ARTICLE IN PRESS

Computers & Industrial Engineering xxx (2014) xxx–xxx

Contents lists available at ScienceDirect

Computers & Industrial Engineering

journal homepage: www.elsevier.com/locate/caie

Customer satisfaction and loyalty analysis with classification algorithms and Structural Equation Modeling

Adnan Aktepe [a,*], Süleyman Ersöz [a,1], Bilal Toklu [b,2]

[a] Kirikkale University, Faculty of Engineering, Department of Industrial Engineering, 71450 Yahsihan, Kirikkale, Turkey
[b] Gazi University, Faculty of Engineering and Architecture, Department of Industrial Engineering, Maltepe, Ankara, Turkey

ARTICLE INFO

Article history:
Available online xxxx

Keywords:
Classification algorithms
Customer satisfaction
Customer loyalty
Structural Equation Modeling
WEKA
LISREL

ABSTRACT

Businesses can maintain their effectiveness as long as they have satisfied and loyal customers. Customer relationship management provides significant advantages for companies especially in gaining competitiveness. In order to reach these objectives primarily companies need to identify and analyze their customers. In this respect, effective communication and commitment to customers and changing market conditions is of great importance to increase the level of satisfaction and loyalty. To evaluate this situation, level of customer satisfaction and loyalty should be measured correctly with a comprehensive approach. In this study, customers are investigated in 4 main groups according to their level of satisfaction and loyalty with a criteria and group based analysis with a new method. We use classification algorithms in WEKA programming software and Structural Equation Modeling (SEM) with LISREL tools together to analyze the effect of each satisfaction and loyalty criteria in a satisfaction–loyalty matrix and extend the customer satisfaction and loyalty post-analysis research bridging the gap in this field of research. To convert developed conceptual thought to experimental study, white goods industry is exemplified. 15 criteria are used for evaluation in 4 customer groups and a satisfaction–loyalty survey developed by experts is applied to 200 customers with face-to-face interviews. As a result of the study, a customer and criteria grouping method is created with high performance classification methods and good fit structural models. In addition, results are evaluated for developing a customer strategy improvement tool considering method outcomes.

1. Introduction

In order to achieve sustainable competition advantage in the market, it is necessary to provide and improve customer satisfaction (CS). CS Analysis is used for measuring customer satisfaction levels, taking counter actions for the low satisfaction points and improving high satisfaction points. When the customer becomes the focus of organization and if it gains more satisfied customers, then high satisfaction contributes in both internal and external processes of a company (Ersoz, Yaman, & Birgoren, 2008; Gale & Wood, 1994). High satisfaction brings many advantages, for example, customer oriented organizations can achieve high financial performance (Johnson, 2000). Therefore, CS analysis is conducted by many of the firms for gaining several competitive advantages in the market (Kengpol & Wangananon, 2006). In addition, to retain customer, organization structuring is to be established in accordance with customer satisfaction (Kotler & Armstrong, 1994).

In the literature, there are several approaches for CS analysis with various satisfaction criteria. Successful and nation-wide applications in this field consider CS analysis as a cause-and-effect model. In CS analyses, different types of customer evaluations cannot be measured directly, so they are modeled as latent variables (variables that affect CS or affected by CS but cannot be measured directly). Therefore, CS analysis becomes meaningful and powerful when analyzed with antecedents and consequences (Ciavolino & Dahlgaard 2007; Fornell, 1992; Fornell, Michael, Eugene, Jaesung, & Barbara, 1996; Grigoroudis & Siskos 2002; Liu, Zeng, Xu, & Koehl, 2004; Martensen, Kristensen, & Grönholdt, 2000; Turkyılmaz and Ozkan (2007); Shao-I, Ching-Chan, Tieh-Min, & Hsiu-Yuan, 2011). Grigoroudis & Siskos 2004 also give a list of studies based on cause and effect models of satisfaction. One of the most addressed consequent of CS in the literature is customer loyalty (CL). CL is can be expressed as the likelihood to recommend company to other customers, the likelihood to

* Corresponding author. Tel.: +90 318 357 42 42/1047; fax: +90 318 357 24 59.
E-mail addresses: aaktepe@gmail.com (A. Aktepe), sersoz40@hotmail.com (S. Ersöz), btoklu@gazi.edu.tr (B. Toklu).
[1] Tel.: +90 318 357 42 42/1008; fax: +90 318 357 24 59.
[2] Tel.: +90 (312) 582 38 51.

http://dx.doi.org/10.1016/j.cie.2014.09.031
0360-8352/© 2014 Elsevier Ltd. All rights reserved.

repurchase or revisit of customers (Anderson & Mittal, 2000). As many researchers indicate, there is a significant relationship between customer satisfaction and customer loyalty. Kumar, Pozza, and Ganesh (2013) state that the association between customer satisfaction and loyalty is highly variable depending on some factors as the industry, customer segment studied, the nature of the dependent and independent variables, and the presence of numerous factors that serve as mediators. The authors also give a list of studies conducted on satisfaction–loyalty relationship. Cronin and Taylor (1992), Garbarino and Johnson (1999), Ngobo (1999), Cronin, Brady, and Hult (2000), Churchill and Halpern (2001), Lam, Venkatesh, Krishna, and Bvsan (2004), Homburg and Furst (2005), Anderson and Mittal (2000), Vesel and Zabkar (2009), Deng, Lu, Wei, and Zhang (2010), Chen (2012), and Orel and Kara (2013) discover strong linear relationships between customer satisfaction and customer loyalty in various sectors and industries. However, customer satisfaction (CS) does not completely determine customer loyalty (CL) (Chen, 2012; Deng et al., 2010; Gerpott, Rams, & Schindler, 2001; Johnson, 2000; Kumar et al., 2013; Lam et al., 2004; Orel & Kara, 2013). The effect from CS to CL is not always fully determined. This means, in CS analysis there are some group of customers who are lowly satisfied–highly loyal and highly satisfied–lowly loyal. So, we can categorize customers into mainly 4 different groups: Group 1: Low Satisfaction–Low Loyalty, Group 2: Low Satisfaction–High Loyalty, Group 3: High Satisfaction–Low Loyalty and Group 4: High Satisfaction–High Loyalty. These 4 groups construct 4 different section of the CS–CL matrix which is presented in Section 2 in detail.

The major deficiency of post analysis methods on CS and CL is the lack of quantitative calculation method of evaluation in a systematic way. The model proposed in this study has the advantage of bridging the gap in this area by presenting an integrated approach using data mining and structural models together.

In this study we extend the CS and CL analysis by integrating relationship results with CS and CL criteria in a CS–CL matrix. This matrix is used for both creating customer segments and positioning CS and CL criteria to the related part of the matrix. This improvement helps us to discover a more comprehensive CS–CL relationship and develop strategies for increasing the total share of 4th Group customers. Thus, we can develop prudential strategies for increasing total share of 4th Group.

Another innovation point in this study is that we analyze CS and CL together with an algorithm using data mining classification algorithms and Structural Equation Modeling (SEM) together. Here we use decision trees produced after classification applications to create customer groups and to find breaking points of the CS–CL matrix. Ngai, Li, and Chau (2009) conduct a detailed literature study on the use of data mining algorithms in customer relationship management. They state that classification algorithms are used for customer segmentation and customer development. In this study we contribute to the literature in this field by determining breaking points of CS–CL matrix with classification algorithms. For classification applications WEKA data mining tool is used developed by Hall et al. (2009). By using classification tool we uncover meaningful and hidden patterns by using data mining techniques in customer data. Results of the study have potential inputs for many customer-focused applications.

A further extension in CS–CL analysis in this study is customer strategy development according to matrix-based model results with SEM. SEM is used to discover CS and CL criteria groups and their relations in the developed structured models. In strategy development process, main objective is not only increasing CS level but also increasing the number of loyal customers and maintaining customer retention in the long term with satisfied and loyal customers. The model helps us to discover these hints for strategy development. As Chikara and Takahashi (1997), Grigoroudis, Samaras, Matsatsinis, and Siskos (1999), and Grigoroudis and Siskos (2002) state that the most important part of CS analysis is building a post-analysis method to create future directions for companies. In this study we evaluate matrix results with classification and SEM results and offer criteria-based customer group strategies. To build structural equation models SIMPLIS language of LISREL 8.80 (by Jöreskog & Sörbom, 2006) software is used.

