ORIGINAL RESEARCH

CrossMark

Customer Behavior Mining Framework (CBMF) using clustering and classification techniques

Farshid Abdi[1] · Shaghayegh Abolmakarem[1]

Abstract

The present study proposes a Customer Behavior Mining Framework on the basis of data mining techniques in a telecom company. This framework takes into account the customers' behavior patterns and predicts the way they may act in the future. Firstly, clustering technique is used to implement portfolio analysis and previous customers are divided based on socio-demographic features using k-means algorithm. Then, the cluster analysis is conducted based on two criteria, i.e., the number of hours the telecom services used and the number of the services selected by customers of each group. Six groups of customers are identified in three levels of attractiveness according to the results of the customer portfolio analysis. The second phase has been devoted to mining the future behavior of the customers. Predicting the level of attractiveness of newcomer customers and also the churn behavior of these customers are accomplished in the second phase. This framework effectively helps the telecom managers mine the behavior of their customers. This may lead to develop appropriate tactics according to customers' attractiveness and their churn behavior. Improving managers' abilities in customer relationship management is one of the obtained results of the study.

Keywords Customer attractiveness · Customer behavior · Customer portfolio analysis · Segmentation · Churn prediction · Data mining

Introduction

Customer relationship management (CRM) has turned into a significant field in telecommunication business. CRM is assumed as an intangible asset which provides the potential of competition for the organizations (Ryals 2002). So, the companies try to identify and analyze the behavior of their customers. Technological improvement has enabled the telecom companies to store record of customers. Analyzing the historical data help the companies to discover the behavioral patterns of existing customers which could result in a significant impact on predicting the future customers' behavior.

Customer portfolio analysis (CPA) is an effective tool to investigate the customer behavior. The aim of CPA is to segment customers into groups (Thakur and Workman 2016). Customer segmentation is the use of past data to divide customers into similar groups based on various features (Hsu et al. 2012). Using the customer segmentation process, the company will be able to identify the customers who are strategically important and profitable. These customers can be categorized into two main classes including high future lifetime value customers or high volume customers (Buttle and Maklan 2015).

Competitors are ready to provide the same services and products with higher quality and lower prices. Customers simply will leave the company for lower costs or higher quality (Keramati et al. 2014). Losing customers also leads to opportunity costs because of decreases in sales (Verbeke et al. 2011). The previous studies have shown that retaining the current valuable customers of the organization is much cheaper than attracting new ones. A little increase in customer retention can lead to a remarkable increase in profits. Hence, paying attention to the retention of the customers is considered as one of the most crucial strategic components of profitability in the companies (Verbeke et al. 2011). If the obtained data from the analysis of the customer

✉ Farshid Abdi
farshidabdi@azad.ac.ir

[1] Department of Industrial Engineering, South Tehran Branch, Islamic Azad University, Tehran, Iran

Springer

behavior is correctly applied, then the companies can increase the rate of customer retention and turn more profits by taking proper actions.

In this paper, a Customer Behavior Mining Framework (CBMF) consisting of two phases is proposed in order to analyze the past, current and future behavior of telecom company's customers. The first phase of CBMF relates to CPA. Data mining techniques such as clustering method are used to carry out portfolio analysis phase. The main modules of the first phase are as follows: segmenting customers according to their socio-demographic features, analyzing the clusters based on customers' behavioral features, forming the portfolio matrix to help the company identify different types of customer. In the second phase, the future customers' behavior is predicted according to the past data. The type of the new customers and the churn behavior of them are predicted using rule-based classification methods in the second phase. Finally, different CRM tactics are provided according to customer attractiveness and churn behavior.

The main contributions of this paper are summarized as follows:

- CBMF is a quite functional framework that can be used by many customer-centric businesses. It can analyze the behavior of old customers, takes the pattern of their behavior and uses it to predict the future. This framework takes into account two of the most important factors in customer relationship management, namely customer attractiveness and churn rate.
- Defining attractiveness criterion and create attractiveness matrix is related to the definition of each business about customer attractiveness, So it's a flexible framework.
- This framework has capability of predicting the level of newcomer customers' attractiveness based on their socio-demographic features at the beginning of their entry to the company.
- Customer churn prediction in CBMF is conducted in two phases. Firstly, at the time of the arrival of the customers to the company and secondly after a month and based on their behavioral features. This framework will help businesses track their customers' behavior during their life cycle.
- This framework addresses the managerial tactics given the level of attractiveness of each customer and the possibility of customer churn.

The remaining parts of the paper are organized as follows. "Literature review of Past works" section presents a literature review on previous relevant research works. The proposed CBMF is proposed in "Theoretical background" section. The case study and results are presented in "Proposed method" section. "Case study and results" section

presents the conclusion remarks and future research directions.

Literature review of past works

Customer portfolio management

Portfolio of customers of a company consists of clients that are clustered according to one or more strategically important variables. Each cluster of clients has a different value for the company. So, they should be managed in different ways. The aim of customer portfolio analysis is to divide customers into mutually exclusive clusters in order to identify profitable and valuable customers (Buttle and Maklan 2015). Consequently, the companies can apply marketing tactics to retain and develop valuable customers (Thakur and Workman 2016).. A summary of previous studies on customer portfolio analysis is mentioned in Table 1.

Customer segmentation

Customer segmentation divides customers into groups with similar characteristics, requirements and behaviors. Segmenting customers can be done based on different classes of features such as user attributes and usage attributes. User attributes include demographic attributes (i.e., age, gender, job status, and marital status), geographic attributes (i.e., country, region) and psychographic attributes (i.e., lifestyle). Usage attributes contain information about the frequency and extent of purchase and buyers' behavior (Mohammadi et al. 2013; Buttle, and Maklan 2015). Customer clustering is used to build the customers' profiles which makes up the core of a customer-centric information system (Bose and Chen 2015). Several authors used various segmentation criteria and various clustering techniques in order to group customers. A summary of previous studies on customer segmentation is highlighted in Table 2.

It can be concluded from Table 2 that various clustering methods were used for grouping customers in different fields including banking, telecommunications, restaurant, textile manufacturing, hospital and tourism. In the present study, customer segmentation is firstly conducted according to customers' socio-demographic features and afterward cluster analysis is implemented based on customers' behavioral features.

Churn prediction using data mining techniques

Recently, many researchers have paid much attention to the topic of customer churn in various service industries (Chen et al. 2012).

Table 1 Selected customer portfolio analysis literature

Researcher(s)	Description	Suggested criteria
Thakur and Workman (2016)	Used the customer portfolio management model and proposed a method that enable companies to define the value of customers and segment these customers into portfolios	Cost to serve Value of the customer for the company
Ritter and Andersen (2014)	Developed a Six-Pack portfolio that is a useful tool for managers to achieve a better understanding of their customer portfolio and the business potential of this customer portfolio	Customer profitability Customer commitment Growth potential
Ryals (2010)	Analysis for risk adjusted value of the firm's customer portfolio	Returns from customer portfolio Risk from customer portfolio
Rangan et al. (1992)	Developed a buying-behavior-based framework suitable for micro-segmenting customers in mature industrial markets	Cost to serve Price
Shapiro et al. (1987)	This model focuses on customer only as a profit center, and they are categorized only according to their profitability	Cost to serve (presale costs, production costs, distribution costs, post-sale costs) Net price
Fiocca (1982)	Fiocca (1982) proposed a two-step customer portfolio analysis. Fiocca explained various factors associated with the customer buying behavior and supplier relationships	Step I Strategic importance of the account Difficulty in managing the account Step II Customer's business attractiveness Relative strength of the relationship

Customer defection means customer withdrawal from further cooperation with the company. The term "Churn" refers to the customers who shift from one service provider to another one (Farquad et al. 2014). The aim of churn prediction models is to identify customers who are most likely to abandon the service provider. This prediction may be accomplished on the basis of socio-demographic or behavioral features (Chen et al. 2012). Models of churn prediction are usually based on historical data collected from customers (Guelman et al. 2012). Using various data mining techniques for churn prediction has been widely reported in previous studies. A summary of previous studies on churn prediction is presented in Table 3.

