Nilashi, Mehrbakhsh, Abumalloh, Rabab Ali, Almulihi, Ahmed, Alrizq, Mesfer, Alghamdi, Abdullah, Ismail, Muhammed Yousoof, Bashar, Abul, Zogaan, Waleed Abdu and Asadi, Shahla ORCID logoORCID: https://orcid.org/0000-0002-8199-2122 (2023) Big social data analysis for impact of food quality on travelers’ satisfaction in eco-friendly hotels. ICT Express, 9 (2). pp. 182-188. doi:10.1016/j.icte.2021.11.006

Official URL: http://dx.doi.org/10.1016/j.icte.2021.11.006
DOI: http://dx.doi.org/10.1016/j.icte.2021.11.006
EPrint URI: https://eprints.glos.ac.uk/id/eprint/12743

Available online at www.sciencedirect.com

ScienceDirect

ICT Express 9 (2023) 182–188

www.elsevier.com/locate/icte

Big social data analysis for impact of food quality on travelers' satisfaction in eco-friendly hotels

Mehrbakhsh Nilashi[a,*], Rabab Ali Abumalloh[h], Ahmed Almulihi[b], Mesfer Alrizq[c], Abdullah Alghamdi[c], Muhammed Yousoof Ismail[d], Abul Bashar[e], Waleed Abdu Zogaan[f], Shahla Asadi[g]

[a] Centre for Global Sustainability Studies (CGSS), Universiti Sains Malaysia, 11800, George Town, Malaysia
[b] Department of Computer Science, College of Computers and Information Technology, Taif University, P.O. Box 11099, Taif 21944, Saudi Arabia
[c] Information Systems Dept., College of Computer Science and Information Systems, Najran University, Najran, Saudi Arabia
[d] Department of MIS, Dhofar University, Oman
[e] College of Computer Engineering and Science, Prince Mohammad Bin Fahd University, Al-Khobar, Kingdom of Saudi Arabia
[f] Department of Computer Science, Faculty of Computer Science and Information Technology, Jazan University, Jazan, 45142, Saudi Arabia
[g] Centre of Software Technology and Management, Faculty of Information Science and Technology, Universiti Kebangsaan Malaysia, 43600, Bangi, Selangor, Malaysia
[h] Computer Department, Applied College, Imam Abdulrahman Bin Faisal University, P.O. Box. 1982, Dammam, Saudi Arabia

Received 29 July 2021; received in revised form 29 September 2021; accepted 10 November 2021
Available online 17 November 2021

Abstract

Revealing customer satisfaction through big social data has been an interesting research topic in tourism and hospitality. Big data analysis is an effective way to detect customers' behaviors in their decision-making. This study aims to perform big social data analysis to reveal whether food quality impacts the relationship between hotel performance criteria and travelers' satisfaction. A two-stage methodology is developed to address the objectives of this study. The findings demonstrated that there is a positive relationship between eco-friendly hotels' performance criteria and satisfaction. The results and implications for managers and future research directions are discussed.

Keywords: Online customers' reviews; Customers' satisfaction; Machine learning; Food quality; Hotel performance criteria; Big social data

1. Introduction

Analyzing tourists' behaviors in a particular destination regarding food selection is an important issue that needs to be investigated through different approaches. While different researches examined tourists' behaviors in various settings [1], for example, travel and accommodation, the impact of destination food is often neglected. Many studies have been performed for analyzing customers' satisfaction through online customers' reviews [1]. Although the quality of food has been explored in previous literature [2], the moderating role of food quality has not been investigated in both online customers' reviews analysis and survey-based quantitative approaches. In addition, although few studies have been performed to investigate the destination food by the use of survey-based quantitative approaches [3], the use of big social data, which are available in social networking sites, is rarely explored in the area of destination food, particularly its impact on customers' satisfaction [4].