The paper is organized as follows: In Section 2, we present the scope and purpose of study. We define the capabilities of developed CS–CL matrix here. In Section 3, we present CS–CL analysis algorithm step by step. The application data are collected by a survey in white-goods industry. In Section 4, we discuss data collection procedure and application results of the model in white-goods industry. And in the final section of the study we discuss results, findings, advantages and future directions of this study.

2. Scope and purpose of the study

The considered problem in this is study is developing a new post-analysis method for customer satisfaction (CS) and customer loyalty (CL) analysis. The importance of post analysis methods in customer satisfaction evaluation is emphasized by Hill (1996), Chikara and Takahashi (1997), Grigoroudis et al. (1999), and Grigoroudis and Siskos (2002). In these papers authors emphasize that a reasonable post evaluation of CS results is very important for future strategies at least the CS analysis itself. The model developed in this study integrates data mining tools with Structured Equation Modeling (SEM) technique and produces beneficial results for creating customer strategies as a CS and CL post-analysis guide. This bridges a significant gap in this area of research. Application of the model is conducted in white-goods sector in Turkey.

In this study we propose a new matrix-based approach for CS and CL analysis. The model developed in the study is a kind of customer satisfaction evaluation that uses data mining (discovering unknown patterns) advantages of classification algorithms and cause-and-effect modeling advantage of Structural Equation Modeling (SEM). This model is not only a pure evaluation of CS and CL but also an interactive matrix-based procedure that investigates CS and CL with post analyses.

For CS–CL analysis, in some of the studies, authors develop a matrix-based approach. Gerpott et al. (2001) develops customer satisfaction–customer loyalty and customer loyalty–customer retention matrices. They discuss relationship among customer retention–customer loyalty and customer loyalty–customer satisfaction in telecommunications market. They define some properties of customer groups in each matrix and build Structural Equation Model (SEM) which is created independently from matrix. Ersoz et al. (2008) use artificial network networks for classification of customers according to a CS matrix developed by Aktas et al. (2000). In these studies authors develop some customer groups and finds out distribution of customer groups in the matrix. Then they evaluate type and distribution of customers. However, in our study the main objective is not only segmentation of customers but also assigning each CS and CL criteria to the related part of the developed CS–CL matrix and creating a criteria-based matrix. Thus, developed CS–CL matrix shows customer groups and most effective criteria on CS and CL together. Here we use the results of best performing classification algorithms. Additionally we offer a customer strategy development tool by integrating the results of classification decision tress with SEM analysis.

CS–CL matrix developed in the study is given in Fig. 1 below. The diagonal blue line shows the target of strategies which aims to create customer retention at the end of high CS and high CL

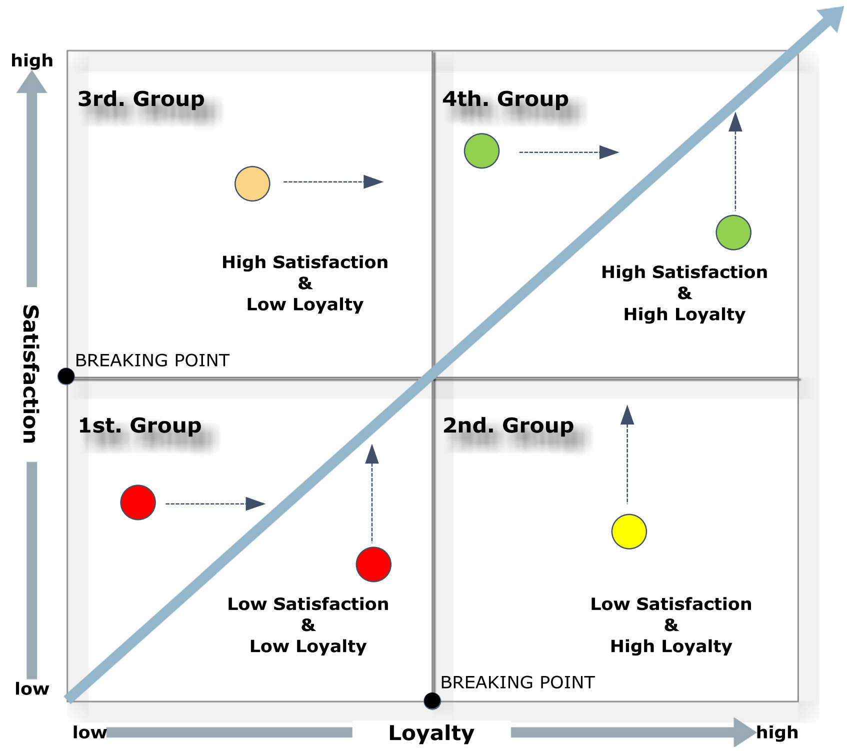


Fig. 1. Customer group and criteria based CS–CL matrix.

by moving all customers to 4th group and keeping there. In this matrix there are 4 different types of customer groups and criteria. 1st group shows low satisfaction–low loyalty group of customers with red-colored criteria. These red-colored criteria are the most important ones for first group and they are used for determining priorities of 1st group to move them to 4th group 2nd group shows low satisfaction–high loyalty group of customers with yellow-colored criteria. These yellow-colored criteria are the most important ones for second group and they are used for determining priorities of 2nd group to move them to 4th group. 3rd group shows high satisfaction–low loyalty group of customers with yellow-colored criteria. These orange-colored criteria are the most important ones for third group and they are used for determining priorities of 3rd group to move them to 4th group. And 4th group shows high satisfaction–high loyalty group of customers with green-colored criteria. These green-colored criteria are the most important ones for forth group and they are used for determining priorities of 4th group to keep them at least the current situation.

The lowest value for both satisfaction and loyalty is "1" which is the lowest value of 5-point Likert scale. The highest value for satisfaction and loyalty is "5" which is the highest value in 5-point Likert scale. With CS–CL matrix in Fig. 1 we find out answers for the questions of how to distinguish satisfied and dissatisfied, loyal and uncommitted customers by using decision trees created as a result of classification algorithms. Then, CS and CL criteria are positioned in the CS–CL matrix according to CS and CL scores. Finally, CS–CL relationship is created with SEM analysis to create future strategies for each group. CS–CL matrix method is depicted in the next section of the paper step by step.

3. CS–CL analysis method: developing CS–CL matrix method with classification algorithms and Structural Equation Modeling (SEM)

In this section we discuss the methodology that is developed for matrix-based CS–CL analysis step by step:

3.1. Step 1. Confirmatory factor analysis

Confirmatory Factor Analysis (CFA) is a statistical technique used to verify the factor structure of a set of observed variables. CFA allows the researcher to test the hypothesis that a relationship between observed variables and their underlying latent constructs exists (Schumacker & Lomax, 1996). After customer data are collected from customers, variables that measure CS or CL is determined with CFA. Here we define two different latent variables for satisfaction criteria (namely Sat.a and Sat.b) and similarly we define two different latent variables for loyalty criteria (namely Loy.a and Loy.b). These 4 latent variables (4 different group of measurement variables) is distributed to 4 sections of the matrix (Sat.a and Sat.b belong to CS and Loy.a and Loy.b belong to CL). CFA helps us to find which measurement variable represents a latent variable more effectively. CFA is used to estimate the model parameters and examine the factor structure. CFA model is built by using the maximum likelihood estimation method developed by Chou and Bentler (1996) which is the most commonly used approach in Structural Equation Modeling (SEM). Performance of CFA model is checked with overall model fit indices in LISREL software.

3.2. Step 2. Running classification algorithms and selecting best performing algorithm

Classification is one of the problem solving techniques used in data mining concept. Classification finds a model for class attribute as a function of the values of other attributes. Given a collection of records or training set, the goal is to assign a class as accurately as possible for previously unseen records or test set. There are several algorithms that are used for different classification purposes which are: Decision Tree based Methods, Rule-based Methods, Memory based reasoning, Neural Networks, Naive Bayes, Bayesian Belief Networks and Support Vector Machines (Tan, Steinbach, & Kumar, 2005; Witten, Frank, & Hall, 2011). Classification variables used in classification algorithms are CS and CL criteria of which groups are defined in Step 1. In this study, we have two types of classes: CS classes and CL classes. We find final class attribute of a record by using an intersection approach (presented in Step 4) because we have we have 2 main classes (CS and CL) and 4 sub-classes (Low CS–Low CL, Low CS–High CL, High CS–Low CL and High CS–High CL) according to matrix structure. Here we use several classification algorithms that create decision trees with same customer data and select the best performing algorithm according to performance results. In this step, best performing classification algorithm is selected among algorithms that are run in Step 2 according to highest correctly classified instance ratio and lowest mean absolute error (MAE) values.

