J. Basic. Appl. Sci. Res., 3(6)1018-1026, 2013

ISSN 2090-4304
Journal of Basic and Applied Scientific Research
www.textroad.com

Providing a Customer Churn Prediction Model Using Random Forest and Boosted TreesTechniques (Case Study: Solico Food Industries Group)

Sadaf Nabavi, Shahram Jafari

Department of Computer Science and Engineering
Shiraz University

ABSTRACT

In order to succeed in the global competition, organizations need to understand and monitor customers' behavior, so that they could retain them by predicting their preference and behavior before others. Recently, marketing strategies have been changed from product-oriented strategies to customer-oriented strategies and most organizations have focused on customer relationship management. In fact, more organizations have found out that retention of their present customers, as their most valuable asset, was very important. Therefore, with the aim of describing data mining abilities in churn management, and designing and implementation of a customer churn prediction model using a standard CRISP-DM (Cross Industry Standard Process for Data Mining) methodology based on RFM(Recency, Frequency, Monetary) and random forest and boosted trees techniques, the database of one of the biggest holdings of the country, Solico food industries group, is explored. Using this model, the customers tending to turn over are identified and effective marketing strategies will be planned for this group. Customer behavior analysis indicated that length of relationship, the relative frequency and the average inter purchase time were among the best predictors.
KEYWORDS:-Churn Model; CRISP-DM Methodology; RFM Model;Judgment-based RFM; Classification; Random Forest;Boosted Trees.

I. INTRODUCTION

The increasing competition and decreased customer loyalty rate in recent decades have led to emergence of a paradigm in marketing which encourages the transition of organizations from adopting a product-oriented approach to a customer-oriented one [1]. In such circumstances, a mere reliance on advanced technology and high quality is not enough for customer retention, because competitors are advancing rapidly as well and soon will reach the same level. Moreover, customers continuously obtain their required information from various communication channels. Through this increased knowledge, the customers would become aware of new options in the market and thus their loyalty to the organization would decrease.

Losing customers not only causes additional costs due to the sales losses, but also it will create the need for acquisition of new customers [2], while the cost of new customer acquisition is 5 to 6 times more than the cost of retaining the present customers [3].The cornerstone of customer relationship management strategies, in general, and customer retention, as a part of it, is customer knowledge. Customer knowledge is the pattern that reveals his real preferences [4].This knowledge in small businesses with limited customers is the outcome of business owner's business intelligence. However, in the largebusinesses which use vast information systems, data mining takes the role of systematic generation of business intelligence [5].

Previously marketers only sought to find customers and the sales group sought to hunt for new customers but in today's vision, marketing means customer growth, paying attention to his satisfaction and quality in his viewpoint. Finally, the art of marketing today is that customers of an organization are cooperators within and supporters outside the organization. Today, making the customer committed has gained a special status. Issues like loyalty, brand loyalty and customer loyalty has been addressed by many researchers. Loyalty creates a positive perception in the minds of the listener. Marketing knowledge experts have also considered many advantages for loyalty, the most notable of which are as follows:

a) Reducing costs and attracting new customers: These costs include the costs incurred by the company to attract each customer for the first time. This price includes costs of direct advertising, sales commissions, sales forces, promotion, etc. That is spent to attract these customers.

b) Income growth: In fact, the actual profit begins when the customer makes purchases with more amounts and types and shows the values resulted from his/her loyalty.

c) References: Higher predictive power of corporations through loyal customers and lower customer risk to return for buying will result to improved return on investment (ROI) and higher operating profit for the company.

d) Reducing customer sensitivity to prices and changes.

e) Benefits of customer lifetime value.

*Corresponding Author: Sadaf Nabavi, Department of Computer Science and Engineering, Shiraz University
Email: sadaf.nabavi@gmail.com (098)125306150

f) Positive performance through increasing predictive power.

g) Increasing barriers for new competitors to entry. [6]

To achieve above objectives, this study develops a model for predicting customer loyalty. The results of the study by Verhoef and Donkers on the articles published from 1995 to 2002 suggest that adaptive cross-tabulation methods were the most common methods of predictive segmentation and modeling in the past and then RFM-based methods had the largest contribution to this modeling. Based on these studies, the Linear Regression Analysis is a common technique for predictive modeling. Interestingly, three methods of CART, Logit Regression and Neural Networks which are now very common, accounted for the last three ranks in the predictive models discussion at that time period. Meanwhile, by reviewing over 89 published articles on customer behavior modeling in the period from 2000 to 2005, Haden suggests that RFM is used more in the definition of the variables used in modeling and not the practical technology, and adaptive cross-tabulation methods were excluded as old and useless methods and were replaced by more advanced and accurate methods. The results show that Regression Analysis is the main method used in most research in this area and then Decision Trees, Markov Models and Neural Networks are in the following ranks. In another study by Haden et al., they investigated 46 articles from 2005 to 2010 and found that the simple regression method is the most used method in customer behavior modeling to date. However, the application interval of this method from other data mining methods has been substantially reduced in recent years but the application of decision trees methods - such as Random Forests and Boosted Trees - has been greatly increased due to the ease of use and accurate interpretation in binary classification. [7, 8]

The scientific contributions of this paper are mentioned below:

a) Developing a comprehensive customer churn model with empirical testing with a large sample of actual customer payment transactions.

b) Measuring customer churn risk based on customer behavioral characteristic as prediction variables

c) Modeling customer churn based on new decision tree techniques such as random forest and boosted trees.

d) Combining existing models and using hybrid prediction model to increase mode accuracy and to achieve reliable results.

