Customer Churn Prediction Model using Data Mining techniques

Ibrahim M.M.Mitkees
Department of managment information system
Modern Academy
Cairo,Egypt
elnegmmasry@yahoo.com

Asist. Prof. Sherif M Badr
Department of computer science
Modern Academy
Cairo,egypt
sherif_badr@afmic.com

Dr. Ahmed Ibrahim Bahgat ElSeddawy
Department of Information Systems
Arab Academy for Science Technology & Maritime Transport (AASTMT)
Cairo, Egypt
Ahmed.AlSeddawy@aast.edu

***Abstract*— A big problem that encounters businesses, especially telecommunications business is 'customer churn'; this occurs when a customer decides to leave a company's landline business for another cable competitor. Therefore, our aim beyond this study to build a model that will predict churn customer through defining the customer's precise behaviors and attributes. We will use data mining techniques such as clustering, classification and association rule. The accuracy and preciseness of the technique used is so essential to the success of any retention attempting. After all, if the company is not aware of a customer who is about to leave their business; no proper action can be taken by that company towards that customer.**

***Keywords*— Data Mining; Customer Churn; clustering; classification;association rule.**

I. Introduction

One of the most direct and effective approaches to keep the current customers is that the company should be able to foresee potential churn in time and react to it quickly. Recognizing the indications of potential churn; satisfying customer needs, restoring and re-establishing loyalty are actions supposed to help the organization minimize the costs of gaining new customers.

Churn prediction modelling systems serve as a means of understanding the customer's exact behaviour and work as an alert to the danger and timing of customer churn. The precision of the strategy used is considered essential to the achievement of any proactive retention intention. After all, if the decision maker cannot anticipate customer's intention to leave their company, there can be no proper decision taken concerning that customer.

In section 2 of this paper, we will expose some previous studies conducted on churn prediction, together with showing out the different techniques used with the different algorithms applied. In section 3, the author will explain the proposed method and the way of setting up the experiments, the collected dataset, data mining techniques and the results obtained. In section 4, we will sum up the research and give proposals for future exploration; Section 5 of this paper deals with suggestions and conclusions.

II. Literature Review

(Cimpoeru and Anca, 2014) discovered a way to classify clusters and find out predictions through commercial ways in a relational database management system. The researchers have identified two models of predicting customers who have a high possibility to leave. They are 'decision tree' and 'naïve Bayes classification' models.

(Manjit Kaur et al., 2013) provided a detailed guideline for changing a customer's raw data of a bank into useful data and then converting it into useful information through using data mining techniques. They have got the data for chosen attributes from the customer's raw data for a selected group of 2000 customers. They exploited naive Bayes, decision tree and support vector machine classifier to recognize significant customer attributes to predict churn. Yet, the prediction success rate of Churn class is more than the prediction success rate of loyal class. This research work aims at predicting the future churn of bank customers which can be addressed by intervention, in the hope of reducing the lost revenue. Thereby, with a better comprehending of these characteristics, a customized approach can be developed by the bank in the context of their Customer Relationship Management strategy.

(Sharma et al., 2013) proposed a neural network (NN) - based approach to predict customer churn in subscription of cellular wireless services. Furthermore, it was found that

978-1-5386-4266-5/17/$31.00 ©2017 IEEE

medium- sized NNs performance is essential for the customer churn prediction when different neural network's topologies were experimented.

(Correa Bahnsen et al., 2015) have presented a new cost-sensitive framework for customer churn predictive modeling. First, they have proposed a new finance- based measure for assessing the effectiveness of a churn campaign, taking into consideration the available portfolio of offers, their individual financial cost and probability of offer acceptance depending on the customer profile. Then, using a real-world churn dataset, they compared different cost insensitive and cost-sensitive classification algorithms and measured their usefulness, based on their predictive power and the cost optimization. The results have shown that using a cost-sensitive approach causes an increase in cost savings of up to 26.4%.

