Construction of Customer Classification Model Based on Inconsistent Decision Table

Li Ju, Xu Wenbin, and Zhou Bei, Member, IACSIT

Abstract—The rough set since proposed has obtained the successful application in each domain. It theoretically has the profound significance and the application. without doubt is one kind of new challenge. This thesis studied the reduction method which can process the inconsistent policy-making table directly. This thesis has studied customer classification forecast model which is based on the rough set. Mainly take the rough set theory as the foundation, first gains the data from the CRM system, and convert them into relevant decisiontable. Secondly, we complete and discretize the data in the decisiontable. Thirdly, we reduce the attribute and value. Lastly, we can conclude rules for making decision and establish logic ratiocination system. This thesis validates and analyzes the conclusion and rules which are established by rough set. This thesis studied the reduction method which can process the inconsistent policy-making table directly. This thesis has studied customer classification forecast model which is based on the rough set. Mainly take the rough set theory as the foundation, first gains the data from the CRM system, and convert them into relevant decisiontable. Secondly, we complete and discretize the data in the decisiontable. Thirdly, we reduce the attribute and value. Lastly, we can conclude rules for making decision and establish logic ratiocination system. This thesis validates and analyzes the conclusion and rules which are established by rough set.

Index Terms—Rough Set, Attribute Reduction, Value Reduction, Rule Reduction, Customer Classification.

I. INTRODUCTION

Philip Kotler pointed out that customer-centric enterprises not only need to build the product, more importantly, is to build customers. A business has a loyal customer base; we can effectively improve the quality of customer relationship management, to achieve the purpose of improving the competitiveness of enterprises.

Currently used in the CRM system, largely based on knowledge discovery decision trees, neural networks, association rules, clustering, K neighbors and other algorithms, these algorithms have some limitations [2][3][4], such as genetic algorithms [5], requires too many parameters, many of the problems codes difficulty, a satisfactory solution rather than the optimal solution, computationally intensive. For example: neural networks [6], opaque, cannot explain how the results is generated.

II. THE BASIC CONCEPT OF ROUGH SET

Rough set theory has the ability that uses incomplete information or knowledge to deal with some uncertain phenomena or disaggregate data according to some uncertainty results.

A. Rough Set features

1) Rough Sets does not require a priori knowledge. Fuzzy sets and evidence theory based on probabilistic methods are a common method of dealing with uncertain information, but these methods require some additional information or a priori knowledge, such as fuzzy membership function, the information sometimes is not easy to get. Rough set analysis methods uses only information provided by the data itself, without any a priori knowledge.

2) Rough set is a powerful data analysis tool. It can express and deal with incomplete information: simplify the data and obtain smallest expression of knowledge in the premise of keeping critical information; identify and assess the dependencies between data, reveal the simple concept model; obtained the rule knowledge from empirical data.

3) Rough sets and fuzzy sets, respectively, depict the two sides of incomplete information: rough set based on the indiscernibility relation, focusing on classification, fuzzy set stress ambiguity of collection itself, (vagueness).

4) From the view of rough sets, rough sets can not be clearly defined because of the lack of adequate knowledge of the domain, but you can approach with a pair of clear sets. Rough set and evidence theory based on probabilistic methods also have some overlap, in practical applications they can complement each other.

5) Rough set theory has some unique perspective. These views make rough set particularly suitable for data analysis. The members of the rough set is an objective calculation, and thus can effectively improve the quality of customer relationship management, to achieve the purpose of improving the competitiveness of enterprises.

Manuscript received April 9, 2011; revised July 26, 2011.

Li Ju is with School of Computer Science and Engineering, Chang Shu Institute of Technology, Changshu, Jiangsu, China, phone: 15851528453; e-mail: liuju284532@163.com

Xu Wenbin is with School of Computer Science and Engineering, Chang Shu Institute of Technology, Changshu, Jiangsu, China, phone: 13962377895, e-mail: 50003580@qq.com

Zhou Bei is with School of Computer Science and Engineering, Chang Shu Institute of Technology, Changshu, Jiangsu, China, phone: 18915608898, e-mail: 327136379@qq.com
denoted the target objects under the
attribute set, which includes all of the reduction.

Definition 2: In decision-making system
\[ S = (U, C \cup D, V, f) \], \( a \in C \) is defined as
\[ \sigma_{(C, D)}(a) = \frac{\gamma_c(D) - \gamma_{C \setminus \{a\}}(D)}{\gamma_c(D)} = 1 - \frac{\gamma_{C \setminus \{a\}}(D)}{\gamma_c(D)} \].