3.3. Step 3. Finding breaking points (b_{CS} and b_{CL}) in CS–CL matrix

Breaking point can be defined as the distinction point in CS–CL matrix. Breaking points have a numeric value between the lowest point and highest point in Likert scale used in data collection. We separate CS groups according to CS breaking point and CL groups according to CL breaking point. b_{CS} is breaking point for *Y*-axis and b_{CL} is breaking point for *X*-axis in CS–CL matrix. This characteristic of the matrix is very important to find out distribution of customer groups.

In this step, for finding out breaking points, classification algorithms that create a decision tree are compared. Decision tree of best selected classification algorithm is used here to find breaking points in the CS–CL matrix. Breaking points are found according to branching result of decision trees. The highest branching value in the tree is determined as the breaking point. Selecting the highest branching score in the decision tree brings two important advantages: i. A higher verge is set for being included in the best group (4th Group) and ii. A dynamic incentive value has been created for companies for achieving a level of CS and CL.

3.4. Step 4. Creating customer groups and constructing customer based CS–CL matrix

After breaking points found in Step 3, we define customer groups with an intersection of main groups. According to CS criteria we have 2 different groups. First group is low satisfaction (G12: below CS breaking point) group and second one is high satisfaction group (G34: above CS breaking point). Similarly 2 different groups are created according to CL criteria. First group is low loyalty (G13: left side of CL breaking point) and second one high loyalty (G24: right side of CL breaking point). Here we use two different type of classification algorithm: i. First one is for determining CS groups (G12-Under CS breaking point and G34-Above CS breaking point), ii. Second set of classification algorithms are used for determining CL groups (G13-Left side of CL breaking point and G24-Right side of CL breaking point). Finally, the group of a customer in CS–CL matrix (shown in Fig. 1) is found according to an intersection approach which is given in Table 1.

3.5. Step 5. Positioning CS and CL criteria in CS–CL matrix with average satisfaction and loyalty score

After defining customer groups in Step 4, in this step we position CS and CL criteria in the matrix. The positioning of each criterion in the matrix is important in terms of customer group and criteria integration. Firstly, satisfaction and loyalty score is calculated by taking average of satisfaction and loyalty responses respectively in CS–CL questionnaire. Then, the position is determined according to breaking points discovered in Step 3. As an example, if a satisfaction score is higher than b_{CS}, then this criterion is evaluated in the 3rd or 4th group. Here we take the advantage of tree-based classification algorithms and we use classification not only for customer segmentation but also for evaluating decision trees. However, final group of the variables is not yet determined in this step. Final group of each criteria group is determined in the last step (Step 6) of the method after analyzing relational structure among criteria.

3.6. Step 6. Creating final CS–CL matrix with Structural Equation Modeling (SEM) results

In the last step, firstly the CS–CL matrix is completed by determining the final group of each criteria group with Structural Equation Modeling (SEM) method. SEM is used to find structural relations among latent variables. A latent variable represents a cluster of observed variables (Bollen, 1989; Kline, 2005). In this study SEM is used for defining relations among CS and CL criteria and according to SEM results we produce concrete strategies for each group of customers. We bridge the results of customer classification (found out in Step 4) with criteria matrix (found out in Step 5). Then, the effect of each CS and CL latent variable is analyzed with SEM strategy. Finally structural model is adopted for generating criteria-based strategy development for customer groups which are discussed in the last section of the paper.

4. Application of CS–CL matrix method

Application of the model is conducted in white-goods industry. Necessary data for testing the developed model collected with survey applications. Face-to-face interviews are carried out with white-goods customers. The survey questionnaire is applied especially to 31+ age women who mostly utilize white goods in daily housework. Sample profile and data collection method is presented in the first sub-section of this section. The results of the application produce powerful insights for CS and CL analysis. This shows the applicability of the model with high performance classification and SEM results. These are discussed in the second sub-section of this section.

4.1. Data collection

Application data are collected via face-to-face interviews conducted by 20 professionals to 205 customers who use white-goods in daily life. Regardless of brand type, refrigerator, washing machine and dishwashing machine in white appliances product

Table 1
Intersection and definition of customer groups.

Classification result	Intersection	Group characteristic	Final group of a customer
G12 in CS and G13 in CL	G**12** ∩ G**13**	Low CS and Low CL	1st group
G12 in CS and G24 in CL	G**12** ∩ G**24**	Low CS and High CL	2nd group
G34 in CS and G13 in CL	G**34** ∩ G**13**	High CS and Low CL	3rd group
G34 in CS and G24 in CL	G**34** ∩ G**24**	High CS and High CL	4th group

Please cite this article in press as: Aktepe, A., et al. Customer satisfaction and loyalty analysis with classification algorithms and Structural Equation Modeling. *Computers & Industrial Engineering* (2014), http://dx.doi.org/10.1016/j.cie.2014.09.031

Table 2
Sample profile.

Variable		Count	%
Gender	Male	70	35
	Female	130	65
	Total	200	100
Age	20–25	75	37.5
	26–30	20	10.0
	31+	105	52.5
	Total	200	100
Education level	Primary school	24	12
	High school	76	38
	Bachelor's degree	91	45.5
	Master's degree	9	4.5
	Total	200	100

Table 3
Variables used in the survey and conceptual classification structure before factor analysis.

Question	Variable name	CS criterion	CL criterion
Q.1	Service network	√	
Q.2	Energy consumption	√	
Q.3	Functional properties	√	
Q.4	Quality level compared to costs	√	
Q.5	Price	√	
Q.6	Listening to other customers' voice	√	
Q.7	Campaigns	√	
Q.8	Distance to store	√	
Q.9	Brand's general qualifications		√
Q.10	Physical appearance		√
Q.11	TV and internet ads of Brand		√
Q.12	New technologies of Brand		√
Q.13	Brand name		√
Q.14	Trust on information provided by manufacturing company		√
Q.15	General trust level to quality		√

group are evaluated by customers who use one of 5 big brands in Turkey. Of 205 customer data, 5 customers' data are not taken into analysis because of blank and inconsistent answers. The application is carried out with 200 customer data.

4.1.1. Sample profile

Table 2 provides information about sample characteristics. Of the 200 total number of respondents, 70 (35%) were male and 135 (65%) were female customers. This gender and age composition is a reasonable representation of the white-goods users in Turkey. The majority of the respondents were above age of above 31 (52.5%). Because, generally women and married people utilize white appliances at home more than other group of users. In addition, the majority (88%) of the respondents had a high school degree or higher, which we believe is another important characteristic of the customer group who can make reasonable evaluations of satisfaction and loyalty questions in the survey. The reliability of the data is checked by conducting reliability analysis in SPSS statistical package (SPSS Inc, 2007). Most reliability scores were within the suggested levels (>.70) in the literature.

4.1.2. Questionnaire design

Table 3 provides information about survey questions. The survey questions are reviewed and prepared with 5 marketing experts in good white-goods sector and 3 academicians whose area of specialization is on service systems and industrial engineering. Application survey questionnaire is attached to this paper in Appendix.