(Praveen et al., 2015) have used data mining technique and R package to predict the results of churn customers on the benchmark Churn dataset. They have evaluated, the number of churns using the classification technique J48 tree. The R tool represents the large dataset churn in the form of graphs which depict the outcomes clearly in a unique pattern visualization manner. The Churn Factor is used in many functions to describe the various areas or scenarios when the churn rate is high. The study proposes that there is a huge deviation in the graph of churners when customer service calls are considered.

(Devi P. U. and Madhavi S. 2012) have given a detailed guideline for changing a customer's raw data of a bank into useful data and then converting it into useful information, using data mining techniques. They have expounded the concept of dynamic timeline that should be considered while converting raw data into useful data. They have extracted the data for chosen attributes from a customer's raw data of a selected set of 1,484 customers. It has been found out that in these 1,484 samples, 1,163 customers have the status of active while 311 customers have the status of churn. They used CART and C5.0 to identify significant customer characteristics to predict churn.

TABLE I. COMPARISON OF PREVIOUS RESEARCH

RESEARCH	DATA MINING TECHNIQUE	ALGORITHMS
• Cimpoeru and Anca, 2014	classification	the Decision Tree and Naïve Bayes
• Manjit Kaur et al., 2013	classification	naive Bayes, decision trees and support vector machine
• Sharma et al., 2013	classification	neural network
• Correa Bahnsen et al., 2015	classification	decision tree logistic regression
• Praveen et al., 2015	classification	random forest decision tree
• Devi P. U., Madhavi S. 2012	classification	decision tree

III. PROPOSED METHODOLOGY

The framework can be subdivided into three phases. They are input, processing and output phases. The description of the framework is as follows:

Phase1: Data pre-processing

Sample data set in the IBM Watson Analytics community, this data set gives information to help predict behavior so as to retain customer, since a telecommunications company is concerned about the customers who leave their landline business for cable competitors.

Then, the data is entered into another level which is the pre-processing where data is converted and cleaned in a way that makes it suitable for use.

Phase2: Processing

Using the Data Mining techniques, the authors will focus on the technique where "hidden information" can be assembled from the dataset by using data mining tools. We can make prediction through the DM classification model. Several experiments will be conducted for feature selection and classification from selected customer churn dataset to compare its usefulness among the different feature selections and classifications, by using a data mining tool and comparing them with each other to choose the one that gives the most accurate results. And detection through DM clustering and association clustering models will be used to classify customers into clusters with shared characteristics, by using different algorithms and comparing the results. The purpose of association rule extraction is to define significant relationships between items or features that occur repeatedly in the dataset. The algorithms used in each model will be discussed later after the analysis.

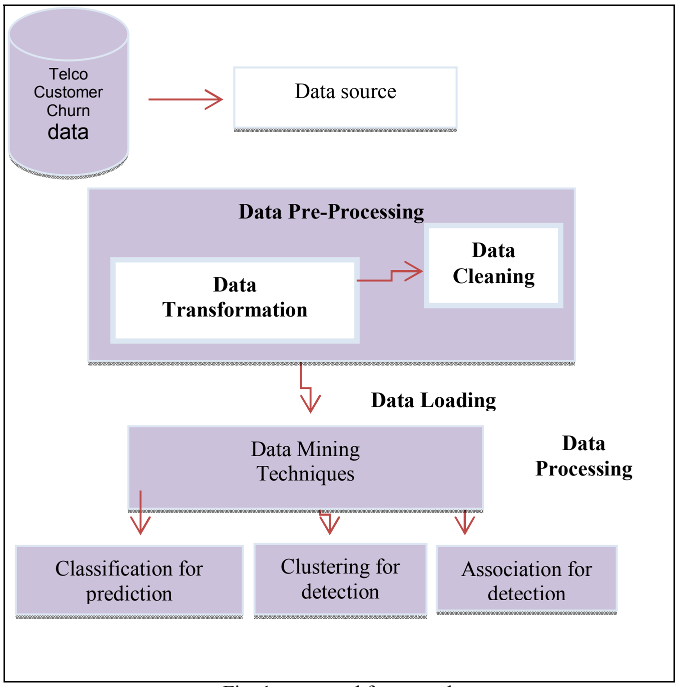

Fig. 1. proposed framework

IV. CASE STUDY

The main aim of this study is to design a model for early prediction of the customer churn through exploiting Data mining techniques, to find a near optimal solution that could help decision makers, in a business firm, to take the right decision in time and this, in its turn, will be effective in lowering churn rate. Also, recognizing the churn beforehand and taking proper measures to keep the customers would increase the firm's overall profitability. Losing customers usually results in opportunity cost, due to reduced sales and an increased need for attracting new customers which is five to six times more costly than customer retention.