In decision-making system \( S = (U, C \cup D, V, f) \), the
attribute importance of \( P \subseteq C \) is defined as
\[ \sigma_{(C, D)}(P) = \frac{\gamma_c(D) - \gamma_{C \setminus P}(D)}{\gamma_c(D)} = 1 - \frac{\gamma_{C \setminus P}(D)}{\gamma_c(D)} \].

Definition 3: the domain and dependence properties. With
decision-making system \( S = (U, C \cup D, V, f) \), \( C, D \) denote condition attributes and decision attribute respectively,
positive region of the decision attribute under the condition
attribute can be defined as \( \text{POS}_c(D) = \bigcup_{x \in C} \text{c}(X) \), \( \text{POS}_c(D) \) denoted the division according to C's knowledge
\( U/C \), can be classified exactly U/D.

The dependence of decision attribute D on condition
attribute C is defined as,
\[ k = \gamma_c(D) = \frac{|\text{POS}_c(D)|}{|U|} \]

dependence \( \gamma_c(D) \) denoted the target objects under the
condition attribute C that can be classified exactly U/D accounted for the ratio of the total object in the domain,
expressed dependence of decision attribute on the condition attribute, denoted \( C \Rightarrow_k D \).

III. CUSTOMER CLASSIFICATION MODEL

In CRM, knowledge discovery based on rough set can be applied in the following areas:

A. Customer profitability analysis.

Customers to buy directly related to profits of enterprise. Using data mining techniques to build customer profitability return prediction model. We can classify customers, high-value customers, low-value customers and worthless customer. For low-value and worthless customers, we can take certain measures to make low-value or worthless customers into profitable customers; Enterprises can place the limited energy and resources on the most profitable place. For high-value business customers, Enterprises can take different promotions and incentives to maintain customer loyalty to the enterprise. We can identify the trend of the customer return size from the prediction model. Accordingly, the company can targetedly take appropriate measures to promote good conversion or avoid the bad conversion.

B. One to one marketing.

One to one marketing is to understand each client, and build lasting relationships. Through the characterization and classification, knowledge discovery technology based on rough set can classify a lot of customers into different classes, customers in each class have similar attributes, rather than similar properties in the different customers. Through data mining to understand the different customer preferences, to
provide targeted products and services, which can greatly improve the satisfaction to enterprise and products.

C. Cross-selling.

Cross-selling means a marketing process that you provide new products and services to your existing customers. Through relevant analysis, data mining can help you get the best reasonable sales match. Relevant analysis results can be used in two aspects of cross-selling: For the higher purchase frequency of the merchandise mix, identify those who purchased most commodities of a portfolio. Abjective adverb fell the “missing” goods; Another aspect is that for each customer to find more suitable relevant laws, sell them the corresponding goods series.

D. Obtain new customers.

In most commercial areas, the main indicators of business development, including the ability to obtain new customers. Including the discovery of new customers who do not understand your products, and also the customers who received services of competitors before.

By the breakdown of these users, it can help companies to filter potential customers. The aims of customer classification is not only to achieve effective identification of customers, often used to guide the strategic resource allocation of enterprise customers manage.

E. Customer classification

Decision trees, clustering is a common tool for customer segmentation, in accordance with different criteria, such as: according to the customer's spending habits, consumer psychology, purchase frequency, demand for the product or the product profit contribution, which is divided into different user groups, in order to achieve the targeted customer service and develop targeted product to improve customer satisfaction, maximum digging customers lifetime value to enterprise. Through the breakdown, it helps the companies to provide different services according to the different characteristics of the customers, enables companies with minimum investment for maximum return.

IV. CONSTRUCTION OF CUSTOMER CLASSIFICATION PREDICTION MODELS

A. Data acquisition and representation

The data stored in the database due to human or physical, and other factors, there are interference factors such as noise data, vacancy data, and inconsistent data, while a large number of raw data in the database is in the concept of a lower level, not suitable for knowledge discovery. Therefore, the preprocessing of customer data is necessary. The experimental data in this paper come from membership experimental data in this paper come from membership details of consumption of a store. Through data preprocessing, the cluster analysis of 30 selected customers, finding the most valuable customers accounted for 19.2% of total, sales profits accounted for 63.7%, the general value customer accounted for 34.6% of total sales profits accounted for 22.9%, low-value customer accounted for 46.04% of total, sales profits accounted for 13.25%. 12 customer detail property of cluster as condition property, the customer group
from cluster analysis as a decision attribute. NS discrete decision table, using this method discreted data, obtained customer discretization results, shown in Table I.