Solico Food Industrial group was set up in 1973 and soon after becomes a leading manufacturer particularly in Dairy, Meat, Ice cream and Beverages with some well established brands such as Kalleh, Maac, Castle, Pemina, Koochin, etc. Over 40 years of experience Solico has been awarded many quality and certifications such as ISIRI, ISO9001:2000, HACCP and it has been ranked as an Iranian top exporter. This group with more than 1.2 billion dollars of annual turnover and over 8500 personnel is a top-notch brand in food industries. The group has focused on improving the level of wellness, innovation and satisfaction in customers by the products and in the mind of employees by the company culture. In the present study, attempt has been made to design a model for prediction of churning customers in the organization by studying the database of Solico food industries group using a classification technique based on REM (Recency, Frequency, Monetary) model and CRISP-DM(Cross Industry Standard for Data Mining) methodology. This study is organized into five sections. The second section reviews the literature on the concept of customer churn. The methodology used in this study including how to determines the target group, prediction model, classification techniques, data sets and prediction variables are examined in the third section. The fourth section presents the model results and their evaluation. Finally, the fifth section discusses the research and its limitations and suggestions for future studies.

II. THEORETICAL FOUNDATION AND BACKGROUND

A. The concept of customer churn

With the emergence of e-commerce, customers can more easily be aware of the market and its opportunities. Their expectations become more and more and they become more willing to change their current supplier and choose a new one. This caused a phenomenon called customer churn. Minguel defined the customer churn as the annual leaving of supplier by customers [9]. Van Den Poel believes that the customer churn management is affected by two separate approaches:

1. Passive approach: in this approach, the measures for encouraging the customer to remain are taken when he requests to end the relationship with the organization.
2. Active approach: in this attitude, the company chooses an approach to identify the groups prone to churn and implement special promotional or incentive programs with the aim of churn prevention [10].

The managers' experience indicates that, the likelihood of effectiveness of promotional programs of second approach for customer retention is more than first one. Moreover, the final costs imposed on the company would be lower. Part of the costs due to customer loss, would be the potential profit of the rival companies. Furthermore, the repetition of unsuccessful interaction by the unsatisfied customer, can also affect the effectiveness of customer acquisition strategies [10].Therefore, the accurate definition of churn concept has great importance. In all of the previous studies, it is observed that the definition of churn is by agreement and there is no identical and common definition which would be applicable to all industries. In fact, the churn phenomenon is defined through agreement and by the experts of a specific industry or organization by considering the characteristics of the environment of that industry and organization [11].

The study of Wu and Chen shows that, the customers that have recently purchased from the organization are more likely to remain active, compared to the customers that much time has elapsed since their last contact with the organization [12].The results obtained from most of the previous studies indicate that, the less the regency amount would be, the more is the likelihood of customer loyalty. Since the customers are the ultimate resource for business growth, the assessment of their behavior is very important. There are various methods for calculation of customer lifetime value. Some of the researchers partitioned the customers based on current value, potential value and loyalty [13, 14].

In the process of prediction of loyal and churn customers, it is necessary to identify the process during which the customer lose his loyalty and become prone to churn. In this research, since the accurate time of churn is not determined, the organization needs to ensure that the relationship with the customer is still alive. In such industries, the customer starts the relationship with the organization in irregular time intervals and might repeat his purchase irregularly. Therefore, in this research the comparison of customer's regency in his previous purchases is used for binary classification of customers into two groups of churn and loyal. Such that, if the customer's regency is more than the maximum time interval between two successive purchases in the past, that customer has changed his recorded pattern and is prone to churn.

III. METHODLOGY

Data mining is a data analysis approach which discovers knowledge. Different algorithms and procedures are developed to do this. One of these algorithms, which can serve analyst well in knowledge extraction, is the standard CRISP-DM algorithm. Since this method is data independent and can be used in order to analyses data in general; in this research, it is used as the basis for data analysis in the organization. This algorithm consists of six phases as business understanding, data understanding, data preparation, modeling, evaluation and deployment. Overview of CRISP-DM methodology and target group identification levels are illustrated in Fig.1.

