AJMT

Asia-Pacific Journal of Management and Technology

Online ISSN: 2652-6840

journal homepage www.ajmt.org

Original Article

A New Approach in Marketing Research: Identifying the Customer Expected Value through Machine Learning and Big Data Analysis in the Tourism Industry

Ali Ghasemian Sahebi, Rahil Kordheydari, Mohammad Aghaei*

Department of Management, Faculty of Management and Economics, Tarbiat Modares University, Tehran, Iran.

*Corespondance Email: Alighasemiansahebi@modares.ac.ir.; K.rahil@modares.ac.ir. ; M_aghaei@modares.ac.ir.**

Abstract

With the advent of the digital age and its entry into the business environment, marketing trends have undergone several changes, such as the expansion of data-driven marketing and the digitalization of the business environment that requires further investigation. As a result, the aim of this study is to present a novel approach in marketing research and identify value components among the large volume of customer feedback in virtual networks using machine learning, big data analysis, and a predictive marketing strategy. This research is applied in terms of purpose, qualitative-quantitative (mixed) in terms of method and descriptive-survey in terms of data collection that has used an inductive approach. For this purpose, the Iranian tourism industry and tourist areas of Tehran province was selected as a case study and 9325 comments from customers about hotels used in their travels to Tehran were collected from the Internet between the summer of 2020 to winter 2021 and using "data clustering" and "Association Rules Extraction" methods, the value components were extracted and RapidMiner software and Python programming language were used to perform data mining, text mining and big data analysis processes. In summary, the findings demonstrate that by employing big data analysis and machine learning, the process of "marketing research" can be performed with greater speed, accuracy, and extensiveness, as well as at a lower relative cost. The findings and suggestions of this study are used to inform marketing researchers, as well as to raise the awareness of managers, especially in hotel industry.

Keywords: *Marketing Research; Customer Expected Value; Big Data; Machine Learning; Tourism Industry*

Introduction

Today the dynamic conditions of marketing have caused the organizations to use the new technologies because marketing is experiencing deep reformations and developments at the present time to the extent that most economists believe that a new era namely the digital age is coming forth now (Gupta *et al*., 2017). Today, companies face challenges such as making quick decisions to improve productivity in the face of the big data problem; Because most of the production systems suffer from a lack of analytical smart facilities, they are not apt to manage their huge amounts of data (Lee, Kim & Lee, 2017). The new age describes a world that empowers people to move between digital and offline reality to live and manage it using

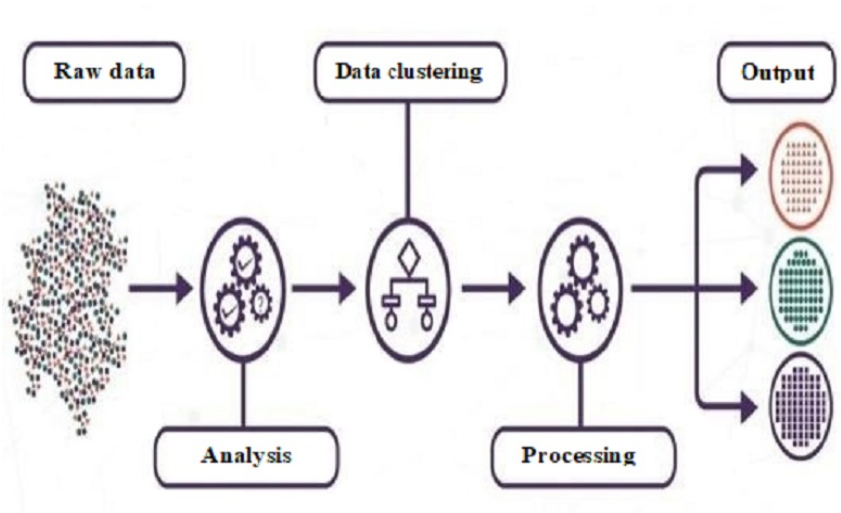

Figure 2: ***Schematic Display of Clustering (Separation) of Text Document Topics***