**TABLE I. NS DISCRETIZATION RESULTS**

<table>
<thead>
<tr>
<th>Symbol</th>
<th>means</th>
<th>Discrete values</th>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>Sex</td>
<td></td>
<td>Female</td>
<td>Male</td>
<td></td>
<td></td>
</tr>
<tr>
<td>B</td>
<td>Marital status</td>
<td></td>
<td>unmarried</td>
<td>married</td>
<td></td>
<td></td>
</tr>
<tr>
<td>C</td>
<td>Region</td>
<td>XIA GUAN area QI XIA area</td>
<td>GU LOU area XUAN WU area</td>
<td>JIAN YE area BAI XIA area QIN HUAI area XUAN WU TAI area</td>
<td></td>
<td></td>
</tr>
<tr>
<td>D</td>
<td>Revenue</td>
<td>Under 20000</td>
<td>20000-50000</td>
<td>50000-100000</td>
<td>On 100000</td>
<td></td>
</tr>
<tr>
<td>E</td>
<td>Cultural level</td>
<td>High school</td>
<td>University</td>
<td>Graduate student</td>
<td>Others</td>
<td></td>
</tr>
<tr>
<td>F</td>
<td>VIP Card Type</td>
<td>Card</td>
<td>Bronze</td>
<td>Silver card</td>
<td>Gold card</td>
<td></td>
</tr>
<tr>
<td>G</td>
<td>has the VIP Card numbers</td>
<td>0-1</td>
<td>1-3</td>
<td>Above 3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>H</td>
<td>Number of private cars</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>Above 2</td>
<td></td>
</tr>
<tr>
<td>J</td>
<td>My Last Days Since</td>
<td>50-60 days</td>
<td>30-50 days</td>
<td>10-30 days</td>
<td>0-10 days</td>
<td></td>
</tr>
<tr>
<td>K</td>
<td>Month of the customer's purchase frequency</td>
<td>0-1 time</td>
<td>1-4 times</td>
<td>4-8 times</td>
<td>On 8 times</td>
<td></td>
</tr>
<tr>
<td>L</td>
<td>get the profits in a month from Customers</td>
<td>0-40 yuan</td>
<td>40 yuan-100 yuan</td>
<td>On 100 yuan</td>
<td></td>
<td></td>
</tr>
<tr>
<td>M</td>
<td>Housing</td>
<td>Poor</td>
<td>General</td>
<td>Better</td>
<td>Very</td>
<td></td>
</tr>
<tr>
<td>N</td>
<td>Decision</td>
<td>Low-value</td>
<td>General-value</td>
<td>High-value</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**B. Condition attribute reduction**

In order to mine classification rules from the large amount of collected data, reduce workload, under the condition required to maintain the condition property relative to the classificaition ability of decision attribute, minimize the redundant rules, remove the repeated condition attributes in the decision table. Further reduction to table I, we get the initial decision table, as shown in Table II. In Table II, there is exactly the same value of the property records of the different customers, as redundant objects merge. Then calculated the condition attribute reduction relative to decision-making attributes, based on reduction of rough set theory and method. I used a program written in MATLAB redu.m, to achieve reduction of rough functions. Specific procedures are as follows:

```matlab
function y = redu(c,d,x) % Defined function
    clear all;
    close all;
    y = core(c,d,x);
    q = ind(d,x);
    p = ind(c,x);
    pos_cd = pospq(p,q);
    re = y;
    red = ind(y,x);
    pos_red = pospq(red,q);
    while pos_cd ~= pos_red
        cc = setdiff(c,re);
        [c1,c2] = size(cc);
        for i = 1:x2,yy(i) = sgf(cc(i),re,d,x);
        end
        cd = setdiff(c,y);
        [d1,d2] = size(cd);
        for i = d2:-1:1,yy(i) = [];
        end
        [zz,ii] = sort(yy);
        for v1 = c2:-1:1
            v2 = ii(v1);
            re = cat(2,re,cc(v2));
            red = ind(re,x);
            pos_red = pospq(red,q);
        end
    end
end
end
[y1,y2] = size(y);
j = 1;
for i = 1:y2,[y,j] = redu2(j,y,c,d,x);
end
function[y,j] = redu2(i,re,c,d,x)
    yre = re;
    [re1,re2] = size(re);
    for qi = re2:-1:1
        if ismember(re(qi),core(c,d,x)),y = re;
            break
        end
        re = setdiff(re,re(qi));
        red = ind(re,x);
        pos_red = pospq(red,q);
        if pos_cd == pos_red,y = re;
            break
        end
        y = yre;
        j = i+1;
        break
    end
end
```

After running MATLAB reduction function redu, we obtained attribute reduction results by a total of 5: (D, J, K, L), (F, J, K, L), (H, J, K, L), (I, J, K, L) and (M, J, K, L). This reduction results demonstrate three important value attributes.
Here only one of the attribute reduction results (F, J, K, L) to illustrate the classification rule extraction process, eliminate redundant attributes and merge the same line, the decision table simplified as shown in Table III.

