

Investigating the Service Quality of Kocaeli Tram Service Using Artificial Neural Networks

Selim DÜNDAR¹

ABSTRACT

Service quality is one of the main issues of today's world. Firms operating in the passenger transportation sector are also trying to improve the quality of the services they provide to their passengers. It is crucial to determine the passenger service quality perceptions and priorities to evaluate and improve the service in this context. In this study, Kocaeli tram user service quality perceptions have been evaluated by applying a survey consisting of 20 questions and user satisfaction levels from different service dimensions. Then, an artificial neural network model was developed using the demographic data of the users and their responses to the survey questions to mimic their service quality satisfaction. The artificial neural network model developed has been examined to understand the importance that tram users give to service quality. Using the developed “change of score” method, how the changes to be made in the tram system will affect the quality of service and, hence, how the opinions of different user groups will be affected can be examined in detail. The artificial neural network model prediction capability was compared with that of the multiple linear regression model, and found its superiority. Based on the developed Change of Score Method, the most frequent user attaches the highest importance to the service dimensions of the convenience to pay for the tram, getting his/her destination on time, and reducing environmental pollution.

Keywords: Service quality, public transportation, artificial neural networks, tram services, transportation planning, transportation modelling.

1. INTRODUCTION

The rapid growth of the population and the intense migration from the village to the city are some of the main reasons for the increasing traffic, especially in the big cities of developing countries. The public enterprises and municipalities are working together and creating plans to minimize the adverse effects of traffic problems. Transportation planning is needed to decrease the existing problems such as accidents, stress, environmental pollution, and energy use brought by increasing traffic.

Note:

- This paper was received on August 20, 2020 and accepted for publication by the Editorial Board on April 9, 2021.
- Discussions on this paper will be accepted by November 20, 2022.
- <https://doi.org/10.18400/tekderg.783110>

¹ Istanbul Okan University, Civil Engineering Department, Istanbul, Turkey - selim.dundar@okan.edu.tr
<https://orcid.org/0000-0003-4433-1998>

Developed communities are formed in environments where people can easily travel without chaos and turmoil. Since economic prosperity creates mobility for various social activities, it has become imperative to establish transportation systems as a requirement created by the desire to travel. Public enterprises have the task of examining and evaluating travel demands and finding solutions for the prospective problems.

Consumer-oriented *Service Quality* (SQ) perception has become one of the most critical urban transportation dynamics [1]. Due to the public transport sector structure, competition occurs with private car use and within the sector itself. This competition requires the sector's components to continually improve the quality of the service and achieve the comfort level provided by personal automobiles.

When the SQ of public transportation is improved, some citizens could prefer public transportation instead of private automobiles. This situation leads to increased mobility with less marginal costs such as gas emissions or noise pollution. Determining the criteria that affect the SQ perceptions of public transportation system users is of great importance for increasing the general SQ.

This study investigates the SQ of Kocaeli tram service from the most frequent user type perspective, using *the Artificial Neural Networks* (ANN) method. Although some studies in the literature focus on SQ from the users' perspective, the presented study differs from them by developing a new method to investigate each SQ dimension's effect from various perspectives. The study is the first to investigate the SQ perception of various tram users' perspectives using ANN to the author's knowledge. The developed *Change Of a Score* (COS) *Method* is a simple but powerful method that uses a trained ANN to evaluate the possible effects of the changes planned or already applied to the tram system. The decision-makers and/or operators can also analyze various user groups' needs thoroughly and develop solutions focusing on their needs easily using this method. For this reason, it will not only fill a gap in the literature, state agencies or companies can also use it to prepare their investment plans more effectively.

This paper's organization is as follows; in the next section, a brief literature review is presented. In the third section, the methodology of the study is described. In the fourth and fifth sections, the study results and discussions about them are presented, respectively. The paper concludes in the sixth and the final section.

2. LITERATURE REVIEW

There are various studies in the literature regarding the SQ of public transport and users' perception of the system.

Dell'Olio et al. [2] investigated the differences between perceived quality and systematic expectations of users from the public transportation agencies in terms of SQ. The users valued public transportation service variables, such as waiting time, cleanliness, and comfort the most, but the magnitude of the values varied according to the user type. Users gave less weight to the variables such as driver courtesy, bus occupancy, and journey. For potential users, the essential variables when determining the quality expected from public transport were waiting time, travel time, and above all, the occupancy level of the vehicles.

Friman and Felleson [3] compared the agencies providing public transportation service by examining public transportation service performance in Europe over 6000 participants. They collected data from six different European cities. They considered three objective service performance measures for each city from the UITP Millennium Database. A total of 6021 participants evaluated three subjective satisfaction criteria obtained from the Benchmarking in European Service of Public Transport [4]. They also used general satisfaction as a subjective measure, in addition to the subjective criteria of quality. Correlational analyses show that the relationship between satisfaction and service performance in public transport is far from perfect.

Eboli and Mazzulla [5] proposed a tool to measure customer satisfaction in public transportation. The study created a structural equation model to investigate the relationship between global customer satisfaction and SQ features. The public transportation service investigated is a bus service used by Calabria University students in Cosenza's (Southern Italy) urban area to reach the campus. To calibrate the model, they used the data collected from a survey conducted on the students. They defined the proposed model as useful for transport agencies and planners to examine the relationship between qualities of service qualifications and identify more appropriate qualifications to improve the service provided.

Noor et al. [6] presented criteria that affect the bus SQ. They defined 24 criteria under three main headings: comfort, accessibility, and safety. After that, they applied the specified criteria to the city of Kinabalu in Malaysia. The study revealed a small difference in satisfaction with the minibus and bus services in the city. The essential features that affected user satisfaction emerge as feeling overcrowded and insecure during the night. Islam et al. [7] conducted a customer satisfaction survey with 300 passengers in Kadeh. According to the results, safety, driver performance, and bus condition were the criteria that affect SQ the most. According to Directorate General Mobility and Transport [8], bus SQ's most important criteria were frequency, cleanliness, and safety. The results show that 70% of the passengers are satisfied with the bus frequency and punctuality. Verbich and Geneidy [9] identified different passenger groups such as the disabled, luggage-carrying. They determined that different types of passenger groups have different needs and priorities. While information is a critical priority for disabled passengers, available spaces for those carrying luggage are essential.

Ardıç and Sadaklıoğlu [10] conducted a local study on customer-based satisfaction levels and demands regarding the measurement and development of SQ levels of institutions that show intercity passenger transportation activities. In the study, they aimed to reveal the SQ levels for the target groups. The research determined seven different attitudes and behaviors of employees that affect customer satisfaction as bus features, timeliness, accommodation and break, reservation procedures, office operations, shuttle service, and luggage operations. Imre and Çelebi [11] have linked user satisfaction to several factors for public transportation systems. They expressed that transportation public services and customer satisfaction should be designed in a connected way to increase public transportation use. The study examined the effect of comfort and convenience on public transportation using qualitative and quantitative data. The research provides a new method to evaluate comfort in public transport systems. It has developed a new index based on a series of qualitative and quantitative indicators. It also provides a better understanding of passenger attitudes towards public transport and the perception of comfort in public transport services.

Özuysal et al. [12] have investigated the effects of utility-based accessibility, one of the SQ criteria on transportation mode choice. They have found that linear approaches like correlation and elasticity analysis, the accessibility indices like residential and social infrastructure, are effective but not enough to predict aggregate mode choice ratios. On the other hand, ANN, as a nonlinear approach, has high prediction capability, especially for private car choices. Doğan and Özuysal [13] have investigated the factors affecting waiting time in public transportation in terms of reliability, passenger information system, and physical condition. As a result of the constituted linear multiple regression models, they have found that passenger information system and reliability have reduced the effect on waiting time by 2 and 1.3 minutes, respectively. Besides, it is concluded that the passenger information systems are more efficient if implemented properly for the stops respecting their geometry and the ones used by the lines coming from the city center.

The bus system is the type of public transportation service preferred by many city administrations due to its low investment cost. For this reason, in many cities, public transportation services are provided by a network of bus lines. Therefore, studies on the SQ of public transport have focused primarily on the SQ of buses. Kahraman and Yıldız [14] aimed to measure the quality of service offered by bus companies that provide public transportation services by associating customers' expectations in addition to their perceptions and demographic characteristics. As a result of the research, they observed that a significant relationship exists between customer satisfaction and demographic characteristics, such as gender, age, education, and income level. In the research conducted by Koçoğlu and Aksoy [15] on 274 customers, they investigated the relationship between the demographic characteristics and their satisfaction with an agency that provides public bus transportation. As a result of the research, there was a relationship between the customer age and satisfaction, while there was no relationship between gender, educational status, income status, and customer satisfaction. In their research, Gökaşar et al. [16] aimed to evaluate Istanbul's bus services. As a result of the online surveys conducted on 2177 people, they determined that the users were least satisfied with the frequency of the buses, comfort of the seats, the availability of sufficient space, and ease of moving within the bus, while they were most satisfied with the convenience of the payment method, the clean clothing of the employees and the benefit of the buses to reduce environmental pollution. Gökaşar et al. [17] created a multiple linear regression model using factor analysis to evaluate the service quality of the buses using 2177 satisfaction surveys. When the two factors that emerged were analyzed, they revealed that users attach more importance to service access than comfort. They compared the user survey results with the service providers' opinions and observed only a 1.733% difference between the two stakeholder groups' opinions.