Data set description:

It embodies information about:

- Customers who left within the previous month – the column is called Churn.
- Services each customer has signed up for, such as phone, multiple lines, internet, online security, online backup, device protection, tech support, and streaming TV and movies.
- Customer account information such as the period they've been a customer, contract, payment method, paperless billing, monthly charges, and total charges
- Demographic information about customers including gender, age range, and if they have partners and dependents.

V. APPLIED EXPERIMENT

This section describes how the experiment will be applied through the following phases

5.1 Data Pre-processing:

The data set passes through different levels in pre-processing stage:

First: Converting the data from string values to binary values

customerID	gender	SeniorCitizen	Partner	Dependents	tenure	PhoneService	MultipleLines	InternetService	OnlineSecurity	OnlineBackup	DeviceProtection	TechSupport
7590-VHVEG	Female	0	Yes	No	1	No	No phone service	DSL	No	Yes	No	No
5575-GNVDE	Male	0	No	No	34	Yes	No	DSL	Yes	No	Yes	No
3668-QPYBK	Male	0	No	No	2	Yes	No	DSL	Yes	Yes	No	No
7795-CFOCW	Male	0	No	No	45	No	No phone service	DSL	Yes	No	Yes	Yes
9237-HQITU	Female	0	No	No	2	Yes	No	Fiber optic	No	No	No	No
9305-CDSKC	Female	0	No	No	8	Yes	Yes	Fiber optic	No	No	Yes	No
1452-KIOVK	Male	0	No	Yes	22	Yes	Yes	Fiber optic	No	Yes	No	No
6713-OKOMC	Female	0	No	No	10	No	No phone service	DSL	Yes	No	No	No
7892-POOKP	Female	0	Yes	No	28	Yes	Yes	Fiber optic	No	No	Yes	Yes
6388-TABGU	Male	0	No	Yes	62	Yes	No	DSL	Yes	Yes	No	No
9763-GRSKD	Male	0	Yes	Yes	13	Yes	No	DSL	Yes	No	No	No
7469-LKBCI	Male	0	No	No	16	Yes	No	No	No internet service	No internet service	No internet service	No internet service
8091-TTVAX	Male	0	Yes	No	58	Yes	Yes	Fiber optic	No	No	Yes	No
0280-XJGEX	Male	0	No	No	49	Yes	Yes	Fiber optic	No	Yes	Yes	No
5129-JLPIS	Male	0	No	No	25	Yes	No	Fiber optic	Yes	No	Yes	Yes
3655-SNQYZ	Female	0	Yes	Yes	69	Yes	Yes	Fiber optic	Yes	Yes	Yes	Yes
8191-XWSZG	Female	0	No	No	52	Yes	No	No	No internet service	No internet service	No internet service	No internet service
9959-WOFKT	Male	0	No	Yes	71	Yes	Yes	Fiber optic	Yes	No	Yes	No
4190-MFLUW	Female	0	Yes	Yes	10	Yes	No	DSL	No	No	Yes	Yes
4183-MYFRB	Female	0	No	No	21	Yes	No	Fiber optic	No	Yes	Yes	No
8779-QRDMV	Male	1	No	No	1	No	No phone service	DSL	No	No	Yes	No
1680-VDCWW	Male	0	Yes	No	12	Yes	No	No	No internet service	No internet service	No internet service	No internet service
1066-JKSGK	Male	0	No	No	1	Yes	No	No	No internet service	No internet service	No internet service	No internet service
3638-WEABW	Female	0	Yes	No	58	Yes	Yes	DSL	No	Yes	No	Yes

Fig. 2. Sample of the data before preprocessing.

