Chapter 7
Intelligent Knowledge Acquisition and Application in Customer Churn

7.1 Introduction

Almost all of the entrepreneurs desire to have brain trust generated decision to support strategy which is regarded as the most critical factor since ancient times. With the coming of economic globalization era, followed by increasing competition, rapid technological change as well as gradually accrued scope of the strategy. The complexity of the explosive increase made only by the human brain generates policy decision-making appeared to be inadequate.

Extension theory is a new discipline engaged in studying the extension properties of things as well as its law and methods (Cai et al. 2003; Cai 1994). Bibliography (Han and Micheline 2006) combined the Extension theory with artificial intelligence, database technology, and software engineering to come up a software system called Extension Strategy Generating System (ESGS) which enable computer to mock human strategy generation. The idea of ESGS is an inevitable trend towards scientific and intelligent decision-making. However, There would have strategic explosion which would definitely lead to an increase in optimal evaluation of artificial workload, If it was not controlled appropriated during the procedure of strategy generation on computer.

In recent years, Data mining as an important instrument for knowledge discovery has been widespread concerned by scholars from all over the world (Han and Micheline 2006; Olson and Shi 2007). It has already been employed in Finance and insurance (Olson and Shi 2007), Marketing (Chen and Hu 2005), bio-medical treatment (Larry et al. 2004), Internet customer analysis (Nie et al. 2006) etc. However, the pattern knowledge obtained from data mining is only the description of characteristic of things, which requires a further combination of expert experience, and finally come out the strategy for solution relying on analysis of business experts.

A website is one of China's major portal sites, "one of the four major portals in China". the company is always maintain the leading position of the industry in china in the development of Internet applications, services and other technology. Since its inception in June 1997, by virtue of advanced technology and high quality service, it is welcomed by the majority of Internet users, and named as China's top

Y. Shi et al., *Intelligent Knowledge,* SpringerBriefs in Business,
DOI 10.1007/978-3-662-46193-8_7

ten sites two times by the China Internet Network Information Center (CNNIC). in 2010 its turnover is 5.7 billion Yuan. Now it provides the online game, e-mail, news, blog, forum, search engine and virtual community services.

Although the website company own a large number of charge-mail registered users. However, some customers are lost due to the intense competition and other objective reasons. The acquirement of the 245 rule is through applying decision tree data mining algorithm to divide user into "the existing user, the freezing user and the lost user" and predict the user type. However, It cannot acquire strategy which promote user transformation from those rules, Actually, the freeze user and the normal user can transform into each other in certain condition. Finding transformation knowledge among different users will provide subordinate strategy for strategy.

7.2 The Data Mining Process and Result Analysis

To get intelligent knowledge for prevention of the customers churn through data mining, project group launched the four phases of work according to the following six steps.

The Four phases are:

1. Data Exploration Phase
2. Experimental Mining Phase
3. Data mining Phase
4. Validation and Follow-ups

The six steps are:

1. Understand the process flow, and the distribution of the data, build up the data map.
2. Discuss the selected dataset, compile data dictionary
3. Pre selection of the data fields
4. Determine the process method for the data mining field
5. Determine the data mining plan.
6. Using MCLP software, input integrated and cleaned data in the table of TXT file, then make dataset partition, modeling, get the scoring model step by step, and visualize the results collated. The following figure is data map for customer churn prevention (Fig. 7.1):

Data collection and consolidation

1. A list is selected by the data field of the data dictionary, and data retrieval based on the data retrieval method
2. Compile the log processing program, transform the data into structured data
3. Label the structured data based on the labels from the service department, authenticated by the technical department
4. Data consolidation
5. Cleaning transformation and discretization.