As shown in Table 3, in previous studies, various data mining techniques have been used for customer churn prediction in different fields including telecommunications, banking, e-commerce, newspaper Company and supermarket. In this paper, churn prediction is performed b according to customers' socio-demographic features as well as their behavioral features.

Theoretical background

Clustering algorithms

The k-means clustering algorithm is the simplest unsupervised learning algorithm. In this method, the dataset is divided into the number of predetermined clusters. The main idea in this algorithm is firstly define the k initial cluster center, and k is the number of the clusters.

The best choice for the clusters' centers in this algorithm is placing them (centers) more far from each other. Afterward, each record in the dataset is assigned to the nearest cluster center. After allocating all the records to one of the clusters formed for each cluster, a new point is calculated as the center. Again, each record in the dataset is attributed to the cluster where its center has the smallest distance to that record. The steps for determining cluster centers and assigning the records to the nearest cluster will be repeated until no change is made in the cluster centers. (Mehmanpazir, and Asadi 2017).

Clustering evaluation metrics

Davies–Bouldin

The best number of clusters can be determined by measuring the evaluation metrics such as Davis-Bouldin Index. This index is a function of total diffraction ratio into the cluster to the distance between the clusters, which is referred to below. (Mitra et al. 2010).

$$DB = \frac{1}{n} \sum_{\substack{i-1 \\ i \neq j}}^{n} \max \left\{ \frac{\Delta(Q_i) + \Delta(Q_j)}{\delta(Q_i.Q_j)} \right\} \tag{1}$$

where n is the number of cluster, $\Delta(Q_i)$ is the intra-cluster distance, and $\delta(Q_i.Q_j)$ is the inter-cluster distance. The

Table 2 Brief literature of customer segmentation researches

Researcher(s)	Description	Dataset/features	Clustering method
Dursun and Caber (2016)	Profiling profitable hotel customers by RFM analysis	Hotel customers/RFM	Self-organizing maps (SOM) and k-means
Rezaeinia and Rahmani (2016)	Presented a new method to increase the accuracy and quality of recommendations associated with filtering systems	Wholesale center in Tehran/weighted RFM	EM-algorithm
Safari et al. (2016)	Proposed a new method of segmenting customer	IT company in Iran/RFM	Fuzzy clustering (fuzzy c-means) and fuzzy AHP (as a ranking method)
Weng (2016)	The paper aims to understand the book subscription characteristics of the students at each college and help the library administrators to conduct efficient library management plans for books in the library/integrates RFC model and ARM technique	library users/Recency-Frequency-College (RFC) model	Association rule mining
(Wang et al. 2016)	Proposed a novel bi-clustering-based market segmentation method using customer pain points	Chinese famous mobile phone company/ 40 major customer pain points for the registered members to vote	Optimal bi-clustering algorithm
		Additional information about participants, such as gender, age, city, education, occupation and monthly income and so on	
Bose and Chen (2015)	Developed the extended FCM algorithm to detect significant changes in behavioral patterns of customers of mobile services over time	Mobile telecommunications services provider/usage and revenue attributes	Fuzzy c-means clustering algorithm
Güçdemir and Selim (2015)	Propose a method for segmenting business customers that combining clustering and multi-criteria decision making.	Business customers of an international original equipment manufacturer (OEM)/RFM	Hierarchical and partitional clustering algorithms
QuelhasBrito et al. (2015)	Used data mining (DM) techniques to solve marketing and manufacturing problems in the fashion industry	Manufacturer of custom made shirts/ bivolino.com/demographic and biometric, geographic, psychographic, behavioral	Clustering and subgroup discovery: k-medoids and CN2SD
Dzobo et al. (2014)	Presented a multi-dimensional customer segmentation model for reliability-worth analysis of power systems	Electricity customers/economic size, economic activity and energy consumption	Hierarchical clustering technique
Chuang, Chia, and Wong (2013)	Provide a data mining approach to classify Taiwanese healthcare institutions based on customer value assessment	Pharmaceutical marketing in Taiwan/ FMC	C5.0 decision tree was then used to generate the behavioral rules of various segmentations
Wei et al. (2013)	In this article, customers have been clustering. Also appropriate strategies have been developed for each cluster	Hair salon in Taiwan/RFM	Self-organizing maps (SOM) and k-means methods

small value of Davis–Bouldin's index represents a valid clustering.

Classification algorithms

•Neural networks

A neural network includes a layered network and feedforward and is completely composed of the artificial neurons connected to nodes. The nature of the network feedforward limits the network flow in one direction and prevents the creation of a looping or cycling. A neural network involves two or more layers. Neural networks usually have an input layer, a hidden layer, and an output layer.

A neural network is completely interconnected. In other words, each node in each layer is connected to all subsequent nodes while it is not connected to other nodes of its layer. Every connection between the nodes has a weight. First, the values between zero and one are randomly

Table 3 Selected customer churn prediction literature

Researcher(s)	Description	Dataset/features	Method
Backiel et al. (2016)	This study investigates the incorporation of social network information into churn prediction models to improve accuracy, timeliness and profitability	Telecommunications/customer information and call details	Cox proportional hazards, LR, ANN
Fathian et al. (2016)	Compared single baseline classifiers with ensemble classifiers for churn prediction	Duke University	DT, ANN, KNN, SVM
	Developed a total of 14 prediction models that were classified in four categories:	76 features including eight nominal features and 68 numerical features	Bagging, Boosting
	(1) basic Classifier; (DT, ANN, KNN, SVM)		
	(2) Classifier with SOM + basic classifier;		
	(3) Classifier with SOM + reducing features with PCA + basic classifier; and		
	(4) Classifier with SOM +reducing features with PCA +bagging and boosting ensemble classifier		
Moeyersoms and Martens (2015)	Add high-cardinality attributes to the churn prediction model	Energy supplier in Belgium/	C4.5, Logit, SVM
		Age, average cost of bills, contacts with company	
		Gender, type of contract, package, payment method	
		Family names, ZIP codes or bank account numbers	
Farquad et al. (2014)	Proposed the hybrid approach composing of three phases:	Bank credit card customer/	SVM + Naive Bayes Tree
	1. Selecting subset of features using SVM-RFE	Two groups of variables:	
	2. Extracting Support vector using SVM	Socio-demographic and behavioral data	
	3. Generating rules using Naive Bayes Tree		
	Further, employed various standards balancing approaches to balance the data and extracted rules		
Keramati et al. (2014)	Proposed a hybrid methodology by combining DT, ANN, KNN and SVM	Telecommunication company/	DT, ANN, KNN, SVM
		Call Failure (CF), number of Complaints (Co), Subscription Length (SL), Charge Amount (CA), Seconds of Use (SU), Frequency of use (FU), Frequency of SMS (FS), Distinct Calls Number (DCN), Age Group (AG), Type of Service (TS), Status (St), and Churn (Ch)	
Kim et al. (2014)	Improved churn prediction in telecommunication industry by analyzing a large network and proposed a new procedure of the churn prediction by examining the communication patterns among subscribers and considering a propagation process in a network based on call detail records which transfer churning information from churners to non-churners	Mobile telecommunication company	Spreading activation (SPA) model
		personal information data and the CDR data	
Lin, Tsai and Ke (2014)	Used dimensionality and data reduction in telecom churn prediction to improve performance.	Telecom dataset	ANN
		173 different features	
	Developed totally eight prediction models that were classified under five categories:		
	Category 1: the baseline/Category 2: feature selection/Category 3: data reduction/ Category 4: feature selection data reduction/ Category 5: data reduction feature selection		

Table 3 (continued)

Researcher(s)	Description	Dataset/features	Method
Verbeke et al. (2014)	Proposed a social network information for customer churn prediction	Mobile telecommunication network variables	Relational and non-relational classification model
Huang, and Kechadi (2013)	Proposed a system combining a modified *k*-means clustering algorithm with a classic rule inductive techniques used to predict the customer future behavior	Telecom dataset 121 attributes, including 11 symbolic attributes and 110 continuous attributes. Demographic profiles/Account information/ Call details	Benchmarked to DT, LR, *k*-NN, SVM, OneR, PART, SePI, Boosted C5.0
Huang et al. (2012)	Presented a new set of features for the customer churn prediction in the telecommunication	Telecoms of Ireland/ Demographic profiles, Information on grants, Customer account information, Service orders, Henley segments, Telephone line information, Complaint information, The record information of bills and payments, Call details, Incoming calls details	Linear classifications, LR, NB, DT, MLP neural network, SVM, Evolutionary Algorithm
Verbeke et al. (2011)	Applied two data mining methods to churn prediction modeling	Wireless telecom operator 21 features	Ant-Miner +/ALBA Benchmarked to traditional rule induction techniques such as C4.5 and RIPPER

assigned to these weights. In neural networks, a composite function (usually the summation $\sum$) produces the linear composition of node inputs and connection weights as a scalar value. Then, this value takes as an input of an activation function (such as sigmoid). The sigmoid activation function is expressed as:

$$y = \frac{1}{1 + e^{-x}} \tag{2}$$

Neural networks represent a supervised learning. For this purpose, these networks require a big Train set from full records including the target variable. Since each observation of the train dataset is processed by the network, an output value is produced from the output node. Then, this value is compared to the actual value of the target variable and the error value is calculated.