As a result of the increasing population and travel demand in recent years, the investments in urban rail systems have also increased. Therefore, in recent years studies investigating the SQ of rail systems have also been started to increase. In their study, Girginer and Cankuş [18] examined the customer satisfaction of public trams. The study carried out in Eskişehir examined the quality of service expectations of students from the public transport companies using data obtained from approximately 300 students. They used a latent variable at the binomial level related to student satisfaction in their model. As a result of the applied Binomial Logistic Regression Analysis, all the independent variables included in the model had adverse effects on students' satisfaction from the Eskişehir tram system (Estram). Hemedoğlu [19] aimed to determine the SQ of the agencies that provide urban public

transportation services and determine the user satisfaction levels regarding a metro line. In the research, they interviewed 761 consumers using the face-to-face survey technique. As a result of the research, customer expectations for all SQ variables were significantly higher than their perceptions. The variable with the lowest difference between the level of SQ consumers expect and the level of SQ they perceive is the presence of direction signs in the stations. According to the research, the most significant difference between the level of SQ that consumers expect and the level of SQ they perceive was the crowdedness of vehicles. Seçilmiş et al. [20] tried to evaluate the SQ of public transportation vehicles that provide rail service by investigating customer expectations regarding comfort, accuracy, timed, and service facilities. In the results of the study, they stated that there is a significant and positive relationship between the "comfort", "service", "information and accuracy", and "time" sub-dimensions of the problems regarding SQ in railways. Akyıldız et al., [21] aimed to measure the quality of service of *the high-speed rail system* (HSRS) serving between the two cities in Turkey. They prepared the questionnaire to ask users to rate satisfaction ratings from 61 features created for seven SQ dimensions for public transport systems: passenger information, fare level and type, accessibility, station environment, vehicle environment, service delivery and security. By analyzing the results, they identified the most problematic areas of the service provided by HSRS. According to the model results, the system operator should provide improvements, especially in the vehicle seat design and the time and frequency of the service. Šojat et al. [22] analyzed the tram network of Zagreb using two hypothetical tram priority scenarios. Study results show that by implementing yellow line enforcement, 7% better operating speed, 7% saving in passenger spaces, and 22 million € investment cost savings can be obtained. However, if absolute tram priority is applied, the operating speed is 41% better, with 29% saving in passenger spaces and 94 million € of investment cost savings, resulting in a significant increase in the tram's SQ. Vujicic and Prester [23] applied the P-Transqual model to public tram transport in Zagreb to assess its service quality. Their study investigated trams service quality using four quality dimensions: comfort, tangible, personnel, and reliability. They have observed 50, 8, 4, and 23 defects in terms of comfort, tangible, personnel, and reliability, respectively, in 10 days. However, they stated that the results were still positive because, in 9 of 10 cases (90%), public tram transport quality was satisfactory. Khelf et al. [24] investigated the tram service quality and its impact on the passenger modal choice in Constantine City (Algeria). The study results show that despite the tram's different advantages, the population is not satisfied with its exploitation and service quality. Almost 40% of passengers declared that they preferred to use their cars or taxis to using tram. They also provided some suggestions to increase the performance and the SQ of the tram in the study.

ANN is a useful method to predict people's decisions [25]. In SQ studies, ANN is used to estimate which criteria are more important than others. For example, De Oña and De Oña [25] proposed a model that uses ANN to analyze the SQ perceived by passengers of a public transport system. In 2007, they carried out a customer satisfaction survey on the Granada metropolitan bus transport system. They developed an ANN model to investigate customer satisfaction. They used to determine each feature's relative contribution and applied a statistical analysis to each method's outputs to identify groups of features that differ significantly in their relative importance. Statistical results have shown that frequency is the most useful feature in SQ. Islam et al. [26] proposed various ANN approaches to define which feature of the SQ criteria is more important than others. They presented a comparative

study using *Generalized Regression Neural Network* (GRNN), *Probabilistic Neural Network* (PNN), and *Pattern Recognition Neural Network* (PRNN). According to the results, punctuality and reliability, frequency of service, free seats, and travel experience were essential.

3. METHODOLOGY

Within the scope of this study, a 20-question survey created by the *International Bus Benchmarking Group* (IBBG) [27] was adapted for the *light rail transit* (LRT) system. User preferences were derived from the answers given to the first 19 questions using the ANN model. The ANN model developed can predict passenger SQ perception using the survey results and how a change in the public transport system's feature affects SQ accurately. Therefore, the *Perturb Method* suggested by De Oña and De Oña [25] has been adapted to develop a new and more straightforward estimation method, namely the *Change Of a Score* (COS) *Method*. The ANN model created using this method can also predict the SQ perception of different user groups, so decision-makers can focus on meeting various customer groups' needs.

Kocaeli, located in the Marmara Region of north-western Turkey, is the tenth most populous city of Turkey with a population of 1,953,035. The city is one of the largest industrial and commercial cities in the country. Kocaeli is the most developed city after Istanbul, Ankara, and Izmir according to the Socio-Economic Development Ranking of Regions – SEGE 2017 [28] ranking. It is a neighbor of important cities, such as Istanbul and Yalova in the west, Sakarya in the east, and Bursa in the south. Due to these features, it is a critical travel generation and attraction center.

Kocaeli Metropolitan Municipality presented the tram service to the public in 2017. They determined the most suitable route by evaluating passenger demand, integration with other transportation systems, and construction costs. The tramline named *Akçaray*, consists of 15 stations along 14.8-kilometer line between Otogar and Sekapark. The route of Akçaray is presented in Figure 1.



Figure 1 - Route of Akçaray [29]

The research was carried out in the survey model to examine user satisfaction with Kocaeli tram system. The research subjects are defined in the conditions they exist and are not attempted to be changed in any way. The study sample consists of individuals over the age of 15 who live in Kocaeli in May 2019 and benefit from the tram system in public transport. According to the *Turkish Statistical Institute* (TSI) address-based information system [30], the number of persons with the age of 15 or over living in Kocaeli Centre in 2019 are 289577. Meyer [31] and Fox et al. [32] suggest the values shown in Table 1 for the population range and the corresponding sample size. For the universe of non-homogeneous subjects, the required sample size was calculated as $n = 384$. To reach more general results, the sample size was selected as 500 people.

Table 1 - Population range and approximate sample size [31, 32]

Population Range	Approximate Sampling Size (n)
Infinity	384
500000	384
100000	383
50000	381
10000	370
5000	357
3000	341
2000	322
1000	278

In the research, the data were collected using the surveys that were conducted online in May 2019. The survey was published on a web site and was advertised using various channels such as social media and text messages to members of non-governmental organizations located in Kocaeli. The questionnaire form consists of two parts. The first part consists of 7 questions that aim to obtain the participants' demographic information, such as age and gender. Descriptive statistics are shown in Table 2. The most of the respondents were males, with a percentage of 65.6. The respondents were from various education levels, but mostly with a bachelor's degree. The percentage of married and single (or divorced) respondents are close to each other with percentages of 54.2 and 45.8, respectively. 47.2% of the respondents were in the age range of 18-30, and 47.6% had a monthly income between 2000-5000 TL. Most of the respondents (53%) used the tram service less than once a month, and 46.4% used it for hobby/socializing journeys. In the second part, a questionnaire consisting of 20 questions to measure the SQ of buses developed by the International Bus Benchmarking Group was adapted to the tram and used. The answers to the survey questions were evaluated using the 5-point Likert scale ranging from 1 to 5. One indicates the worst case (strongly disagree), and five indicates the best case (strongly agree). 44 of 500 participants left some questions empty and contained missing information. Therefore, these surveys were not taken

into consideration in developing the ANN model. Table 3 shows the average of the scores given to each survey question.

