Integration of correlations with standard deviations for determining attribute weights in multiple attribute decision making. *Mathematical and Computer Modelling*, *51*(1–2), 1–12. https://doi.org/10.1016/j.mcm.2009.07.016

Xie, J., Zhang, S., & Lin, L. (2022). Prediction of network public opinion based on bald eagle algorithm optimized radial basis function neural network. *International Journal of Intelligent Computing and Cybernetics*, *15*(2), 260–276. https://doi.org/10.1108/IJICC-07-2021-0148

Yalcin, A. S., Kilic, H. S., & Delen, D. (2022). The use of multi-criteria decision-making methods in business analytics: A comprehensive literature review. *Technological Forecasting and Social Change*, *174*, 121193. https://doi.org/10.1016/j.techfore.2021.121193

Yang, C., Lan, S., & Tseng M. (2019). Coordinated development path of metropolitan logistics and economy in Belt and Road using DEMATEL–Bayesian analysis. *International Journal of Logistics Research and Applications*, *22*(1), 1–24. https://doi.org/10.1080/13675567.2018.1439907

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

Yu, S.-M., Wang, J., Wang, J.-Q., & Li, L. (2018). A multi-criteria decision-making model for hotel selection with linguistic distribution assessments. *Applied Soft Computing Journal*, *67*, 741–755. https://doi.org/10.1016/j.asoc.2017.08.009

Zardari, N. H., Ahmed, K., Shirazi, S. M., & Yusop, Z. Bin. (2015). Weighting Methods and their Effects on Multi-Criteria Decision Making Model Outcomes in Water Resources Management. In *SpringerBriefs in Water Science and Technology* (Issue Mcdm). Springer.

Zhang, L., Ma, B., & Cartwright, D. K. (2013). The impact of online user reviews on cameras sales. *European Journal of Marketing*, *47*(7), 1115–1128. https://doi.org/10.1108/03090561311324237