

# Evaluating User Experience in Home Health Software Applications: An Online Reviews Analytics Study

Mahmud Alrahhah<sup>a,\*</sup>, Ferhat Bozkurt<sup>a</sup>

<sup>a</sup> Atatürk University, Computer Engineering Department, Yakutiye, Erzurum, Türkiye, Postal Code: 25240

\* Corresponding author email address: [alrahhah24@gmail.com](mailto:alrahhah24@gmail.com)

## Abstract

The growing population has increased the demand for people to benefit from home health care (HHC) agencies, therefore, raising the need to use home health software (HHS) products to improve the quality of service and management of HHC companies. HHS products play a key role in managing and scheduling appointments, which leads to saving time and improving the productivity of HHC companies. Analyzing online reviews of HHS products helps the decision-makers in improving their products which leads to more satisfied customers (i.e. HHC agencies) and increases the quality of the HHS applications. In this study, a total of 5338 online reviews were gathered from the SoftwareAdvise website using a customized crawler. After data preprocessing and imputation, machine learning (ML) techniques including the Latent Dirichlet Allocation (LDA) topic model were used to reveal the major topics that influenced the customers' perception. Self-organized maps (SOM) were utilized to segment the customers according to their ratings behavior which eventually led to six groups. Finally, classification and regression trees (CART) were used to predict customer outcomes based on criteria ratings. The results of this study revealed that customers have similar rating behaviors in terms of ease of use and functionality criteria in HHS. In addition, users of HHS were interested mostly in ten features including scheduling, management, support, privacy, integration, ease of use, accessibility, controllability, flexibility, and efficiency. These outcomes help researchers define their features while delivering surveys for evaluating customers' satisfaction with HHS products.

**Keywords:** Home health, online reviews, machine learning, customer satisfaction.

## 1. Introduction

The increasing ratio of the aging population has raised the need to employ home health care (HHC) and home health software (HHS) services in several countries to alleviate the pressure on the limited traditional medical model (G. Du et al., 2022). Especially during the COVID-19 pandemic, the health industry was back on its heels. HHC visits including doctor, nurse, prenatal care, and reproductive visits are fundamentally important to promote individuals' quality of life and well-being, especially for senior citizens (Koeleman et al., 2012; Pahlevani et al., 2022), patients with chronic diseases (Nilsson et al., 2023), and women. HHC organizations are growing over time, particularly in Europe and North America (Van Eenoo et al., 2018). For a brief period of time following hospitalization, HHC services offer nurse and/or treatment sessions to aid a patient's recovery at home (Russell et al., 2023). The majority of HHC beneficiaries are senior citizens (85%), have either four or more chronic illnesses (66%), and women (63%) (Bankole et al., 2023). About five million

Medicare recipients got HHC services from more than 10,500 providers in 2019 (Bankole et al., 2023). HHC agencies provide health care services at either individuals' homes or home-like settings (Cappanera & Scutellà, 2015; Harrison et al., 2020). HHC agencies deploy HHS services to communicate with customers and schedule orders. These services are software applications or platforms designed specifically for HHC agencies to manage their clinical, operational, and administrative tasks. This indicates the importance of such services in alleviating the pressure on the management of the HHC workload. Consumers' purchase decision is heavily affected by online reviews (Eslami et al., 2018). Therefore, analyzing user-generated content (UGC) in HHS is essential to provide deep insights for the HHS owners to enhance their services according to the revealed customers' perceptions which leads to customer satisfaction.

The increasing number of HHC recipients raises the need to employ HHSs which are platforms that help manage the increasing data as well as mitigate pressure

on the management staff. This points to the importance of revealing customer satisfaction with HHS using machine learning (ML) approaches, and UGC shared on the SoftwareAdvice website to improve both the quality of service and function of HHC. SoftwareAdvice is a website that provides reviews, ratings, comparisons, and research on different software products and tools for businesses including HHS, accounting, customer relationship management (CRM), project management, human resources (HR), and more (softwareadvise). The UGC presented on this platform stands as an important source for researchers to perform data analyses and reveal customers' satisfaction dimensions that can help the targeted enterprises improve their services.

As the number of similar products on the market continues to grow, companies are increasingly competing with each other, leading to fierce rivalries (Li et al., 2023). Consequently, considering the customers' preferences provided by online reviews is essentially critical for businesses to enhance their services accordingly. With the fast advancements in social media networks, the number of people sharing their recommendations in the form of online reviews has increased making it an important representative of customers' needs (Y. Du et al., 2022). Therefore, the researchers have concentrated on information extraction and mining of online reviews using natural language processing (NLP) techniques in the last few years (Afrizal et al., 2019; Alrahhhal & Bozkurt). Examining these online reviews using ML approaches enables revealing the customers' satisfaction dimensions in order to present services and products that align with the users' needs (J. Zhang et al., 2021). The research in the literature has shown that UGC is important for helping customers to find new information, reduce doubts about

specific services, influence the decisions of others, and save time in their search (Amatulli et al., 2019; Bi et al., 2020; Gavilan et al., 2018; Jia, 2020; Zhang & Hanks, 2018). Information provided by users' complaints in the form of UGC frequently exposes issues related to product features or service quality. Consequently, examining the UGC using ML approaches is critical for helping companies upgrade their products and services (Li et al., 2023) according to the consumers' demand. Although there are several studies examined UGC in several fields (Ahani et al., 2021; Lucini et al., 2020; Nilashi, Abumalloh, et al., 2021; Nilashi et al., 2018; Nilashi, Yadegaridehkordi, et al., 2019; Rita et al., 2022; Y. Song et al., 2022; Wang et al., 2018), to the best of our knowledge, analyzing UGC in the context of HHSs using ML techniques hasn't been explored yet. Some works in the literature investigated the customers' satisfaction by employing 2 stage methodologies including ML and survey-based approaches (Nilashi, Abumalloh, et al., 2021; Nilashi, Abumalloh, et al., 2022). In this study, ML approaches were investigated for understanding customer perception using UGC. The main contribution of this study is to reveal the customers' satisfaction dimensions and examine their rating behavior in the context of HHS by analyzing the UGC using ML techniques.

The rest of this paper is structured as follows: The literature review is provided in Section 2. The deployed methods in this study are discussed in detail in Section 3. The results of the data analyses are presented in Section 4. The discussion of this study is provided in 5. In Section 6 the conclusion of this study and recommendations for future work are provided.

A list of abbreviations used in this study is given in Table 1.

**Table 1.** List of abbreviations

Abbreviation	Full Term
ML	Machine Learning
LDA	Latent Dirichlet Allocation
HHS	Home Health Software
HHC	Home Health Care
NLP	Natural language Processing
HOSVD	Higher-Order Singular Value Decomposition
SOM	Self Organized Maps
CART	Classification and Regression Trees
UNWTO	United Nations World Tourism Organization
UGC	User-Generated Content
NLP	Natural Language Processing
PHCT	Palliative Home Care Teams
PCA	Principal Component Analysis

## 2. Background

### 2.1. Online Reviews

UGC is content that is produced by consumers, including discussions and social interactions among users, numerical product ratings, "Like", and detailed

review comments (Zhao et al., 2022). As such, this content is present on a variety of mediums such as social media platforms, application distribution channels, third-party review websites, marketplaces, and brand-owned websites or applications (Ray et al., 2021). UGC is a significant information source for service providers,

potential clients, and scholars. UGC-based recommendation systems help customers with the decision of whether to consider a specific product or not (Cascio et al., 2015; Nilashi, Samad, et al., 2022; Nilashi, Samad, et al., 2021; Zablocki et al., 2019). In order to improve their products, service providers can benefit from the UGC as it offers comprehensive feedback on customer experience regarding their offerings (Jiménez & Mendoza, 2013). In addition, UGC offers the possibility of investigating a considerable amount of user experiences in a structured manner, which can produce scientific knowledge and understanding (Hu & Trivedi, 2020; Huang, 2017). As such, UGC can be harnessed to analyze customer satisfaction (Szymanski & Hise, 2000). For instance, UGC shared on the TripAdvisor website has been widely used in academic research, hotel and tourism management, and information systems (Xiang et al., 2015). The transition from traditional word-of-mouth (WOM) to electronic word-of-mouth (e-WOM) has brought about a change in the research approach (Marine-Roig, 2019). Instead of relying on paper surveys and in-person interviews, researchers in the tourism and hospitality sector have begun to collect UGC from tourism platforms and social media to gain a deeper understanding of customers' viewpoints and choices (Marine-Roig, 2019). The authors (Herrero et al., 2015) investigated how users perceive hotel reviews and found that users rely on three factors when making decisions: the credibility of the source, the perceived value of the information provided, and the similarity between the review generators and the users. The research in the literature has emphasized the importance of analyzing online reviews to reveal customers' satisfaction in many fields (Bian et al., 2022; Sezgen et al., 2019; Zibarzani et al., 2022). However, there is a lack of studies in terms of revealing customers' satisfaction in HHS using ML approaches.

