Analysis of Health tourist's Behaviors Using Data Mining Method (International Tourists)

Masoumeh Sheikhhassan[1], Mehdi Sadeghzadeh[2*], Mahdi Faghihi[3]

[1] Student of information technology management, department of management, faculty of management and economics, Tehran Science and research branch, Islamic Azad University, Tehran, Iran

[2] Assistant professor of computer engineering, department of computer, faculty of electrical and computer engineering, Mahshahrbranch, Islamic Azad University, Mahshahr, Iran

[3] Assistant professor of information technology management, department of management, faculty of management and economics, Tehran Science and research branch, Islamic Azad University, Tehran, Iran

*Corresponding authors: Mehdi Sadeghzadeh, Assistant professor of computer engineering, department of computer, faculty of electrical and computer engineering, Mahshahrbranch, Islamic Azad University, Mahshahr, Iran. Email:m.sadeghzadeh@mhriau.ac.ir

Abstract

Background and Objectives: Currently, attracting tourists and keeping them in a competitive environment, especially in the health tourist industry, have a high priority. Given the different behaviors of health tourists, the analysis of their behaviors is important.

Methodology: This research is applied in terms of purpose and descriptive method. Data mining methods and powerful Python tools were used to analyze the data. The combined method of clustering algorithm and association rules were adopted to identify the pattern of tourists' behaviors.

Results: After analyzing the data, the motivation of foreign tourists to travel to Iran was determined and the selected destinations were prioritized based on importance. In addition, according to the behaviors and different needs of health tourists, different travel packages were designed.

Conclusion: By analyzing the behaviors of tourists and clustering them according to the type of behaviors in each cluster, it is possible to achieve strategies tailored to their needs, predict the future behaviors of tourists through their past behaviors, and be informed of the threats, opportunities, strengths, and weaknesses of the system.

Keywords: Cluster analysis; Data mining; Facilitator companies; Health behavior

1. Introduction

Currently, with the increasing growth of information technology and smart technologies, competitiveness of the tourist market, and growth in the level and information of the customers, the marketing process is becoming more complex and sensitive every day. Furthermore, paying attention to the issue of customer relationship management is becoming more important every day and the budget allocated to it is increasing in organizations. Leading organizations always try to design new products according to the needs of their customers and offer customized goods based on preferences for each group of customers with common needs. This would be possible by the analysis of customers' tastes and behaviors (1).

Customer retention is also important in today's competitive world as its costs may be many times lower than attracting new customers. It is especially valid for the group of customers with the most profit for the organization. The most important prerequisite for knowing customers and discovering their usefulness is market clustering. During customer classification, they are divided into different groups. In this way, similar customers are placed in the same group. More accurate clusters and more similar group members lead to more accurate customer knowledge and as a result more services.

Because tourists have different cultures and behaviors, it is essential to recognize the pattern of tourist behaviors. With regard to the many determinants affecting the supply and demand of tourism, the perception of the behavioral patterns of tourists can lead to desirable results. Therefore, the recognition of the behavioral patterns of tourists (e.g., where they come from, how much they spend, how much time they spend, and when they spend) helps to develop strategy and marketing. Furthermore, it would increase profits for service providers.

Companies and tourism service providers have large sets of data related to their customers. If the data are stored in an organized and integrated manner in the database, they can be used for marketing decisions. The identification of important variables and their relationships is often a hard task that can be enhanced with the help of data mining. In this regard, the present study tried to extract the behavioral

groups of tourists and their dominant characteristics. For this purpose, a hybrid method based on clustering algorithms and association rules was used to suggest travel packages based on different behaviors of tourists according to new approaches and policies extracted (2). This will increase the number of visitors, retain tourists, attract investors, and improve Iran economy. To date, numerous studies have been carried out in this field concerning the tourism industry. One of the requirements of this study is that due to the complex behaviors of the new tourist generation and increasing demand for different products, it has become very difficult to understand tourists' behaviors.

For this purpose, by analyzing the behaviors of tourists using dimensions, such as culture, lifestyle, religion, social relations, and type of disease, it is possible to meet the needs and expectations of tourists in each sector and have better management to accordingly choose a tourist destination. In the present study, a recommender algorithm was designed to introduce different travel packages based on different tourist behaviors. Accordingly, the current study analyzed the behaviors of tourists using data mining techniques.

