# Customer Churn Prediction Using Local Linear Model Tree for Iranian Telecommunication Companies

**Mehdi Fasanghari [1] and Abbas Keramati [1]**
[1] Department of Industrial Engineering, College of Engineering, University of Tehran, Tehran, Iran
(Received 16 November 2010, Accepted 11 July 2011)

## Abstract

For winning in global competition, companies need to recognition and monitoring of customer's behavior to forecast their behavior and desires earlier than competitors. This research tries to recognize the attributes which lead to customer churn. For this, behavior of 3150 subscribers of an Iranian mobile operator, has observed during one year and trends of them has analyzed by a customized LLNF model. For this purpose, the application of the locally linear model tree (LOLIMOT) algorithm, which integrates the advantage of neural networks, tree model and fuzzy modeling, was experimented.
Results suggest that dissatisfaction of customer, his/her usage from services and demographic attributes have significant effect on decision to churn or retention. Furthermore, the active or inactive subscriber situation has mediation effect on his/her retention.

**Keywords**: LLNF, Customer churn, LOLIMOT, Fuzzy Logic, Neural network, Prediction, Mobile service provider

## Introduction

In age of communication, it is natural that telecom industry possesses one of the highest growth indicators among other industries. Besides, among different telecom industries, one of the fastest rates of growth belongs to the mobile service industry that its contribution in everyday communication of people is extraordinary increasing and growing on the verge of overtaking fixed phone. This growth speed is not only owes innovation in Mobile Communication technology, but also owes intense competition and ruthless operators under the minimum regulations governing the market of this industry. Mobile service market along with the growth of its industry shows the remarkable growth that cannot be due only to increase in the number of subscribers and the increasing diversity of new services and ignore competitive arena of them. This surprising and integrated development of industry and market caused the operators despite being a young industry, trying to get new customers and start their marketing blitz gradually and attract customers of other operators [1, 2]. Obtaining new customers is much harder and more expensive than keeping existed customers, and according to the study has been done in 2004, the cost is equivalent to 300 dollars versus 20 dollars in 1995 and 300 dollars versus 25 dollars in 2004 [3]. Part of this fact is due to the suppliers that have significant and valuable information about their customers and analyze it to understand their behavior and preferences. Additionally, in a developed market, attracting new customers needs to ward off them from other operators and having a more attractive stimulus.

Keeping old customers, especially in-service markets such as mobile cell, except the cost of attracting customers, have the opportunity value for mobile operators. It means that the provider can provide additional and new services to customer and earn more money. For this reason, the loss of existing customers not only reduces revenues and imposes costs of attracting new customers, but also leads to the loss of potential revenues.

Therefore, with the growth of telecom companies in the global market and maturation of these services, customer churns management has become a

* Corresponding author: Tel: +98- 21- 82081067 Fax: +98- 21- 88013102 Email: keramati@ut.ac.ir

fundamental concern of these companies. Currently, the churn rate of customers in such companies in Europe, North America and East Asia has been reported in a range of 20% to 40% [1, 2, 4]. This high rate of customer churn is in an industry that its cost to obtain a new customer is five to 10 times more than maintaining an existing one [5]. Calculated values in the mobile service market in America are equivalent to 300 dollars versus 25 dollars [6].

Living in the global village and taking the path of fast globalization has made communication development, and especially mobile services, for developing countries such as Iran necessary. Providing mobile services in Iran began in 1991 by Telecommunication Company of Iran as the first operator. Telecommunication Company was a 12-year monopolized in this industry, until entering Taliya Company as the first private mobile service provider in June 2003 with the first series of MMC card, and ended the monopolized market. Irancell Company started its services as the second operator of mobile services in Iran, with the participation of MTN Company of South Africa (the largest mobile operator in Africa and the Middle East now). So far, 15 foreign companies from Canada, Great Britain, Russia, Malaysia, China and the UAE stated that they would have been ready to participate in the development of the third mobile operator, according to the program to cover 17 million users. Now considering a competition in similar global markets, as well as population growth and the growing needs of mobile communication, strong competition among operators to increase their penetration and also increasing their contribution in the Iranian market can be predicted.

Imparting the general policies of Article 44 of the constitution which quickly accelerated privatization in Iran, Iranian mobile service market will experience stronger competition, particularly in the field of MMC card. The first operator is on the verge of privatization, and after that management ability, true policies and timely marketing make survival of this great organization possible in the absence of strong government institutions.

In this strong competition, the winner will be someone who protects her customers as valuable assets while trying to attract new customers, and use scientific methods and useful tools in this way.

Considering challenges and thriving future of this market, a case study research on one of the active companies in mobile services in Iran has been done, with the help of survival analysis tools to understand features that lead to customer churn.

## Literature Review

Although according to some researchers, studies carried out in customer churn have not been focused and organized so far [1], but they can generally be divided in two categories:

1. Articles and researches had been done to show what data mining can do in customer relationship management (which henceforth will be mentioned briefly by CRM) and provide reasons to prove fitness of data mining tools to analyze CRM processes and also articles proposed in order to introduce specific data mining methods for solving problems facing the CRM in the organization.
2. Field researches that analysis customer churn with respect to customer behavior in different organizations.

### A. *Importance of data mining in Customer Relationship Management*

In this article, CRM is defined as a multiple layer concept that data mining is one of its layers. In his study, considering the results of data mining analysis about customers is required for an efficient CRM system. Also two examples of data mining techniques were introduced and their applications in two case studies of CRM were described accurately. Valletti and Cave, in 1998, analyzed expanded competition in the mobile industry from the perspective of different operators' strategies [7]. Fullerton also followed the same goal in his research about the telecommunications market in America, studying the relationship between market structure and market performance and has been shown that how understanding the market structure can help predicting market performance and reduce the customer churn to its minimum.

