# Evaluation of service quality using SERVQUAL scale and machine learning algorithms: a case study in health care



Serkan Altuntas
*Department of Industrial Engineering, Yıldız Technical University, Istanbul, Turkey*

Türkay Dereli
*Office of the President, Hasan Kalyoncu University, Gaziantep, Turkey, and*

Zülfiye Erdoğan
*Department of Industrial Engineering, Iskenderun Technical University, Iskenderun, Turkey*

Received 8 October 2020
Revised 22 January 2021
Accepted 21 February 2021

## Abstract

**Purpose** – This study aims to propose a service quality evaluation model for health care services.

**Design/methodology/approach** – In this study, a service quality evaluation model is proposed based on the service quality measurement (SERVQUAL) scale and machine learning algorithm. Primarily, items that affect the quality of service are determined based on the SERVQUAL scale. Subsequently, a service quality assessment model is generated to manage the resources that are allocated to improve the activities efficiently. Following this phase, a sample of classification model is conducted. Machine learning algorithms are used to establish the classification model.

**Findings** – The proposed evaluation model addresses the following questions: What are the potential impact levels of service quality dimensions on the quality of service practically? What should be prioritization among the service quality dimensions and Which dimensions of service quality should be improved primarily? A real-life case study in a public hospital is carried out to reveal how the proposed model works. The results that have been obtained from the case study show that the proposed model can be conducted easily in practice. It is also found that there is a remarkably high-service gap in the public hospital, in which the case study has been conducted, regarding the general physical conditions and food services.

**Originality/value** – The primary contribution of this study is threefold. The proposed evaluation model determines the impact levels of service quality dimensions on the service quality in practice. The proposed evaluation model prioritizes service quality dimensions in terms of their significance. The proposed evaluation model finds out the answer to the question of which service quality dimensions should be improved primarily?

**Keywords** Health care, Service quality, Evaluation model, Classification, Machine learning, Case study

**Paper type** Case study

The authors would like to thank the three anonymous reviewers for their insightful comments and suggestions that have significantly improved the paper.



0368-492X
DOI 10.1108/K-10-2020-0649

## 1. Introduction

Nowadays, evaluating, determining and improving the service quality are some of the most crucial fields of study, regardless of the sector. Hitherto, various methods have been suggested for several purposes in a wide range of sectors, such as transportation (Deb and Ahmed, 2018; Lee and Yu, 2018), health care (Aggarwal *et al.*, 2018; Cullen *et al.*, 2018; Jennings *et al.*, 2015), economy (Fragoso and Espinoza, 2017) and web services (Oriol *et al.*, 2014; Sá *et al.*, 2016; Somu *et al.*, 2018). However, artificial intelligence-based methods have the highest preferability among the approaches in the literature. Moreover, machine learning algorithms are one of the most preferred algorithms in the field of artificial intelligence. The studies that are based on machine learning algorithms generally focus on the economy, agriculture, health care and engineering. Many approaches, such as experimental design and quality control charts have been proposed to examine service quality in health care in the literature that considers their advantages and disadvantages. The main advantage of the approaches, which are used for data processing, is that they have a dynamic structure. These methods enable to analyze all of the data so that they do not disregard any items and samples that could be significant. Because of these features, the data processing methods come to the forefront since they are considered the most appropriate method in sectors such as health care services where the elimination of the errors is vital.

Service quality has become a significant issue since the service industries started competing for traditional sectors such as manufacturing and production (Javed *et al.*, 2019). The ultimate goal of service systems is to meet and exceed customer requirements and to increase service quality in practice (Altuntas and Kansu, 2019). The proposed service quality evaluation model based on the service quality measurement (SERVQUAL) scale and machine learning algorithm in health care help decision-makers and managers to fulfill this ultimate goal. Providing services to patients on the basis of their expectations and needs is a necessary and important step in offering high-quality services for the success of an organization to remain competitive in the market (Aghamolaei *et al.*, 2014). Hospitals have a very strategic role in accelerating the enhancement of public health (Kadir *et al.*, 2017). The demand for better service quality is rising due to the increased aspiration level of customers with an increase in their per capita income (Singh and Prasher, 2019). The dimensions that lead to unsatisfied customers can be easily defined through the SERVQUAL scale (Altuntas and Kansu, 2019). The SERVQUAL scale, which is a comprehensive service quality measurement scale, is empirically examined for its potential usefulness in a setting of hospital service (Babakus and Mangold, 1992). Hence, the SERVQUAL scale is extensively used in the health care service quality assessment (Pekkaya *et al.*, 2019). Hospitals, in particular, aim to provide excellent clinical care and quality services to their patients for providing high-quality services, which is of key importance in the management of service organizations (Teshnizi *et al.*, 2018). Level of the patient satisfaction could help decision-makers and managers to identify specific areas of improvement in public sector hospitals (Hussain *et al.*, 2019). In addition, public hospitals play a key role in Turkey for enabling the access of population to health services. Therefore, the proposed approach in this study is performed to a public hospital in Turkey. A service quality evaluation model, which is based on the SERVQUAL scale and ensemble machine learning algorithm in health care services, has not been conducted in the literature so far.

Similar to the studies that use only the SERVQUAL scale for the measurement of service quality, the use of the SERVQUAL scale has been combined with various methods including, multi-criteria decision-making methods (Ocampo *et al.*, 2019; Singh and Prasher, 2019) and fuzzy logic (Behdioglu *et al.*, 2019; Riono, 2017) to increase the efficiency for the measurement of the service quality in the hospitals. Therefore, in this study, a service quality evaluation model based on SERVQUAL scale and machine learning algorithms was proposed for health care services.

The primary contribution of this study is threefold. These are presented as follows:

- The proposed evaluation model determines the impact levels of service quality dimensions on the service quality in practice.
- The proposed evaluation model prioritizes service quality dimensions in terms of their significance.
- The proposed evaluation model finds out the answer to the question of which service quality dimensions should be improved primarily?

In this study, it is not only discussed measurement of service quality. The prioritization of improvement activities is considered. This study has two main aims. The first aim is to propose factors that will increase general service quality for the improvement activities to managers in health care. The second aim is to test the validity of the use of machine learning algorithms to predict the service quality. The patient testimonials are considered for the improvement activities to analyze items that affect service quality. The item scores are determined using patient testimonials. Afterward, the effects of these items on the general service quality are detected. These items are collected under various factors using factor analysis. The factors that have the highest gap value between effect value on the general service quality and item score value are evaluated in detail. Thus, the factors that would increase the service quality at the highest level are determined for improvement activities. As a result, the budget allocated for the improvement activities can be directed to the factors that will increase the service quality at the highest level.

In addition, a sample of service quality classification model based on ensemble machine learning algorithms was performed for health care services in this study. Unlike other service quality evaluation models, machine learning techniques have a dynamic structure. As long as the data flow keeps on, the model improves itself continuously. Machine learning techniques are interested in generating algorithms and computer systems that machines can learn from previous experiences (Izenman, 2008). Ensemble machine learning techniques train multiple learners that can solve the same problems. Because, these techniques aim to obtain an ensemble global model that achieves more reliable forecasts (Erdoğan, 2017; Maimon and Rokach, 2005). Because of these features, the use of the ensemble machine learning techniques generally provides better results than those of individual machine learning techniques. Machine learning techniques have three functions, including classification, clustering and association rules. In this study, classification algorithms are used to develop a service quality evaluation model in health care. The classification is a process of constructing a model that identifies and categorizes data classes or concepts to forecast the classes of objects with unknown class labels (Han and Kamber, 2001). The classification techniques are supervised learning algorithms. These algorithms have advantages over clustering and association rules, which are unsupervised learning algorithms. Because the performance values of models obtained using supervised learning algorithms can be calculated in practice. Thus, the best-fitted algorithm for the data set can be determined considering these performance values. There are values of output in the data that is used in this study. Therefore, supervised learning algorithms are more suitable for the existing data.

The service quality of a hospital can be evaluated using machine learning technique. Through this application, the patients could be able to make their hospital preferences more accurately and hospital managers could be able to assess satisfied or dissatisfied patient masses. Moreover, the hospital managers would be able to determine the service gaps that need to be improved through assessing these results. The service quality evaluation model, which has been developed in this study, is beneficial both for patients and service providers. The service quality evaluation model enables patients to determine the most appropriate hospital without experiencing the provided service quality. In addition to this advantage, the hospital management saves time for the necessary improvements and prevents the occurrence of any possible dissatisfaction. It is expected that the results of this study would guide patients, companions and hospital managers to provide satisfactory service quality.

The rest of the paper is organized as follows. The literature review is provided in Section 2. The proposed approach is introduced in Section 3. The case study is explained in Section 4. The results of the proposed approach are given in Section 5. Finally, conclusions are provided in Section 6.

## 2. Literature review

### *2.1 Service quality*

To maintain the existence of a service system in an increasingly competitive environment, companies need to ensure customer satisfaction. Customer satisfaction can be achieved in companies those having an adequate level of service quality. To increase customer satisfaction, customer requirements should be considered by using several tools, such as quality function deployment (Parezanović *et al.*, 2019) and SERVQUAL scale. In the literature, the studies related to service quality were broadly carried out in various sectors. Among these studies, Lee and Yu (2018) used user-generated online survey data to assess airport service quality based on Google reviews and performed sentiment analysis. Fragoso and Espinoza (2017) analyzed the service qualities of two banks using a modified version of the SERVPERF mode. They examined the service quality using samples obtained from the branches in four cities in Mexico. Deb and Ahmed (2018) aimed to explore the service quality of the city bus by taking perceptions and expectations of the users and data was analyzed by a combination of statistical tools comprising of factor analysis, linear regression analysis and structural equation modeling. Sá *et al.* (2016) developed a methodology to assess the qualities of local e-government online services based on an empirical study using the Delphi process. Somu *et al.* (2018) used multi-level hypergraph coarsening based robust heteroscedastic probabilistic neural networks to forecast the reliability of the service applications based on cloud technologies. Oriol *et al.* (2014) analyzed 47 quality models of web services from 65 papers to evaluate the state of the art of the proposed quality models for web services. Berry *et al.* (2019) intended to find the answer to the question of how do customers perceive the organization following the service and discussed the key concepts of service organization brand, namely, service marketing and service quality, to answer the question.

Innovative applications such as the use of information technologies highly influence the quality provided in health services. The use of information systems have provided a remarkable contribution to health care services (Mudavadi *et al.*, 2016). In the literature, there are various studies related to the use of information technologies in health services. Topacan *et al.* (2008) evaluated the determinants related to the adoption of health information services in practice. Behkami and Daim (2011) proposed an analysis model for an assessment of the adoption of health information technologies. Behkami and Daim (2012a) measured the effects of health information technologies on the delivery of care in patient-centered medical homes. Furthermore, Behkami and Daim (2012b) discussed the adoption of health information technology and highlighted that the use of health information technology provides lower cost and better patient experience. Behkami and Daim (2016) explored technology adoption in the case of the patient-centered medical home using structural equation modeling. They found that the use of health information technology is associated with lower cost and higher care quality. Mudavadi *et al.* (2016) used the analytical hierarchical process model to find the importance of perceived benefit, perceived ease of use and external factors concerning physicians' adoption of electronic health records.

The studies about service quality in health care services are given in Table 1.

Several studies on the topic of SERQUAL scale is given in Table 2.

*2.2 Machine learning algorithms*

When the literature is reviewed, it is noticed that machine learning techniques were widely used for the prediction of diseases in the health care services. Among these studies, Soni *et al.* (2011) used decision tree, Bayesian classification, K-nearest neighbors and neural networks classification for heart disease prediction and found that decision tree is a suitable method for heart disease prediction. Dangare and Apte (2012) conducted decision trees, Naïve Bayes and neural networks for heart disease prediction and found that neural networks provide accurate results as compared to decision trees and Naive Bayes. Besides, Vijayarani and Dhayanand (2015) predicted liver diseases using classification algorithms, namely, Naïve Bayes and support vector machine, and found that the support vector machine is a better classifier to predict liver diseases. Esteva *et al.* (2017) classified skin cancer with deep neural networks. Shah and Jivani (2013) compared three classification algorithms, namely, decision tree, Bayesian network and K-nearest neighbor algorithms, to predict breast cancer and found that Naïve Bayes is a superior algorithm. In addition to these studies, there are also studies dealing with data mining applications in the health care services and examining simple applications on these issues (Durairaj and Ranjani, 2013; Koyuncugil and Özgülbaş, 2019; Tomar and Agarwal, 2013). Furthermore, the use of large-scale data in health care has been discussed in the literature (Kaur and Wasan, 2006).