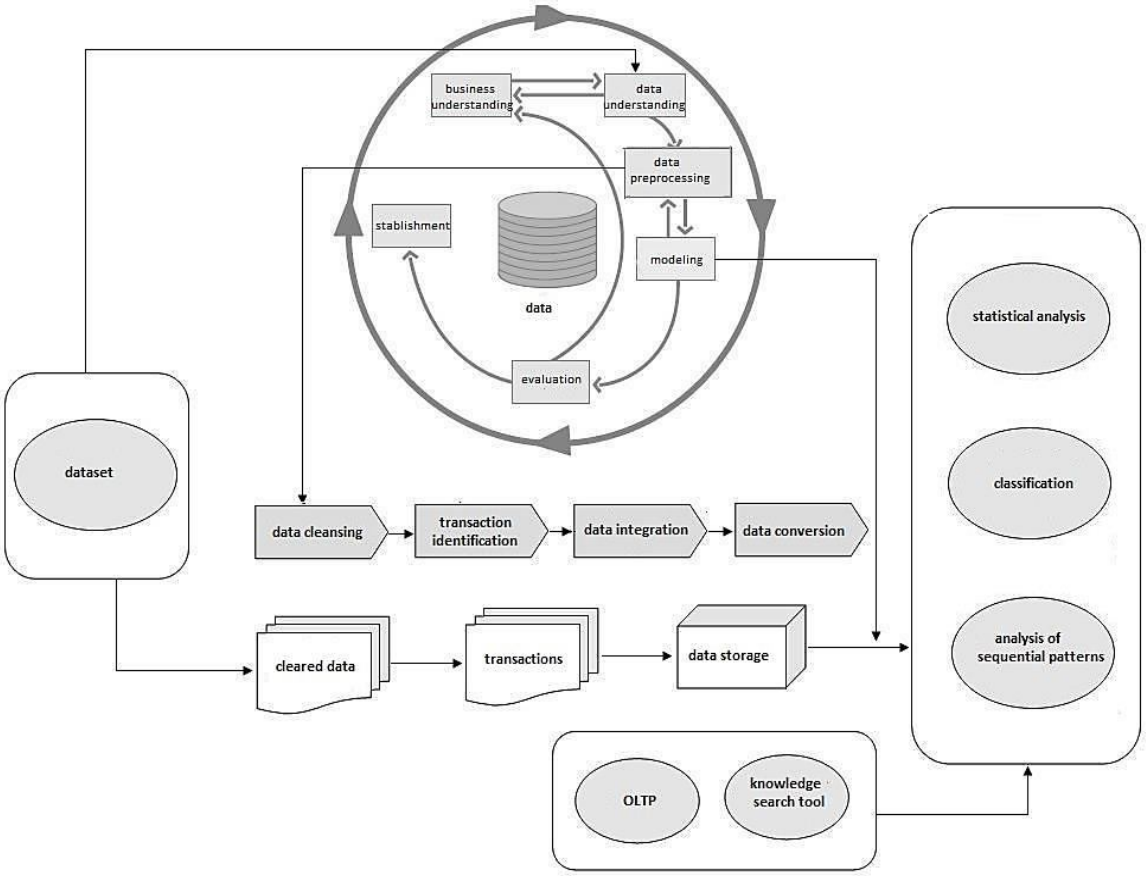

Figure1. Overview of CRISP-DM methodology and target group identification levels

A. Dataset Integration

Solico group as one of the top holdings in the country has special emphasis on customer orientation and the importance of understanding the customer needs as its fundamental principles, during its operation years. Moreover, considering the nature of the goods supplied by the organization and large dispersion in the customers purchase time and the uncertainty about return time, the identification of churn customers is necessary. The Urmia distribution company which is one of the distribution companies of Solico groups and has direct interaction with customers is selected as the target group of this study.

The dataset of this research consists of 5079 customers data acquired from 1024170 sales invoices, in the time interval from 8th April 2010 to 8th September 2012.In this step of data mining process that falls into the pre-processing phase, completion of null values, smoothing of disturbed data and eliminating the data incoherence is performed. In order to reduce the effect of different scales of different variables, the Z score method is used for data normalization in this study [15].

B. Modelling and target group identification

After finishing the data preprocessing phase, the modeling stage, with two sub-stages: high-value customers group identification model and churn prediction model, start.

1) High-value customers group identification model

In the process of customer churn management, it is not necessary to focus on all of the customer databases; because according to the performed researches, usually 80 percent of the profit of an organization is provided by only 20 percent of its customers. This rule is known as Pareto principle, or 80/20 rule [4]. Therefore all the customers don't have identical values and it is not necessary to maintain long-term relationship with them [7]. Thus, in order to identify the loyal and churn customers, the focus of this research is on high-value customers. Customer valuation of organization is performed based on RFM model; and in this partitioning customer valuation is calculated based on regency features (time elapsed from the last transaction of customer with organization), purchase frequency (number of customers transactions in a specified time interval) and monetary value (the amount of money that customer spent in the organization in a specified time interval) [16].

The RFM features are appropriate predictors for customer lifetime value. These variables are used in various ways for customer valuation. In one of these methods, each one of the R, F and M variables have the same value. In a more recent approach, a weight is assigned to each variable; then, the customer valuation is done based on desired weights. Since the weights are determined based on experts' opinion, the approach which is used in this study is called judgment-based RFM [16, 17].