### B. *Analysing the behaviour of customers and their churn*

The second category includes those researches that analyze issues related to customer churn in mobile phone services and have been highly regarded in recent years. Song and Kim, in 2001, evaluated the effect of change in mobile market structure of the Korean on customer churn, utilizing the simulation approach [8]. Choi, Lee and Chung (2001) also examined the effect of business strategies of five major mobile service provider companies in Korea's on customer loyalty [9]. Kim and Kwon in 2003 began to study factors that customers consider when choosing their mobile operator [10]. The research results indicate the effect of discount of intra network calls and the quality of communication on selecting the operator. In 2004, Kim and Ion trace on 973 users of five major mobile service operators in South Korea and identified subscriber churn and loyalty features [2]. Their study has shown that the probability of modifying operator by a customer depends on a satisfaction level related to his operator service features, including conversation quality, tariff levels, the device provided by operator and brand reputation. Factors such as conversation quality, type of device offered by the operator and brand reputation also impress the customer loyalty, and this loyalty is measured by a customer intend to recommend the operator to the others. Not existing significant relationship between subscription duration and customer loyalty, indicated "lock-in" effect among customers that they can be named as a sham loyal, those who don't end their subscription only because of avoiding additional costs.

The survey research of Gerpott, Rams and Schindler, in 2001, hold on the mobile market in Germany showed that maintaining customer, his loyalty and satisfaction are correlated [11].

In their study, structural differences in three concepts of customer maintenance, customer satisfaction and customer loyalty and also their correlation had been investigated. Their analysis is based on a two-stage model in which overall customer satisfaction significantly affects the loyalty, and the loyalty affects his intent to leave or stay in touch with the operator. Price of mobile services, perception of benefit from services and portability of numbers between different operators were considered as variables related to the provider that has a lasting impact on customer churn.

In 2006, the telecommunication industry's research was published in South Korea [12]. This study, entitled " *customer churn Analysis: churn factors and the effect of transition stage churn on Korean Mobile Communication Industry*" by Ahn, Hana, and Lee that studied churn factors based on data related to transactions and payment of subscribers and claimed that cost of modification and service consuming also affect the decision of leaving or sustain. Most of the previous studies focus on the direct impact of independent variables on customer churn while this study states that status of customer is as intermediaries between the customers churn signs and complete churn. Another study, by Seo, Ranganathan, and Babad was published in 2008 that focuses on understanding the factors that led to the customer churn [6]. To understand these issues two basic questions come in mind: 1- How the customer satisfaction and cost of modification factors, such as subscription time, complexity of services and quality of communication, help customers sustain and 2- How a customer demographic characteristics such as age and sex affect their choices and lead to differences in their approach of leave or sustain.

Customer churn analysis in the mobile market of Iran is a time consuming and complex problem which has many criteria for consideration. According to the literature review, the statistical analysis methods are useful for statistical analysis of the customer churn data. Thus, it is recommended in this paper to propose a customized neural network to come across the customer churn complexity in the mobile market of Iran as a new contribution.

One of the restrictions in this paper is the limitation for access to all the customer churn parameters in the mobile market of Iran. Most of the parameters are not available in this industry, and some of the parameters are not clear enough for considering in customer churn analysis. Therefore, it is necessary to constitute the customer churn criteria framework in the mobile market of Iran according to the available information.

## Research methodology

Customer churn prediction in Iranian mobile operators company is done in this paper according to the proposed methodology in Figure 1. As presented in this methodology, the data should be selected from a database which is belonged to a mobile operators company in Iran, and the dataset should be large enough and have the mentioned criteria of the proposed customer churn framework as the attribute of each customer data in dataset.

The dataset was picked up from the database of a mobile service provider in Iran. According to the variables which can be extracted from the selected dataset, and as for the developed customer churn framework in this research, randomly, 5000 records are extracted from the available dataset, in which 3150 records were prepared for use in this research as the other records was not complete in all the attributes assigned for each record.

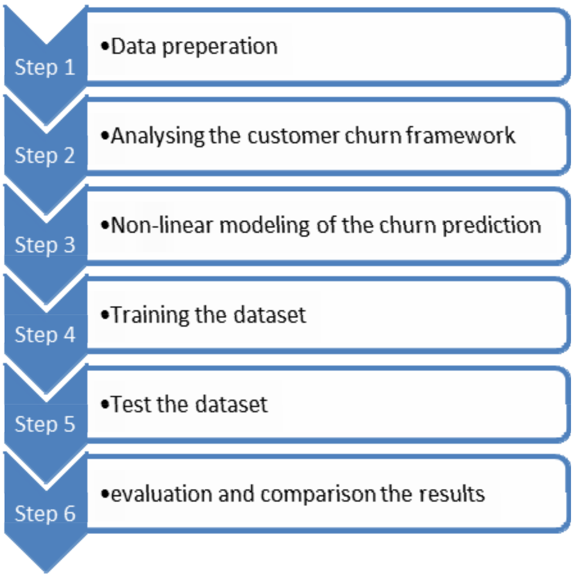


**Figure 1: The proposed research methodology**

We customized a local linear neural network for modeling the customer churn prediction with LOLIMOT learning method. 80% (2520 records) of data was assigned for the training process of the developed neural network, and the rest (630 records) was assigned to the test process. Then, for evaluating the developed method, the multi-layer perceptron was selected for comparison. Finally, the obtained results were comprised with the most popular evaluation criteria (RMSE and $R^2$).