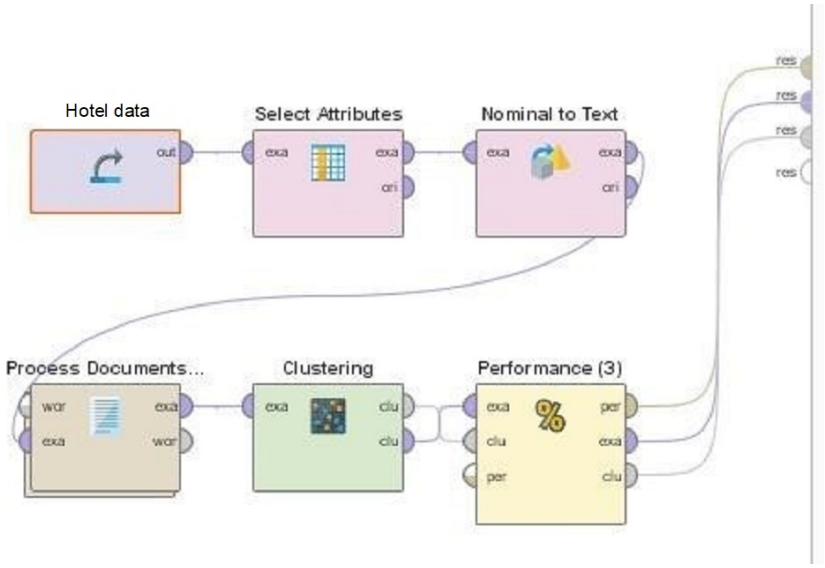

Figure 3: ***Clustering of Comments by Rapidminer Software***

In this study, to determine the number of clusters and clustering validation, the "Dunn index" method has been used due to better performance than other methods in big data (Kassambara, 2017). The "Dunn" index is presented with the aim of determining the amount of compactness (distance of members within the cluster with each other) and separation (distance of clusters from each other) and calculates the amount of compactness and separation with two criteria of "cluster distance" and "diameter". This criterion is defined by Equation 1, in which d (C_i, C_j) is the distance between clusters i and j and diam (C_l) represents the diameter of cluster l (Faezi Rad & Pouya, 2016).

Equation 1.

$$DI = \min_{i+1\le j\le K}\left\{\min_{i+1\le j\le K}\left\{\frac{d(C_i.C_j)}{\max\limits_{1\le i\le K} diam\,(C_i)}\right\}\right\}$$

Association Rules

Association rules are an approach to data mining that is used to discover repetitive patterns, correlations, relationships, and causal structures in a variety of databases and to analyze and predict user behavior. By discovering the connections between data sets, these rules show the conditions that occur frequently in a data set and explain the presence of some items (entities) based on other items (entities) (Shalini & Lal, 2016). The purpose of association rules in text mining is to discover the relationships between words and to identify the rules that require the presence of a set of words in the text and the existence of other words.

In a database, if $I = \{I_1, I_2, \ldots, I_m\}$ is a set of items, D is a set of transaction databases, each of which has an identifier, as well as a set of items such as T ⊆ I.

If A is a set of items, a transaction like T contains A, if we have: $A \subseteq T$. An association law is expressed as $A \Rightarrow B$ in which $A \subset I$, $B \subset I$, $A \cap B = \varphi$, $B \neq \varphi$, $A \neq \varphi$. An association rule states in its general form that if event *A* occurs, then event *B* also occurs. Symbol $A \Rightarrow B$ is used to represent an association rule, and *A* is called the introduction of the rule and *B* is the result of the rule. The source of the production of association rules are also times that have been repeated enough that they are called "collection of recurring items" (Han, Pei & Kamber, 2011). Each rule of $A \Rightarrow B$ has an indicator called "Support" that indicates the repetition rate of the concurrency. The Support value of rule $A \Rightarrow B$ is calculated using Equation 2. The numerator is equal to the number of concurrences *A* and *B*, and the denominator refers to the total number of events.

Equation 2. $Support\ (A \Rightarrow B) = N(A\&B)/N$

Once the rules have been generated with a minimum value of Support, the probability of the rules being correct can also be calculated. This indicator, called the "Reliability", indicates what percentage of the time that Event *A* occurred, Event *B* also occurred. The Reliability of rule of $A \Rightarrow B$ is calculated using Equation 3 (Lee, Kim & Lee, 2017).