Table 2 - Descriptive statistics

Groups	Frequency (n)	Percentage (%)
Sex		
Male	328	65.6
Female	172	34.4
Education		
Elementary School	27	5.4
High School	69	13.8
Associate	51	10.2
Bachelor's	279	55.8
Graduate	70	14.0
Post Graduate	4	0.8
Marital status		
Married	271	54.2
Single (or divorced)	229	45.8
Age Range		
<18	2	0.4
18-30	236	47.2
31-40	153	30.6
41-50	55	11.0
51-60	37	7.4
61-70	15	3.0
>70	2	0.4
Monthly Income Range		
< 2000 TL	98	19.6
2000-5000 TL	238	47.6
5001-7000	94	18.8
>7000 TL	70	14.0

Table 2 - Descriptive statistics (Continued)

Frequency of Tram Use	
Less than once a month	265 53.0
Between once a month and once a week	83 16.6
Between once a week and thrice a week	82 16.4
Between thrice a week and six times a week	57 11.4
Everyday	13 2.6
Purpose of Tram Use	
Hobby/socializing journeys	232 46.4
Work/school journeys	160 32.0
Shopping journeys	127 25.4
Doctor/hospital journeys	32 6.4
Other	20 4.0

3.1. ANN Model

ANN is an artificial intelligence method developed based on the human brain's fundamental functions such as learning, remembering, generalizing, and memorizing. The ANN consists of interconnected processing units called neurons, developed in the form of layers. Learning in ANN is done by adjusting the values of the connections (weights) between neurons. There are different network structures developed in the literature. In this study, the feed-forward neural network type, a *multilayer perceptron (MLP)* procedure, and error backpropagation algorithm were used to develop the ANN model. In backpropagation algorithm, which is a type of supervised learning, the difference (error) between the actual (observed) result and the result produced by ANN is propagated backward by updating the weight values in the network towards the previous layers to reduce the obtained error [33].

In ANN, a black box modeling technique, instead of deriving a mathematical relationship between input values and output values, aims to produce appropriate output values from the given input values. Therefore, what mathematical operations are performed between input and output or how they are performed is not essential. For this reason, the number of hidden layers between the input and output layers and the number of neurons in the hidden layer(s) are found in a way that optimizes the performance of the model using the *trial-and-error method*. It is also crucial for the model to be as simple as possible, reducing the amount of necessary calculations. For this reason, model development is carried out by taking into account the trade-off between model simplicity and performance [34].

An ANN model has been developed to mimic the satisfaction of Akçaray users. In the developed model, the demographic variables and the first 19 questions asked to obtain the users' SQ satisfaction level were used as the input values. The only output value of the ANN was the answer to the 20th question of the survey, "How satisfied are you overall with the

tram services in the city?”. In other words, the ANN model tries to estimate the level of satisfaction of the users from the system by using the demographic data of the users and the satisfaction levels of different aspects of service offered by the tram.

Table 3 - Average score given to survey questions

Nr.	Question	Average Score
1	The tram service is usually reliable	3.798246
2	The trams are frequent	3.464912
3	The tram route is convenient for me	3.405702
4	It is easy for me to get on and off the tram	3.688596
5	It is easy to move inside the tram	3.541667
6	It is convenient to pay for the tram	3.800439
7	It is easy to get information about the tram services	3.392544
8	It is easy to find out how well the trams are running	3.559211
9	If there are problems, I can easily find an alternative route	3.015351
10	The tram gets me to my destination on time	3.642544
11	Staff are well dressed	3.710526
12	Staff are helpful	3.480263
13	It is easy to resolve problems or complaints	3.368421
14	The seats are comfortable, and there is enough space	2.980263
15	The tram is well driven and gives a comfortable ride	3.541667
16	The tram is clean	3.587719
17	The tram is quiet, well lit, ventilated, and at an adequate temperature	3.592105
18	The tram is a secure place for me	3.640351
19	The tram does not cause too much pollution	3.760965
20	How satisfied are you overall with the tram services in the city?	3.578947

Since only gender and marital status belong to 2 different classes among the input values of ANN, they were represented by a single neuron. For example, the input value is 0 when the user was male and 1 when the user was female. For other demographic data, input neurons as many as the number of classes were used. Education level was represented by 6, the age range was represented by 7, the monthly income range was represented by 4, frequency and purpose of tram use were represented by 5 neurons each. For each demographic data, the

neuron representing this class takes the value 1, and the others take the value 0. For example, for a 45-year-old user, the 4th of the neurons showing the age range takes the value 1, while the remaining ones take the value of 0. A single neuron was used for each of the 19 questions of the survey. Since the answers to the questions ranged from 1 to 5, this range was normalized between 0 and 1 using Eq. 1. In other words, 0 indicates the lowest level of satisfaction, and 1 indicates the highest level of satisfaction. The answers to the survey's 20th question, which constitute the single output of the ANN model, were also normalized between 0 and 1, just like the other 19 questions representing the survey's other questions.

$$\text{Normalized Value} = \frac{\text{Answer given to the question} - 1}{4} \quad (1)$$

In the ANN model, which was developed using the Matlab® Neural Networks Toolbox, 48 neurons were forming the input and a single neuron forming the output, while the number of hidden layers and the number of neurons in the hidden layer were determined by trial-and-error approach suggested by Murat and Başkan [35] due to its simplicity. The sigmoid function in Eq. 2 was selected as the activation function in the hidden layer and a linear function on the output layer. Levenberg Marquardt (LM) algorithm was preferred in the training of ANN because of its fast convergence and its need for fewer iterations. The parameters of the LM algorithm were selected as initial $\mu = 0.001$, μ decrease factor = 0.1, μ increase factor = 10, maximum $\mu = 10000000000$, which are also the default values of MATLAB®. The ANN model developed can be seen in Figure 2.

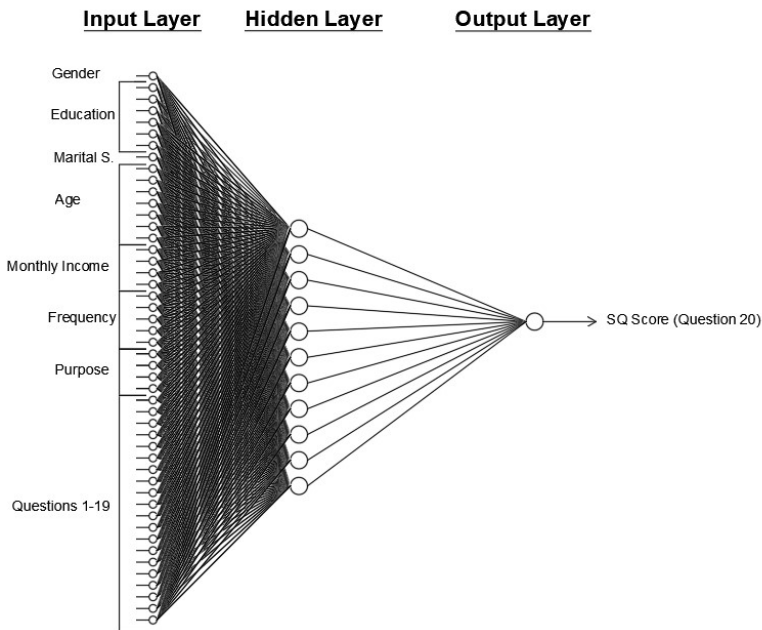


Figure 2 - The ANN model developed

$$\sigma(x) = \frac{1}{1+e^{-x}} \tag{2}$$

The outputs obtained from the ANN model were denormalized between 1 and 5 using Eq. 3 for the evaluation of its performance.

$$\text{Denormalized Value} = \text{Normalized Value} \times 4 + 1 \tag{3}$$

The three-way data split method by Osuna [36] was used to prevent the over-fitting problem and determine the best network architecture. The flow chart of the three-way data split methodology is given in Figure 3. First, the data were divided into three parts as training (80%), test (10%), and validation (10%) sets. Then the best network architecture was investigated considering the combinations of the data sets given in Table 4. For each partition, the network architecture was trained and tested with the test dataset that was not used in training. The best network architectures obtained from different trials and partitions and corresponding error values are given in Table 5. For each run, the best network architecture was selected with the best network performance. One and two hidden layers with a maximum of 15 neurons were tried at each layer.

