

A Novel Framework for Customer Churn Prediction Using Rough Set Theory

Ankur Vashist,

Department of Computer Science,
MRIIRS, Faridabad,
Haryana, India

Yash Aggarwal

Department of Computer Science,
MRIIRS, Faridabad,
Haryana, India

Smriti Gupta

Department of Computer Science,
MRIIRS, Faridabad,
Haryana, India

Vinay Jain

Accendere Knowledge
Management Services Pvt. Ltd.,
India

Abstract

Churn prediction is emerging as a great influential field in the category of customers. Companies are investing a huge amount of capital in the process of holding the customers, i.e., preventing them from a churn. Every company likes to increase its number of customers as it is directly proportional to the profit and sales of the company. But at the same time the company prefers to hold the customers who have purchased their product at least once, for that company requires some kind of prediction that depicts when the customers will churn. Accuracy, comprehensibility, and justifiability are three key aspects of a churn prediction model. This paper proposed a rough set based framework for Customer Churn Prediction using on customer reviews dataset. The utilization of presented framework help companies in making intelligent decision support system.

Keywords: Churn prediction, data mining, E-commerce, Sentiment analysis, Rough sets

Introduction

Churn prediction can be defined as the probability of a customer to stop doing the business with the particular company/vendor in the given time period (Chandar et al., 2006). Every company likes to increase its number of customers as it is directly proportional to the profit and sales of the company. But at the same time the company prefers to hold the customers who have purchased their product at least once, for that company requires some kind of prediction that depicts when the customers will churn (Ahn et. al., 2011). If the company gets an idea about the churning of the customer it might make an offer for the customer to hold him/her. Since, we have data in every field of work and also we have the customers or the users in every business, therefore predicting their churn becomes a necessary task. Customer Relationship Management (CRM) is one of the most comprehensive strategies to enhance the relations with the customers (Pang & Lee, 2008). It is broadly acknowledged and is being used today in the fields like telecommunication (Yu-Teng, 2009), e-commerce etc (Keramati et al., 2009).

According to Hangxia (2009), accuracy, comprehensibility, and justifiability are three key aspects of a churn prediction model. An accurate model permits to correctly target future churners in a retention marketing campaign, while a comprehensible and intuitive rule-set allows identifying the main drivers for customers to churn, and to

develop an effective retention strategy in accordance with domain knowledge(Ahn et. al.,2011).

The theory of rough sets has been under continuous development and a fast growing group of researchers and practitioners are interested in this methodology. The theory has discovered numerous fascinating applications in solution, pharmacology, business, marketing, statistical surveying, building plan, meteorology, vibration investigation, exchanging capacity, strife examination, picture preparing, Human computer interaction, simultaneous framework investigation, choice examination, character recognition, and different heterogeneous fields.

This paper provides comprehensible customer churn prediction models based on comprehensible and intuitive rule-set using data mining techniques.

Related Work

There are two most popular approaches to churn modeling are machine learning techniques and survival analysis. Every modeling technique requires distinct data structures and feature selection procedures(Ahn et. al.,2011). Ultimately, there is no single churn methodology that is proven to work in most situations; either machine learning models or survival regression could be appropriate based on the application.

It is shown that only limited attention has been paid to the comprehensibility and the intuitiveness of churn prediction models. The definition of churn is totally dependent on your business model and can differ widely from one company to another.

Table 1 Churn prediction models for different domains

Authors	Domain Datasets	Prediction Technique
Smith <i>et al.</i> (2000)	Insurance company	Neural Network; Logistic regression; Decision tree
Mozer <i>et al.</i> (2000)	Wireless Telecom Industry	Neural Network; Logistic regression
Wei and Chiu(2002)	Taiwan mobile Company	Decision tree
Chiang <i>et al.</i> (2003)	Network banking	Association Rule
Kim and Yoon(2004)	Korea mobile carriers	Binomial logit model
Larivie`re and Van den Poel (2005)	Belgian financial Services	Random forests; Regression forests; Linear regression; Logistic regression
Nie <i>et al.</i> (2006)	Charge Email	Decision tree
Luo <i>et al.</i> (2007)	Personal Handy- phone System Service	Neural Network; Decision tree
Burez and Van den Poel(2008)	Pay-TV company	Random forests Survival analysis
Tsai and Lu (2009)	American Telecom Companies	Hybrid Neural Network;

Rough Set Theory

Rough set theory can be regarded as the tool for the imperfect data analysis. This theory has paved its way in many fields like, engineering, decision support, banking, pharmacy etc. The assumption that paved the way for this theory was that every data set in this universe contains some information/data associated with it.