Equation 3. $Confidence\ (A \Rightarrow B) = N(A\&B)/N(A)$

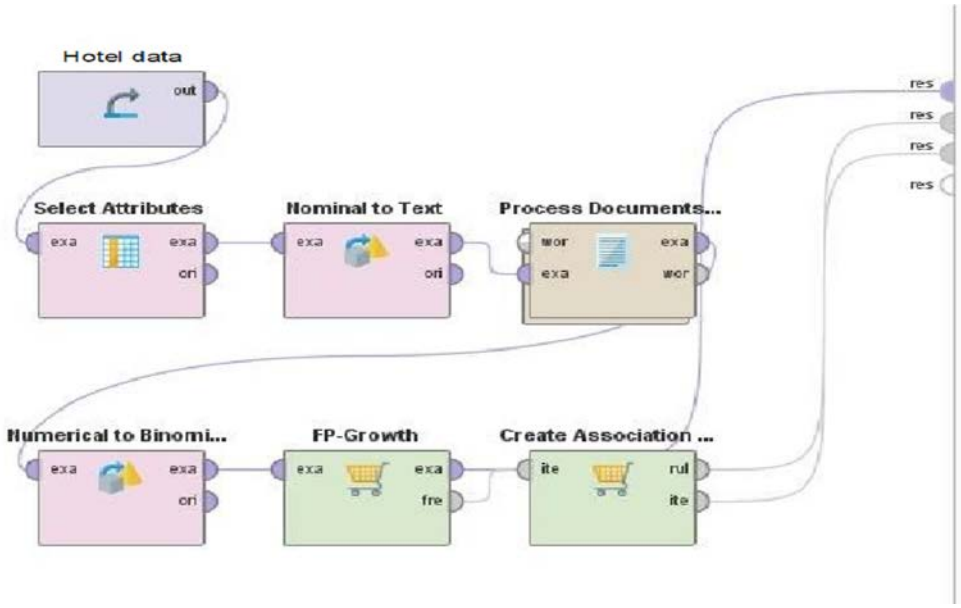

Figure 4: ***Extraction of Association Rules by Rapidminer***

Apriori and FP-Growth are two of the most important and practical methods of association rule extraction algorithms. In this study, after examining these two methods and comparing

their efficiency and speed, the FP-Growth method was used to extract the association rules (Esmailipour, 2016). Figure 4 shows how RapidMiner performs the process of extracting association rules in this research.

Results and Discussion:

In this section, the findings of the text mining of customers' comments about the hotel in Iran's tourism destinations are presented in the form of two parts: "Clustering of comments" and "Extraction of association rules".

Findings of the "clustering" process

In this section, to extract the value components expected by customers, the collected comments are clustered using the k-means algorithm, and to identify the topic or feature mentioned in each cluster, the most frequently worded words in that cluster are identified. The naming of each cluster is done according to the logical relationship between the words in each category of topics that have the most repetition (probability) (Guo, Barnes, & Jia, 2017).

The results of clustering tourists' comments along with eight "words" with the most repetition (probability) in each cluster to identify the topic or feature desired by the customer in each cluster are shown in Table 1. By applying the clustering algorithm to the tourists' comments, 20 important clusters were identified. Examining each cluster and the words that have the most repetition in that cluster, the topic of discussion in that cluster is determined. Of the 20 topics or features identified in this collection, after sorting the clusters, the first 11 topics refer to the positive components of value and the next 9 topics refer to the negative components of value from the tourists' point of view in using the hotel. It should be noted that the components identified in Table 1 are not presented in order of priority and are only located in different clusters, which include two parts of positive and negative components.

Table 1: ***Topics Extracted from the Clustering of Comments and Eight Words with a Higher Probability in Each Cluster***

Topics from Clustering Comments		**Words" with the most repetition (probability) in each cluster"**							
Cluster 1	Other comments about the hotel	Comments	Friends	site	Hotel	room	Selection	Chic	Star
Cluster 2	Cleanliness and general hygiene of the hotel	room	Hotel	Service	Antifungal	Breakfast	Clean	Health	Restaurant
Cluster 3	Personnel behavior	Hotel	Thanks	Mr	Lady	Staff	Behavior	Management	Good
Cluster 4	Proximity to shopping and entertainment centers	Hotel	on foot	Market	Beach	Center	Minutes	Close	Buy
Cluster 5	Proper bathroom and toilet	Bathroom	Service	Towel	Clean	Shampoo	Sanitary	WC	Good
Cluster 6	Room cleanliness (bedding, tablecloths, etc.)	room	Good	Clean	Arranged	bed	White	Bed linen	Table
Cluster 7	The beauty of the space and the interior layout	Hotel	Excellent	Beautiful	room	Restaurant	Space	Regular	Chic
Cluster 8	Regional facilities for recreation	Bar	View	Rial	Restaurant	Ship	Coffee Shop	Sea	Kish