2. Theoretical foundations and review of previous studies

Organizations are trying to gain a greater share by understanding the needs and behaviors of tourists in competitive environments. The tourist experience of a destination is a fundamental product in tourism. For this reason, the tourist destination is in the center of the competition. The features of a tourist destination, including natural resources, cultural resources, human resources, facilities, medical services, and other activities, should be limited to the tourist visit and offered to others to visit the destination. Therefore, it is not enough for the destinations to be in the minds of the customers; however, it should be unique and distinctive to be selected in the final decision of the tourist. The mental image can represent a simplified set of fragmented perceptions containing a lot of information about a place (3).

The destination image also identifies the strengths and weaknesses perceived by the tourist. In this regard, e-tourism helps tourists to have a targeted program led by the government, travel agencies, and centers in the shortest time, with the least facilities and costs, and identifies the tourism capacity of a country. In this regard, the provision of a comprehensive tourism system, development of channels and public educational space for tourists to facilitate the process of booking and choosing a tourist destination, and use of historical, natural, and environmental attractions in one step without frequent visits to offices and organizations are very impressive.

Advertising on channels is carried out both offline and online. Offline advertising is an individual's experience from his/her previous trip and announcement of the level of his/her satisfaction with the trip to another person. In online advertising, information is virtually received through social networks and websites. As a result of improving information and communication, the provision of facilities and discounts have a significant role in advertising. For this purpose, in the present study, the behaviors of tourists were analyzed using data mining techniques.

Data mining is the exploration and analysis of large volumes of data to discover meaningful patterns and rules. The process of data mining is sometimes called knowledge discovery (4). Data mining methods are divided into descriptive and predictive categories. Methods for predicting the value of a particular attribute based on other attributes. The predicted feature is called the goal and is dependent on other features; nevertheless, the goal is to use the extraction techniques of descriptive patterns. Prediction methods include division, regression, and time series. Furthermore, descriptive methods include clustering and rules of the association. In the present study, clustering methods and association rules were utilized to extract and analyze customers' behaviors. In the clustering method, an irregular population is divided into a set of regular subgroups and objects are grouped based on the most similarity between the members of each cluster and the least similarity between different clusters. The criterion of similarity when all the features are continuous is usually expressed by the Euclidean distance. Otherwise, a suitable criterion has been considered for it.

Clustering methods are divided into two categories, namely segmentation methods and hierarchical methods. In this study, the clustering method was used. In other words, the data are clustered in group K where each group should have at least one object and each object should belong to only one group. The K-Means algorithm is a common and efficient method in clustering taking k (i.e., the number of clusters) as input and dividing the set of n objects into k clusters. The algorithm firstly works randomly, selects an object k as the center of the initial clusters, and then assigns each object to the clusters due to its maximum similarity to the center of the clusters. Refer to the second step until there is no change in the cluster (5).

In addition, the RFM variable has been used to segment customers, which is one of the most popular methods in analyzing customer value. The RFM is based on three factors, namely novelty (R), frequency (F), and financial value (M). These three variables mean freshness and it refer to the time interval from the last visit, number of visits means the number of visits in a certain period and financial value means paying money in a certain period. Another task of data mining is to discover the rules of the association. Using communication rules, dependencies and connections between data can be discovered in a database. The law of inference association is in form x and then y; accordingly, x and y are incompatible sets of cases. This rule of thumb means that also x trades probably contain y. Each association law has two criteria of trust and support which are as follows according to Equation 1:

(Equation 1)

$$support = \frac{(X \cup Y).count}{n}$$

$$confidence = \frac{(X \cup Y).count}{X.count}$$

They also use recommender systems which are a subset

of decision support systems to analyze past behaviors, make recommendations for current issues, and assist the user in making decisions to find the best option among the available options.

3. Background research

A study carried out by (6) has studied and identified the behavioral pattern of tourists, implemented the laboratory model of the destination management system, and presented the appropriate response to passenger demand using data mining techniques. Another study conducted by (7) predicts customer behaviors in the online method. In a study performed by (8), the behaviors of tourists have been extracted using the data mining method in the telecommunication company and the clustering method has been used to analyze customer behavior (9). In addition, (10) examined the inconsistencies in tourism marketing strategies. The current study aimed to identify discourse differences in expressing the position of regional tourism marketing strategy documents and prioritize the identity of a specific destination.

A study carried out by (11) provides a forecasting framework that uses machine learning and internet search indicators for the prediction of tourist arrivals. Accurate forecasting of tourist numbers can help tourism professionals optimize and manage capacity. Furthermore, (12) have identified components and content analysis for medical tourism facilitator (MTF) websites. In another study conducted by (13), the Apriori algorithm, with great weakness in time calculations, has been investigated. Moreover, due to repeated studies, it has to scan the database repeatedly for each combination, which is very time-consuming. The present study investigated the effect of the clustering algorithm on the Apriori algorithm and more information can be produced in the shortest time by combining these two algorithms.