### C. *Analysis of Customer Churn Framework*

The usefulness of a theory depends on appropriate replication, development, and generalization to create new insights and add it to existing knowledge collections [13]. Reviewing the literature and existing models in customer churn, the model that was proposed by Ahn and her colleagues (2006) was the most comprehensive model that has been studied and was chosen as a foundation for developing the framework of this study because in addition to involving all factors that had been studied in other researches [12], the effect of some new churn factors is also examined. In addition to measuring the direct effect of these factors on customer churn, the effect of customer status is also investigated and illustrated in **Figure 2**.

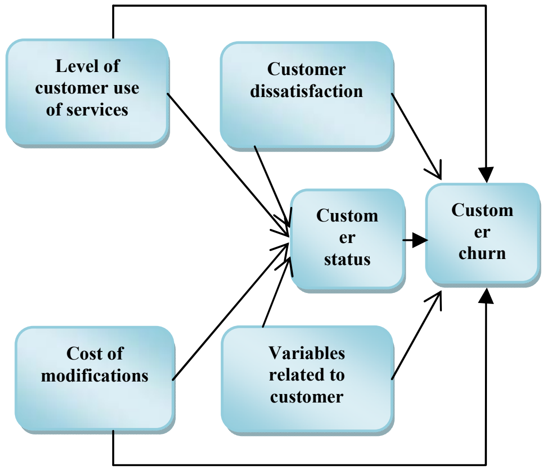


**Figure 2: A conceptual model of customer churns based on Ahn and her colleagues [12]**

According to this model, features that may have a relationship with customer churn in an organization have been divided into four main categories:

1. Customer dissatisfaction
2. Level of customer use of services
3. Cost of modifications that faces customer churn
4. Customer demographic characteristics

In all reviewed studies about features that lead to customer churn, customer dissatisfaction is always considered as one of the most important options, and the possibility of changing the operator depends on the level of customer satisfaction from her operator service features including communication quality, tariff levels and reputation of the brand [2]. Different definitions of customer satisfaction are presented by marketing theorists. Cutler defined customer satisfaction as the degree that actual performance of a company meets customer expectations. In Cutler opinion if company performance meets customer expectations, customer feels happy, otherwise she feels unhappy.

Another definition proposed by Burly, Martin and Quintana state the customer satisfaction is a result of the comparison of expected performance before buying with actual performance customer percepts [14]. Studies that are based on data collected directly from customers, consider the customer satisfaction as an experience oriented concept that represent satisfaction of customer expectations from individual or total service functions has been purchased. In these studies, higher quality of network, low price, customer care and more personal profits gained from mobile services provides more customer satisfaction [11].

In studies that based their analysis on real data instead of questionnaires, network quality and conversation quality considered as main factors of satisfaction, and variables like the number of disconnected calls, the number of abortive calls and the number of customers complaints have been considered for its measurement [12]. Moreover, the quality of contact and the subscription life are recognized as the most important factors in the customer satisfaction attractions [6].

As this research is based on data mining method, interpretation of customer behavior is possible only from existing data in the service provider database, and customer assessment or perception of operator services cannot directly be extracted or available to use in research. Accessible data from mobile subscriber database fit in two different groups: data related to contract and data related to details of conversations [5].

Thus, among the variables examined in previous researches, it is possible to extract variables that place in those two categories. Variables to measure customer dissatisfaction were the rate of disconnected calls (call that has been disconnected in the middle of conversation), the rate of abortive calls, the number of customer complaints (any customer contact to operator for removing defects and problems), and duration of customer relationship with the operator.

Among these variables with regard to the information existed in the operator call center as well as limitations in providing data, three variables, number of abortive calls, customer relationship duration and customer complaints were extracted to measure customer satisfaction. Since in the case studied 97% of plaintiff subscribers had complained only once, the variable of a number of customer complaints changed to the variable of having customer complaints. This variable is set to zero if a customer did not record any complain, and it is set to one if at least one complain is recorded in the operator call center database. Hypotheses related to customer dissatisfaction are defined as follows:

1. number of abortive calls have positive effect on customer churn
2. duration of subscriber relationship with the operator have negative effect on his churn
3. subscriber complaints from operator have positive effect on his churn

To measure this trait, variable the amount of monthly bills, the amount of deferred payments, number of unpaid bills, duration of monthly calls, number of calls per month and number of different phone numbers that customer calls them in each month were extracted from previous researches [2] [11] [12] [6].

In addition, the variable number of SMS sent by the operator was added and variable the amount of monthly bills varies to the amount of charge per month due to credit services provided by the operator. Hypotheses about the level of customer use of services are defined as follows:

1. The amount of credit charge by a subscriber has negative effect on his churn.
2. Duration of conversations conducted by a subscriber has negative effect on his churn.
3. Number of calls taken by a subscriber has negative effect on his churn.
4. Number of SMSs sent by a subscriber has negative effect on his churn.
5. Total number of different phone numbers that subscriber called them, has negative effect on his churn.

