International Journal of Hospital Research 2018, 7(1): 36-56
http://ijhr.iums.ac.ir
doi 10.34172/ijhr.2018.xx

Research Article

IJHR

Open Access

Providing a Model Based on Recommender Systems for Hospital Services: Case of Shariati Hospital of Tehran

Mehrdad Kargari [1*], Kobra Akbari[1]

[1]Faculty of Systems and Industrial Engineering, Tarbiat Modares University, Tehran, Iran

Abstract

Background and Objectives: In the increasingly competitive market of the healthcare industry, the organizations providing health care services are highly in need of recommender systems to meet the clients' needsand to to identify the factors affecting patient satisfaction focus. The purpose of this study, then, was to provide a model based on recommender systems in order to increase patient satisfaction with the quality of hospital services.
Methods: In order to conduct the model, we used the data related to satisfaction forms of 556 discharged patients from Shariati Hospital in Tehran. By estimating the accuracy of the predictions of the model based on the mean absolute error criterion and the mean squared error, the values were respectively obtained as 40% and 49%.
Findings: In this study, through weighting the characteristic for different groups of patients, the more important services were identified. Considering the number of 148 test data, it was determined that the model of the most important dimensions of the service for each cluster are correctly determined. Therefore, the hospital can decrease dissatisfaction of the new patients in each group through reinforcing the important services in each group, after discharge.
Conclusions: Information technology can provide the possibility of moving towards better services by analyzing customer preferences and tailoring the content and process of service provision according to customer needs. On the other hand, the personalization of products and services is one of the most important factors affecting customer satisfaction.
Keywords: Patient Satisfaction, Service Quality, Personalization, Recommender Systems, Clustering, Feature weighing

Background and Objectives

Patient satisfaction as an outcome index of providing health care is widely regarded as an important index for measuring the quality of health care and has been accepted as an important component in improving clinical performance and efficacy.[1] Getting feedback from the patient is an essential element of political planning and a helpful knowledge provider for efficient management of care and provides important information about what patients are expecting and how patients perceive quality.[2] In the increasingly competitive healthcare market, healthcare managers should focus on achieving patient satisfaction to improve service quality. Therefore, patient satisfaction, which is used as a tool for assessing the quality of services and also focusing on higher-value features, requires initiating a number of recovery strategies in other areas of services that are not good from the perspective of patients.[3]

Due to the process of globalization, technology advancements and demagraphical changes, the healthcare marketplace faces major changes, including increasing demand of patients, changing disease patterns, increasing competition between hospitals and healthcare providers, and reducing government spendings.[4] Personalization of products and services is one of the most important factors affecting customer satisfaction. Recommender systems are widely used in e-commerce to support the personalization of services. Decision strategies for appropriate provision and delivery of services and use of information technology systems are important since they can provide the possibility of moving toward better services by analyzing customer preferences and tailoring the content and the serving process to the needs of the customer.[5]

With regard to increasing demands of customers and

*Corresponding Author: Mehrdad Kargari, Faculty of Systems and Industrial Engineering , Tarbiat Modares University, Tehran, Iran, Tel/Fax: +982182884955, Email: m_kargari@modares.ac.ir

their complex and different needs and expectations, the organizations are seeking to introduce and apply new IT solutions, under which customer relationship management solutions play a significant role.[6]

In recent years, in addition to e-commerce fileds, recommender systems have also been used in the field of healthcare. For example, there are recommender systems with educational purposes, diet, activity assistance and appropriate hospital recommendation. Recommender systems can provide useful information for the users and through these systems users can search for an appropriate service provider using their location, expertise and reputation.

The purpose of this study is to adapt a model using the discharged patients' views to be able to predict appropriate and relevant services to the new patient. In fact, the goal of this model is to take a step forward in increasing patients' satisfaction through personalizing hospital services. Thus, we will first have a review of literature on the patient's satisfaction and personalization of services and recommender systems; then we deal the methodology of the research and the proposed method and finally, the results of the study are presented and then we will discuss and conclude.

Patient Satisfaction

According to Hafezi and Esmaeili[7] "Satisfaction is the existence of a positive feeling which is ultimately developed by the consumer or the recipient". Essentially, this feeling comes about because of customer expectations and supplier performance. In fact, satisfaction is the degree to which a customer believes that the organization or company that produces goods or services meet his needs. Improving and maintaining the quality of care in hospitals is a persistent and enduring challenge.[8] In terms of business, patients represent the main clients of the hospital who receive and feel the healthcare services directly and in reality. Patient satisfaction is a key to maintaining hospital profits; because selecting patients from an optimal hospital is often based on their examination of health information and the experiences of their friends, family members, or colleagues.[9] The administers of the American University of Health, reported in their annual review that patient satisfaction is one of the 10 most important concerns of hospital executives. On the other hand, meeting the needs of patients is the first step towards having loyal patients, therefore, hospitals that are trying to make their patients fully satisfied are likely to be more successful.

Patients' experience of the quality of hospital care can provide clear feedback in order to improve quality in necessary areas. Some of patients' experiences may be good, while some not good. Some experiences are probably more effective for patients in shaping the overall level of their satisfaction in comparison to other experiences. Therefore, when they had a positive experience of the important factors, their overall satisfaction is probably good. On the other hand, if they had a negative experience of important factors, they are more probably dissatisfied.[10]

Various studies have emphasized that various dimensions of the quality of health services are involved in forming the overall patient satisfaction, and these dimensions variously affect patient satisfaction. Therefore, identifying the relative weight of the various dimensions of the quality of health services which are used together to determine the patients' satisfaction is very important.

Based on a conducted scaling, the factors affecting patient satisfaction are classified into 2 categories:

1. Many studies have examined the relationship between demographic factors such as age, gender, health status and level of education with patient satisfaction; however, these factors are not significant predictors of overall patient satisfaction; because these factors can not be changed for healthcare managers who desire to improve patient satisfaction. Nevertheless, patient characteristics should be considered for the fair regulation of patient satisfaction studies to be used in comparative assessments with other health care institutions.
2. The researchers have discussed extensively about the multidimensional characteristics of healthcare environments. Healthcare administrators need more focus on higher-value features and start a series of recovery strategies in poor areas of services related to hospitalization (for example, facilities in the patient's room) or those related to the organization and service providers.