Big social data are massive sizes of data, which are produced from a large number of social media platforms [5]. Big data analysis refers to the analysis of huge amounts of data by utilizing several sophisticated methods and tools to get valuable inferences and to address business needs [6]. To solve the emerging business obstacles, machine learning approaches have been evolved and utilized effectively in several contexts [6]. In fact, utilizing data sources from social media can be effectively used in the analysis of the customers' behaviors and perceptions of the selected destination [7]. Focusing solely on survey-based quantitative approaches, that utilize a limited number of customers' perceptions may not be an

* Corresponding author.
E-mail address: nilashidotnet@hotmail.com (M. Nilashi).
Peer review under responsibility of The Korean Institute of Communications and Information Sciences (KICS).

https://doi.org/10.1016/j.icte.2021.11.006

effective way to assess the destination food quality and its impact on customers' satisfaction and purchase intention. It is found that the deployment of big data analysis to investigate customers' satisfaction levels can be more efficient in terms of generalizability and quality of findings [1].

In this research, revealing customer satisfaction through the analysis of big social data was investigated. We aimed to perform big social data analysis to reveal whether food quality impacts the relationship between hotel performance criteria and travelers' satisfaction. We found that the analysis of big data collected from customers' reviews can be complementary to the survey-based quantitative approaches in customers' satisfaction analysis. Accordingly, a two-stage methodology was developed to meet the aims of this research. In the first stage of our methodology, we employed cluster analysis for customer segmentation and used the supervised learning technique for customers' satisfaction analysis through tourists' data on social networking sites. To perform data segmentation, Learning Vector Quantization (LVQ) was applied to find the most effective groups of travelers with similar preferences. We then applied Classification and Regression Trees (CART) on each segment of LVQ for the prediction of travelers' preferences through a set of eco-friendly hotels' performance criteria. In the second stage of our methodology, we develop a hypothetical model to investigate the paths between the performance criteria and travelers' satisfaction through Partial Least Squares Structural Equation Modeling (PLS-SEM). In this stage, the moderating effect was investigated through a set of measurement items. The method was evaluated on the TripAdvisor data, in which data were collected from 216 customers who have previously booked eco-friendly hotels through the TripAdvisor platform.

2. Material and methods

This study uses a two-stage methodology to address the aim of the article. The proposed framework of the study is presented in Fig. 1. In the first step of our methodology, a hybrid method through the machine learning technique is developed. The method uses customers' online reviews concerning food quality. In addition to food quality criteria, other hotel performance criteria, such as sleep quality, value (cost–benefit), service, location, rooms, and cleanliness, are also used for travelers' preference prediction [8]. To better detect travelers' preferences, cluster analysis is performed through a supervised learning technique. To do so, we use the LVQ technique which is a robust learning technique to discover effective clusters from big datasets. This is done through the classification by LVQ on the datasets which include class labels. In the next step of our methodology, the CART approach is performed on each cluster generated by LVQ. This study uses an ensemble approach of CART for traveler' preference prediction through big social data. In the next stage of our methodology, we develop a hypothetical model to analyze the relationships between the performance criteria and satisfaction through partial least squares structural equation modeling (see Fig. 2). This part includes the stages which occur in the partial least squares structural equation modeling such as evaluation criteria for reflective and formative models and evaluation criteria for a structural model. In the research model, we propose the following hypotheses:

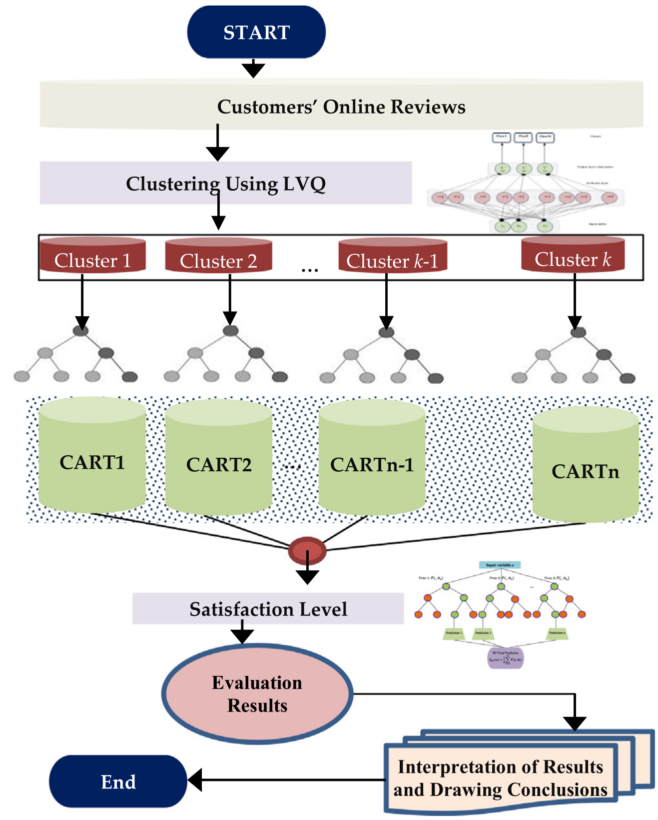

Fig. 1. Research method.