The error back propagation algorithm calculates the prediction error for a record and distributes the error in the network. Such an algorithm assigns a share of error to each connection. Then, the weights of these connections are adjusted using the gradient descent method for reducing the error (Markopoulos et al. 2016).

●**Decision tree**

Decision tree induction refers to learn and construct decision trees from training tuples that have target class. The goal of decision tree is to create a classification model that predicts the value of a *target attribute* (often called *class* or *label*) based on several input attributes of the dataset.

A decision tree, as its name implies, has a tree structure similar to a flowchart. Each internal node in this tree displays the test on an attribute, and a class label is kept on each leaf node or end node. The highest node in a tree is the root node.

Decision trees generated by recursive partitioning. Recursive partitioning means repeatedly splitting on the values of attributes.

The attribute selection measures are used to choose an attribute to divide the tuples into separate categories as best as possible. Information gain is considered as one of the most famous measures (Han et al. 2012).

●**Attribute selection measures (*Splitting Criterion*)**

Information gain is a measure that provides a set of same-rank training records for each feature. The criterion of "Entropy" which is widely used in information theory is also used to determine information gain (Han et al. 2012).

The expected information required to classify a tuple in D is given by Eq. (3):

$$\text{Info}(D) = -\sum_{i=1}^{m} p_i \log_2(p_i) \tag{3}$$

Here, D is the set of training records, p_i is the probability of the belonging of a record (in D) to class C_i, and m is the number of classes. In this formula, $\text{Info}(D)$ represents the average required information for the identification of the label of a record class in D. When records of D are partitioned on the basis of a feature such as A (which has v distinct values of $a_1, a_2, \ldots, a_v$), Eq. (4) is used.

$$\text{Info}_A(D) = \sum_{j=1}^{v} \frac{|D_j|}{|D|} \times \text{Info}(D_j). \tag{4}$$

$\frac{|D_j|}{|D|}$ is the weight of partition j. $\text{Info}(D)$ is the expected information required to classify a tuple from D according to the partitioning on A.

Information gain index can be defined as the difference between initial required data (defined on the basis of proportion of parts) and new required data (which are determined after the partition on the basis of A).

$$\text{Gain}(a) = \text{Info}(D) - \text{Info}_A(D) \tag{5}$$

Proposed method

Dataset description

The dataset used in this paper contains information of the customers of a telecom company. The dataset consists of 25 variables, with 24 predictor variables and 1 target variable. It consists of 1000 instances with 274 records

labeled churned and 726 non-churned ones. The variables are divided into two groups: socio-demographic and behavioral attributes which are described in Table 4.

Structure of proposed Customer Behavior Mining Framework

The proposed framework as shown in Fig. 1 is composed of two phases including clustering and classification. In the first phase, customers are clustered according to socio-demographic features (see Table 4). Afterward, the customers of each cluster are ranked according to their attractiveness for the company. The second stage addresses the behavior mining of the future customer. In this stage, classification methods are employed.

Case study and results

All algorithms of this research have been implemented using Rapid Miner Studio software, version 7.0.001 running on a 2.4 GHz CPU with 8 GB RAM and windows 7-

Table 4 Research variables

| Class of variable		Variable	Description	Type
Socio-demographic attributes		Region	The region where the customer lives	Nominal
		Age	The age of customer	Numeric
		Marital	Marital status: 1: Yes, 0: No	Binominal
		Address	The number of years of residence in current location	Numeric
		Income	The customers' income	Numeric
		Education	The customers' education: 1-Diploma, 2: AS 3: BS 4:MS, 5: PhD	Nominal
		Employment	Years of employment	Numeric
		Retire	Retired or not?: 1: Yes, 0: No	Binominal
		Gender	Gender of customer: 1: Male, 0: Female	Binominal
Behavioral attributes	Hours of usage	Longmon	Hours of using service 1 per month	Numeric
		Tollmon	Hours of using service 2 per month	Numeric
		Equipmon	Hours of using service 3 per month	Numeric
		Cardmon	Hours of using service 4 per month	Numeric
		Wiremon	Hours of using service 5 per month	Numeric
	Selected services	Multiline	Is customer has a multiline phone: 1: Yes, 0: No	Binominal
		Voice	Has voice service or not?: 1: Yes, 0: No	Binominal
		Pager	Has pager or not?: 1: Yes, 0: No	Binominal
		Internet	Has internet or not?: 1: Yes, 0: No	Binominal
		Callid	Has caller ID or not?: 1: Yes, 0: No	Binominal
		Callwait	Has call waiting service or not?: 1: Yes, 0: No	Binominal
		Forward	Has call forwarding service or not?: 1: Yes, 0: No	Binominal
		Confer	Has conference service or not?: 1: Yes, 0: No	Binominal
		Callcard	Has contact card or not?: 1: Yes, 0: No	Binominal
		Wireless	Has wireless system or not?: 1: Yes, 0: No	Binominal
Label		Churn	Churner or Non-churner?: 1: Yes, 0: No	Binominal

Fig. 1 Framework of the research
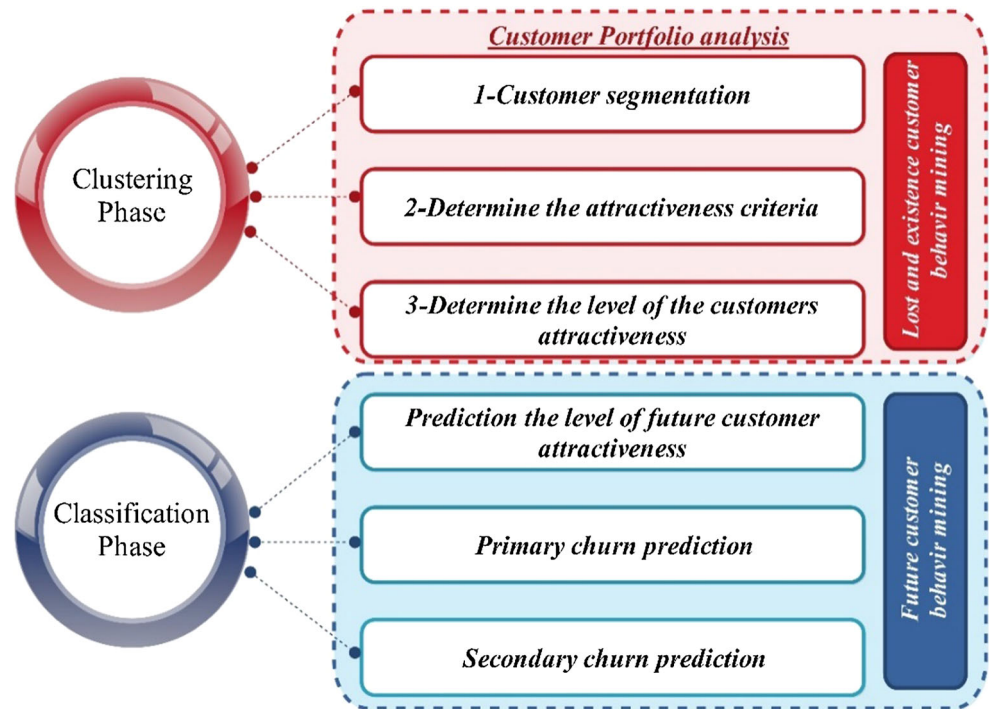
64 bit operating system. The data were pre-screened before clustering and classification phases. Using the control charts and considering 3σ (σ is the standard deviation) as the threshold, outliers were detected and removed. Missing values were also estimated by k-nearest neighbor imputation (k-NNI) mechanism.