Several studies on the topic of data-driven analysis in service systems are given in Table 3.

As can be seen from the literature provided above, the literature, which is reviewed in this paper, is grouped into two parts: studies related to service quality and machine learning algorithms in health care. The studies in the first group are related to determining the quality of service in health and SERVQUAL scale. In the second group, there are studies with applications of machine learning algorithms in the field of health.

In this study, a model based on the SERVQUAL scale and machine learning algorithm is proposed to evaluate the service quality in health care services. Items that affect the service quality are determined in the proposed approach. Then, the mean score of these items is calculated. Subsequently, the impact of these items on service quality is determined. Thus, weak items, that is to say regarding poor health care services of the hospital, could be identified. Among these items, the items that have the highest impact on service quality are prioritized. Thus, the resources allocated for the improvement of related activities can be managed optimally. This evaluation model is important in terms of time, cost, and patient satisfaction. As far as we know, this study differs from the previous researches in that it uses from the machine learning algorithms in health care services based on the data obtained from patients by a survey. The proposed approach is a service quality evaluation model based on the SERVQUAL scale and machine learning algorithm in health care services. In this study, ensemble machine learning methods were used to eliminate the inadequate aspects of individual machine learning techniques and to enhance the evaluation performance. It should be noted that the ensemble machine learning techniques intends to establish a global ensemble model by putting the strengths of individual machine learning techniques to the forefront. Because of these features, the use of the ensemble machine

**Table 1.**
Studies about service quality in health care services

| Author(s) (year) | Method | Aim of study |
|---|---|---|
| Badrick *et al.* (2018) | The quality control techniques | To improve patient care quality |
| Roy *et al.* (2018) | A rough strength relational DEMATEL model | To analyze the key success factors of hospital service quality |
| Bonner *et al.* (2019) | Statistical analysis | To increase patient satisfaction and to assess the treatment quality |
| Jiang and Liao (2019) | A linear programming method based on probabilistic linguistic Kolmogorov-Smirnov distance | To evaluate the quality of service of the hospital |
| Martín-Martínez *et al.* (2019) | Delphi methodology | To evaluate the quality of care in the management of patients with rheumatoid arthritis |
| Mirzaeia *et al.* (2019) | Content validity, face validity and exploratory factor analysis | To measure consumers' perceptions of service quality in community pharmacies |
| Tuzkaya *et al.* (2019) | Interval-valued intuitionistic fuzzy-PROMETHEE | To evaluate the hospital service quality |
| Aburayya *et al.* (2020) | Principle component analysis, Pearson correlation coefficient, and multiple regression analyses | To examine the impact of TQM elements on hospital service quality |
| Alkafaji and Al-Sharmey (2020) | A fuzzy assessment model | To develop an assessment model based on fuzzy inference to assess the service quality |
| Bayati and Emadi (2020) | Panel data analysis | To investigate the factors affecting hospital death rate as a indicators of inpatient services quality |
| Firouzi Jahantigh and Ostovare (2020) | PROMETHEE-II and DEA | To evaluate the performance of teaching hospitals |
| S. Jiang *et al.* (2020) | A large group linguistic Z-DEMATEL approach | To determine key performance indicators in hospital performance management |
| Nemati *et al.* (2020) | HEALTHQUAL model | To compare hospital service quality based on the HEALTHQUAL model and trusting nurses at university and non-university hospitals |
| Shirazi *et al.* (2020) | FAHP-PROMETHEE hybrid approach | To rate hospitals in Sari city of Iran in terms of patient satisfaction during the outbreak of COVID-19 |
| Suresh *et al.* (2020) | Fuzzy logic approach | To measure the leanness of a hospital |
| Yucesan and Gul (2020) | Pythagorean fuzzy AHP and fuzzy TOPSIS | To evaluate the hospital service quality |

**Notes:** DEMATEL = The decision making trial and evaluation laboratory; DEA = data envelopment analysis; HEALTHQUAL = healthcare service quality and AHP = analytic hierarchy process

| Author(s) (year) | Method |
|---|---|
| Al-Neyadi *et al.* (2018), Ali (2018); Ali *et al.* (2018), Gullu *et al.* (2017); Nyandwe *et al.* (2017), Rehaman and Husnain (2018); Shuv-Ami and Shalom (2017), Ting *et al.* (2019); Singh *et al.* (2020a), Vanichchinchai (2020), Zarei *et al.* (2020) | SERVQUAL |
| Behdioglu *et al.* (2019) | Fuzzy SERVQUAL |
| Rasouli and Zarei (2016), Altuntas *et al.* (2020) | SERVQUAL and statistical quality control charts |
| Shafiq *et al.* (2017) | SERVQUAL and structural equation modeling |
| Singh and Prasher (2019) | Fuzzy SERVQUAL and fuzzy AHP |
| Souri *et al.* (2018) | Grey SERVQUAL |
| Alam and Mondal (2019) | SERVQUAL and AHP |
| Stevic *et al.* (2019) | SERVQUAL and BMW |
| Gundogdu and Kahraman (2021) | SERVQUAL and fuzzy AHP |
| Perera and Dabney (2020) | SERVQUAL, principal component analysis, confirmatory factor analysis and Gap analysis |
| Singh *et al.* (2020b) | SERVQUAL scale and net promoter score |
| Farhadi *et al.* (2020) | SERVQUAL, fuzzy DEMATEL and analytic network process (ANP) |
| Hatam *et al.* (2020) | SERVQUAL, DEMATEL and Andersen-Petersen (AP) |

**Note:** AHP = analytic hierarchy process

**Table 2.** Several studies on the topic of SERVQUAL scale

| Author(s) (year) | Method |
|---|---|
| Shah *et al.* (2019) | Deep learning approach |
| Akhyani *et al.* (2020) | Selectability/rejectability measures approach |
| Beura *et al.* (2020) | Associativity functional network, genetic programming and step-wise regression |
| Deng *et al.* (2020) | Multinomial logistic model, K-means algorithm and Markov chain model |
| Fattore *et al.* (2020) | Neural networks |
| Isak-Zatega *et al.* (2020) | Logistic regression method |
| Moro *et al.* (2020) | Text mining and topic modeling |
| Saleem *et al.* (2020) | Fuzzy AHP and fuzzy mean clustering |
| Shokouhyar *et al.* (2020) | Kano model, SERVQUAL scale and RFM clustering technique |
| Son *et al.* (2020) | Dynamic neural network and genetic algorithm |
| Tan and Yan (2020) | Linear regression, text classification and text pattern recognition |
| Vicente *et al.* (2020); Wang *et al.* (2020) | Fuzzy clustering approach |
| Lucini *et al.* (2020) | Text mining |
| Eldeeb and Mohamed (2020) | Latent class choice model and error components interaction model |
| Golmohammadi *et al.* (2020) | Neural networks, sensitivity analysis |
| Mukherjee *et al.* (2020) | linear discriminant analysis, K-means clustering |
| Rallis *et al.* (2020) | Unsupervised learning, classification and regression trees |

**Note:** RFM = recency, frequency, and monetary

**Table 3.** Data-driven analysis in service systems

learning techniques generally provides better results than those of individual machine learning techniques. Thus, evaluation results with a higher accuracy rate can be obtained. In addition to improving evaluation performance, the selection of the best performing algorithm is also crucial. Because specific performance values of different algorithms might give better results, hence, the synthesis index (SI) value (Chou *et al.*, 2014; Erdogan and Namli, 2019) was used to obtain a single algorithm with the best performance. Ensemble machine learning techniques and SI values have been preferred in this study, taking these situations into account.

## 3. The proposed approach

In this section, the proposed service quality evaluation model based on the SERVQUAL scale and machine learning algorithm in health care is introduced in detail. As seen in Figure 1, the proposed approach consists of four parts.

The first part of the proposed approach is the preparation and preliminary test. This part consists of three steps. First, the SERVQUAL scale was established. Afterward, a pilot study was carried out. A preliminary test was performed to increase the comprehensibility of the participants throughout the application of the survey. Within the scope of the pilot

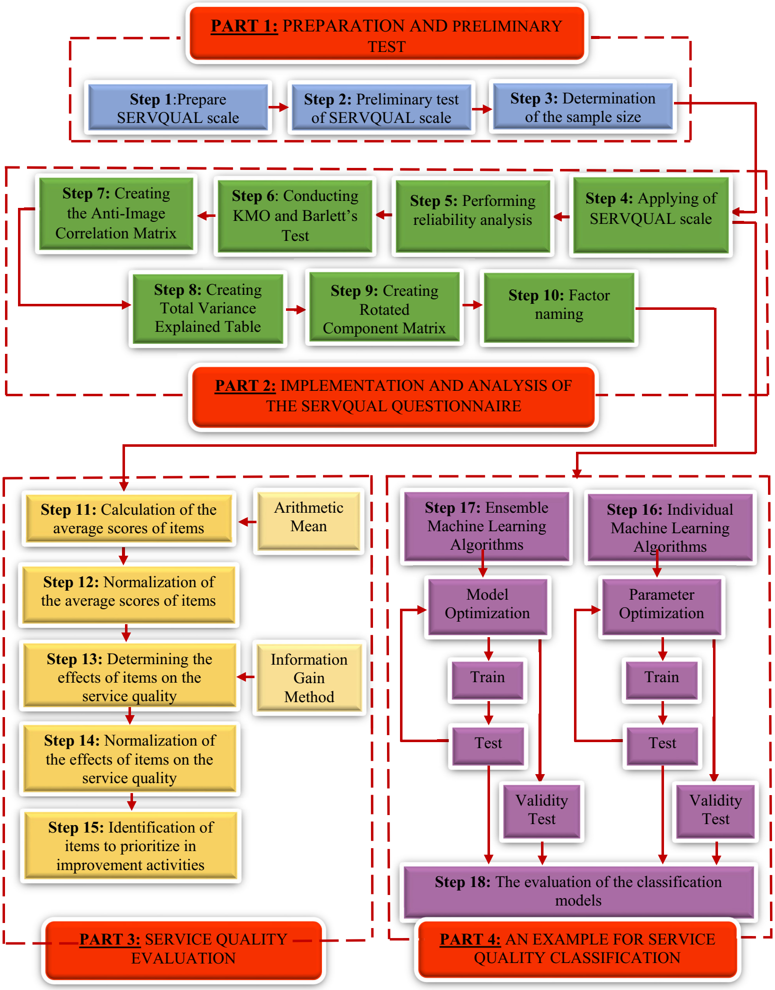


**Figure 1.**
Proposed approach

study, 50 people participated in the survey. Finally, the sample size was determined. The second part of the proposed approach is the implementation and analysis of the SERVQUAL questionnaire. This part consists of seven steps (Step 4–Step 10). In Step 4, the SERVQUAL scale was applied. In Step 5, the reliability analysis of the applied SERVQUAL scale was conducted. In Step 6, Kaiser–Meyer–Olkin (KMO) measure of sampling adequacy and significant Bartlett's test of sphericity was performed in SPSS. This test reveals whether factor analysis is required. Furthermore, the suitability of the sample adequacy for factor analysis is determined using this test. In Step 7, the image correlation matrix was generated to decide the suitability of the study for factor analysis. It is examined whether the values in the diagonal of this matrix are less than 0.5. Values that are less than 0.5 must be removed from the scale. In this study, principal component analysis with correlation matrix and varimax method as the rotation method were used for factor analysis. In Step 8, the total variance explained table were created. This table shows that how many factors the scale will involve. Subsequently, rotated component matrix was created in Step 9. In the last stage of the second part, factor naming was carried out to represent the factors obtained in the last stage of the second part optimally. The third part of the proposed approach consists of five steps (Step 11–Step 15). First, the mean scores of items were calculated in Step 11.