The cost of shirk the supplier can prevent the customer churn. This cost can be due to the call number change or buying a phone device supporting the new network standards; although, this cost is changes according to the telecommunication service providers strategy [2]. Cost modification and customer satisfaction are related. The high cost of modification can dissuade the somewhat dissatisfaction customer from migration from the current operator into a new operator. However, we should have considered that the customer tolerance is limited, and he or she can move to a new operator with the low-quality services delivered from the current operator regardless of the imposed cost of the migration [6]. The cost modification is disported to three main groups:

1- The learning cost related to the difficulties that the customer should bears to achieve the convenience sense of the service of the last operator. This kind of cost is for the mobile phone set type delivered to the customer from the operator company in the mobile services market. We pass up this cost in our research because this incentive is not offered to the customers by any Iranian operator companies.

2- The cost of customer churn imposed from the operator company, which contractually is determined from the operator company as the migration cost. The mobile service market of Iran is a competitive market, and all the operator companies of Iran are delivering the same services. Thus, the risk of potential customer churn is not accepted with any of them, and they do not put this constrain in their contract with the new customers; therefore, this cost is not attractive for use in the mobile service market of Iran.

3- The transactional change cost consisted of the financial cost due to ending the relation with the current mobile service provider and switching to a new service provider. This cost is attractive in Iran since the operator companies set their tariff cost according to the loyalty of their customers and give some loyalty bonus according to the kind of service usage for their customers. This makes some economic advantages for the customers which arrest them from switching to another mobile service provider.

In the last researches, some factors related to the customer specifications were considered, in which the variables such as the age of customers, the sexuality of customers, the payment method, and the rank of customers (in case of customer ranking done with the operator company) are selected as the most popular research variables. In this research, according to the situation of the mobile service providers of Iran, the age of the customer was selected as the variable of the variables related to customer as a multi dimension variables illustrated in Table 1.

| Age group | Age interval (x) |
|---|---|
| 1 | x≤15 |
| 2 | 15<x≤30 |
| 3 | 30<x≤45 |
| 4 | 45<x≤60 |
| 5 | 65<x≤75 |

**Table 1: The customer age segmentation in this research**

### D. *Artificial neural network (ANN)*

Moreover, the main characteristic of a Neural Network (NN) consisted of many neurons is to compute and sort information. ANN procedure is designed as the human brain; accordingly, ANNs consist of many nodes, which each node can have many inputs and many outputs. Just as the chemical connections among brain neurons, a weight value is assigned to each input parameter to the each neuron reflecting the strength of its link [15]. That is the reason of great usage of ANN to solve non-linear and ill-structured problems [16]. Therefore, ANN is used frequently for forecasting nonlinear time series events [17, 18]. Each node $j$ sums its weighted input [15]:

$$net_j = \sum_{j=1}^{n} W_j X_j \,. \qquad (1)$$

A nonlinear function is used as the output of a node ($Y_j$):

$$Y_j = f\,(net_j). \qquad (2)$$

Although there are many models for ANN which are developed for application of different area, multilayer perceptron learning with feed forward back propagation (FFBP) is a famous model which is used in many time series forecasting cases [19, 20], which is introduced by Rumelhart, Hinton, and Williams [21, 22]. The parameters of FFBP model are: the number of nodes in input layer, the number of hidden layers, the number of nodes in each hidden layer, the transfer function, and the learning rate. In fact, these parameters are problems dependent which should be adjusted with try and error method to fit a suitable model for a mentioned time series event. This model consisted of tree phases: learning, test, and validation, in which the last two phases are combined to each other in most of the researches. A set of weights are obtained in learning phase and the errors of the structured learned model, as the difference of the outputs of the model and the target outputs, are calculated in test phase, which shows the best model with the lowest error in this phase. The optimization method for reaching the best objective, the minimum deference of the outputs of the model and target outputs, is steepest descent approach with the learning rate and momentum parameter using to propagating the errors back through the network to adjust the weights [23, 24].

### E. *LLNF method*

The main feature of the LLNF model is for function approximation through dividing the input space into small linear subspaces with fuzzy validity function $\phi_i(u)$ [25-27]. These functions describe the validity of each linear model in its region [27]. Each local linear subspace with its validity function is called a fuzzy neuron; therefore, LLNF is a neurofuzzy network with two layers consisted of one hidden layer and a linear (1) neuron in the output layer for computing the weighted sum of the locally linear models as[27]:

$$\hat{y}_i = \omega_{i0} + \omega_{i1}u_1 + \omega_{i2}u_2 + ... + \omega_{ip}u_p, \qquad (2)$$

$$\hat{y} = \sum_{i=1}^{M} \hat{y}_i \phi_i(u), \qquad (3)$$

where $u = \left[u_1, u_2, ..., u_p\right]^T$ is the model input, $M$ is the number of MML neurons, and $\omega_{ij}$

denotes the LLM parameters of the $i^{th}$ neuron.

According to the applied validity function, the interpolation quality is changed. As normalization is necessary for a proper interpretation of validity functions [26, 27], we applied the validity function of the normalized Gaussian function as [26, 27]:

$$\mu(x) = e^{\left(-\frac{(x-c)^2}{2\sigma^2}\right)} \tag{4}$$

where $c$ is the center and $\sigma$ is the standard deviation of the Gaussian function. Thus,

$$\phi_i(u) = \frac{\mu_i(u)}{\sum_{j=1}^{M}\mu_j(u)}, \tag{5}$$

such that

$$\begin{aligned}\mu_i(u) &= e^{\left(-\frac{1}{2}\left(\frac{(u_1-c_{i1})^2}{\sigma_{i1}^2}+\ldots+\frac{(u_p-c_{ip})^2}{\sigma_{ip}^2}\right)\right)} = \\ &e^{\left(-\frac{1}{2}\frac{(u_1-c_{i1})^2}{\sigma_{i1}^2}\right)} \times \cdots \\ &\times e^{\left(-\frac{1}{2}\frac{(u_p-c_{ip})^2}{\sigma_{ip}^2}\right)}.\end{aligned} \tag{6}$$