## 2.2. Text Mining Applications on Online Reviews

Text-mining methods have been utilized in the previous literature in order to examine various factors that impact customers' satisfaction either positively or negatively using UGC (Abumalloh et al., 2024; Ahani, Nilashi, Ibrahim, et al., 2019). A fuzzy evaluation approach was established for measuring customer satisfaction via online reviews (Liang et al., 2021). The Technique for Order of Preference by Similarity to Ideal Solutions (TOPSIS) was deployed to examine the satisfaction of tourists as an ML approach in the Canary Islands (Ahani, Nilashi, Yadegaridehkordi, et al., 2019). According to (Xu, 2021), the satisfaction of customers for takeaway restaurants is influenced by the drivers' performance. In (Bilgihan et al., 2018), the authors observed three discernible types of customer perceptions that were evident in their reviews of restaurants: mechanical, functional, and humanistic. The study in (Wang et al., 2019) identified six emotions, joy,

anger, sadness, trust, fear, and disgust, expressed in reviews that could indicate the level of satisfaction or dissatisfaction of the reviewers. Another study (Yan et al., 2015) examined UGC from a Chinese online community and found that variables such as atmosphere, price, and value mentioned in restaurant reviews could predict customers' intention to revisit. The outcomes of their study validated the outcomes of analogous research that applied sentiment analysis to reviews on Google Maps (Mathayomchan & Taecharungroj, 2020). The price wasn't found to be a key indicator of users' experience regarding their data analysis of Airbnb (Li et al., 2020).

Topic modeling is one of the most used ML approaches in the context of revealing customers' satisfaction using UGC. Humanly understanding the factors that impact customers' satisfaction is impossible due to the increase in number of the UGC (Heng et al., 2018). Topic modeling is an effective approach that utilizes natural language processing to uncover topics from an extensive collection of unstructured reviews (Wu & Chang, 2020). Topic modeling has been utilized to identify significant emotions or opinions related to important characteristics pertinent to electronic gadgets (Ma et al., 2013), digital assistants (Sánchez-Franco et al., 2021), customers' perception of the service quality in the context of rental accommodation (Ding et al., 2020) as well as in airline hotels (Moro et al., 2020) and restaurants during COVID-19 outbreak (Zibarzani et al., 2022). It has also been used to uncover the desires and worries of customers regarding amusement parks (Luo et al., 2020), and to assess satisfaction of customers using pharmacy reviews (Liu et al., 2020). In the study by (Nasiri & Shokouhyar, 2021), Latent Dirichlet Allocation (LDA) was deployed as a text-mining approach in order to reveal the perceived value of online reviews in the context of online retailing.

## 2.3. ML Applications in the Health Sector

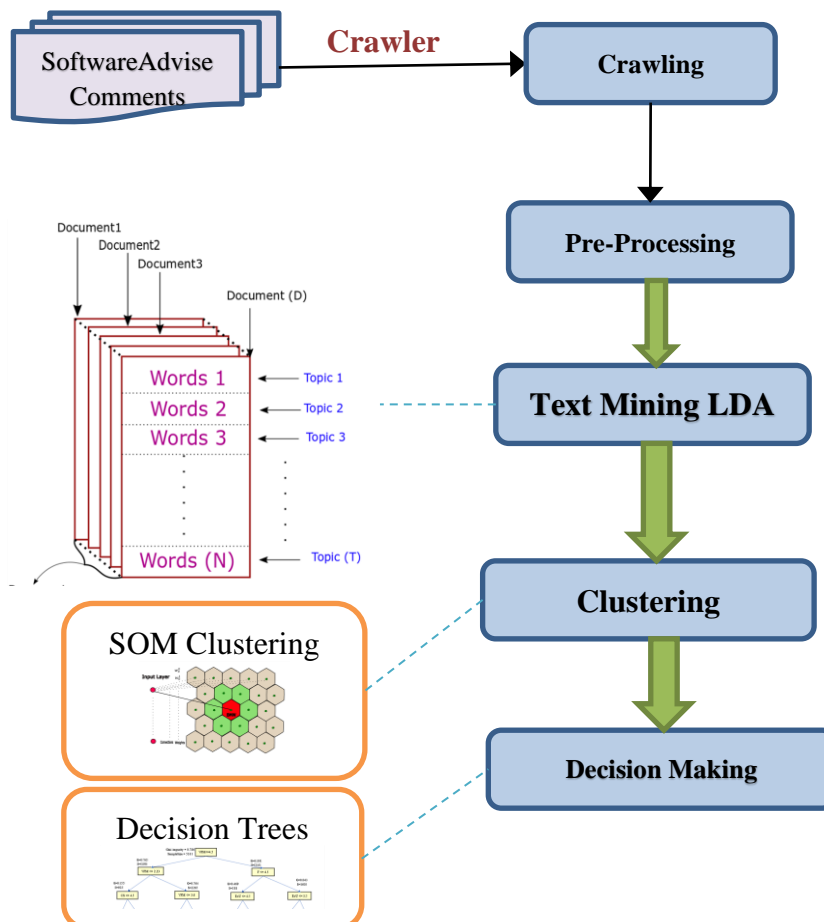
The recent development of ML and deep learning approaches has allowed researchers to conduct many experiments in several fields including the health sector (Tiryaki, 2023) (McArthur et al., 2020) (J. Song et al., 2022). In this (J. Song et al., 2022), the authors deployed ML approaches in their risk models. Their results revealed that the integration of clinical notes with structured data enhanced the identification of patients who are at risk of experiencing emergent medical events that require urgent care during HHC. The authors in this (Topaz et al., 2020) study deployed predictive models based on ML and text mining to anticipate the probability of patient hospitalization or emergency department visits by analyzing the content of HHC clinical notes. In this study (Song et al., 2021), by comparing three ML approaches using clinical notes, risk factors were extracted and utilized to develop predictive models to define patients' probability of experiencing a wound infection during their HHC

treatment. In (Jones et al., 2018), in order to predict hospital utilization and emergency department for patients who receive HHC, the authors compared three ML approaches against traditional statistical methods. In this (Kuspinar et al., 2019), a ML approach was utilized to prevent first-time falls among elderly individuals receiving Home HHC, thereby mitigating the substantial personal and financial costs associated with such falls. In this (McArthur et al., 2020), the authors deployed an ML method to predict the occurrence of hip fractures in home care residents over one year. They aimed to develop and validate this method using the Resident Assessment Instrument for Home Care (RAI-HC), which is a standardized assessment tool used to evaluate the needs and strengths of individuals who receive home care services. In this (Zhu et al., 2007), the authors compared a ML approach and a Clinical Assessment Protocol (“ADLCAP”) in order to predict the rehabilitation potential. Their findings demonstrated how a machine-learning algorithm could improve the quality of clinical decision-making. Another ML approach was deployed by (Nabal et al., 2014) in order to predict the probability of patient survival at 7 and 30 days using symptoms detected by palliative home care teams (PHCTs). The study suggests that classification schemes and decision trees based on specific symptoms

can assist clinicians in predicting the survival of patients with advanced cancer treated by PHCTs. In this (Yang et al., 2021), the authors analyzed healthcare-related data of advanced cancer patients who were under home care using a data mining algorithm. They compared 3 ML algorithms: support vector machine, random forest algorithm, and neural network, and found that the most appropriate model for predicting both patient survival and medical expenses was the random forest algorithm.

### 3. Methodology

This section presents the details of our methodology depicted in Fig. 1. In order to achieve the objective of this research, UGC was gathered from the SoftwareAdvice website using a crawling technique. Afterward, a data imputation step was performed using HOSVD (Nilashi, Ahani, et al., 2019; Nilashi et al., 2014). LDA was utilized to extract the major topics from the UGC. The self-organized maps (SOM) clustering technique was applied to partition the customers into similar clusters according to their satisfaction behaviors. Finally, the Classification and Regression Tree (CART) prediction approach was deployed for the prediction of new customers.