H1: The cleanliness of the hotel has a direct influence on tourists' satisfaction.
H2: The location of the hotel has a direct influence on tourists' satisfaction.
H3: The room of the hotel has a direct influence on tourists' satisfaction.
H4: The presented services have a direct influence on tourists' satisfaction.
H5: Sleep quality has a direct influence on tourists' satisfaction.
H6: The perceived value has a direct influence on tourists' satisfaction.
H7a: The quality of the food has a moderating influence on the link between location and tourists' satisfaction.
H7b: The quality of the food has a moderating influence on the link between presented services and tourists' satisfaction.
H7c: The quality of the food has a moderating influence on the link between value and tourists' satisfaction.

2.1. CART

As a nonparametric modeling and effective learning technique, CART is widely used for regression problems. It has been demonstrated as a computational–statistical algorithm for discovering nonlinear relationships without variable transformations [9]. In CART, recursive binary partitioning is performed to construct the decision trees [10,11]. The decision trees are formed by a collection of rules according to the values of certain input and output variables in the modeling dataset [12]. Usually, the modeling algorithm stops if the user reaches the maximum tree depth or if splits cannot be created any longer [13]. Outliers have less impact on the CART results compared to the other prediction learning techniques. Although CART has shown high accuracy in many prediction tasks, its ensemble learning, Random Forests (RF), which trains several individual decision trees in parallel through bootstrapping approach followed by aggregation can present more robust outcomes in training samples and reducing the noise in the training dataset and exhibit good generalization. The main procedure of CART is shown in Algorithm 1.

2.2. LVQ

Neural Networks (NNs) are divided, according to the learning process, into two kinds: unsupervised and supervised [14]. The difference between unsupervised and supervised learning techniques lies in how the networks are trained through a set of data to recognize and categorize the objects. Teuvo Kohonen in the mid-1980s developed LVQ as a supervised artificial neural network model for statistical classification [15]. In many real-world applications, to enhance the network training, LVQ has been used in combination with the unsupervised ANN such as SOM or Self-Organizing Map. The LVQ is a competitive network [16], which includes a hidden competitive layer and a linear output layer. In the Kohonen layer of LVQ, several clusters can be generated through network training. LVQ uses the Euclidean distance to discover the winner unit [16]. The main procedure of LVQ is shown in Algorithm 2.

Algorithm 1: Random Forest Algorithm

i. Applying bagging sampling method on original training set M to generate q training sets. For each training set, a total of N samples are considered.
ii. Training the models for q training sets to obtain q decision tree models by CART.
iii. For the individual CART's decision tree models, each time the selection of the optimal division attribute of the current node is performed based on the GINI index [12] to generate branch nodes, and accordingly a single CART decision tree is obtained.
iv. The obtained q decision trees by CART are formed into a RF to get the final output. The final output can be obtained by: $\hat{\mathcal{F}}_{RF}(x) = \frac{1}{q}\sum_{l=1}^{q}\hat{\mathcal{F}}(x, \Theta_l)$.

Algorithm 2: LVQ Algorithm

i. Initializing the learning rate α and the codebook vectors W_i.
ii. Selecting an input vector X.
iii. Finding the winner unit closest to X which is based on Euclidean distance of codebook vector W_c with regard to X):
iv. $\|X - W_c\| = \min_k \|X - W_k\|$ i.e. $c = arg \min_k \|X - W_k\|$
v. Modifying of the winner units' weights:
vi. If the classification has been correct i.e. W_c and X belong to the same class:
vii. $W_c(t+1) = W_c(t) + \alpha(t)[X(t) - W_c(t)]$
viii. If the classification has not been correct i.e. W_c and X belong to different classes:
ix. $W_c(t+1) = W_c(t) - \alpha(t)[X(t) - W_c(t)]$
x. Adjusting the learning rate α.
xi. Checking the number of iterations, if the number of iterations exceeds, the iteration ends, otherwise, refer to the second step.