As illustrated in Eq. (6) and Eq. (5), the nonlinear hidden layer of the LLNF has $2Mp$ parameters ($c_{ij}, \sigma_{ij}$) as the parameters of validity functions and the rule-consequent parameters of the locally linear models ($\omega_{ij}$) which should be adjusted respectively with a learning method and optimization method. On view of this point, the least squares optimization method is used for fine-tuning the $\omega_{ij}$ and the $c_{ij}$ and $\sigma_{ij}$ are adjusted through the Locally Linear Model Tree (LOLIMOT) training method which is described below. Now, we have a complete parameter vector:

$$\begin{aligned}\omega = [&\omega_{10}, \omega_{11}, \ldots, \omega_{1p}, \omega_{20}, \omega_{21}, \\ &\ldots, \omega_{M0}, \ldots, \omega_{Mp}],\end{aligned} \tag{7}$$

And an associated regression matrix $X$ for $N$ measured data samples:

$$X = [X_1, X_2, \ldots, X_M], \tag{8}$$

where

$$X_i = \begin{bmatrix} \phi_i(u(1)) & \cdots & u_p(1)\phi_i(u(1)) \\ \phi_i(u(2)) & \cdots & u_p(1)\phi_i(u(2)) \\ \vdots & \vdots & \vdots \\ \phi_i(u(N)) & \cdots & u_p(1)\phi_i(u(N)) \end{bmatrix}. \tag{9}$$

Hence:

$$\begin{aligned}\hat{y} &= X\hat{\omega}, \\ \hat{\omega} &= \left(X^T X + \alpha I\right)^{-1} X^T y, \quad \alpha \ll 1,\end{aligned} \tag{10}$$

(4)

where $\alpha$ is a regularization parameter for avoiding any near singularity of matrix $X^T X$ [26, 27]. Utilizing the training method, which is described in next section, is depended to the data sets that should be linearly normalized to the interval $[-1,1]$. (5)

## F. Locally Linear Model Tree (LOLIMOT)

The LOLIMOT algorithm proposed by Nelles and Isermann [25, 28, 29] is based on the idea to approximate a nonlinear function with piece-wise linear model [25, 28, 30, 31]. (6)

Replacing the outputs layer weights with a linear function of the network input putting out the LOLIMOT, in which each neuron represents a local linear model with its related validity function. Partitioning the input space by axis-orthogonal splits introduced the LOLOMOT learning method as an incremental tree-construction. In each iteration, a new rule or local linear model (LLM) is added to the model. Using the local weighted least square method for optimizing the corresponding rule consequents and comparing the validity functions related to the actual partitioning of the input space is used in each iteration. Generally, two loops can be designed in LOLIMOT method as an outer loop for upper-level adjustment and inner loop as lower level tuning that (7) respectively determines the parameters for nonlinear partitioning of the input space and estimates the parameters of those local linear models [25, 26, 28, 31].

LOLIMOT can be done in five stages [25, 26, 28, 31]:

1- Starting with an initial model: A single LLM is selected so that over the input space with $\phi_1(u)=1$, it be a global linear model. Then, let $M=1$ and select the initial structure if there is a priori input space partitioning.

2- Finding the worst LLM: Find the worst performing LLM with mean square error (MSR) measure.

3- Checking all divisions: Select the worst LLM as further refinement. The hyper rectangle (more than a three-dimensional rectangle or cube) of this LLM is split into two halves with an axis orthogonal split. Divisions are tried in all dimensions and in each $p$ division.

4- Finding the best division: The best division which is checked in step 3 is selected for constructing its related validity functions and LLMs.

5- Testing the termination condition: If the termination condition ($\|y-\hat{y}\| \le \varepsilon$) is not occurred, go to step 2; else, stop the training procedures.

*G. Evaluation criteria*

Even though the basic mean square error index is used in many of the past research as a useful evaluation criterion for specifying the correct model, there are several evaluation criteria reducing the disadvantages of the basic mean square error since this measure does not consider the absolute value of the errors. Thus, the negative errors neutralize the positive errors. In this paper, we validate our prediction results of different applied models with RMSE, MAPE, and $R^2$ as bellows:

1- Root mean square error (RMSE):

$$RMSE=\sqrt{\frac{\sum\left(\tilde{y}_t-y'_t\right)^2}{n}}, \tag{11}$$

where $y'_t$ is the target (real) output value at time point $t$ and $\tilde{y}_t$ is the estimated output value through our applied method at time point $t$. As well as, $n$ is the number of predicted value through the applied model.

2- Mean absolute percentage error (MAPE):

$$MAPE=\frac{\sum\left|\frac{\tilde{y}_t-y'_t}{\tilde{y}_t}\right|}{n}. \tag{12}$$

3- The square of the Coefficient Of Correlation:

$$R^2=1-\frac{\sum\left(\tilde{y}_t-y'_t\right)^2}{\sum\tilde{y}_t^{\,2}}. \tag{13}$$

## Experimental results

The dataset is provided by one of the mobile service providers in Iran. The database contains datasets of mature subscribers who were with the company for at least one year. There are 2520 variables for calibration and 630 customer profiles for test. A total of 11 variables is included, one for churn indication, and 10 for prediction. The churn response is coded as a dummy variable with churn = 1 if the customer churns, and churn = 0 otherwise. Potential predictors are calculated based on whether the customer left the company during the 31-60 days after the customer was originally sampled. In this case, practitioners would have a one-month delay to target and retain customers before churn. A delay of one month in measuring the churn variable is reasonable as the implementation of proactive customer retention campaigns requires some time.