<table>
<thead>
<tr>
<th>Record Number</th>
<th>Condition attribute</th>
<th>D</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0 0 2 1 2 0 0 1 1 3 0 1</td>
<td>0 1 0 1</td>
</tr>
<tr>
<td>2</td>
<td>0 0 1 2 0 3 2 0 1 3 1 2 0</td>
<td>2 0 2</td>
</tr>
<tr>
<td>3</td>
<td>0 0 2 1 0 0 2 0 1 0 1 0 0</td>
<td>0 1 0 0</td>
</tr>
<tr>
<td>4</td>
<td>0 0 1 2 0 0 1 1 1 0 1 0 1</td>
<td>0 1 0 1</td>
</tr>
<tr>
<td>5</td>
<td>0 1 0 0 1 2 0 0 1 1 0 1 0 1</td>
<td>0 1 0 1</td>
</tr>
<tr>
<td>6</td>
<td>1 1 0 0 0 0 1 0 1 0 0 0 0 0</td>
<td>0 1 0 0</td>
</tr>
<tr>
<td>7</td>
<td>0 1 0 1 2 3 1 0 1 2 1 0 1 1 0</td>
<td>0 1 0 1</td>
</tr>
<tr>
<td>8</td>
<td>1 1 0 0 0 0 1 0 1 0 0 0 0 0</td>
<td>0 1 0 0</td>
</tr>
<tr>
<td>9</td>
<td>1 0 0 0 1 0 0 0 0 2 3 2 2 0 2</td>
<td>0 1 0 1</td>
</tr>
<tr>
<td>10</td>
<td>1 0 5 1 1 0 0 0 0 2 3 2 2 1 0 1</td>
<td>0 1 0 1</td>
</tr>
<tr>
<td>11</td>
<td>0 1 0 0 1 2 0 0 1 1 1 1 1 0 0</td>
<td>0 1 0 1</td>
</tr>
<tr>
<td>12</td>
<td>0 1 0 1 2 0 3 0 1 1 1 1 0 0 0</td>
<td>0 1 0 1</td>
</tr>
<tr>
<td>13</td>
<td>0 1 0 2 1 2 1 0 1 0 3 2 2 1 1 2</td>
<td>0 1 0 1</td>
</tr>
<tr>
<td>14</td>
<td>1 0 0 1 1 2 1 1 1 0 3 1 2 1 1 1</td>
<td>0 1 0 1</td>
</tr>
<tr>
<td>15</td>
<td>1 0 0 1 1 2 1 1 1 0 3 1 2 1 1 1</td>
<td>0 1 0 1</td>
</tr>
<tr>
<td>16</td>
<td>0 1 0 1 1 2 2 0 0 0 0 0 0 0 2 0</td>
<td>0 1 0 1</td>
</tr>
<tr>
<td>17</td>
<td>0 0 0 1 2 0 2 0 0 0 1 1 1 1 1 1</td>
<td>0 1 0 1</td>
</tr>
<tr>
<td>18</td>
<td>0 1 0 1 1 2 2 0 0 0 0 0 0 0 2 0</td>
<td>0 1 0 1</td>
</tr>
<tr>
<td>19</td>
<td>0 0 2 1 0 2 0 0 2 0 1 0 0 1 0 0</td>
<td>0 1 0 1</td>
</tr>
<tr>
<td>20</td>
<td>0 0 0 1 2 0 2 0 0 0 1 1 1 1 1 1</td>
<td>0 1 0 1</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Rule</th>
<th>Number</th>
<th>Condition attribute</th>
<th>D</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rule 1</td>
<td>2</td>
<td>3 3 1 2 2</td>
<td></td>
</tr>
<tr>
<td>Rule 2</td>
<td>14,15</td>
<td>2 3 1 2 1</td>
<td></td>
</tr>
<tr>
<td>Rule 3</td>
<td>11,12</td>
<td>2 1 1 1 0</td>
<td></td>
</tr>
<tr>
<td>Rule 4</td>
<td>17,20</td>
<td>0 1 1 1 1</td>
<td></td>
</tr>
<tr>
<td>Rule 5</td>
<td>1</td>
<td>0 3 0 1 1</td>
<td></td>
</tr>
<tr>
<td>Rule 6</td>
<td>3,19</td>
<td>0 0 1 1 0</td>
<td></td>
</tr>
<tr>
<td>Rule 7</td>
<td>4,5</td>
<td>2 1 0 1 1</td>
<td></td>
</tr>
<tr>
<td>Rule 8</td>
<td>11,12</td>
<td>2 1 1 1 0</td>
<td></td>
</tr>
<tr>
<td>Rule 9</td>
<td>13</td>
<td>2 3 2 2 2</td>
<td></td>
</tr>
<tr>
<td>Rule 10</td>
<td>14,15</td>
<td>2 3 1 2 1</td>
<td></td>
</tr>
<tr>
<td>Rule 11</td>
<td>9</td>
<td>0 3 2 2 2</td>
<td></td>
</tr>
<tr>
<td>Rule 12</td>
<td>10</td>
<td>0 3 2 2 2</td>
<td></td>
</tr>
</tbody>
</table>