In a continuous purchase period, the customers that have continuous profitability for organizations have higher values as well. During the active lifetime of their relationship with organization, these customers will have larger amount of purchase and can perform effective verbal advertising as well. The combination of F and M indices indicate that the customers that visit the organization continuously and frequently have high monetary value for organization [16]. In this study in order to identify the high value customers, the M variable is ignored; because the organization experts believe that the customers visiting to the organization frequently but creating little monetary values, and actually being potentially high-value, fall into the category of high-value customers. The R variable is not used in customer valuation, because it's possible that the loyalty of customers would be decreasing. In fact this variable is used for identification of churn customers.

Among the 5077 customers purchasing from the organization, 4618 customers have frequent purchase or in other words, $F_{customer} > 1$. These customers constitute93 percent ofthe entire database. In this database 459 customers only oncehave purchased from the organization, this number which constitute 7 percent of entire database, are ignored. By performing aggregationqueries on the present database, it was determined that the average purchase frequency of the customers that have repeated purchase is 40. Thus, the customers that their purchase frequency is more than this average value are identified as high-value. However, by investigating the database it was determined that 1584 customer fall into the category of high-value customers. This group constitutes 24 percent of the entire database and 34.30 percent of the customers that had repeated purchase.Based on the studies performed on high-value customers' data and investigation of their status in the organization it was determined that the average of purchase frequency in the high-value customers group is 91, which is doubled compared to the customers with repeated purchase and is 2.57-fold compared to the average of the entire database.

Moreover, the money spent in the organization by the high-value customers group constitutes 76.43 percent of the total money spent by the customers with repeated purchase. This value is 76.11 percent of the total money spent in the entire database. Thus, as it was mentioned before, 76 percent of the total profit of organization is provided by 24 percent of its customers, which indicates their high value.

2) The churn prediction model

In the considered industry the database of customers is consists of a series of active customers which visit the organization regularly and a series of inactive customers. Therefore, determining the border between active and inactive customers is not trivial; because the inactive customers might become active in future. Thus, the study of churn conditions has significant importance. Previous studies and researches indicate that the customers who had recently purchased from the organization, are more likely to remain active compared to the customers that a long time has elapsed from their last purchase. Thus, the less the

recency value is, the more would be the likelihood of customer's loyalty [12, 18].The churn prediction model is illistrated in Fig.2.

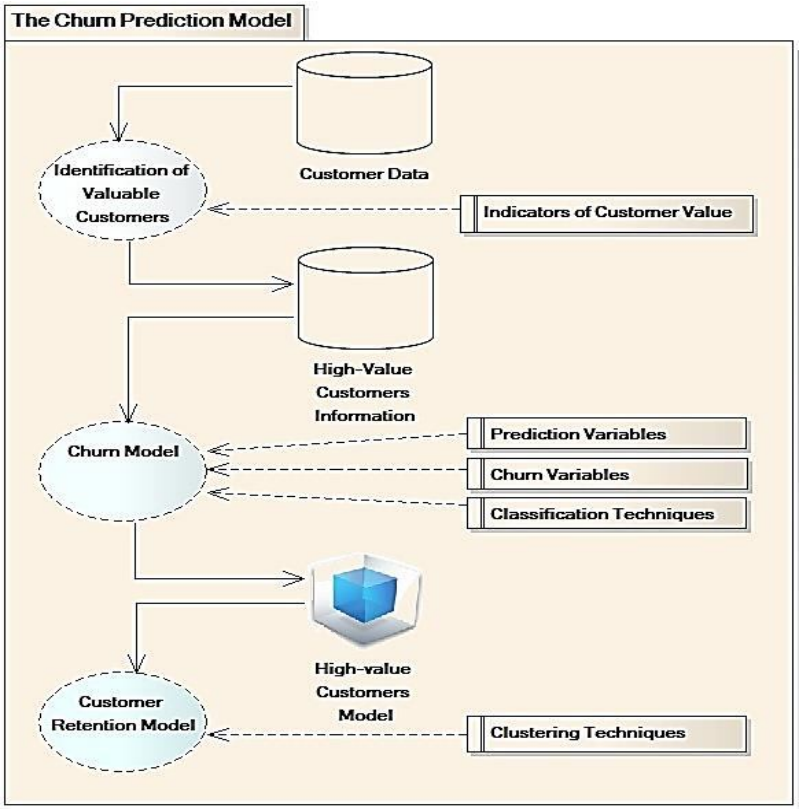

Figure2. The churn prediction model

According to the churn definition in this study, 469 customers among 1584 high-value customers are considered as churn. This number is equivalent to a churn rate of 29.06 percent; from these the share of customers in the supermarkets is 90 percent.

3) Classification technique

Segmentation of customers who lost their loyalty can be done using any of the classification methods. Since the final objective is to categorize customers into two retained and churn groups, a predictive model is required. Predictive data mining offers a model of system that includes the use of variables and fields in data warehouse in order to predict unknown values. Several methods have been used to predict customer churn behavior in literature. In this study, two effective techniques of random forest and boosted trees have been compared for segmentation of churn customers[19].