4. Research method

In this study, the data mining method was used to analyze the behaviors of foreign tourists and Python's powerful tool was utilized to analyze the data. The statistical population of the current study consisted of tourism-related organizations, such as hotels, restaurants, hospitals, and tours. In addition, the statistical sample of
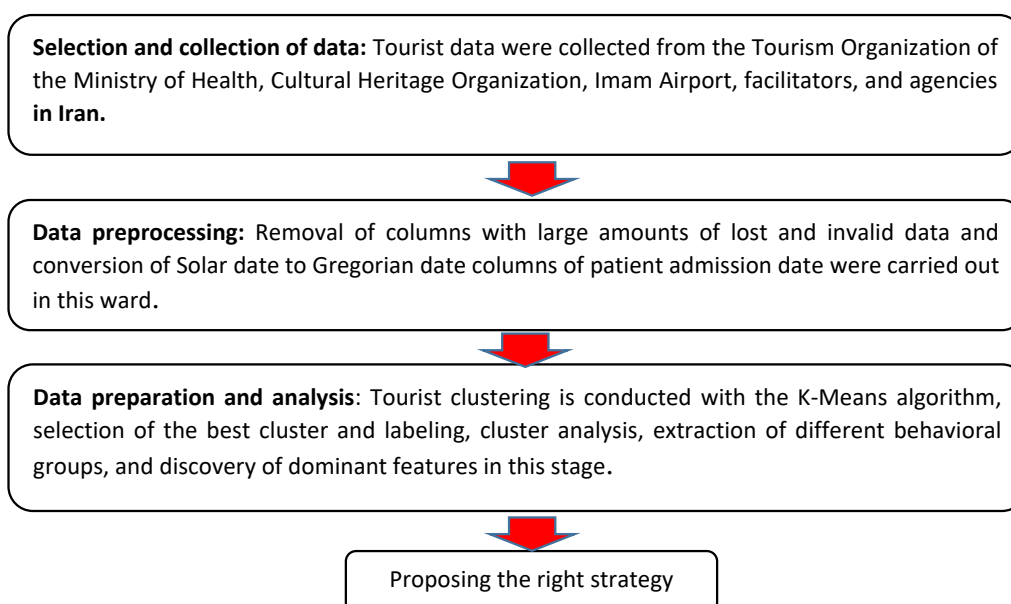
Selection and collection of data: Tourist data were collected from the Tourism Organization of the Ministry of Health, Cultural Heritage Organization, Imam Airport, facilitators, and agencies **in Iran.**

Data preprocessing: Removal of columns with large amounts of lost and invalid data and conversion of Solar date to Gregorian date columns of patient admission date were carried out in this ward.

Data preparation and analysis: Tourist clustering is conducted with the K-Means algorithm, selection of the best cluster and labeling, cluster analysis, extraction of different behavioral groups, and discovery of dominant features in this stage.

Proposing the right strategy

Table 1. Sample information of foreign tourists in 2020

Row	University	Hospital	Gender	Residential status	Nationality	Date of acceptance	Type of admission	Release date	Total cost
1	Shahid Beheshti University of Medical Sciences, Tehran, Iran	Seyed Al-Shohada	Female	Short stay	Afghanistan	1398/08/22	Null	2020/01/12	22772114
2	Tehran University of Medical Sciences, Tehran, Iran	Dr. Shariati	Male	Short stay	Afghanistan	1398/10/27	Null	2015/01/18	2919664
3	Tehran University of Medical Sciences, Tehran, Iran	Dr. Shariati	Female	Short stay	Iraq	1398/01/26	Null	2016/03/14	8402560

this study was considered the data set of the Ministry of Health of Iran and interviews with facilitating companies.

The method begins with the help of RFM criteria. The scores of tourists were calculated, among which 111 categories were the best tourists. These tourists were examined regarding the three aspects, namely last travel time, maximum travel, and maximum payment. Then, the data were clustered using the K-Means algorithm. In this section, firstly, the Davies-Bouldin index (DBI) was used to find the best number of clusters to categorize the data set per year and then the association rules. The cluster, including the category of 111 customers, was examined for the determination of the rules related to top tourists using the Apriori algorithm to attract more tourists.