In the application, data of the problem are based on the customers' judgments as in most of the customer satisfaction research. The customers' replies constitute input data used in the CS–CL matrix method developed in this study. This is a multivariate evaluation problem that CS and CL are evaluated with several variables (which are also called CS and CL criteria). Outputs of the model are creating customer groups, creating criteria groups and relating them to customer groups and discovering strategies to improve

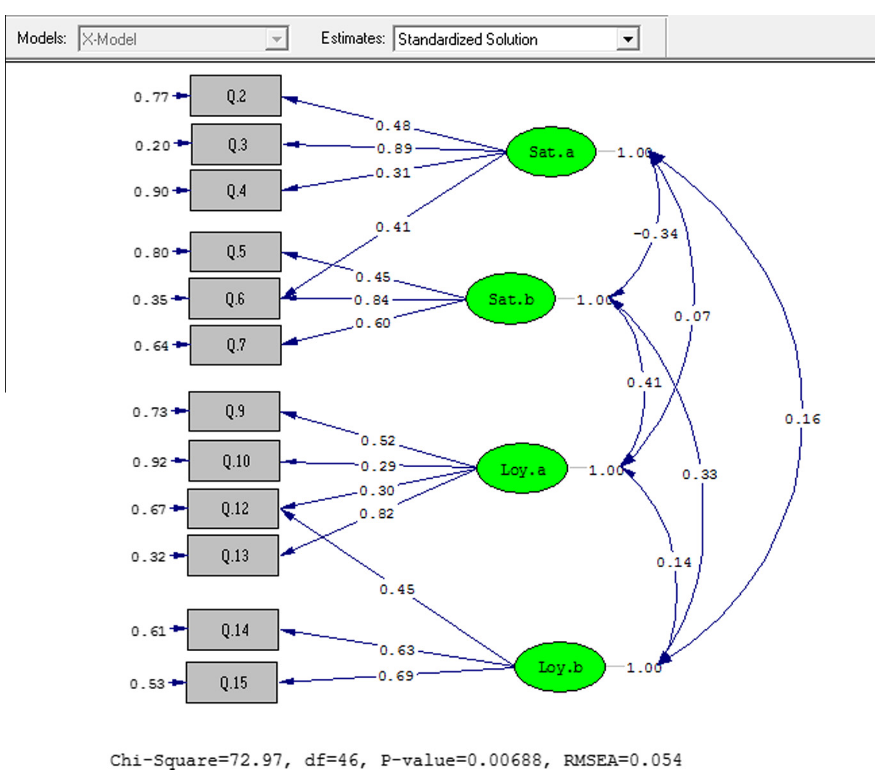


Fig. 2. CFA model of application in LISREL.

Table 4
Goodness of fit statistics about CFA results.

Goodness of fit statistics	Value	Statistical Fit
Normal theory weighted least squares Chi-square	72.97 (df = 46, $p = 0.00688$)	Good fit
Root mean square error of approximation (RMSEA)	0.054	Good fit
Normed fit index (NFI)	0.85	Acceptable
Non-normed fit index (NNFI)	0.91	Good fit
Comparative fit index (CFI)	0.94	Good fit
Incremental fit index (IFI)	0.94	Good fit
Goodness of fit index (GFI)	0.94	Good fit
Adjusted goodness of fit index (AGFI)	0.90	Good fit
Root mean square residual (RMR)	0.055	Good fit
Model CAIC	274.51	Acceptable

Table 5
Performance of classification algorithms for satisfaction analysis.

Classification algorithm type	Correctly classified instances (%)	Incorrectly classified instances (%)	Mean absolute error (MAE)	Test option
ADTree	94	6	0.1958	Using training set
BFTree	90.5	9.5	0.1480	Using training set
Decision Stump	73.5	26.5	0.3471	Using training set
FT	93.125	6.875	0.1310	Percentage split %20
J48 and J48-Graft	**95**	**5**	**0.0881**	Using training set
LAD	94.5	5.5	0.1388	Using training set
LMT	89.375	11.625	0.1476	Percentage split %20
NBTee	90	10	0.2000	Using training set
RandomForest	84.375	15.625	0.2550	Percentage split %20
RandomTree	79.375	20.625	0.2063	Percentage split %20
REPTree	87.5	12.5	0.2118	Using training set
SimpleCart	70.625	29.375	0.3275	Percentage split %20

Bold values are the best results in the table.

CS and CL evaluation. In the next sub-section we present the application results of the model step-by-step as defined in Section 3.

4.2. Application results

4.2.1. Application of Step 1. Confirmatory factor analysis results

We conduct CFA by using LISREL 8.80 (Linear Structural Relations) software created by Jöreskog & Sörbom, 1993, 2006). LISREL is a statistical language that interfaces with statistical applications. We use SIMPLIS codes in LISREL application for CFA application.

In white-goods sector application, firstly in accordance with the opinion of experts questions between 1 and 8 were associated with customer satisfaction and questions between 9 and 15 were associated with customer loyalty. Satisfaction criteria are divided into 2 groups: Satisfaction on Functional and technical properties (Sat.a) and Satisfaction on Price and campaigns (Sat.b). Similarly loyalty criteria are divided into 2 groups: Brand loyalty (Loy.a) and Loyalty and Trust on Quality (Loy.b).

CFA model to test conceptual model created by experts is depicted in Fig. 2. CFA was first used to estimate the model

Table 6
Performance of classification algorithms for loyalty analysis.

Classification algorithm type	Correctly classified instances (%)	Incorrectly classified instances (%)	Mean absolute error	Test option
ADTree	94	6	0.1832	Using training set
BFTree	92	8	0.1247	Using training set
Decision Stump	74.5	25.5	0.3784	Using training set
FT	94.5	5.5	0.1121	Percentage split %20
J48 and J48-Graft	**94.5**	**5.5**	**0.0958**	Using training set
LAD	94	6	0.1339	Using training set
LMT	92.7	7.3	0.1287	Percentage split %25
NBTee	91	9	0.1679	Using training set
RandomForest	85.625	14.375	0.2437	Percentage split %20
RandomTree	83.75	16.25	0.1625	Percentage split %20
REPTree	90	10	0.1588	Using training set
SimpleCart	83.125	16.875	0.1797	Percentage split %20

Bold values are the best results in the table.

parameters and find the latent variables of prediction. The measurement models are estimated using the maximum likelihood estimation method in LISREL which is the most commonly used approach in SEM (Orel & Kara, 2013). In Fig. 2, standardized solution result of final CFA model is shown. After several trials of CFA model we found out that Q.1, Q.8 and Q.11 does not produce statistically significant results according to t-value statistics (LISREL produces estimates, standardized solution, *t*-values and modification indices for each model developed). So Q.1, Q.8 and Q.11 are excluded from model and the rest of application is implemented with 12 variables.

In CFA model: There are 4 factors (Sat.a, Sat.b, Loy.a and Loy.b); 2 factors of both CS and CL. Q.2, Q.3 and Q.4 represent Sat.a latent variable and Q.5, Q.6, Q.7 represent Sat.b latent variable in CS group. Q.9, Q.10, Q.12 and Q.13 represent Loy.a latent variable and Q.14 and Q.15 represent Loy.b latent variable in CS group.

The CFA model created produces acceptable results according to statistical fit indices. The threshold values for good fit statistics is defined in the studies of Bentler (1980), Bentler and Bonett (1980), Byrne (1998), Jöreskog and Sörbom (2006), Simsek (2007) and Cokluk, Sekercioglu, and Buyukozturk (2012). Statistical fit results are shown in Table 4 and all of them are in acceptable ranges.

Modification indices (another output property of Lisrel) created as an output of the model create useful suggestions for a better fit model. Application of these suggestions decreases the value of Chi-square which is one of the most important fit statistics in Structural Equation Modeling (SEM) (Byrne, 1998). Considering the results of modification indices, Q.6 is connected to Sat.a and Q.12 is connected to Loy.b. This helped us to decrease in Chi-Square statistics and Root Mean Square Error of Approximation (RMSEA).

4.2.2. Application of Step 2. Running classification algorithms and selecting best performing algorithm

Before classification application firstly data are divided into 2 parts according to factors created in Step 1 by CFA. Satisfaction data (Data of sample replies on Q.2, Q.3, Q.4, Q.5, Q.6 and Q.7) and Loyalty data (Data of sample replies on Q.9, Q.10, Q.12, Q.13,

Please cite this article in press as: Aktepe, A., et al. Customer satisfaction and loyalty analysis with classification algorithms and Structural Equation Modeling. *Computers & Industrial Engineering* (2014), http://dx.doi.org/10.1016/j.cie.2014.09.031