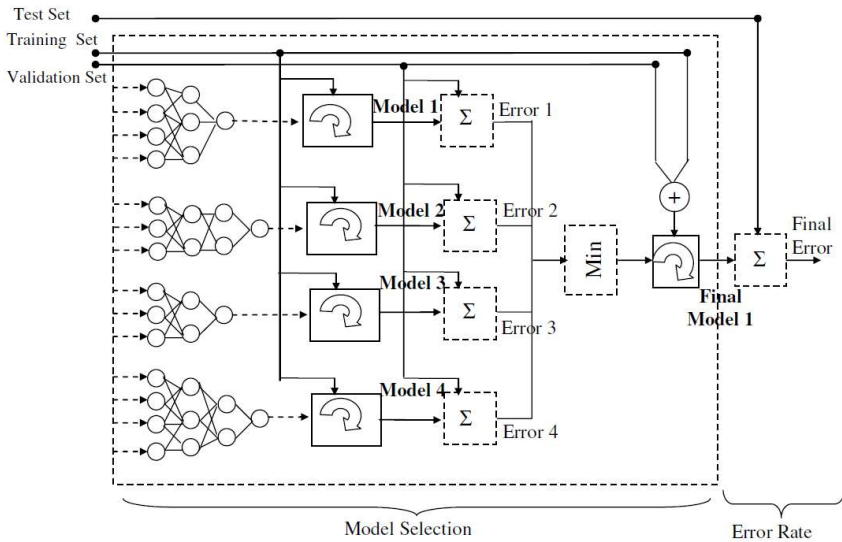


Figure 3 - Flow chart of the three-way data split validation method [37]

Table 4 - Partitions of the data used

Partition I	Training				Validation	Test
Partition II	Training			Validation	Test	Training
Partition III	Training		Validation	Test	Training	
Partition IV	Training	Validation	Test	Training		
Partition V	Validation	Test	Training			

Table 5 - Samples from network architectures searching and corresponding errors

Partition no	Best network architecture	Train errors	Validation Errors	Test Errors
I	$48 \times 10 \times 1$	0.348	0.324	0.434
II	$48 \times 15 \times 1$	0.391	0.427	0.425
III	$48 \times 9 \times 1$	0.361	0.410	0.434
IV	$48 \times 11 \times 1$	0.302	0.358	0.391
V	$48 \times 12 \times 1$	0.371	0.272	0.322

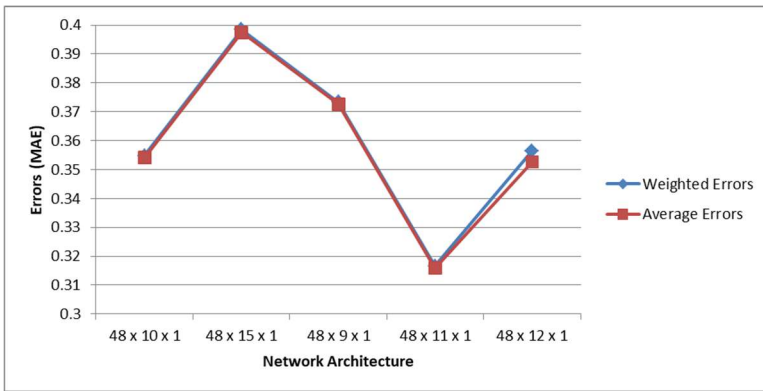


Figure 4 - Weighted and average error values for different network architectures.

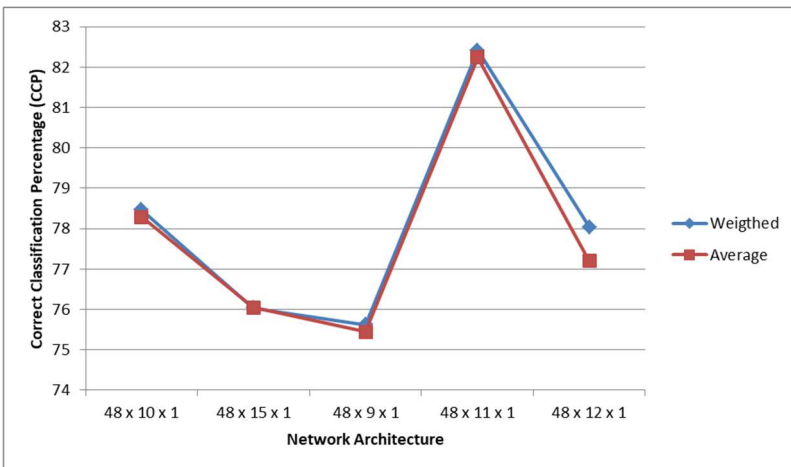


Figure 5 - Weighted and average correct classification percentages for different network architectures.

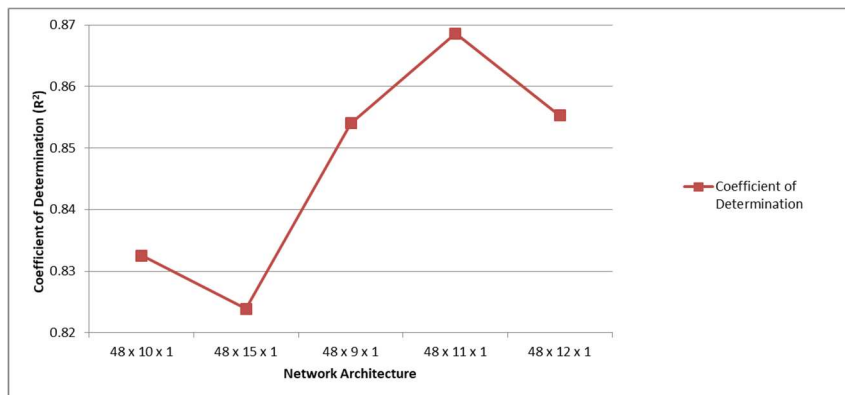


Figure 6 - Coefficient of determination values for different network architectures.

Three criteria, the correct classification percentage (CCP), the Mean Absolute Error (MAE) values, and coefficient of determination (R^2) values, were taken into consideration for searching the best network architecture using the train, test, and validation data. The ultimate network architecture was determined as $[48 \times 11 \times 1]$ based on the evaluations. The MAE, CCP, and R^2 values of different network architectures are indicated in Figures 4, 5, and 6, respectively.

Since most of the respondents (53%) were not regular users and used the tram service for less than once a month, another ANN has been developed and trained using the data of the 223 users who use the system at least once a week. For this reason, the input neuron that corresponds to non-regular users has been removed from the ANN. The same method has been applied to the development of the ANN for regular users. The best network architecture for every partition was obtained as $[47 \times 12 \times 1]$ based on the evaluations. The error values obtained for each partition and MAE CCP, and R^2 values of the network architecture are indicated in Table 6.

Table 6 - Errors of ANN developed using regular users

Partition no	I	II	III	IV	V
Train errors	0.087	0.076	0.088	0.095	0.091
Validation errors	0.154	0.194	0.039	0.066	0.108
Test errors	0.027	0.076	0.126	0.050	0.037
Weighted error	0.087	0.087	0.087	0.087	0.088
Weighted CCP	0.960	0.959	0.960	0.964	0.959
Overall					
Average error			0.087		
Average CCP			0.960		
R^2			0.943		

Although the ANN developed using regular user data only seems to provide better results; it cannot predict non-regular users' evaluations. The predicted values for both ANN models have been compared with each other and are shown as a scatter diagram in Figure 7. R^2 value between the two ANNs predicted values was found as 0.825, and both models predicted the same scores for 77% of the regular users. These values show that although most of the respondents were non-regular users due to the sampling method used, ANN can still learn and mimic every user type's opinions.

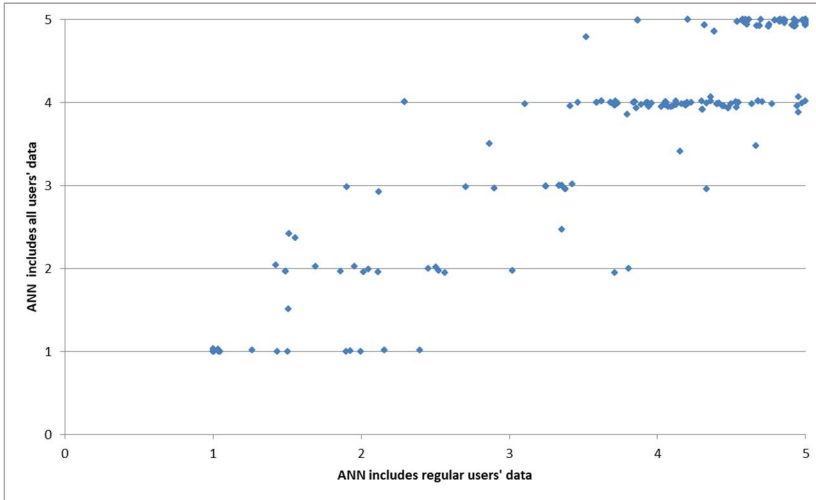


Figure 7 - Scatter diagram between the two ANN models' predicted values

3.2. Multiple Linear Regression Model

A *Multiple Linear Regression (MLR)* model has been developed to compare the ANN model performance. Although some machine learning methods such as *Support Vector Machines (SVM)*, *k-nearest neighbor*, and *random forest* may provide better classification performance, MLR was used in performance evaluation because it provides continuous output results just like the ANN model. For this reason, the COS method described in the next subsection can also be applied to the MLR model. The data split method was also used in the MLR model to evaluate the model prediction performance. However, this time the data was split into two-ways as training and test set. Hence, the validation set of each partition was included in the test set. The performance obtained for each partition using the ANN and MLR models are given in Table 7. Results show that the ANN model performs better in estimating the SQ perception of tram system users than the MLR model.

