The Application of Multi-Rough Set Sorter Model to the Character Identification

Hong Xianggong[1], Wang Ping[1], Zhao An[1]

1.Information Engineering Academy of Nanchang University, Nanchang, Jiangxi Province, 330000

Abstract: The Article presents a kind of multi-rough set sorter model and use it in the field of character identification. First, it offers the procedures and steps of building a multi-rough set sorter model: many reduction sets come of the same training sample set through attribute reduction, every reduction set brings forth a regulation assemble, and every regulation assemble corresponds with a sorter. Then the article introduces the methods of Decision fusion of three kind of multi-rough set sorters, then compares and analyzes them by experiments, finally, choose one of them in practical objects. Practice proves this method receives a good identification result.

Keywords: Rough set, Attribute reduction, Rule extraction, Character recognition, Decision fusion

1 Introduction

The pattern recognition is more and more widely used as the application of computer technologies universalize in all professions, for example, character recognition, voice recognition, or face recognition. The pattern recognition system is composed of following parts: Pretreatment, feature extraction and decision recognition, any of them directly affects the final result of recognition. This article mainly studies and discusses the 3rd link-decision recognition. The explanation of following introduction and experiments take the recognition of character of airplane tickets as example.

Most temporary character recognition technologies use the Baye decision-making method and the neural network recognition methods which are more or less insufficient. For example, the Baye decision-making method requires the independence among attributes; the neural network recognition methods have the shortcomings of long-time training and non-intuitional results. Compared to other methods, the rough set-based sorter has the following advantages: (1) rough set sorter does not require Apriori knowledge such as the probability distribution of Statistics, or the Probability evaluation of the Dempster – Shafer Theory, etc; (2) Rough set can be more simplified and accurate through the reduction of attribute and rule redundant value, which makes decision-maker more easily to accept and comprehend.

Some researchers used rough set method to solute classification problem, for example, Literature [9] and [10], but they used the decision plan of single sorter, that is, reduces the attribute set and comes into a reduction set, from which extract rule set, and build a sorter based on the rule set. System of single sorter is built on one “good” reduction set. One of the insufficiencies is that much of the centralized hidden information is missed, the other is that there is no objective criteria can measure the quality of a reduction set, which makes the choice more subjective. The plan presents in this paper that based on multi-rough set sorter not only keeps the advantage of itself, but also overcomes the shortage of single rough set sorter, and received excellent result in practical objects through the combination of neural network-based isomerism sorter.

2 The Imagery Prcessing & Feature Extraction in the Recgnition of Airplane Tickets

The purpose of airplane tickets recognition is to identify the alphanumeric characters that printed on it. Airplane ticket has the following features when scanned

978-1-935068-01-3

by scanner into character image: (1) the typeface and size of the characters is various; (2) the characters on the bill are of big noise, always are broken, lack of side or full of dots. The characteristic extraction algorithm mainly analyzes according to constitution, which is based on the description method of geometry shape of characters, the features would be badly affected if noise point exists. Therefore we use the process which is offered by diagram 2.1 to deal with the images.

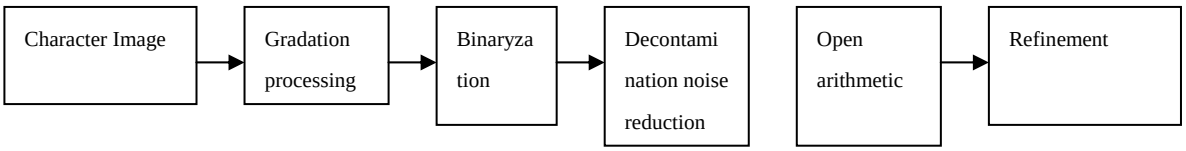

Diagram2.1 Processing of character image

The images processed by the method above will turn into a refined, noiseless and break point-less character image. For details of methods of image processing, please refer to pre-paper[4].

We use 14 features to state an input character. These 14 features can be divided into two categories: first is the extraction of four frames which apply template matching; the other use the method of constitution analyzing. For the way to extract and its introductions, please refer to related articles[3] and pre-papers[4].

We concluded 14 features after imagery processing and feature extraction, and saved training set in the form of two-dimensional form. The number of ranks are the number of training samples, and rows are the 14 features and their character values.

3 The Basic Algorith m of Building Multi Rough Set Sorter

The building of a multi rough set sorter requires following steps: (1) find core sets in the sample. (2)generate many reduction set through attribute reduction. (3) extract rule set which used to classify based on reduction set, then build multiple classifier system.