The process of creating Customer Behavior Mining Framework (CBMF) is depicted in Fig. 2. As can be seen, CBMF contains two phases namely "clustering" and "classification." Phase 1 starts with customer segmentation. Socio-demographic features are considered as segmentation variables, and customers are clustered using the k-means algorithm. The clustering results are evaluated by Davies–Bouldin index. Analysis of the clusters is then performed based on the attractiveness criterion. The attractiveness criterion is composed of behavioral features (see Table 4). Labeling the clusters is performed based on the average amount of attractiveness criterion.

Phase 2 is related to classification. First, the level of the future customer attractiveness is predicted based on their socio-demographic features. The customer churn prediction is also conducted in two phases. Primary churn prediction is conducted based on customers' socio-demographic characteristics. Secondary prediction is made after 1 month and based on customer's behavior. Rule-based classification algorithms are applied to predict whether the customer is likely to churn. In order to validate the prediction model, split-validation is used, and the
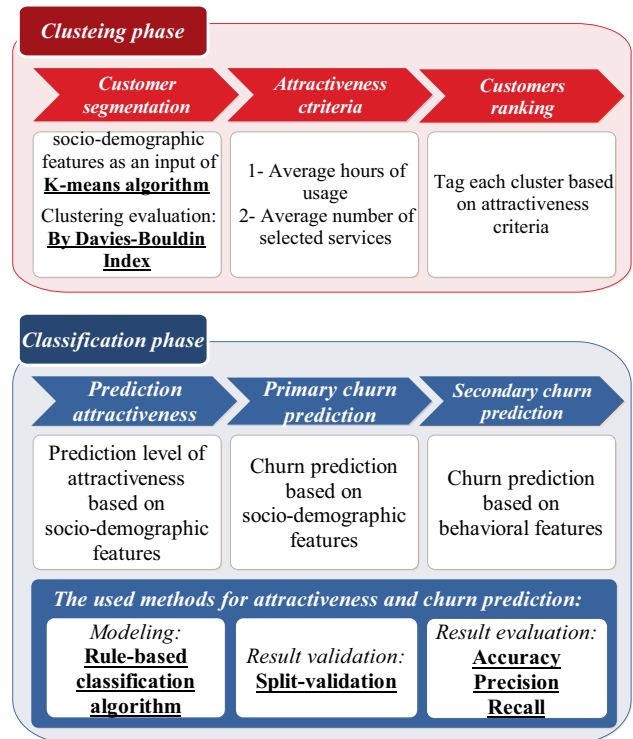
Fig. 2 Design process of CBMF

factors of accuracy, precision and recall are calculated for each model in order to evaluate the prediction performance of models.

Clustering phase

Customer segmentation

In this stage, k-means clustering algorithm is employed in order to manage customer portfolio. Customers' socio-demographic features are taken as an input of k-means algorithm. The optimum number of clusters is determined by calculating Davies–Bouldin Index.

The index was calculated for $k \in [2, 12]$. The results of clustering evaluation are represented in Table 5. According to Davies–Bouldin values, the optimum number of cluster is equal to 6. Therefore, six different groups of customers are identified.

The results of clustering and the number of records of each cluster are represented in Table 6. As can be seen, 120 customers out of 363 customers placed in cluster 5 have been churned. Cluster 2 contains the lowest number of customers with 38 persons.

Defining attractiveness criterion

In this subsection, the customer attractiveness criteria are determined. The behavioral attributes of the customers consist of 15 features that are used to define attractiveness criteria. Five features are related to hours of usage and ten features are related to the number of selected services. Accordingly, these attributes are structured into two dimensions including (a) average hours of using telecom services; and (b) average number of the selected telecom services. These two dimensions are considered as a measure of customers' attractiveness. In this way, the customers who have used more hours of telecom services, as well as those who have chosen more telecommunication services, are more attractive. Table 7 represents the linguistic term and thresholds of the criteria.

It can be concluded from Table 7 that if the average hours of usage are either greater than 60, or less than 30 or between 30 and 60, the high (H), low (L) and medium (M) labels are assigned to the corresponding variable, respectively. Also if the average number of selected services is greater than 4, less than 2 and between 2 and 4, the high (H), low (L) and medium (M) labels are assigned to the corresponding variable, respectively.

A matrix consists of nine cells is created by combining these two dimensions regarding the fact that each of the dimensions consists of three levels (High, medium and low). Figure 3 provides the aforementioned matrix as well as the ranking of the attractiveness criteria.

It can be concluded from Fig. 3 that five levels of attractiveness including highly attractive, attractive, normal, low attractive and unattractive are defined. The customers who are placed in cell number 1 of the matrix are highly attractive customers for Telecom Company. Attractive customers are placed in cells number 2 and 4. Moreover, cells number 3, 5 and 7 contain normal customers. Cells number 6 and 8 contain low attractive customers and the customers of cell number 9 are unattractive.

Cluster analysis based on attractiveness criteria

The total average of behavioral attributes in each cluster is calculated. The results are represented in Table 8. It can be concluded from Table 8 that cluster 2 has the highest average hours of usage and also the highest average number of the selected services.

Figure 4 represents the attractiveness matrix of the studied telecom company. Cluster numbers 1, 2, 3 and 4 in both dimensions have the high (H) label and are placed in

Table 5 Davies–Bouldin Values

K	2	3	4	5	6	7	8	9	10	11	12
Davies–Bouldin	0.63	0.74	0.89	0.91	**0.88**	0.94	1.00	0.99	1.08	1.09	1.09

Table 6 Clustering results

Cluster	Cluster_0	Cluster_1	Cluster_2	Cluster_3	Cluster_4	Cluster_5
No. of records	103	144	38	58	251	363
Non-churner	89	113	29	45	186	227
Churner	14	31	9	13	51	136

Table 7 Grading attractiveness criteria

Criteria	Low (L)	Medium (M)	High (H)
Average hours of usage (h)	< 30	[30–60]	> 60
Average number of selected services (No.)	< 2	[2–4]	> 4