MLR developed for partition III provides the best performance as seen in Table 5. For this reason, this model was taken into consideration. The formulation of the MLR model is given in Eq. 4 where SQ is dependent variable, β 's are the coefficients and x_i 's are the independent variables.

$$SQ = \beta_0 + \beta_1 \times x_1 + \beta_2 \times x_2 + \dots + \beta_n \times x_n \quad (4)$$

The goodness of fit statistics, analysis of the variance, and MLR model parameters are given in Tables 8, 9, and 10, respectively.

Table 7 - Comparison of ANN and MLR methods

Partition no	Weighted MAE		Average MAE		Weighted CCP		Average CCP		R ²	
	ANN	MLR	ANN	MLR	ANN	MLR	ANN	MLR	ANN	MLR
I	0.355	0.418	0.354	0.414	78.466	72.382	78.289	73.210	0.833	0.808
II	0.398	0.424	0.398	0.424	76.031	66.006	76.046	65.755	0.824	0.804
III	0.373	0.411	0.373	0.410	75.616	73.900	75.439	73.632	0.854	0.797
IV	0.317	0.428	0.316	0.427	82.419	69.738	82.237	69.475	0.869	0.806
V	0.357	0.410	0.353	0.409	78.038	70.395	77.193	70.131	0.855	0.809

Table 8 - Goodness of fit statistics of MLR model

Observations	365
Sum of weights	365
DF	320
R ²	0.825
Adjusted R ²	0.801
MSE	0.328
RMSE	0.573
MAPE	14.038
DW	2.102
Cp	45.000
AIC	-364.768
SBC	-189.273
PC	0.224

Table 9 - Analysis of variance of MLR model

Source	DF	Sum of squares	Mean squares	F	Pr > F
Model	44	495.532	11.262	34.323	0.000
Error	320	105.000	0.328		
Corrected Total	364	600.532			

Table 10 - MLR model parameters

Source	β Value	Standard error	t	Pr > t	Lower bound (95%)	Upper bound (95%)
Intercept	0.187	0.673	0.277	0.782	-1.138	1.512
Sex	-0.077	0.081	-0.951	0.342	-0.237	0.083
Elementary School	0.166	0.456	0.363	0.717	-0.732	1.064
High School	0.389	0.435	0.893	0.372	-0.467	1.245
Associate	0.102	0.435	0.234	0.815	-0.754	0.959
Bachelor's	0.261	0.426	0.613	0.541	-0.577	1.100
Graduate	0.357	0.431	0.827	0.409	-0.492	1.206
Post Graduate	0.000	0.000				
Marital	-0.001	0.086	-0.009	0.993	-0.171	0.169
<18	-0.239	0.642	-0.372	0.710	-1.501	1.024
18-30	0.273	0.451	0.605	0.546	-0.614	1.159
31-40	0.207	0.445	0.465	0.642	-0.668	1.081
41-50	0.247	0.452	0.546	0.586	-0.643	1.137
51-60	0.245	0.455	0.540	0.590	-0.649	1.140
61-70	-0.175	0.476	-0.368	0.713	-1.112	0.762
>70	0.000	0.000				
<2000	0.132	0.137	0.968	0.334	-0.136	0.401
2000-5000	0.088	0.116	0.756	0.450	-0.141	0.317
5001-7000	0.028	0.117	0.237	0.813	-0.202	0.258
>7000	0.000	0.000				
<once a month	-0.592	0.195	-3.036	0.003	-0.975	-0.208
oncem.-oncew.	-0.459	0.209	-2.193	0.029	-0.871	-0.047
oncew.-thricew.	-0.319	0.196	-1.623	0.105	-0.705	0.068
thricew.-everyday	-0.303	0.198	-1.528	0.127	-0.692	0.087
everyday	0.000	0.000				
hobby	-0.043	0.093	-0.460	0.646	-0.226	0.140
workschool	0.100	0.088	1.146	0.253	-0.072	0.272
shopping	-0.015	0.086	-0.175	0.861	-0.185	0.154
medical	0.039	0.140	0.277	0.782	-0.236	0.313
other	-0.063	0.172	-0.368	0.713	-0.401	0.274

Table 10 - MLR model parameters (continue)

Source	β Value	Standard error	t	Pr > t	Lower bound (95%)	Upper bound (95%)
Q1	-0.107	0.066	-1.620	0.106	-0.237	0.023
Q2	0.041	0.047	0.872	0.384	-0.051	0.133
Q3	0.215	0.051	4.202	0,000	0.114	0.316
Q4	-0.134	0.058	-2.326	0.021	-0.247	-0.021
Q5	0.029	0.054	0.528	0.598	-0.078	0.135
Q6	-0.052	0.069	-0.756	0.450	-0.189	0.084
Q7	0.038	0.058	0.648	0.517	-0.077	0.152
Q8	0.047	0.063	0.746	0.456	-0.077	0.170
Q9	-0.068	0.041	-1.667	0.097	-0.148	0.012
Q10	0.167	0.063	2.633	0.009	0.042	0.292
Q11	0.074	0.079	0.938	0.349	-0.082	0.231
Q12	0.174	0.086	2.031	0.043	0.005	0.342
Q13	-0.157	0.077	-2.033	0.043	-0.308	-0.005
Q14	-0.030	0.042	-0.713	0.476	-0.111	0.052
Q15	-0.078	0.069	-1.124	0.262	-0.214	0.059
Q16	0.172	0.081	2.137	0.033	0.014	0.331
Q17	0.078	0.077	1.010	0.313	-0.074	0.229
Q18	0.250	0.078	3.217	0.001	0.097	0.403
Q19	0.260	0.067	3.866	0.000	0.128	0.392

The results of the MLR indicated the predictors explained 82.5% of the variance ($R^2 = 0.825$, $F(2,320) = 34.323$, $p < 0.05$). The analysis shows that using the tram less than once a month ($\beta = -0.592$, $t(364) = 0.003$, $p < 0.05$), using the tram between once a month and once a week ($\beta = -0.459$, $t(364) = 0.029$, $p < 0.05$), Question3 ($\beta = -0.215$, $t(364) = 0.000$, $p < 0.05$), Question4 ($\beta = -0.134$, $t(364) = 0.021$, $p < 0.05$), Question10 ($\beta = 0.167$, $t(364) = 0.009$, $p < 0.05$), Question12 ($\beta = 0.174$, $t(364) = 0.043$, $p < 0.05$), Question13 ($\beta = -0.157$, $t(364) = 0.043$, $p < 0.05$), Question16 ($\beta = 0.172$, $t(364) = 0.033$, $p < 0.05$), Question18 ($\beta = 0.250$, $t(364) = 0.001$, $p < 0.05$) and Question19 ($\beta = 0.260$, $t(364) = 0.000$, $p < 0.05$) did significantly predict the SQ value. The value of β indicates the order of importance of the independent variables. The variable with the highest β value is relatively most important independent variable [38]. Positive β values indicate that increasing the independent variable also increases the SQ value, whereas negative β values indicate that increasing the independent variable decreases the SQ value.

3.2. Change of Score Method

The developed ANN model was then used to examine to what extent each of the dimensions related to the tram system's SQ affected the general level of satisfaction from itself. For this purpose, first of all, the most frequent user profile has been defined. The highest frequency features were selected from the demographic variables (shown in Table 2) while creating the profile. However, the most frequent user was defined as a person who uses the tram every day. Therefore, the profile determined as a standard user is a male, with an undergraduate degree, married, between the ages of 18-30, with a monthly income of 2000-5000 TL, using the tram every day, and mostly for hobby/socialization journeys. The score that such an individual would give to each SQ dimension was determined by taking the mean of the total scores given to them. This individual's characteristics and preferences were presented in the ANN model. An overall satisfaction score of 4.42 was obtained as the overall SQ score of the most frequent user. Since none of the 456 respondents fitted precisely into the defined profile, such a user's actual opinion could not be obtained.