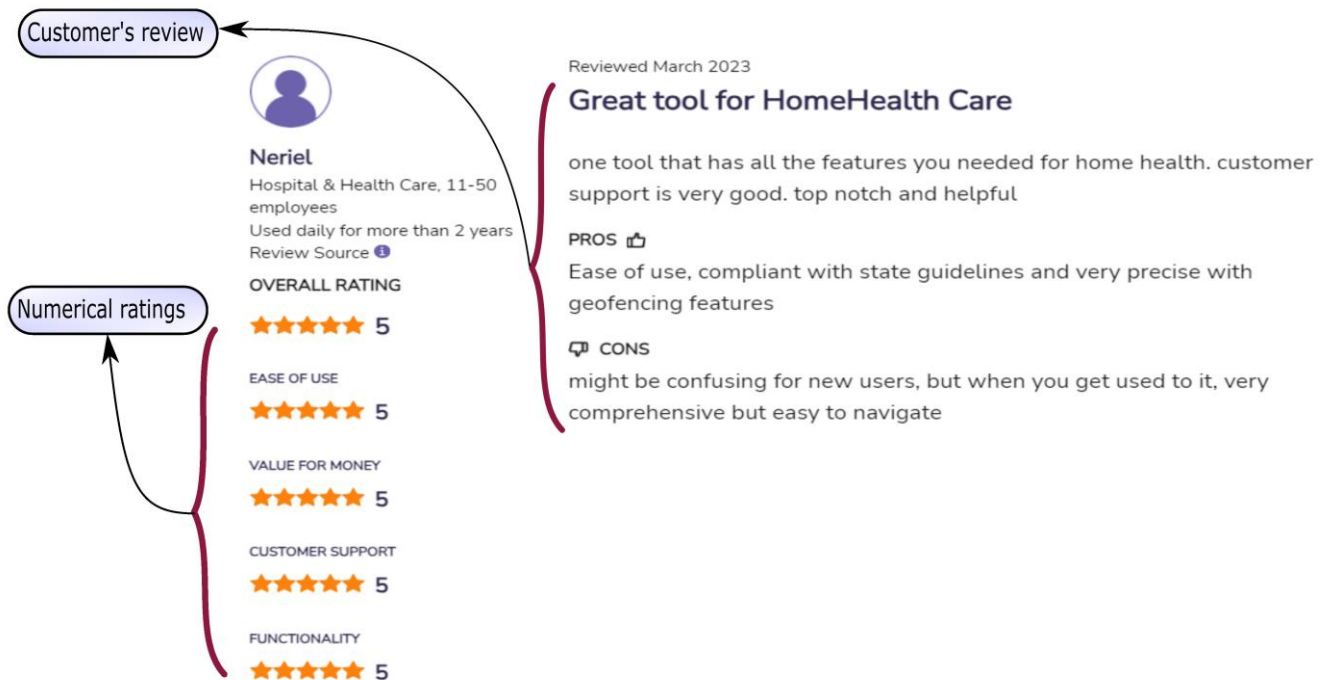


**Fig. 1.** Proposed research method for automated analysis of customer reviews

### 3.1. Data collection and *p* reprocessing

UGC is a powerful kind of data that includes important information for both health applications' users and service providers (Haggag et al., 2022). This information provides the service providers understanding of the major factors that impact the customers' satisfaction in order to help them update their product regularly, increase publicity, and refine their marketing strategy (Qiao et al., 2018). The data collection stage was performed to collect the UGC located on the SoftwareAdvise. We developed a customized crawler

in order to gather the customers' reviews. The crawler was developed with Python programming language using the Selenium and BeautifulSoup library. Using this crawler, we collected 5338 reviews of textual data and criteria ratings expressed by users. An example of the gathered data is presented in Fig. 2. The numerical data consists of 5 features: overall rating, ease of use, value for money, customer support, and functionality. The customers express their level of satisfaction by rating the HHS products ranging from 0 to 5 on the aforementioned features. Data cleaning procedures were utilized to remove unnecessary information from the dataset.

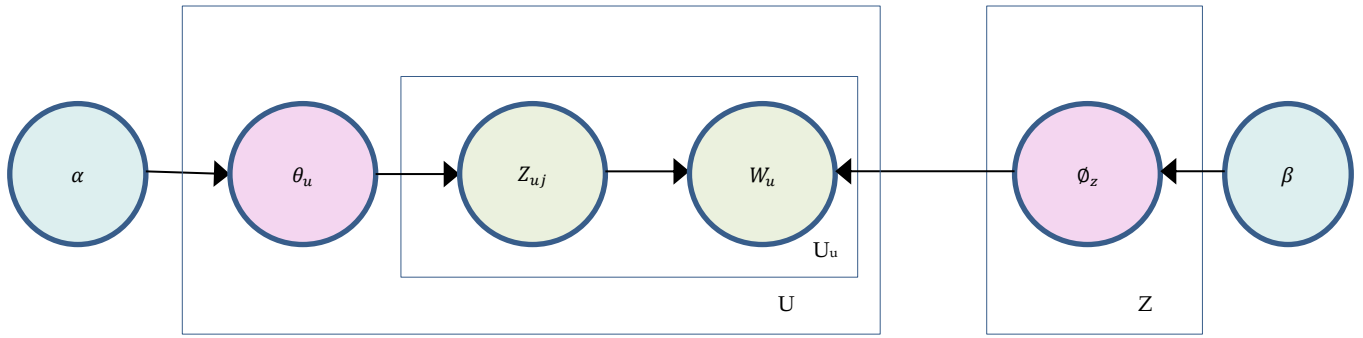


**Fig. 2.** A customer's review on SoftwareAdvise website

### 3.2. LDA topic modeling

The topic modeling technique aims to reveal latent information from unstructured textual data and explore hidden themes. LDA is a probabilistic model that supposes topics to be a probability distribution across a group of words, while a document is considered a combination of topics (Jung & Suh, 2019). The topic numbers should be

defined in the LDA. Scholars have indicated that assigning 20 topics have given the best performance (Williams & Betak, 2018). LDA was utilized widely in text mining studies in the literature to investigate the UGC and inspect the customers' perception in several fields including the accommodation business (Sim et al., 2021), marketing (Huang et al., 2022), and online education (Wei & Taecharunroj, 2022). We present the generative model of the LDA in Fig. 3.



1. For each topic  $z \in Z$ 
  - Draw a multinomial distribution  $\phi_z \sim \text{Dir}(\vec{\beta})$ .
2. For every user  $u \in U$ ,
  - Draw a multinomial distribution  $\theta_u \sim \text{Dir}(\vec{\alpha})$ .
  - For every Word  $w \in D_u$ ,
    - (a) Draw a topic  $z \sim \text{Multinomial}(\vec{\theta}_u)$ .
    - (b) Draw a word  $w \sim \text{Multinomial}(\vec{\phi}_z)$ .

**Fig. 3.** LDA Generative Procedure

### 3.3. Missing values imputation using HOSVD

In several real-world applications, tensor as well as matrix factorization has widely been used (Ahani, Nilashi, Ibrahim, et al., 2019). In this study, the higher-order singular value decomposition (HOSVD) which is an extended version of SVD, was utilized in missing data imputation. By breaking down tensors into their fundamental components, this technique reduces the dimensions of the data, enabling the calculation of similarity measures (Nilashi, Abumalloh, et al., 2022). Data imputation using HOSVD entails estimating missing values in a tensor by approaching it with a dimensionally reduced tensor. Forms that do not fit in a single column should be centered on the entire page.

### 3.4. SOM clustering

Following the data imputation stage, a clustering method was employed in order to segment the customers with similar behaviors based on their reviews. Consequently, we applied the SOM clustering approach which is a well-known clustering technique that tackles large input data sets. Kohonen initially proposed SOM (Kohonen, 1990), which is an unsupervised type of artificial neural network (Qiang et al., 2023). During training, SOM learns to create a low-dimensional output grid from high-dimensional input data using nonlinear methods (Zhang et al., 2023).

The SOM clustering approach consists of two layers; the input layer and the output layer. The SOM is a fully connected artificial neural network (ANN) because each input vector in the input layer is connected to the entire units inside the lattice. As presented in Figure 4, the SOM receives

the training data through the input layer which consists of  $m$ -dimensional vectors that stands for the features of the dataset. The output layer is a lattice; made of interconnected units which are typically arranged in a  $u \times v$  grid pattern. As presented in Fig. 4, the units are connected to their adjacent units using a neighborhood that can be either rectangular or hexagonal in shape. Each neuron in the output layer is associated with a weight vector that connects it to the input neurons. In order to initialize the values of the weights, an available initialization method can be used to randomly generate the initial weights of the SOM clustering method (Attik et al., 2005).

As illustrated in Fig. 4, every neuron in the input layer which is represented as a red circle is connected to the units in the SOM lattice. The Best Matching Unit (BMU), represented by the red unit, is connected to the input unit by bold lines. The weights are denoted by  $w_0^0, w_0^1, \dots, w_{m-1}^n$ . The green units denote the neighborhood of the red BMU. The BMU is identified using a similarity measure like cosine distance, Euclidean, and absolute value (Wan et al., 2002). After identifying the BMU, the weights of the BMU and its neighboring units are updated through a process that considers their topological proximity. During this updating process, the input vector 'pulls' the BMU and its neighbors towards itself. The weights are updated during the iteration regarding the neighborhood size and the learning rate.