Random forest: Among implementation mechanisms of a model for customers classification based on their behavior, decision tree is the considered option. Decision trees are well known due to their ease of use and interpretation in binary classifications [19]. Application of these trees allows the usage of predictor variables with different scales. A problem in using decision trees is their instability and creation of local optimal solutions. One way to confront these problems is the new classification technique called random forest. Regarding the prediction performance, it is one of the best classification techniques. The classification outputs of this approach are very powerful than confused and deviated data. This method provides useful information on the importance of each variable, thus can determine those variables that have the greatest impact on the dependent variable [20, 21]. In this technique, a set of decision trees are grown and each tree votes for the most popular class, then the votes of different trees are integrated and a class is predicted for each sample. In this approach which is designed to increase the accuracy of the decision tree, more trees are produced to vote for class prediction. This approach is an ensemble classifier composed of some decision trees and the final result is the mean of individual trees results. Some copies of dataset are made in the learning phase and a decision tree without pruning is separately created from each copy, thus the result of each test is the mean of prediction result of all the trees [22]. Following the method proposed by Breiman, a huge number will be considered for the produced trees. Using a set of trees in this method can be effective in significant increase of prediction accuracy [22].

Boosted trees: Boosted trees method is shown as one of the most powerful predictive data mining techniques over the past few years. This technique is developed through application of boosting method in decision trees. In boosting method a weight is assigned for each learning object and training is repeated for a set of K classifier. After training the $M_i$classifier, the weight of

learning objects is updated to allow M_{i+1} classifier pay more attention to objects not accurately classified by $M_i$classifier. In this method a weight is assigned to the vote of each classifier based on its performance power. Less the classifier error rate, more the weight assigned to its vote. The final classifier, M^*, combines the votes of each classifier. For each class, the weight of classifiers which voted for object X of that class is added. The class which gains greater total weight will be the winner class and will be introduced as the predictive class for object X[23].

4) Predictive variables

In the real world, data seldom are gathered for data mining purpose. In such cases, usually we are confronted with a huge volume of data and want to use them for a specific application. Thus, dimensionality reduction is one of the required measures. Feature generation, extraction and selection are among the effective data reduction strategies. In this study using the observed behavior of customers in the past, a number of variables are selected for churn prediction. These variables are widely available and known as powerful and efficient predictive factors [24]. Following each variable is reviewed:

Purchase Frequency: in this study the RFM model is used for churn customer identification. In this model purchase frequency is one of the important parameters in identification of loyal customers. This variable is the number of customer's visits to the organization.

Length of Relationship (LOR): since the considered organization is among the non-contractual industries, and it's possible that customers refer to the organization irregularly, this index indicates the time that customer has started his relationship with the organization. In fact this variable indicates the time interval between the first and last purchase of customer in the observed interval.

Relative Purchase Frequency: in this study the relative purchase frequency of the customer is studied as well. This variable is calculated as the ratio of the number of customer's visits to the active relationship length of the customer with the organization.

Average Inter-purchase Time: one of the most effective variables in the identification of churn customers is the average inter-purchase time; which is the average time elapsed between two successive purchases of the customer. Due to irregularity of the customers' behavior in the time intervals of visits to the organization, this variable has great effect in identification of customers.

Variance of Inter-purchase Time: in order to study the regularity of customers' behavior in visits to the organization, and evaluation of the data dispersion around mean, the variance of inter-purchase time is used as another predictive variable.

Last Inter-purchase Time Difference: in order to be able to have better results from the customer's behavior, the variations of his recency in the last purchase are used.

Monetary value: one of the variables used in the RFM model is the amount of money spent in the organization by the customer, based on which in combination with customers purchase frequency, the high-value customers can be identified. Moreover, this variable can be used in the prediction of customers purchase pattern in the future.

Relative Monetary value: this variable indicates the amount of customer's purchase compared to his interaction time with the organization.

Weight of Sold Items: based on the organization experts' opinion, the total weight of customer's purchased goods is used as one of the customer's behavior study variables.

Customer Behavior in Product Category: in this study this variable is used in order to study the customer's behavior in different product categories. In the Solico food industries, there are different products such as dairy, drinks, sauces, meat, packaging, trading, and undefined; which are sold to the customers by the Urmia distribution company. Therefore, in order to represent the customer's purchase from the considered product group, 7 binary variables are used.

Number of Categories: this variable indicates the total number of product groups purchased by the customer.

Customer Type: in this study the customers are divided into two general categories: real and legal. In the legal category there is another categorization. The legal customers are divided into subgroups such as hotels, supermarkets, chain stores, restaurants, schools, villages and etc. determining that the loyal customers belong to which group, increases the possibility of marketing and providing better services in order to encourage them.