3. Results

The data was collected from TripAdvisor. The data includes six criteria along with an output for travelers' satisfaction. This

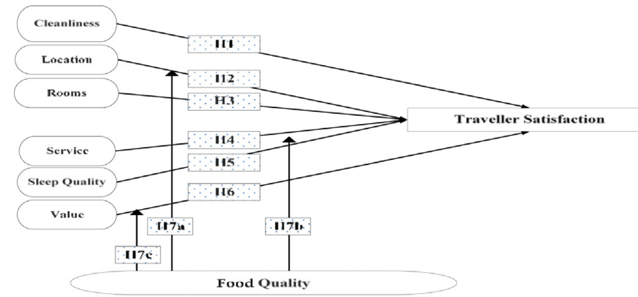

Fig. 2. A research model for SEM-PLS.

is mainly used as overall ratings to show the travelers' overall satisfaction through their ratings on the eco-friendly hotels' performance criteria. We only consider the reviews which include travelers' concerns on the food quality of the destination. Totally, 3218 ratings were collected from January 2021 to June 2021. After data pre-processing, only 2654 ratings were considered for further data analysis. These ratings were used in LVQ for data clustering. In LVQ, we considered 2 clusters for each class of the dataset. We finally apply the CART technique to the LVQ segments. Accordingly, ten clusters were generated, in which the majority of ratings belonged to Segment 2 (LVQ1_2: 599) followed by Segment 4 (LVQ2_2: 542), Segment 5 (LVQ3_1: 333), Segment 1 (LVQ1_1: 255), Segment 6 (LVQ3_2: 192), Segment 10 (LVQ5_2: 188), Segment 9 (LVQ5_1: 174), Segment 8 (LVQ4_2: 172), Segment 7 (104) and Segment 3 (LVQ2_1: 95). In Segment 1, the segment' centroids are 1.568627, 4.552941, 4.392157, 4.337255, 3.274510, and 3.862745 respectively for Rooms, Value, Location, Service, Cleanliness, and Sleep Quality. This shows that Value, Location and Service have been more important for this group of tourists. In Segment 2, the segment's centroids are 3.951586, 4.400668, 4.415693, 4.440735, 3.928214, and 2.913189, indicating that Value, Location, and Service are more important criteria for the tourists of this group. Similarly, the segment's centroids show that Rooms, Sleep Quality, Cleanliness, Location, and Sleep Quality have been the most important criteria for Segments 3–10.

The results were in the form of IF-THEN rules in different segments. In fact, the rules could be easily understood and the relative importance of each criterion could be found from its relationship with customer satisfaction. For example, in Rule 5 of Segment 1 which was IF Food Quality = [Very High] AND Rooms = [Very High] AND Value = [High] AND Location = [High] AND Service = [High] AND Cleanliness = [High] AND Sleep Quality = [Very High] THEN Satisfaction = [Very High], it is found that if Food Quality is in high level, customers' satisfaction is very high. The CART models were trained on training sets and tested on test data in each cluster. The performance of the models was evaluated through the Area Under the Curve (AUC) score. It was found that the performance of CART models measured by AUC scores on test datasets is good in all clusters (AUC between 0.94–0.98). We also compared the results of this study with the Support Vector Machine (SVM) classifier. The results showed that

Table 1
Demographic results of the participants ($N = 216$).

Feature	Item	Frequency	Percentage
Age	18–20	55	25.5
	21–30	101	46.8
	>30	60	27.7
Marital status	Married	155	71.8
	Single	61	28.2
Occupation	Employee	73	33.8
	Employer	52	24.1
	Student	25	11.6
	Retired	57	26.3
	Other	9	4.2
Usage of TripAdvisor	Once	50	23.1
	2–3 Times	100	46.3
	More than 3 times	66	30.6
Mode of travel	Family	111	51.4
	Solo	50	23.1
	Friends	55	25.5

LVQ can provide better AUC compared with the SVM (AUC between 0.91–0.93).