The SOM approach modifies the weight vectors of the BMU and its neighboring to minimize the distance between the weight vectors and the current sample. This process is repeated until convergence (Vesanto & Alhoniemi, 2000).

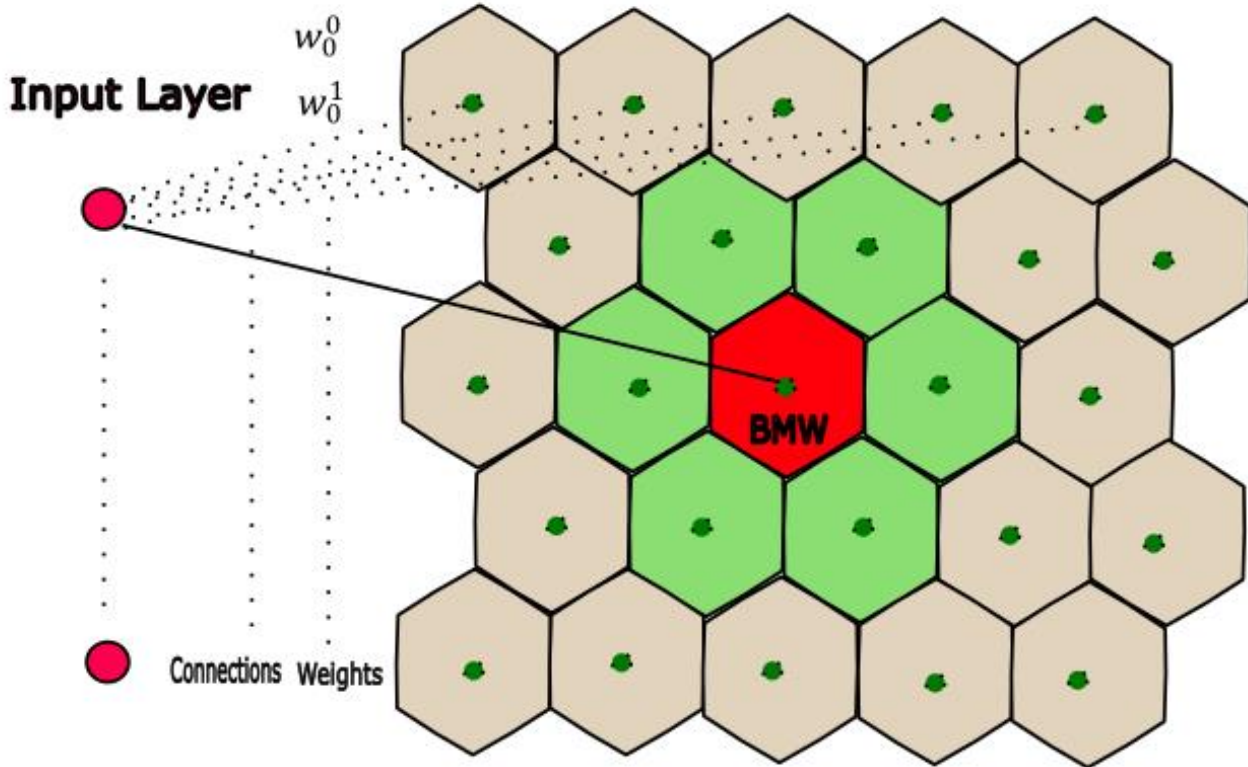


Fig. 4. The SOM Approach Lattice

### 3.4. CART approach

CART is an effective way to solve regression problems due to its ability to identify non-linear relationships without transforming variables (Breiman, 1984; Steinberg, 1997). As such, this method is widely employed in decision-making systems to establish the relationship between inputs and outputs. Initially, the algorithm partitions the input variables into value intervals and then constructs a tree structure consisting of nodes that test the value of each interval, as well as leaves that correspond to the predicted value (Galicia et al., 2019).

The CART approach measures the quality of a split in the decision tree using the Gini impurity measure, which is a criterion used to evaluate the homogeneity of the target variable in each subset of data after being split. The Gini impurity calculates the probability of misclassifying a randomly chosen item from the dataset if it were labeled randomly, by taking into account the weight of the probability of each label in the dataset (Chen et al., 2023). The higher the Gini impurity, the higher the probability of misclassification. The Gini impurity can be calculated as:

$$\text{Gini impurity} = 1 - \sum_{i=1}^c (p_i^2)$$

where  $i$  ranges over all classes in the dataset,  $c$  indicates the number of classes in the dataset, and  $p_i$  refers to the probability of the  $i$ th class label in the dataset.

## 4. Data Analyses

Social websites are important resources for data collection. In this study, we gathered 5338 online reviews as well as ratings from SoftwareAdvise website using the customized crawler. The Data analyses using ML approaches were applied to reveal customers' satisfaction with the HHS applications. The process of analyzing the data entails four major steps: Data imputation and pre-processing, revealing customers' dimensions of satisfaction, segmenting the customers into different groups according to their behaviors of ratings, and finally, applying the customer prediction approach in order to predict the new customers regarding their level of satisfaction.

Data cleaning and, and data imputation processes were applied for textual reviews as well as criteria ratings. The former entails deleting stop words, emails, single quote characters, and consecutive whitespaces (e.g. tabs, newlines, spaces) and replacing them with a single space character. The latter involves the missing data imputation using HOSVD for numerical criteria ratings.

The LDA topic modeling technique was applied to the textual data in order to reveal the main topics that impacted the customers' satisfaction in HHS. Python programming language and genism library were utilized in this process. Using Pyldavis we visualized the extracted topics as

presented in Fig. 5. Meaningful topic models should exhibit large, non-overlapping clusters in the visualization, indicating that the topics are distinct from one another and easily separable (i.e., the clusters are “well-separated”). Such models will be more interpretable than those with small, overlapping, or scattered clusters (Nicolas et al., 2021; Sievert & Shirley, 2014). We modified the number of topics until we achieved data segments that do not overlap with each other (N. Zhang et al., 2021). Consequently, obtained 10 topics that are segregated from each other. In figure. 5, on

the left side, the 10 topics were visualized as circles; larger circles indicate more prevalence of a topic. On the right side of the Fig. 5, the top 30 most significant expressions related to a particular topic were displayed. From these expressions, it is obvious that the customers were interested in features such as ease of use, time, and helpfulness. In Fig. 6, we present a sample word cloud of a particular topic. From this word cloud, it is clear that terms such as schedule, time, online, efficiency, accessible, and flexible were prevalent and caught the customers’ attention in the context of HHS.

