Data Mining Using Fuzzy Theory for Customer Relationship Management

K.W. Wong

School of Information Technology
Murdoch University, Australia.
E-mail: k.wong@murdoch.edu.au

ABSTRACT

Customer Relationship Management (CRM) initiatives have gained much attention over the past few years. Although CRM involves technology, the important success factor involves strategy. This is the strategy of building business around the customer. However, with the aid of data mining techniques, business can formulate the strategy easier. Fuzzy theory allows human expertise and decisions to be modelled more closely, thus it is suggested that it can be used in the CRM model. A case study presented in this paper has examined two area of a typical CRM model, where the decision making process can be improved by using fuzzy theory.

Keywords: Data Mining, Fuzzy Logic, Fuzzy Clustering, Customer Relationship Management, and Segmentation

INTRODUCTION

When businesses first used computers to store data, data mining technology started to evolve as a new technology in navigating through the database. Its purpose is mainly helping businesses to focus on important and useful information, by extracting the hidden predictive information from large databases. Basically, data mining techniques perform these predictive features based on modelling. The known situation is used to build the model, and then it is applied to another situation where it is not known. The objective of the data mining algorithm is to automate the detection of relevant patterns in a large database. The commonly used techniques in data mining are artificial neural networks (Bigus, 1996), decision trees (Cherkauer and Shavlik, 1996; Gaines and Compton, 1995), genetic algorithms (Turney, 1995; Scott et al, 1997), nearest neighbour method (Hastie and Tibshirani, 1995), and rule induction (Agrawal and Srikant, 1994).

In recent years, Customer Relationship Management (CRM) initiatives have gained much attention. Although CRM involves technology, the important success factor involves strategy. This is the strategy of building business around the customer (Ward, 2001). However, with the aid of data mining techniques, business can formulate the strategy easier. This strategy is not to make sure that the customer is always right but to transform the business into a customer centric model. As the term CRM suggests, three main areas need to be taken care of: (1) Customer, (2) Relationship, and (3) Management (McLaughlin and Erickson, 2001; Nykamp, 2001). Customer is the most important person in a business, who requires a lot of effort in convincing them that the product and service are what they want. Data mining technology can enhance the understanding of the needs and background of the customer (Thearling, 1999). Relationship information is used by an organization to provide effective communications with the customer. They can be further broken down into channel of communication, segmentation, and differentiated treatment. Data mining techniques have to
be designed and used carefully in this area to identify and explore complex relationships that are difficult to realise through intuitive analysis. Management is required to effectively making use of the knowledge of the customer, the relationship established, and any other data collected.

One of the major parts in the relationship is the identification of segments. In marketing research, this is normally known as market segmentation (Wedel and Kamakura, 1998). Market segmentation normally tries to break down a heterogeneous market into a number of smaller homogeneous markets where special treatment and care can be used to address a more precise satisfaction factor of the customer needs. Segmentation can normally be classified as a-priori and post-hoc approaches (Wind, 1978). In this paper, data mining tool using fuzzy theory has been proposed in the segmentation stage as well as part of the CRM building block. A case study will be used to demonstrate how the fuzzy theory can be used in a typical CRM system.

DATA MinING IN CRM

In order to understand a customer, it has to start from analysing all the relevant data belonging to the customer, thus data mining is the intelligence behind a successful CRM strategy (Linoff, 1999). This technology is to transform data into useful information for business to focus on customers. There are basically two main types of data mining: descriptive and predictive. Descriptive data mining generates information about the data so that we can realise some interesting underlying information. Predictive data mining makes use of past patterns and information in predicting what a customer will buy in the future. With most of the present techniques as mentioned in the introduction and those presented in Berson et. al (2000), it is difficult to simultaneously perform these two analyses in the same model. All these methods are also insufficient to address uncertainties in the data.

There are five main steps in the process of implementing a successful data mining solution for CRM (Siragusa, 2001): setting goals, data collection, data preparation, analysis and prediction, and measurement and feedback. When setting the goals, identifying the market segmentation model is important. It can allow reasonable goals under each segment to precisely address the issues like retention, risk avoidance as well as possible cross selling (Siragusa, 2001). In data collection and preparation, it is important to address the issues like feature selection, parameter identification and handling of missing data (Siragusa, 2001). When building analysis and prediction models, different methods may have to be used in each different segment to meet the intended goals. The most crucial point in gaining business confidence in establishing a model is to avoid a total “black box” method that eliminates the contributions of the expert in the business.