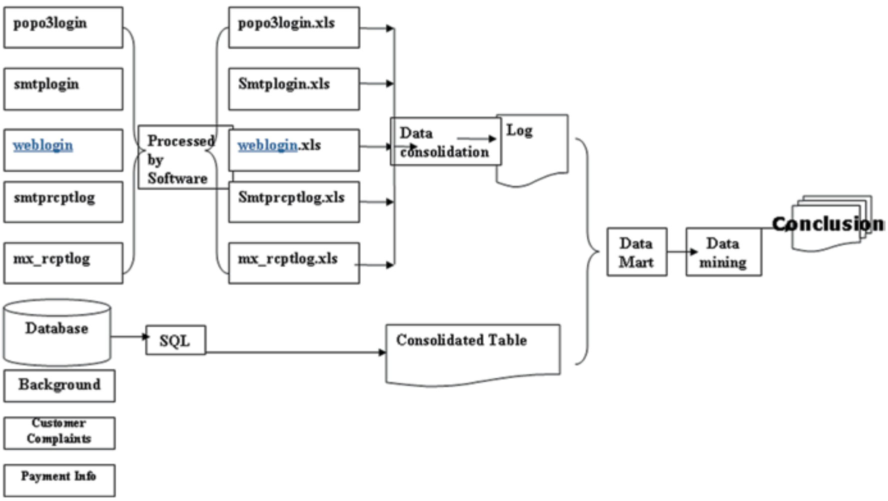
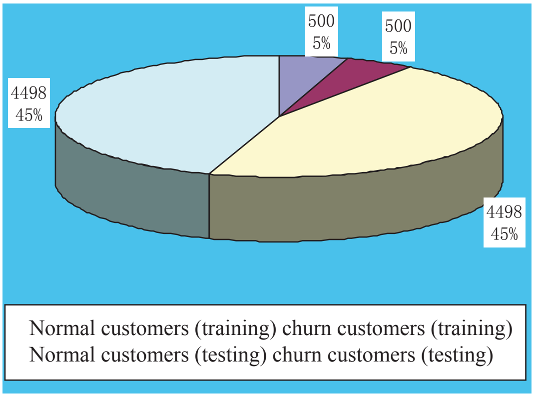


Fig. 7.1 Data map for customer churn prevention

Fig. 7.2 Selections for training samples and test samples



By above five steps, we get 4998 records of churn customers and 4998 records of normal customers. Then we used cross-validation method for ten times, each time we selected 500 records randomly to constitute the training sample from the two data sets respectively, and the remaining data as test sample. Finally we got 10 groups of training samples and test samples. As shown in Fig. 7.2:

Data mining Modeling.

We selected 10 groups on the basis of sample data to set up the evaluation model, the training and testing results show as follows, smooth lines shows our scoring models with high stability. The results are as shown in Fig. 7.3:

In this project, we use cross validation algorithm to generate the 9 groups of score voting machines, their predictive accuracy as shown in Table 7.1:

The combination of the 9 votes consisted 10 MCLP score models, if score[i]>5 it will be judged as normal, score[i]<=5 will be judged for the churn of customers.

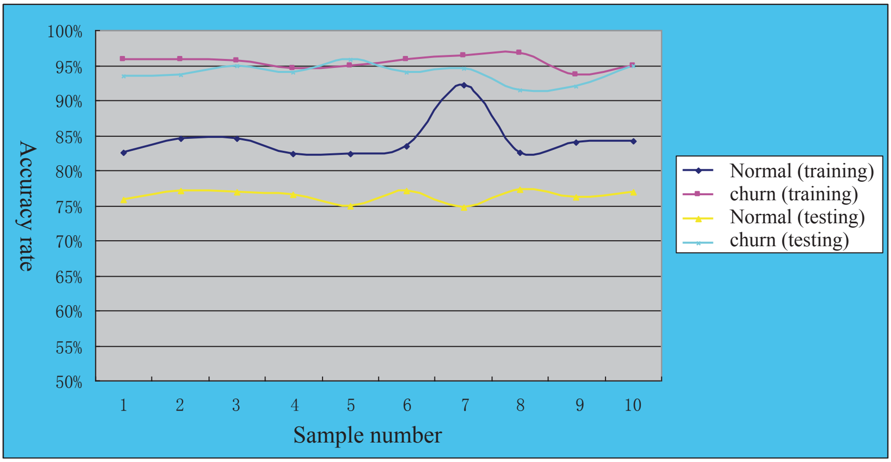


Fig. 7.3 Training and testing results

Table 7.1 Cross validation table

Cross validation	Testing data sets(3382churn+65493 normal)			
	Churn	Accuracy (%)	Normal	Accuracy (%)
Dataset 1	2506	74.0982	46,777	71.4229
Dataset 2	2451	72.4719	47,336	72.2764
Dataset 3	2518	74.4530	46,940	71.6718
Dataset 4	2505	74.0686	46,728	71.3481
Dataset 5	2509	74.1869	46,844	71.5252
Dataset 6	2467	72.945	46,951	71.6886
Dataset 7	2565	75.8427	46,534	71.0510
Dataset 8	2535	74.9556	46,518	71.0274
Dataset 9	2475	73.1815	46,496	70.9938

The score precision measurement has a lot of kinds of methods, including the cumulative distribution structure of the KS value method. It was confirmed to be able to identify more efficient data sets, which are widely applied in the field of credit risk management. We explain the MCLP score model performance from the view of distribution density and cumulative distribution as following.