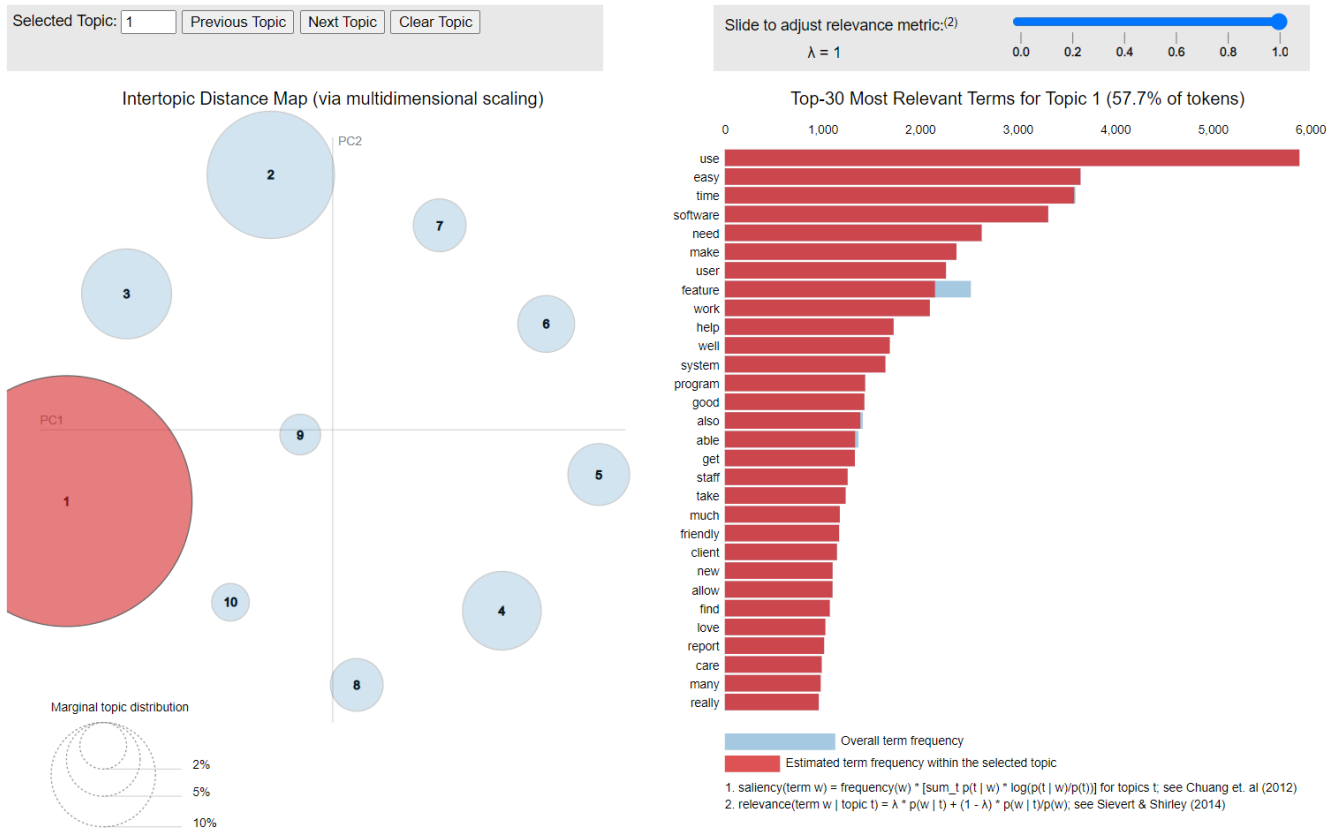


Fig. 5. Topics Distribution

Following the data imputation and LDA topic modeling approach, in order to understand the customers’ satisfaction with each feature of the criteria ratings, the SOM clustering technique was applied. Using the ‘minisom’ library in Python for the SOM clustering approach, the customers were segmented into six different groups each group representing a particular level of satisfaction of the customers. The customer segmentation process is essentially important for predicting the satisfaction of new customers. We present six segments of customers in Table 2. It is obvious in the Table 1 that customers in Segment1 have shown a very low level of satisfaction with the HHS in general, especially for the ease of use feature. In Segment 2 and Segment 3, the customers expressed a moderate level of satisfaction in general and were more satisfied with the value-for-money feature compared to the remaining three features. In Segment4 and Segment5 high level of satisfaction was expressed. Finally, the customers in Segment 3 were highly satisfied, especially with the first three features displayed in the Table. 1.

In order to better inspect the customers’ rating behavior for each feature and visualize them in 2-dimensional figures, we applied the Principal Component Analysis (PCA) approach which is a technique used for dimensionality reduction. The outcomes of the customers’ rating behavior are illustrated in Fig. 7. As presented in Fig. 7, it is obvious that the customers’ had similar rating behavior in the features of ease of use and functionality.

Finally, for the prediction of the new customers we deployed CART. Prediction to which cluster the customer belongs is important for the HHS owners. The CART approach splits the dataset into smaller datasets to decrease the Gini impurity which indicates the homogeneity of the resulting subsets. We have presented the first four steps of the CART algorithm in Fig. 8, which illustrates the process of constructing a decision tree for the dataset. These steps involve recursively partitioning the data based on feature values, calculating the Gini impurity of each resulting subset, selecting the feature that produces the greatest reduction in impurity, and repeating the process until a



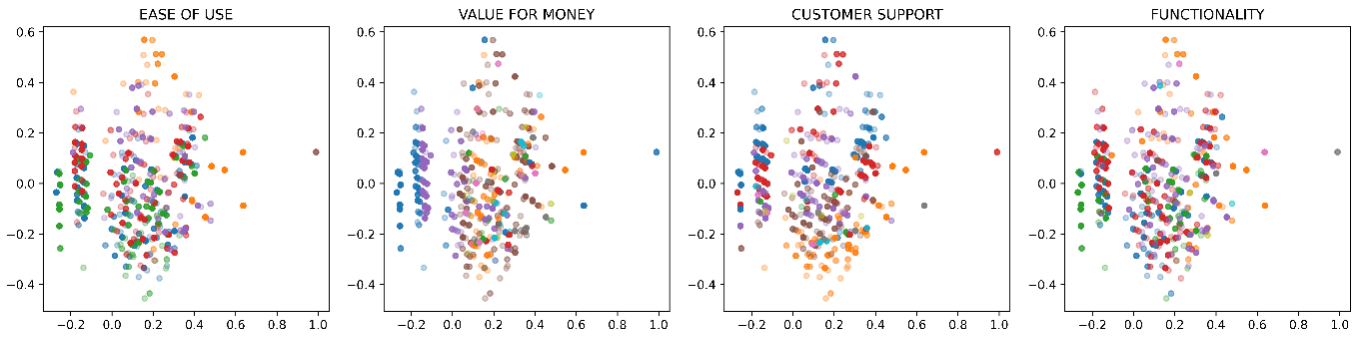


Fig. 7. Customers' Rating Behaviors

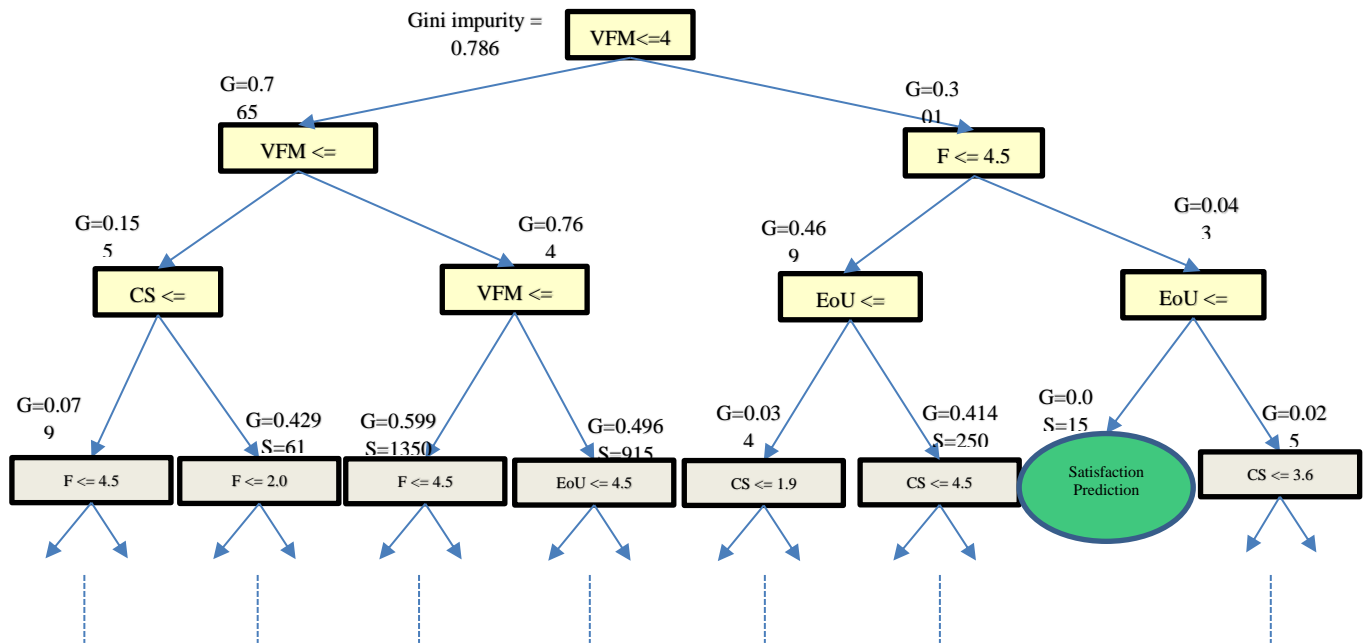


Fig. 8. CART Structure

5. Discussion

Customer loyalty is influenced by customer satisfaction, which is evidenced by positive comments, repeat purchases, and the recommendation of the product or service to others (Forgas et al., 2010; Nilashi, Abumalloh, et al., 2022). On the other hand, customers may express their dissatisfaction by sharing negative reviews, which can harm a business's reputation or image (Namukasa, 2013). ML approaches can be deployed in order to examine the customers' online reviews as well as criteria ratings to understand the factors that affect customers' satisfaction. Although ML approaches have been utilized in previous literature to reveal customers' satisfaction in the context of healthcare (Chatterjee et al., 2021), revealing customers' satisfaction in HHS using textual reviews as well as criteria ratings hasn't been investigated widely. In addition, we deployed ML approaches that are different from the aforementioned work in order to inspect the data from a different viewpoint. With the increasing population throughout the world, HHC has become essentially important to mitigate the pressure on the

health sector, especially during the COVID-19 pandemic. HHS plays a key role in supporting the management of the HHC business enterprises and constructing a bridge between the patients and HHC providers. Hence this study focuses on examining customers' satisfaction to contribute to the HHC by examining the online reviews of HHSs posted on SoftwareAdvise website.

