

Fuzzy data mining for customer loyalty analysis

Jay B.Simha
ABIBA Systems, Bangalore, INDIA.
jay.b.simha@abibasystems.com

Iyengar S.S
Department of Computer Science,
Louisiana State University, Baton Rouge,
USA., iyengar@bit.cse.lsu.edu

Abstract

One of the major conceptual analyses in analytical CRM is customer loyalty. Current methods use user defined and statistical methods for loyalty analysis. One of the major limitations of these methods is that they assign every customer into exactly one segment, which may be counter intuitive. In this paper a novel approach using fuzzy logic has been discussed. The proposed method uses fuzzy context model to extract the customer membership to a segment of interest. A real life case study shows promising results for the proposed approach.

1. Introduction

Telecom companies are intensely customer centric and depend heavily on analytical Customer Relationship Management (CRM) for increasing the profit. Increasing the loyalty of the customer is known to be better at improving the bottom lines. One of the methodologies widely used to analyze the loyalty is segmentation. Currently used methods have some limitations and in this paper an attempt has been made to demonstrate a novel approach to loyalty analysis using fuzzy logic.

2. Literature survey

Bharatheesh [1] used fuzzy data mining with context approach to identify the critical machines for maintenance. The approach has reported to improve the maintenance effectiveness significantly.

Meier et.al [3] used a similar approach for customer loyalty analysis and it has been incorporated into an e-commerce product. However they have not reported any real world application results and this is the motivation for this research.

3. Fuzzy data mining

For segmentation to be effective, an algorithm should be able to handle linguistic or fuzzy variables. This is because the ability to do so would allow some interesting patterns to be more easily discovered and expressed. For example, if crisp boundaries are defined for “high revenue”, there is a possibility that the records selected may dependent to a large extent on the definitions of boundaries. Despite its importance, many segmentation algorithms (e.g., k-means) were not developed to handle fuzzy data. They were used mainly to deal with categorical and quantitative attributes. In particular, when dealing with quantitative attributes, their domains are usually divided into distance based bins automatically, sometimes even allowing overlapping bins. In most cases, the resulting intervals are not too meaningful and are hard to understand.

Even though algorithms like fuzzy c-means can mitigate this problem to certain extent, by allowing partial membership to multiple segments, it will not capture the semantic intent of the query and relations in line with the current knowledge of the domain expert, there by creating a cognitive gap.

In this paper we use the context model. The main advantage of this approach compared to classical one is that the classification of the customer is based on the meta knowledge provided by the domain expert and it still allows the partial membership to multiple classes. The notion of partial membership gives a much better description of the classified elements and also helps to find out the potential or possible weakness of the considered element.

Fuzzy data mining using a context model is a relation over the attributes and their contexts [Bharatheesh]. A linguistically arrived segment over a context relation is a relation of the form,

$$F(K^*) = E \in \prod K_{ji}$$

Where, K^* defines the context of interest and K_{ji} defines the terms of the different attributes.

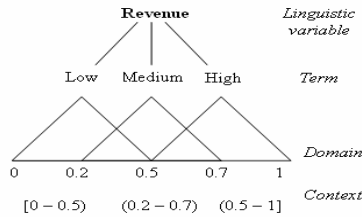


Fig 4. Context model

With this approach to data mining using context model, linguistic variables and membership functions, the classification space becomes fuzzy. Such a fuzzy partition will result in disappearance of sharp boundaries and establishes continuous transition. This results in any given customer belonging to multiple segments in varying degrees of membership. This precise information about the customer allows for better judgment of decisions.

The algorithm for identifying a record's (customer's) membership in a conceptual segment defined over the contextual terms is given in Fig 5. Complete details of the algorithm and experiment are available in [2].

Algorithm FDM:

Inputs:

- A set of records D
- Context details K
- Query Q

Outputs:

- Graded list of records D_g

Algorithm:

- For each attribute in D, normalize the data using Max division.
- For each record in D compute the membership in each of the terms in Q over the context K.
- Aggregate the approximate membership of each record to the context K.

Fig 5. Fuzzy data mining algorithm

4. Case study

In order to validate the proposed approach, data from a mobile telecom operator was used. The following contexts and the linguistic terms for the dimensions discussed in [4] were used as variables.

- *Behavior*: customer revenue
- *Commitment*: the age on network
- *Involvement with network*: number of services used.

For simplicity only three contexts for each of the linguistic variable were defined and simple triangular membership functions were considered for transforming the crisp values to relevant fuzzy subspaces. The telco wanted to know the customer behavior of its customer base to understand their

loyalty to the company. Based on these behavioral attributes, the profile of each of the context combinations were extracted using the fuzzy data mining algorithm. The results were compared with a classical segmentation using user defined crisp ranges. The results are shown in Table 1. It can be seen that the lift using proposed method outperforms the classical method on all counts.

Table 1. Comparative results

	Number of customers	Normalized expected revenue	Percentage lift
Context : High revenue, High AON and High Involvement			
Classical segmentation	1600	0.62	-
Fuzzy data mining	2100	0.85	37
Context : Low revenue, High AON and High Involvement			
Classical segmentation	4500	0.57	-
Fuzzy data mining	6200	0.78	36.8
Context : High revenue, Low AON and High Involvement			
Classical segmentation	1200	0.6	-
Fuzzy data mining	1425	0.7	16.6

5. Conclusions

Customer loyalty is one of important aspects in CRM. Classical methods of exploring the customer loyalty have some limitation in extracting the hidden relationships as they can not fully satisfy the semantic intent of the query. In this paper an attempt has been made to propose an alternative and novel approach to customer loyalty exploration using fuzzy data mining. The proposed method uses context model to explore the customer base and experiments on real world data sets are promising.

6. References

[1] Bharatheesh T.L., 2003, "Intelligent decision support and data mining for service resource planning", Unpublished Ph.D. thesis, Bangalore University

[2] Jay B.Simha, Iyengar S.S., 2006, "Fuzzy data mining for CRM", Working paper, Department of Computer Science, Louisiana State University, Baton Rouge, LS, USA

[3] Meier A, Werro N, Albrecht M, Sarakinos M, 2005, "Using a Fuzzy Classification Query Language for Customer Relationship Management", VLDB 2005: Trondheim, Norway, 1089-1096

[4] Shailendra K, 2005, "Understanding the Relationship Between loyalty and Churn", BI Reports, December 2005