Table 7.2 below is a distribution density table based on MCLP scoring systems with two types of customers, the first column score is in the range of [1, 10], the second column LOST is the scores of all the number and percentage of the churn customers, the third bar CURRENT shows scores of all the number and the percentage of the normal customers. As can be seen from the charts, the 5382 churn of customers mainly in the low field; for the 69,473 regular customers, the score paragraph mainly in high field. Among them, the churn of customers' score gathered in the scores of 1, of the total churn number 55.797 %, and the normal customer vscore gathered in the scores of 10, accounting for all the normal customer number of 46.196 %.

Table 7.2 The predictive accuracy of Churn prediction

Score	Churn (5382 records)	Normal (69,473 records)
	Percentage	Percentage
1	55.797101	13.196924
2	7.785210	4.115982
3	4.923820	2.961789
4	3.994797	3.118842
5	3.493125	3.269969
6	3.010033	3.923370
7	3.530286	4.588624
8	3.660349	6.107300
9	4.998142	12.521299
10	8.807135	46.195902

A more intuitive density distributions figure is shown in Fig. 7.4, the yellow line represents the distribution of churn customers, blue lines represent the distribution of normal customers, as we can see, the yellow line represents the churn customers and the blue line for normal customers are basically linear separable, thus the MCLP method has good applicability to solve churn problems. (Table 7.3)

Sum up the above distribution density data, we get the distribution function (cumulative distribution):

From the table, we can see the maximum separation of the churn customers and normal customers appears in the arrow pointing to the score = 5 position. that is to say, if our model on the customer's score is $score[i] > 5$, then it can be assumed that the customer loyalty is high, do not churn; if our model's score is $score[i] <= 5$, then it can be assumed that the customer is in the state to be loss, we need to take measures to let them stay.

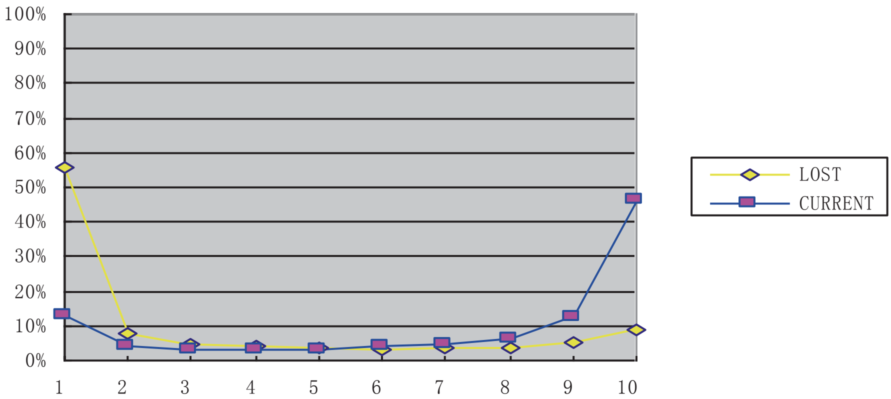


Fig. 7.4 Density distributions figure

Table 7.3 distribution function lists

SCORE	Churn ACCUMULATE	Normal ACCUMULATE	Absolute degree of separation PERCENTAGE
0	0.000000%	0.000000%	0.000000%
1	55.797101%	13.196924%	42.600177%
2	63.582311%	17.312907%	46.269405%
3	68.506132%	20.274695%	48.231436%
4	72.500929%	23.393537%	49.107392%
5	*75.994054%*	*26.663506%*	*49.330548%*
6	79.004088%	30.586876%	48.417212%
7	82.534374%	35.175500%	47.358874%
8	86.194723%	41.282800%	44.911923%
9	91.192865%	53.804098%	37.388767%
10	100.000000%	100.000000%	0.000000%

The following Fig. 7.5 is for the KS graphical display:

Set the origin of coordinates (0, 0), we see that, for the yellow line marked churn customers, there's a big jumps for cumulative distribution in scores 1, growing from 0 to 55.797 %, shows that a large number of customers are accumulated in the scores 1, from 1 to 10 the growth is with relative ease; and for the blue line marking the normal customers, cumulative distribution from scores of 1–9 with relative ease, when increased from 9 to 10, the cumulative distribution grows from 53.804098 to 100 %, a large number of customers is statistically in the numerical. While the two