In order to reveal the factors that caught the attention of the customers LDA topic modeling was utilized to extract 10 topics from the data. It is apparent from the word cloud sample presented in Figure. 6. That customers were most interested in terms and functions including scheduling which plays an important role in HHS for scheduling appointments for patients. Besides, customers were interested in other features of the product such as accessibility, flexibility, and efficiency. Following topic modeling, to understand the rating behavior of customers in separate clusters, customers were partitioned into 6 variant groups. In segment 1 customers were more satisfied with functionality and dissatisfied with the ease of use. In segments 1 and 2,

customers were relatively low satisfied with the HHS. Segment 3 and 4 shows a moderate level of satisfaction in general. In segment 5, the customers' level of satisfaction was high, especially for value for money and customer support. The customers in Segment 6 were highly satisfied with all features. As presented in the Figure 7. our findings revealed that customers have similar rating behavior with the ease of use and functionality features. The extracted features and topics from the data are significant in helping product

## 6. Conclusion, Limitations, and Future Work

Online reviews and criteria ratings posted on social websites play a key role in understanding customers' behavior, as well as their satisfaction. As such, these ML approaches can be applied to this data to gain a bird's-eye view of customers' preferences and dimensions of satisfaction and dissatisfaction. Companies should leverage the power of ML to analyze data and reveal hidden insights about their customers' points of view as it helps them enhance their software products, which leads to increased software quality, profits and customer satisfaction. By deploying the ML approach to examine UGC in HHS posted on SoftwareAdise website, firstly, we revealed the dimensions that impacted the customers' satisfaction using LDA. The LDA approach found 10 topics of the textual data. The topics were well-separated using Python programming language and the Genism library which indicates that the approach was successful in finding the topics. LDA revealed several dimensions of satisfaction that impacted the customers' satisfaction such as efficiency, accessibility, and flexibility. Secondly, we partitioned the customers into 6 various clusters to understand the customers' preferences based on the features of the ratings. These clusters are important in predicting new customers' level of satisfaction. The levels of satisfaction consist of 6 levels including very low, low, moderate, moderate close to high, high, and very high. Our findings turned out that the customers have similar rating behavior with ease of use and functionality. Finally, we deployed the CART approach to predict new customers' level of satisfaction. This study has few limitations in terms of data collection and methodology. In order to have a broader understanding, the data can be collected not only using UGC but also from companies' records. Having a big dataset help researchers to reveal more accurate results as well as broader insights. The UGC increases and differs over time. Researchers can analyze the data based on specific periods of time for future research. In terms of methodology, using fuzzy logic and ANFIS can be deployed in the new customer prediction stage to predict the customer's level of satisfaction. In future research, questionnaire-based approaches such as PLS-SEM and smartPLS can be deployed to analyze the customers' satisfaction dimensions revealed by LDA. Besides, the customers can be partitioned using a clustering technique based on the extracted 10 topics by LDA. These can give a broader insight into the customers' satisfaction in the context of HHS.

owners as well as HHC providers in their decision-making process. Therefore, they can improve the flexibility, efficiency, accessibility as well as service quality of their service or product. In order to help in the decision-making process, we proposed a CART method that can predict the level of satisfaction of the customers. This can help business enterprises to have a better understanding of their customers' satisfaction.

## References

- Abumalloh, R., Alrahal, M., El-Haggar, N., Alsumayt, A., Alfawaer, Z., & Aljameel, S. (2024). Exploring Individuals' Experiences with Security Attacks: A Text Mining and Qualitative Study. *Emerging Science Journal*, 8(1).
- Afrizal, A. D., Rakhmawati, N. A., & Tjahyanto, A. (2019). New filtering scheme based on term weighting to improve object based opinion mining on tourism product reviews. *Procedia Computer Science*, 161, 805-812.
- Ahani, A., Nilashi, M., Ibrahim, O., Sanzogni, L., & Weaven, S. (2019). Market segmentation and travel choice prediction in Spa hotels through TripAdvisor's online reviews. *International Journal of Hospitality Management*, 80, 52-77.
- Ahani, A., Nilashi, M., Yadegaridehkordi, E., Sanzogni, L., Tarik, A. R., Knox, K., Samad, S., & Ibrahim, O. (2019). Revealing customers' satisfaction and preferences through online review analysis: The case of Canary Islands hotels. *Journal of Retailing and Consumer Services*, 51, 331-343.
- Ahani, A., Nilashi, M., Zogaan, W. A., Samad, S., Aljehane, N. O., Alhargan, A., Mohd, S., Ahmadi, H., & Sanzogni, L. (2021). Evaluating medical travelers' satisfaction through online review analysis. *Journal of Hospitality and Tourism Management*, 48, 519-537.
- Alrahal, M., & Bozkurt, F. Analyzing Big Social Data for Evaluating Environment-Friendly Tourism in Turkey. *Journal of Intelligent Systems: Theory and Applications*, 6(2), 130-142.
- Amatulli, C., De Angelis, M., & Stoppani, A. (2019). Analyzing online reviews in hospitality: Data-driven opportunities for predicting the sharing of negative emotional content. *Current Issues in Tourism*, 22(15), 1904-1917.
- Attik, M., Bougrain, L., & Alexandre, F. (2005). Self-organizing map initialization. *Artificial Neural Networks: Biological Inspirations—ICANN 2005: 15th International Conference*, Warsaw, Poland, September 11-15, 2005. Proceedings, Part I 15,
- Bankole, A. O., Girdwood, T., Leeman, J., Womack, J., & Toles, M. (2023). Identifying unmet needs of older adults transitioning from home health care to independence at home: A qualitative study. *Geriatric Nursing*, 51, 293-302.
- Bi, J.-W., Liu, Y., Fan, Z.-P., & Zhang, J. (2020). Exploring asymmetric effects of attribute performance on customer satisfaction in the hotel industry. *Tourism Management*, 77, 104006.
- Bian, Y., Ye, R., Zhang, J., & Yan, X. (2022). Customer preference identification from hotel online reviews: A neural network based fine-grained sentiment analysis. *Computers & Industrial Engineering*, 172, 108648.
- Bilgihan, A., Seo, S., & Choi, J. (2018). Identifying restaurant satisfiers and dissatisfiers: Suggestions from online reviews. *Journal of Hospitality Marketing & Management*, 27(5), 601-625.
- Breiman, L., Friedman, J.H., Olshen, R.A., Stone, C.J. (1984). *Classification And Regression Trees*. CRC Press.
- Cappanera, P., & Scutellà, M. G. (2015). Joint assignment, scheduling, and routing models to home care optimization: A