The calculated average scores of items were normalized from 0 to 1 in Step 12. In Step 13, the information gain method was used to determine the effects of items on service quality. Information gain method is an entropy-based feature selection method that is widely used in the field of machine learning. The application motivation of the information gain method is to maximize information between the class label and the given features (Dhir *et al.*, 2007; Lei, 2012). In this study, class label and features are service quality and items, respectively. After the effects of items on service quality are calculated, these values are normalized between 0 and 1 in Step 14. Thus, items to be prioritized for recovery activities are determined in Step 15. Items with a high impact on service quality and low average scores are prioritized.

A sample classification model is conducted in the last part. The data used in this study have class labels. Therefore, this data type is suitable for the use of supervised learning methods. Individual classification algorithms including various algorithms such as decision trees, statistical algorithms, artificial neural networks are used in Step 16. A trial and error method based on the use of different parameter settings was used to select the best performing algorithm among all these individual classification algorithms. Besides, the grid search method was used to determine parameter settings. According to these test results, the individual classification algorithms that give the best performance on the available data are reduce error pruning (REP) tree, random tree and J48 algorithms. Decision tree algorithms are individual classification algorithms that are most suitable for the data type used in this study. The advantages of decision trees are that they are comprehensible and interpretable. In this way, the reliability of the model for diagnostics can be verified by using both test data and expert knowledge (Yan *et al.*, 2016). Furthermore, the J48 algorithm ignores missing values and missing values are estimated using attribute values of other records (Patil and Sherekar, 2013). Thus, the quality of the classification performance is increased. Decision tree algorithms can run on numerical and categorical data. Decision tree algorithms can work efficiently in little data preparation and also can perform well when the number of features is big and unstable (Erdogan and Namli, 2019). Hence, ensemble machine learning methods were preferred in this study, to eliminate the missing parts of individual machine learning techniques and to improve the classification performances of these techniques. Ensemble machine learning techniques train multiple learners that can solve the same problems. Because these techniques aim to obtain an ensemble global model that achieves more reliable forecasts (Erdoğan, 2017; Maimon and Rokach, 2005). Because of these features, the use of the ensemble machine learning techniques generally provides better results than those of individual machine learning techniques. Random subspace and multi-class classifier algorithms that have the best estimation performance on the data used in the study among ensemble machine learning techniques were preferred in Step 17. Weston and Watkins (1998) and Vapnik (1998) proposed a multi-class classifier method in the literature. The multi-class classification approach can be examined under two titles. The first title includes algorithms that can be extended to handle multi-class cases. The second title includes methods that involve the reduction of multi-class classification problems to binary ones. The random subspace method (*RSM*) was proposed by Ho (1998). The RSM is based on the selection of a random feature subset in the training of each ensemble classifier. This method relies on an autonomous operation to select small number sizes from a given features field randomly. In each iteration, a selection is made and a subspace is fixed. Then, all samples are reflected in this subspace and a classifier is trained using the anticipated training examples. The methods with the highest classification value are preferred as sub-classifiers. These sub-classification methods are rep tree, random tree and J48 algorithms. *The J48 algorithm* is a learning algorithm equipped with additional features for solving problems that the ID3 algorithm cannot overcome. This algorithm was proposed by Quinlan (1996). *Random tree algorithms* are a collection of tree models. The collection of the tree models is called a forest. The random tree is performed to obtain a tree considering *k* items randomly determined at each node. Random selection means that the probability of selection for each tree in the forest has a uniform probability (Zhao and Zhang, 2008). The REP tree algorithm is one of the fastest classification algorithms. The information gain is used as the splitting criterion to obtain a tree (MeeraGandhi, 2010; Zhao and Zhang, 2008). The data is divided into training and test data during the model creation (Han and Kamber, 2001). The various methods have been used to divide data into two parts as test and training data. In this study, a 90% split ratio and 10-fold cross-validation methods were used. In the 90% split ratio method, 90% of the data set was determined as training data. The classification model was established using the 90% of the data set. Then, the remaining 10% of the data was used to predict the label value in the case study and to assess the classification model. In the K-fold cross-validation method, the data is divided into K different sets. These sets are approximately the same size. Subsequently, K-1 of K observations is used as a training set. After the model is established, an observation is used to test the obtained model. This procedure is run for each of the K observations (Fernandez, 2010; Maimon and Rokach, 2005).

The various individual classification algorithms are used to establish the classification model of the hospital quality in the literature. Ensemble classification algorithms are proposed in this study. The proposed ensemble classification algorithms are multi-class classifier and the RSM. In the literature, the performance values are used for the evaluation of the classification algorithms. The performance values used in this study are accuracy, precision, recall, f-measure, mean absolute error (MAE), root mean squared error (RMSE), relative absolute error (RAE), root relative squared error (RRSE) and SI in Step 18. These values are the parameters that are used in the selection of classification algorithms. However, these values are not adequate to evaluate the performance of algorithms. Different performance values of various algorithms give better results. Therefore, SI value was used to obtain a single algorithm with the best performance. SI value was calculated using equation (1) (Chou *et al.*, 2014):

$$SI = \frac{1}{m}\sum\nolimits_{i=1}^{m}\frac{P_i - P_{min,i}}{P_{max,i} - P_{min,i}} + \frac{1}{n}\sum\nolimits_{j}^{n} 1 - \left(\frac{P_j - P_{\min,j}}{P_{max,j} - P_{min,j}}\right) \tag{1}$$

where i shows measurements expected to give high value such that accuracy, precision, recall and f-measure. j implies measurements expected to give low value such that MAE, RMSE, RRSE and RAE.

## 4. Application of the proposed approach

In this section, a real-life case study, which is conducted in a public hospital in Kocaeli province of Turkey, is explained elaborately. To carry out the study, a protocol has been signed between the Kocaeli Provincial Directorate of Health and our University.

In the first part of the proposed approach, the SERVQUAL scale was prepared within the scope of the Preparation and preliminary test section. In the literature, the use of five-point Likert scale for the application of the SERVQUAL is suggested based on the management team's experience with previous surveys, which indicated that the five-point format would reduce the frustration level of the respondent patients, and would thereby increase the response rate and the quality of the responses (Babakus and Mangold, 1992). Service quality of hospitals is widely measured with scales that gauge patients' perspective (Shafiq *et al.*, 2017). Five-point Likert scale is extensively used and accepted in the literature for the application of the SERVQUAL to hospital service quality by many researchers, such as Altuntas *et al.* (2012), Altuntas and Yener (2012), Altuntas *et al.* (2020), Altuntas and Kansu (2019), Shafiq *et al.* (2017), Rai *et al.*, 2019; Aghamolaei *et al.* (2014) and Li *et al.* (2015). Therefore, a five-point Likert scale was used in the survey study. The items of the questionnaire were prepared based on the SERVQUAL scale. The items used in the questionnaire are presented in Table A1. Then, preliminary test was conducted for pilot application. The sample size was calculated using equation (2):

$$n = \frac{pX_{(1-p)}X\left(Z_{\propto/2}\right)^2}{e^2} \tag{2}$$

Where $n$ is required minimum sample size, $p$ is percentage picking a choice and e is error margin.

In this study, the sample size was calculated by taking 0.5 $p$-value for 95% confidence interval. $n$ value was found to be 385. Thus, 410 people were surveyed in the study. Of the surveys, 390 were found eligible to use in the study.

In the second section of the proposed approach, the SERVQUAL scale was applied to the Kocaeli Public Hospital in Turkey. A survey study was conducted to obtain data from patients in this hospital between August 15th, 2016 and October 28th, 2016. The survey was conducted in 14 separate department of the hospital and the data was obtained from 390 patients. These departments are that brain surgery, internal medicine, physical therapy and rehabilitation, general surgery, chest diseases, eye, cardiology, otorhinolaryngology, neurology, orthopedics and traumatology, plastic surgery, urology, chest cardiovascular surgery and infectious diseases. After the survey was conducted, reliability analysis was carried out. Cronbach's alpha value was found to be 0.931. Item-total statistics study was carried out to determine whether deleting an item will increase the reliability of the questionnaire. Item-total statistics is given in Table A2. As can be seen from Table A2, there is no need to remove any item from the questionnaire. KMO measure of sampling adequacy and significant Bartlett's test of sphericity was performed to determine whether factor analysis is required. As can be seen from Table 4, $p$ is less than 0.05.

| | | |
|---|---|---|
| KMO measure of sampling adequacy | | 0.890 |
| Bartlett's test of sphericity | Approx. chi-square | 8,613.281 |
| | df | 946 |
| | Sig. | 0.000 |

**Table 4.** KMO and Barlett's test

This value indicates that the test result is meaningful and factor analysis is required. The result of KMO measure of sampling adequacy was found to be 0.890. This means that the sample size is adequate for factor analysis. Anti-image correlation matrix was established to determine whether items were adequate for factor analysis. All values in the diagonal of this matrix were examined. There is no value less than 0.736. Considering the obtained results, it is concluded that the items are suitable for factor analysis. In this study, principal component analysis with correlation matrix and varimax method as the rotation method was used for factor analysis. Total variance explained is given in Table 5.

As can be seen from Table 5, the scale consists of 10 sub-sections. These dimensions were accounted for the 62.601% of the total variance. Subsequently, the rotated component matrix was generated. Rotated component matrix is presented in Table 6.

Finally, factor naming was carried out. As can be seen in Table 7, the items are allocated in 10 factors.

Mean scores for each item were also calculated based on a five-point Likert scale. Table 8 shows mean scores for items. As can be seen from Table 8, Item 1 (modern physical appearance and medical equipment) has the lowest mean score among the 44 items. While the highest mean score is recorded at Item 12 (employees have a neat appearance).

Then, the info gain method was used for the assessment of the effects of the items on the service quality. Table 9 shows the effects of items on the label value. Eight items (item no 18, 33, 26, 28, 16, 15, 12 and 22) did not have any contribution to the classification model. These ineffective items were removed from the data set. These ineffective items did not have a positive or negative impact on the classification performance.

Figure 2 illustrates the percentage effects of items on the service quality. The top 5 items having the highest effects on service quality are Item 43, Item 1, Item 25, Item 7 and Item 5.

It is crucial to focus on items by considering item percentage effects of items on the service quality (Figure 2) and the mean score of items (Table 8). Figure 3 shows the scores of patients on items and the effect of items on service quality. It should be noted that item scores are based on five Likert-scale. The percentage of gap for factors is given in Table 10.

In this study, it was considered not only item scores but also the effects of items on the service quality. The factors that have the highest value between item score and effect value on the service quality were proposed for improvement activities. First, item scores and the effects of items on the service quality were normalized between 0 and 1. Because these two values should be comparable. The reason why these two results were given together was so the hospital could notice more easily which item should be improved. Hospital managers can group patients based on the scores they give to the items and find out to which items the patient group attach importance most. As can be seen from Figure 3, the lowest scores are assigned to Items 1, 6, 9, 10, 23, 25, 27 and 38. These items have scores that are less than 4 out of 5. Which of these items needs to be improved first can be determined using the proposed approach. The effects of items shown with the red circle in Figure 3 are higher than the item score given by patients. Therefore, it can be concluded that the decision-makers should consider items shown with the red circle to improve service quality in the future. As it can be seen from Table 10, the items in Factors 4 and 7 should be improved to provide higher perceived service quality in practice. Items having the highest gap between the effect of item on the service and item score are 5 (Factor 4), 6 (Factor 4), 25 (Factor 1), 23 (Factor 2), 27 (Factor 2), 38 (Factor 8). The first 10 items having the highest gap between the effects of items on the service quality and item score among the items are 1, 25, 43, 7, 5, 9, 23, 10, 27 and 38. These items represent six factors, namely, Factor 1, Factor 2, Factor 4 and Factor 7, Factor 8 and Factor 9. Item 1 has the highest gap between the effects of items on the service quality and item score among the items. Hence, hospital managers and decision-