According to the developed customer churn framework, the hierarchical structure of the customer churn variables illustrated in Figure 3, is picked up for customer churn analysis.

Based on our research, it is obvious that churn prediction is a nonlinear classification problem. Thus, we utilize a locally linear model (11) for churn index modeling such that several local lines can be fitted for each region of the dataset. In order to obtain the maximum generalization, the appropriate number of LLMs should be selected for prediction. We compute and plot the error criteria with respect to number of LLMs. In this experiment, the appropriate number of LLMs is where the mean square error (MSE) in the validation set start to increase. To evaluate the relative performance of the different methods, we apply the LOLIMOT algorithm to the balanced calibration sample. As a benchmark, we apply the other estimation model namely: multi-layer perceptron neural networks on the same sample. Later, we apply the estimated models to the test sample in order to obtain churn predictions for each of the customers belonging to this latter sample. The output of the LOLIMOT algorithm is a number between zero and one. Furthermore, the threshold for detecting a customer as a churner is 0.5.

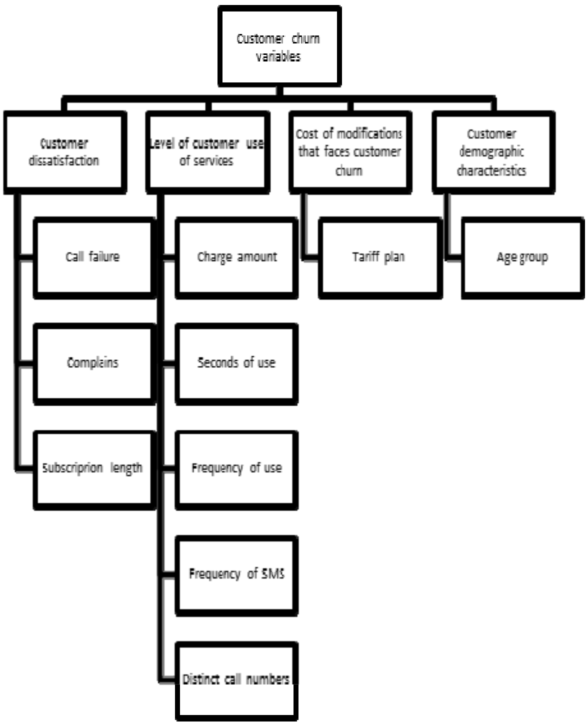


**Figure 3: The hierarchical structure of the customer churn framework**

In this paper, the proposed LLNF was run for $1 \le \alpha \le 2$ and $20 \le \text{LLN} \le 40$ illustrated in Figure 4. The point of $\alpha = 1.6$ and $\text{LLN} = 39$ has the minimum RMSE values, which indicate the best parameters that can be adjusted for the proposed LLNF in this paper.

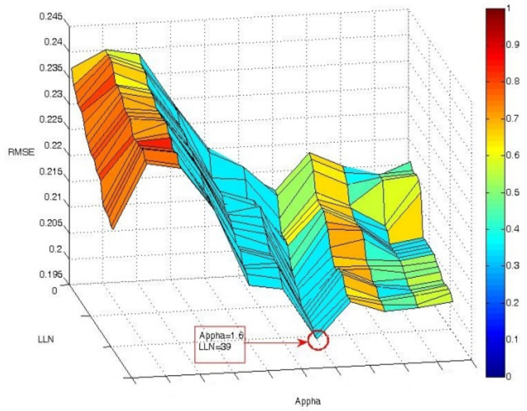


**Figure 4: The RMSE values for 1≤α≤2 versus 20≤LLN≤40**

$R^2$ is another evaluation criterion chosen in this paper for comparing the obtained results of the proposed LLNF. According to the results, the parameters of $\alpha = 1.6$ and $\text{LLN} = 39$ are the best values in view point of the $R^2$ criterion.

Thus, we set the proposed LLNF for $\alpha = 1.6$ and $\text{LLN} = 39$, and we run the model for calculating the RMSE and $R^2$ values of the training data and test data. Obviously, the evaluation criteria have the best values for the designed LLNF. Now, we can run the multi-layer perceptron neural network model and compare their results to make sense about the designed algorithm for predicting and estimating the customer churn in telecommunication industry of Iranian market.

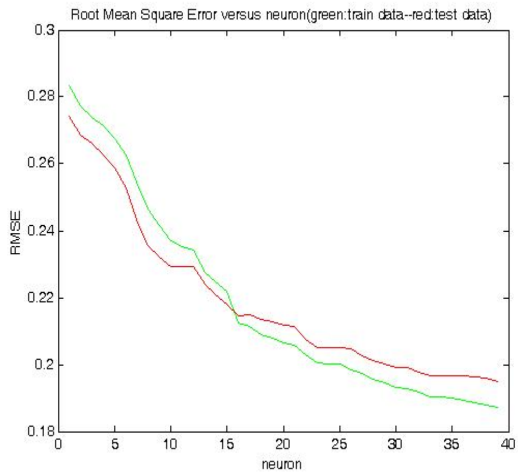


**Figure 5: The RMSE values for α=1.6 versus LLN=39**

We developed a two-layer perceptron neural network with mutation rate of 0.1 as we found these parameters the best values for the Iranian telecommunication company dataset. As illustrated in Figure 6, the 15th epoch is the best point for the neural network stop point, in which the validation data and the test data find the ascending gradient in this epoch.