IV. EVALUATION AND RESULT ANALYSIS

1) Quantitative assessment

When using functions such as classification which is predictive in nature, assessment of extracted models will be necessary. For this purpose, the data of valuable customers are randomly divided into two learning data, test data sets in a 70 to 30 ratio, and then the model is derived from it. Then the accuracy of the model is measured from evaluation data set. In order to verify the performance accuracy and effectiveness of the prediction model, confusion matrix and cumulative gain curve were used in this study. The results presented in Table 1 show that the prediction of valuable customers churn is a critical strategy. Our results indicate that both techniques of random forest and boosted trees have similar function in prediction of customers churn and there is no significant difference among them. The results of this comparison show the advantage of decision tree techniques to predict the customers churn behavior.

TABLE I. RANDOM FOREST & BOOSTED TREES RESULTS

Model	*Accuracy*	*Sensitivity*	*Specificity*	*Precision*
Random Forest	76.64%	80.48%	64.06%	88.43%
Boosted Trees	63.32%	67.81%	66.85%	86.76%

The results show that the prediction of valuable customers churn is considered as an essential strategy. Fig.3 and Fig.4 show the cumulative gain curve for customer apt to churn in random forest and boosted trees techniques. The diagonal line in cumulative gain curve, called the base line, represents a non-purposeful activity that randomly selects a subset of customers as potential churn. If the non-purposeful activity randomly selects x% of total customers, this subset will contain x% of total actual churns. [25] In this chart, the x axis shows the selection percentage of whole customers (in this case 1584 customers) and the y axis shows the inclusion percentage of whole churns (in this case 469 churns). As these charts show, application of this model will result in more benefit than the baseline.

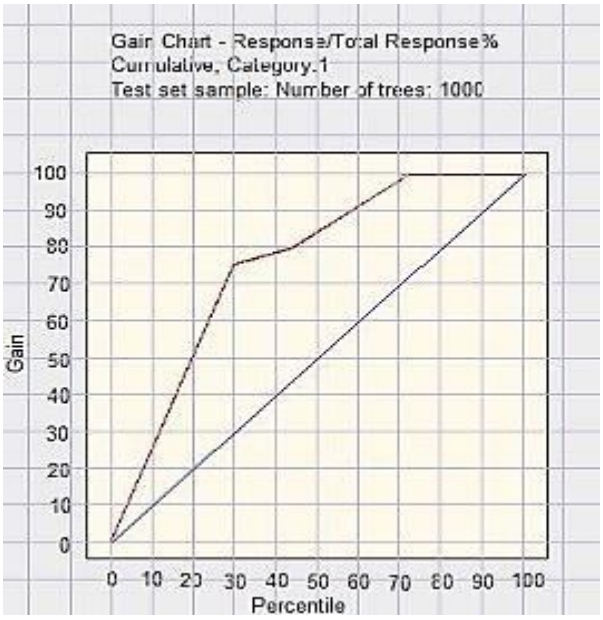

Figure 3. Cumulative gain curve in random forest technique

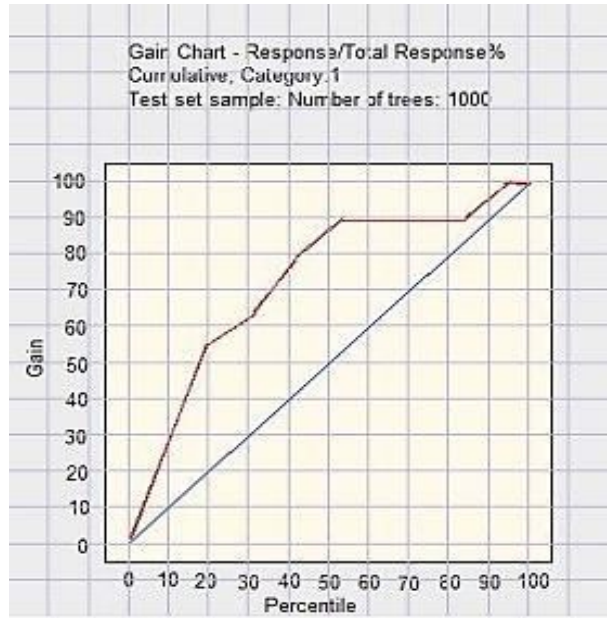

Figure 4. Cumulative gain curve in boosted trees technique

2) Qualitative Assessment

The performance of classification techniques for predicting churn can be studied regarding the false alarm rate and miss alarm rate. False alarm rate is the percentage of non-churns that are falsely presented as churns. Miss alarm rate is the percentage of predictions of the applied technique that failed in determining the churns. If there is a high false alarm rate, the retention cost of loyal customers augments; and if there is a high miss alarm rate, the opportunity cost will increase due to lack of retention and customers churn. [25] According to the study, the cost of customer retention is much less than the cost of new customer attraction in the studied organization. The objective of the method used to predict the churn is to gain the minimum achievable miss alarm rate, while the false alarm rate is kept at an acceptable level.