To measure the effect of dimensions on the overall SQ, *the Change Of a Score (COS)* method, similar to the Perturb Method suggested by De Oña and De Oña [25], has been developed. In the Perturb Method, small amounts of noise are applied to each input neuron until 50% of the original value is perturbed, while the remaining input neurons keep unaltered. The change in the Mean Square Error (MSE) assesses each attribute's relative importance; therefore, the larger the MSE for each input perturbation, the more the corresponding attribute's relative importance. In the developed COS method, instead of perturbing the original values by a small amount of noise, the values given by the most frequent user to each dimension were increased and decreased by 1 point, respectively. The most frequent user's modified preferences are presented to the ANN model, and their output is considered as the overall SQ score. By dividing the difference between these two values into two, how the 1-point change in SQ dimensions will change the overall satisfaction level has been calculated. This method can be considered as a modified and simplified version of the perturb method. An example of the COS method is as follows:

The most frequent user's overall score for the SQ, in other words, the ANN model's output is 4.42283. Similarly, his score for the first dimension, the tram service is usually reliable, is 3.79824. First of all, the first-dimension score is increased by 1 point, keeping every other input value constant, obtaining a new score of 4.79824. The most frequent user's new opinion is supplied to the ANN model, obtaining an overall score of 4.52438. It means that 1 point of an increase in dimension 1 increases the SQ perception of the most frequent user by 0.10155. Then, the first-dimension score is decreased by 1 point, obtaining a new score of 4.28902. Furthermore, this shows that 1 point of a decrease in dimension 1 decreases the SQ perception of the most frequent user by 0.13380. The average of both values, 0.11768, shows the impact of dimension 1 on the overall SQ.

4. RESULTS

Figure 8 shows the scatter diagram between user satisfaction with SQ and trained ANN's predicted values. When the figure is examined, it is seen that the outputs of the ANN model is not only composed of integers between 1 and 5 but also includes intermediate

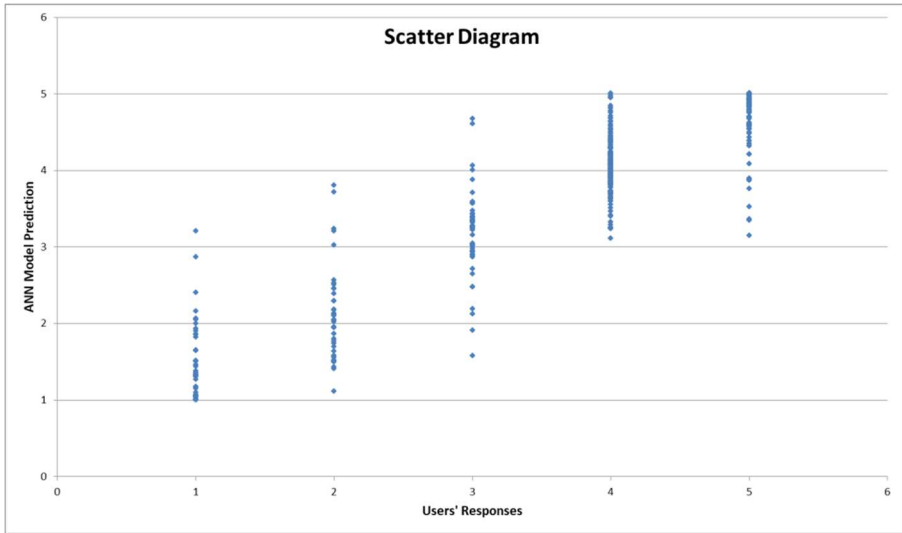


Figure 8 - Scatter diagram between user satisfaction with SQ and ANN's predicted values

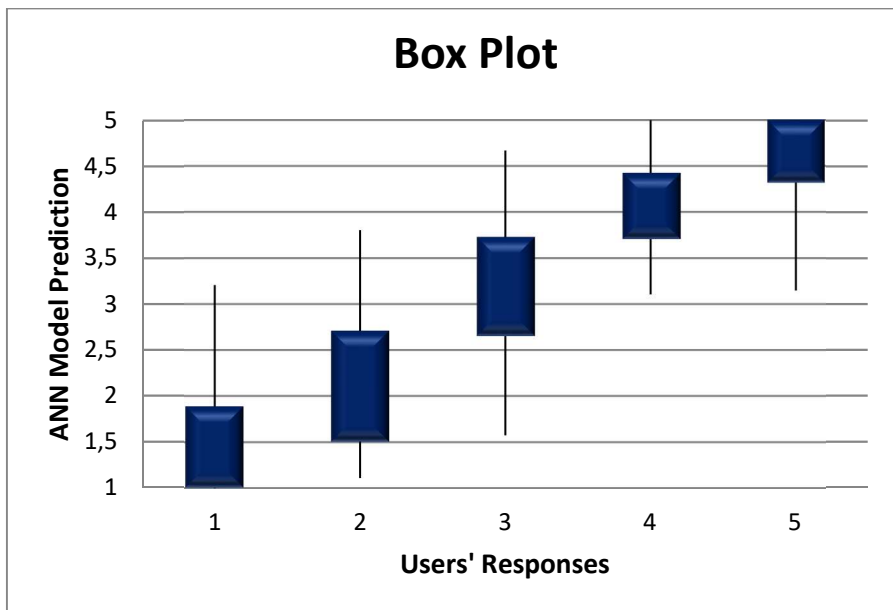


Figure 9 - Box plot showing the mean and the standard deviations estimated by ANN for each respondent with respect to their overall SQ satisfaction

values. This feature also investigates how the changes of SQ dimensions related to the tram system affect their overall satisfaction with the tram service. R^2 value between the users SQ scores and ANN's predicted values was found as 0.8604 and Pearson correlation statistic (r) as 0.928. The box plot that shows the mean and the standard deviations estimated by ANN for each respondent regarding overall SQ satisfaction can be seen in Figure 9. The developed ANN model correctly estimated the satisfaction score of 82% of the respondents with an error margin of ± 0.5 and 97% of the respondents with an error margin of ± 1 . This performance stands out as an appropriate value, especially considering the stochastic nature of human thoughts and behaviors and also that different people with similar characteristics may have different evaluations.

Table 11 shows the most frequent user scores on SQ dimensions and each dimension's effect on the overall SQ. The values shown in the table show that the user is most satisfied with the convenience of paying for the tram, the tram's reliability, and the trams being useful in reducing environmental pollution. On the other hand, he/she is least satisfied with the comfort of the seats and the space inside the trams, finding an alternative route in case of problems and ease to resolve problems or complaints. Furthermore, the dimensions that most affect the overall satisfaction level of the most frequent user from tram service are the convenience to pay for the tram, the tram getting him/her to his destination in time, and the benefits of trams in reducing the environmental pollution. The cleanliness of the trams, their frequency, and dress of the staff affect the level of general satisfaction the least. Figure 10 shows the most frequent user score of each dimension and its impact on overall SQ.

Table 11 - The most frequent user scores on SQ dimensions and each dimension's effect on the overall SQ

Dim.	Question	Most Freq. User Score	Impact on Overall SQ
6	It is convenient to pay for the tram	3.800439	0.42324
10	The tram gets me to my destination on time	3.642544	0.32150
19	The tram does not cause too much pollution	3.760965	0.30908
18	The tram is a secure place for me	3.640351	0.29437
4	It is easy for me to get on and off the tram	3.688596	0.20247
3	The tram route is convenient for me	3.405702	0.17710
9	If there are problems, I can easily find an alternative route	3.015351	0.15540
7	It is easy to get information about the tram services	3.392544	0.13452
12	Staff are helpful	3.480263	0.12301
1	The tram service is usually reliable	3.798246	0.11768
17	The tram is quiet, well lit, ventilated, and have adequate temperature	3.592105	0.11643

Table 11 - The most frequent user scores on SQ dimensions and each dimension's effect on the overall SQ (continue)

Dim.	Question	Most Freq. User Score	Impact on Overall SQ
5	It is easy to move inside the tram	3.541667	0.11368
8	It is easy to find out how well the trams are running	3.559211	0.10693
15	The tram is well driven and gives a comfortable ride	3.541667	0.10401
13	It is easy to resolve problems or complaints	3.368421	0.07868
14	The seats are comfortable, and there is enough space	2.980263	0.06630
11	Staff are well dressed	3.710526	0.04185
2	The trams are frequent	3.464912	0.04025
16	The tram is clean	3.587719	0.03815

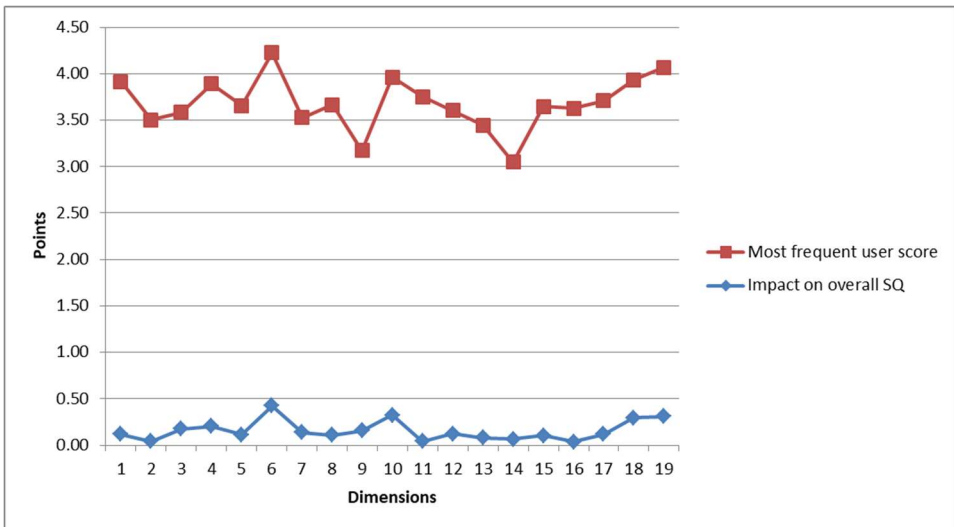


Figure 10 - The most frequent user score of each dimension and its impact on overall SQ.