In the second stage of this research, the distribution of the questionnaire was distributed among tourists by utilizing social media portals. We clarified that the participants' responses will only be used for research goals. We managed to collect 216 valid answers, which were kept for further analysis. The questionnaire has three basic parts: (1) a preface to present the goal of the study to participants, (2) participants' demographic data, and (3) the survey's main questions. The data collection procedure was conducted within a period of four months from January 2021 to April 2021. The demographic data are displayed in Table 1.

In order to evaluate the measurement model's reliability and validity, several measures were utilized. Cronbach's alpha is utilized to evaluate the reliability of all the factors, in which the outcomes range from 0.754 to 0.854 (Table 2), which meet the level of the required threshold (0.70) as indicated by Hair Jr, et al. [17]. We utilized the factor loadings, composite reliability, and Average Variance Extracted (AVE) of all factors to check the convergent validity assessment. The composite reliabilities (CR) of the factors met the required threshold (0.8), demonstrating that the scales have acceptable internal consistency reliability [17]. The AVE for each factor was above 0.5 indicating an acceptable outcome of convergent validity of the scales [18].

Besides, we checked the Discriminant Validity (DV) by referring to the square root of each factor's AVE and comparing it with its coefficients of correlation with other factors. If the square root of a factor's AVE is the largest in comparison with its correlation coefficients with other factors in this model, the factor has acceptable DV [18]. All factors in our model meet this condition, indicating adequate DV outcome. The DV of the factors was also inspected by analyzing the loadings and cross-loadings of items. A cross-loading table must present measurement indicators that load highly on their theoretically assigned factors and not high on other variables [19], which was confirmed in the results of the analysis.

Table 2
Constructs' reliability and validity.

Constructs	CR	Cronbach's Alpha	AVE
CLN	0.899	0.85	0.691
FQ	0881	0.832	0.597
LOC	0.864	0.763	0.68
RO	0.867	0.769	0.686
SAT	0.896	0.854	0.634
SERV	0.846	0.762	0.579
SQ	0.85	0.766	0.587
VAL	0.844	0.754	0.576

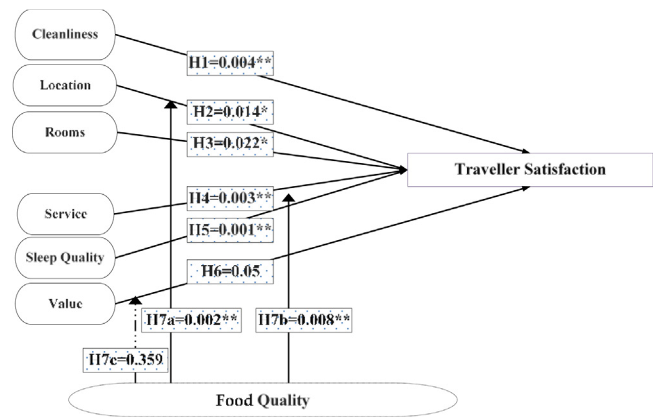

Fig. 3. Final research model.

In order to test hypothesized links, we referred to the outcomes resulted from paths' assessment (β, t-value, and p-value) for accepting or rejecting a particular path. The strength of the links among the independent and dependent factors can be evaluated with path loadings [17], in which the presented model could explain 76.9% of the variance in satisfaction. The outcomes of the hypotheses test are presented in Table 3. All direct hypotheses are positive and statistically significant in the expected direction. As shown in Table 3, the results shown that eight of the suggested hypotheses were significant except one of the hypotheses (H7c) that was not significant. According to obtained results, cleanliness, location, rooms, service, sleep quality and value had significant effects on satisfaction. Regarding the moderating effects of food quality on each of location, service, and value and satisfaction, we found that food quality moderates the relationship between each of location and satisfaction and service and satisfaction. But food quality did not moderate the influence on value and satisfaction. The final research model is presented in Fig. 3.

4. Discussion

This article aimed to explore tourists' perceptions towards several criteria that can influence tourists' satisfaction, including cleanliness, location, rooms, service, sleep quality, and value. These criteria were explored based on the comments posted on the TripAdvisor portal during the COVID-19 crisis and using the text mining approach. The data included six criteria along with an output of travelers' satisfaction, which

Table 3
Outcomes of paths evaluation.