FUZZY THEORY AND FUZZY CLUSTERING

Fuzzy theory works on the basis derived from fuzzy sets (Zadeh, 1973; Klir and Yuan, 1995). A fuzzy set allows for the degree of membership of an item in a set to be any real number between 0 and 1, this allows human observations, expressions and expertise to be modelled more closely. The membership function of a fuzzy set A is denoted by:

$$A : X \rightarrow [0,1]$$

Once the fuzzy sets have been defined, it is possible to use them in constructing rules for fuzzy expert systems and in performing fuzzy inference. Fuzzy system can produce more accurate results based on the basic idea of the defuzzification. A defuzzification technique is used to calculate the conclusion by evaluating the degree of matches from the observation that
triggered one or several rules in the model. This will lead to a better result by handling the fuzziness in the decision making. Thus, the fuzzy technique can improve the statistical prediction in certain cases.

Fuzzy sets allow human expertise and decisions to be modelled more closely, thus it is suggested in this paper that it can be used in the CRM model. A set of example data or knowledge from the business analyst is used as the basic knowledge available to build the fuzzy rule base. Using knowledge from the business analyst, fuzzy rules can be hand-coded into the CRM model. However, with the availability of the vast amount of data, it will be useful to extract knowledge from the data directly. This has the advantage of discovery new knowledge or relations underlying the data. In extracting fuzzy rules from the data, the first step is to translate all the available data into linguistic fuzzy rules using linguistic labels. The following algorithm outlines the steps in extracting the fuzzy linguistic rules from the available data. For \( k \) inputs, the given input-output data pairs with \( n \) patterns:

\[
(x_1, ..., x_{i}^k; \ y_1) \\
(x_{2}^1, ..., x_{2}^k; \ y_2) \\
\vdots \\
(x_{n}^1, ..., x_{n}^k; \ y_n)
\]

The number of linguistics terms \( T \) and the distribution of data in the regions of the whole domain are first determined. For ease of interpretation and computational simplicity, the shape of the membership function used in this algorithm is triangular. In this case, we will obtain for every \( x \in X \),

\[
A_i \in F(x) \rightarrow \{0,1\} \quad \text{for all} \ i \in T \quad (2)
\]

After the fuzzy regions and membership functions have been set up, the available data set will be mapped. If the value cuts on more than one membership function, the one with the maximum membership grade will be assigned to the value:

\[
R_n \Rightarrow [x_1^n (A_{i1}, \ \text{max}), ..., x_k^n (A_{ik}, \ \text{max}) : y^n (B_t, \ \text{max})] \quad (3)
\]

After all the data sets have been assigned with a fuzzy linguistic label, Mamdani type fuzzy rules are then formed and centroid defuzzification is used.

After fuzzy rules have been generated from each data point, repeated rules are removed. In the event that there are repeated fuzzy rules, the number of repetitions of the fuzzy rules and the firing strengths of the rules will be examined to resolve conflicts.

Besides using fuzzy theory in the data mining process of the CRM model, fuzzy clustering can also be used to segment the market. The fuzzy clustering can be used to modify a segmentation technique by generating a fuzzy score for each customer. This provides a more precise measure to the company in delivering value to the customer and profitability to the company. Given a set of data, clustering techniques partition the data into several groups such that the degree of association is strong within one group and weak for data in different groups. Classical clustering techniques result in crisp partitions where each data can belong to only one partition. Fuzzy clustering extends this idea to allow data to belong to more than one group. The resulting partitions are therefore fuzzy partitions. Each cluster is associated with a membership function that expresses the degree to which individual data belongs to the cluster. Fuzzy C-Means (FCM) clustering has been very reliable and popular in performing fuzzy clustering (Bezdek, 1981).

Given a set of data, FCM clustering iteratively search for a set of fuzzy partitions and the associated cluster centres that represent the structure of the data. The FCM clustering
algorithm relies on the user to specify the number of clusters present in the set of data to be clustered. Given the number of cluster \( c \), FCM clustering partitions the data \( X = \{x_1, x_2, \ldots, x_n\} \) into \( c \) fuzzy partitions by minimizing within group sum of squared error objective function using the following equation:

\[
J_m(U, V) = \sum_{k=1}^{n} \sum_{i=1}^{c} (U_{ik})^m \| x_k - v_i \|^2, \quad 1 \leq m \leq \infty \quad (4)
\]

where \( J_m(U, V) \) is the sum of squared error for the set of fuzzy clusters represented by the membership matrix \( U \), and the associated set of cluster centres \( V \). \( \| \cdot \| \) is some inner product-induced norm. In the formula, \( \| x_k - v_i \|^2 \) represents the distance between the data \( x_k \) and the cluster centre \( v_i \). The squared error is used as a performance index that measures the weighted sum of distances between cluster centres and elements in the corresponding fuzzy clusters. The number \( m \) governs the influence of membership grades in the performance index. The partition becomes fuzzier with increasing \( m \) and it is proven that the FCM clustering converges for any \( m \in (1, \infty) \).