Fig. 3 Attractiveness matrix
and customers' attractiveness
ranking
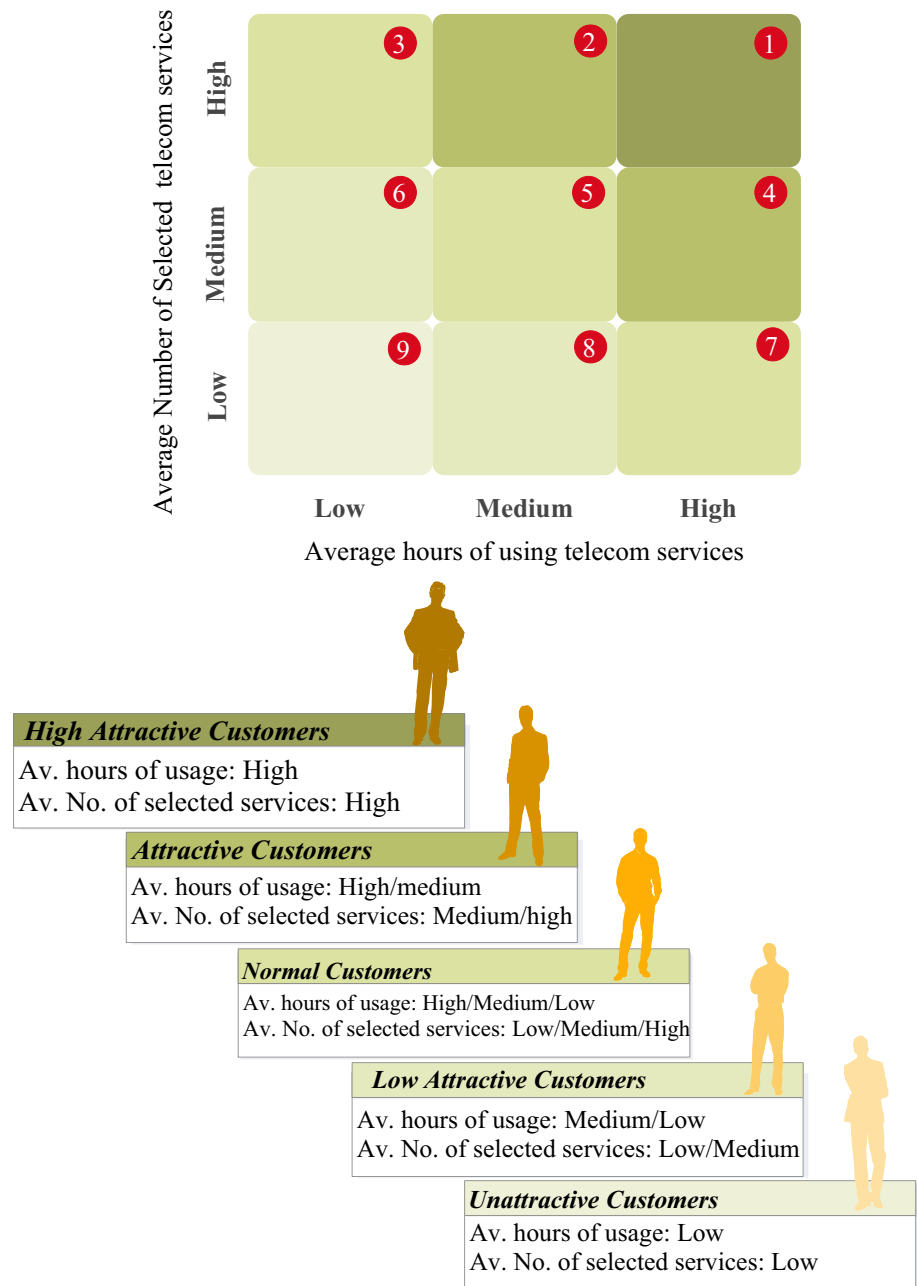
the category of high attractive customers. These clusters include 491 customers that 118 persons of whom have been churned.

Cluster number 5 in both dimensions has the medium (M) label and is placed in the category of normal customers. Customers of cluster 0 have the medium level in terms of average hours of usage and high (H) label and in terms of average of selected services. Thereby, they have been labeled as attractive customers. No customer is located in the low attractive and unattractive levels. Therefore, based on the result of cluster portfolio analysis, six groups of customers in three levels of attractiveness have been identified.

As mentioned before, attracting new customers is more expensive than retaining the existing customers. Therefore, companies should have a plan to retain their existing customers. Customers, who are placed in high attractive cell, are more important for the company. Retention of high attractive customers is significantly important.

Churn prediction problem is one of the results of the CRM application. Having access to a system with the ability of the prediction customer churn, service provider can gain a more accurate understanding and an efficient insight about the customers. In this way, company can formulate better plans and strategies to retain the existing

Table 8 The results of cluster analysis based on attractiveness criteria

Variable/cluster		Cluster_0	Cluster_1	Cluster_2	Cluster_3	Cluster_4	Cluster_5
Average hours of usage	Average (cardmon)	16.33	15.86	19.42	17.26	15.05	9.40
	Average (equipmon)	8.29	14.27	13.99	15.28	16.58	14.29
	Average (longmon)	16.06	14.87	16.20	14.63	11.50	7.54
	Average (tollmon)	12.74	15.36	22.38	20.35	13.62	9.45
	Average (wiremon)	7.23	14.28	19.88	15.95	13.19	8.55
Total		60.6	74.6	91.9	83.5	69.9	49.2
Grade		High	High	High	High	High	Medium
Average number of Selected services	Average (callcard)	0.796	0.785	0.895	0.793	0.704	0.518
	Average (wireless)	0.184	0.326	0.447	0.379	0.335	0.251
	Average (multline)	0.476	0.593	0.553	0.574	0.506	0.374
	Average (voice)	0.165	0.299	0.424	0.345	0.342	0.279
	Average (pager)	0.155	0.313	0.447	0.276	0.267	0.229
	Average (internet)	0.204	0.403	0.289	0.448	0.362	0.411
	Average (callid)	0.379	0.553	0.553	0.614	0.484	0.415
	Average (callwait)	0.427	0.555	0.605	0.569	0.490	0.419
	Average (forward)	0.476	0.535	0.579	0.581	0.486	0.433
	Average (confer)	0.495	0.569	0.592	0.603	0.494	0.423
Total		3.76	4.93	5.38	5.18	4.47	3.75
Grade		Medium	High	High	High	High	Medium
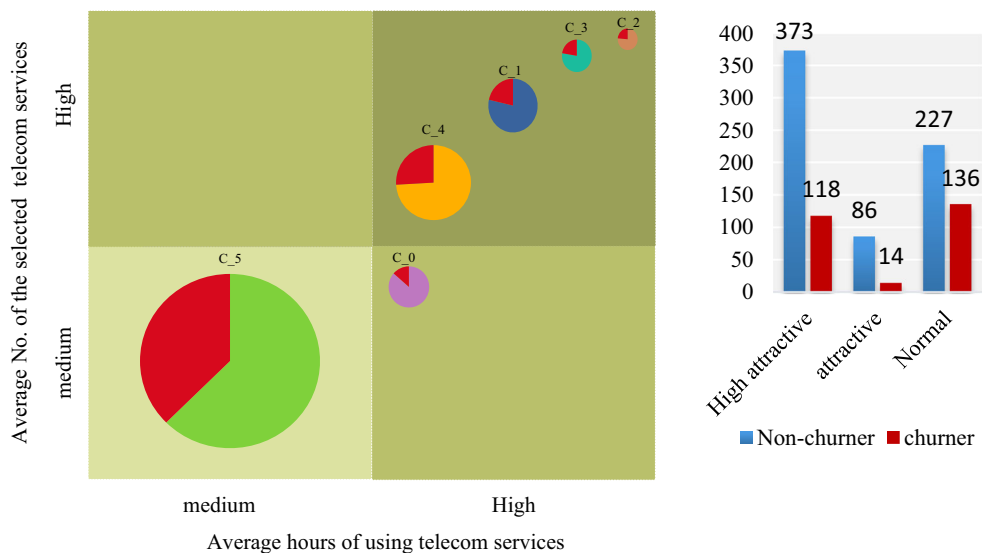
Fig. 4 Attractiveness matrix of the studied telecom company and distribution of customers

customers. Due to the importance of customer churn and retention, these will be discussed in the following sections.

Classification phase

Validation and evaluation of the results

In this phase, the level of the new customers' attractiveness and the churn behavior of them are predicted using several classification methods. In order to evaluate the validation of the models, split-validation method is used. Data are divided into the training set (70%) and testing set (30%).

Confusion matrix is used to evaluate classification performance. Accuracy, precision, and recall evaluation metrics can be explained with respect to a confusion matrix as shown in Table 9 (Kittidecha and Yamada 2018).

These metrics are calculated, respectively, using Eqs. (6), (7) and (8). (Wang and Ma 2012)

Table 9 Confusion matrix

		Prediction	
		Positive	Negative
Actual	Positive	True negative (TP)	False negative (FN)
	Negative	False positive (FP)	True negative (TN)

$$\text{Accuracy} = \frac{TN + TP}{TP + TN + FP + FN} \quad (6)$$

$$\text{Precision} = \frac{TP}{TP + FP} \quad (7)$$

$$\text{Recall} = \frac{TP}{TP + FN} \quad (8)$$

Classification algorithms and parameter setting

Neural networks and decision tree algorithms have been used in order to implementing the classification phase.

- Neural networks

Feedforward neural network trained by a back propagation algorithm (multilayer perceptron) using for prediction the level of the newcomer customers' attractiveness. Neural network parameter settings are shown in Table 10.

- Decision trees

Decision tree algorithm using for prediction the level of the newcomer customers' attractiveness.

The parameters examined in the decision tree algorithm are presented in Table 11.

Prediction of the level of the future customer attractiveness

In this stage, the clusters obtained in the previous step are considered as class variables and the level of customers' attractiveness is predicted using neural networks and decision trees algorithms at the time of the arrival of the customers to the company and based on their socio-demographic features.