The Framework and Process of Personalization

The first step in personalization is collecting data. This is done through a variety of approaches. Features, interests, desires and needs of the user are collected directly and indirectly. Some applications directly capture user data through surveys, questionnaires, personal information logged in, and etc. In these cases, the content can be displayed to the user according to the user's selections and preferences. Some other applications, based on user activity and without direct intervention, create a user profile. In this case, the appropriate content is displayed to the user based on the number of times a webpage is visited (accessed) or tracking of clickthroughs on a website.[11]

Some researchers argue that personalization is a one-on-one marketing in which the organization must recognize its customers and treat them individually. They believe that personalization is a process that has four steps: 1) identifying potential customers; 2) determining customer needs and customer lifetime value; 3) interacting with the user to learn more about him/her; 4) delivering goods, services or content based on user characteristics.[12] Thus, it can be said that personalization has 4 dimensions: value dimension, knowledge dimension, dimensional orientation, and the quality of communication.

Customer value dimension is the most basic dimension to describe the difference between customers. This dimension takes into account all the quantitative-qualitative and internal-external criteria that complete the customer purchasing experience. In general, value analysis for the personalized offers should include content and interactive features, and then value refers to customer expectations of these two features.[13]

Knowledge dimension considers customer's knowledge and expertise, and degree of familiarity (for example, with the Internet). Multi-layer personalization increases the amount of information that needs to be processed by the customer, and this also reflects the difference among customers based on their information processing capabilities. So, with this in mind, you can have a deeper understanding of the customer. Researchers emphasize that in the digital environment, there is a relationship between the personality of the user, technology and the application of communication technology, and they can be categorized into 2 groups of novice and experts. Accordinly, users with different expertise levels have different information and communication needs.

The orientation dimension deals with customer mindset and experience. Researchers believe that people with different cognitive orientations have different needs and are looking for different benefits. They divide people into target-oriented and experience-oriented groups. Target-oriented people focus on transactions when they interact with an e-commerce website, avoiding social interactions and other contacts. On the other hand, experience-orineted people, besides addressing the target, also do other things like search.

The fourth dimension is the quality of a relationship. According to the researchers, satisfied customers always have long-lasting connections and long-lasting benefits. Sometimes good customer relationships can lead to asymmetry of the information regarding the competitors. In interaction with the customer, his consent and trust must always be monitored. It can be said that the quality of the relationship is the result of customer satisfaction and trust.[14]

In 2007, researchers presented a framework in which the ideal personalization process is introduced based on user and customer grouping based on the four above mentioned dimensions and personalized content offerings for each group. The proposed framework has a two-step approach for grouping customers. In the first step, customers are grouped according to the four dimensions introduced, and in the second step, the required value of the content placed as the grouping criteria. In fact, collecting information about the four dimensions mentioned above allows the user to provide personalized content and interaction to the user. Obtaining feedback from the user can improve decisions and lead to better communication.[15] There is another framework for personalization, according to which the personalization can occur both by the system and by the user. In this framework, content, interface or communication channel can be personalized. It should be taken into account that this framework determines that personalization can be done individually or for a group of people.[16]

Recommender Systems

Recommender systems are effective tools for helping decision makers choose items that fit their preferences and interests. These systems are used in many ways to personalize applications by recommending items such as books, movies, sound, news, articles, and more. For example, this system is used in e-commerce to learn from the customer and recommending products that are more valuable to the customer than other products. Amazon is the most famous recommender system in the field of e-commerce.[17]

Research on recommender systems as an independent research field started in the mid-1990s. The purpose of developing recommender systems was to reduce information overload through recycling the most relevant data and services from a large amount of data and therefore, the provision of personalized services to improve customer relationship management. The most important feature of a recommender system is its ability to guess the preferences and interests of the user through analyzing the behavior of the user in order to create customized recommendations. The most commonly used methods of recommendation are: collaborative filtering, content-based filtering, and knowledge-based methods. Each method has advantages and disadvantages. For example, while collaborative filtering has overspecialized recommendations, it involves problems of cold-start, scalability and sparseness. To solve these problems, more advanced recommender approaches, such as social network-based recommender systems, fuzzy recommender systems, context awareness recommender systems and group recommender systems have been addressed in the literature, and are considered as new and proposed topics.[18]

The motivations and goals of the various recommender systems can be summarized as Table 1.

According to the classification in the applications of the recommender systems fall into eight main domains: e-government, e-business, e-commerce, e-library, e-learning, e-tourism, electronic resources, and electronic group activities.

There are five phases to create a recommender system[19]: 1. Data collection; 2. File creation for the user; 3. Calculation of similarity; 4. Neighborhood selection; 5. Prediction and recommendation.

There are 2 ways to collect data about user preferences, implicit (system monitors user behavior in an indirect way) and explicit (users express their preferences or choices clearly and directly).[20] If the user file is a collection of properties extracted from the descriptions of the user-interested items, then the recommender system is content-based and if the user's file is a list of user-supplied ratings he/she provided previously for the items, the recommender system is collaborative.

Recommender systems have been extensively studied for providing the user with items such as movies, music, and favorite books, though the current recommender systems mostly focus on business activities. In recent years, the recommender systems in health area have been considered as a solution to many health problems.[21]

Recommender systems in healthcare area have been introduced as complementary tools in decision making processes in healthcare services and to reduce the information overload. Recommender systems are currently used also in health services for educational purposes, diet, and as personal health counseling tools. Therefore, recommender systems in healthcare area play an important role in filtering information for self-diagnostic searches of the users on the Web and the diagnostic and educational purposes of doctors.[22]

Table 1. Recommender Systems Goals and Motivations

Motivation	Goal
Which products will better meet customer preferences?	Recommend the products with the highest possible chance
Recommender systems	Recomment products suitable for the recommnder person
Which products can reach a high level of satisfaction after using it?	Recommend for products with higher satisfaction levels after use

Methods

Collaborative Recommender System

Collaborative filtering is a well-known algorithm whose prediction and recommendation is based on rating or the behavior of other users in the system. The main assumption behind this approach is that the opinions of other users can be selected and collected as a reasonable prediction provider from the target user preferences. In fact, if users agree on the suitability and quality of some items, they probably agree on other items. The range of information for the collaborative filtering system includes users who have expressed their preferences for different items. Preferences expressed by the user for an item is called as rating and is often indicated in triple (user, item, rating).[23]