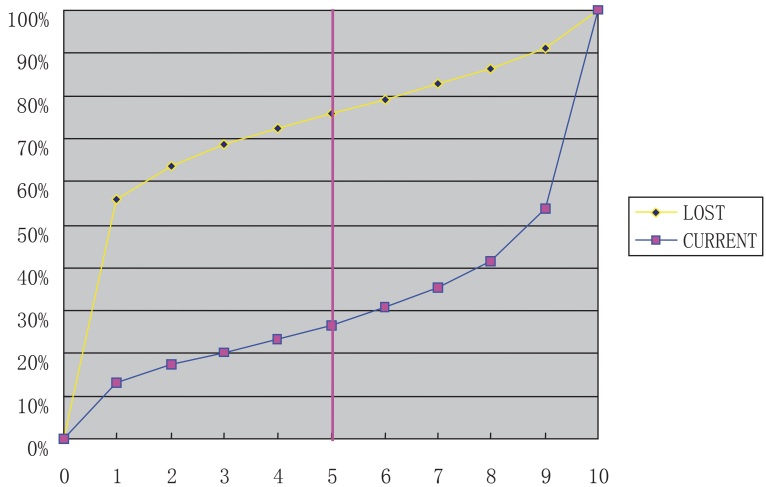


Fig. 7.5 KS score chart

customer maximum separation values appeared in scores of 5 (pink line), the churn customers in scores of 5 have accumulation of 75.994054 %, normal customers for 26.663506 %, KS = |75.994054 % − 26.663506 %| = 49.330548 %, MCLP based vote scoring models of two types of customers is distinguished more clearly.

Conclusion and Management Recommendations for the first stage:

1. Determine whether a customer is churn based on the score of the customer.
2. Analyzing the MCLP model document, find out the characteristics of churn customers by See5 and SPSS.
3. Actions are taken based on the predicted churn customer list and the churn characteristics.

The application of the project result is a complicated, systematic task. The log files are retrieved from the functional department, and processed into structured documents. The functional department designs the application measures, and forwards the feedback to the technical department, mining the new data and build up a reinforcing cycle.

The following is the second stages of the intelligent knowledge management.

7.3 Theoretical Analysis of Transformation Rules Mining

7.3.1 From Classification to Transformation Strategy

Generated by classification can provide a reference for strategy. The existing category mining is based on classical set theory and fuzzy set theory. The classical theory requires each object in Data Universe must belong to one set and only one or the other. The classical collection uses the characteristic function of range {0, 1} to qualitatively describe whether one thing has certain property, A corresponding classification is "within class is the same, among class is the different", But it cannot be used to quantitatively describe the degree of possessing certain property things (Cai et al. 2003). The fuzzy set uses the membership function of range {0, 1} to describe the degree of certain things which are undergoing differences during the intermediate transition. But for the membership degree on the domain object to 0 or 1 are also indiscriminate, is still "within the class is the same" expression. Neither classical set nor fuzzy set study the changes among the categories of things, therefore both of them cannot directly describe the conversion between "non" and "is" in certain condition (Cai 1994). Frankly speaking, many things can divided into two parts according to P property. The part which does not have the nature of P can be further divided into two categories, one is "can be convert into holding P property" and the other is "cannot be transformed into possessing the nature of P" under certain condition. In the actual production, For instance, unqualified products can turn to be qualified after some processing. For example, The number of axles of a workshop production require qualified diameter D = 50 ± 0.1 mm,the unqualified diameter d > 50.1 cm, Considering the "re-lathe processing" of transformation measures, are then taken to turning the product into qualified ones.

In recent years, Extension theory has been initially applied in the field of data mining, and achieved good result. Bibliography (Li et al. 2004) have had outlook-style description on extension application in Data mining. Bibliography (Chen et al. 2006) proposed the conception of Extension Knowledge, which regards the expanding mode as the basic knowledge of the extension knowledge, The Transform Implication of Transmission theory of Extension is the change of knowledge, Extension introduce correlation function to alter contradiction problem into quantitative knowledge by quantification procedure. Bibliography (Zhang and He 2001) exploit methods of the corresponding potential knowledge by utilizing properties of matter element including divergence, relevance, and implication, arousing our concerns on potential information category. Bibliography (Li et al. 2006) analyzed the existing problem caused within the data mining process based on extension theory, and also proposed a new data mining application based on extension transformation through establishing matter-element set in enterprise data. Bibliography (Huang and Chen 2006) further proposed measures for fundamental improvement of data quality promoted by consulting data mining, which promotes the development of transformation of the extension domain in matter-element set. Bibliography (Gao 2002) use extension transformation to turn false proposition into true proposition, infeasibility to feasibility, and come out the idea transform infeasibility to feasibility from the point of change, It also comes out the conception and theories of change knowledge through the study of extension transformation. Bibliography (Tan et al. 2000) gives out the conception and assumption of extension classification, The above reference laid the foundation for our further research. Decision tree classification which has stronger interpretability of classification principle is one of the most commonly used data mining measures (Han and Micheline 2006; Nie et al. 2006). The paper studied the acquisition methods of transformation strategies among one category of thing from the basic idea and methods of extension set, and also engaged in mining rules and knowledge of transformation strategies among one category of thing, and go further to design the implementation of Algorithm.