Second: Data cleaning through removing any missing values from any records

customerID	gender	SeniorCitizen	Partner	Depender	tenure	PhoneSer	MultipleLi	InternetS	OnlineSec	OnlineBac	DevicePrc	TechSupp	Streaming	Streaming	Contract	PaperlessBilling	PaymentMethod
7590-VHVEG	0	0	1	0	1	0	0	1	2	1	2	2	2	2	0	1	5
5575-GNVDE	1	0	0	0	34	1	2	1	1	2	1	2	2	2	1	0	6
3668-QPYBK	1	0	0	0	2	1	2	1	1	1	2	2	2	2	0	1	6
7795-CFOCW	1	0	0	0	45	0	0	1	1	2	1	1	2	2	1	0	3
9237-HQITU	0	0	0	0	2	1	2	2	2	2	2	2	2	2	0	1	5
9305-CDSKC	0	0	0	0	8	1	1	2	2	2	1	2	1	1	0	1	5
1452-KIOVK	1	0	0	1	22	1	1	2	2	1	2	2	1	2	0	1	4
6713-OKOMC	0	0	0	0	10	0	0	1	1	2	2	2	2	2	0	0	6
7892-POOKP	0	0	1	0	28	1	1	2	2	2	1	1	1	1	0	1	5
6388-TABGU	1	0	0	1	62	1	2	1	1	1	2	2	2	2	1	0	3
9763-GRSKD	1	0	1	1	13	1	2	1	1	2	2	2	2	2	0	1	6
7469-LKBCI	1	0	0	0	16	1	2	2	0	0	0	0	0	0	2	0	4
8091-TTVAX	1	0	1	0	58	1	1	2	2	2	1	2	1	1	1	0	4
0280-XJGEX	1	0	0	0	49	1	1	2	2	1	1	2	1	1	0	1	3
5129-JLPIS	1	0	0	0	25	1	2	2	1	2	1	1	1	1	0	1	5
3655-SNQYZ	0	0	1	1	69	1	1	2	1	1	1	1	1	1	2	0	4
8191-XWSZG	0	0	0	0	52	1	2	2	0	0	0	0	0	0	1	0	6
9959-WOFKT	1	0	0	1	71	1	1	2	1	2	1	2	1	1	2	0	3
4190-MFLUW	0	0	1	1	10	1	2	1	2	2	1	1	2	2	0	0	4
4183-MYFRB	0	0	0	0	21	1	2	2	2	1	1	2	2	1	0	1	5
8779-QRDMV	1	1	0	0	1	0	0	1	2	2	1	2	2	1	0	1	5
1680-VDCWW	1	0	1	0	12	1	2	2	0	0	0	0	0	0	1	0	3
1066-JKSGK	1	0	0	0	1	1	2	2	0	0	0	0	0	0	0	0	6
3638-WEABW	0	0	1	0	58	1	1	1	2	1	2	1	2	2	2	1	4

Fig. 3. Sample of the data after preprocessing.

5.2 Data processing:

The main objective of this study is to build a model for early prediction of the customer churn by using Data mining techniques, to produce a near optimal solution that could help managers to take a suitable decision at the right time, thereby reducing churn rate. To a firm, anticipating the customer churn beforehand and taking action to keep the customers would increase the firm's overall profitability, while losing customers would cause opportunity cost owing to reduced sales and the high cost of attracting new customers, which is five to six times more expensive than customer retention.

5.2.1 Classification:

Classification is a technique concerned with data examination. It is used in extracting model describing critical data classes. This model is called classifier. They anticipate absolute (discrete, unordered) class labels. Specialists have proposed various classification methods such as machine learning, pattern recognition, and statistics. Most of the algorithms are memory occupant, thereby accepting a little data measure. A late data mining research, based on such work, has created versatile classification and prediction techniques equipped for taking care of a highly massive amount of data. Classification has different applications such as extortion discovery, target promoting, execution prediction, assembling, and medicinal analysis. Data classification is a process carried out in two steps; they are a 'learning step' (where a classification model is structured) and a 'classification step' (where the model is used to predict class labels for given data). (Han et al., 2011)

The authors have used different algorithms to test and train the data using cross validation method with the aid of Matlab 2016 data mining software which are multi-layer perceptron, back propagation, decision tree, logistic regression, and support vector machine.