Cluster 9	Quality and variety of breakfast	Cheese	Good	Egg	Jam	Butter	Breakfast	Lentil feed	Sausage
Cluster 10	Transfer services and delivery room	Room	Clock	Hotel	Fast	Delivery	Transfer	baggage	Airport
Cluster 11	Ancillary facilities	room	Phone	Refrigerator	Tea	Herbal Tea	Internet	Speed	Coffee
Cluster 12	Poor hygiene hotel	room	Dirty	Service	Bathroom	WC	Spoiled	Antifungal	Towel
Cluster 13	Power outages	room	Electricity	Rooms	darkness	Cut	lamp	Region	Electrical
Cluster 14	The hotel star does not fit with its quality	Hotel	room	Star	Behavior	Quality	Service	Bag	Residence
Cluster 15	flight delay	Clock	Late	Rials	Delay	Uncertainty	Airport	Delivery	Time
Cluster 16	Existence of annoying insects in the room	Bedbug	Peace	Morning	Insect	Annoying	Beetle	Night	Ant
Cluster 17	Low quality sanitary ware	Towel	Quality	Soap	Toothbrush	Shampoo	Bad	Female	Dirty
Cluster 18	Annoying noise coming from outside	room	sound	Floor	the outside	Tall	window	Voice	Annoying
Cluster 19	Expensive entertainment outside the hotel	Rials	Bar	Price	Ship	Restaurant	Ticket	Kish	Expensive
Cluster 20	High price	Rial	Cost	agency	room	Flight	Tour	Money	Expensive

Findings of the "Association Rules" process

The purpose of identifying association rules is to explore the relationships between words in a large set of texts (customer feedback). In this method, hidden knowledge is extracted from the analysis of words or phrases that customers have used together and abundantly in their comments. In Table 2, the rules are extracted, and two indicators of Support and Reliability are presented.

Table 2: ***Results of Extracting Association Rules and the Value of Two Indicators of Support and Reliability***

N	Introduction (Law)	Result	Confidence level	Backup amount	Law
1	Health	Accommodation, hotel	0.9626	0.2317	(Health) → (Accommodation, hotel)
2	Behavior	Accommodation, hotel	0.9650	0.3901	(Behavior) → (Accommodation, hotel)
3	Service	Accommodation, hotel	0.9679	0.2221	(Service) → (Accommodation, hotel)
4	room	Accommodation, hotel	0.9686	0.5174	(Room) → (Accommodation, hotel)
5	Restaurant	Accommodation, hotel	0.9689	0.3008	(Restaurant) → (Accommodation, hotel)
6	Breakfast	Accommodation, hotel	0.9734	0.4023	(Breakfast) → (Accommodation, hotel)
7	Star	Accommodation, hotel	0.9847	0.2726	(Star) → (Accommodation, hotel)

Association rules are defined as "if ⟶ then". In other words, by using this method, hidden relationships and dependencies in the big data are revealed and the discovery of important and useful rules facilitates the decision-making process of managers by providing appropriate information. For example, according to the findings in Table 2, if the hotel is hygienic, then tourists will choose to stay in that hotel.

Discussion:

Today, every business operates on the basis of a "business model". One of the most important components of a business model is "value proposition to the customer", which represents the product or service that the company is committed to delivering to customers (Sheehan & Bruni-Bossio, 2015). On the other hand, the most important component that can provide optimal value proposition to customers and create a competitive advantage for the organization, is to recognize the values expected by customers and strive to achieve it. In general, it can be said that recognizing and providing the right "value" to the customer is one of the most important challenges for the success and profitability of any business. This issue, with the introduction of the business environment into the digital age and the change of tastes and expectations of customers and the intensity of competition, has become more important and needs further investigation. With the growth of the tourism industry, the issue of the values expected by customers to choose a destination and a hotel, has become very important, which has faced challenges in this area. Also, the prevalence of Covid-19 has added to the evolution of customer expectations and necessitated research in this area. Past research that has linked marketing research to new technologies in the tourism industry is very limited. Tourism and hoteling industry in Iran, especially in Tehran province, despite its high potential, has low growth; therefore, due to the lack of research in this field and considering the new technologies and approaches that have entered the field of marketing to assess the value from the perspective of customers, this study was done with the aim of identifying the values expected by customers using new marketing research approaches (including machine learning and big data analysis) in the tourism industry of Iran and the tourist areas of Tehran.

The results of this study has been obtained through the two methods of machine learning (including clustering of comments and extraction of association rules), and big data analysis to explore features or value components commensurate with the expectations of customers (tourists), using the analysis of their online comments on the expected values of hotels in Iranian tourist destinations through RapidMiner software and Python programming language. The results and suggestions of this study is applicable in informing marketing researchers with how to use new approaches in this field to better understand customers and provide optimal services to them and also in increasing awareness of managers, especially in hotel industry about the expected benefits of customers and how to provide optimal services to gain competitive advantage.