| Item | Initial eigenvalues Total | % of variance | Cumulative (%) | Rotation sums of squared loadings Total | % of variance | Cumulative (%) |
|---|---|---|---|---|---|---|
| 1 | 12.256 | 27.854 | 27.854 | 4.195 | 9.534 | 9.534 |
| 2 | 2.767 | 6.288 | 34.142 | 3.336 | 7.582 | 17.117 |
| 3 | 2.453 | 5.575 | 39.716 | 3.020 | 6.863 | 23.980 |
| 4 | 1.984 | 4.509 | 44.225 | 2.884 | 6.554 | 30.534 |
| 5 | 1.747 | 3.971 | 48.196 | 2.735 | 6.216 | 36.750 |
| 6 | 1.483 | 3.370 | 51.566 | 2.647 | 6.015 | 42.765 |
| 7 | 1.372 | 3.118 | 54.684 | 2.609 | 5.930 | 48.695 |
| 8 | 1.241 | 2.820 | 57.505 | 2.548 | 5.791 | 54.486 |
| 9 | 1.163 | 2.644 | 60.148 | 2.040 | 4.637 | 59.122 |
| 10 | 1.079 | 2.453 | 62.601 | 1.531 | 3.479 | 62.601 |
| 11 | 0.975 | 2.216 | 64.817 | | | |
| 12 | 0.970 | 2.205 | 67.022 | | | |
| 13 | 0.905 | 2.056 | 69.078 | | | |
| 14 | 0.843 | 1.915 | 70.993 | | | |
| 15 | 0.827 | 1.879 | 72.872 | | | |
| 16 | 0.780 | 1.773 | 74.645 | | | |
| 17 | 0.728 | 1.654 | 76.300 | | | |
| 18 | 0.692 | 1.572 | 77.872 | | | |
| 19 | 0.644 | 1.464 | 79.335 | | | |
| 20 | 0.612 | 1.390 | 80.725 | | | |
| 21 | 0.603 | 1.371 | 82.096 | | | |
| 22 | 0.576 | 1.309 | 83.405 | | | |
| 23 | 0.563 | 1.280 | 84.684 | | | |
| 24 | 0.550 | 1.250 | 85.934 | | | |
| 25 | 0.534 | 1.215 | 87.149 | | | |
| 26 | 0.488 | 1.108 | 88.257 | | | |
| 27 | 0.461 | 1.048 | 89.305 | | | |
| 28 | 0.451 | 1.025 | 90.330 | | | |
| 29 | 0.413 | 0.938 | 91.267 | | | |
| 30 | 0.383 | 0.870 | 92.138 | | | |
| 31 | 0.370 | 0.842 | 92.979 | | | |
| 32 | 0.355 | 0.808 | 93.787 | | | |
| 33 | 0.321 | 0.729 | 94.516 | | | |
| 34 | 0.306 | 0.696 | 95.211 | | | |
| 35 | 0.300 | 0.683 | 95.894 | | | |
| 36 | 0.282 | 0.642 | 96.536 | | | |
| 37 | 0.277 | 0.630 | 97.165 | | | |
| 38 | 0.253 | 0.576 | 97.741 | | | |
| 39 | 0.246 | 0.560 | 98.301 | | | |
| 40 | 0.191 | 0.435 | 98.735 | | | |
| 41 | 0.171 | 0.389 | 99.125 | | | |
| 42 | 0.163 | 0.370 | 99.494 | | | |
| 43 | 0.117 | 0.266 | 99.761 | | | |

**Table 5.**
Total variance explained

makers should focus on the physical appearance of medical equipment (Item 1) in the hospital.

In the literature, service quality was assessed by using the concordance and discordance tests (Nacer *et al.*, 2015), multi-criteria decision-making methods (AHP, Intuitionistic Fuzzy (IVIF)-Preference Ranking Organization Method for Enrichment Evaluation (PROMETHEE) IVIF-Technique For Order Preference by Similarity to Ideal Solution (TOPSIS), Fuzzy AHP, etc.) (Akdag *et al.*, 2014; Maghsoodia *et al.*, 2019; Mudavadi *et al.*, 2016; Shafii *et al.*, 2016;

| Item | Factor 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
|---|---|---|---|---|---|---|---|---|---|---|
| 20 | 0.795 | | | | | | | | | |
| 9 | 0.761 | | | | | | | | | |
| 21 | 0.677 | | | | | | | | | |
| 24 | 0.556 | | | | | | | | | |
| 26 | 0.539 | | 0.358 | | | | | | | |
| 32 | 0.455 | | 0.417 | | | 0.399 | | | | |
| 25 | 0.424 | | | | | | | 0.344 | | |
| 13 | 0.423 | | | | | | | | 0.392 | |
| 43 | 0.321 | | | | | | | | | |
| 22 | | 0.713 | | | | | | | | |
| 29 | | 0.705 | | | | | | | | |
| 30 | | 0.625 | | | | | | | | 0.352 |
| 23 | | 0.580 | | | | | | | | |
| 34 | | 0.528 | | | | 0.406 | | | | |
| 35 | | 0.487 | | | | | | | | |
| 27 | | 0.392 | | | | | | | | |
| 28 | | | 0.623 | | | | | | | |
| 12 | 0.321 | | 0.609 | | | | | | | |
| 41 | | | 0.580 | | | | | 0.440 | | |
| 4 | | | 0.577 | | | | | | | |
| 31 | | | 0.543 | | | 0.451 | | | | |
| 33 | 0.327 | | 0.444 | | | | | | 0.324 | |
| 6 | | | | 0.831 | | | | | | |
| 8 | | | | 0.786 | | | | | | |
| 1 | | | | 0.654 | | | | | | |
| 5 | | | | 0.592 | | | | | | |
| 15 | | | | | 0.905 | | | | | |
| 16 | | | | | 0.836 | | | | | |
| 14 | | | | | 0.794 | | | | | |
| 17 | 0.332 | | | | 0.403 | | | | | |
| 37 | 0.331 | | | | | 0.718 | | | | |
| 36 | | | | | | 0.664 | | | | |
| 18 | | | | | | 0.509 | | | −0.344 | |
| 11 | | | | | | | 0.875 | | | |
| 9 | | | | | | | 0.854 | | | |
| 10 | | | | | | | 0.823 | | | |
| 39 | | | | | | | | 0.777 | | |
| 38 | | | | | | | | 0.752 | | |
| 40 | | | | | | | | 0.568 | | |
| 42 | | 0.360 | | | | 0.408 | | 0.450 | | |
| 4 | | | | | | | | | 0.719 | |
| 7 | | | | 0.406 | | | | | 0.628 | |
| 2 | | | | | | | | | | 0.658 |
| 3 | | | | | | | | | | 0.553 |

**Table 6.** Rotated component matrix

Tuzkaya *et al.*, 2019), structural equation modeling (Safari *et al.*, 2019). However, the methodology used in these studies is static. The use of machine learning algorithms provides real-time assessment and dynamic evaluation of provided service quality. The classification model is constructed by using machine learning algorithms. Because these algorithms continue to improve themselves as long as the flow of the continues. Furthermore, it is well-documented in the literature that ensemble machine learning algorithms provide better results than

| Factors | Item no |
|---|---|
| 1: Being ready to serve in the hospital | 13, 19, 20, 21, 24, 25, 26, 32, 43 |
| 2 : Adequacy of health care staff | 22, 23, 27, 29, 30, 34, 35 |
| 3 : Act ethically | 12, 28, 31, 33, 41, 44 |
| 4: General physical conditions | 1, 5, 6, 8 |
| 5: Reliability in services | 14, 15, 16, 17 |
| 6: Accessibility | 18, 36, 37 |
| 7: Food services | 9, 10, 11 |
| 8: Information and communication | 38, 39, 40, 42 |
| 9: Cleaning | 4, 7 |
| 10: Physical condition of the rooms | 2, 3 |

**Table 7.**
Factor naming

| Item no | Score* | Item no | Score* | Item no | Score* | Item no | Score* |
|---|---|---|---|---|---|---|---|
| *1* | *3.6360* | *12* | *4.7872* | 23 | 3.8051 | 34 | 4.2231 |
| 2 | 4.4026 | 13 | 4.4462 | 24 | 4.5051 | 35 | 4.3538 |
| 3 | 4.0538 | 14 | 4.0487 | 25 | 3.9205 | 36 | 4.5128 |
| 4 | 4.2077 | 15 | 4.2487 | 26 | 4.6949 | 37 | 4.5436 |
| 5 | 4.1333 | 16 | 4.3667 | 27 | 3.8641 | 38 | 3.8205 |
| 6 | 3.9154 | 17 | 4.5231 | 28 | 4.7821 | 39 | 4.1103 |
| 7 | 4.1589 | 18 | 4.1333 | 29 | 4.4487 | 40 | 4.3487 |
| 8 | 4.1077 | 19 | 4.5026 | 30 | 4.4051 | 41 | 4.7487 |
| 9 | 3.6897 | 20 | 4.5795 | 31 | 4.7026 | 42 | 4.2513 |
| 10 | 3.9692 | 21 | 4.5180 | 32 | 4.6000 | 43 | 4.2000 |
| 11 | 4.0744 | 22 | 4.0436 | 33 | 4.6718 | 44 | 4.7333 |

**Note:** *Mean score (five-point Likert scale)

**Table 8.**
Mean scores of items

individual machine learning algorithms in practice, as ensemble machine learning algorithms could obtain optimum global models. However, despite all these advantages, ensemble machine learning algorithms have not been used for the evaluation of service quality of hospitals in the literature so far. Service quality evaluations performed by neglecting the interactions among the items cannot provide the necessary information to the hospital managers. However, the use of ensemble machine learning algorithms considers interaction among the items. Hence, these algorithms provide a comprehensive and factual evaluation of the service systems. Therefore, a sample case study related to use of machine learning techniques to the evaluation of the service quality in health care was given in the last section of the proposed approach, in the study. The data used in this study involve class label. The distribution of the label value among the participants is given in Table 11. Approximately, 30% of the patients assigned the highest scores to the quality of the hospital. A great majority of patients (46%) assigned 4 points to the quality of the hospital. When the scores given by the patients to the hospital quality were assessed in general, the mean score of the hospital was 4.02. This score indicates that the service quality of the hospital is good. Nevertheless, this service quality score is not adequate, as the analyzed sector is a health care sector. To enhance the score of service quality to 5, items with low scores should be identified. Among these items, improvements of the items that have the most effect on the label value should be prioritized.

The parameter settings providing the best performance values are given in Table 12. The classification models were established by considering these parameter settings. The grid

| Item no | Rank value | Item no | Rank value | Item no | Rank value | Item no | Rank value |
|---|---|---|---|---|---|---|---|
| 43 | 0.1769 | 36 | 0.0939 | 21 | 0.0711 | 31 | 0.0529 |
| 1 | 0.1588 | 19 | 0.0929 | 9 | 0.0697 | 14 | 0.0529 |
| 25 | 0.1526 | 10 | 0.0926 | 3 | 0.0686 | 41 | 0.0522 |
| 7 | 0.1425 | 37 | 0.0874 | 44 | 0.0639 | 18 | 0 |
| 5 | 0.1385 | 34 | 0.0836 | 6 | 0.0634 | 33 | 0 |
| 35 | 0.1339 | 8 | 0.0828 | 2 | 0.0628 | 26 | 0 |
| 30 | 0.1230 | 17 | 0.0819 | 27 | 0.0624 | 28 | 0 |
| 29 | 0.1154 | 11 | 0.0817 | 39 | 0.0591 | 16 | 0 |
| 42 | 0.1098 | 23 | 0.0768 | 4 | 0.0572 | 15 | 0 |
| 24 | 0.1073 | 32 | 0.0756 | 40 | 0.0555 | 12 | 0 |
| 20 | 0.0979 | 13 | 0.0722 | 38 | 0.0553 | 22 | 0 |