In this theory with every rough set, a pair of precise set known as the lower and the upper approximation is associated. All the objects that are in the lower approximation mark their presence in the set for sure and those who are in the upper approximation set are dicy about their existence in the set. The difference between the upper and the lower approximation gives rise to the boundary

region of the rough set. Approximation is one of the fundamental blocks of the rough set theory.

Analysis starts with the formation of the table known as the decision table. In this table the columns represents the attributes and the rows represents the objects. Attributes are further classified as condition and decision attributes. Each row generates a new rule which is termed as decision rule, which specifies the decision if some conditions stand positive. If the rule determines a decision in a unique manner in terms of the conditions provided then it is called certain rule else it is termed as uncertain by Miao et al.(2009).

In 1991, Rough Set(RS) Theory was given by Pawlak (1991). In an Information System there is an existing

Equivalence Class Partition(ECP) among their relationship this is analyzed by RS . To constitute the basis of the theory Two core concept are created, discernibility relation and approximation. Through RS theory an approximate equivalent IS with smaller scale are extracted for significant attributes. Processes to find concise equivalent IS to the original one are done by RS algorithm. Finding reducts is the process known. RS based applications in artificial intelligence decision tables are used with this. Reduction as patterns of decision table based on machine learning algorithm can construct classifiers to categorize new objects. However, NP-hard problem is finding minimal reduction of decision table given by Pawlak (1991).

Table represents objects and their attributes, each row of which is an object(e.g., a man or woman, an observed object, a test) and some values based on a certain measurement describes each column as an attribute. The table is called an information system. Formally, an information system is a pair $I = (U, A)$, where U and A are both non-empty, finite sets. U is a set of objects, called a universe. A is a set of attributes. $\forall a \in A, a: U \rightarrow V_a$, where V_a is the set of all possible values of a .

Redundant information, such as repeated rows, indiscernible rows, dependable attributes(attributes can induced by these),are contained by them and information system built on real world data. In practice it is very useful to infer a more concise equivalent information system

given by Qian et al.(2008). There are fewer rows and columns in equivalent information system that can be considered a symbol of original one. Reduct is the name given to this pattern in Rs theory. There are two concepts to obtain reducts:

- 1) Indiscernibility relation
- 2) Set of approximation.

An equivalence relation is a indiscernibility. The set of attributes is defined by equivalence relation. Formally, let $I = (U, A)$ be an information system. $\forall B, B \subseteq A$, defines an associated binary relation $I(B)$ on $U: I(B) = \{(x, y) \in U \times U \mid \forall a \in B, a(x) = a(y)\}$, which is called B-indiscernible.

If $(x_1, x_2) \in I(B)$, then we say x_1 and x_2 are indiscernible with respect to B . The equivalence class of the B-indiscernible are denoted by $B(x)$ or $[x]_B$ (Tripathy et al.,2011).

Data Collection

Data collection is the basic step towards any prediction(Jain & Kumar,2015).It lays the foundation for the prediction or can say that it initiates the process for the prediction. For the data collection, data scrapping technique has been used for collecting the data of customer reviews from Mouthshut.com].