To have a better view of the designed multi-layer perceptron neural network, the regression of the training data, test data, and validation data for the designed multi-layer perceptron neural network, and the $R^2$ values are shown in Figure 7.

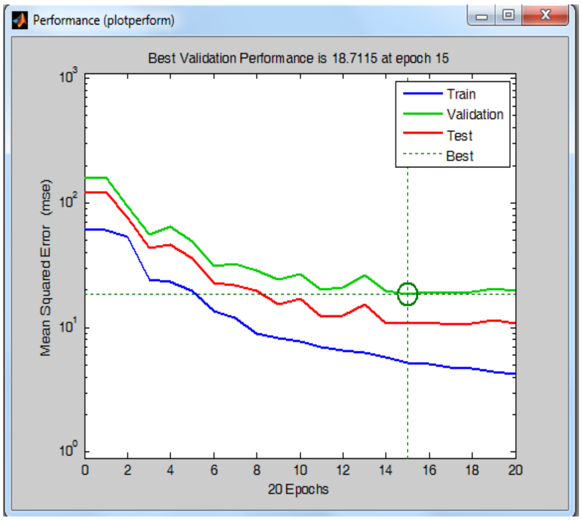


**Figure 6: The MSE criterion value for training data, test data, and validation data**

We calculate, although, the $R^2$ parameter, illustrated in Figure 8, and the RMSE parameter of the designed multi-layer perceptron neural network. Clearly, the best parameter for the multi-layer perceptron neural network is Neuron=25 and the initial point of start=20.

Obviously, the RMSE value and the $R^2$ value of the designed LLNF model are better than the designed multi-layer perceptron neural network values. Thus, the designed system is more applied that the other traditional neural network models for predicting the customer churn in Iranian mobile service provider companies.

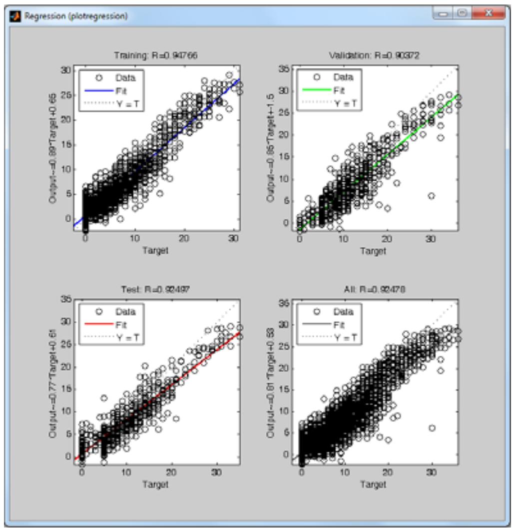


**Figure 7: The regression of the training data, test data and validation data for the designed multi-layer perceptron neural network and the $R^2$ values of it**

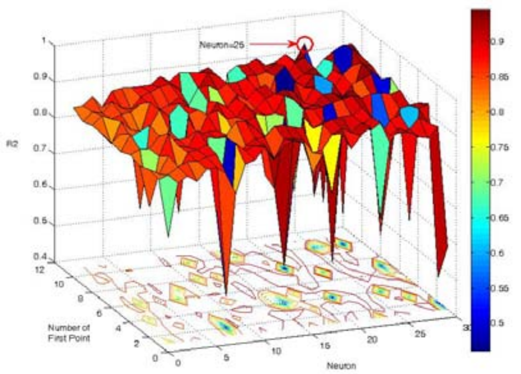


**Figure 8: The $R^2$ parameter values for the designed multi-layer perceptron neural network**

## Conclusion

Many prediction models and churn detection techniques have been presented to date. However, a more robust algorithm is required to distinguish churners from non-churners considering the extremely unbalanced dataset. In this paper, locally linear fuzzy model and the LOLIMOT learning algorithm were introduced to the marketing literature. Since the algorithm is easy to implement and all the parameters are automatically calculated, it is a practical approach for marketing practitioners and academics. Results indicate that compared to other frequent methods in the literature, the locally linear models are an effective way to solve binary classification problems such as churn, where the aim is to detect a small group out of the customer base. One limitation facing this research was the memory usage by Matlab, which could not easily handle the large number of data in this experiment. For the future, we plan to introduce a bias correction technique to handle the bias caused by balancing the data. Such method is required for the LOLIMOT algorithm similar to the bagging and boosting approach. This enhancement could lead to perfect result as in many other fields of an experiment with LOLIMOT. .

## References:


1- Kim, M.K. and Jeong, D.H. (2004). “The effects of customer satisfaction and switching barriers on customer loyalty in Korean mobile telecommunication services.” *Telecommunications Policy*. Vol. 28 No. 2: PP. 145-159.

2- Kim, H.-S. and Yoon, C.-H. (2004). “Determinants of subscriber churn and customer loyalty in the Korean mobile telephony market.” *Telecommunications Policy*. Vol. 28 No.: PP. 751-765.

3- Brown, K. (2004). “Holding on to customers.” *Wireless Week*. Vol. 15 No.: PP. 6-13.

4- Barret, J. (2003). *US Mobile Market Intelligence*. [cited 2003 August]; Available from: http://www.parkassociates.ecnext.com.