**Table 9.**
Ranking values of the items

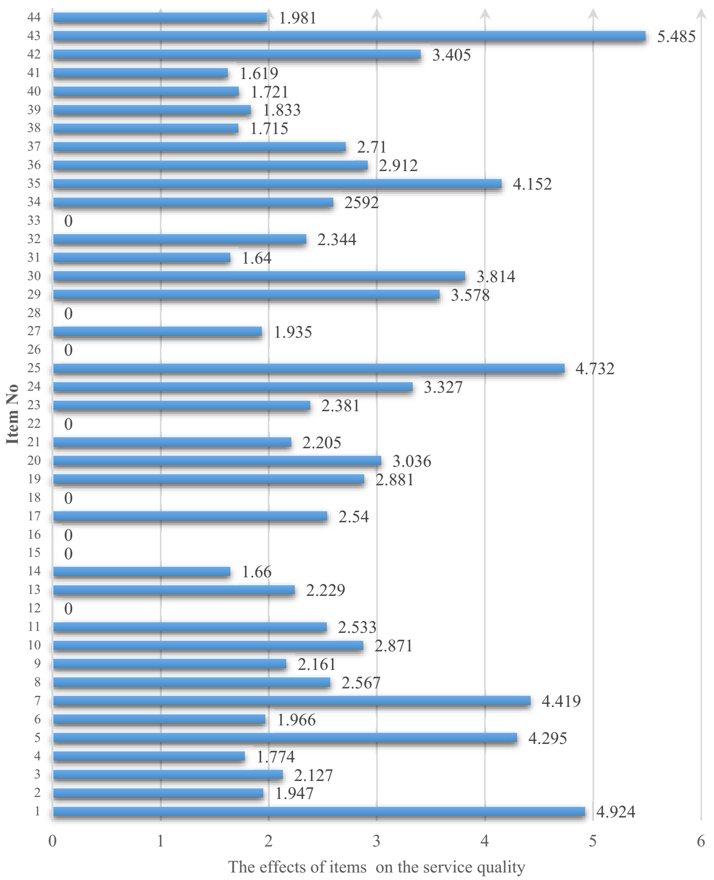


**Figure 2.**
Percentage effects of items on the service quality

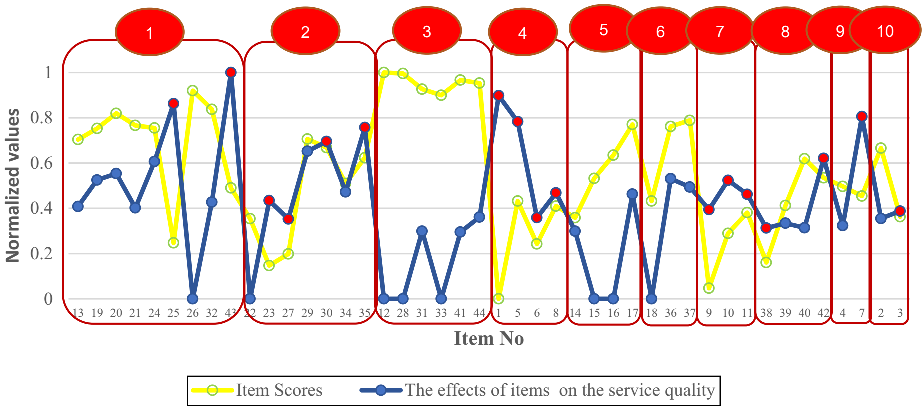


**Figure 3.**
Comparative analysis of item scores and the effects of items on the service quality

| Factor | No. of items | No. of items having gaps | (%) of gap |
|---|---|---|---|
| Factor 1 | 9 | 2 | 22 |
| Factor 2 | 7 | 4 | 57 |
| Factor 3 | 6 | 0 | 0 |
| *Factor 4* | *4* | *4* | *100* |
| Factor 5 | 4 | 0 | 0 |
| Factor 6 | 3 | 0 | 0 |
| *Factor 7* | *3* | *3* | *100* |
| Factor 8 | 4 | 2 | 50 |
| Factor 9 | 2 | 1 | 50 |
| Factor 10 | 2 | 1 | 50 |

**Table 10.**
Percentage of gap for factors

| The service quality of hospital (label) | Sample value | (%) |
|---|---|---|
| 5 | 116 | 29.7436 |
| 4 | 180 | 46.1538 |
| 3 | 85 | 21.7949 |
| 2 | 5 | 1.2821 |
| 1 | 4 | 1.0256 |
| Total | 390 | 100 |

**Table 11.**
The distribution of the label value among the hospital

search method was used to determine parameter settings given in Table 12. Classification models were obtained by using the software of WEKA (Witten *et al.*, 2011).

The performance values of the classification models used in this study are given in Table 13. As can be seen from Table 13, the multi-class classifier that uses the J48 algorithm as a sub-classifier performed the best with respect to the accuracy, precision, recall and f-measure values for 90% split ratio validity method. The RSM that uses a random tree algorithm as a sub-classifier performed the best regarding the RMSE, RRSE values for 90% split ratio validity method. The random tree algorithm performed the best with respect to MAE, RAE values for 90% split ratio validity method. Different performance values of different algorithms gave better results.

| Model | Parameter | Setting |
|---|---|---|
| Multi-class classifier | Batch size | 100 |
| | Classifier | Respectively, J48, random tree, REP tree |
| | Random width factor | 2 |
| | Method | Exhaustive correction code |
| RSM | Batch size | 100 |
| | Classifier | Respectively, J48, random tree, REP tree |
| | Num execution slots | 1 |
| | Num iterations | 10 |
| | Subspace size | 0.5 |
| | Batch size | 100 |
| J48 | Confidence factor | 0.2 |
| | Min number object | 2 |
| | Seed | 1 |
| | K-value | 0 |
| Random tree | Batch size | 100 |
| | Max depth | 0 |
| | Min number | 1 |
| | Min variance prop | 0.001 |
| REP tree | Batch size | 100 |
| | Initial count | 0 |
| | Max depth | −1 |
| | Min number | 2 |
| | Min variance prop | 0.001 |

**Table 12.** Parameter settings of ML models

Therefore, the SI value was used to obtain a single algorithm with the best performance. In this study, the RSM that uses random tree algorithm as sub-classifier provides the best SI value. The SI value is calculated regardless of the significance levels of these performance criteria. However, each of the performance values can be of different importance in practice. Thus, it is necessary to determine whether the deviation value is more important than the accuracy value. In this study, the significance level of performance values in Table 13 was regarded as equal. Also, the multi-class classifier and the RSM were used instead of the use of only one algorithm. Classification models established in this study have proved that machine learning (ML) techniques are appropriate methods for the classification of service quality in the health sector. This study provides a roadmap that takes the impact of the items on the quality of service into account for hospital managers.

## 5. Conclusion

Evaluation of service quality in health care is a popular and hot topic. The aim of this study is to propose a service quality evaluation model based on the SERVQUAL scale and machine learning algorithm in health care services. A real-life case study is performed to reveal how the proposed evaluation model works in practice. A survey study was conducted to obtain data from patients in a public hospital in Turkey. Then, mean score for items are calculated. Subsequently, the impact of items on service quality is calculated using information gain method. Mean score for items and the impact of items on service quality are normalized for comparison. Items that have lower mean score and higher impact are prioritized for improvement activities. The fact that which of these items should be improved primarily could be determined using the proposed approach. Later, items having a high effect value and a low score are determined based on the perspective of patient. In this study, the factors having the highest gap value between the effects value on the service

| Classification algorithms | Validation test | Accuracy | Precision | Recall | F-measure | MAE | RMSE | RAE (%) | RRSE (%) | SI |
|---|---|---|---|---|---|---|---|---|---|---|
| Multi-class classifier (random tree) | 90% split ratio | 71.7949 | 0.7170 | 0.7180 | 0.6970 | 0.2908 | 0.3668 | 116.5140 | 107.1940 | 1.1248 |
| | Fold 10 | 53.3300 | 0.5480 | 0.5330 | 0.5370 | 0.2965 | 0.3762 | 113.5110 | 104.2640 | 0.5841 |
| Multi-class classifier (REP tree) | 90% split ratio | 58.9744 | 0.5480 | 0.5900 | 0.5520 | 0.2969 | 0.3725 | 118.9380 | 108.8410 | 0.6385 |
| | Fold 10 | 55.6410 | 0.5736 | 0.5560 | 0.5424 | 0.3026 | 0.3799 | 115.8590 | 105.2890 | 0.6112 |
| Multi-class classifier (J48) | 90% split ratio | *76.9231* | *0.7800* | *0.7690* | *0.7600* | 0.2897 | 0.3641 | 116.0518 | 106.3980 | 1.3163 |
| | Fold 10 | 60.0000 | 0.5945 | 0.6000 | 0.5955 | 0.2969 | 0.3736 | 113.6800 | 103.5520 | 0.7778 |
| The RSM (random tree) | 90% split ratio | 71.7949 | 0.6150 | 0.7180 | 0.6600 | 0.1802 | *0.2909* | 72.1801 | *84.9912* | *1.6534* |
| | Fold 10 | 61.2821 | 0.6290 | 0.6130 | 0.6157 | 0.2103 | 0.3240 | 80.5427 | 89.7945 | 1.2979 |
| The RSM (REP tree) | 90% split ratio | 71.7949 | 0.6960 | 0.7180 | 0.6890 | 0.2118 | 0.3076 | 84.8500 | 89.8871 | 1.5720 |
| | Fold 10 | 53.3333 | 0.5762 | 0.5330 | 0.5173 | 0.2298 | 0.3330 | 88.0035 | 92.2797 | 0.9567 |
| The RSM (J48) | 90% split ratio | 66.6667 | 0.6560 | 0.6670 | 0.6490 | 0.1993 | 0.3031 | 79.8605 | 88.5772 | 1.4887 |
| | Fold 10 | 60.2564 | 0.6390 | 0.6030 | 0.6075 | 0.2173 | 0.3275 | 83.2053 | 90.7584 | 1.2477 |
| Random tree | 90% split ratio | 58.9744 | 0.6410 | 0.5900 | 0.5880 | *0.1641* | 0.4051 | *65.7433* | 118.3740 | 1.1304 |
| | Fold 10 | 44.8718 | 0.4540 | 0.4490 | 0.4510 | 0.2210 | 0.4699 | 84.6310 | 130.2240 | 0.3085 |
| The REP tree | 90% split ratio | 56.4103 | 0.5360 | 0.5640 | 0.5490 | 0.2185 | 0.3602 | 87.5454 | 105.2450 | 0.9126 |
| | Fold 10 | 61.0256 | 0.6744 | 0.6100 | 0.6208 | 0.2114 | 0.3251 | 80.9950 | 90.1261 | 1.3248 |
| J48 | 90% split ratio | 58.9744 | 0.6960 | 0.7180 | 0.6890 | 0.2185 | 0.3602 | 87.5454 | 105.2450 | 1.2889 |
| | Fold 10 | 51.2821 | 0.5313 | 0.5130 | 0.5120 | 0.2036 | 0.4120 | 77.9442 | 114.1790 | 0.7495 |

**Table 13.**
Performance values of the classification models

quality and item scores are Factor 4 (general physical conditions) and Factor 7 (general physical conditions). The gap corresponding to Factor 3 (act ethically), Factor 5 (reliability in services) and Factor 6 (accessibility) are 0. The reason for this case is that these factors have a high-quality value from the perspective of patient. Therefore, the improvements that will be carried on these factors will not provide a high effect level on the general service quality. As result, the factors that have a high effect value on the service quality and the low-quality score value are proposed for improvement activities. Thus, the budget allocated for improvement activities can be directed to the right items. In addition, a sample service quality classification model is presented in this study. Sample classification models, which have the best performance values, were selected using SI value among all of the established classification models. The classification algorithm, which has the best performance values, is the RSM that use the random tree algorithm as a sub-classifier. The accuracy value and SI value of this algorithm are 71.7949% and 1.6534, respectively. The classification algorithm, which has the highest performance values (accuracy, precision, recall and f-measure), is the multi-class classifier algorithm that use the J48 algorithm as a sub-classifier. In this study, the classification models have proved that ensemble machine learning algorithms are an appropriate approach for classification of service quality in the health care services. The proposed classification model achieved a successful predict with a rate of 76.923%.

By using the proposed classification model, the level of hospital service quality can be improved in practice. Necessary preventive actions can be taken by monitoring the fluctuations in the predicted hospital service quality. Thus, before the decrease in the quality level reaches an irreversible level, it could be possible to intervene in advance.

There are two limitations in this study. First, SERVQUAL results are valid only for one hospital because data is collected from one hospital. Second, it is assumed that each SERVQUAL item has equal importance in practice.