Table 2 A sample of reviews and their corresponding sentiment/polarity

Customer	Rating	Review	Polarity	Review Sentiment	Positivity	Negativity
1	1	Horrible experience. Booked a Volvo Bus from Manali to Delhi through RedBus.	0.5	negative	0.2	0.8
2	1	Red Bus is totally a fraud company, they have recruited best liars, dumb and fraud from India in their company.	0.1	neutral	0.5	0.5
3	4	I was entering the redbus.in so speed of server and give the details about the buses also good.	0.8	positive	0.1	0.9
4	1	terrible sevies in providind seat alloatment, never to trust redus and ND travels	0.4	neural	0.3	0.7
.	.	.	.	.	.	.
.	.	.	.	.	.	.
.	.	.	.	.	.	.

n	1	I traveled to Porbandar from Ahmedabad in sudama mahasagar travels.pathetic bus design .I booked J AND K upper sleeper. Hopeless seat it is. sleep there.it was like nightmare.maximum small blowers has broken so they were making your respiration more difficult .don't go by sudama and don't book J and k seat	0.5	neutral	0.5	0.5
---	---	---	-----	---------	-----	-----

Proposed framework

The proposed method is based on the identification of relevant attributes for churn prediction and analyzed using Rough set theory for identification of churners. The major steps for the proposed framework illustrated in Figure 1 have been given below.

Step 1: Data Collection

Data collection is the basic step towards any prediction. It lays the foundation for the prediction or can say that it initiates the process for the prediction. For the data collection, data scrapping technique has been used for collecting the data of customer reviews from Mouthshut.com.

Step 2: Data Preprocessing

Every word present in customer review is prominent in

decision making. Thus, pre-processing is required to decrease the noise present in review and filter out irrelevant words (Khan et al., 2016). Pre-processing steps such as tokenization, stop word removal, stemming, lemmatization, feature weighting, dimensionality reduction and frequency based methods were applied for normalization of the data.

Step 3: Sentiment Analysis

Sentiment analysis is a technique that helped us to analyze the reviews of the customers for the dedicated field of data. A review can be classified as the positive, negative or neutral based upon the experience of the customer. Sentiments of the reviews play a major dominant role in the prediction of the churn for the customer. Table 2 shows the polarity of the review, which is also a major factor in churn prediction.

Fig 1. Proposed framework for customer churn prediction using Rough sets



Step 4: Rough set based analysis

Consider a customer information system represented in Table 3 of 'n' customer/users 'ci' where, $i = (1, 2, \dots, n)$ as the set of objects of the universe with a set of attributes given in Table 3. The attribute churn is considered as the decision attribute with value=1, represent churn and with

value=0, represent not churn. For particular, customer 'c1' is characterized in the table by the attribute value set (sentiment, positive), (Rating, <3), (Polarity, <0.5) etc which form the information about the particular user. The reduced customer information system is presented in Table 3.

Table 3 Customer information system

Comments	a1	a2	a3	an-2	an-1	an	Churn
c1	y	n	y	n	y	y	1
c2	n	y	n	n	n	n	0
c3	y	y	n	n	y	n	0
c4	y	n	y	n	n	y	1
c5	y	y	n	y	n	n	0
c6	y	y	n	n	y	y	1
c7	n	n	y	n	n	n	0
:	y	y	y	y	y	n	1
:	y	y	n	n	n	n	1
cn	y	n	n	y	y	y	1

Step 5: Rule generations

Decision rules are found using Jhonson algorithm and Genetic algorithm using Rosetta tool (Hvidsten, 2013).

Step 6: Churn prediction

Proposed approach provides rules for identification of customer who will be in position of churn.

Application of Customer Churn Prediction

Seat and Capacity Optimization

Prescient investigation enables carriers to foresee the quantity of cancelations and no-appears for each flight and to offer recently opened seats at balanced costs. What's more, aircrafts can arrange this data with freight dealing with to hold sacks of no-appears and get flights noticeable all around on time. By taking a gander at traveler characteristics, flight classes, air terminal terminals, plane writes, time of day, and flight limit, investigation offers absent likelihood go for each flight, helping income bookkeeping offices make and take after an educated overbooking strategy.

Price Optimization

Prescient examination takes a gander at deals execution designs and takes into consideration both short and long haul valuing procedures, and permits carrier income divisions to inspect the impacts of advancements after some time. By taking a gander at the advancement channels, directed clients, and that's just the beginning, aircrafts can grow more complex procedures and keep away from squandered endeavors later on. This joined with request expectation, regular patterns, and different elements makes for a considerably more individualized understanding for each kind of explorer, and empowers up-to-the-minute evaluating changes that stay lined up with general organization system.