5.2.2 Clustering:

In this technique, the groups are not previously known. Instead, the algorithms are used to analyse the input data patterns and identify the natural groupings of records or cases. When new cases are scored by the generated cluster model they are assigned to one of the revealed clusters.

The authors applied the k-mean and DB-scan algorithms, using Rapidminner data mining software.

5.2.3 Association rule:

This technique belongs to the class of unsupervised modelling. They do not contain direct prediction of a single field. As a matter of fact, all the fields included have a double role. They act as inputs and outputs at the same time. Association model detects associations between discrete events, products, or attributes. Sequence models detect associations over time.

The algorithms which are used are Apriori-type and FP-Growth algorithms applied on Weka data mining software.

VI. RESULTS

6.1 Classification:

6.1.1 Multi-layer perceptron:

TABLE II. MULTI-LAYER PERCEPTRON RESULTS OF THE EXPERIMENT

Epoch	**Hidden layers**	**Perf.**	**feature**	**Correl /per.**	**feature**	**PCA/ per.**	**feature**
2500	[9 6 3]	86.8 %	19	86.4%	12	85.5%	8
	[30 1]	86.5 %		86.2%		84.4%	
	[10 1]	86.9 %		85.8%		86.6%	
	[10 5 3]	86.4 %		85.5%		85%	
2000	[9 6 3]	87%		86.8%		85.3%	
	[30 1]	86.5 %		85.2%		82.7%	
	[10 1]	86.5 %		86.3%		86.2%	
	[10 5 3]	86%		86.1%		85.4%	
1500	[9 6 3]	85.5 %		86.4%		85.1%	
	[30 1]	85.4 %		85.1%		86%	
	[10 1]	86.4 %		86.1%		85%	
	[10 5 3]	86%		86.3%		85.9%	

6.1.2 Back propagation:

TABLE III. BACK PROPAGATION RESULT OF THE EXPERIMENT

Epoch	**Hidden layers**	**Perf.**	**Original feature**	**Correl/per.**	**feature**	**PCA /per.**	**feature**
2500	[9 6 3]	85.1 %	19	84.6 %	12	85.9 %	8
	[30 1]	86.1 %		84.8%		85.7 %	
	[10 1]	86.3 %		84.9%		84.7 %	
	[10 5 3]	86.2 %		84.8%		85.8 %	
2000	[9 6 3]	86.2 %		84.7%		85.9 %	
	[30 1]	86.3 %		84.9%		85.7 %	
	[10 1]	86.2 %		84.8%		85.8 %	
	[10 5 3]	86.4 %		84.7%		85.7 %	
1500	[9 6 3]	86.3 %		84.6%		85.7 %	
	[30 1]	86.2 %		85%		85.7 %	
	[10 1]	85.2 %		84.7%		85.7 %	
	[10 5 3]	86.9 %		84.6 %		85.6 %	

6.1.3 Decision tree (DT):

TABLE IV. DECISION TREE RESULTS OF THE EXPERIMENT

performance	**Original feature**	**Correlation/per.**	**feature**	**PCA /per.**	**feature**
78.3%	19	76%	12	78.5 %	8

6.1.4 Support Vector Machine(SVM) :

TABLE V. SVM RESULTS OF THE EXPERIMENT

performance	**Original feature**	**Correlation/per.**	**feature**	**PCA /per.**	**feature**
79.9%	19	77.6%	12	79.3 %	8

6.1.5 Logistic regression:

TABLE VI. LOGISTIC REGRESSION RESULTS OF THE EXPERIMENT

performance	Original feature	Correl./ per.	feature	PCA/per.	feature
80.1%	19	77.5%	12	79.1%	8

6.2 Clustering:

6.2.1 DB-Scan clustering:

TABLE VII. CLUSTER 0 USING DB-SCAN

	Churn		
Cluster	**no**	**yes**	**Grand Total**
cluster_0	4474	1250	5724
Grand Total	4474	1250	5724