Hypothesis	Path coefficient	T statistics	P values	Results
H1: Cleanliness -> Satisfaction	0.169	2.887	0.004**	Accepted
H2: Location -> Satisfaction	0.15	2.475	0.014*	Accepted
H3: Rooms -> Satisfaction	0.148	2.294	0.022*	Accepted
H4: Service -> Satisfaction	0.192	3.018	0.003**	Accepted
H5: Sleep Quality -> Satisfaction	0.226	3.491	0.001**	Accepted
H6: Value -> Satisfaction	0.162	1.964	0.05	Accepted
H7a: Location × Food quality -> Satisfaction	0.132	3.146	0.002**	Accepted
H7b: Service × Food quality -> Satisfaction	0.13	2.677	0.008**	Accepted
H7c: Value × Food quality -> Satisfaction	−0.053	0.917	0.359	Not-Accepted

*Significance level = < 0.05.
**Significance level = < 0.01.

are: "Location", "Sleep Quality, "Value", "Service", "Rooms", and "Cleanliness". The result indicated the significant role of these criteria on travelers' satisfaction during the COVID-19 outbreak. This outcome is supported in the 10 clusters derived from travelers' ratings of eco-friendly hotels. These results are consistent with previous literature [20]. The textual reviews also indicated that food quality is an important factor in travelers' assessment of a particular destination. Even though travelers' preferences have been investigated in previous literature [21], this issue is not well explored, particularly in the context of a worldwide health crisis such as COVID-19.

To confirm the findings of the first step of the proposed approach, we used the six factors to design a research model to explore travelers' satisfaction. The research model was evaluated using a survey-based approach based on a questionnaire that we distributed among travelers who had previous experiences with the TripAdvisor portal. In the research model, we hypothesized that each of "Location", "Sleep Quality, "Value", "Service", "Rooms", and "Cleanliness" has a positive influence on travelers' satisfaction. Additionally, we hypothesized that the quality of the food has a moderating influence on the relationship between "Value", "Location", and "Service" of the hotel and tourists' satisfaction. The outcomes of the structural model analysis highlighted the significant role of the hypothesized factors on consumers' satisfaction. Hence, the outcomes of the research showed that the influences of "Services", "Cleanliness", and "Sleep quality" on consumer satisfaction are the highest among main research paths (H1–H6). This outcome has been indicated in previous literature by many studies. Cleanliness was indicated as a significant driver of travelers' satisfaction in a study by Bhatnagar and Dheeraj [22]. On the other hand, the influence of the location of the hotel on guest satisfaction was also supported [23]. Service quality was also indicated as an important driver of customers' satisfaction in previous literature [24]. Besides, the moderating influence of food quality on the relationship between "Location" and "Service" of the hotel and tourists' satisfaction was supported. As the research outcomes presented, food quality is among the essential drivers of consumers' satisfaction. Following the COVID-19 crisis, food and beverage aspects should not be evaluated in a traditional manner, as new views related to COVID-19 preventive measures gained tourists' concerns. Particularly, during the COVID-19 crisis, tourists are putting new aspects related to hygiene and safety at the top of their priorities. Even though the presentation of the food in a way that maintains social distancing measures is considered as a significant aspect. The investigation of the quality of the food during this emerging crisis is an important issue, which is considered a new topic that should be explored in future research broadly [25].

5. Conclusion, limitations, and future work

The goal of this study was to explore tourists' preferences and satisfaction through the data crawled from TripAdvisor. By integrating machine learning and survey-based methods, the electronic ratings and reviews were analyzed. Following that, the hypothesized research model was presented and evaluated. We developed a novel approach that entails machine learning to rank tourists' choices for the eco-friendly hotels' characteristics and cluster tourists based on their levels of satisfaction.

Exploring tourists' experiences based on traditional approaches is no longer efficient in the existing and changeable hotel market, particularly with the current COVID-19 crisis context. Emerging technology allows the usage of valuable information through electronic reviews to present novel experiences to tourists. Hence, this study tries to utilize electronic reviews to assess travelers' choices and levels of satisfaction. There is a need to understand what triggers consumer satisfaction or dissatisfaction with presented services [26]. Besides, future researches should also investigate the influence of online comments on tourists' general impression of the quality of the presented services. Future researches may explore how online comments and ratings can be joined together to explore consumers' levels of satisfaction more precisely. The influence of the moderating impact of food quality on travelers satisfaction presents future research directions, as other the impact of other variables on travelers' satisfaction should be explored.