Loyalty Program Optimization

One carrier, seeing that its reliability program individuals were leaving in huge numbers, looked to realize why this was. They contributed broad time and cash on manual research, and found that reliability individuals didn't feel they had anything to pick up by remaining on – however this was just piece of the photo. By applying prescient examination, they could conjecture stir before it happened and take educated activities by understanding what kind of a reaction each activity would get. This kind of two-advance forecast does ponders for diminishing beat, and can likewise enable carriers to make extreme conclusions, for example, which faithfulness clients to offer open seats to.

Campaign Management

Prescient investigation is a reference point of light for aircraft income bookkeeping divisions. It requires a lot of information, be that as it may, and also information refinement and examination. By banding together with an IT seller that can accumulate, store, and parcel information, aircrafts will be extraordinarily situated to use prescient investigation and increment benefit.

Understand customer behavior

Through prescient examination we would analysis be able to it's anything but difficult to foresee whether a client will agitate or not and the organization dealing with it can without much of a stretch find a way to lessen it and make helpful measure and influence their client to tie up with their item. It additionally encourages different organizations to check whether the client is agitating or not if yes then they can offer arrangements for them and make client.

Conclusion

In this paper rule generation techniques of Rough Set Theory has been are applied for identification of customer churn using customer review data. Multiple relevant words (attributes) reflect customers churn in different ways so it becomes difficult to filter out the most relevant attributes which predict churn. The paper proposed a framework which can be used as a decision system for identification of customer who will be at the position of churning. The utilization of presented framework help companies in making intelligent decision support system.

References

- Ahn, H., Ahn, J.J., Oh, K.J., et al. (2011) Facilitating Cross-Selling in a Mobile Telecom Market to Develop Customer Classification Model Based on Hybrid Data Mining Techniques. *Expert Systems with Applications*, 38, 5005-5012
- Hangxia Ma, Min Qin, Jianxia Wang. (2009), "Analysis of the Business Customer Churn Based on Decision Tree Method", *The Ninth International Conference on Control and Automation*, Guangzhou, China.
- Hvidsten, T. R. (2013). A tutorial-based guide to the ROSETTA system: A Rough Set Toolkit for Analysis of Data.
- Keramati, A., Jafari-Marandi, R., Aliannejadi, M., et al. (2014) Improved Churn Prediction in Telecommunication Industry Using Data Mining Techniques. *Applied Soft Computing*, 24, 994-1012.

- Khan, K., Ullah, A. & Baharudin, B. (2016). Pattern and semantic analysis to improve unsupervised techniques for opinion target identification. *Kuwait Journal of Science*, 43(1):129-149.
- Jain, V. K. & Kumar, S. (2015). An effective approach to track levels of influenza-A (H1N1) Pandemic in India Using Twitter. *Procedia Computer Science*, 70(1):801–807.
- Miao, D., Duan, Q., Zhang, H. & Jiao, N. (2009). Rough set based hybrid algorithm for text classification. *Expert Systems with Applications*, 36(5):9168–9174.
- Pang, B. & Lee, L. (2008). Opinion mining and sentiment analysis. *Foundations and Trends in Information Retrieval*, 2(1):1–135.
- Pawlak, Z. (1991). *Rough sets: theoretical aspects of reasoning about data. Theory and decision library*. Kluwer Academic Publishers. Norwell, MA, USA.
- Qian, Y. H., Liang, J. Y., Li, D. Y. et al. (2008). Measures for evaluating the decision performance of a decision table in rough set theory. *Information Sciences*, 178(1):181-202.
- Tripathy, B. K., Acharjya, D. P. & Cynthia, V. (2011). A framework for intelligent medical diagnosis using rough set with formal concept analysis. *International Journal of Artificial Intelligence & Applications*, 2(2):1-14.
- Xia, G. (2010) Research on Current Situation and Development of Customer Churn Prediction. *Application Research of Computers*, 27, 413-416.
- Yu-Teng Chang, “Applying Data Mining To Telecom Churn Management”, *IJRIC*, 2009 67–77.