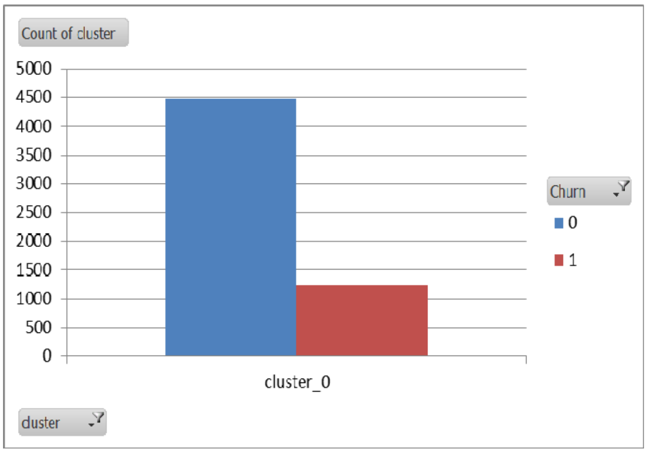

Fig. 4. Cluster_0 using DB-scan

TABLE VIII. CLUSTER_1 USING DB-SCAN

	Churn		
cluster	**no**	**yes**	**Grand Total**
cluster_1	700	619	1319
Grand Total	700	619	1319

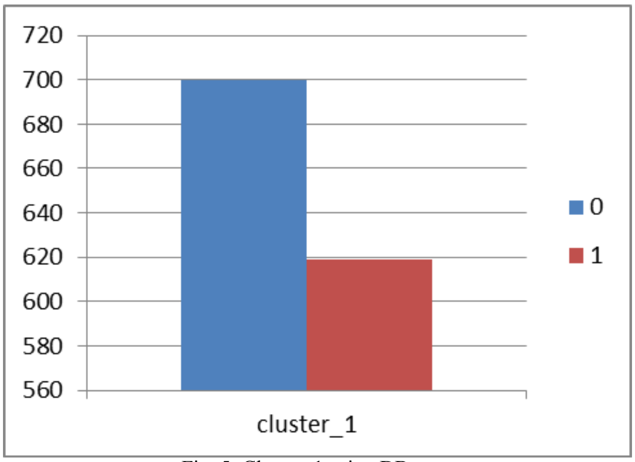

Fig. 5. Cluster_1 using DB-scan

TABLE IX. DB-SCAN RESULTS

	Churn		
Count of cluster	**No**	**Yes**	**Grand Total**
cluster_0	4474	1250	5724
cluster_1	700	619	1319
Grand Total	5174	1869	7043

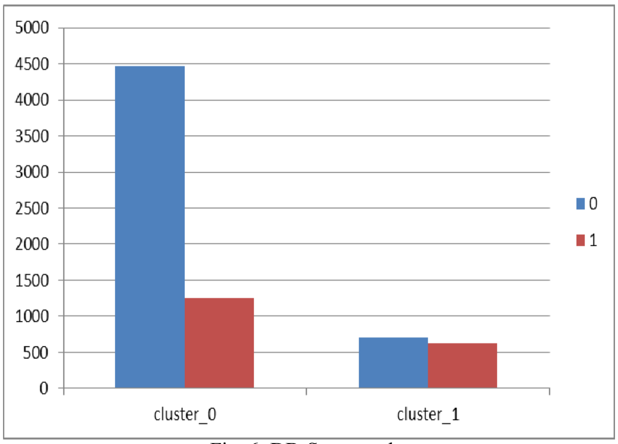

Fig. 6. DB-Scan results

6.3 Kmean clustering:

TABLE X. CLUSTER_0 USING K-MEAN CLUSTERING

	Churn		
cluster	**no**	**yes**	**Grand Total**
cluster_0	1768	321	2089
Grand Total	1768	321	2089