For the prospective studies, other service quality measurement scales can be used for the data obtaining process. Fuzzy logic-based service quality classification and evaluation can be used in future research. Expert opinion can be taken into consideration during the service quality evaluation process to prioritize factors for improvement activities. The patient satisfaction can be analyzed for pre-improvement and post-improvement so the success of the improvement activities can be tested. The number of observations in the data can be increased to obtain the higher predictive performance. In this study, it was observed that a model with a predictive performance of 76.923% was obtained by using 90% of 390 observations. If the number of observations is increased, k-fold validation methods providing a more robust performance evaluation can be used in future studies.

## References


Aburayya, A., Alshurideh, M., Al Marzouqi, A., Al Diabat, O., Alfarsi, A., Suson, R., Bash, M. and Salloun, S.A. (2020), "An empirical examination of the effect of TQM practices on hospital service quality: an assessment study in UAE hospitals", *Systematic Reviews in Pharmacy*, Vol. 11 No. 9, pp. 347-362.

Aggarwal, A., Aeran, H. and Rathee, M. (2018), "Quality management in health care: the pivotal desideratum", *Journal of Oral Biology and Craniofacial Research*, Vol. 9 No. 2.

Aghamolaei, T., Eftekhaari, T.E., Rafati, S., Kahnouji, K., Ahangari, S., Shahrzad, M.E., Kahnouji, A. and Hoseini, S.H. (2014), "Service quality assessment of a referral hospital in Southern Iran with SERVQUAL technique: patients' perspective", *BMC Health Services Research*, Vol. 14 No. 1, p. 322.

Akdag, H., Kalaycı, T., Karagöz, S., Zülfikar, H. and Giz, D. (2014), "The evaluation of hospital service quality by fuzzy MCDM", *Applied Soft Computing*, Vol. 23, pp. 239-248.

Akhyani, F., Birjandi, A.K., Sheikh, R. and Sana, S.S. (2020), "New approach based on proximity/remoteness measurement for customer classification", *Electronic Commerce Research*, pp. 1-32, doi: 10.1007/s10660-020-09402-7.

Alam, M.S. and Mondal, M. (2019), "Assessment of sanitation service quality in urban slums of Khulna city based on SERVQUAL and AHP model: a case study of railway slum, Khulna, Bangladesh", *Journal of Urban Management*, Vol. 8 No. 1, pp. 20-27.

Ali, M. (2018), "How patients perceive health care services: a case of Ayub teaching hospital, Abbottabad–Pakistan. SERV service QUAL quality", *International Journal of Healthcare Management*, Vol. 11 No. 1, pp. 52-59.

Ali, S.S., Basu, A. and Ware, N. (2018), "Quality measurement of Indian commercial hospitals – using a SERVQUAL framework", *Benchmarking: An International Journal*, Vol. 25 No. 3, pp. 815-837.

Alkafaji, M.K. and Al-Sharmey, E.S. (2020), "A fuzzy assessment model for hospitals services quality based on patient experience", *Karbala International Journal of Modern Science*, Vol. 6 No. 3, pp. 314-321.

Al-Neyadi, H.S., Abdallah, S. and Malik, M. (2018), "Measuring patient's satisfaction of healthcare services in the UAE hospitals: using SERVQUAL", *International Journal of Healthcare Management*, Vol. 11 No. 2, pp. 96-105.

Altuntas, S. and Kansu, S. (2019), "An innovative and integrated approach based on SERVQUAL, QFD and FMEA for service quality improvement", *Kybernetes*, Vol. 49 No. 10, pp. 2419-2453.

Altuntas, S. and Yener, E. (2012), "An approach based on TRIZ methodology and SERVQUAL scale to improve the quality of health care service: a case study", *Ege Akademik Bakış Dergisi*, Vol. 12 No. 1, pp. 95-104.

Altuntas, S., Dereli, T. and Kaya, İ. (2020), "Monitoring patient dissatisfaction: a methodology based on SERVQUAL scale and statistical process control charts", *Total Quality Management and Business Excellence*, Vol. 31 Nos 9/10, pp. 978-1008.

Altuntas, S., Dereli, T. and Yilmaz, M.K. (2012), "Multi-criteria decision making methods based weighted SERVQUAL scales to measure perceived service quality in hospitals: a case study from Turkey", *Total Quality Management and Business Excellence*, Vol. 23 Nos 11/12, pp. 1379-1395.

Babakus, E. and Mangold, W.G. (1992), "Adapting the SERVQUAL scale to hospital services: an empirical investigation", *Health Services Research*, Vol. 26 No. 6, pp. 767-786.

Badrick, T., Cervinski, M. and Loh, T.P. (2018), "A primer on patient-based quality control techniques", *Clinical Biochemistry*, Vol. 64, pp. 1-5.

Bayati, M. and Emadi, M. (2020), "Factors affecting hospital mortality rate in Iran: a panel data analysis", *BMC Research Notes*, Vol. 13 No. 1, pp. 1-5.

Behdioglu, S., Acar, E. and Burhan, H.A. (2019), "Evaluating service quality by fuzzy SERVQUAL: a case study in a physiotherapy and rehabilitation hospital", *Total Quality Management and Business Excellence*, Vol. 30 Nos 3/4, pp. 301-319.

Behkami, N.A. and Daim, T.U. (2011), "An analysis model for health information technology adoption", *First International Technology Management Conference*, pp. 468-474.

Behkami, N.A. and Daim, T.U. (2012a), *Adoption of Health Information Technologies*, Technology Management for Emerging Technologies (PICMET).

Behkami, N.A. and Daim, T.U. (2012b), "Research forecasting for health information technology (HIT), using technology intelligence", *Technological Forecasting and Social Change*, Vol. 79 No. 3, pp. 498-508.

Behkami, N.A. and Daim, T.U. (2016), "Exploring technology adoption in the case of the patient-centered medical home", *Health Policy and Technology*, Vol. 5 No. 2, pp. 166-188.

Berry, L.L., Parish, J.T. and Dikec, A. (2019), "Creating value through quality service", *Organizational Dynamics*, Vol. 49 No. 3.

Beura, S.K., Kumar, K.V., Saman, S. and Bhuyan, P.K. (2020), "Service quality analysis of signalized intersections from the perspective of bicycling", *Journal of Transport and Health*, Vol. 16

Bonner, A., Havas, K., Tam, V., Stone, C., Jennifer, A., Barnes, M. and Douglas, C. (2019), "An integrated chronic disease nurse practitioner clinic: service model description and patient profile", *Collegian*, Vol. 26 No. 2, pp. 227-234.

Chou, J.S., Tsai, C.F., Pham, A.D. and Lu, Y.H. (2014), "Machine learning in concrete strength simulations: multi-nation data analytics", *Construction and Building Materials*, Vol. 73, pp. 771-780.

Cullen, W.K., West, D. and Grant, S.W. (2018), "Evaluating quality in clinical care", *Surgery (Oxford)*, Vol. 36 No. 9, pp. 497-502.

Dangare, C.S. and Apte, S.S. (2012), "Improved study of heart disease prediction system using data mining classification techniques", *International Journal of Computer Applications*, Vol. 47 No. 10, pp. 44-48.

Deb, S. and Ahmed, M.A. (2018), "Determining the service quality of the city bus service based on users' perceptions and expectations", *Travel Behaviour and Society*, Vol. 12, pp. 1-10.

Deng, Y., Luo, X., Hu, X., Ma, Y. and Ma, R. (2020), "Modeling and prediction of bus operation states for bunching analysis", *Journal of Transportation Engineering, Part A: Systems*, Vol. 146 No. 9, p. 146.

Dhir, C.S., Iqbal, N. and Lee, S.-Y. (2007), "Efficient feature selection based on information gain criterion for face recognition", *2007 International Conference on Information Acquisition*, pp. 523-527.

Durairaj, M. and Ranjani, V. (2013), "Data mining applications in healthcare sector: a study", *International Journal of Scientific and Technology Research*, Vol. 2 No. 10, pp. 29-35.

Eldeeb, G. and Mohamed, M. (2020), "Quantifying preference heterogeneity in transit service desired quality using a latent class choice model", *Transportation Research Part A: Policy and Practice*, Vol. 139, pp. 119-133.

Erdoğan, Z. (2017), *Yaşam Kalitesi Endeksi Tabanlı Bileşik Makine Öğrenme Teknikleriyle Yaşam Alanı Tahmin Modeli*, İstanbul Üniversitesi.

Erdogan, Z. and Namli, E. (2019), "A living environment prediction model using ensemble machine learning techniques based on quality of life index", *Journal of Ambient Intelligence and Humanized Computing*, pp. 1-17, doi: 10.1007/s12652-019-01432-w.

Esteva, A., Kuprel, B., Novoa, R.A., Ko, J., Swetter, S.M., Blau, H.M. and Thrun, S. (2017), "Dermatologist-level classification of skin cancer with deep neural networks", *Nature*, Vol. 542 No. 7639, pp. 115-118.

Farhadi, P., Niyas, M., Shokrpour, N. and Ravangard, R. (2020), "Prioritizing factors affecting health service quality using integrated fuzzy DEMATEL and ANP: a case of Iran", *The Open Public Health Journal*, Vol. 13 No. 1, pp. 263-272.

Fattore, U., Liebsch, M., Brik, B. and Ksentini, A. (2020), "AutoMEC: LSTM-based user mobility prediction for service management in distributed MEC resources", *Proceedings of the 23rd International ACM Conference on Modeling, Analysis and Simulation of Wireless and Mobile Systems*, pp. 155-159.

Fernandez, G. (2010), *Statistical Data Mining Using SAS Applications*, 2nd ed., CRC Press.

Firouzi Jahantigh, F. and Ostovare, M. (2020), "Application of a hybrid method for performance evaluation of teaching hospitals in Tehran", *Quality Management in Health Care*, Vol. 29 No. 4, pp. 210-217.

Fragoso, J.T. and Espinoza, I.L. (2017), "Assessment of banking service quality perception using the SERVPERF model", *Contaduría y Administración*, Vol. 62 No. 4, pp. 1294-1316.

Golmohammadi, D., Parast, M.M. and Sanders, N. (2020), "The impact of service failures on firm profitability: integrating machine learning and statistical modeling", *IEEE Transactions on Engineering Management*, pp-1-15, doi: 10.1109/TEM.2020.3015771.

Gullu, O., Tekindal, M., Tekindal, M.A. and Yazıcı, A.C. (2017), "Evaluation of expected and perceived of quality of service with the SERVQUAL scale: the case of a private physical therapy and rehabilitation center", *Biomedical Research*, Vol. 28 No. 2, pp. 711-715.

Gundogdu, F.K. and Kahraman, C. (2021), "Hospital performance assessment using interval-valued spherical fuzzy analytic hierarchy process", *Decision Making with Spherical Fuzzy Sets*, Springer, pp. 349-373.

Han, J. and Kamber, M. (2001), *Data Mining: Concepts and Techniques*, 2nd ed, in Morgan Kaufmann, San Francisco, ISBN 10: 1-55860-901-6.

Hatam, N., Sadeghi, A., Shojaee, P., Jafari, H. and Ghorbanian, A. (2020), "Effective factors on improving the quality of hospital services in South of Iran", *The Journal of the Pakistan Medical Association*, Vol. 70 No. 10, pp. 1709-1713.

Ho, T.K. (1998), "The random subspace method for constructing decision forests", *EEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 20 No. 8, pp. 832-844.

Hussain, A., Sial, M.S., Usman, S.M., Hwang, J., Jiang, Y. and Shafiq, A. (2019), "What factors affect patient satisfaction in public sector hospitals: evidence from an emerging economy", *International Journal of Environmental Research and Public Health*, Vol. 16 No. 6, pp. 16

Isak-Zatega, S., Lipovac, A. and Lipovac, V. (2020), "Logistic regression based in-service assessment of mobile web browsing service quality acceptability", *EURASIP Journal on Wireless Communications and Networking*, pp. 1-21, doi: 10.1186/s13638-020-01708-2.

Izenman, A.J. (2008), *Modern Multivariate Statistical Techniques: Multidimensional Scaling and Distance Geometry*, 1st ed., Springer.

Javed, S.A., Liu, S., Mahmoudi, A. and Nawaz, M. (2019), "Patients' satisfaction and public and private sectors' health care service quality in Pakistan: application of grey decision analysis approaches", *The International Journal of Health Planning and Management*, Vol. 34 No. 1, pp. 168-182.