6. Research implications

Considering methodological implications, this study utilized online reviews in analyzing tourists' destination choices and satisfaction. The basic contribution of this research is to present a novel method based on two stages in the tourism context as follows: In the first stage of our approach, a hybrid method was advanced using clustering and prediction learning techniques. To perform data clustering, LVQ was performed to locate the most efficient sets of tourists with related choices. We then performed CART on each segment of LVQ for the prediction of travelers' choices through a set of eco-friendly hotels' performance criteria. Second, we developed a hypothetical model to analyze the relationships between the performance criteria and satisfaction through PLS-SEM. The method was evaluated based on two sources of data: the TripAdvisor crawled data and the data collected from 216 customers who had previously booked eco-friendly hotels through the TripAdvisor platform.

This research has also some significant implications for tourism management and advancement and introduces directions to exhibit how emerging conditions can influence the market in the future. The practical implications are observable, as this research emphasized that TripAdvisor is a significant data source, particularly during this emerging situation, that aids hotel managers to improve their visibility to tourists and form robust links with tourists through presented comments and ratings. At the same time, understanding various folds in tourists' choices aids businesses to advance various policies for each cluster, thus presenting a competitive advantage. Hence, decision-makers should optimize their resources for tourism advancement and develop integrated business policies. The outcomes of this research not only define the various characteristics that impact tourists' choices but also, classify the features for each cluster. The research outcomes aid researchers to define the main features of services and rank them based on their importance. Accordingly, to advance suitable business policies, hotel managers should consider the variations between clusters. Another significant implication of this research is aiding decision-makers to follow suitable measures to address the flaws in services to present a more satisfactory experience. Decision-makers can concentrate on the low-rated items for each cluster and address related issues, which can promote travelers to post more positive opinions on the TripAdvisor portal and decrease negative opinions by dissatisfied tourists. The outcomes of this study will aid decision-makers to face bad experiences to enhance tourists' satisfaction. Besides, the outcomes of this research can aid managers to enhance their brand reputation by enhancing their services. These main features are significant as they indicate the basic factors that trigger a traveler to choose a specific destination. Even though it is significant for managers to concentrate on appropriate advertising policies, it is also significant to focus on understanding and keeping the quality aspects of eco-friendly hotels.

CRediT authorship contribution statement

Mehrbakhsh Nilashi: Supervision, Conceptualization, Methodology, Investigation, Software, Data curation, Formal analysis, Writing – original draft, Writing – review & editing, Validation. **Rabab Ali Abumalloh:** Conceptualization, Methodology, Investigation, Software, Data curation, Formal analysis, Writing – original draft, Writing – review & editing, Validation. **Ahmed Almulihi:** Methodology, Writing – original draft, Writing – review & editing, Validation. **Mesfer Alrizq:** Writing – review & editing, Validation. **Abdullah Alghamdi:** Investigation, Writing – review & editing, Visualization. **Muhammed Yousoof Ismail:** Investigation, Writing – review & editing, Visualization. **Abul Bashar:** Investigation, Writing – review & editing, Visualization. **Waleed Abdu Zogaan:** Investigation, Writing - review & editing, Visualization. **Shahla Asadi:** Investigation, Writing - review & editing, Visualization.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

This research was made possible by a genours fund from the deanship of scientific research at Taif University, Taif, Saudi Arabia under Taif University researchers supporting project No. TURSP-2020/344.