7.3.2 Theoretical Analysis of Transformation Rules Mining

The basic extension theories of transformation rules mining contain three parts: First and foremost is the extension theory of matter-element, which indicates that everything has the possibility of development and change, The second is the extension transformation theory, that is the property of things will change through certain transformation; The last comes to the theory of extension set, which reflects the degree of transformation of the nature of things. Its definition can be stated as follows:

Definition1 Set U as universe, u is any element in U, k is a mapping of U to real domain I, the given transformation of $T=(TU, Tk, Tu)$ refers to the following equation.

$$\tilde{E}(T) = \{(u, y, y') \mid u \in T_U U, y = k(u) \in I, y' = T_k k(T_u u) \in I\}$$

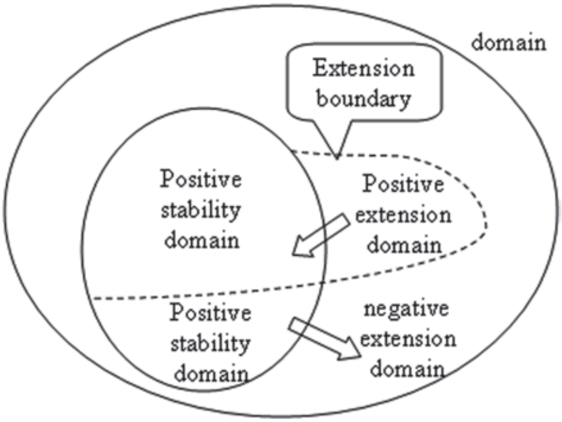


Fig. 7.6 Domain division of element transformation under extension set

The above set is a extension set of U universe, $y = k(u)$ is correlation function of $\tilde{E}(T)$. $y' = T_k k(T_u u)$ is extension function of $\tilde{E}(T)$. T_U T_k T_u is separately the transformation of universe U, correlation quasi-function K, and element u.

On the conditions of $T \neq e$, $\underset{\cdot}{E}+(T)$ is the positive extension domain of $\tilde{E}(T)$, $\underset{\cdot}{E}-(T)$ is the negative domain of $\tilde{E}(T)$, $\dot{E}+(T)$ is positive stable domain of $\tilde{E}(T)$, $\dot{E}-(T)$ is negative stable domain of $\tilde{E}(T)$, $J_0(T)$ is extension bounding of $\tilde{E}(T)$.

Positive extension domain indicates that element which initially does not belong to E turn to a part of E after the implementation of A transformation; while the Negative extension domain reflects element which initially does not belong to E remain the same after A transformation; Positive stability domain refers to element which originally belong to E remain belongs to E after the implementation of A transformation, Similarly, Negative stability domain refers to element which originally does not belong to E is still not part of E after A transformation. Extension boundary refers to element which existed at the border of extension transformation and its extension function is zero. Extension boundary describe the qualitative change point which indicates that elements surpass the point will definitely produce qualitative change. The diagram is shown in Fig. 7.6.

Extension utilize extension set to represent the degree of things holding certain property, and indicates things of certain property can be changed into things that lack of such property, Similarly, things absent of certain property can be altered into things holding such property. In order to make the description more clear and accurate, you should use two definitions together to describe the variability and the process of quantitative and qualitative change of elements.

Bibliography (Nie et al. 2006) transform T from element extend to transformation of correlation function or universe. Element transformation (including affair element and matter element transformation), correlation function transformation and universe transformation, are collectively known as extension transformation. And therefore, make further effort to dig rules of classification transformation on the basic of connecting classification with transformation through extension set

Rule mining which is the extension and expansion of data mining refers to Decision-making strategy come from the process of rule knowledge transformation obtained from data mining, then go through extension transformation and finally enter into category things transformation. According to different business and variable demands, matter element set can alter into another set through special extension transformation. For instance, the churned client might transform to the loyal ones under certain condition. Similarly, the normal customer might turn to be the churned ones under certain condition.

Decision tree as data set classifier has been widely used in data mining. However, it resulted in a static subset of classification of data which only describe the characteristics of different branches of leaves. The branch number is too large, although it produces numerous rules. The proportion of rules truly interested is too small when dealing with the complex data. And therefore effective measures aiming to prevent the loss of customers only relied on classification rules are invalid. Using matter element extension set to represent the results of Decision tree mining so that it can transform the static rule set into changing, dynamic extension rule set, It can transform static descriptive knowledge discovery to dynamic strategy generation. The following describe the acquirement of strategy based on Decision tree extension transformation.