5. Suggested method

In the data selection and collection stage, the information was collected from tourism organizations. Table 1 tabulates an example of the information collected in 2019 related to foreign tourists.

In the data preprocessing phase, there was no acceptance type column in all the years except for 2020 and 2021. Therefore, this column was removed except for the two years and was not used in data analysis stages, such as clustering and creating association rules.

In the data preparation and analysis phase, the Python Quartiles library was used to implement the RFM technique for grading columns at the rates of return, frequency, and amount of money for the tourists of different nationalities each year. This is performed to grade the recency column showing the values below 0.25 ° 1, below 0.5 ° 2, below 0.75 at 3, and the rest at 4. The nationalities that accept the first rank have the least delay in referring to the medical centers of Iran. This scenario will be reversed for the frequency and monetary value columns. After this step, a table will be shown in which the RFM scores for the tourists of each nationality are specified.

The number of tourist trips from different countries in the RFM algorithm group was equal to 111. The group with the highest number of repetitions, lowest travel delay, and highest travel costs had the highest profit for Iran. Table 2 shows the output of Python software related to different categories.

In the clustering stage, firstly, the DBI was used to find the best number of clusters to categorize the data set. The year 2020 had five clusters, and due to the limited number of clusters 1, 2, 3, 4, only 111 degrees in the zero cluster to the number of 83,132 was examined because the number of repetitions is important to reach the rules and it was not possible to discover the rules with a small number. In addition, in 2017, 2018, 2019, and 2021 according to 2020, the degree of 111 in each cluster was examined.

Table 2. Sample of the RFM score of the foreign tourist's software output of 2020

Country	Recency	Frequency	Monetary value	R Quartile	F Quartile	M Quartile	RFM class
Iraq	1.0	51554	2755334884676183	1	1	1	111
Iran	16.0	197	965928384183	1	1	1	111
Afghanistan	1.0	66343	5644893918743	1	1	1	111
Azerbaijan	3.0	5006	250366422196	1	1	1	111
Oman	21.0	4169	239421286785	1	1	1	111
Bahrain	4.0	938	101509260085	1	1	1	111
Kuwait	12.0	820	84116273227	1	1	1	111
Pakistan	1.0	1631	82510341281	1	1	1	111
Saudi Arabia	20.0	167	73699235981	1	1	1	111
Qatar	17.0	416	37347259042	1	1	1	111

Table 3. Number of selected clusters of Python software in 2020 based on Davies-Bouldin index criterion

Number of cluster	Davies-Bouldin index score	Python codes
3	7.11	Kmeans= kmeans(n_clusters=3, random_state=1).fit(data.end) Labels=kmeans.lables_print(davies_bouldin_score(data_end, lables)) 7.114988407616791e-06
4	1.53	Kmeans= kmeans(n_clusters=4, random_state=1).fit(data.end) Labels=kmeans.lables_print(davies_bouldin_score(data_end, lables)) 1.9265004615564522e-05
5	0.16	Kmeans= kmeans(n_clusters=5, random_state=1).fit(data.end) Labels=kmeans.lables_print(davies_bouldin_score(data_end, lables)) 0.16164501888849642
6	0.18	Kmeans= kmeans(n_clusters=6, random_state=1).fit(data.end) Labels=kmeans.lables_print(davies_bouldin_score(data_end, lables)) 0.18050865705076888

After reviewing the DBI score briefly for the number of different clusters, it was observed that except for 2018 and 2020, the best number of clusters was equal to 3 for the rest of the data set. For 2018 and 2020, this number equaled 6 and 5, respectively. Table 3 tabulates a list of Python software output clusters selected in 2020.

In this study, the K-Means method was used for clustering and the clustering results were reviewed based on the data from different years. In addition, in the present study, the number of frequent visits was investigated using the Word Cloud technique. Most duplicate images were larger in word size indicating that they had the most views. Furthermore, the fewest duplicates were smaller indicating that they had the least views.

6. Analytical recognitions

After calculating the RFM score of each nationality and clustering the foreign tourist data set, the policies resulting from the implementation of the rules related to the values of min_support (0.01) and min_confidence (0.85) in the last 5 years were reviewed. Table 4 shows an example of the relevant rules for foreign tourists in 2020.