The results obtained using the first method (data clustering) show 20 important components (subject or characteristics) in the perceived values of tourists, which include 11 positive components of value and 9 negative components of value. Positive features or components of perceived value of tourists are: 1) Opinions of others about the hotel, 2) Cleanliness and general hygiene of the hotel, 3) Good treatment of hotel staff, 4) Proximity of the hotel to shopping and entertainment centers, 5) Proper and clean bathroom and toilet , 6) cleanliness of hotel rooms (bedding, towels, tablecloths), 7) beauty of the hotel space and interior layout, 8) regional facilities of the hotel for entertainment, 9) quality and variety of breakfast, 10) transfer service and room delivery and 11) Additional facilities such as telephone, refrigerator, high speed internet and so on.

Also, the negative characteristics or components of the perceived value of tourists are: 1) lack of hygiene and not cleaning the room, etc., 2) power outages in the rooms, 3) inadequacy of the hotel star with its quality, 4) delay in transportation, 5 Existence of annoying insects in the room, 6) Low quality of personal hygiene items such as towels, shampoo, etc., 7) Intrusion of annoying noises from outside into the room, 8) High cost of

entertainment outside the hotel and 9) High and inappropriate price of the hotel.

The results obtained using the second method (association rules) include 7 important rules that have been obtained by extracting hidden knowledge from the relationships between the expressions used in tourists' comments. The findings of this section show that 7 issues are more important for customers about staying in a hotel, which are: 1) general health of the hotel, 2) attitude of hotel staff, 3) how and quality of hotel service to tourists, 4) quality and conditions of rooms, 5) Quality and conditions of the hotel restaurant, 6) Quality and variety of breakfast and 7) Number of hotel stars. The results of extracting association rules provide the opportunity for service providers to be able to use the opinions of their customers and attract more satisfaction of customers.

Finally, the use of machine learning in the analysis of customer data showed that this method can analyze very quickly, which is a very important component in today's competitive world. Also, due to the lack of human intervention in the processes, the accuracy of the analysis has been very high. Extensibility is another advantage of this method, because it is possible to analyze big data from this method, if humans are not capable of this level of analysis. If this method is used extensively and in large volumes, the cost of analysis will be reduced, which is a very vital factor for many organizations.

The results of the study on the "importance of customer service quality" are consistent with the research of Sánchez-Franco, Navarro-García and Rondán-Cataluña (2019). The results of Chong *et al*.'s (2018) research show the importance of the opinions discussed on the Internet about the hotel and indicate that most travelers consider the opinions of previous travelers as a source of information for decision making. The results of the present study also confirm this and one of the positive elements obtained in the clustering section is the review of previous travelers' comments about the hotel. Also, the results of this research are in some way in line with the results of (Chong *et al*., 2018; Hair Jr & Sarstedt, 2021; Amirkhani and Mottaqi, 2018; Kauffmann *et al*., 2020; Soltani Zanozi, 2019; Krafft, Sajtos & Haenlein, 2020; Ma & Sun, 2020; Tukiran, Tan & Sunaryo, 2021; Liu, 2021; Hsu *et al*., 2021; Kumar, Venkatesh & Rahman, 2021). Regarding the importance of using the opinions of previous tourists in cyberspace and using the potential capabilities of data analysis in this field, it should be said that the research of Raguseo, Neirotti and Paolucci (2017) and Sánchez-Franco, Navarro-García and Rondán-Cataluña (2019) are consistent with the approach of the present study. The research of Gil-Soto et al. (2019) shows that customers do not have a complete understanding of the level of environmental obligations of the hotel, which was not observed in the results of this study and the analysis of Iranian tourists.

Conclusion

According to the results of the study, it is suggested to marketing managers, especially in the field of tourism and hotel management to pay high attention to the components identified in this study as key indicators of customer satisfaction and gaining a competitive advantage and to use all the forces and capital of their organization in order to realize the value-creating components and prevent the barriers to value creation. Managers and entrepreneurs can use the big data to check the importance and validity of their new ideas and use it to identify and predict possible customer feedback. Also, due to value being relative from the customers' point of view, organizations should measure and present the values expected by their business customers, and more specifically, the desired values of their target market, and avoid looking at customer values in general. Due to the rapid changes in customer demands and expectations, these organizations should always and periodically evaluate these values in their target market to avoid changing customer expectations. It is recommended that they use social media data to validate their ideas, attitudes, and analyzes to identify existing weaknesses and provide new services. Business managers are advised to focus on the most important value dimensions that can be developed by examining their strengths and weaknesses, if they do not have the necessary strength and capital to achieve all the values expected by customers and to create a competitive advantage for themselves in that field. In other words, to gain a corner of the

market by focusing on that segment. Organizational marketing and R&D researchers are also encouraged to make extensive use of artificial intelligence processes, machine learning approaches, and big data analysis in their research and customer-related fields; Also, they should always be up-to-date and know the latest technologies and approaches, learn and use applied approaches in this field so with higher accuracy, faster examinations, a bigger scope of research, reduced costs, they can present optimal goods and service in accordance with the demand and expectations and take steps to improve the performance of the marketing unit and the organization as a whole.