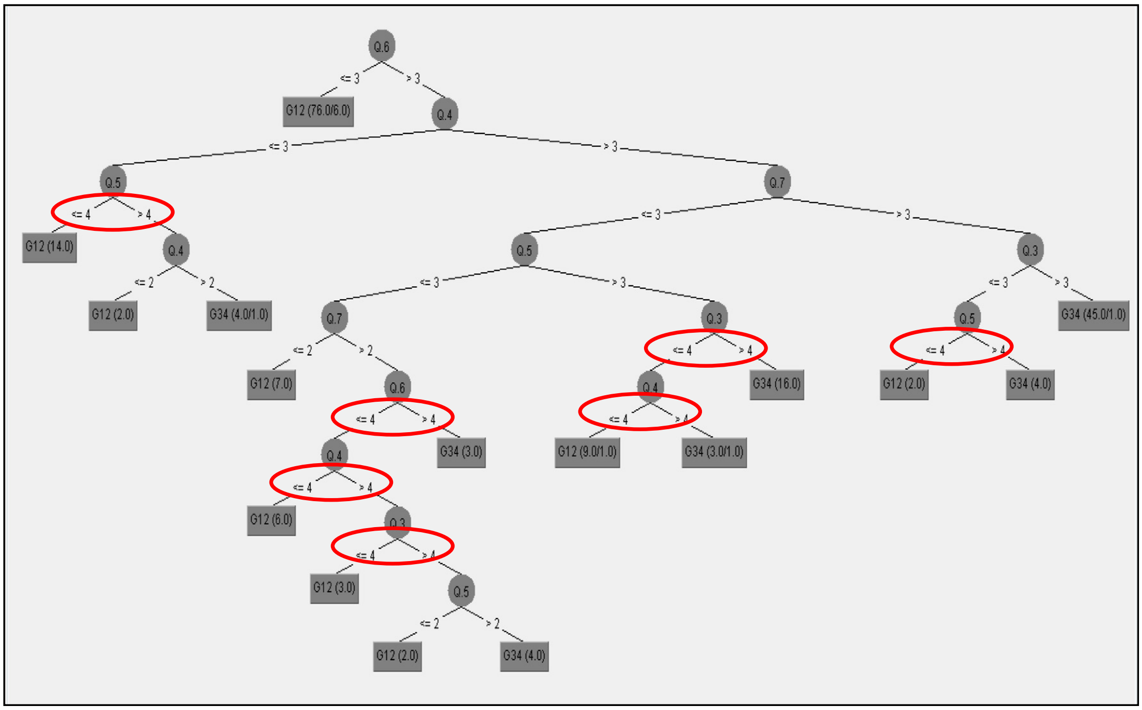


Fig. 3. J48 decision tree for CS criteria.

Q.14 and Q.15). In WEKA programming database (Hall et al., 2009) classification algorithms that create decision tree are used for classification of data.

Comparison of performances of classification algorithms is given Tables 5 and 6 for CS and CL respectively. In both of the cases J48 algorithm produces the best performance results according to correctly classified instances and MAE values. Test option is a property of WEKA programming and we tried several test options for each tree algorithm and in Tables 5 and 6, we show the best results of each trial for each algorithm type.

The initial training set for 20 customers is determined from the study of expert team. Then final training set for developed classification algorithm is extended according to new method presented in this study.

4.2.3. Application of Step 3. Finding Breaking Points (b_{CS} and b_{CL}) in CS–CL matrix

Best performing tree algorithm both for CS and CL variables is J48 that was discovered in Step 2. J48 decision tree results are depicted in Figs. 3 and 4 for satisfaction and loyalty criteria

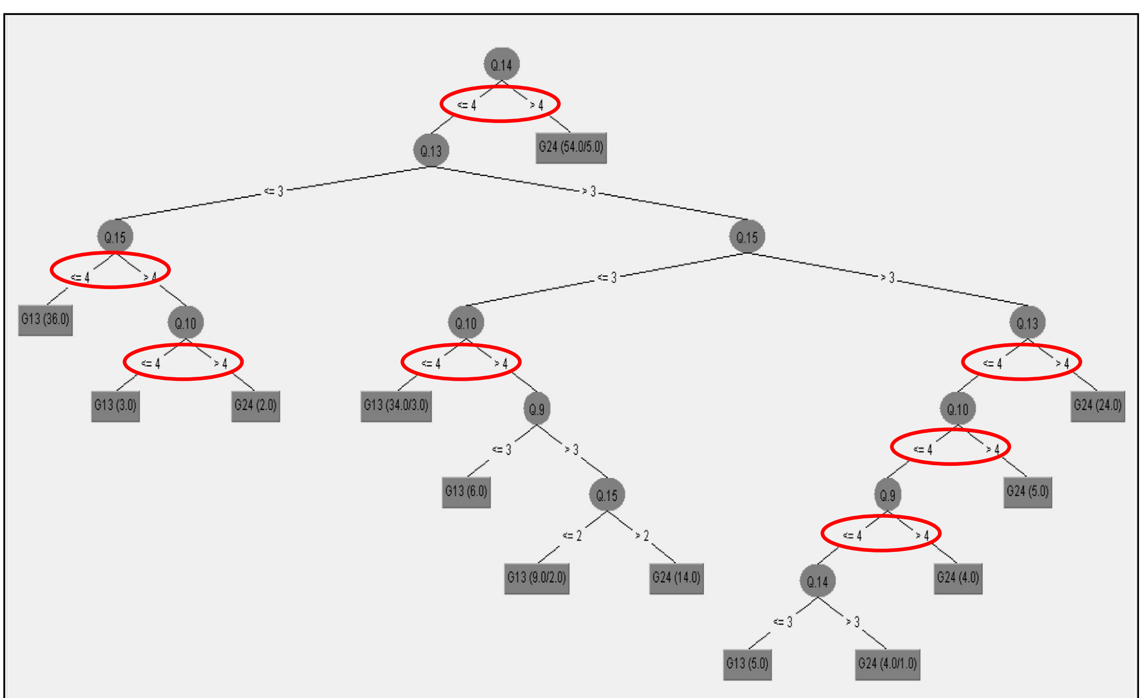
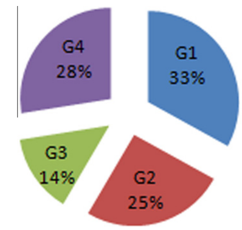


Fig. 4. J48 decision tree for CL criteria.

Please cite this article in press as: Aktepe, A., et al. Customer satisfaction and loyalty analysis with classification algorithms and Structural Equation Modeling. *Computers & Industrial Engineering* (2014), http://dx.doi.org/10.1016/j.cie.2014.09.031

Table 7
Intersection and definition of customer groups.

Intersection	Group	Number of customers in this group	Percentage	Distribution of customer groups
G**12** ∩ G**13**	1st group (G1)	66	33	
G**12** ∩ G**24**	2nd group (G2)	51	25.5	
G**34** ∩ G13	3rd group (G3)	28	14	
G**34** ∩ G**24**	4th group (G4)	55	27.5	

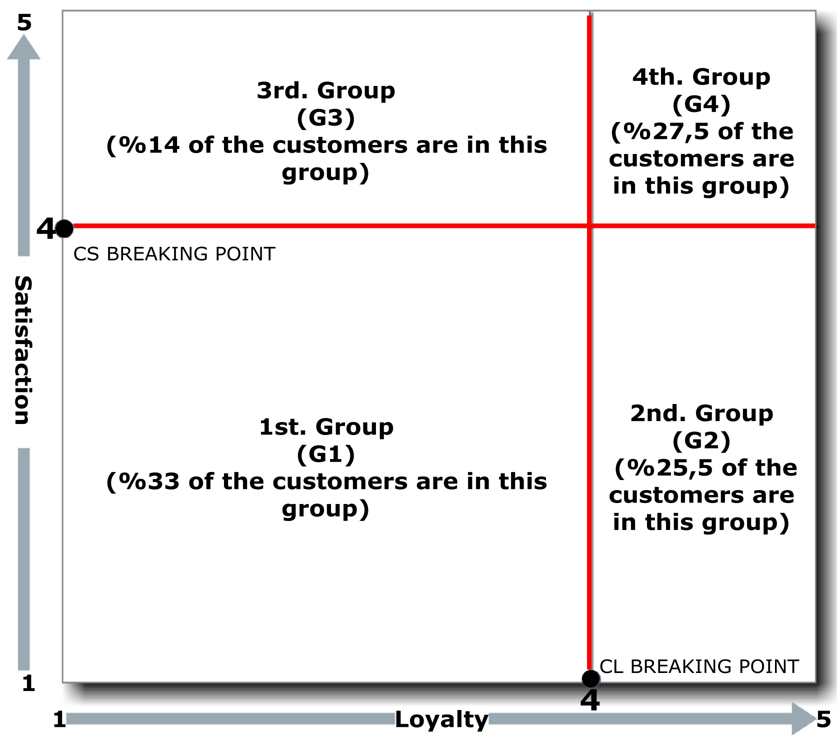


Fig. 5. Customer groups in CS–CL matrix.

respectively. Branching values range from 1 (lowest scale in 5-point Likert scale) to 5 (highest scale in 5-point Likert scale). The branching value on each tree shows that 4 is the highest value. Therefore we set breaking values of CS–CL matrix (b_{CS} *and* b_{CL}) as "4".