Jennings, N., Clifford, S., Fox, A.R., O'Connell, J. and Gardner, G. (2015), "The impact of nurse practitioner services on cost, quality of care, satisfaction and waiting times in the emergency department: a systematic review", *International Journal of Nursing Studies*, Vol. 52 No. 1, pp. 421-435.

Jiang, L. and Liao, H. (2019), "A linear programming method based on probabilistic linguistic Kolmogorov-Smirnov distance for hospital service quality evaluation", *2019 IEEE Symposium Series on Computational Intelligence (SSCI)*, pp. 3195-3200.

Jiang, S., Shi, H., Lin, W. and Liu, H.-C. (2020), "A large group linguistic Z-DEMATEL approach for identifying key performance indicators in hospital performance management", *Applied Soft Computing*, Vol. 86, pp. 1-12.

Kadir, A.R., Kamariah, N., Saleh, A. and Ratnawati (2017), "The effect of role stress, job satisfaction, self-efficacy and nurses' adaptability on service quality in public hospitals of Wajo", *International Journal of Quality and Service Sciences*, Vol. 9 No. 2, pp. 184-202.

Kaur, H. and Wasan, S.K. (2006), "Empirical study on applications of data mining techniques in healthcare", *Journal of Computer Science*, Vol. 2 No. 2, pp. 194-200.

Kaya, S. (2014), *Evaluation of Inpatients Perceptions by SERVQUAL Based Analysis and Quality Function Deployment: A Case Study for Eskişehir State Hospital*, Eskisehir Osmangazi University.

Koyuncugil, A.S. and Özgülbaş, N. (2019), "Veri Madenciliği: Tıp ve Sağlık Hizmetlerinde Kullanımı ve Uygulamaları", *Bilişim Teknolojileri Dergisi*, Vol. 2 No. 2, pp. 21-32.

Lee, K. and Yu, C. (2018), "Assessment of airport service quality: a complementary approach to measure perceived service quality based on Google reviews", *Journal of Air Transport Management*, Vol. 71, pp. 28-44.

Lei, S. (2012), "A feature selection method based on information gain and genetic algorithm", *2012 International Conference on Computer Science and Electronics Engineering*, pp. 355-358.

Li, M., Lowrie, D.B., Huang, C.Y., Lu, X.C., Zhu, Y.C., Wu, X.H., Shayiti, M., Tan, Q.Z., Yang, H.L., Chen, S.Y., Zhao, P., He, S.H., Wang, X.R. and Lu, H.Z. (2015), "Evaluating patients' perception of service quality at hospitals in nine Chinese cities by use of the ServQual scale", *Asian Pacific Journal of Tropical Biomedicine*, Vol. 5 No. 6, pp. 497-504.

Lucini, F.R., Tonetto, L.M., Fogliatto, F.S. and Anzanello, M.J. (2020), "Text mining approach to explore dimensions of airline customer satisfaction using online customer reviews", *Journal of Air Transport Management*, Vol. 83, p. 101760.

Maghsoodia, A., Saghaeia, A. and Hafezalkotob, A. (2019), "Service quality measurement model integrating an extended SERVQUAL model and a hybrid decision support system", *European Research on Management and Business Economics*, Vol. 25 No. 3, pp. 151-164.

Maimon, O. and Rokach, L. (2005), *Data Mining and Knowledge Discovery Handbook*, 2nd ed., Springer.

Martín-Martínez, M.A., Andreu-Sanchez, J.L., Sanchez-Alonso, F., Corominas, H., Perez-Venegas, J.J., Roman-Ivorra, J.A., Alperi, M., Blanco-Alonso, R., Caliz, R., Chamizo-Carmona, E., Grã na-Gil, J., Hernández, B., Marras, C., Mazzucchelli, R., Medina Luezas, J.A., Naranjo-Hernández, A., Ortiz, A., Roselló, R., Sanchez-Nievas, G., Sanmartí, R. and Vela-Casasempere, P. (2019), "A composite indicator to assess the quality of care in the management of patients with rheumatoid arthritis in outpatient rheumatology clinics", *Reumatología Clínica (English Edition)*, Vol. 15 No. 3, pp. 156-164.

MeeraGandhi, G. (2010), "Machine learning approach for attack prediction and classification using supervised learning algorithms", *International Journal of Computer Science and Communication*, Vol. 1 No. 2, pp. 11465-11484.

Mirzaeia, A., Cartera, S.R., Chen, J.Y., Rittsteuer, C. and Schneidera, C.R. (2019), "Development of a questionnaire to measure consumers' perceptions of service quality in community pharmacies", *Research in Social and Administrative Pharmacy*, Vol. 15 No. 4, pp. 346-357.

Moro, S., Lopes, R.J., Esmerado, J. and Botelho, M. (2020), "Service quality in airport hotel chains through the lens of online reviewers", *Journal of Retailing and Consumer Services*, Vol. 56, p. 102193.

Mudavadi, C., Hogaboam, L. and Daim, T.U. (2016), "A hierarachical decision model (HDM) for exploring the adoption of electronic health records", *In 2016 Portland International Conference on Management of Engineering and Technology (PICMET)*, pp. 2770-2781.

Mukherjee, S., Choi, T., Islam, M.T., Choi, B.-Y., Beard, C., Won, S.H. and Song, S. (2020), "A supervised-learning-based spatial performance prediction framework for heterogeneous communication networks", *ETRI Journal*, Vol. 42 No. 5, pp. 689-702.

Nacer, A.A., Bessai, K., Youcef, S. and Godart, C. (2015), "A multi-criteria based approach for web service selection using quality of service (QoS)", *In* 2015 IEEE International Conference on Services Computing, pp. 570-577.

Nemati, R., Bahreini, M., Pouladi, S., Mirzaei, K. and Mehboodi, F. (2020), "Hospital service quality based on HEALTHQUAL model and trusting nurses at Iranian university and non-university hospitals: a comparative study", *BMC Nursing*, Vol. 19 No. 1, pp. 1-9.

Nyandwe, J., Mapatano, M., Lussamba, P., Kandala, N.-B. and Kayembe, P. (2017), "Measuring patients' perception on the quality of care in the democratic republic of congo using a modified, service quality scale (SERVQUAL)", *Archives of Science*, Vol. 1 No. 2, pp. 1-6.

Ocampo, L., Alinsub, J., Casul, R.A., Enquig, G., Luar, M., Panuncillon, N., Bongo, M. and Ocampo, C.O. (2019), "Public service quality evaluation with SERVQUAL and AHP-TOPSIS: a case of Philippine government agencies", *Socio-Economic Planning Sciences*, Vol. 68.

Oriol, M., Marco, J. and Franch, X. (2014), "Quality models for web services: a systematic mapping", *Information and Software Technology*, Vol. 56 No. 10, pp. 1167-1182.

Parezanović, T., Petrović, M., Bojković, N. and Pamučar, D. (2019), "One approach to evaluate the influence of engineering characteristics in QFD method", *European J. Of Industrial Engineering*, Vol. 13 No. 3, pp. 299-331.

Patil, T. and Sherekar, S. (2013), "Performance analysis of Naive Bayes and J48 classification algorithm for data classification", *International Journal of Computer Science and Applications*, Vol. 6 No. 2, pp. 256-261.

Pekkaya, M., Pulat İmamoğlu, Ö. and Hayriye, K. (2019), "Evaluation of healthcare service quality via SERVQUAL scale: an application on a hospital", *International Journal of Healthcare Management*, Vol. 12 No. 4, pp. 340-347.

Perera, S. and Dabney, B.W. (2020), "Case management service quality and patient-centered care", *Journal of Health Organization and Management*, Vol. 34 No. 5, pp. 551-568.

Quinlan, J.R. (1996), "Improved use of continuous attributes in C4.5", *Journal of Artificial Intelligence Research*, Vol. 4, pp. 77-90.

Rai, N.K., Tyrrell, H., Carey, C. and Tiwari, T. (2019), "Patient perceptions in quality of care: report from university veterans clinic", *BMC Oral Health*, Vol. 19 No. 1.

Rallis, I., Markoulidakis, I., Georgoulas, I. and Kopsiaftis, G. (2020), "A novel classification method for customer experience survey analysis", *Proceedings of the 13th ACM International Conference on PErvasive Technologies Related to Assistive Environments*, pp. 1-9.

Rasouli, O. and Zarei, M.H. (2016), "Monitoring and reducing patient dissatisfaction: a case study of an Iranian public hospital", *Total Quality Management and Business Excellence*, Vol. 27 Nos 5/6, pp. 531-559.

Rehaman, B. and Husnain, M. (2018), "The impact of service quality dimensions on patient satisfaction in the private healthcare industry in Pakistan", *Journal of Hospital and Medical Management*, Vol. 4 Nos 1/4, pp. 1-8.

Riono, A. (2017), "Analysis of healthcare services quality using SERVQUAL-fuzzy method", *International Journal of Economics and Management Sciences*, Vol. 6 No. 6, doi: 10.4172/2162-6359.1000485.

Roy, J., Adhikary, K., Kar, S. and Pamucar, D. (2018), "A rough strength relational DEMATEL model for analysing the key success factors of hospital service quality", *Decision Making: Applications in Management and Engineering*, Vol. 1 No. 1, pp. 121-142.

Sá, F., Álvaro, R. and Cota, M.P. (2016), "Potential dimensions for a local e-government services quality model", *Telematics and Informatics*, Vol. 33 No. 2, pp. 270-276.

Safari, Y., Cheshmeh-Kaboodi, A.M. and Yousefi, B. (2019), "The data on the quality of services, satisfaction, psychological commitment and oral advertising in clinical centers in Kermanshah", *Data in Brief*, Vol. 23, p. 23.

Saleem, S., Haider, H., Hu, G., Hewage, K. and Sadiq, R. (2020), "Performance indicators for aquatic centres in Canada: identification and selection using fuzzy based methods", *Science of the Total Environment*, Vol. 751, p. 141619.

Shafii, M., Rafiei, S., Abooee, F., Bahrami, M.A., Nouhi, M., Lotfi, F. and Khanjankhani, K. (2016), "Assessment of service quality in teaching hospitals of Yazd university of medical sciences: using multi-criteria decision making techniques", *Osong Public Health and Research Perspectives*, Vol. 7 No. 4, pp. 239-247.

Shafiq, M., Naeem, M.A., Munawar, Z. and Fatima, I. (2017), "Service quality assessment of hospitals in Asian context: an empirical evidence from Pakistan", *INQUIRY: The Journal of Health Care Organization, Provision, and Financing*, Vol. 54, pp. 1-12.

Shah, A.M., Yan, X., Shah, S.A.A. and Mamirkulova, G. (2019), "Mining patient opinion to evaluate the service quality in healthcare: a deep-learning approach", *Journal of Ambient Intelligence and Humanized Computing*, Vol. 11 No. 7, pp. 2925-2942.

Shah, C. and Jivani, A. (2013), "Comparison of data mining classification algorithms for breast cancer prediction", *2013 Fourth International Conference on Computing, Communications and Networking Technologies (ICCCNT)*, pp. 1-4.

Shirazi, H., Kia, R. and Ghasemi, P. (2020), "Ranking of hospitals in the case of COVID-19 outbreak: a new integrated approach using patient satisfaction criteria", *International Journal of Healthcare Management*, Vol. 13 No. 4, pp. 1-13.

Shokouhyar, S., Shokoohyar, S. and Safari, S. (2020), "Research on the influence of after-sales service quality factors on customer satisfaction", *Journal of Retailing and Consumer Services*, Vol. 56.

Shuv-Ami, A. and Shalom, T. (2017), "Demographic differences of perceived service quality in emergency rooms of hospital organizations", *International Journal of Organizational Analysis*, Vol. 25 No. 2, pp. 282-294.

Singh, A. and Prasher, A. (2019), "Measuring healthcare service quality from patients' perspective: using fuzzy AHP application", *Total Quality Management and Business Excellence*, Vol. 30 Nos 3/4, pp. 284-300.

Singh, A., Prasher, A. and Kaur, N. (2020a), "Assessment of hospital service quality parameters from patient, doctor and employees' perspectives", *Total Quality Management and Business Excellence*, Vol. 31 Nos 13/14, pp. 1467-1486.