The present study is based on the opinions of customers in the Iranian tourism industry and tourist areas of Tehran province and researchers are recommended to conduct this research globally and compare the results. Also, it is suggested that by combining classical methods of analyzing and predicting consumer behavior with new methods of data science such as machine learning and deep learning, they use the potential capabilities of big data analysis in market research and other branches related to business management.

Regarding the limitations of the research, we can mention the use of non-supervisory method for clustering comments in the research method. To solve this problem, researchers are suggested to process the text and compare the results by combining two methods of supervised and unsupervised. Also, the limitation of the statistical population due to the relative lack of comments in the field of Iranian tourism on the Internet, creates the possibility of sampling errors.

Acknowledgement:

The authors are thankful to the tourists of Iran and Tehran province, for collaborating with this study, for data collection.

Funding:

The author(s) received no financial support for the research, authorship, and/or publication of this article.

Conflict of interest:

The authors have no conflicts of interest to declare.

References

Aaker, D. A. (2005). *Marketing Research: Marketing Research: The Pacific Rim Edition*. John Wiley & Sons.

Aghaei, M., Ghasemian Sahebi, A., & Kordheydari, R. (2020). The effect of COVID-19 on marketing innovations and corporate social responsibility (case study: active companies in food industry). *International Journal on Customer Relations*, *8*(2), 15-26.

Aghaei, M., Sahebi, A. G., & Kordheydari, R. (2021a). The Effect of Covid-19 on Sustainable Consumption Behavior in Chain Stores. *Asia-Pacific Journal of Management and Technology (AJMT)*, *2*(2), 35-51.

Aghaei, M., Sahebi, A. G., & Kordheydari, R. (2021b). Design and explain the pattern of destination brand value creation strategies in Iranian tourism industry. *Journal of Tourism Quarterly*, *3*(2), 68-97.

Ahmad, A. K., Jafar, A., & Aljoumaa, K. (2019). Customer churn prediction in telecom using machine learning in big data platform. *Journal of Big Data*, *6*(1), 1-24.

Amado, A., Cortez, P., Rita, P., & Moro, S. (2018). Research trends on Big Data in Marketing: A text mining and topic modeling based literature analysis. *European Research on Management and Business Economics*, *24*(1), 1-7.

Amirkhani, A., & Mottaqi, V. (2018). Presentation of product production forecasting system on big data using machine learning, *National Conference on New Research in Electrical Computer Engineering, Information Technology*, Mobarakeh, Isfahan, Iran. https://civilica.com/doc/827992/

An, M.A., & Han, S.L. (2020). Effects of experiential motivation and customer engagement on customer value creation: Analysis of psychological process in the experience-based retail environment. *Journal of Business Research*, *120*, 389-397.

Beam, A. L., & Kohane, I. S. (2018). Big data and machine learning in health care. *Jama*, *319*(13), 1317-1318.

Chong, A. Y. L., Khong, K. W., Ma, T., McCabe, S., & Wang, Y. (2018). Analyzing key influences of tourists' acceptance of online reviews in travel decisions. *Internet Research*.

Czinkota, M. R., Kotabe, M., Vrontis, D., & Shams, S. M. (2021). Marketing Research and

Information. In *Marketing Management* (pp. 177-235). Springer, Cham.

Dolnicar, S., & Ring, A. (2014). Tourism marketing research: Past, present and future. *Annals of Tourism Research*, *47*, 31-47.

Dwivedi, Y. K., Ismagilova, E., Hughes, D. L., Carlson, J., Filieri, R., Jacobson, J., ... & Wang, Y. (2021). Setting the future of digital and social media marketing research: Perspectives and research propositions. *International Journal of Information Management*, *59*, 102168.

Esmailipour, M. (2016). *Step-by-step Data Mining Training with RapidMiner*. *Ati Negar*, Tehran, Iran. https://www.adinehbook.com/gp/product/6006004815

Faezi Rad, M., & Pouya, A. (2016). Clustering of online stores from the supplier's point of view with the help of optimizing the number of clusters in the two-stage SOM algorithm. *Iranian journal of Industrial Management Studies*, 14(43), 109-134.