4.2.4. Application of Step 4. Creating customer groups and constructing customer based CS–CL matrix

In this step of application, we consider results of J48 algorithm found in Steps 2–3 and define customer groups with an intersection of main CS and CL groups. G12 (under CS breaking point) low satisfaction and G34 (above CS breaking point) is high satisfaction group. G13 (left side of CL breaking point) is low loyalty and G24 (right side of CL breaking point) is high loyalty group. Distribution of customer groups is shown in Table 7 and CS–CL matrix presentation is depicted in Fig. 5.

Customer classification results show that most of the customers are in G1 (worst group in the matrix) and least number of customers appears in G3 (high satisfaction and low loyalty group). 27.5% of the customers appear in the target group (G4) and the rest of the customers are in G2 (low satisfaction and high loyalty group).

A quick interpretation of the distribution scheme provides us some insights about customer profile:

- Target group has a share more than a quartile which shows that a good level of CS and CL is achieved for 28% of customers. But still majority of the customers (33%) appear in the worst group (low CS and low CL) which shows that majority of the customers are unhappy and we need strong and applicable strategies for them.
- Other groups (G2 and G3) constitutes remaining 39% share. Customer strategies that are developed for these groups are at least as important as strategies developed for G4 and G1. The characteristics of each group, customer behaviors and strategies developed by integrating other outcomes of the model are discussed in the last section of the study in detail.

4.2.5. Application of Step 5. Positioning CS and CL criteria in CS–CL matrix with average satisfaction and loyalty score

In this step of application, each CS and CL criteria are positioned in the matrix according to average CS and CL scores respectively. In this application, there are two different main group of consideration. Combining CFA results found in Step 1 and classification results found in Step 2, we determine the region of each latent variable in CS–CL matrix. CS criteria are divided into two sub-classes: First one is "Satisfaction on Functional and Technical Properties

Table 8
CS and CL average satisfaction scores.

CS and CL criteria groups	Latent variable	Average CS score	Average CL score	Position in the matrix
Satisfaction on Functional and technical properties	Sat.a	3.513	–	G13 (G1 or G3) region
Satisfaction on Price and campaigns	Sat.b	4.393	–	G13 (G1 or G3) region
Brand loyalty	Loy.a	–	3.675	G24 (G2 or G4) region
Loyalty and Trust on Quality	Loy.b	–	4.039	G24 (G2 or G4) region

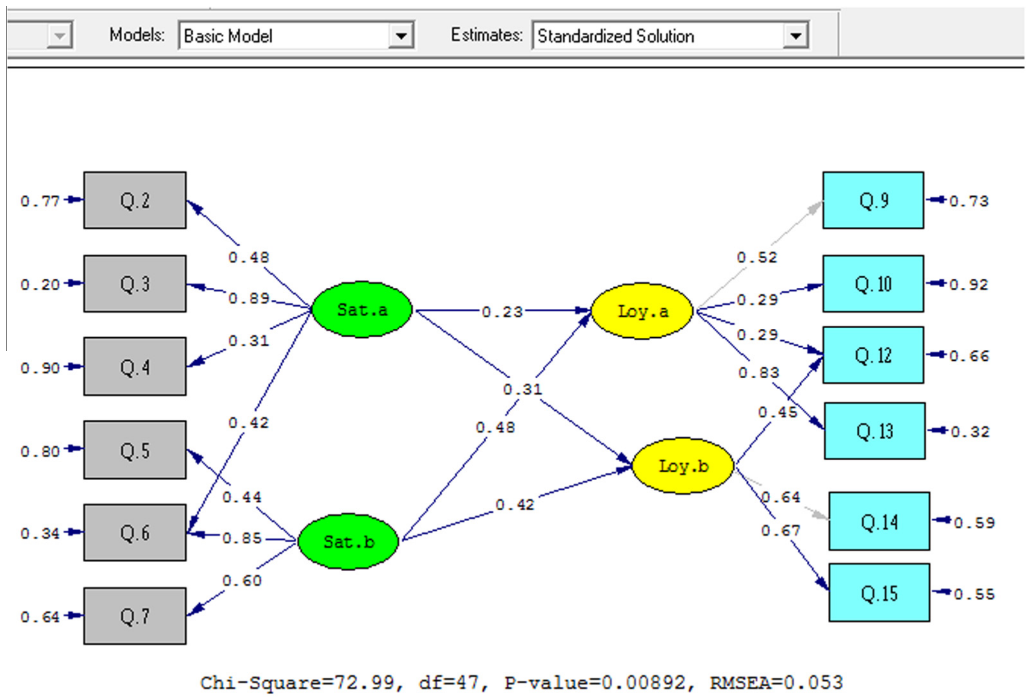


Fig. 6. Standardized solution estimate of structural equation model.

Table 9
Goodness of fit statistics about SEM results.

Goodness of Fit Statistics	Value	Statistical Fit
Normal theory weighted least squares Chi-square	72.99 (p = 0.00892)	Good fit
Root mean square error of approximation (RMSEA)	0.053	Good fit
Normed fit index (NFI)	0.85	Acceptable
Non-normed fit index (NNFI)	0.91	Good fit
Comparative fit index (CFI)	0.94	Good fit
Incremental fit index (IFI)	0.94	Good fit
Goodness of fit index (GFI)	0.94	Good fit
Adjusted goodness of fit index (AGFI)	0.90	Good fit
Root mean square residual (RMR)	0.056	Good fit
Model CAIC	268.24	Acceptable

Table 10
Assignment procedure for CS and CL latent variables.

Relations	Path coefficient	Total effect from satisfaction latent variable to loyalty latent variable	Final group assignment
Assignment procedure for satisfaction (CS) latent variables			
Sat.a → Loy.a	0.23	0.23 + 031 = 0.54	Sat.a → G1
Sat.a → Loy.b	0.31		(0.54 < 0.90: Low satisfaction area)
Sat.b → Loy.a	0.48	0.48 + 0.42 = 0.90	Sat.b → G3
Sat.b → Loy.b	0.42		(0.90 > 0.54: High satisfaction area)
Relations	Path coefficient	Total effect to loyalty latent variable from satisfaction latent variable	Final group assignment
Assignment procedure for loyalty (CL) latent variables			
Sat.a → Loy.a	0.23	0.23 + 0.48 = 0.71	Loy.a → G2
Sat.b → Loy.a	0.48		(0.71 < 0.73: Low satisfaction area)
Sat.a → Loy.b	0.31	0.31 + 0.42 = 0.73	Loy.b → G4
Sat.b → Loy.b	0.42		(0.73 > 0.71: High satisfaction area)

(Sat.a)" and second one is "Satisfaction on Price and Campaigns (Sat.b)". CL criteria are divided into two sub-classes: "Brand loyalty (Loy.a)" and "Loyalty and Trust on Quality (Loy.b)". Average score for each CS and CL criteria group are given in Table 8 below:

4.2.6. Application of Step 6. Creating final CS–CL matrix with Structural Equation Modeling (SEM) results

In this Step of application, firstly we determine the final position of each latent variable (we consider a latent variable as a group of observed variables) in CS–CL matrix according to quantity of influence by comparing path coefficients. In Fig. 6 standardized solution estimate of final structural model is depicted.

Goodness of statistical fit results is presented in Table 9. The model produces good fit results which are higher than lower bounds defined by Bentler and Bonett (1980), Byrne (1998), Jöreskog and Sörbom (2006), and Cokluk and Sekercioglu (2012).

The assignment procedure of positioning latent variables in CS–CL matrix is shown in Table 10. Here we determine final group of CS and CL criteria by evaluating the total effect (sum of path coefficients) in structural model depicted in Fig. 6.

According to final assignment results shown in Table 10, final CS–CL matrix is built and depicted in Fig. 7.