7.3.3 The Algorithm Design and Implementation of Transformation Knowledge

Decision tree is a tree structure similar to flow chart, where each internal node represents a test on attribute, each branch represents output of test, and each leaf node represents a category. Decision tree is mainly based on the summarized data attribute values, from the tree top-level node(root node)to leaf node traversal which store the forecast sample, for example,

Rule 1: Total Types of Mail <= 0

Average No. of logon by POP3 on the second week <= 0
Total No. of log on by Web <= 0
Standby mail service status = not selected
Percentage of Service Period 7 <= 0.25
Total Payment in the past 3 months = 0
→Churn [0.736]

Rule 2: Total No. of logon by WEB <= 0

Standby mail service status = not selected
Percentage of payment method 11 > 0.2941177
Total Payment in the past 3 months = 0
Contact mailbox = No
Contact method = No
ID Number = No
→Freeze [0.757]

Through this procedure can transform the decision tree into a "if-then" form of classification rules. Take rules obtained by see5 measures for example, In the following form:

"Rule 2: (198/14, lift 2.7)

whether use mobile-mail services=0
POINTS <=6
The length of occupied time >92
Type = 6
→ class 0 [0.925]"

In the above form, 198 of rule 2 represents the record number which meet the rule in training set, while 14 represents the record number which does not meet the rule in training set, Predicting Accuracy Rate (PAR)=(198−14+1)/(198+2)=0.925, Enhance degree of lift=PAR/the relative frequency of the occurrence of such class in training set=0.925/0.343=2.7.

7.3.3.1 The Method of Obtaining Transformation Knowledge

In order to obtain strategies of classification transformation rules turn to change through extension transformation mining On the foundation of decision tree classification rule connecting with extension set theory. Take A, B two types of conversion as example.

Set up A as:

$$\left\{D_{\cdot+}(T)\right\} = \left\{D_i \middle| D_i = \begin{bmatrix} I_i, & d_1, & u_{i1} \\ & d_2, & u_{i2} \\ & d_r, & u_{ir} \end{bmatrix}, K_i < 0, K_i \cdot K_i(T) < 0, i \in J_{D_{\cdot+}} \right\}$$

J_D is index set of information unit D_i which meet the condition of $K_{ip}^{\circ+} < 0, K_{ip} \cdot K_{ip}(p) < 0$, the later sign is similar, so no long explain

$$\left\{I_{\cdot+}(T)\right\} = \left\{I_i \middle| I_i = (O_i, c_j, v_{ij}), D_i \in \left\{D_{\cdot+}(T)\right\}, i \in J_{D_{\cdot+}}\right\}$$

Set all of relevant characteristics of $c_j(j=1,2,\cdots,m)$ and $d_p(p=1,2,\cdots,r)$ as $\{j_0\}$, change the property value of rules hold the same property in the two types of rule based on A<-B replacement transformation; there must be transformation set called T_{AB} which make A=>B by transforming the rule existed in B but not in A through adding transformation;

The reliability is $\frac{\left|D_{\cdot+}(T)\right|}{\left|D_{-}(T)\right|}$, support degree is the transformation knowledge of $\frac{\left|D_{\cdot+}(T)\right|}{\left|D\right|}$, which indicates that about $j \subset \{j_0\}$, if $I = (O, c_j, v_j)$ have $v_j \in V_{\cdot+j}(T)$, then $D_i = \begin{bmatrix} & d_1, & u_{i1} \\ I_i, & d_2, & u_{i2} \\ & d_r, & u_{ir} \end{bmatrix}$ which originally belong to E will turn to not belong to E after the implement of T transformation, among which $V_{\cdot+j}(T) = \begin{bmatrix} Min v_{ij} & Max v_{ij} \\ i \in J_{D_{\cdot+}} & i \in J_{D_{\cdot+}} \\ j \in \{j_0\} & j \in \{j_0\} \end{bmatrix}$

For example:

Through original knowledge

"Rule 2: (198/14, lift 2.7)

whether using mobile-mail services=0

POINTS <=6

the length of occupied time >92

Type =6

→class B[0.925]

Rule 3: (6, lift 2.7)

whether using mobile-mail services =1

POINTS <=6

the length of occupied time <=795

→ class A [0.875]"

Obtain transformation rule knowledge

"Rule6: (6/9) support: 4.25 % ID: 240–235

Under:

POINTS <=6(same)

92 < occupied time <=795(same)

Trans:

whether using mobile-mail services =1 to =0

Add: none

→class A to B [61.70%]"

ID: 240–235 is the source rule number generating transformation rule.

Below the word "under" will list rules existed in category A but not in category B. (add "same" signal in the rear)

Below the word "Trans:" will list rules had the same attribute but different value, and convert the antecedent value among condition category based on target category value

"Add: "refers to copy rules existed in B but not in A as an additional transformation condition.