Ghasemaghaei, M., & Calic, G. (2020). Assessing the impact of big data on firm innovation performance: Big data is not always better data. *Journal of Business Research*, *108*, 147-162.

Ghasemian Sahebi, A., Moshabaki, A., & Khodadad Hosseini, H. (2018). Investigating Brand loyalty through Customer Engagement in Online Brand Communities (A Case study of Instagram users). *Brand Management*, *5*(1), 13-34.

Gil-Soto, E., Armas-Cruz, Y., Morini-Marrero, S., & Ramos-Henríquez, J. M. (2019). Hotel guests' perceptions of environmental friendly practices in social media. *International Journal of Hospitality Management*, *78*, 59-67.

Grimmer, J. (2015). We are all social scientists now: How big data, machine learning, and causal inference work together. *PS: Political Science & Politics*, *48*(1), 80-83.

Guo, Y., Barnes, S. J., & Jia, Q. (2017). Mining meaning from online ratings and reviews: Tourist satisfaction analysis using latent dirichlet allocation. *Tourism Management*, *59*, 467-483.

Gupta, M. S., Keen, M. M., Shah, M. A., & Verdier, M. G. (Eds.). (2017). *Digital Revolutions in Public Finance*. International Monetary Fund.

Hair Jr, J. F., & Sarstedt, M. (2021). Data, measurement, and causal inferences in machine learning: opportunities and challenges for marketing. *Journal of Marketing Theory and Practice*, *29*(1), 65-77.

Halawani, F. M., Soh, P. C., & Muthaiyah, S. (2019). The Effect of Social Media on Hotels' Business Performance in the Lebanese Hotel Sector: Effect of Social Media on Hotels' Business Performance. *Journal of Electronic Commerce in Organizations (JECO)*, *17*(3), 54-70.

Han, J., Pei, J., & Kamber, M. (2011). *Data Mining: Concepts and Techniques*. Elsevier.

Hastie, T., Tibshirani, R., & Friedman, J. (2009). *The Elements of Statistical Learning: Data Mining, Inference, and Prediction.* 2nd edition. Springer, Germany.

Heidari, A., Valipour, A., & Bakhtiyari, B. (2017). Marketing research trend in Iran: An analytical review. *Management Research in Iran*, *21*(3), 97-119.

Hsu, C. W., Chang, Y. L., Chen, T. S., Chang, T. Y., & Lin, Y. D. (2021). Who Donates on Line? Segmentation Analysis and Marketing Strategies Based on Machine Learning for Online Charitable Donations in Taiwan. *IEEE Access*, *9*, 52728-52740.

Huang, J. C., Ko, P. C., Fong, C. M., Lai, S. M., Chen, H. H., & Hsieh, C. T. (2021). Statistical Modeling and Simulation of Online Shopping Customer Loyalty Based on Machine Learning and Big Data Analysis. *Security and Communication Networks*, *2021*.

Jiang, Y., & Hong, F. (2021). Examining the relationship between customer-perceived value of night-time tourism and destination attachment among Generation Z tourists in China. *Tourism Recreation Research*, 1-14.

Kassambara, A. (2017). *Practical Guide to Cluster Analysis in R: Unsupervised Machine Learning* (Vol. 1). Sthda.

Kauffmann, E., Peral, J., Gil, D., Ferrández, A., Sellers, R., & Mora, H. (2020). A framework for big data analytics in commercial social networks: A case study on sentiment analysis and fake review detection for marketing decision-making. *Industrial Marketing Management*, *90*, 523-537.

Kim, J. J., Ariza-Montes, A., & Han, H. (2021). The Role of Expected Benefits towards Smart Hotels in Shaping Customer Behavior: Comparison by Age and Gender. *Sustainability*, *13*(4), 1698.

Krafft, M., Sajtos, L., & Haenlein, M. (2020). Challenges and opportunities for marketing scholars in times of the fourth industrial revolution. *Journal of Interactive Marketing*, *51*, 1-8.

Kumar, M. R., Venkatesh, J., & Rahman, A. M. Z. (2021). Data mining and machine learning in retail business: developing efficiencies for better customer retention. *Journal of Ambient Intelligence and Humanized Computing*, 1-13.

Kumar, V., & Reinartz, W. (2016). Creating enduring customer value. *Journal of Marketing*, *80*(6), 36-68.

Lansley, G., & Longley, P. (2016). Deriving age and gender from forenames for consumer analytics. *Journal of Retailing and Consumer Services*, *30*, 271-278.

Lee, E. B., Kim, J., & Lee, S. G. (2017). Predicting customer churn in mobile industry using data mining technology. *Industrial Management & Data Systems*.