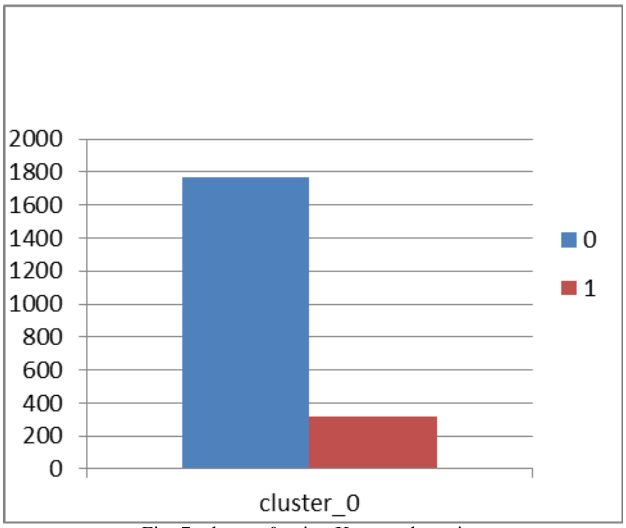

Fig. 7. cluster_0 using Kmean clustering

TABLE XI. CLUSTER_1 USING KMEAN CLUSTERING

Count of cluster	**Churn**		
cluster	**no**	**yes**	**Grand Total**
cluster_1	3406	1548	4954
Grand Total	3406	1548	4954

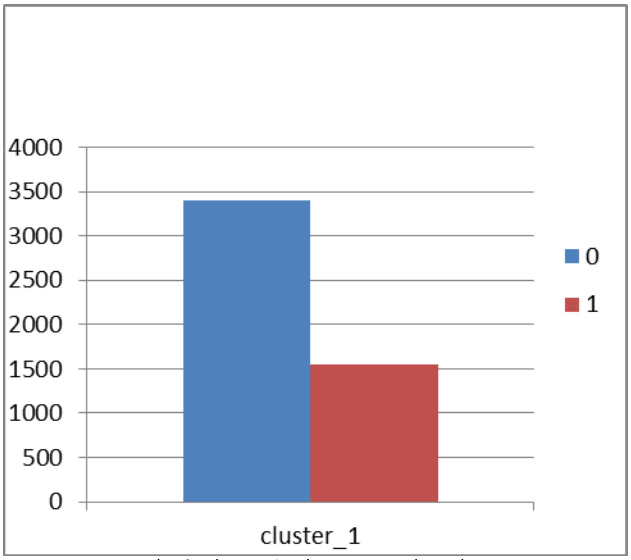

Fig. 8. cluster_1 using Kmean clustering

TABLE XII. KMEAN CLUSTERING RESULTS OF THE EXPERIMENT

Count of cluster	**Churn**		
cluster	**no**	**yes**	**Grand Total**
cluster_0	1768	321	2089
cluster_1	3406	1548	4954
Grand Total	5174	1869	7043

Fig. 9. Kmean clustering results of the experiment

6.4 Association rule:

6.4.1 Aperiori:

1. InternetService= fiber optic 4622 ==> PhoneService=yes 4622 conf :(1)
2. SeniorCitizen=0 InternetService=2 3739 ==> PhoneService=yes 3739 conf :(1)
3. MultipleLines=no 3390 ==> PhoneService=yes 3390 conf :(1)
4. Dependents=no InternetService=fiber optic 3317 ==> PhoneService=yes 3317 conf :(1)
5. InternetService=fiber optic Churn=no 3212 ==> PhoneService=yes 3212 conf :(1)
6. SeniorCitizen=0 MultipleLines=no 3017 ==> PhoneService=yes 3017 conf :(1)
7. MultipleLines=yes 2971 ==> PhoneService=yes 2971 conf :(1)
8. InternetService=fiber optic PaperlessBilling=yes 2841 ==> PhoneService=yes 2841 conf :(1)
9. SeniorCitizen=0 InternetService=fiber optic Churn=no 2727 ==> PhoneService=yes 2727 conf :(1)
10. InternetService=fiber optic Contract=month-to-month 2652 ==> PhoneService=yes 2652 conf :(1)