Table 10 Neural networks parameter setting

Parameter	Setting
Activation function	Sigmoid function
Learning rate	[0.1–1]
Momentum	[0.1–1]
Training cycles	[100–700]
Hidden layers	0, 1, and 2

Table 11 Decision trees parameter setting

Parameter	Setting
Maximal depth	[2–30]
Minimal gain	[0.1–1]
Minimal size for split	[1–20]
Minimal leaf size	[1–20]
Splitting criterion	Information gain, information gain ratio

Table 12 The results of the attractiveness based on socio-demographic features (using neural networks)

Algorithm	Accuracy	Recall	Precision
Neural network	98.61%	96.81%	97.87%

Table 13 The results of attractiveness prediction based on socio-demographic features (using decision trees)

Algorithm	Accuracy	Recall	Precision
Decision tree	98.26%	97.63%	96.69%

- Neural networks

One of the best results of neural networks is shown in Table 12. In this case, the training cycle is equals to 600, the momentum is equal to 0.5, the learning rate is equal to 0.9, and the number of hidden layers is equals 2.

It can be concluded from Table 12 that the accuracy, precision and recall of this model are 98.61%, 96.81% and 97.87%, respectively.

- Decision trees

One of the best results of using the decision tree algorithm in order to prediction the level of attractiveness is presented in Table 13.

In this case, Information gain is considered as splitting criterion. Maximal depth is equal to 23, Minimal gain is equal to 0.3, Minimal size for split is equal to 5, and Minimal leaf size is equal to 2.

It can be concluded from Table 13 that the accuracy, precision and recall of this model are 98.26%, 97.63% and 96.69%, respectively.

Several rules extracted from decision tree algorithm are presented as follows:

- if income > 42.937 and income ≤ 49.500 and age ≤ 53 and age > 31 and employ > 1.500 then ***High Attractive***
- if income > 42.937 and income ≤ 49.500 and age ≤ 53 and age > 31 and employ > 1.5 then ***High Attractive***

- if income > 42.937 and income ≤ 49.500 and age ≤ 53 and age > 31 and employ ≤ 1.500 and income > 47 then **High Attractive**
- if income > 42.937 and income ≤ 49.500 and age > 55 then **Attractive**
- if income ≤ 42.937 and age > 44.500 and address > 11.500 and age > 46 then **Attractive**
- if income ≤ 42.937 and age > 44.500 and address > 11.500 and age ≤ 46 and gender = 0 then **Attractive**
- if income > 42.937 and income ≤ 49.500 and age ≤ 53 and age ≤ 31 and employ ≤ 10 then **Normal**
- if income ≤ 42.937 and age > 44.500 and address ≤ 11.500 and age ≤ 55 and income ≤ 37.500 and address > 5.500 and age > 47 and employ ≤ 4.500 then **Normal**
- if income ≤ 42.937 and age > 45 and address ≤ 11.500 and age ≤ 55 and income ≤ 37.500 and address ≤ 5.500 then **Normal**

The fourth rule states that if "income" for a customer is between 42.937 and 49.5 and age of the customer is above 55, then this customer is likely to be Attractive.

Churn prediction

Customer churn prediction is conducted in two phases using neural networks and decision trees algorithms. In the first phase, the customer churn prediction is accomplished at the time of the arrival of the customers to the company and based on their socio-demographic features. In the second phase, the customer churn prediction is accomplished after a month and based on their behavioral features. Rule-based algorithm is used for modeling.

Primary churn prediction

- Neural networks

At this stage, a primary churn prediction of the customer is provided using Neural networks algorithm based on demographic variables at the time of their arrival to the company. One of the best results is presented in Table 14. In this case, the neural network does not have hidden layer and the training cycle equals to 700, the momentum equals to 0.1, the learning rate is equal to 0.4.

It can be concluded from Table 14 that the accuracy, precision and recall of this model are 67.60%, 76.15% and 80.19%, respectively.

- Decision trees

In this section, a decision tree algorithm is used to implement primary churn prediction of the customers.

The results are presented in Table 15. It can be concluded from Table 15 that the accuracy, precision and recall of this model are 69.00%, 73.20% and 89.90%, respectively.

Several rules extracted from this decision tree are presented as follows:

- if age < 29.800 and address < 11.800 and income < 57 and employ < 9.400 and region = 1 and marital = 0 and education = 1 and retire = 0 and gender = 1 and pager = 0 then **Non-Churner**
- if age = [29- 42] and address < 11.800 and income < 57 and employ = < 9 and region = 1 and education = 1 and marital = 0 then **Non-Churner**
- if age = [42 - 53] and address < 11.800 and income < 57 and employ < 9.400 and education = 1 and gender = 0 then **Non-Churner**
- if age = range4 [53 - 65] and employ < 9.4 and address < 11.8 and education = 1 then **Non-Churner**
- if age > 62 and income < 57 and education = 1 then **Non-Churner**
- if age < 29.800 and address < 11.800 and income < 57 and employ < 9.400 and region = 1 and marital = 0 and education = 2 and pager = 1 then **Churner**
- if age = [29 - 42] and address < 11.800 and income < 57 and employ < 9.400 and region = 1 and education = 1 and marital = 1 and retire = 0 and gender = 0 and pager = 0 then **Churner**
- if age = [42 - 53] and address < 11.800 and income < 57 and employ < 9.400 and education = 1 and gender = 1 and region = 2 then **Churner**

The fifth rule states that if "age" of the customer is greater than 62, "income" for a customer is less than 55 and customer has a diploma degree then this customer is likely to be Non-Churner.

Secondary churn prediction   Secondary churn prediction model is presented after a month and based on customers' behavioral features.

- Neural networks

In this section, a decision tree algorithm is used to secondary churn prediction of the customers. The results of

Table 14 The results of churn prediction based on socio-demographic features (using neural networks)

Learning method	Accuracy	Precision	Recall
Neural network	67.60%	76.15%	80.19%

Table 15 The results of churn prediction based on socio-demographic features (using decision trees)

Learning method	Accuracy	Precision	Recall
decision tree	69.00%	73.20%	89.90%

modeling by neural networks algorithm are presented in Table 16.

In this result, the training cycle equals to 700, the momentum equals to 0.2, the learning rate is equal to 0.3, and the number of hidden layers is equals 2.

It can be concluded from Table 16 that the accuracy, precision and recall of this model are 75.61%, 81.86% and 86.45%, respectively.

- Decision trees

In this section, a decision tree algorithm is used to secondary churn prediction of the customers. The results are presented in Table 17.

It can be concluded from Table 17 that the accuracy, precision and recall of this model are 74.22%, 78.30% and 88.89%, respectively.

In this case, Information gain ratio is considered as splitting criterion. Minimal gain is equal to 0.1, Minimal size for split is equal to 14, and Minimal leaf size is equal to 11.

Several rules extracted from this decision tree are presented as follows:

- if longmon = [0–20.710] and tollmon = [0–34.600] and equipmon = [0–15.540] and cardmon = [0–21.850] and wiremon = [0–22.390] and callcard = 0 and wireless = 0 and multline = 0 and voice = 0 and pager = 0 and internet = 0 and callid = 0 and callwait = 0 and forward = 0 and confer = 0 then ***Non-Churner***
- if longmon = [0–20.710] and tollmon = [0–34.600] and equipmon = [$-\infty$–15.540] and cardmon = [65.550–87.400] then ***Non-Churner***
- if longmon = [0–20.710] and tollmon = [0–34.600] and equipmon = [15.540–31.080] and cardmon = [21.850–43.700] and internet = 0 then ***Non-Churner***
- if longmon = [0–20.710] and tollmon = [0–34.600] and equipmon = [15.540–31.080] and cardmon = [21.850–43.700] and internet = 1 and multline = 0 then ***Churner***

Table 16 The results of churn prediction based on behavioral features (using neural networks)

Learning method	Accuracy	Precision	Recall
Neural network	75.61%	81.86%	86.45%

Table 17 The results of churn prediction based on behavioral features (using decision trees)

Learning method	Accuracy	Precision	Recall
Decision Tree	74.22%	78.30%	88.89%

- if longmon = [20.710–40.520] and tollmon = [0–34.600] and equipmon = [31.080–46.620] and callcard = 0 then ***Churner***
- if longmon = [20.710–40.520] and tollmon = [34.600–69.200] and equipmon = [46.620–62.160] then ***Churner***

Proposed customer relationship management tactics

In this subsection, CRM tactics are proposed based on different customer attractiveness and churn. Attracting customers is the first step in customer relationship management. The first goal of customer attraction is to select the "right" prospects. The "right" prospects are persons who are more likely to be profitable in the future. Companies have commonly attracted new customers with advertising, sales promotion, buzz and social media. Applying appropriate tactics in different periods of customer life cycle is an important issue.