Table 7.4 Statics of transformation rule's record

	The record number	The record number which correspond to antecedent	The record number which correspond to consequent
A	$\lvert D(A)\rvert$	Fa	Ra
B	$\lvert D(B)\rvert$	Fb	Rb
Subtotal	$\lvert D\rvert$	F	R

"POINTS<=6", "92< the length of occupied time <=795" and "whether use mobile-mail services=1" are the antecedent of rule knowledge, "whether use mobile-mail services=0" and "none" are the consequent of rule knowledge.

7.3.3.2 Evaluation Index of Transformation Knowledge

Set the record number of $\{A\}\cup\{B\}$ in database table as $|D|$, set the record number which correspond to antecedent in $\{A\}$ as Fa, set the record number which correspond to consequent in $\{A\}$ as Ra, set the record number which correspond to antecedent of transformation rule knowledge T_{iAB} in $\{B\}$ as Fb, set the record number which correspond to consequent of this in $\{B\}$ as Rb, set all the record number which meet antecedent of rule set as F, and all the record number which meet consequent of rule set as R, set the record number which accord with A rule set in $\{D\}$ as $|D(A)|$, set the record number which accord with B rule set in $\{D\}$ as $|D(B)|$, The following table 1 clearly reflect those evaluation index (Table 7.4). the accuracy rate of rules

$$P_{iAB} = (\mathrm{Rb}+1)/(\mathrm{Rb}+\mathrm{Ra}+2) \tag{7.1}$$

anticipative conversion rate

$$Tr = Fa / F \tag{7.2}$$

the support degree of the rule

$$S = \frac{|F|}{|D(B)|} \tag{7.3}$$

the reliability

$$R = \frac{|Rb|}{|R|} \tag{7.4}$$

For instance, In the "Rule6: (6/9) support: 4.25% ID: 240−235", 6 refers to the record number which meet transformation condition. 9 refer to the record number which meet antecedent of A "under" condition. support: 4.25% refers to reliability.

7.3.3.3 Implementation Steps

1. **Read in the original rule set**

Take See5 (http://www.rulequest.com.) decision tree software for example, the initial rule set saved in out text file as the form of text file, Rule format as shown in the above example, in which 198 of rule2 represents the record number which meet the rule in training set, 14 represents the record number which does not meet the rule in training set, Predicting accuracy rate $=(198-14+1)/(198+2)=0.925$, Enhance degree of lift $2.7=$prediction accuracy rate/the relative frequency of occurrence of such class in training set. Classification rules will be read into database in turn, stored into the rule table.

2. **Pretreatment of rule set**

Expurgate the same rules generated by rereading in the process, establish keyword of full-text index and so forth.

3. **Set mining parameters**

Set the following parameters by user:

1) ,"Mining rules transform from class__ into class__", such as class0, and class1,etc, shown in the mentioned example.
2) ,"The number of rules have the same content > = ——", such as "POINTS < = 6"and"92 < the length of occupied time< = 795" in rules 2 and rule 3.
3) ,"The number of rules have the same content< =——", such as the different value of antecedent in" whether use mobile-mail services" in rule 2 and rule 3 in the example.
4) ,"The predicting transformation rate of extension rule> =——%", the conversion rate of applying predicting extension rule = the record number consisted with transformation rule in rule set/all the record number consisted with antecedent of rule set.

4. **Tule Mining**

Search for rules have many similarity and less discrepancies, by comparing the output generated by transformation rules

5. **Rule evaluation index calculation**

In order to evaluate the practicality and novelty of extension rule, you should calculate the indicators, such as accuracy rate, predicting transformation rate, support degree and credibility.

6. **Demonstrate results report**

The results of mining provide the list of transformation rules and the summary report of mining condition.

7.3.3.4 Mining Algorithm of Transformation Knowledge

The following shows the brief algorithm of transformation knowledge mining:

Input: The result set based on decision tree data mining (two class are respectively represented by A and B), and the minimum record number n from elements of the two set.

Output: matter element of A might transform into strategy of B under the condition of TkK(TRR).

Method:

(1) The elements in A, B should be respectively represented as multi-dimensional matter element w1 and w2, R1m and R2m analysis indicate the first m matter element in W1 an W2.
(2) The number of matter element for i = 0 to A
(3) The number of matter element for j = 0 to B
(4) Set integer total equal to 0
(5) Set the dimension of R1i as iN, set the dimension of R2j as jN;
(6) For k = 0 to iN
(7) For kk = 0 to jN
(8) If R1ir is the K-dimension feature, then value is the same as the value of KK-dimension of R2j
(9) Total = total + 1
(10) End k,kk circulation
(11) If total is greater or equal to the system input value N, then output one transformation knowledge of R1i and R2j
(12) End i, j circulation
(13) End.