In this regard, the policies adopted in 2020 are as follows:

Policy 1: Iraqi individuals have traveled to Tehran more frequently, chosen Mashhad, Ahvaz, Urmia, Shiraz, Qom, Kermanshah, and Tabriz, Iran, as their travel destinations, and mostly referred to the aforementioned cities for eye treatment. In addition, by analyzing the behaviors of tourists, it can be concluded that Iraq has a dry climate and low rainfall, increased pollution, and reduced river

water. In this regard, Rasht, Mazandaran, and Tabriz, Iran, which, unlike Iraq, have mountainous, green, and pleasant places, can be recommended to Iraqi tourists. Moreover, considering that the religion of 60% of Iraqis is Shiite and they are very interested in visiting religious places, such as the shrine of Imam Reza and go to that city for further treatment, it is important to equip those places for tourists.

Policy 2: Omani individuals travel to Shiraz and undergo Rhinoplasty in Shahriar, Shiraz, and are more confident in ophthalmology at Dr. Mir Hosseini hospital in Shiraz. In this regard, in order to increase the income of Shiraz hospitals and clinics, along with Rhinoplasty, Botox, teeth whitening, and other cosmetic surgeries can be offered to tourists with a free visit or a special discount. In addition to the treatment package, due to the proximity of Shiraz to Isfahan, it can be recommended as a travel package for Omani tourists.

Policy 3: Azerbaijan firstly targeted Tabriz for treatment and then Tehran, Ardabil, Gilan, Urmia, and Qom, Iran. In addition, by analyzing the behaviors of Azerbaijani tourists, it can be concluded that most of them, due to their proximity to Tabriz and speaking the same language, firstly traveled to Tabriz for tourism and treatment. Because 93% of Azeris are Shiite, they also visit holy shrines in several cities, such as Qom and Mashhad. The proposed package of travel to Azeris is accompanied by a tour of Tehran because the gross domestic product of Azerbaijan is 50 thousand dollars and it is financially good. They can buy many products from Tehran.

Policy 4: Most short-term tourists to Shahid Beheshti Medical Center in Qom, Jam hospital in Tehran, Bina Ophthalmology hospital in Mashhad, Imam Sajjadi
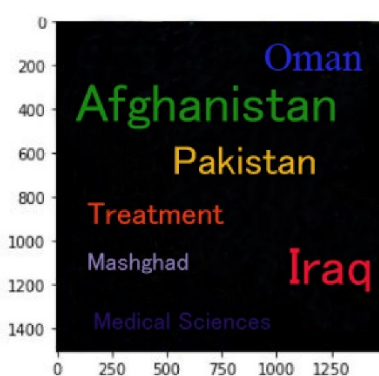

Figure 1. Sample of Word Cloud extracted results in 2020

Table 4. An example of association rules for foreign tourists in 2020

Support	Confidence	Country	Service center	Cluster
0.036	0.940	Iraq	Urmia University of Medical Sciences, Urmia, Iran	0
0.040	0.995	Iraq	Ahvaz Jondishapur University of Medical Sciences, Ahvaz, Iran	1
0.013	0.995	Iraq	Kermanshah University of Medical Sciences, Kermanshah, Iran	1
0.010	0.984	Iraq	Urmia Healing hospital, Urmia, Iran	3
0.011	0.994	Iraq	Ahvaz Mehr hospital, Ahvaz, Iran	4
0.026	0.982	Iraq	Bina Ophthalmology hospital, Mashhad, Iran	2
0.031	0.942	Iraq	Jam hospital, Tehran, Iran	1

Shahriar hospital for delivery, Mashhad Mother hospital for rhinoplasty, Izadi Medical Center in Qom for delivery, and Firoozabadi Hospital in Ray, Iran, have referred for normal birth. In addition, in the long run, most patients have been referred to Nekouei, Hedayati, and Forghani hospitals in Qom for delivery and then to Hashemi hospital, Umm Al-Banin hospital in Mashhad, and Shahid Beheshti hospital in Qom.

Policy 5: Afghanistan in the cities of Tehran, Shahid Akbarabadi Hospital with 2952 patients for delivery and Imam Sajjad Shahriar Hospital with 1831 patients for delivery and in Qom Nekouei, Hedayati, Forghani Educational and Medical Center with 4893 patients Delivery, acute appendicitis, neonatal dyspnea, and jaundice and Shahid Beheshti Hospital in Qom with 3416 patients for treatment of pneumonia and atherosclerotic heart disease, stroke, acute appendix and Shahid Hasheminejad Hospital in Mashhad with 1644 patients for delivery, cataracts account for the largest number of patients. As a result, more Afghan tourists have traveled to Tehran, Qom, and Mashhad to give birth. Because most Afghan tourists are Muslim and they want the medical staff to be female physicians. Furthermore, before and after the operation, according to their tastes, they underwent manicures or pedicures to reduce stress, relax the muscles, and increase blood circulation, and after the operation, they have performed dominoplasty and abdominoplasty. For this purpose, along with and after treatment, the patient and his/her companions can choose the tour packages of Kish and Qeshm to benefit from the services. Because Afghans have very limited visits to Kish and Qeshm and because Islam is the main religion of Afghanistan, they can have a variety of entertainment there while preserving Islamic customs. Therefore, visiting these cities is very attractive to Afghan tourists.