Leong, L. Y., Hew, T. S., Ooi, K. B., & Lin, B. (2019). Do electronic word-of-mouth and elaboration likelihood model influence hotel booking?. *Journal of Computer Information Systems*, *59*(2), 146-160.

Liu, P. (2021, March). Research on Customer Value Measurement. In *6th International Conference on Financial Innovation and Economic Development (ICFIED 2021)* (pp. 74-78). Atlantis Press.

Ma, L., & Sun, B. (2020). Machine learning and AI in marketing–Connecting computing power to human insights. *International Journal of Research in Marketing*, *37*(3), 481-504.

Mansouri Moayyed, F., kordheydari, R., & Ghasemian Sahebi, A. (2020). The Role of Knowledge Brokers in Developing of technical knowledge marketing model for Knowledge-Intensive Business Service. *Management Research in Iran, 24*(2), 35-60. URL: http://mri.modares.ac.ir/article-19-31309-fa.html

Miklosik, A., & Evans, N. (2020). Impact of big data and machine learning on digital transformation in marketing: A literature review. *IEEE Access*, *8*, 101284-101292.

Moro, S., Rita, P., & Vala, B. (2016). Predicting social media performance metrics and evaluation of the impact on brand building: A data mining approach. *Journal of Business Research*, *69*(9), 3341-3351.

Raguseo, E., Neirotti, P., & Paolucci, E. (2017). How small hotels can drive value their way in infomediation. The case of 'Italian hotels vs. OTAs and TripAdvisor'. *Information & Management*, *54*(6), 745-756.

Raschka, S., & Mirjalili, V. (2017). Python Machine Learning: Machine Learning and Deep Learning with Python. *Scikit-Learn, and TensorFlow. Second Edition ed*.

Saleem, F., Khattak, A., Ur Rehman, S., & Ashiq, M. (2021). Bibliometric analysis of green marketing research from 1977 to 2020. *Publications*, *9*(1), 1.

Sanayei, A. (2017). *The Fourth Industrial Revolution*, Isfahan, University Jihad, University of Isfahan, Iran. https://isma.org.ir/book-4th_industrial_revolution/

Sánchez-Franco, M. J., Navarro-García, A., & Rondán-Cataluña, F. J. (2019). A naive Bayes strategy for classifying customer satisfaction: A study based on online reviews of hospitality services. *Journal of Business Research*, *101*, 499-506.

Shalini, S., & Lal, K. (2016, April). Improved pseudo-association rules technique. In *2016 International Conference on Computing, Communication and Automation (ICCCA)* (pp. 890-895). IEEE.

Sheehan, N. T., & Bruni-Bossio, V. (2015). Strategic value curve analysis: Diagnosing and improving customer value propositions. *Business Horizons*, *58*(3), 317-324.

Soltani Zanozi, M. (2019). A Study of the Role of Big Data in Marketing and Business, *The Second National Conference on New Thoughts in Business Management*, Tehran, Iran. https://civilica.com/doc/950800/

Suarniki, N. N., & Lukiyanto, K. (2020). The role of satisfaction as moderation to the effect of relational marketing and customer value on customer loyalty. *International Journal of Innovation, Creativity and Change*, *13*(4), 108-122.

Tukiran, M., Tan, P., & Sunaryo, W. (2021). Obtaining customer satisfaction by managing customer expectation, customer perceived quality and perceived value. *Uncertain Supply Chain Management*, *9*(2), 481-488.

Wang, Y., Deng, Q., Rod, M., & Ji, S. (2021). A thematic exploration of social media analytics in marketing research and an agenda for future

inquiry. *Journal of Strategic Marketing*, *29*(6), 471-491.

Weiss, S.M., Indurkhya, N., & Zhang, T. (2015). *Fundamentals of Predictive Text Mining*. Springer.

Xie, G., Qian, Y., & Wang, S. (2021). Forecasting Chinese cruise tourism demand with big data: An optimized machine learning approach. *Tourism Management*, *82*, 104208.

Xu, M., David, J. M., & Kim, S. H. (2018). The fourth industrial revolution: Opportunities and challenges. *International Journal of Financial Research*, *9*(2), 90-95.

Yen, C. H., Teng, H. Y., & Tzeng, J. C. (2020). Innovativeness and customer value co-creation behaviors: Mediating role of customer engagement. *International Journal of Hospitality Management*, *88*, 102514.

Zhou, L., Pan, S., Wang, J., & Vasilakos, A. V. (2017). Machine learning on big data: Opportunities and challenges. *Neurocomputing*, *237*, 350-361.