7.4 Conclusions

We briefly analyzes the measures of the acquisition strategy, combined extension theory with research result to come up with strategy knowledge measures of acquiring customer transformation through data mining and extension transformation, and implement through designing algorithm programming. The practicality of this method is confirmed by preliminary test.

First of all, import all the decision tree rules into rule base, as is shown in Fig. 7.7.

Then set the parameters (such as transform users from freeze user to normal user) to engage in mining strategy to come out dozens of strategies, the rule 6 of the former Sect. 3.2 is the case in point, from which indicates that users among the scope of POINTS<=6 and the length of occupied time between 92 and 795 can reduce their loss, as long as advising them not to use mobile-mail services. This intui-

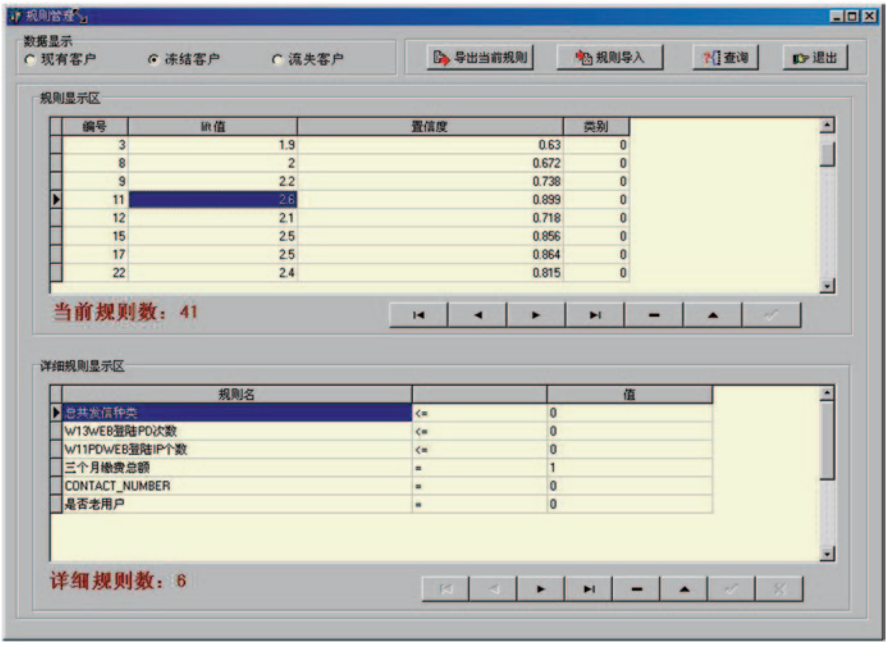


Fig. 7.7 Example software library of transformation knowledge acquisition

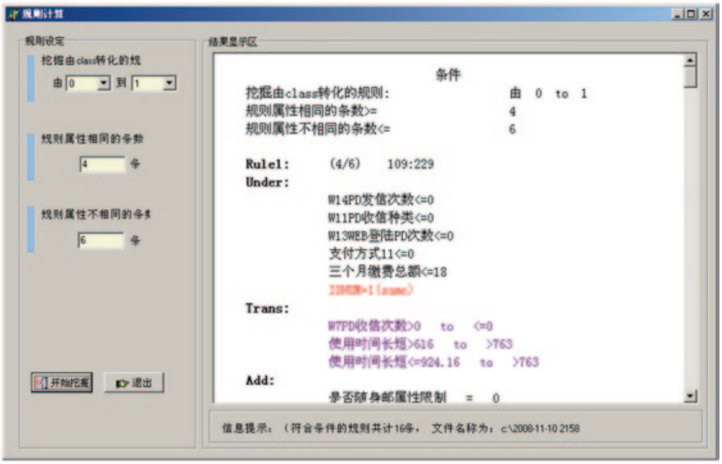


Fig. 7.8 Extension strategy mining interface

tive transformation knowledge plays a pivotal role on taking effective operational measures. The interface is shown as in Fig. 7.8:

We found that there are two paths for transformation knowledge mining by combining decision tree method with extension set theory.

1. The indirect rule mining method refers to further digging on the basis of traditional decision tree rule. After generating a static rule set through data mining of decision tree, coming the second excavation of the rule set.
2. Extension strategy direct mining method refers to directly dig out transformation knowledge on original data base by improving traditional decision tree algorithm.

The chapter mainly based on the first path to achieve the acquisition of transformation knowledge, the second path has more practicality and need further research. In the research process, we also found that transformation of classification of things need certain condition, the definition of transformation condition for qualitative and quantitative analysis is also research direction. there would be great application prospection by taking advantage of the result of extension theory research which connect traditional data mining with extension set, and with extension transformation as well as extension logical theory to dig out "can't to can, not to yes" strategy by using methods of extension data mining.