Policy 6: Ophthalmology is very important in Iran because 90% of tourists go to Iran for eye treatment and choose destinations, such as Tehran, Mashhad, Tabriz, and Shiraz. Therefore, by attracting the best physicians and specialists, building ophthalmology treatment centers, increasing the awareness of the treatment staff,

and providing the necessary facilities and equipment, we can take a step in this important direction and achieve profitability.

Policy 7: Most foreign tourists have traveled to Iran from Afghanistan, Iraq, Azerbaijan, Oman, and Pakistan for treatment, respectively. In Pakistan, given that 97% of Pakistanis are Muslim and 20% of them are Shiite, the majority of them travel to Qom and Mashhad for treatment and pilgrimage. Pakistanis do not know much about Iran and do not have a high level of funding for entertainment, sightseeing, and even treatment in Iranian cities. It should be noted that most of the advertising in these countries is performed by word of mouth and offline. According to the behavior style of Pakistani people, traveling to Tehran is offered as a travel package to visit the tourist attractions and places of interest of that city.

According to the above-mentioned information, the existence of facilitation companies is of great importance. Figure 1 depicts four interactive models of the main players in health tourism with international patients in pioneer countries in this industry.

Mode 1: It is the first and simplest mode of medical tourism. A direct medical tourism model in which patients communicate with the medical center mainly through word-of-mouth advertising or websites.

Mode 2: Medical tourism is provided by MTF companies. In this case, patients use services that are provided by those facilitation companies cooperating with foreign hospitals in the destination countries; the coordination of treatment affairs, logistics services, accommodation, insurance affairs, and follow-up of patients during recovery, all can be included in these services.

Mode 3: In order to use lower-cost treatment services, patients are referred to foreign hospitals or facilitators by employers and insurance companies and encouraged to seek treatment from foreign hospitals.

Mode 4: It is the transfer of patients from hospitals in the country of origin to foreign hospitals. Outsourcing medical services in this way is defined by sending patients directly from hospital to hospital. Johns Hopkins Hospital in the United States, for instance, uses this method based
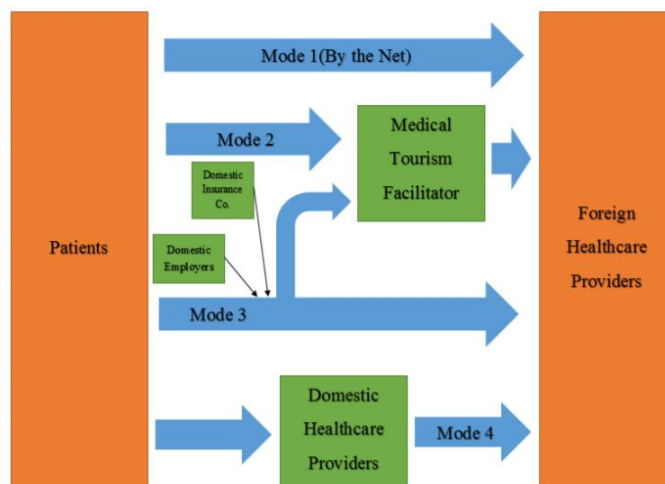

Figure 2. Four interactive models of the main players in health tourism

on counseling patients and referring them to hospitals that have a contract with them, such as hospitals in India and Singapore (21).

7. Discussion

Iran, despite its vast geographical location, along with the diversity of climate, existence of many historical attractions and places, and prominent role of religion, is still unknown to the people of the world. After clustering and selecting the best tourists and surveys conducted on tourism data, the results obtained from this study indicated that Mashhad and Qom have many capabilities to attract religious tourists. However, religious tourism in Iran, despite 8,919 holy religious sites, still lacks specialized and centralized organization, and this unsettled situation and lack of planning can be observed even in important cities, such as Mashhad and Qom. In this regard, proper planning has not been performed for the countries with the most cultural and religious similarities with Iran. They cannot plan properly for other markets, such as European countries.