It can be concluded from Fig. 5 that the major tactics in relation to high attractive customers who are reluctant to churn is the development of relationships. Customer development is the process of growing the value of retained customers.

A major action in order to develop customer relationships cross /up-selling. The attractive customers who are unlikely to churn can be converted into highly attractive customers using suitable CRM tactics.

A retention tactics concentrates on "Highly attractive", "Attractive" and "Normal" customers who use the company's services, but are likely to churn in the near future. The goal of this tactic is to strengthen and retain their relationships.

A win-back campaign concentrates on churner customers that no longer produce turnover for the company that are placed in highly attractive and attractive level. The aim of this tactic is to get the customer back. Low attractive and unattractive churners that are not strategically important for company, can be ignored or be invested less to be retained. Delighting customers and adding customer perceived value are the two suggested tactics for retention customers and prevent them to switch toward competitors.

Companies can create additional value for their customers by establishing customer clubs, providing loyalty schemes and providing sales promotions. Moreover, customizing services and offering latest services and technologies are good strategies for increasing customer perceived value. Delighting customer can be also performed by providing high level of services, improving service quality and creating price stability.
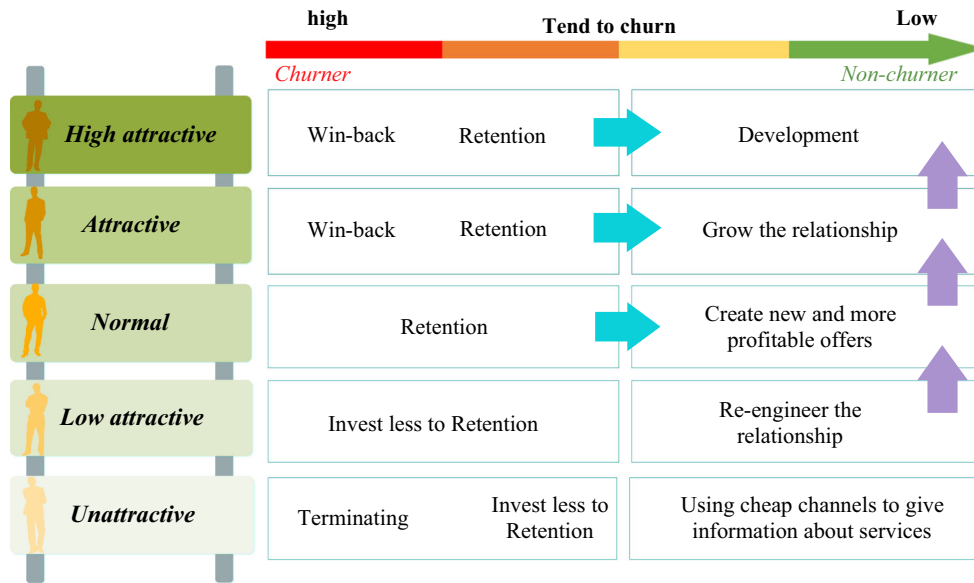
Fig. 5 Proposed tactics for each group of customers

As it can be seen in Fig. 5, some retention tactics are presented to decrease the propensity to churn. According to Fig. 5, normal customers who are willing to churn may migrate up one level by creating new and more profitable offers. Unattractive customers can be provided with information about various types of services using cheap channels such as sending an email or making a call.

Quantitative evaluation of the proposed framework

For a quantitative evaluation of the proposed framework in the paper, key performance indicators for customer retention programs, such as churn rate, can be used and check this out whether there is any significant difference between the average churn rate of customers of the company before using the framework, with the average churn rate of customers after using the framework?

The quantitative evaluation process of the impact of the proposed framework on the implementation of the customer relationship management system is considered as follows.

1. Firstly, the churn rate is calculated in the telecommunication company.

    Assume that this company has N number of customers at the beginning of the month. At the end of the month, it lost a number of its customers. The churn rate is calculated as Eq. (9)

$$\text{Churn rate} = \frac{a \text{ lost customers}}{N \text{ initial customers}} = \left(\frac{a}{N}\right) \times 100 \tag{9}$$

2. Compare the average of churn rate in the year in which the framework has been used with years that have not been used.

So, assume that the churn rate of customers is calculated at the end of each month of the year. Consider Table 18.

As can be seen in Table 18, in year A, the proposed framework has not been used and in year B this framework has been used. $a_1, a_2, \ldots, a_{12}$ are the churn rate of customers at the end of each month of year A, and $b_1, b_2, \ldots, b_{12}$ are the churn rate of customers at the end of each month of year B.

In order to determine whether there is any significant difference between the average churn rate before applying the framework and the average churn rate after

Table 18 Calculated churn rate for two years considered

Churn rate per month of year A (before using the framework)	Churn rate per month of year B (After using the framework)
a_1	b_1
a_2	b_2
a_3	b_3
a_4	b_4
a_5	b_5
a_6	b_6
a_7	b_7
a_8	b_8
a_9	b_9
a_{10}	b_{10}
a_{11}	b_{11}
a_{12}	b_{12}

applying the framework, Independent t test in SPSS can be used, and the test hypothesis can be considered as follows:

$$H_0 : \mu_{\text{YearA}} = \mu_{\text{YearB}}$$
$$H_1 : \mu_{\text{YearA}} \neq \mu_{\text{YearB}} \qquad (10)$$

If the sig value in the output table of Independent t test is lower than Type I error, H_0 hypothesis is rejected, meaning that there is a significant relationship between the means.

In order to determine which group has a lower mean, the upper bound and the lower bound of the obtained confidence interval can be used as the following:

1. If the upper bound and the lower bound are both positive, the average of the first group is larger than the average of the second group.
2. If the upper bound and the lower bound are both negative, the average of the second group is larger than the average of the first group.
3. If one of the upper bound and the lower bound is positive and the other is negative, the means of the two groups are not significantly different.

It is expected that with the continuous use of the proposed framework, churn rate will be improved every year. Therefore, in order to compare the average churn rate in more than 2 years, one-way analysis of variance (ANOVA) in SPSS can be used.

According to Table 19, consider the churn rate per month of K different years. The one-way analysis of variance (ANOVA) is used to determine whether there are any statistically significant differences between the means of two or more independent (unrelated) groups.

So, the tests hypothesis can be considered as follows:

$$H_0 : \mu_1 = \mu_2 = \cdots = \mu_K \qquad (11)$$

H_1: At least between two groups, the averages are different.

Also in the output of this test, through comparison of the value of sig and Type I error, it can be determined whether there are significant differences between the averages of the groups. If the sig value is lower than Type I error, it means that there is a significant difference between the average churn rates in different years. Tests such as LSD, Duncan, Tukey, etc., can be sued for specifying that which years this difference is related to.

Conclusion

The growth of the telecommunications industry and providing various services by telecom companies will increase the probability of losing valuable customers. Rapid growth of information technology in various businesses including the telecom industry yields to the generation of large databases. Suitable analysis can be used to analyze the behavior of customers. This will lead to develop customer relationships systems successfully in order to satisfy, attract and retain customers.

Companies require analysis of customer behavior to survive in the competitive global market, which will help them retain customers and build a long-term relationship.