Hence, in order to reduce the miss alarm rate and increase the efficiency of prediction system, two prediction models of random forest and boosted trees with a logical operator were combined. In this case, if at least one of the prediction models for the studied customer assigns the churn label, the prediction system will introduce that object as a churn risk customer. Therefore, the miss alarm rate will reduce and the false alarm rate will increase. The results indicate enhanced performance of combined prediction system in identifying customers willing to churn. This means that the application of hybrid system reduces the number of churn customers who were falsely introduced as loyal by the model, by about 7.2%. In this case, the likelihood of retention of customers willing to churn increases through implementation of proper strategies, which in turn will have significant impact on the profitability of the organization. Hybrid prediction model is illistrated in Fig. 5.In the hybrid prediction model, customer data are independently examined by each of the random forests and boosted trees techniques. According to characteristics and processing of input data, each model introduces a churn or loyal customer, so that if a customer is identified as loyal in the studied model, it will be assigned the tag $P_M = 1$, otherwise $P_M = 0$. After processing of both techniques and customer tagging, a hybrid system is used to select votes. If both models introduce a churn customer, then the hybrid model assigns the churn tag to the customer; and if both model introduce the customer as loyal, then it introduces the customer as loyal. If at least one of the prediction models identifies the customer as churn, the hybrid model will introduce the customer at the risk of leaving.

The importance of each predictor in random forest technique is shown in Fig. 6 As can be seen, the top 10 variables are behavioral variables. Meanwhile, length of relationship, the variance of purchases interval, relative frequency, and monetary value are the major predictors for segregation of loyal customers from customers tend to churn.

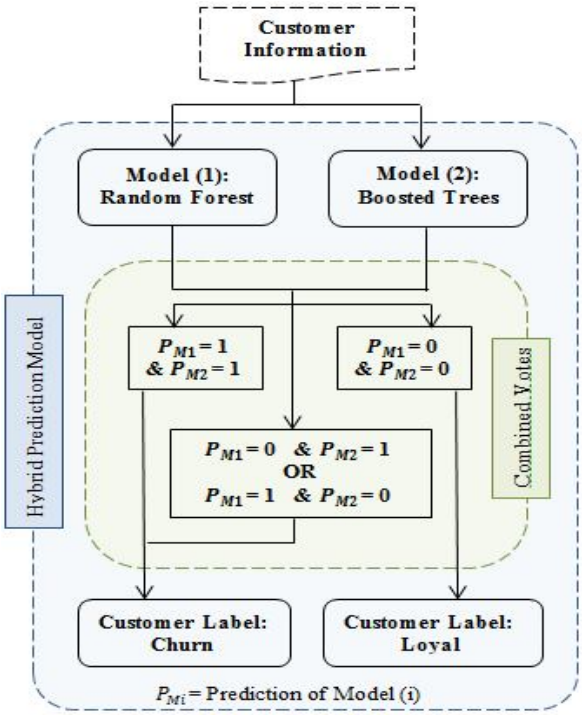

Figure 5. Hybrid prediction model

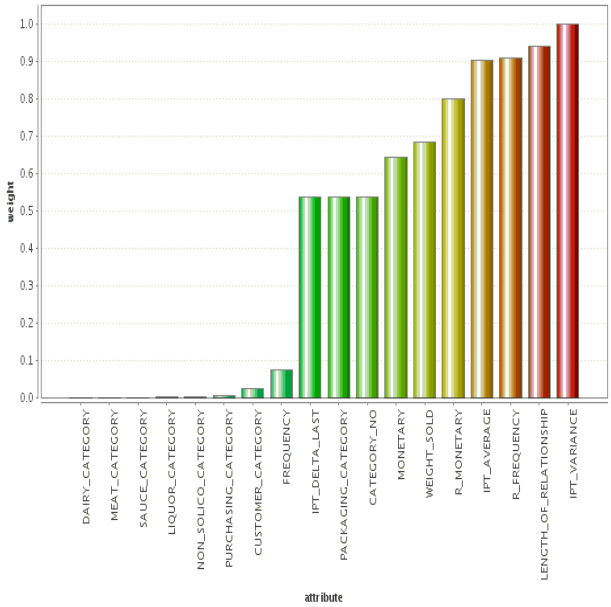

Figure 6.Importance of prediction variables resulted from Rapid Miner software

V. DISCUSSION AND CONCLUSION

All organizations are dependent on their customers, so they should try to satisfy them and maintain their position in the market. Therefore, focus on customer will result in quick response of the organization to market opportunities. Thus identification of churn customers is worthwhile for the organization. This study examined the data of Solico group of food industries in nonconventional industries context in which the effective factors in customer churn prediction was studied and instead of using customer mental desire for churn, modeling has been done through observation of its actual behavior. This process makes a deeper understanding of customer's behavior. To identify the churn indices, customer's behavioral and demographic variables were used; by development of these indices based on customer perception, deep insights can be gained from his behavior. The studied data were collected within a period of two years; by increment of this period, its impact on the model can be evaluated. In this model, only the indices affective on churn prediction were studied; this issue is the starting point of churn management. In future studies, strategies to prevent customer loss can be explored through analysis of a retention model and determination of churn factors and clustering of customers based on these factors.