**CASE STUDY**

This case study examined a typical CRM model used by a business. The model was examined and possible areas that fuzzy theory can be used as a data mining technique were also investigated. Due to confidentiality, the nature of the business and the business name cannot be presented here. The building blocks of the CRM model are to focus on providing a framework to deliver value to the customer and profitability to the company. This framework starts with differentiating customers along two dimensions, by their value to the firm, and by their wants and needs. The have focused on three different measurements by differentiating the customers. First, they examine the customer valuation measurement by understanding the value a customer represents to the company. This measurement identifies revenue growth potential from repeat purchases, referrals, and expanding scope of business. The second measurement looks at the customer preference to determine what are the wants and needs of customers. This measurement examines the results that a customer is trying to achieve, and determine whether the company can bundle products to deliver the results. Lastly, customer loyalty measurement is used to determine the customer’s loyalty. In this measurement, the feedbacks from the customers are very important to gain their loyalty. Basically, these three measurements can fit into the CRM model (Nykamp, 2001) as shown in Figure 1. Please take note that in a CRM process, it is a cycle. Each time you repeat the cycle, more data and experience will be accumulated for the analysis model. This will put your company in a better position to meet customers’ needs and increase revenue.

<table>
<thead>
<tr>
<th>Understanding &amp; Differentiate</th>
<th>Develop &amp; Customise</th>
<th>Interact &amp; Deliver</th>
<th>Acquire &amp; Retain</th>
</tr>
</thead>
<tbody>
<tr>
<td>Examine customer’s needs</td>
<td>Differentiate based on customer’s needs, characteristics, and behaviours</td>
<td>Develop products to meet customer’s needs</td>
<td>Customised by customer segment</td>
</tr>
<tr>
<td>Customer preference measurement</td>
<td>Customer valuation measurement</td>
<td>Customer loyalty measurement</td>
<td></td>
</tr>
</tbody>
</table>

Figure 1: CRM Process Model
It is impossible to examine all the data mining techniques used in the whole CRM process adopted by the company. The investigation performed in this paper has narrowed down to just two broad areas, the segmentation and targeting, and the customer loyalty management. Due to confidentiality, their present techniques used are not discussed here.

For segmentation and targeting approach, FCM clustering is first used to perform an analysis on the available information and to perform unsupervised clustering. The FCM clustering identified different segments by using input parameters contributed by quantifying customers’ needs, characteristics and behaviours. At the end of the clustering, each customer will be assigned to some fuzzy scores that contribute to the whole distribution as well as to the different segments. After the FCM clustering has identified the segments with fuzzy partition, a set of fuzzy rules can also be extracted based on the clustering information. This will allow the business analyst to gain knowledge on how the segments are determined. Business analyst can examine the set of fuzzy rules to modify the behaviour of the segmentation model, as well as to incorporate their own knowledge and experience. Beside the advantage of providing a set of understandable fuzzy rules, the fuzzy partitions together with the fuzzy scores generated for each customer will allow mixing model to be constructed easily and efficiently. With the extracted information, a customer’s needs and expectations can be understood better. As a human, we do not always belong to a segment. However, depending on the situations and circumstances, we may also have a certain degree of belonging in other segments at some time. With the assistance of fuzzy rules, fuzzy partition and fuzzy scores, this characteristic can be modelled more precisely.

For customer loyalty management, this company makes use of call centre to measure the satisfaction and loyalty of the customers. The call centre interviews the customer segments that are important to the business. They use an online reporting system to link the data mart to the real time access of the voice from customer data. This will be used for tactical decision making. The decision support system in the data mart will then make decision based on the criteria by measuring the customer valuation, satisfaction and loyalty as shown in Figure 2. If the support system decided that the customer needs attention, it will signal the client manager.

The decision support system can be constructed using fuzzy memberships and fuzzy rules. In the very beginning, the business analyst has to hand-code the fuzzy rules based on their experience and knowledge. After more data have been collected, fuzzy rules can be constructed directly from the data. Business analyst will then need to verify the fuzzy rule base. The advantage of using fuzzy theory in mining the pool of surveyed customer is that, it will alert the client manager even though the warning signal is uncertain. The client manager can then decide on the seriousness of the warning signal based on the fuzzy firing strength and by examining relevant parameters.
Besides generating the warning signal, the overall customer voice data can be used in constructing segmented fuzzy rule bases to signal any change of trends in customer perception of service quality and product satisfaction. As the fuzzy rule bases work on fuzzy memberships, the degree of changes can be model more closely to produce a more reliable analysis of the market.

CONCLUSION

This paper has examined the possibility of using fuzzy theory and fuzzy clustering in the CRM model. The case study presented has highlighted two areas in a typical CRM model where the use of fuzzy theory can improve the decision making process. The advantage of using fuzzy theory in CRM is that the business analyst can gain in-depth understanding into the data mining model. With the understanding of the model, the analyst can modify and add-on knowledge and experience into the model. Besides, fuzzy theory can handle uncertainties in the data more efficiently than traditional data mining techniques.

REFERENCES