In fact, the most common input for a collaborative recommender system is the users' rating of the items.[24] The collaborative recommender systems are usually assigned into one of the two categories of memory-based or model-based systems.[25] In memory-based filtering, the entire dataset is considered for prediction that is the entire dataset is scanned for finding a similar set of users for a particular user; then, the final rating of the user is predicted for an item, based on their preferences. The memory-based collaborative filtering method is divided into two categories of user-based and item-based. In the user-based approach, a similar group of users is searched for prediction, whereas the similar items are searched in the item-based method.[26]

A user-based collaborative filtering algorithm, finds users with similar preferences to the target user through the collected information and preferences from the users. Then the value of rating for the items which the user has not yet rated is predicted based on the ratings provided by his neighborhood.[27] Memory-based systems have a scalability problem. A solution to this problem is the model-based methods that uses rating data to train a model and then uses the created model to guide the recommender.[28] Developing the model is time consuming, but it greatly escalates production of recommendations. In model-based algorithms, machine learning methods such as clustering, Bayesian network, decision tree, neural networks, etc. are used for the recommendation.

Subject clustering has already been used in recommender systems.[29] With clusters of users, the common algorithms of collaborative filtering can be applied on clusters instead of the entire user-item matrix. By reducing the size of the user-item matrix and avoiding the dispersion problem, better recommendation results can be obtained regarding accuracy and improve the efficiency of the recommender algorithms in online mode. The application of the clustering methods is the reduction of dispersion and improvement of the scalability of the recommender system; because similarity is only calculated for users in the same clusters.

There are 2 approaches to use clustering in recommender systems: (1) Cluster-based and (2) Cluster-only.

The efficiency of the system increases in both approaches, since clustering stage is done offline, the first approach is more common and only focuses on improving time efficiency, and clustering is used to find target users' neighbors. Then, generating recommendations for them using collaborative memory-based filtering is conducted on the part of the input data which determines the most similar cluster. The final precision of the recommender is lower than that of the memory-based method. The second approach uses clustering as the main unit of the recommender system. By applying clustering to the input data, a model is created and then the calculations are done only on this model. The final accuracy of the recommender in this method is also lower compared to the memory-based methods.

In most research, users' data and item clustering have been used to re-use data, while other additional information such as the relationship among users (for example, friendship and trust) and the relationship among items (for example, referrals between publications) are ignored. It has been proven that additional information is useful in some specific areas of application. Therefore, a clustering method based on social relationships for recommending has been proposed. In this way, users are clustered on the basis of a social network created from different relationships. Methods for clustering networks can be divided into graph partitioning and blocking modeling or hierarchical clustering.

Another method, based on the K-means clustering algorithm, aims to cluster users based on their interests and on the voting system predictions are made. The matrix is considered as the input of the K-means algorithm and users are clustered based on their interests. In this method, instead of using the similarity criterion that is commonly used in refinement, Minkowski distance is used. The problem with this method is that this algorithm is effective in finding spherical clusters, but in some areas, such as e-commerce, there is no need for clusters to always be spherical. Another limitation of the algorithm is its noise sensitivity and that the number of clusters must be already specified.

Another proposed method is based on the Chameleon hierarchical clustering algorithm along with voting system to predict the customer's current rating for each item. Predicted ratings can be used later to decide if a particular item is being offered to the customer. In this way, users are clustered based on their characteristics (for example, age, gender, and occupation) using the Chameleon hierarchical clustering algorithm. Then, a voting system is used to predict a user's item. In a poll, one item is selected, then the number of each rank (1 to 5) given by the group's users is calculated, followed by the rank with the highest number for a particular user of that group. If the selected rank is good, then the item is offered to the user. Therefore, the voting system, by predicted ratings, helps to decide whether to recommend a product or not. The Chameleon hierarchical clustering algorithm is not sensitive to noise and can find clusters of different shapes, and clustering is based on relative proximity and relative correlation between clusters. Unlike the k-means algorithm, the number of clusters in the Chameleon hierarchical clustering algorithm is defined automatically.

Another proposed model uses a set of data mining techniques to develop a product recommender system for online retail customers. In this model, customers are firstly clustered on the basis of the characteristics of the RFM model, relying on a life-value-based segmentation approach, with relative consideration of preferences. Then, using a two-step suggestion structure, various offers are presented at two distinct levels from the product classification to each target customer. In the first step, using association rules method, the customer transactions of each cluster are investigated at the level of the product group and by extracting the hidden patterns and dependencies in the data, the valid recommendation rules are extracted and a list of proposed products is presented to each target customer. In the second step, using the cooperative refinement approach and considering the outputs of the previous step, customer preferences are proposed at the level of the items of the proposed products category.

In order to design a web-based recommender system, a new web-based search engine approach is proposed to automatically predict web pages according to user interests, which uses a fuzzy clustering algorithm to categorize similar user sessions; then, in order to extract the recommendation model, association rules method that describe relationship among the pages is used.

One of the most important problems for collaborative recommender systems that has been heavily considered by the researchers, is the cool start, since the new user has not yet rated any item, it is not possible to find users similar to him. In 3 following strategies are presented:

1. ***Using additional data sources:*** The main idea behind this category is to use additional resources such as demographic data, users' comments and social tags in order to be the best choice for new

user's neighbors.

2. ***Choosing the most prominent groups of analogous users:*** The idea behind this category is to improve the methods of identifying analogous users without the need for additional data sources and using clustering algorithms and decision tree.

3. ***Enhancing the prediction using hybrid methods:*** The idea behind this is to use hybrid methods to calculate similarities and predict ratings.