5. DISCUSSIONS

Within the scope of this study, general satisfaction level of tram services were tried to be estimated using the demographic characteristics of tram users in Kocaeli and their evaluation for tram SQ dimensions. The ANN model developed for this purpose has an accurate classification level of 82%. The values obtained indicate that the developed ANN model has a good performance level when the stochastic and complex human thoughts and behaviors

are considered. Thanks to the ANNs generalization feature, by which the developed model can predict how different users would evaluate the SQ.

After developing the ANN model, a most frequent user profile was created using the surveyed users' dominant demographic features. The most frequent user opinions about the dimensions that make up the quality of the service were also formed by taking the average of the participants' scores to the survey questions. The most frequent user seems to be entirely satisfied with the city's tram service, with a score of 4.42/5 obtained from ANN.

In the survey applied to the tram system users, only the satisfaction levels of 19 dimensions related to the tram system were considered, and how vital these levels were for them if ignored. The most frequent user's score for each dimension has been altered by 1 point width, and these preferences have been provided as an input to the ANN to determine each dimension's importance. The average of the changes in SQ perception when the dimension is improved and disturbed by 1 point shows the impact of the subject dimension into the overall SQ level, or in other words, it shows how much the tram users attach importance to this feature.

According to the survey results applied, the most frequent user is most satisfied with the convenience to pay for the tram, reliability of the tram service, and the positive environmental effects of the tram. In recent years, the implementation of smart cards instead of paying with money, ticket, or coin in all transportation systems throughout the city and the integrated payment system seems to be the primary reason behind the satisfaction. A high satisfaction level from the tram service reliability shows that the tram operations do not deviate much from the announced time table. It also shows that the waiting time, which is the service variable that the users valued most, according to Dell'Olio et al. [2], is also low or predictable. This finding is also interesting from another point of view. According to Akyıldız et al. [21], the HSRS operator should improve the servicetime and frequency. Both studies reveal that the reliability of the rail system services is an essential aspect for their users. Since Kocaeli is an industrial city and its Dilovası district is the most polluted district of Turkey, it is not surprising that users are satisfied with trams' help in reducing the environmental pollution. The finding is similar to the study conducted by Dell'Olio et al. [2] since cleanliness is one of the service variables that users valued most.

Dimensions that most frequent user is least satisfied with are the seats being comfortable and providing enough space in vehicles, the possibility to find an alternate route in case of service disruption, and easiness to resolve problems or complaints. Comfort level in the vehicles is also found significant by Noor et al. [6] and Seçilmiş et al. [20]. However, this dimension was ranked as the fourth lowest in affecting the most frequent user's general SQ level, which is an interesting finding. Since the tram does not have an alternate route as parallel lines, the low satisfaction from finding an alternate route in case of problems is not surprising. Still, the passengers can use other public transportation systems, such as the city's bus transit services, in case of service disruption. Unlike the bus system, users do not communicate with the driver, or other staff of the tram system. Therefore, it is not surprising that their satisfaction with easiness to resolve problems or complaints is low.

The most frequent user's most important dimension in tram service is the convenience of paying for the tram. It is an interesting finding that this dimension ranked first in terms of both satisfaction and importance. It reveals that the operator knows their customers'

expectation, or in other words, their passengers' will and is successful in satisfying them, in some way. The most frequent user assigns the second highest importance to get his destination on time. This finding is not surprising considering that Kocaeli is the tenth most populated and fourth most developed city in Turkey, hence, people pay high importance to the time spent on the transportation system. It can be thought that the dimension to which the most frequent user attaches the third highest importance in terms of tram service is the beneficial in reducing the environmental pollution of trams. Because Kocaeli is an industrial city, and especially the high level of air pollution in the city affects this evaluation.

The trams' cleanliness, frequency, and dress of the staff stand out as the most frequent user's least important dimensions. However, the low importance of cleanliness assigned is contrary to Dell'Olio et al. [2] and Directorate General Mobility and Transport [8]. This finding reveals that while cleanliness is an important aspect for European cultures, it loses its importance in Turkey and some Asian countries like Malaysia [6] and Afghanistan [7]. The second lowest important dimension for the most frequent user, the tram frequency, is somewhat a controversial issue. One may consider that this finding is contrary to Directorate General Mobility and Transport [8] and Gökaşar et al. [16]. Akçaray provides services with an interval of 6 minutes at peak hours and 10-15 minutes at off-peak hours. Since the most frequent user is using the tram service every day and for hobby/socializing journeys, it reveals that waiting a few minutes less or more does not significantly affect his SQ perception. It is not surprising that the most frequent user attaches low importance to the staff's dress since the driver and other staff of the tram system have no or minimum level of communication with the passengers. It should be noted that this is the contrary to the bus system, from which the questionnaire has been adapted .

6. CONCLUSIONS

Using the COS method developed within this study, it can be examined to what extent a change of 1 point in each dimension that constitutes SQ will change the perception of general service quality. This 1-point change may be in the direction of improving or reducing the dimension of SQ. Due to this feature, the study differs from other studies evaluating the SQ in the literature. Furthermore, the developed method enables the decision-makers to determine which dimensions need betterment to improve the tram system users' general SQ perception. They can also evaluate to what extent improvements in dimensions will change the SQ perception of society. In this way, they can plan the investments to be made in the tram system and perform a benefit/cost analysis that will reveal the benefits to be provided in return for these investments. Also, it can be ensured that limited resources are optimally used. Besides, decision-makers can use this method to create different user profiles instead of most frequent user and measure the SQ perception of different users and develop improvement solutions specific to user types. In this way, it can be ensured that the vulnerable part of the society, especially the children, the elderly, and the disabled, benefit from this service more efficiently. Thus, it can be stated that the improvements to be put forward through the developed method will also support social justice.

There are some weak points of the method applied in this study. For example, the survey applied to tram users may have some questions developed by IBBG for the bus system that is not suitable for the tram. Developing a global benchmarking survey on tram would be

beneficial for more appropriate and objective assessments. Another issue that can be considered the weakness is that some user groups may be more dominant in this survey data, obtained from the internet users by using a random sampling method. Developing a method that will ensure uniform distribution of each group may provide a more appropriate evaluation method. However, thanks to ANN's ability to mimic user through its abilities to learn and generalize, it also reduces the applied method's weakness. Similarly, 53% of respondents using the tram less than once a month may create the impression that a group of non-regular users has the dominant assessment. In contrast, the developed ANN has been trained to include the opinions of users of each group. In the subsequent evaluation, a person who uses the system every day has been identified as the most frequent user. This way, the opinion of the users who know the system better has been evaluated. However, it is natural that users who use the system less than once a month also have opinions about the service, as they may have experienced and dissatisfied with the system and/or may have relatives who regularly use the tram. Thanks to the method applied, the opinions of these people were able to be added to the model. In this way, the kind of improvements applied to increase their tram usage can also be evaluated. This way, the decision-makers can develop user-oriented investment plans and strategies.

In future studies for tram user satisfaction, the development and implementation of a global survey structure specific to the tram system will enable more appropriate evaluations. Periodic application of such a survey will help assess the change of user opinions or the effects of different service implementations on the system. Besides, thanks to the developed method, evaluations specific to certain user groups such as women, children, and older adults can also be made. In this way, special applications can be made to increase each group's satisfaction and thus usage of tram system.

Acknowledgement

The author would like to thank the graduate thesis students Damla Yaşar and Cihan Bahadır Yaşar for their help in collecting and evaluating the survey data.

References

- [1] Yaşar, D., Kocaeli tramvay sistemi kullanıcı memnuniyetinin incelenmesi, M.Sc. Thesis, İstanbul Okan Üniversitesi, 2020 (In Turkish).
- [2] Dell'Olio, L., Ibeas, A., Cecin, P., The quality of service desired by public transport users. *Transport Policy*, 18(1), 217-227, 2011.
- [3] Friman, M., Fellesson, M., Service supply and customer satisfaction in public transportation: The quality paradox. *Journal of Public Transportation*, 12(4), 57-69, 2009.
- [4] Benchmarking in European Service of public Transport (BEST), Results of the 2004 survey, <http://benchmarkingpublictransport.org/content/download/292/1328/file/Report%20BEST%20Survey%20-%202004.pdf>, 2004. Accessed 08 August 2020.