- pattern-based approach. *Transportation Science*, 49(4), 830-852.
- Cascio, C. N., O'Donnell, M. B., Bayer, J., Tinney Jr, F. J., & Falk, E. B. (2015). Neural correlates of susceptibility to group opinions in online word-of-mouth recommendations. *Journal of Marketing Research*, 52(4), 559-575.
- Chatterjee, S., Goyal, D., Prakash, A., & Sharma, J. (2021). Exploring healthcare/health-product ecommerce satisfaction: A text mining and machine learning application. *Journal of Business Research*, 131, 815-825.
- Chen, H., Wang, P., Ren, T., Pan, Z., Liu, J., & Ma, Y. (2023). NT-Com: A combined machine learning model for picking up first arrival. *Computers & Geosciences*, 173, 105321.
- Ding, K., Choo, W. C., Ng, K. Y., & Ng, S. I. (2020). Employing structural topic modelling to explore perceived service quality attributes in Airbnb accommodation. *International Journal of Hospitality Management*, 91, 102676.
- Du, G., Tian, Y., & Ouyang, X. (2022). Multi-resources co-scheduling optimization for home healthcare services under the constraints of service time windows and green transportation. *Applied Soft Computing*, 131, 109746.
- Du, Y., Liu, D., & Duan, H. (2022). A textual data-driven method to identify and prioritise user preferences based on regret/rejoicing perception for smart and connected products. *International Journal of Production Research*, 60(13), 4176-4196.
- Eslami, S. P., Ghasemaghaei, M., & Hassanein, K. (2018). Which online reviews do consumers find most helpful? A multi-method investigation. *Decision Support Systems*, 113, 32-42.
- Forgas, S., Moliner, M. A., Sánchez, J., & Palau, R. (2010). Antecedents of airline passenger loyalty: Low-cost versus traditional airlines. *Journal of Air Transport Management*, 16(4), 229-233.
- Galicia, A., Talavera-Llames, R., Troncoso, A., Koprinska, I., & Martínez-Álvarez, F. (2019). Multi-step forecasting for big data time series based on ensemble learning. *Knowledge-Based Systems*, 163, 830-841.
- Gavilan, D., Avello, M., & Martinez-Navarro, G. (2018). The influence of online ratings and reviews on hotel booking consideration. *Tourism Management*, 66, 53-61.
- Haggag, O., Grundy, J., Abdelrazek, M., & Haggag, S. (2022). A large scale analysis of mHealth app user reviews. *Empirical Software Engineering*, 27(7), 196.
- Harrison, K. L., Leff, B., Altan, A., Dunning, S., Patterson, C., & Ritchie, C. S. (2020). What's happening at home: a claims-based approach to better understand home clinical care received by older adults. *Medical care*, 58(4), 360.
- Heng, Y., Gao, Z., Jiang, Y., & Chen, X. (2018). Exploring hidden factors behind online food shopping from Amazon reviews: A topic mining approach. *Journal of Retailing and Consumer Services*, 42, 161-168.
- Herrero, Á., San Martín, H., & Hernández, J. M. (2015). How online search behavior is influenced by user-generated content on review websites and hotel interactive websites. *International Journal of Contemporary Hospitality Management*.
- Hu, F., & Trivedi, R. H. (2020). Mapping hotel brand positioning and competitive landscapes by text-mining user-generated content. *International Journal of Hospitality Management*, 84, 102317.
- Huang, J. (2017). The dining experience of Beijing Roast Duck: A comparative study of the Chinese and English online consumer reviews. *International Journal of Hospitality Management*, 66, 117-129.
- Huang, S., Zhang, J., Yang, C., Gu, Q., Li, M., & Wang, W. (2022). The interval grey QFD method for new product development: Integrate with LDA topic model to analyze online reviews. *Engineering Applications of Artificial Intelligence*, 114, 105213.
- Jia, S. S. (2020). Motivation and satisfaction of Chinese and US tourists in restaurants: A cross-cultural text mining of online reviews. *Tourism Management*, 78, 104071.
- Jiménez, F. R., & Mendoza, N. A. (2013). Too popular to ignore: The influence of online reviews on purchase intentions of search and experience products. *Journal of Interactive Marketing*, 27(3), 226-235.
- Jones, A., Costa, A. P., Pesevski, A., & McNicholas, P. D. (2018). Predicting hospital and emergency department utilization among community-dwelling older adults: Statistical and machine learning approaches. *PLoS One*, 13(11), e0206662.
- Jung, Y., & Suh, Y. (2019). Mining the voice of employees: A text mining approach to identifying and analyzing job satisfaction factors from online employee reviews. *Decision Support Systems*, 123, 113074.
- Koeleman, P. M., Bhulai, S., & van Meersbergen, M. (2012). Optimal patient and personnel scheduling policies for care-at-home service facilities. *European journal of operational research*, 219(3), 557-563.
- Kohonen, T. (1990). The self-organizing map. *Proceedings of the IEEE*, 78(9), 1464-1480.
- Kuspinar, A., Hirdes, J. P., Berg, K., McArthur, C., & Morris, J. N. (2019). Development and validation of an algorithm to assess risk of first-time falling among home care clients. *BMC geriatrics*, 19(1), 1-8.
- Li, H., Liu, Y., Tan, C.-W., & Hu, F. (2020). Comprehending customer satisfaction with hotels: Data analysis of consumer-generated reviews. *International Journal of Contemporary Hospitality Management*, 32(5), 1713-1735.
- Li, Q., Yang, Y., Li, C., & Zhao, G. (2023). Energy vehicle user demand mining method based on fusion of online reviews and complaint information. *Energy Reports*, 9, 3120-3130.
- Liang, D., Dai, Z., & Wang, M. (2021). Assessing customer satisfaction of O2O takeaway based on online reviews by integrating fuzzy comprehensive evaluation with AHP and probabilistic linguistic term sets. *Applied Soft Computing*, 98, 106847.
- Liu, J., Zhou, Y., Jiang, X., & Zhang, W. (2020). Consumers' satisfaction factors mining and sentiment analysis of B2C online pharmacy reviews. *BMC Medical Informatics and Decision Making*, 20(1), 1-13.
- Lucini, F. R., Tonetto, L. M., Fogliatto, F. S., & Anzanello, M. J. (2020). Text mining approach to explore dimensions of airline customer satisfaction using online customer reviews. *Journal of Air Transport Management*, 83, 101760.
- Luo, J. M., Vu, H. Q., Li, G., & Law, R. (2020). Topic modelling for theme park online reviews: Analysis of Disneyland. *Journal of Travel & Tourism Marketing*, 37(2), 272-285.
- Ma, B., Zhang, D., Yan, Z., & Kim, T. (2013). An LDA and synonym lexicon based approach to product feature extraction from online consumer product reviews. *Journal of Electronic Commerce Research*, 14(4), 304.
- Marine-Roig, E. (2019). Destination image analytics through traveller-generated content. *Sustainability*, 11(12), 3392.
- Mathayomchan, B., & Taecharungroj, V. (2020). "How was your meal?" Examining customer experience using Google maps reviews. *International Journal of Hospitality Management*, 90, 102641.
- McArthur, C., Ioannidis, G., Jantzi, M., Adachi, J. D., Giangregorio, L., Hirdes, J., & Papaioannou, A. (2020). Development and validation of the fracture risk scale home care (FRS-HC) that predicts one-year incident fracture: an electronic record-linked longitudinal cohort study. *BMC Musculoskeletal Disorders*, 21, 1-9.