Demographic Recommendation Systems

The basis of the demographic recommendation systems (DRSs) dates backs to 1979. These systems have been used in everyday life. Also, the inherent problems of many recommender systems can be solved using DRSs. Age, gender, income, nationality, occupation, and other demographic information are essential for a large number of applications. For example, when offering movies, age groups are very important, whereas income is very important for proposing tourist destinations. In marketing, the purchasing needs of men and women are totally different, and it is not possible to target some products regardless of the intended user's gender. The DRS first classifies users based on their demographic characteristics; then uses user comments for system items. Both collaborative and demographic systems use user-to-user correlations, but based on different data. In demographic information-based recommender systems, there is no need for the rating lists to create a file for the user, and it is necessary only to calculate the similarity of the target user with other users. This will strengthen the system against the cool start problem of the new user. Since the user file in this system has fewer fields than the ratings, this system is fast; this is very important when the number of users is abundant. In other systems, the accuracy of the system depends to a large extent on the number of ratings, because more user ratings in the system lead to higher quality suggestions. Nonetheless, this is not the case in DRS, since the user file remains constant for a long time. On the other hand, DRS's main weakness is its sensitivity to security and privacy issues, especially for e-commerce applications. Also, recommendations from demographic groups may be very general. Demographic characteristics may have only one or a certain amount of value.

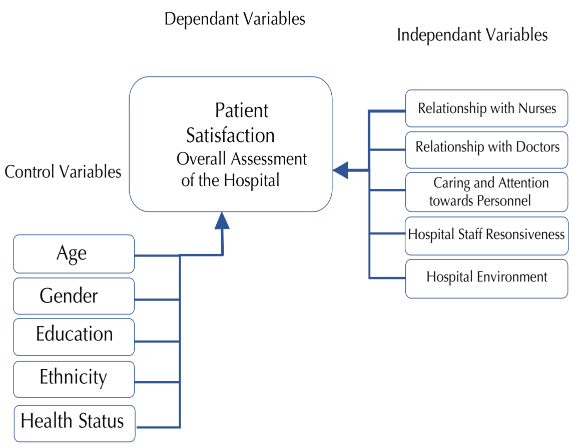

Figure 1. Conceptual Framework of Patient Satisfaction.

Suggested Method

In this study, using the five phases mentioned above to design a recommender system and based on the survey data of hospitalized patients, a model based on recommender systems is designed to increase the satisfaction of new patients from hospital services.

The purpose of designing this model is to increase the satisfaction of new patients and its users are hospital staff. After the discharge of patients from hospitals, their satisfaction from hospital services are evaluated. In this method, the views of other users can be selected and aggregated as providing reasonable predictions of target users' preferences. In recent years, healthcare recommender systems have also been used as complementary tools in decision-making processes in health services, therefore, according to the previous sections, in order to predict new patient's opinions and views, a user-based collaborative filtering method seems appropriate. Considering that new patients have not yet rated any service at the time of arrival, and their files only have the personal information, we face a cold startup problem due to the arrival of a new patient, where we can use additional data sources that is the demographic characteristics. Therefore, a hybrid method is used to design the model.

The purpose of designing this model is to identify the most important services affecting the overall satisfaction of patients in each cluster based on individual information

Table 2. Factors Affecting Patient Satisfaction

Demographic Features	Clinical and Support Services
✓ Age ✓ Sex ✓ Education ✓ Health status ✓ Race	✓ Skill, Attention, and the Doctor's manner ✓ Skill, Attention, and the nurse's manner ✓ Employees' attention and manner ✓ Reception and discharge process ✓ Environment and equipment ✓ Accessibility of the data ✓ Preclinical services ✓ Management factors

and using the views of patients in each cluster, and then by identifying the cluster to which the new patients are assigned, the rating of hospital services is predicted based on the views of the mentioned cluster. In order to increase the satisfaction of new patients based on their predicted opinions, a proper recommend is made about important services in the desired cluster (Figure 1). Finally, the prediction accuracy of the model is evaluated. In order to conduct this study, the WEKA software package was used in MATLAB program.

Innovations of the Proposed Model

According to the reviewed literature, most of the work conducted on the basis of customer ratings are in the field of recommender systems in business areas, e-learning and tourism. In recent years, recommender systems have also been used in the customized personaliztion. For example, there are recommender systems with educational purposes, diet, activity assistance, and recommendation of appropriate hospital using hospital ratings as well as location. Current collaborative recommender systems mostly focus on business activities and are investigated to provide items such as movies, music, and favorite books of the users.

In this model, a collaborative user-based filtering method is used to predict opinions and ratings. Considering that new patients have not yet received any service at the time of entry and their records are only personal information, a demographic method based on the recommendation of a new patient is used to solve the cold starting problem, and the calculation of similarity is based on the demographic data. Also, clustering method was used in order to improve the scalability of the model; thus, unlike the memory-based method, it is not necessary to determine the number of neighbors or thresholds of similarity.

The research originality includes:

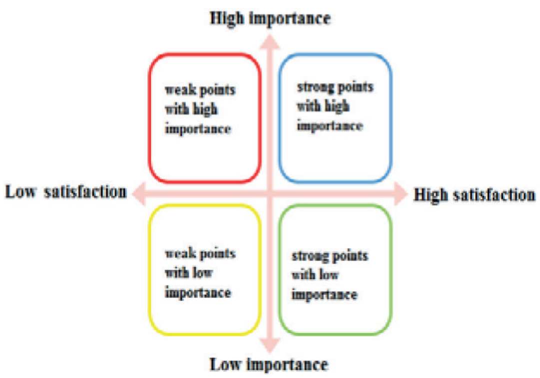

Figure 2. Satisfaction-Importance Matrix.

To use patients' opinions to rate hospitals and use the opinions of the discharged patients in order to predict the views of new patients and to make appropriate recommendations based on the predicted views for the hospital in order to provide appropriate and important services to the new patient. Users of this system are also service providers.

The second novelty is the use of a hybrid approach, which eliminates both the cold start problem and the scalability.

The third novelty is that in this model unlike other models, services with low level of prediction are included in the list of recommendations. Also, another criterion which is considered for making recommendation, is the importance of the dimensions for the cluster that the patient is assigned to.

Steps of Designing the Proposed Model

This consists of the following 4 steps.

First Stage

As previously stated, based on the conducted study in the factors affecting satisfaction are categorized in two categories of demographic characteristics and clinical and support services. Table 2 lists the variables of each category.