- [5] Eboli, L., Mazzulla, G., Service quality attributes affecting customer satisfaction for bus transit. *Journal of public transportation.*, 10 (3), 21-34, 2007.
- [6] Noor, H.M., Nasrudin, N., Foo, J., Determinants of Customer Satisfaction of SQ: City Bus Service in Kota Kinabalu, Malaysia. *Procedia - Social and Behavioural Sciences.*, 153, 595-605, 2014.
- [7] Islam, R., Chowdhury, M.S., Sarker, M.S., Ahmed, S., Measuring Customer's Satisfaction on Bus Transportation. *American Journal of Economics and Business Administration.*, 6(1), 34-41, 2014.
- [8] Directorate General Mobility and Transport and Co-ordinated by the Directorate General for Communication. Europeans' Satisfaction with Urban Transport, England, Transport for NSW., https://ec.europa.eu/commfrontoffice/publicopinion/flash/fl_382b_en.pdf, 2014, Accessed 08 August 2020.
- [9] Verbich, D., Geneidy, A., The pursuit of satisfaction: Variation in satisfaction with bus transit service among riders with encumbrances and riders with disabilities using a large-scale survey from London, UK. *Transport Policy.*, 47(1), 64-71, 2015.
- [10] Ardiç, K., Sadaklıoğlu, H., Şehirlerarası yolcu taşımacılığında hizmet kalitesinin ölçümü: Tokat örneği. *Atatürk Üniversitesi İktisadi ve İdari Bilimler Dergisi.*, 23(3), 167-190, 2009 (In Turkish).
- [11] İmre, Ş., Çelebi, D., Measuring comfort in public transport: a case study for İstanbul. *Transportation Research Procedia.*, 25, 2441-2449, 2017.
- [12] Özuysal, M., Tanyel, S., Oral, M.Y., Fayda esaslı erişilebilirliğin ulaşım türü seçimi üzerindeki etkisi. *İMO Teknik Dergi.*, 23(113), 5987-6016, 2012 (In Turkish).
- [13] Doğan, G., Özuysal, M., Toplu ulaşımında bekleme süresini etkileyen faktörlerin incelenmesi: Güvenilirlik, yolcu bilgilendirme sistemi ve fiziksel koşullar. *İMO Teknik Dergi.*, 28(3), 7927-7975, 2017 (In Turkish).
- [14] Kahraman, Ç., Yıldız, M.S., Şehirlerarası otobüs işletmelerinde hizmet kalitesinin ölçülmesi ve bir uygulama. *Hacettepe Üniversitesi İktisadi ve İdari Bilimler Fakültesi Dergisi.*, 23(2), 121-144, 2005 (In Turkish).
- [15] Koçoğlu, C.M., Aksoy, R., Hizmet kalitesinin servperf yöntemi ile ölçülmesi: otobüs işletmeleri üzerinde bir uygulama. *Akademik Bakış Dergisi.*, 29(1), 1-25, 2012 (In Turkish).
- [16] Gökaşar, I., Dündar, S., Buran, B., Yolcu ihtiyaçlarının incelenmesi. *İETT Örneği. İBB Transist 2017 Bildiri Kitabı.*, 2-4 Kasım, İstanbul, 386-395, 2017 (In Turkish).
- [17] Gökaşar, I., Buran, B., Dündar, S., Kent içi otobüs memnuniyet anketi verileri ve faktör analizinden yararlanılarak otobüslerin hizmet kalitesinin modellenmesi: İETT örneği. *Pamukkale Üniv Müh Bilim Derg.*, 24(6), 1079-1086, 2018 (In Turkish).
- [18] Girginer, N., Cankuş, B., Tramvay yolcu memnuniyetinin lojistik regresyon analiziyle ölçülmesi: Estram örneği. *Yönetim ve Ekonomi: Celal Bayar Üniversitesi İktisadi ve İdari Bilimler Fakültesi Dergisi.*, 15(1), 181-193, 2008 (In Turkish).

- [19] Hemedoğlu, E., Toplu taşımacılık sektöründe hizmet kalitesini ölçme: Algılanan hizmet kalitesi ve müşterinin arzuladığı hizmet kalitesi üzerindeki etkileri (Master Thesis). Yıldız Technical University, Graduate School of Natural and Applied Sciences, Istanbul, Turkey, 2010 (In Turkish).
- [20] Seçilmiş, C., Kaşlı, M., Kılıçlar, A., Sarı, Y., The effect of quality at railway services on customer satisfaction in terms of fee paid. *Ege Academic Review.*, 11(4), 573-586, 2011 (In Turkish).
- [21] Akyıldız Alçura, G., Kuşakçı, S.Ş., Gölbaşı Şimşek, G., Gürsoy, M., Tanrıverdi, S.C., Impact score technique for analyzing the service quality of a high-speed rail system. *Transportation Research Record: Journal of the Transportation Research Board.*, 2541(1): 64-72, 2016.
- [22] Šojat, D., Brcic, D., Slavulj, M., Analysis of public transport service improvements on tram network in the City of Zagreb. *Tehnički vjesnik.*, 24, 217-223, 2017.
- [23] Vujičić, M., Prester, J., Assessing service quality of public tram transport in Zagreb city using P-TRANSQUAL mode. *Zbornik Ekonomskog fakulteta u Zagrebu.*, 17, 19-31, 2019.
- [24] Khelf, M., Boukebbab, S., Bhouri, N., Boulahlib, M.S. Tram Service Quality and Its Impact on the Passengers' Modal Choice in Constantine City (Algeria), Selected Papers from the 18th International Conference on Reliability and Statistics in Transportation and Communication, RelStat'18, 17-20 October 2018, Riga, Latvia, 2018.
- [25] De Oña, R., De Oña, J., Neural Networks for Analyzing Service Quality in Public Transportation. *Expert Systems with Applications.*, 41(15), 6830–6838, 2014.
- [26] Islam, M.R., Hadiuzzaman, M., Banik, R., Hasnat, M.M., Musabbir, S.R, Hossain, S., Bus Service Quality Prediction and Attribute Ranking: a Neural Network Approach. *Public Transport.*, 8(1), 295–313, 2016.
- [27] International Bus Benchmarking Group (IBBG), <https://busbenchmarking.org>. Accessed 08 August 2020.
- [28] Acar, S., Bilen Kazancık, L., Meydan, M.C., Işık, M., İllerin ve bölgelerin sosyo-ekonomik gelişmişlik sıralaması araştırması SEGE-2017., *Kalkınma Ajansları Genel Müdürlüğü Yayını Sayı: 3* 2018 (In Turkish).
- [29] KBB Ulaşım Dairesi Başkanlığı., Akçaray güzergay haritası, 2017 <https://twitter.com/kocaeliulasim/status/876024582746066944/photo/1> Accessed 08 August 2020.
- [30] Turkish Statistical Institute TSI, Outcomes of Address Based Population Registration System 2018. TUIK, Ankara, Turkey, 2019 <https://biruni.tuik.gov.tr/medas/?kn=95&locale=tr>. Accessed 08 August 2020.
- [31] Meyer, J.T., *Fundamental Research Statistics for the Behavioural Sciences.*, New York: Holt Rinehart & Winston, 1979.
- [32] Fox, N., Hunn, A., Mathers, N., *Sampling and Sample Size Calculation. The NIHR RDS for the East Midlands / Yorkshire & the Humber*, 2007.

- [33] Dündar, S., Şahin, İ., Train re-scheduling with genetic algorithms and artificial neural networks for single-track railways. *Transportation Research Part C: Emerging Technologies* Volume 27(1), 1-15, 2013.
- [34] Yardım, M.S., Değer Şitilbay, B., Dündar, S., Modelling the effects of hydrated lime additives on asphalt mixtures by fuzzy logic and ANN. *Teknik Dergi*, Vol. 30(6), 9533-9559, 2019.
- [35] Murat, Y.Ş., Başkan, Ö., Modelling vehicle delays at signalized junctions: Artificial neural networks approach. *Journal of Scientific & Industrial Research*, Vol. 65(1), 558-564,2006.
- [36] Osuna, R.G., CS790: Selected Topics in Computer Science, Lecture Notes, Texas A&M University, 2002.
- [37] Murat, Y.Ş., Comparison of fuzzy logic and artificial neural networks approaches in vehicle delay modeling. *Transportation Research Part C* Vol. 14(1), 316–334, 2006.
- [38] Kaya Uyanık, G., Güler, N., A study on multiple linear regression. *Procedia - Social and Behavioral Sciences*, Vol. 106, 234-240, 2013.