Afghanistan and then Iraq has the largest number of foreign tourists to Iran indicating that agencies can use Afghanistan and Iraq to boost tourism. One of the challenges facing Iraqi patients is a lack of trust in authorized and reputable health tourism companies. Negative publicity from agencies and intermediaries has led them to believe that these companies make a lot of money from them. Therefore, they prefer to go to hospitals immediately which leads to problems, such as the profitability of intermediaries, lack of proper services, and incorrect translation of treatment. In this regard, building trust between intermediaries should be performed in the source of advertising. It is necessary to attract patients to medical centers through tourism companies and transfer patients from the moment of arrival to the end of the trip by introducing and following the treatment in equipped medical centers with professional translators in a correct and planned cycle.

8. Conclusion

According to the result of the present study, tourism officials should focus their programs on the target markets of medical tourism, especially for neighboring countries, such as Afghanistan, Iraq, Azerbaijan, Oman, and Pakistan, which are more culturally, religiously, and linguistically similar to Iran. Furthermore, medical services should meet their expectations. This is undoubtedly important in the current economic situation as the urgent need of the country for the inflow of currency.

Moreover, most of the problems for which foreign tourists have referred to the hospital are gynecology, ophthalmology, beauty, cardiovascular diseases, and orthopedics because the cost of treatment is low, the quality of services is high, and good physicians and specialists are working in this field. The equipment of hospitals with complete facilities, such as specialized surgical devices, hospital beds, and other facilities in the operating room, and development of startups for remote consulting and surgery play an important role in enhancing the field of tourism and health.

In cities, such as Shiraz, Isfahan, Tehran, Qom, Mashhad, Tabriz, Rasht, and Mazandaran, due to natural attractions, historical sites, handicrafts, and pilgrimage sites, medical services have the most visitors. Creating appropriate infrastructures (e.g., communication networks), designing smart campaigns, building restaurants, hospitals, and hotels in pleasant-weather places with easy local access and special discounts, providing banking facilities to increase the quality of services, turning tourist attention into attractions, and profitable deals with minimal costs can motivate individuals to travel and increase the number of visitors.

The results of the present study are consistent with the findings of a study carried out by Harandi et al. In the aforementioned study (14), the results showed that motivation, communication activities, information resources, infrastructures, and potentials have a significant impact on the choice of the destination. The findings of the present study are also consistent with the results of a study conducted by Reza Minaei et al. (15) indicating that channels and information sources are very effective in advertising and motivating tourists, leading to attract more tourists, provide more services, and ultimately achieve more profitability. For this reason, most tourists, based on the experience and mentality of their previous trips, offer their visit experience to others using word-of-mouth and offline advertising. Therefore, the present study corresponds to the dimensions of the proposal and visit in the variable of the destination image of a study by Taghi Pourian et al. (16).

In addition, according to a study carried out by Tahmasebi and Roshanian (17), interfering factors, including social, economic, psychological, cultural, and environmental factors, are influential on the variable of tourism behavior that is consistent with the variable of motivation in the current study and a study by Slabbert (18); however, the findings of the present study are inconsistent with other infrastructural factors, destination image, and information sources. In another study performed by (19), the combined clustering method and association rules have been used to analyze customer behavior, and finally in each cluster a policy model is presented that is fully consistent with the findings of the current study. Furthermore, offering different travel suggestions according to the different behaviors of tourists is in line with the results of a study carried out by Bagheri et al. (20) in creating tourism recommendation systems.

9. Recommendations and restrictions

Therefore, preparing short films by radio and television of Iran in connection with the tourism industry in Iran and sharing on social networks, creating strong teams to respond to offline and word-of-mouth electronic advertisements of tourists on websites, holding training courses for foreign tourists, and designing different advertising programs for tourism industry intermediaries are some of the practical suggestions of the current study in order to introduce travel packages based on the cultural heritage by tourism organizations.

Moreover, in this research to analyze the behavior of tourists, characteristics such as religion, culture,

lifestyle, type of treatment by the big data of the tourism organization of the ministry of health, cultural heritage organization, and interview with a facilitator company. It is recommended to analyze the behavior of tourists using technology and income based on social and cultural class, principles, and laws in the country in future studies using data mining techniques and clustering algorithms. The current study analyzed domestic tourists with different styles, such as ecotourism, urban tourism, adventure, economist, and interest in architecture, nature, tranquility, population, and history. It is also suggested to investigate the styles of Iranian tourists who have traveled abroad in future studies. In addition, in the field of health, the behaviors of tourists in neighboring countries, such as Pakistan, Azerbaijan, Afghanistan, Oman, and Iraq, have been studied; however, in future studies, it is possible to analyze the behaviors of health tourists in European countries and other countries around the world.