Singh, A., Tewari, E. and Ravi, P. (2020b), "SERVQUAL (service quality) vs NPS (net promoter score): a comparative study of private and public hospitals in Sikkim", *Indian Journal of Marketing*, Vol. 50 Nos 10/11, pp. 23-39.

Somu, N., Mr, G.R., Kalpana, V., Kirthivasan, K. and Vs, S.S. (2018), "An improved robust heteroscedastic probabilistic neural network based trust prediction approach for cloud service selection", *Neural Networks*, Vol. 108, pp. 339-354.

Son, L.H., Ciaramella, A., Huyen, D.T.T., Staiano, A., Tuan, T.M. and Van Hai, P. (2020), "Predictive reliability and validity of hospital cost analysis with dynamic neural network and genetic algorithm", *Neural Computing and Applications*, pp. 1-12.

Soni, J., Ansari, U., Sharma, D. and Soni, S. (2011), "Predictive data mining for medical diagnosis: an overview of heart disease prediction", *International Journal of Computer Applications*, Vol. 17 No. 8, pp. 43-48.

Souri, M.E., Sajjadian, F., Shaikh, R. and Sana, S.S. (2018), "Grey SERVQUAL method to measure consumers' attitudes towards green products-a case study of Iranian consumers of LED bulbs", *Journal of Cleaner Production*, Vol. 177, pp. 187-196.

Stevic, Z., Dalic, I., Pamucar, D., Nunic, Z., Veskovic, S., Vasiljevic, M. and Ilija, T. (2019), "A new hybrid model for quality assessment of scientific conferences based on rough BWM and SERVQUAL", *Scientometrics*, Vol. 119 No. 1, pp. 1-30.

Suresh, M., Vaishnavi, V. and Pai, R.D. (2020), "Leanness evaluation in health care organizations using fuzzy logic approach", *International Journal of Organizational Analysis*, Vol. 28 No. 6.

Tan, H. and Yan, M. (2020), "Physician-user interaction and users' perceived service quality: evidence from Chinese mobile healthcare consultation", *Information Technology and People*, Vol. 33 No. 5.

Teshnizi, S.H., Aghamolaei, T., Kahnouji, K., Teshnizi, S.M.H. and Ghani, J. (2018), "Assessing quality of health services with the SERVQUAL model in Iran. A systematic review and Meta-analysis", *International Journal for Quality in Health Care*, Vol. 30 No. 2, pp. 82-89.

Ting, L.C., Moorthy, K., Kee, H.W., Yee, C.W., Yee, L.W., Ni, O.A. and Ting, Y.W. (2019), "Service quality and outpatients satisfaction in public hospitals in Malaysia", *International Journal of Public Policy and Administration Research*, Vol. 6 No. 1, pp. 57-73.

Tomar, D. and Agarwal, S. (2013), "A survey on data mining approaches for healthcare", *International Journal of Bio-Science and Bio-Technology*, Vol. 5 No. 5, pp. 241-266.

Topacan, U., Basoglu, A.N. and Daim, T.U. (2008), "Exploring the success factors of health information service adoption", *PICMET'08-2008 Portland International Conference on Management of Engineering and Technology*, pp. 2453-2461.

Tuzkaya, G., Sennaroglu, B., Kalender, Z.T. and Mutlu, M. (2019), "Hospital service quality evaluation with IVIF-PROMETHEE and a case study", *Socio-Economic Planning Sciences*, Vol. 68.

Vanichchinchai, A. (2020), "Priority nonconformity and service quality analysis of hospitals in Thailand: a care provider perspective", *The TQM Journal*.

Vapnik, V.N. (1998), *Statistical Learning Theory*, Wiley.

Vicente, P., Suleman, A. and Reis, E. (2020), "Index of satisfaction with public transport: a fuzzy clustering approach", *Sustainability*, Vol. 12 No. 22.

Vijayarani, S. and Dhayanand, S. (2015), "Liver disease prediction using SVM and Naïve Bayes algorithms", *International Journal of Science, Engineering and Technology Research*, Vol. 4 No. 4, pp. 816-820.

Wang, W., Sun, Z., Wang, L., Yu, S. and Chen, J. (2020), “Evaluation model for the level of service of shared-use paths based on traffic conflicts”, *Sustainability*, Vol. 12 No. 18.

Weston, J. and Watkins, C. (1998), *Multi-Class Support Vector Machines*, Technical Report CSD-TR-98-04, Department of Computer Science, Royal Holloway, University of London, May, pp. 98-04.

Witten, I.H., Frank, E. and Hall, M.A. (2011), *Data Mining Practical Machine Learning Tools and Techniques*, 3rd ed., Morgan Kaufmann.

Yan, R., Ma, Z., Zhao, Y. and Kokogiannakis, G. (2016), “A decision tree based data-driven diagnostic strategy for air handling units”, *Energy and Buildings*, Vol. 133, pp. 37-45.

Yucesan, M. and Gul, M. (2020), “Hospital service quality evaluation: an integrated model based on Pythagorean fuzzy”, *Soft Computing*, Vol. 24 No. 5, pp. 3237-3255.

Zarei, E., Bagheri, A., Daneshkohan, A. and Khodakarim, S. (2020), “Patients’ views on service quality in selected Iranian hospitals: an importance-performance analysis”, *Shiraz E-Medical Journal*, Vol. 21 No. 9.

Zhao, Y. and Zhang, Y. (2008), “Comparison of decision tree methods for finding active objects”, *Advances in Space Research*, Vol. 41 No. 12, pp. 1955-1959.

**Corresponding author**
Serkan Altuntas can be contacted at: serkan@yildiz.edu.tr

## Appendix

| Item no | Statement |
|---|---|
| 1 | Has modern physical appearance and medical equipment |
| 2 | Equipment in patient rooms such as TV, nurse call bell, lamp and bed are working |
| 3 | Has adequate patient room temperature |
| 4 | Clean patient room |
| 5 | Peaceful and quiet hospital environment |
| 6 | Has ideal number of inpatients in patient rooms |
| 7 | Clean WC/bathroom |
| 8 | Has enough number of WC/bathroom |
| 9 | Gives meals which are mouth-pleasing |
| 10 | The food is hot |
| 11 | Meals are satisfying |
| 12 | Employees are neat in appearance |
| 13 | Performs the service right the first time |
| 14 | Has less waiting time in radiology (film, x-ray, ultrasound) and laboratory (blood, urine analysis) services |
| 15 | Provides analysis reports related to radiology and laboratory services on time |
| 16 | Never occur unsatisfied services in radiology and laboratory such as loss of results, faulty and incomplete results |
| 17 | Has less waiting time for bureaucratic procedures in a hospital (referral opening-closing, admission-exit procedures [. . .]) and these procedures are completed smoothly |
| 18 | Act urgent inspection in case of emergency |
| 19 | Doctors ready to serve at any time |
| 20 | Health staff ready to serve at any time |
| 21 | Shows interest in inpatient problems and sincere interest to solve inpatient problems |
| 22 | Has experienced doctors in all branches |
| 23 | Has knowledgeable and experienced doctors and nurses at the weekend as well |
| 24 | Provides services expected by companions and inpatient relatives |
| 25 | The frequency of doctor visits to patients is sufficient |
| 26 | Has patient visiting hours and duration of visit convenient to inpatient relatives |
| 27 | Has the time allocated for visit operations convenient |
| 28 | Provides food services on time |
| 29 | Has doctors who are knowledgeable and experienced |
| 30 | Has health staff who are knowledgeable and experienced |
| 31 | Has employees who are respect in-patient privacy |
| 32 | Has health staff who are polite, gentle and respectful |
| 33 | Performs only the necessary tests and treatments |
| 34 | Provides visit operations fairly and equally for every inpatient |
| 35 | Has employees who gives attention to inpatient security such as the confidentiality of patient information, the physical and monetary security |
| 36 | Easy to reach personnel who are wanted to be consult by companion and inpatient |
| 37 | Has health staff who can be easily asked related to any questions |
| 38 | Tells inpatients and their relatives the procedures, operations, average length of stay in the hospital |

*(continued)*

**Table A1.** Items for hospital service quality (Kaya, 2014)

| Item no | Statement |
|---|---|
| 39 | Tells you information about the procedures they will perform (fever-blood pressure measurement, blood-urine, drugs) |
| 40 | Tells you using an appropriate speech style (not including medical terminology) for intelligibility |
| 41 | Get inpatient approval for the procedures to be performed on the patient |
| 42 | Provides information about patients' situation at any time |
| 43 | Has services planned according to patients' wishes, needs and expectations |
| 44 | Allows everyone to have adequate health care as possible without social security, financial possibilities, ethnic origin and religious beliefs and without financial expectation of all employees |

**Table A1.**

**Table A2.**
Item-total statistics

| Items | Scale mean if item deleted | Scale variance if item deleted | Corrected item-total correlation | Cronbach's alpha if item deleted |
|---|---|---|---|---|
| 1 | 185.1564 | 537.094 | 0.445 | 0.930 |
| 2 | 184.3897 | 547.683 | 0.341 | 0.930 |
| 3 | 184.7385 | 544.523 | 0.382 | 0.930 |
| 4 | 184.5846 | 540.768 | 0.486 | 0.929 |
| 5 | 184.6590 | 535.860 | 0.480 | 0.929 |
| 6 | 184.8769 | 538.926 | 0.372 | 0.931 |
| 7 | 184.6333 | 533.503 | 0.545 | 0.929 |
| 8 | 184.6846 | 538.977 | 0.386 | 0.930 |
| 9 | 185.1026 | 541.563 | 0.348 | 0.931 |
| 10 | 184.8231 | 538.547 | 0.437 | 0.930 |
| 11 | 184.7179 | 542.558 | 0.382 | 0.930 |
| 12 | 184.0051 | 554.931 | 0.420 | 0.930 |
| 13 | 184.3462 | 536.325 | 0.614 | 0.928 |
| 14 | 184.7436 | 542.083 | 0.415 | 0.930 |
| 15 | 184.5436 | 546.007 | 0.401 | 0.930 |
| 16 | 184.4256 | 549.191 | 0.388 | 0.930 |
| 17 | 184.2692 | 547.760 | 0.410 | 0.930 |
| 18 | 184.6590 | 546.374 | 0.330 | 0.931 |
| 19 | 184.2897 | 537.918 | 0.617 | 0.928 |
| 20 | 184.2128 | 538.708 | 0.654 | 0.928 |
| 21 | 184.2744 | 536.200 | 0.640 | 0.928 |
| 22 | 184.7487 | 547.412 | 0.368 | 0.930 |
| 23 | 184.9872 | 541.509 | 0.497 | 0.929 |
| 24 | 184.2872 | 539.105 | 0.592 | 0.928 |
| 25 | 184.8718 | 528.904 | 0.557 | 0.928 |
| 26 | 184.0974 | 548.463 | 0.463 | 0.929 |
| 27 | 184.9282 | 536.550 | 0.445 | 0.930 |
| 28 | 184.0103 | 557.841 | 0.301 | 0.930 |
| 29 | 184.3436 | 543.707 | 0.554 | 0.929 |
| 30 | 184.3872 | 540.870 | 0.605 | 0.928 |
| 31 | 184.0897 | 551.074 | 0.460 | 0.930 |
| 32 | 184.1923 | 542.562 | 0.598 | 0.929 |
| 33 | 184.1205 | 550.970 | 0.449 | 0.930 |
| 34 | 184.5692 | 542.318 | 0.481 | 0.929 |
| 35 | 184.4385 | 541.794 | 0.528 | 0.929 |
| 36 | 184.2795 | 535.564 | 0.617 | 0.928 |
| 37 | 184.2487 | 538.660 | 0.571 | 0.928 |
| 38 | 184.9718 | 537.416 | 0.397 | 0.930 |
| 39 | 184.6821 | 533.729 | 0.516 | 0.929 |
| 40 | 184.4436 | 536.489 | 0.596 | 0.928 |
| 41 | 184.0436 | 550.957 | 0.441 | 0.930 |
| 42 | 184.5410 | 530.681 | 0.591 | 0.928 |
| 43 | 184.5923 | 532.319 | 0.601 | 0.928 |
| 44 | 184.0590 | 550.683 | 0.474 | 0.930 |