5- Junxiang, L. (2003). *Predicting Customer Churn in the Telecommunications Industry: An Application of Survival Analysis Modeling Using SAS*. Overland Park, Kansas: Sprint Communications Company.

6- Seo, D.B., Ranganathan, C. and Babad, Y. (2008). “Two-level model of customer retention in the US mobile telecommunications service market.” *Telecommunications Policy*. Vol. 32 No.: PP. 182-196.

7- (!!! INVALID CITATION !!!). Vol. No.

8- Song, J.-D. and Kim, J.-C. (2001). “Is five too many? Simulation analysis of profitability and cost structure in the Korean mobile telephone industry.” *Telecommunications Policy*. Vol. 25 No. 1-2: PP. 101-123.

9- Choi, S.-K., Lee, M.-H. and Chung, G.-H. (2001). “Competition in Korean mobile telecommunications market: Business strategy and regulatory environment.” *Telecommunications Policy*. Vol. 25 No. 1-2: PP. 125-138.

10- Kim, H.-S. and Kwon, N. (2003). “The advantage of network size in acquiring new subscribers.” *Information Economics and Policy*. Vol. 15 No. 1: PP. 17-33.

11- Gerpott, T.J., Rams, W. and Schindler, A. (2001). “Customer retention, loyalty, and satisfaction in the German mobile cellular telecommunications market.” *Telecommunications Policy*. Vol. 25 No.: PP. 249-269.

12- Ahn, J.-H., Hana, S.-P. and Lee, Y.-S. (2006). “Customer churn analysis: Churn determinants and mediation effects of partial defection in the Korean mobile telecommunications service industry.” *Telecommunications Policy*. Vol. 30 No.: PP. 552-568.

13- Santhanam, R. and Hartono, E. (2003). “Issues in linking information technology capability to firm performance.” *MIS Quarterly*. Vol. 27 No. 1: PP. 125-153.

14- Beerli, A., Martin, J.D. and Quintana, A. (2004). “A model of customer loyalty in the retail banking market.” *European Journal of Marketing*. Vol. 38 No. 1/2: PP. 253-275.

15- Dvir, D., et al. (2006). “Critical managerial factors affecting defense projects success: A comparison between neural network and regression analysis.” *Engineering Applications of Artificial Intelligence*. Vol. 19 No.: PP. 535-543.

16- Ghiassi, M., Saidane, H. and Zimbra, D.K. (2005). “A dynamic artificial neural network model for forecasting time series events.” *International Journal of Forecasting*. Vol. 21 No.: PP. 341– 362.

17-Hill, T., Connor, M. O' and Remus, W. (1996). "Neural network models for time series forecasts." *Management Science*. Vol. 42 No. 7: PP. 1082-1092.

18-Zhang, G. (2003). "Time Series forecasting using a hybrid ARIMA and neural network model." *Neurocomputing*. Vol. 50 No.: PP. 159-175.

19-Adeli, H. (2001). "Neural Networks in Civil Engneering: 1989-2000." *Computer-Aided Civil and Infrastructure Engineering*. Vol. 16 No. 2: PP. 126-142.

20-Chao, L.-C. and Chien, C.-F. (2010). "AModel for Updating Project S-curve by Using Neural Networks andMatching Progress." *Automation in Construction*. Vol. 19 No.: PP. 84–91.

21-Rumelhart, D.E., Hinton, G.E. and Williams, R.J. (1986). *Learning internal representation by back-propagating errors. D. E. Rumelhart, & J. L. McCleland Parallel Distributed Processing: Explorations in the Microstructure of Cognition*. Vol. 1, MIT Press, Boston.

22-Rumelhart, D.E., Hinton, G.E. and Williams, R.J. (1986). "Learning representations by back-propagating errors." *Nature*. Vol. 323 No. 6188: PP. 533- 536.

23-Battiti, R. (1992). " First and second order methods for learning: Between steepest decent and Newton's method." *Neural Computation*. Vol. 4 No. 2: PP. 141-166.

24-Yu, X.H., Chen, G.A. and Cheng, S.X. (1993). "Dynamic learning rate optimization of the back-propagation algorithm." *IEEE Transactions on Neural Networks*. Vol. 6 No. 3: PP. 669- 677.

25-Nelles, O. (2001). *Nonlinear system identification*, Springer Verlag, Berlin.

26-Sharifie, J., Lucas, C. and Araabi, B.N. (2006). "Locally linear neurofuzzy modeling and prediction of geomagnetic disturbances based on solar wind conditions." *Space Weather*. Vol. 4 No.: PP. 1-12.

27-Pedram, A., et al. (2006). "Local Linear Model Tree (LOLIMOT) Reconfigurable Parallel Hardware." *Transactions on engineering, compuring and technology*. Vol. 13.

28-Nelles, O. and Isermann, R. (1996) *Basis Function Network for interpolation of local linear model*. *Proc Decision and Control*, Kobe, Japan.

29-Holzmann, H. and et al. (1999) *Vehicle Dynamic Simulation Based on Hybrid Modeling*. *Proc International Conference on Advanced Intelligent Mechateronics*: IEEE Explorer.

30-Ghorbani, A., Taghiyareh, F. and Lucas, C. (2009) *The Application of the Locally Linear Model Tree on Customer Churn Prediction*. *Proc 2009 International Conference of Soft Computing and Pattern Recognition*.

31-Rezaie, J. and et al. (2006) *Modified LOLIMOT Algorithm for Nonlinear Centralized Kalman Filtering Fusion*. *Proc CIS 2006*: IEEE Xplore.