REFERENCES

[1] S.Y. Hung and D. Yen and H.Y. Wang, "Applying Data Mining to telecom churn management". *Expert System with Applications*, Vol. 31, No. 3, pp. 515-524, Elsevier, 2006.

[2] A.D. Athanassopoulos, "Customer satisfaction cues to support market segmentation and explain switching behavior". *Business Research*, Vol. 47, No. 3, pp. 191-207, Elsevier, 2000.

[3] M.R. Colgate and P.J. Danaher, "Implementing a customer relationship strategy: The asymmetric impact of poor versus exelentexcution". *Academy of Marketing Science*, Vol. 28, No. 3, pp. 375-387, Springer, 2000.

[4] Y.F. Wang and C. Ding-An and H. Mei-Hua, "Arecommender system to avoid customer churn: A case study". *Expert System with Applications*, Vol. 36, No. 4, pp. 8071-8075, Elsevier, 2009.

[5] J. Shahrabi, "Data Mining". *Amirkabir University* Vol. 1, 2007.

[6] P. Kotler and A. Armstrong, "Principles of marketing". *Prentice Hall*, 13th edition ISBN. 0137006691, 2009.

[7] J. Hadden and A. Tiwaria and R. Roya and D. Rutab, "Cumputer Assisted customer churn management: State of the art and future trends". *Computer & Operation Research*, Vol. 34, No. 10, pp. 2902-2917, Elsevier, 2005.

[8] P.C. Verhoef and B. Donker, "Predicting customer potential value: An application in the insurance industry". *Decision Support System*, Vol. 32, pp. 189-199, Emerald, 2004.

[9] A.P. Minguel, "Measuring the impact of data mining on churn management". *Internet Research: Electronic Network Application and Policy*, Vol. 11, No. 5, pp. 375-387, 2001.

[10] E.W.T. Ngai and L. Xiu and D.C.K. Chau, "Application of data mining techniques in customer relationship management: A literature review and classification". *Expert System with Applications*, Vol. 36, No. 4, pp. 2592-2602, Elsevier, 2009.

[11] W.J. Reinartz, "On the profitability of long life customers in a noncontractual setting: An empirical investigation and implications for marketing". *Journal of Marketing*, Vol. 64, pp. 17-35, Elsevier, 2000.

[12] C. Wu and H.L. Chen "Counting your customer's in store decisions, inner purchase time and repurchasing behavior". *European Journal of Operational Research*, Vol. 127, pp. 109-119, 2000.

[13] S.L. Chan and V.Cho, "A model for predicting customer value from perspectives of product attractiveness and marketing startegy". *Expert System with Applications*, Vol. 37, No. 2, pp. 1207-1215, Elsevier, 2010.

[14] H. Hwang and T. Jung and E. Suh, "An LTV model and customer segmentation based on customer value: A case study on the wireless telecommunication industry". *Expert System with Applications*, Vol. 26, No. 2, pp. 181-188, Elsevier, 2004.

[15] D.T. Larose, "Discovering Knowledge in data: An introduction to data mining". *John Wiley & Sons*, ISBN. 0471666572, pp. 35-39, 2004.

[16] W. Buckinx and D. Van Den Poel "Customer Base Analysis: partial defection of behaviourally loyal clients in a non-contractual FMCG retail setting". *European Journal of Operational Research*, Vol. 164, pp. 252-268, 2005.

[17] J.A McCarty and M. Hastak, "Segmentation approach in data mining: A comparison of RFM, CHAID and logistic regression". *Business Research*, Vol. 60, pp. 656-662, Elsevier, 2007.

[18] J. Wu and Z. Lin, "Research on customer segmentation model by clustering". In Proceeding of the Seventh International Conf. on electronic commerce (ICEC'05), 2005.

[19] K. Coussement and V. Den poel, "Churn prediction in subscription services: An application of support vector machines while comparing two parameter selection techniques". *Expert System with Applications*, Vol. 34, No. 1, pp. 313-327, Elsevier, 2008.

[20] A. Cutler and D.R Cutler and J.R Stevens, "Ensemble Machine Learning: Random Forests". *Springer US*, ISBN. 978144199325, 2012.

[21] Z.H. Zhou, "Ensemble methods: Foundation and algorithms".*Taylor& Francis Group*, 3th edition, ISBN. 9781439830031, 2012.

[22] F. Livingston, "Implementation of Breiman's random forest machine learning algorithm". *ECE591Q Machine Learning Conference*, 2005.

[23] J. Han and M. Kamber, "Data mining concepts &techniques".*Morgan Kaufmann Pub.*, 3th edition, ISBN. 97831, 2011.

[24] W. Buckinx and D. Van Den Poel "Customer Base Analysis: partial defection of behaviourally loyal clients in a non-contractual FMCG retail setting". *European Journal of Operational Research*, Vol. 164, pp. 252-268, 2005.

[25] C.P. Wei and I.T. Chiu, "Turning telecommunication call details to churn prediction: A data mining approach". *Expert System with Applications*, Vol. 23, No. 2, pp. 103-112, Elsevier, 2002.