Numerous tools have been developed to measure patient satisfaction. The tools which have been repeatedly referred to in the literature to measure patient satisfaction, are surveys, critical incident technique, and questionnaire. Case studies, interviews and observation are also used to collect data.[30]

Second Stage

A cluster-based approach is used in the proposed model, and the discharged patients are clustered based on their personal characteristics. In fact, clustering is used to find new patients' neighbors. Recommendations for the new patient is conducted by applying collaborative memory-based filtering methods on a part of the input data as the determinant of the most similar cluster. In this study, K-means algorithm is used to perform clustering and the Silhouette Index and Davis-Bouldin Index are used to determine the number of suitable clusters.

After clustering, the importance or weight of service dimensions is determined on overall patient satisfaction. Weighting the process is important for identifying the chief subset of the features. Weighting the feature is important for rating the activity.[31] The weighting of a feature is the attribution of a weight (importance) to each feature, and can be a generalization of choosing the feature, that is, not assigning only binary weights to features, but allocating any arbitrary real number as weight.[32]

Relief algorithm is a weight-based method inspired by sample-based algorithm. This algorithm works well for good correlations, and its time complexity is linear and a function of the number of features and the number of samples in the sample set. Relief algorithm has been widely used as the most successful weighting feature for categorization. This algorithm is efficient for high-dimensional data and can correctly estimate the quality of features even when there is a nonlinear communication between features.

The original relief algorithm is used for numerical and nominal features; however, it can not work with incomplete data and is limited to two class problems. To deal with these issues, the Relief-F algorithm which is an extended version of this algorithm can be used. This version is not limited to two class issues, it is more stable and can work with incomplete and noisy data. In this algorithm, similar to the original version, a sample is selected randomly, then searches k closest neighbor of the same class as the closest collisions and k nearest neighbor of each different class as the closest failures. This algorithm updates the estimated quality for all features depending on their values for the selected sample and the closest collisions and failures. The fundamental difference between this algorithm and the original version is the selection of k collision and failure and more stability in relation to noise. K is a parameter which is determined by the user and controls the range of estimation.

Updating the weights of features after determining the closest collision and closest failure is such that the square of difference between the value of the desired feature in the selected sample and the sample of the closest collision of the weight is reduced from the weight of the feature and the square of the difference between the property value in the selected sample and the closest failure, is added to the weight of the feature. The larger the size of the weight, the desired attribute can better separate the instances of a class from others. Finally, the algorithm eliminates features that weigh less than or equal to one threshold > 0 and returns other features as the subset of the features of the response. The threshold value is set by the user, although it may be determined automatically by a function of the total number of attributes or determined by error and effort. It is also possible to remove features whose weight is negative.

Third Stage

Euclidean distance is used to determine the cluster to which the new patient is assigned; the new patient is assigned to a cluster which has the lowest Euclidean

Table 3: Variables of the study

	Variable	Categorization and Coding the values	Variable Type
Personal factors	Age	0 = under 14 years 1 = 14 to 29 years 2 = 30 to 60 years 3 = over 60 years old	Sequential discrete
	Gender	0 = male 1 = female	Numerical discretion
	Education	0 = illiterate and having the ability to read and write 1 = No university education 2 = university education	Sequential discrete
	Supplementary insurance	0 = No 1 = Yes	Numerical discretion
Service dimensions	1. Doctor 2. Nurse 3. Surgery ward 4. Reception 5. Diagnostic services 6. Environment 7. Management	Satisfaction of the aspects of each dimension: Sequential discrete values 4 = high 3 = average 2 = low 1 = never	Continuous (Satisfaction of any dimension)
Dependent variable	Overall satisfaction	3 = high 2 = average 1 = low	Sequential discrete

distance with it. In the proposed model, after identifying the cluster related to the new patient, in order to predict his views on service dimensions, first, using the cosine similarity (relation 1), the degree of similarity with the members of the cluster is calculated and then his opinions are predicted using equation (2). The prediction for the user is a goal based on estimating the average of the views of other users in the same cluster.

$$\cos\left(u_x,u_y\right)=\frac{\sum_{k=1}^{L}x_{i,k}\times y_{i,k}}{\sqrt{\sum_{k=1}^{L}x_{i,k}^{2}\sum_{k=1}^{L}y_{i,k}^{2}}} \quad (1)$$

$$pr_{x,k}=\frac{\sum_{u_y\in N_x}S\left(u_x,u_y\right)\times r_{y,k}}{\sum_{u_y\in N_x}\left|S\left(u_x,u_y\right)\right|} \quad (2)$$

After determining the cluster into which the new patient is assigned, the proper recommend is created based on how is the predicted rating for him as the most important dimensions of the cluster's service. In fact, a satisfaction-importance matrix can be used to make applied recommendations for the hospital. This matrix consists of four areas which are indicated in Figure 2.

Fourth Stage

Tool recommendation is usually for other purposes such as user satisfaction, increasing the sales, etc. Testing requires considering the goals and measuring the desired effect. Testing the algorithm on a real set of users and measuring the effects can be costly. In addition, measuring some of the desired effects may be impossible and difficult.[33]

The basis of most recommender systems is prediction, that is, the prediction of users' rating items or the possibility of using recommendations. It is assumed that the recommender system which provides more accurate predictions is preferred by users. Therefore, many researchers are looking for algorithms that provide better predictions. The accuracy of prediction is usually independent of the user interface and can therefore be measured in an offline assessment using the previously collected data set from users' ratings.[34] Typically, the mean squared error and the mean absolute error are used to calculate the prediction accuracy. In to calculate RMSE and MAE for M_A, active user and N_i^P predicted for user ***i*** from equations 3 and 4 have been used respectively:

Table 4. Accuracy of Categorizing Different Methods Based on Importance Services in Each Cluster.