6.4.2 FP-Growth:

1. [InternetService= DSL]: 4622 ==> [PhoneService=yes]: 4622 conf :(1)
2. [MultipleLines=no]: 3390 ==> [PhoneService=yes]: 3390 conf :(1)
3. [MultipleLines=yes]: 2971 ==> [PhoneService=yes]: 2971 conf :(1)
4. [InternetService= DSL, PaperlessBilling=yes]: 2841 ==> [PhoneService=yes]: 2841 conf :(1)

5. [InternetService= DSL, Contract=month-to-month]: 2652 ==> [PhoneService=yes]: 2652 conf :(1)
6. [InternetService= DSL, gender=male]: 2322 ==> [PhoneService=yes]: 2322 conf :(1)
7. [InternetService= DSL, OnlineSecurity=no]: 2257 ==> [PhoneService=yes]: 2257 conf :(1)
8. [InternetService= DSL, TechSupport=no]: 2230 ==> [PhoneService=yes]: 2230 conf :(1)
9. [InternetService= DSL, Partner=yes]: 2234 ==> [PhoneService=yes]: 2234 conf :(1)
10. [InternetService= DSL, MultipleLines=no]: 2342 ==> [PhoneService=yes]: 2342 conf :(1)

VII. Experimental evaluation

As already mentioned, this research has applied the prediction model of the proposed customer churn through using different points of view, related to the classification technique. I used different algorithms to measure the accuracy of predicting who of the customers will leave the organization. Back propagation and Multi-layer perceptron have been effective towards giving high rate of preciseness over SVM, DT and LR, different tests have been carried out in each algorithm by changing the hidden layers as well as the number of epochs.

The algorithms that were used in clustering have been effective in categorizing the data set into two clusters. The algorithm that was used to do the clustering is DB scan. It shows that in cluster 0 the number of non-churner is much more than that of churner 78.16% and 21.84% , whereas in K-mean clustering algorithms it reveals that in cluster 1 the non-churners are much more than churners 68.75% and 31.25%.

Concerning the association rule techniques, the authors used the two algorithms; Aperiori and FP- growth algorithms in an attempt to show the customer behaviour to those who offer certain services. Each algorithm served to test ten rules as mentioned previously.

VIII. Conclusion

This research presents a model that can foretell which customer will leave the organization and who will not. To do this, the authors used different data mining techniques.

Using various techniques helped me to accurately predict customer churn as well as defining the causes that lead to customer retention. Each of the techniques was of great help towards detecting and predicting customer churn as regards classification, back propagation and multi-layer perceptron and they gave a high rate of preciseness.

As for the clustering technique, it has proved that DB-scan clustering is suitable and effective in clustering the data set into two categories together with giving high percentage of non-churners over churners at the cluster-0. As regards the association rule angle Aperiori and FP-Growth, they have been applied to specify the behavior of each customer and that is supposed to help the business organization to focus on certain services or attributes to stop customer churn from happening.

References

[1] Catalin CIMPOERU, Anca ANDREESCU.2014. Predicting Customers Churn in a Relational Database, Informatica Economică vol. 18, no. 3:5-16

[2] Manjit Kaur et al., 2013.Data Mining as a tool to Predict the Churn Behaviour among Indian bank customers, International Journal on Recent and Innovation Trends in Computing and Communication, Volume: 1 Issue: 9

[3] Anuj Sharma et al., 2011. A Neural Network based Approach for Predicting Customer Churn in Cellular Network Services, International Journal of Computer Applications (0975 – 8887) Volume 27– No.11

[4] Correa Bahnsen et al. Decision Analytics (2015), a novel cost-sensitive framework for customer churn predictive modeling2:5

[5] Praveen et al., Churn Prediction in Telecom Industry Using R, International Journal of Engineering and Technical Research (IJETR) ISSN: 2321-0869, Volume-3, Issue-5, May 2015

[6] Devi P. U., Madhavi S. - Prediction of Churn Behavior of Bank Customers Using Data Mining Tools, Business Intelligence Journal - January, 2012 Vol.5 No.1

[7] Han, Jiawei, Jian Pei, and Micheline Kamber. Data mining: concepts and techniques. Elsevier, 2011.