References

[1] A. Ahani, et al., Revealing customers' satisfaction and preferences through online review analysis: The case of Canary Islands hotels, J. Retailing Consumer Serv. 51 (2019) 331–343.
[2] T. Zhang, J. Chen, B. Hu, Authenticity, quality, and loyalty: Local food and sustainable tourism experience, Sustainability 11 (12) (2019) 3437.
[3] S. Taylor, R.B. DiPietro, K.K.F. So, Increasing experiential value and relationship quality: An investigation of pop-up dining experiences, Int. J. Hosp. Manag. 74 (May) (2018) 45–56.
[4] M. Nilashi, et al., Big social data and customer decision making in vegetarian restaurants: A combined machine learning method, J. Retailing Consumer Serv. 62 (2021) 102630.
[5] G. Bello-Orgaz, J.J. Jung, D. Camacho, Social big data: Recent achievements and new challenges, Inf. Fusion 28 (2016) 45–59.
[6] J.Z. Zhang, P.R. Srivastava, D. Sharma, P. Eachempati, Big data analytics and machine learning: A retrospective overview and bibliometric analysis, Expert Syst. Appl. (2021) 115561.
[7] S.-E. Kim, K.Y. Lee, S.I. Shin, S.-B. Yang, Effects of tourism information quality in social media on destination image formation: The case of Sina Weibo, Inf. Manag. 54 (6) (2017) 687–702.
[8] E. Yadegaridehkordi, et al., Customers segmentation in eco-friendly hotels using multi-criteria and machine learning techniques, Technol. Soc. 65 (2021) 101528.
[9] L. Breiman, J. Friedman, C.J. Stone, R.A. Olshen, Classification and Regression Trees, CRC Press, 1984.
[10] E.A. Antipov, E.B. Pokryshevskaya, Mass appraisal of residential apartments: An application of random forest for valuation and a CART-based approach for model diagnostics, Expert Syst. Appl. 39 (2) (2012) 1772–1778.
[11] H. Bittencourt, R. Clarke, Feature selection by using classification and regression trees (CART), in: The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, 2004.
[12] R. Timofeev, Classification and Regression Trees (CART) Theory and Applications, Humboldt University, Berlin, 2004, pp. 1–40.
[13] W.Y. Loh, Fifty years of classification and regression trees, Internat. Statist. Rev. 82 (3) (2014) 329–348.
[14] M. Nilashi, et al., Coronary heart disease diagnosis through self-organizing map and fuzzy support vector machine with incremental updates, Int. J. Fuzzy Syst. (2020) 1–13.

[15] T. Kohonen, The self-organizing map, Proc. IEEE 78 (9) (1990) 1464–1480.
[16] R. Li, H. Wang, Y. Cui, X. Huang, Solar flare forecasting using learning vector quantity and unsupervised clustering techniques, Sci. China Phys., Mech. Astron. 54 (8) (2011) 1546–1552.
[17] J.F. Hair Jr., G.T.M. Hult, C. Ringle, M. Sarstedt, A Primer on Partial Least Squares Structural Equation Modeling (PLS-SEM), Sage publications, 2016.
[18] C. Fornell, D.F. Larcker, Structural Equation Models with Unobservable Variables and Measurement Error: Algebra and Statistics, SAGE Publications Sage CA, Los Angeles, CA, 1981.
[19] D. Barclay, C. Higgins, R. Thompson, The partial least squares (PLS) approach to casual modeling: personal computer adoption ans use as an illustration, 1995.
[20] Y. Zhao, X. Xu, M. Wang, Predicting overall customer satisfaction: Big data evidence from hotel online textual reviews, Int. J. Hosp. Manag. 76 (2019) 111–121.
[21] A. Ahani, et al., Revealing customers' satisfaction and preferences through online review analysis: The case of Canary Islands hotels, J. Retailing Consumer Serv. 51 (May) (2019) 331–343.
[22] E. Bhatnagar, N. Dheeraj, Impact of housekeeping services and practices on customer satisfaction and repeat business, Prabandhan: Indian J. Manag. 12 (8) (2019) 46–57.
[23] J. Dong, H. Li, X. Zhang, Classification of customer satisfaction attributes: An application of online hotel review analysis, in: Conference on e-Business, e-Services and e-Society, Springer, 2014, pp. 238–250.
[24] R. Nunkoo, V. Teeroovengadum, P. Thomas, L. Leonard, Integrating service quality as a second-order factor in a customer satisfaction and loyalty model, Int. J. Contemp. Hosp. Manag. 29 (2017) 2978–3005.
[25] K. Byrd, E. Her, A. Fan, B. Almanza, Y. Liu, S. Leitch, Restaurants and COVID-19: What are consumers' risk perceptions about restaurant food and its packaging during the pandemic? Int. J. Hosp. Manag. 94 (2021) 102821.
[26] G. Tontini, G. dos Santos Bento, T.C. Milbratz, B.K. Volles, D. Ferrari, Exploring the nonlinear impact of critical incidents on customers' general evaluation of hospitality services, Int. J. Hosp. Manag. 66 (2017) 106–116.