C. Generation of classification rules

Currently, most rule reduction methods focus on consistent decision-making system, the rule reduction methods applied inconsistent decision-making system is less. However, there is often inconsistent information in the current data storage systems. Therefore, we studied the direct access rules in an inconsistent system. Complete rules From the results of this attribute reduction are shown in Table IV.

D. Classification rules according to customer classification prediction

According to rule 1, the semantic description is as follows: If the Gold Card customers the number of days of since its recent purchase is within 10 days, consumption frequency in one month is between 1 and 4 times, the profits obtained from the customer in one month is more than 100, this customer is divided into the most valuable customers. It is similarly available.

V. Function module

Based on the above analysis, we can customer classification based on rough set prediction system functional design.
System is divided into two parts. The first part is initialized, the second part is data mining of the customer classification prediction algorithm based on rough set.

The first module is to initialization module. Its main function is: Import the data warehouse which is needed mining analysis, it contains a large amount of data records and more detailed customer information. According to the customer property in the data warehouse, develop one or more enterprise solutions for customer classification prediction. The second module is data mining module based on rough set. It is the core algorithm of the system modules.

VI. CUSTOMER CLASSIFICATION PREDICTION MODEL ASSESS

Using supermarket customer's personal information and detailed transaction information stored in the database of a membership supermarket, effectiveness of the method described in this paper has been verified. There are details of the 11,236 member customer in the supermarket database. There are a total of 8350 customers registered for more than two years, this part of the customers are selected for the experimental samples. Take random sampling method, according to the ratio of 7:3, the 8350 test samples were divided into two parts: Part 1 5845 (8350 × 70%) clients as "customers" as a training sample, assuming that part 2 2505 (8350 × 30%) place the customer as a "new customer", that is to say the test sample. In order to verify the validity of the classification prediction model described in this paper. The following comparison method is used in the experiment: Part 2 500 customers are randomly selected from 2505 customers, according to probability, the high-value customers should be 500 × 21% = 105; but which is actually 238 high-value customers. That is, according to the method described in this article had improved 238/105 = 2.27 degrees. Through analysis of experimental results it can be found that compared with a certain blindness of random development strategies of key customers, according to a new level of customer value is getted from the classification prediction model to develop customers will get even better results.

VII. CONCLUSION

Evidence shows that this chapter of the customer classification prediction method is feasible, and through the rough set attribute reduction, do customer classification according to a few of the attributes customer classification, the proposed customer classification prediction model, could help companies predict in advance the new customer or potential customer value level and make the more targeted client development strategy.

Innovation of this paper:
1. For discrete data, studied discretization method based on discernible matrix and Boolean logic combination.
2. In the rule reduction algorithm, this paper study the most general rule reduction approach, when in dealing with inconsistent rule reduction, the decision-making has been redefined equivalence class, this equivalence class should include in the initial decision table with inconsistent with the object-related records, after the revised definition, the reduction algorithm can be directly inconsistent with the rules.

REFERENCES

Li Ju Changshu, Jiangsu 1981, graduated from Nanjing University of Information Science, getted master degree about Rough sets and mage, Nanjing, Jiangsu, China. She taught at the Changshu Institute of Technology, Lecturer.