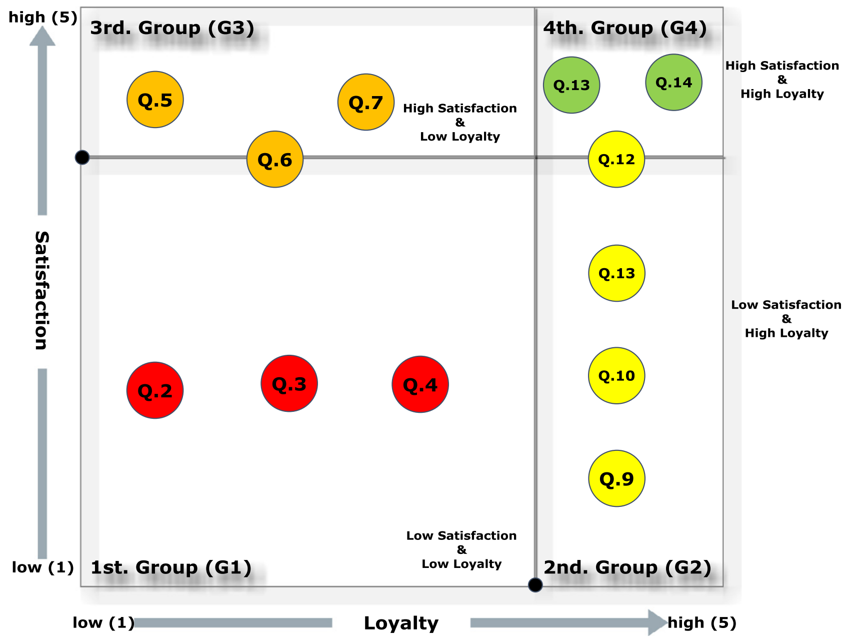


Fig. 7. Final CS–CL matrix.

After final CS–CL matrix is created we bridge the results of classification (found in Steps 2–4) with structural model created in this step. Then, strategies are developed by integrating customer groups with related CS and CL criteria that are placed in the same region of matrix. The interpretation of results and customer strategy improvements based on related literature and expert views are discussed in the next and last section of this study.

5. Conclusion and discussions

In this study, for the evaluation of the customer-related data, an alternative post-analysis method is developed by using classification algorithms and Structural Equation Modeling (SEM). Firstly in order to determine CS and CL criteria, confirmatory factor analysis method is used. This is the basis and first sub-step of SEM approach. Then, we discover customer groups with classification algorithms. Classification algorithms are not only used for customer group segmentation but also for determining breaking points of CS and CL on developed CS–CL matrix. After customer groups and CS–CL criteria positions are assigned to the matrix, finally we use method results to define strategies and priorities according to satisfaction and loyalty criteria and their relationships.

The implementation of the method in customer satisfaction and loyalty analysis offers a quantitative post-analysis evaluation of satisfaction and loyalty levels and integration of customer grouping method with criteria grouping in a matrix based approach. The main objective of strategies is increasing the total share of high satisfaction and high loyalty segment and thus achieving revenue growth and profitability which are the main target of CS analyses (Anderson & Mittal, 2000; Churchill, 2001; Hill, 1996; Hill & Alexander, 2006; Johnson, 2000; Kotler & Armstrong, 1994; Ranaweera & Prabhu, 2003). Customer strategies according to each group, discovered after implementation of method in a real-world application in white-goods example are presented below:

i. *Customer group 1 (CG1-Low satisfaction and low loyalty group)*: This is the most desperate group of customers. Customers in this segment are more effective to express their dissatisfaction to other groups than customers in G3. Mostly related and effective criteria on CG1 determined with 6-step CS–CL method are Q.2, Q.3, Q.4 and Q.6 (Satisfaction on Functional and technical properties). Strategies of top priority for CG1 are; moving to G4 region by;
   - Giving priority to answering to customer requests on functional and technical properties of product (Ersoz et al., 2008; Lai, Xie, Tan, & Yang, 2008; Odabasi, 2000),
   - Moving to G3 region and then to G4 region by giving priority to provide customer needs and then creating brand loyalty.

ii. *Customer group 2 (CG2-Low satisfaction and high loyalty group)*: This is the most complaining group of customers and they are generally loyal customers because of some obligations. Customers' loyalty is mostly based on brand name. Mostly related and effective criteria on CG2 determined with 6-step CS–CL method are Q.9, Q.10, Q.12 and Q.13 (Brand loyalty). Strategies of top priority for CG3 are; moving to G4 region by:
   - Giving priority to increase sustainability of customer,
   - Giving priority to comply with to customers' recommendations (Odabasi, 2000).

iii. *Customer group 3 (CG3-High satisfaction and low loyalty group)*: Gaining and retaining this group is very difficult and losing them is so easy. Mostly related and effective criteria on CG3 determined with 6-step CS–CL method are Q.5, Q.6 and Q.7 (Satisfaction on Price and campaigns). Strategies of top priority for CG3 are; moving to G4 region by:
   - Giving priority to meet to customer standards and customization,
   - Implementing flexible pricing policy to transform price-dependent behavior to brand-dependent behavior.

iv. *Customer group 4 (CG4-High satisfaction and high loyalty group)*: This is the best and target group of customers. They share their satisfaction with other customers more than other groups. Mostly related and effective criteria on CG4 determined with 6-step CS–CL method are Q.12, Q.13 and Q.14 (Loyalty and Trust on Brand Quality). Retaining G4 region and maintaining customer retention by;

- Preventing competitors by innovative marketing applications,
- Giving priority to creating new unique benefits (Ranaweera & Prabhu, 2003).

The main advantages of the CS–CL matrix method are;

- Implementation of a comprehensive CS and CL analysis with a customer satisfaction and customer loyalty matrix constructed by a 6-step data mining approach,
- Discovering hidden patterns of customer data by creating decision trees with best performing classification algorithms and developing a classification approach taking into consideration both customer groups and criteria,
- Finding out relationships among CS and CL criteria, positioning on the CS–CL matrix and integrating structural model results,
- Developing key strategies for CS and CL improvement.

Future research regarding the CS–CL matrix method is mainly focused on comparison analysis with other alternative satisfaction measurement approaches like fuzzy sets and other advanced prediction methods like artificial neural networks. Moreover, the problem of selecting appropriate values for the parameters of the method (breaking point) and its impact to the reliability and stability of the provided results should be studied. Finally, it is interesting to apply the CS–CL matrix method in other industries and comparing results.

Appendix A

Satisfaction and loyalty survey questionnaire on white-good brands

1. Part personnel data.

Gender	M	F		
Age				
Educational Level	Primary	H. School	Bachelor's	Master's

2. Part please rate each question according to 5-point Likert scale given below.

Very Low	Low	Neutral	High	Very High
1	*2*	*3*	*4*	*5*

Survey Questions	*1*	*2*	*3*	*4*	*5*
Q.1. Importance degree of service network in white-goods preference					
Q.2. Importance degree of energy consumption in white-goods preference					
Q.3. Importance degree of functional properties in white-goods preference					
Q.4. Quality level of product compared to costs					
Q.5. Importance degree of price in white good preference					
Q.6. General satisfaction level					
Q.7. Effecting level of campaigns on my decisions					
Q.8. Distance to store is not much more important					
Q.9. Level of Brand's general qualifications that are better than other brands					
Q.10. Importance degree of physical appearance of Brand in white-goods preference					
Q.11. Importance degree of TV and internet ads of Brand					
Q.12. Importance degree of new technologies of Brand in white-goods preference					
Q.13. Importance degree of Brand name in white good preference					
Q.14. Trusting level to information provided by manufacturing company about products					
Q.15. General trust level to the quality of brand					

References

Aktas, H., Akarsu, T., Bicici, S., Gokcin, G., Erkay, H., Gokkus, S., et al. (2000). Musteri Memnuniyeti Yonetimi. KAL-DER Yayinlari, No: 31.

Anderson, E. W., & Mittal, V. (2000). Strengthening the satisfaction–profit chain. *Journal of Service Research, 3*(November 2), 107–120.

Bentler, P. M. (1980). Multivariate analysis with latent variables: Causal modeling. *Annual Review of Psychology, 31*, 419–456.

Bentler, P. M., & Bonett, D. G. (1980). Significance tests and goodness of fit in the analysis of covariance structures. *Psychological Bulletin, 88*, 588–606.

Bollen, K. A. (1989). *Structural equations with latent variables*. New York: Wiley.

Byrne, B. M. (1998). *Structural equation modeling with LISREL, PIRELIS and SIMPLIS: Basic concepts, applications, and programming*. Hillsdale, NJ: Lawrence Erlbaum Associates.

Chen, S. C. (2012). The customer satisfaction–loyalty relation in an interactive e-service setting: The mediators. *Journal of Retailing and Consumer Services, 19*, 202–210.

Chikara, T., & Takahashi, T. (1997). Research of measuring customer satisfaction for information systems. *Computers and Industrial Engineering, 33*(3–4), 639–642.