Method Cluster	Decision Tree (%)	Logistic Regression (%)	Neural Networks (%)
1	80.56	69.44	75
2	100	100	91.67
3	92.31	92.31	76.92
4	70	60	50
5	87.5	75	62.5
6	75	66.67	91.67
7	100	100	60
8	71.43	47.62	33.33
9	81.25	56.25	43.75
11	60	80	80
12	90	100	100

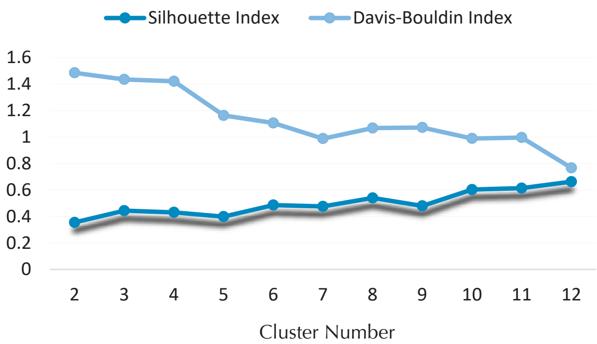

Figure 4. Determining the Proper Number of Clusters for the K-Means Algorithm

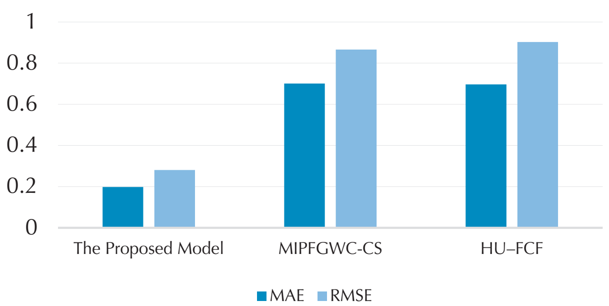

Figure 3. The Comparison of the Proposed Model's Accuracy.

Figure 5. The Importance of Services on Overall Satisfaction.

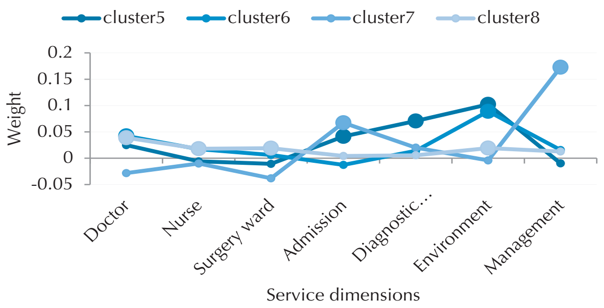

Figure 6. The Importance of Services on Overall Satisfaction.

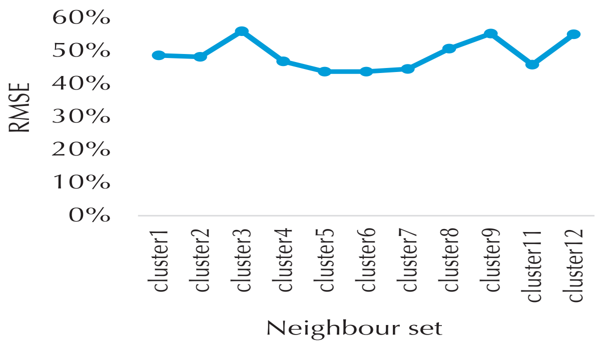

Figure 8. Prediction Accuracy Based on RMSE.

Figure 7. The Importance of Services on Overall Satisfaction.

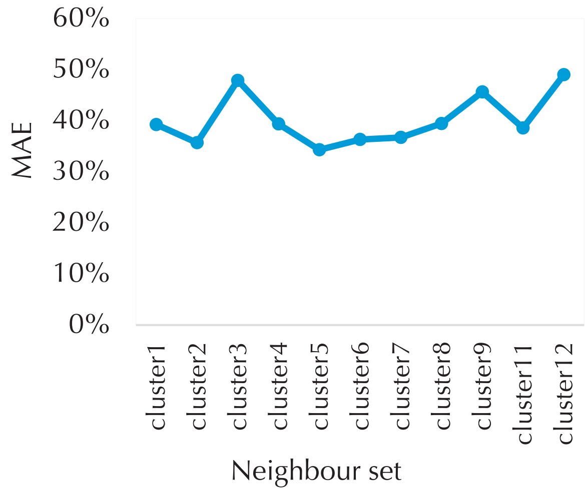

Figure 9. Prediction Accuracy Based on MAE.

$$RMSE = \frac{1}{M_A}\sum_{i=1}^{M_A}\left(\sqrt{\frac{1}{N_i^p}\sum_{k=1}^{N_i^p}\left(p_{i,k}-r_{i,k}\right)^2}\right)$$

$$MAE = \frac{1}{M_A}\sum_{i=1}^{M_A}\left(\frac{1}{N_i^p}\sum_{k=1}^{N_i^p}\left|p_{i,k}-r_{i,k}\right|\right)$$

The operation of the proposed model is in accordance with Figure B. Functioning of the model.

Comparison of the Proposed Model With Other Methods

Compared to memory-based collaborative filtering method, which takes into account the entire dataset for prediction purposes, in the proposed model, since the neighboring population of new patients is limited to the cluster to which it is assigned, this model is more effective in terms of scalability.

By using averaging of the precision obtained from each cluster, the overall accuracy of the model based on the MAE and RMSE criteria is 20% and 29%, respectively. In Figure 3, the values of these criteria for MIPFGWC-CS and HU-FCF algorithms mentioned in Figure 3, which were reviewed on the MovieLens dataset, are compared with the proposed model. Both algorithms, like the proposed model, use demographic data to solve a cold start problem.

Results

Result of Implementation of the Model

In order to implement the model, the data collected from the satisfaction forms filled out by 556 hospitalized patients discharged from Shariati hospital from 2012 to 2013. The collected data includes a range of demographic information and the patient's satisfaction with hospital services, which is rated from 1 to 4. In Table 3, the existing factors which are categorized by the hospital are mentioned in the patient satisfaction form. In order to prepare the data, the variables were also coded. Also, since the focus of the study is on the service dimensions, by averaging the scores given to the aspects of each dimension, the degree of satisfaction from each dimension is obtained. For testing the reliability of the data Cronbach's alpha coefficient of 0.877 was obtained, which is a desirable value.