One of the limitations of the current study is a lack of access to databases of all sectors due to the expansion of the tourism industry and a lack of sufficient theoretical and experimental background on the subject of the research. Other limitations of this study include advertising, media, and educational limitations due to a lack of appropriate infrastructure and platform. Infrastructural constraints, concentration of industries and services in several large cities, and cultural and religious restrictions are other limitations of the present study.

Acknowledgments

References

1. Raffaele. Mapping tourist consumption behavior from destination card data. International Journal of Tourism Research.2017; 48(2), pages 216-229.

2. Cavagnaro, Staffieri. A study of students' Travellers values and needs to establish futures patterns. Journal of Tourism Futures. 2015; pages 94-107.

3. Banki, Ismail, Dalil, Kawu. Moderating role of affective destination image on the relationship between Tourists satisfaction and behavioural intention. Evidence from obudu mountain resort. Journal of Environment and earth science. 2014; pages 47-60.

4. Han, Kamber. Data mining concepts and techniques, 2nd edition. Morgan Kaufmann publishers.2006; pages 56-63.

5. Akhondzadeh, Albadvi. Mining the dominant patterns of customer shifts between segments by using top-k and distinguishing sequential rules. Management decision. 2016; pages 1976-2003.

6. Khaniki, Movahedifar. Analysis of the model of tourist destinations in travel with the help of data mining studied in Mashhad. The second national conference on applied research in computer science and information technology.

2019; https: // civilica .com / doc / 455020.

7. Jalali, Abdolvand, Harandi. Analysis of customer behavior using data mining, new marketing research. Summer. 2018; 2 (29); pages 93-110.

8. Abdi, Shakori. The study of tourists' propensity to revisit Iran. International journal of social sciences. 2017; 6(1), pages 19-31.

9. Rani.Praveen, Vohra.Rajan, Gulia.Anju. Association rule mining in discovering travel pattern in passport data analysis. (IJCSIT) international journal of computer science and information technologies. 2014; 5 (4), pages 5015-5019.

10. Jeuring, discursive contradictions in regional tourism marketing strategies: the case of Fryslan the netherlands1. *Perspectives on proximity tourism in Fryslan* 2017.

11. Shaolonge Sun. forecasting tourist arrivals with machine learning and internet search index 2019; pages 24-38.

12. Taghizadeh, Shami Zanjani, Haghighi, Hashemabadi. Components and content of medical tourism facilitation websites. Tourism management studies quarterly, 2016; pages 1- 18.

13. Dharshini Azmi. Analysis of accuracy K-Means and apriori algorithms for patient data clusters. Journal of Physics: 2019; pages 1230 -12020.

14. Harandi, Mirzaeian Khamseh. Explaining the model of attracting healthy tourists: using the data theory strategy of the classical foundation. Urban tourism, spring 2017; pages 87-98.

15. Minaei, Namamian, Maroofi, Moradi. Model of the effect of e-tourism on the purchase of tourism services and its compliance with the behavior of Iranian tourists. Bi-Quarterly Journal of Social tourism studies, 2020; pages 412-453.

16. Taghi Pourian, Yazdani, Aghaeifar. The role of destination brand image on the behavioral Tendencies of Tourists case study: tourists in the west of Mazandaran province, urban tourism quarterly. Winter 2019; pages 37-50.

17. Tahmasebi, Roshanian. Characteristics and shopping behavior of tourists in Baneh. Quarterly Journal of tourism management studies. 2016; pages 31-59.

18. Slabbert Elmarie. Vuren van travel motivations and behavior of tourist to a South African resort. Book of proceedings vol. l-international conference on tourism & management Studies_Algarve 2011.

19. Rani Praveen, Vohra.Rajan, Gulia.Anju, association rule mining in discovering travel pattern in passport data analysis. International journal of computer science and information technologies. 5 (4), 2014; pages 5015-5019.

20. Bagheri, Shad, Ghaemi, Honarparvar, Ajagh. A Review of Mobile advisory systems in the field of tourism, journal of surveying engineering and spatial information 2018.

21. Saeed Hashemzadeh, Niloufar A'alami, health tourism facilitator companies in Iran: a missing trade, October 2020; https://www.researchgate.net/publication/344928951.