- Moro, S., Lopes, R. J., Esmerado, J., & Botelho, M. (2020). Service quality in airport hotel chains through the lens of online reviewers. *Journal of Retailing and Consumer Services*, 56, 102193.
- Nabal, M., Bescos, M., Barcons, M., Torrubia, P., Trujillano, J., & Requena, A. (2014). New symptom-based predictive tool for survival at seven and thirty days developed by palliative home care teams. *Journal of Palliative Medicine*, 17(10), 1158-1163.
- Namukasa, J. (2013). The influence of airline service quality on passenger satisfaction and loyalty: The case of Uganda airline industry. *The TQM Journal*, 25(5), 520-532.
- Nasiri, M. S., & Shokouhyar, S. (2021). Actual consumers' response to purchase refurbished smartphones: Exploring perceived value from product reviews in online retailing. *Journal of Retailing and Consumer Services*, 62, 102652.
- Nicolas, C., Kim, J., & Chi, S. (2021). Natural language processing-based characterization of top-down communication in smart cities for enhancing citizen alignment. *Sustainable Cities and Society*, 66, 102674.
- Nilashi, M., Abumalloh, R. A., Alghamdi, A., Minaei-Bidgoli, B., Alsulami, A. A., Thanoon, M., Asadi, S., & Samad, S. (2021). What is the impact of service quality on customers' satisfaction during COVID-19 outbreak? New findings from online reviews analysis. *Telematics and Informatics*, 64, 101693.
- Nilashi, M., Abumalloh, R. A., Minaei-Bidgoli, B., Zogaan, W. A., Alhargan, A., Mohd, S., Azhar, S. N. F. S., Asadi, S., & Samad, S. (2022). Revealing travellers' satisfaction during COVID-19 outbreak: moderating role of service quality. *Journal of Retailing and Consumer Services*, 64, 102783.
- Nilashi, M., Ahani, A., Esfahani, M. D., Yadegaridehkordi, E., Samad, S., Ibrahim, O., Sharef, N. M., & Akbari, E. (2019). Preference learning for eco-friendly hotels recommendation: A multi-criteria collaborative filtering approach. *Journal of Cleaner Production*, 215, 767-783.
- Nilashi, M., bin Ibrahim, O., & Ithnin, N. (2014). Multi-criteria collaborative filtering with high accuracy using higher order singular value decomposition and Neuro-Fuzzy system. *Knowledge-Based Systems*, 60, 82-101.
- Nilashi, M., Ibrahim, O., Yadegaridehkordi, E., Samad, S., Akbari, E., & Alizadeh, A. (2018). Travelers decision making using online review in social network sites: A case on TripAdvisor. *Journal of computational science*, 28, 168-179.
- Nilashi, M., Samad, S., Alghamdi, A., Ismail, M. Y., Alghamdi, O., Mehmood, S. S., Mohd, S., Zogaan, W. A., & Alhargan, A. (2022). A new method for analysis of customers' online review in medical tourism using fuzzy logic and text mining approaches. *International Journal of Information Technology & Decision Making*, 21(06), 1797-1820.
- Nilashi, M., Samad, S., Minaei-Bidgoli, B., Ghabban, F., & Supriyanto, E. (2021). Online reviews analysis for customer segmentation through dimensionality reduction and deep learning techniques. *Arabian Journal for Science and Engineering*, 46(9), 8697-8709.
- Nilashi, M., Yadegaridehkordi, E., Ibrahim, O., Samad, S., Ahani, A., & Sanzogni, L. (2019). Analysis of travellers' online reviews in social networking sites using fuzzy logic approach. *International Journal of Fuzzy Systems*, 21, 1367-1378.
- Nilsson, L., Lindblad, M., Johansson, N., Säfström, L., Schildmeijer, K., Ekstedt, M., & Unbeck, M. (2023). Exploring nursing-sensitive events in home healthcare: A national multicenter cohort study using a trigger tool. *International journal of nursing studies*, 138, 104434.
- Pahlevani, D., Abbasi, B., Hearne, J. W., & Eberhard, A. (2022). A cluster-based algorithm for home health care planning: A case study in Australia. *Transportation Research Part E: Logistics and Transportation Review*, 166, 102878.
- Qiang, Z., Hai-Chao, C., Zu-Yuan, L., Bai-Wei, F., Cheng-Sheng, Z., Xide, C., & Xiao, W. (2023). Multi-stage design space reduction technology based on SOM and rough sets, and its application to hull form optimization. *Expert Systems with Applications*, 213, 119229.
- Qiao, Z., Wang, G. A., Zhou, M., & Fan, W. (2018). The impact of customer reviews on product innovation: Empirical evidence in mobile apps. *Analytics and Data Science: Advances in Research and Pedagogy*, 95-110.
- Ray, A., Bala, P. K., & Rana, N. P. (2021). Exploring the drivers of customers' brand attitudes of online travel agency services: A text-mining based approach. *Journal of Business Research*, 128, 391-404.
- Rita, P., Moro, S., & Cavalcanti, G. (2022). The impact of COVID-19 on tourism: Analysis of online reviews in the airlines sector. *Journal of Air Transport Management*, 104, 102277.
- Russell, D., Burgdorf, J. G., Washington, K. T., Schmitz, J., & Bowles, K. H. (2023). "Second set of eyes:" Family caregivers and post-acute home health care during the COVID-19 pandemic. *Patient Education and Counseling*, 109, 107627.
- Sánchez-Franco, M. J., Arenas-Márquez, F. J., & Alonso-Dos-Santos, M. (2021). Using structural topic modelling to predict users' sentiment towards intelligent personal agents. An application for Amazon's echo and Google Home. *Journal of Retailing and Consumer Services*, 63, 102658.
- Sezgen, E., Mason, K. J., & Mayer, R. (2019). Voice of airline passenger: A text mining approach to understand customer satisfaction. *Journal of Air Transport Management*, 77, 65-74.
- Sievert, C., & Shirley, K. (2014). LDAvis: A method for visualizing and interpreting topics. Proceedings of the workshop on interactive language learning, visualization, and interfaces.
- Sim, Y., Lee, S. K., & Sutherland, I. (2021). The impact of latent topic valence of online reviews on purchase intention for the accommodation industry. *Tourism Management Perspectives*, 40, 100903.
- softwareadvice. Retrieved 26.04.2023 from <https://www.softwareadvice.com/>
- Song, J., Hobensack, M., Bowles, K. H., McDonald, M. V., Cato, K., Rossetti, S. C., Chae, S., Kennedy, E., Barron, Y., & Sridharan, S. (2022). Clinical notes: An untapped opportunity for improving risk prediction for hospitalization and emergency department visit during home health care. *Journal of biomedical informatics*, 128, 104039.
- Song, J., Woo, K., Shang, J., Ojo, M., & Topaz, M. (2021). Predictive risk models for wound infection-related hospitalization or ED visits in home health care using machine-learning algorithms. *Advances in Skin & Wound Care*, 34(8), 1-12.
- Song, Y., Liu, K., Guo, L., Yang, Z., & Jin, M. (2022). Does hotel customer satisfaction change during the COVID-19? A perspective from online reviews. *Journal of Hospitality and Tourism Management*, 51, 132-138.
- Steinberg, D., Colla, P. (1997). *CART: Classification and Regression Trees*. Salford Systems; CRC Press.
- Szymanski, D. M., & Hise, R. T. (2000). E-satisfaction: an initial examination. *Journal of retailing*, 76(3), 309-322.
- Tiryaki, V. M. (2023). Mass segmentation and classification from film mammograms using cascaded deep transfer learning. *Biomedical Signal Processing and Control*, 84, 104819.
- Topaz, M., Woo, K., Ryvicker, M., Zolnoori, M., & Cato, K. (2020). Home Health Care Clinical Notes Predict Patient Hospitalization and Emergency Department Visits. *Nursing research*, 69(6), 448.
- Van Eenoo, L., van der Roest, H., Onder, G., Finne-Soveri, H., Garms-Homolova, V., Jonsson, P. V., Draisma, S., van Hout, H., & Declercq, A. (2018). Organizational home care models

- across Europe: A cross sectional study. *International journal of nursing studies*, 77, 39-45.
- Vesanto, J., & Alhoniemi, E. (2000). Clustering of the self-organizing map. *IEEE Transactions on neural networks*, 11(3), 586-600.
- Wan, K. X., Vidavsky, I., & Gross, M. L. (2002). Comparing similar spectra: from similarity index to spectral contrast angle. *Journal of the American Society for Mass Spectrometry*, 13(1), 85-88.
- Wang, X., Tang, L. R., & Kim, E. (2019). More than words: Do emotional content and linguistic style matching matter on restaurant review helpfulness? *International Journal of Hospitality Management*, 77, 438-447.
- 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.
- Wei, X., & Taecharunroj, V. (2022). How to improve learning experience in MOOCs an analysis of online reviews of business courses on Coursera. *The International Journal of Management Education*, 20(3), 100675.
- Williams, T., & Betak, J. (2018). A Comparison of LSA and LDA for the Analysis of Railroad Accident Text. *Procedia Computer Science*, 130, 98-102.
- Wu, J.-J., & Chang, S.-T. (2020). Exploring customer sentiment regarding online retail services: a topic-based approach. *Journal of Retailing and Consumer Services*, 55, 102145.
- Xiang, Z., Schwartz, Z., Gerdes Jr, J. H., & Uysal, M. (2015). What can big data and text analytics tell us about hotel guest experience and satisfaction? *International Journal of Hospitality Management*, 44, 120-130.
- Xu, X. (2021). What are customers commenting on, and how is their satisfaction affected? Examining online reviews in the on-demand food service context. *Decision Support Systems*, 142, 113467.
- Yan, X., Wang, J., & Chau, M. (2015). Customer revisit intention to restaurants: Evidence from online reviews. *Information Systems Frontiers*, 17, 645-657.
- Yang, C., Yu, R., Ji, H., Jiang, H., Yang, W., & Jiang, F. (2021). Application of data mining in the provision of in-home medical care for patients with advanced cancer. *Translational Cancer Research*, 10(6), 3013.
- Zablocki, A., Schlegelmilch, B., & Houston, M. J. (2019). How valence, volume and variance of online reviews influence brand attitudes. *Ams Review*, 9, 61-77.
- Zhang, J., Zhang, A., Liu, D., & Bian, Y. (2021). Customer preferences extraction for air purifiers based on fine-grained sentiment analysis of online reviews. *Knowledge-Based Systems*, 228, 107259.
- Zhang, L., & Hanks, L. (2018). Online reviews: The effect of cosmopolitanism, incidental similarity, and dispersion on consumer attitudes toward ethnic restaurants. *International Journal of Hospitality Management*, 68, 115-123.
- Zhang, N., Liu, R., Zhang, X.-Y., & Pang, Z.-L. (2021). The impact of consumer perceived value on repeat purchase intention based on online reviews: by the method of text mining. *Data Science and Management*, 3, 22-32.
- Zhang, Y., Zhang, Q., Chen, W., Shi, W., Cui, Y., Chen, L., & Shao, J. (2023). Hydrogeochemical analysis and groundwater pollution source identification based on self-organizing map at a contaminated site. *Journal of Hydrology*, 616, 128839.
- Zhao, K., Zhang, P., & Lee, H.-M. (2022). Understanding the impacts of user-and marketer-generated content on free digital content consumption. *Decision Support Systems*, 154, 113684.
- Zhu, M., Chen, W., Hirdes, J. P., & Stolee, P. (2007). The K-nearest neighbor algorithm predicted rehabilitation potential better than current Clinical Assessment Protocol. *Journal of clinical epidemiology*, 60(10), 1015-1021.
- Zibarzani, M., Abumalloh, R. A., Nilashi, M., Samad, S., Alghamdi, O., Nayer, F. K., Ismail, M. Y., Mohd, S., & Akib, N. A. M. (2022). Customer satisfaction with Restaurants Service Quality during COVID-19 outbreak: A two-stage methodology. *Technology in Society*, 70, 101977.