Chou, C. P., & Bentler, P. M. (1996). Estimates and tests in structural equation modeling. In R. H. Hoyle (Ed.), *Structural Equation Modeling* (pp. 37–55). Thousand Oaks, CA: Sage Publications.

Churchill, C. F. & Halpern, S. S. (2001). *Building Customer Loyalty*. (http://www.microfinancegateway.org/gm/document-1.9.27775/2685.pdf).

Ciavolino, E., & Dahlgaard, J. J. (2007). ECSI – Customer satisfaction modeling and analysis: A case study. *Total Quality Management & Business Excellence, 18*(5), 545–554.

Cokluk, O., Sekercioglu, G., & Buyukozturk, S. (2012). *Sosyal bilimler icin cok degiskenli istatistik: SPSS ve LISREL Uygulamalari*. Ankara: Pegem Akademi. ISBN-ISSN: 978-605-5885-67-0.

Cronin, J. J., & Taylor, S. A. (1992). Measuring service quality: Examination and extension. *Journal of Marketing, 56*(3), 55–68.

Cronin, J., Brady, M., & Hult, T. (2000). Assessing the effects of quality, value and customer satisfaction on consumer behavioral intentions in service environments. *Journal of Retailing, 76*(2), 193–218.

Deng, Z., Lu, Y., Wei, K. K., & Zhang, J. (2010). Understanding customer satisfaction and loyalty: An empirical study of mobile instant messages in China. *International Journal of Information Management, 30*, 289–300.

Ersoz, S., Yaman, N., & Birgoren, B. (2008). Modeling and analyzing customer data in customer relationship management with artificial neural networks. *Journal of the Faculty of Engineering and Architecture of Gazi University, 23*(4), 759–767.

Fornell, C. (1992). A national satisfaction barometer: The Swedish experience. *Journal of Marketing, 56*, 6–21.

Fornell, C., Michael, D. J., Eugene, W. A., Jaesung, C., & Barbara, E. B. (1996). The american customer satisfaction index: Nature, purpose and findings. *Journal of Marketing, 60*, 7–18.

Gale, B., & Wood, R. (1994). *Managing customer value: Creating quality and service that customers can see*. New York: The Free Press. ISBN: 0-02-911045-9.

Garbarino, E., & Johnson, M. S. (1999). The different roles of satisfaction, trust and commitment in customer relationships. *Journal of Marketing, 60*(2), 70–87.

Gerpott, T. J., Rams, W., & Schindler, A. (2001). Customer retention, loyalty, and satisfaction in the German mobile cellular telecommunications market. *Telecommunications Policy, 25*, 249–269.

Grigoroudis, E., Samaras, A., Matsatsinis, N. F., & Siskos, Y. (1999). Preference and customer satisfaction analysis: An integrated multi criteria decision aid approach. *International Conference on Integrating Technology & Human Decisions, 2*, 1350–1352.

Grigoroudis, E., & Siskos, Y. (2002). Preference disaggregation for measuring and analysing customer satisfaction: The MUSA method. *European Journal of Operational Research, 143*, 148–170.

Grigoroudis, E., & Siskos, Y. (2004). A survey of customer satisfaction barometers: Some results from the transportation–communications sector. *European Journal of Operational Research, 152*, 334–353.

Hall, M., Frank E., Holmes, G., Pfahringer, B, Reutemann, P., & Witten, I. H. (2009). The WEKA data mining software: An update. SIGKDD explorations, Vol. 11, issue 1.

Hill, N. (1996). *Handbook of customer satisfaction measurement*. Aldershot, Hampshire: Gower Press.

Hill, N., & Alexander, J. (2006). *The handbook of customer satisfaction and loyalty measurement*. Gower Publishing.

Homburg, C., & Furst, A. (2005). How organizational complaint handling drives customer loyalty: An analysis of the mechanistic and the organic approach. *Journal of Marketing, 69*(July), 95–114.

Johnson, M. D. (2000). *Improving customer satisfaction, loyalty, and profit: An integrated measurement and management system*. University of Michigan, Business School Management Series.

Jöreskog, K. G., & Sörbom, D. (1993). *Lisrel 8: Structural equation modeling with the SIMPLIS command language*. Lincolnwood, IL: Scientific Software International.

Jöreskog, K. G., & Sörbom, D. (2006). *Lisrel 8.80 for windows (computer software)*. Lincolnwood, IL: Scientific Software International, Inc.

Kengpol, A., & Wangananon, W. (2006). The expert system for assessing customer satisfaction on fragrance notes: Using artificial neural networks. *Computers and Industrial Engineering, 51*, 567–584.

Kline, R. B. (2005). *Principles and practice of structural equation modeling* (2nd ed.). Guilford Press.

Kotler, P., & Armstrong, G. (1994). *Principles of marketing*. Prentice Hall.

Kumar, V., Pozza, I. D., & Ganesh, J. (2013). Revisiting the satisfaction–loyalty relationship: Empirical generalizations and directions for future research. *Journal of Retailing, 89*, 246–262.

Lai, X., Xie, M., Tan, K. C., & Yang, B. (2008). Ranking of customer requirements in a competitive environment. *Computers and Industrial Engineering, 54*, 202–214.

Lam, S. Y., Venkatesh, S., Krishna, M. E., & Bvsan, M. (2004). Customer value, satisfaction, loyalty and switching costs: An illustration from a business to business service context. *Journal of the Academy of Marketing Science, 32*(3), 293–311.

Liu, X., Zeng, X., Xu, Y., & Koehl, L. (2004). A fuzzy model of customer satisfaction index in e-commerce. *Mathematics and Computers in Simulation, 77*. 512-421.

Martensen, A., Kristensen, K., & Grönholdt, L. (2000). Customer satisfaction measurement at Post Denmark: Results of application of the European Customer Satisfaction Index. *Total Quality Management, 11*(7), 1007–1015.

Ngai, E. W. T., Li, X., & Chau, D. C. K. (2009). Application of data mining techniques in customer relationship management: A literature review and classification. *Expert Systems with Applications, 36*, 2592–2602.

Ngobo, P. V. (1999). Decreasing returns in customer loyalty: Does it really matter to delight the customer. *Advances in Consumer Research, 26*, 469–476.

Odabasi, Y. (2000). *Satista ve Pazarlamada Musteri Iliskileri Yonetimi*. Istanbul: Sistem Yayincilik.

Orel, F. D., & Kara, A. (2013). Super market self-checkout service quality, customer satisfaction and loyalty: Empirical evidence from an emerging market. *Journal of Retailing and Consumer Services*. http://dx.doi.org/10.1016/j. jretconser.2013. 07.002i.

Ranaweera, C., & Prabhu, J. (2003). The influence of satisfaction, trust and switching barriers on customer retention in a continuous purchasing setting. *International Journal of Service Industry Management, 14*(4), 374–395.

Schumacker, R. E., & Lomax, R. G. (1996). *A beginner's guide to structural equation modeling*. Mahwah, NJ: Lawrence Erlbaum Associates, Publishers.

Shao-I, C., Ching-Chan, C., Tieh-Min, Y., & Hsiu-Yuan, H. (2011). Preliminary research on customer satisfaction models in Taiwan: A case study from the automobile industry. *Expert Systems with Applications, 38*(8), 9780–9787.

Simsek, O. F. (2007). Yapisal Esitlik Modellemesine Giris: Temel Ilkeler ve LISREL Uygulamalari, Ekinoks.

SPSS Inc. (Released 2007). SPSS for Windows. Version 16.0, Chicago, SPSS Inc.

Tan, P. N., Steinbach, M., & Kumar, V. (2005). Introduction to data mining. Pearson international edition. (1st ed.). ISBN-10: 1292026154.

Turkyılmaz, A., & Ozkan, C. (2007). Development of a customer satisfaction index model: An application to the Turkish mobile phone sector. *Industrial Management & Data Systems, 107*(5), 672–687.

Witten, I. H., Frank, A., & Hall, M. A. (2011). *Data mining: Practical machine learning tools and techniques* (2nd ed.). Morgan Kaufmann. ISBN 0-12-088407-0.

Vesel, P., & Zabkar, V. (2009). Managing customer loyalty through the mediating role of satisfaction in the retail loyalty program. *Journal of Retailing and Consumer Services, 16*, 396–406.

Please cite this article in press as: Aktepe, A., et al. Customer satisfaction and loyalty analysis with classification algorithms and Structural Equation Modeling. *Computers & Industrial Engineering* (2014), http://dx.doi.org/10.1016/j.cie.2014.09.031