Taking into account the four individual characteristics

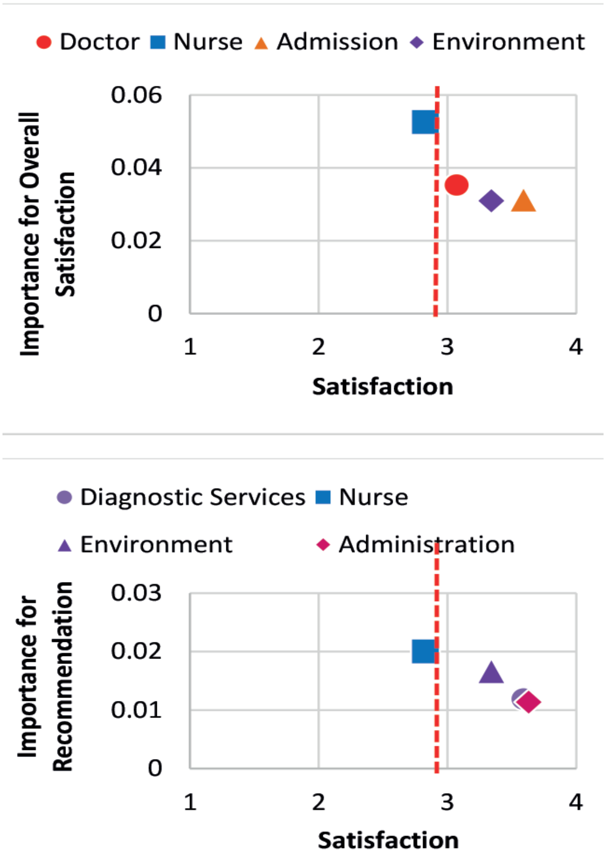

Figure 10. Status of Services of Greater Importance.
Abbreviations

in Table 3, the patients are clustered using the K-means method. According to 2 Silhouette and Davis-Bouldin indexes (Figure 4), the most suitable number of clusters was determined as 12. After creating a cluster model with 12 clusters, the SSE within cluster sum of squared errors was obtained as 178, which was lower in comparison to the low number of clusters. In Supplementary file 1, the created clusters are determined based on the characteristics of the centers and the number of their members, and in Figure C of tthe supplementary Figure B. and Table S1., the clustring result is shown as a tree.

By applying the Relief-F algorithm, the weight of the service dimensions in each cluster is determined which indicates the results of weighting of service dimensions for each cluster in Figures 5-7 (each 4 clusters in a diagram). Considering the average weights in each cluster as the threshold, services with a weight greater than the threshold value for each cluster are specified. As stated, the threshold value should be positive; therefore, in some clusters that the average value was negative, the threshold was considered as zero.

Considering these weights, hospitals prioritize services, and then they need to take corrective actions by predicting the views of new patients to improve and reinforce related aspects that are placed in the second quarter of satisfaction-importance matrix.

$$RMSE = \frac{1}{M_A}\sum_{i=1}^{M_A}\left(\sqrt{\frac{1}{N_i^p}\sum_{k=1}^{N_i^p}\left(p_{i,k}-r_{i,k}\right)^2}\right)$$

Evaluation of the accuracy of the prediction and implementation of the model was based on data on satisfaction forms from 520 clearance patients from Shariati hospital in Tehran. From the existing data set, about 30% (150 records) was randomly selected as the test data. In Figures 8 and 9, the prediction accuracy was indicated based on the neighboring set of cluster test data.

In order to verify the accuracy of the model in identifying the more important dimensions in each cluster, 148 records were randomly selected as the test data from among the available data; then the data for each cluster were determined and three methods of logistic regression, neural networks and Decision tree were applied to each cluster. Table 4 indicates the accuracy of each of the three methods for each cluster.

Discussion and Conclusion

Today, providing good service is one of the basic concerns of all service organizations, and many companies use personalization to increase their loyal customers satisfaction. Marketers and managers should be sure that they provide the right product or service to the right person. In this study, we were also looking forward to see how the hospital can increase the satisfaction of new patients, by personalizing the services. To this end, using the weighting method, the feature of the significance of the examined service dimensions on the overall satisfaction of the different groups of patients was determined, and then a model based on the recommender systems was provided to increase the satisfaction of patients for the quality of the hospital services.

Using this model, the hospital can predict the new patients' opinions after discharge, and if their predicted rating level is low in comparison to the important service dimensions in the group they are assigned to, they will take the necessary actions to provide them with better services and thereby reduce their dissatisfaction after discharge. Using the averaging of the precision obtained from each cluster, the accuracy of the overall prediction based on two criteria of the mean squared error and the mean absolute error are 0.49 and 0.40, respectively. Those results can be compared to the method in Diagram 7 using the set additional data to solve the cold start problem. As it is clear, the accuracy of the proposed model is better. Also, the model's accuracy in identifying important service dimensions was investigated by three methods through testing the model on the test data. As indicated in Table 4, all three methods have a fairly high accuracy and indicate that the model correctly determines the most important service dimensions for each cluster.

In this study, it was also shown that for different groups of patients in Shariati Hospital, the importance of various dimensions of the services was different for overall satisfaction. For this purpose, the weighting feature was used and each dimension had a different weight. In Figure 10, the status of services of greater importance was shown based on the degree of satisfaction of the desired cluster. With this in mind, services of importance to which cluster satisfaction is less than three, should be considered by the hospital in order to increase their satisfaction, and the status of the service with the importance and satisfaction of more than three should be maintained.

Abbreviations

(SQ): Service Quality; (RS): Recommender Systems.

Competing Interests

The authors declare no competing interests.

Acknowledgement

The author thanks the hospital staff who participated in the study. We would like to thank the nursing, the administrative and secretarial staff of Shariati operating room for their contribution to the maintenance of patient record without which this project would have been impossible.

Supplementary Materials

Supplementary file 1 contains Table S1, Figure B and Figure C.