Companies are seeking to create long-term relationships with more profitable customers, and therefore, one of the challenges facing organizations is the ability to predict churn rate of customers. One solution to determine the valuable customers is customer segmentation. Customer clustering and analyzing behavioral patterns of each cluster

Table 19 Calculated churn rate for K years considered

Churn rate per month of year A (before using the framework)	Churn rate per month of the first year using the framework	Churn rate per month of the second year using the framework	…	Churn rate per month of the K_{th} year using the framework
a_1	b_{11}	b_{21}	…	b_{k1}
a_2	b_{12}	b_{22}	…	b_{k2}
a_3	b_{13}	b_{23}	…	b_{k3}
a_4	b_{14}	b_{24}	…	b_{k4}
a_5	b_{15}	b_{25}	…	b_{k5}
a_6	b_{16}	b_{26}	…	b_{k6}
a_7	b_{17}	b_{27}	…	b_{k7}
a_8	b_{18}	b_{28}	…	b_{k8}
a_9	b_{19}	b_{29}	…	b_{k9}
a_{10}	b_{110}	b_{210}	…	b_{k10}
a_{11}	b_{111}	b_{211}	…	b_{k11}
a_{12}	b_{112}	b_{212}	…	b_{k12}

can be important steps in implementing customer relationship management systems.

Data mining techniques can be effectively used to extract hidden information and knowledge in customers' data. Managers can utilize this knowledge in the process of decision making.

In this paper, data mining techniques were proposed to develop a two-stage framework containing portfolio analysis and customer churn prediction. In order to implement portfolio analysis, firstly k-means algorithm was used to conduct customer segmentation. Customers were divided into six groups based on their socio-demographic features. Afterward, the groups of the customers were analyzed on the basis of two attractiveness criteria. Finally, the results indicated three different levels of customers based on their attractiveness.

In the second phase, the classification methods were used in order to predict the level of future customer attractiveness based on their socio-demographic features. Also customer churn prediction was conducted. A last, several CRM tactics were developed by taking into account both customer attractiveness and customer churn behavior.

Application of this framework allows the companies to provide the possibility of analyzing the behavior of past, current and future customers. The proposed framework of this paper also helps managers in successful implementation of CRM systems.

References

Backiel A, Baesens B, Claeskens G (2016) Predicting time-to-churn of prepaid mobile telephone customers using social network analysis. J Oper Res Soc 67:1135–1145

Bose I, Chen X (2015) Detecting the migration of mobile service customers using fuzzy clustering. Inform Manag 52:227–238

Buttle F, Maklan S (2015) Customer relationship management: concept and technology, 3rd edn. Routledge, New York

Chen Z-Y, Fan Z-P, Sun M (2012) A hierarchical multiple kernel support vector machine for customer churn prediction using longitudinal behavioral data. Eur J Oper Res 223:461–472

Chuang Y-F, Chia S-H, Wong JY (2013) Customer value assessment of pharmaceutical marketing in Taiwan. Ind Manage Data Syst 113(9):1315–1333

Dursun A, Caber M (2016) Using data mining techniques for profiling profitable hotel customers: an application of RFM analysis. Tour Manag Perspect 18:153–160

Dzobo O, Alvehag K, Gaunt CT, Herman R (2014) Multi-dimensional customer segmentation model for power system reliability-worth analysis. Electr Power Energy Syst 62:532–539

Farquad MAH, Ravi V, BapiRaju S (2014) Churn prediction using comprehensible support vector machine: an analytical CRM application. Appl Soft Comput 19:31–40

Fathian M, Hoseinpoor Y, Minaei-Bidgoli B (2016) Offering a hybrid approach of data mining to predict the customer churn based on bagging and boosting methods. Kybernetes 45(5):732–743

Fiocca R (1982) Account portfolio analysis for strategy development. Ind Market Manag 11:53–62

Güçdemir H, Selim H (2015) Integrating multi-criteria decision making and clustering for business customer segmentation. Ind Manag Data Syst 115(6):1022–1040

Guelman L, Guillén M, Pérez-Marín AM (2012) Random forests for uplift modeling: an insurance customer retention case. In: Engemann KJ, Gil-Lafuente AM, Merigó JM (eds) Modeling and simulation in engineering, economics and management. Lecture notes in business information processing, vol 115. Springer, Berlin pp 123–133

Han J, Kamber M (2012) Data mining: concepts and techniques, 3rd edn. Morgan Kaufmann Publishers, Burlington

Hsu F-M, Lu L-P, Lin C-M (2012) Segmenting customers by transaction data with concept hierarchy. Expert SystAppl 39:6221–6228

Huang Y, Kechadi T (2013) An effective hybrid learning system for telecommunication churn prediction. Expert Syst Appl 40:5635–5647

Huang B, Kechadi T, Buckley B (2012) Customer churn prediction in telecommunications. Expert Syst Appl 39:1414–1425

Keramati A, Jafari-Marandi R, Aliannejadi M, Ahmadian I, Mozzafari M, Abbasi U (2014) Improved churn prediction in telecommunication industry using data mining techniques. Appl Soft Comput 24:994–1012

Kim K, Jun C-H, Lee J (2014) Improved churn prediction in telecommunication industry by analyzing A large network. Expert Syst Appl 41:6575–6584

Kittidecha C, Yamada K (2018) Application of Kansei engineering and data mining in the Thai ceramic manufacturing. J Ind Eng Int. https://doi.org/10.1007/s40092-018-0253-y

Lin W-C, Tsai C-F, Ke S-W (2014) Dimensionality and data reduction in telecom churn prediction. Kybernetes 43(5):737–749

Markopoulos AP, Georgiopoulos S, Manolakos DE (2016) On the use of back propagation and radial basis function neural networks in surface roughness prediction. J Ind Eng Int 12:389–400

Mehmanpazir F, Asadi S (2017) Development of an evolutionary fuzzy expert system for estimating future behavior of stock price. J Ind Eng Int 13:29–46

Mitra S, Pedrycz W, Barman B (2010) Shadowed c-means: integrating fuzzy and rough clustering. Pattern Recognit 43:1282–1291

Moeyersoms J, Martens D (2015) Including high-cardinality attributes in predictive models: a case study in churn prediction in the energy sector. Decis Support Syst 72:72–81

Mohammadi Nasrabadi A, Hosseinpour MH, Ebrahimnejad S (2013) Strategy-aligned fuzzy approach for market segment evaluation and selection: a modular decision support system by dynamic network process (DNP). J Ind Eng Int 9:1–17

QuelhasBrito P, Soares C, Almeida S, Monte A, Byvoet M (2015) Customer segmentation in a large database of an online customized fashion business. Robot Comp-Int Manuf 36:93–100

Rangan VK, Moriarty RT, Swartz GS (1992) Segmenting customers in mature industrial markets. J Market 56:72–82

Rezaeinia SM, Rahmani R (2016) Recommender system based on customer segmentation (RSCS). Kybernetes 45(6):946–961

Ritter T, Andersen H (2014) A relationship strategy perspective on relationship portfolios: linking customer profitability, commitment, and growth potential to relationship strategy. Ind Market Manag 43(6):1005–1011

Ryals L (2002) Are your customers worth more than money? J Retail Consum Serv 9:241–251

Ryals L (2010) Making customers pay: measuring and managing customer risk and returns. J Strategic Market 11(3):165–175

Safari F, Safari N, Montazer GA (2016) Customer lifetime value determination based on RFM model. Market Intell Plan 34(4):446–461

Shapiro BP, Rangan VK, Moriarty RT, Ross EB (1987) Manage customers for profits (not just sales). Harvard Bus Rev 65(5):101–108

Thakur R, Workman L (2016) Customer portfolio management (CPM) for improved customer relationship management (CRM): are your customers platinum, gold, silver, or bronze? J Bus Res 69(10):4095–4102

Verbeke W, Martens D, Mues C, Baesens B (2011) Building comprehensible customer churn prediction models with advanced rule induction techniques. Expert Syst Appl 38:2354–2364

Verbeke W, Martens D, Baesens B (2014) Social network analysis for customer churn prediction. Appl Soft Comput 14:431–446

Wang G, Ma J (2012) A hybrid ensemble approach for enterprise credit risk assessment based on Support Vector Machine. Expert Syst Appl 39(5):5325–5331

Wang B, Miao Y, Zhao H, Jin J, Chen Y (2016) A biclustering-based method for market segmentation using customer pain points. Eng Appl Artif Intell 47:101–109

Wei J-T, Lee M-C, Chen H-K, Wu H-H (2013) Customer relationship management in the hairdressing industry: an application of data mining techniques. Expert Syst Appl 40:7513–7518

Weng C-H (2016) Knowledge discovery of digital library subscription by RFC item sets. The Electron Libr 34(5):772–788