References

1. Roham M, Gabrielyan AR, Archer NP. Predicting the impact of hospital health information technology adoption on patient satisfaction. Artif Intell Med. 2012;56(2):123-135. doi:10.1016/j.artmed.2012.08.001
2. Carlucci D, Renna P, Schiuma G. Evaluating service quality dimensions as antecedents to outpatient satisfaction using back propagation neural network. Health Care Manag Sci. 2013;16(1):37-44. doi:10.1007/s10729-012-9211-1
3. Al-Abri R, Al-Balushi A. Patient satisfaction survey as a tool towards quality improvement. Oman Med J. 2014;29(1):3-7. doi:10.5001/omj.2014.02
4. Nasiripour A, Gohari M, Nafisi A. Branding and functional indices relationship in Mashhad Ghaem hospital. J Health Adm. 2010;13(41):15-20. [Persian].
5. Liang TP, Chen HY, Du T, Turban E, Li Y. Effect of personalization on the perceived usefulness of online customer services: a dual-core theory. J Electron Commer Res. 2012;13(4):275-288.
6. Šebjan U, Bobek S, Tominc P. Organizational factors influencing effective use of CRM solutions. Proc Technol. 2014;16:459-470. doi:10.1016/j.protcy.2014.10.113
7. Hafezi A, Hafezi A, Esmaeili H. Presentation of the proposed national model of customer satisfaction index. Tehran: Paper presented at the Management and Humanities Research Conference in Iran; 2016:31-45.
8. Beattie M, Lauder W, Atherton I, Murphy DJ. Instruments to measure patient experience of health care quality in hospitals: a systematic review protocol. Syst Rev. 2014;3:4. doi:10.1186/2046-4053-3-4
9. Hung SY, Hung WH, Tsai CA, Jiang SC. Critical factors of hospital adoption on CRM system: Organizational and information system perspectives. Decis Support Syst. 2010;48(4):592-603. doi:10.1016/j.dss.2009.11.009
10. Otani K, Herrmann PA, Kurz RS. Improving patient satisfaction in hospital care settings. Health Serv Manage Res. 2011;24(4):163-169. doi:10.1258/hsmr.2011.011008
11. Sunikka A, Bragge J. Applying text-mining to personalization and customization research literature – Who, what and where? Expert Syst Appl. 2012;39(11):10049-10058. doi:10.1016/j.eswa.2012.02.042
12. Adomavicius G, Tuzhilin A. Personalization technologies: a process-oriented perspective. Communication of the ACM. 2005;48(10):83-90.
13. Chen YS, Cheng CH, Lai CJ, Hsu CY, Syu HJ. Identifying patients in target customer segments using a two-stage clustering-classification approach: a hospital-based assessment. Comput Biol Med. 2012;42(2):213-221. doi:10.1016/j.compbiomed.2011.11.010
14. Pai YP, Chary ST. Measuring patient-perceived hospital service quality: a conceptual framework. Int J Health Care Qual Assur. 2016;29(3):300-323. doi:10.1108/ijhcqa-05-2015-0069
15. Surprenant CF, Solomon MR. Predictability and personalization in the service encounter. J Mark. 1987;51(2):86-96. doi:10.2307/1251131
16. Minvielle E, Waelli M, Sicotte C, Kimberly JR. Managing customization in health care: a framework derived from the services sector literature. Health Policy. 2014;117(2):216-227. doi:10.1016/j.healthpol.2014.04.005
17. Son LH. Dealing with the new user cold-start problem in recommender systems: a comparative review. Inf Syst. 2016;58:87-104. doi:10.1016/j.is.2014.10.001
18. Lu J, Wu D, Mao M, Wang W, Zhang G. Recommender system application developments: a survey. Decis Support

Syst. 2015;74:12-32. doi:10.1016/j.dss.2015.03.008
19. Al-Shamri MYH. User profiling approaches for demographic recommender systems. Knowl Based Syst. 2016;100:175-187. doi:10.1016/j.knosys.2016.03.006
20. Liang TP, Lai HJ, Ku YC. Personalized content recommendation and user satisfaction: theoretical synthesis and empirical findings. J Manag Inf Syst. 2006;23(3):45-70. doi:10.2753/MIS0742-1222230303
21. Wiesner M, Pfeifer D. Health recommender systems: concepts, requirements, technical basics and challenges. Int J Environ Res Public Health. 2014;11(3):2580-2607. doi:10.3390/ijerph110302580
22. Sezgin E, Özkan S. A systematic literature review on Health Recommender Systems. Iași, Romania: IEEE; 2013:12-19. doi:10.1109/EHB.2013.6707249
23. Ekstrand MD, Riedl JT, Konstan JA. Collaborative filtering recommender systems. Human–Computer Interaction. 2011;4(2):81-173.
24. Vozalis M, Margaritis KG. Collaborative filtering enhanced by demographic correlation. Web Intelligence and Agent Systems: An international Journal. 2006;117-138.
25. Chen T, He L. Collaborative Filtering Based on Demographic Attribute Vector. International Conference on Future Computer and Communication. Wuhan: IEEE; 2009;31-43. doi:10.1109/FCC.2009.68
26. Gupta U, Patil N. Recommender system based on hierarchical clustering algorithm chameleon. IEEE International Advance Computing Conference (IACC); 2015:1006-1010.
27. Mobasher B, Burke R, Sandvig JJ. Model-based collaborative filtering as a defense against profile injection attacks. AAAI'06 proceedings of the 21st national conference on Artificial intelligence; 2006:1388-1393.
28. Pham MC, Cao Y, Klamma R, Jarke M. A clustering approach for collaborative filtering recommendation using social network analysis. J Univers Comput Sci. 2011;17(4):583-604.
29. Kużelewska U. Clustering algorithms in hybrid recommender system on movielens data. Stud Log Gramm Rhetor.2014;37(1):125-139.
30. Debono D, Travaglia J. Complaints and patient satisfaction: a comprehensive review of the literature. Centre for Clinical Governance Research in Health, University of New South Wales; 2009:145-156.
31. Yu H, Oh J, Han WS. Efficient feature weighting methods for ranking. 18th ACM Conference on Information and Knowledge Management; 2009:311-349.
32. Robnik-Šikonja M, Kononenko I. Theoretical and empirical analysis of ReliefF and RReliefF. Mach Learn. 2003;53(1-2):23-69. doi:10.1023/a:1025667309714
33. Divandari A, Delkhah J. Modeling and designing a model for measuring customer satisfaction in the banking industry and measuring the satisfaction of customers of Mellat Bank based on it. Quarterly Journal of Business Research. 2005;37(3):185-223.
34. Shani G, Gunawardana A. Evaluating recommendation systems. Microsoft Research; 2009:310-324.

Please cite this article as:
Kargari M, Akbari K. Providing a Model Based on Recommender Systems for Hospital Services: Case of Shariati Hospital of Tehran. Int J Hosp Res. 2018;7(1):36-48. doi:10.34172/ijhr.2018.xx.