3.1 The Algorithm of Core Set

Mainly find the core attribute from the conditional attributes in the decision form.

Algorithm 1: Algorithm of Core Set

Input: $S=(U,A,V,f)$, $A=C\cup D$

Output: core(A)

Start:

（1）$core(A)=\{\ \}$

（2）For every c attribute of A

（3） If $sig(c)\neq 0$

（4） $core(A)=core(A)\cup\{c\}$

（5） End if

（6）End For

End.

3.2 The Algorithm of Attribute Reduction

The so-called attribute reduction is a subset of original reduction set, which has the same classification ability as the whole attribute set. The rough set theory is widely used in the data reduction and data analyzing, which can help reducing the complexity of calculation training process, and remove those redundant attributes, finally we get the rightsizing and efficient knowledge.

Algorithm 1: Algorithm of Attribute Reduction

Input： $S=(U,A,V,f)$, the core set of A core(A)

Output：a attribute reduction set RED（A）

Start：

（1）$B=B-core(A)$，RED（A）=A，

$C^* = core(A)$

（2）Arrange the elements of B in a certain order， $B^* = \{c_1^*, c_2^*, ..., c_k^*\}$

（3）For i=1 to $\|B^*\|$

（4） $C^* = C^* \bigcup \{c_i^*\}$

（5） If $POS_{C^*}(D) = POS_C(D)$

（6） $RED(A) = C^*$

（7） Break

（8） End If

（9）End For

End.

3.3 The Algorithm of Generating Multiple Reduction Set

From Algorithm 2, we can conclude that in (2), different ways of arrangements of B^* will result in different reduction sets. In order to make our decision system concise and also covers the implied information as much as possible, we choose the reduction set according to the following principles: (1) the reduction set contains the minimum reduction set which used the attribute importance as heuristic information. (2) the composite set of the reduction set should cover the original attributes as many as possible.

Algorithm 3: Algorithm of Multiple Reduction Set Generation

Input: $S = (U, A, V, f)$， the core set of A-CORE(A)

Output: the set of multiple reduction set RED.

Start:

（1） $B = B - CORE(A)$.

（2）descending sort the elements of B according to their importance，get B_1^*。

（3）according to Algorithm 2，use B_1^* generate the reduction set RED1。

（4） $K = RED1 - CORE(A)$

（5）For i=1 to $\|k\|$

(6) $B_i^* = A - CORE(A) - \{k_i\}$

(7) according to Algorithm 2，use B_i^* generate the reduction set RED_i

(8) $RED = RED \bigcup RED_i$

(9) If RED has all the attributes of A

(10) Exit circulation.

(11) End If

(12) End For

End.

3.3 The Algorithm of Extracting Rule Set

The rough set theory takes the rule sets in the decision form as a kind of mapping relationship from the condition attribute space to decision attribute space. Different reduction sets can generate different rule sets, every rule set can be taken as the implement of the decision form knowledge (Literature [8]). Generally, the algorithm of extracting rule set is gained from the algorithm of indecomposable matrices which requires creating and processing the $n \times n$ matrix, in which n refers to the saved objects stored in the form, obviously, this method does not fit the large datasets.

The search algorithm used in this article is span-prior heuristic algorithm [7]. Its basic principle is that, extending the nodes in turns, layer by layer, with the degree of approaching the starting node, then re-order and cut the nodes according to heuristic information, until the

978-1-935068-01-3 © 2009 SciRes.

problem is solved or the space is fully researched.

Algorithm 4: Algorithm of Extracting Rule Set

Input： $S=(U, A, V, f)$， $A=C\cup D$。

Output： rule set R'

Introduction： In this algorithm, open form is used to save the node of the search tree, close form is used to save the rules that has been extracted.

Start：

（1） For all the objects in U, x_i

（2） $open=\bigcup_{a\in A}\{a\}$， $close=\{\ \}$。

（3） If $open=\{\ \}$

（4） Continue

（5） End if

（6） Let Q be the first term of the open set， remove it from open set， and judge if it is in close form. If exists, turn to（3）. Otherwise, continue.

（7） If $[xi]_Q\subseteq[xi]_D$

（8） $R'=R'\cup\{Des([u]_Q)\Rightarrow Des([u]_D)\}$

（9） $close=close\bigcup\{Q\}$, turn to（3）

（10） Else

（11） Extending nodes, join the terminal of open according to following ways：for all $c'\in C$并且$c'\notin Q$， so $C'=\{c'\}\cup Q$， see if the subset of the new node C' is in the close form，if not exist, $open=open\cup\{C'\}$ ， turn to（3）

（12） End If

End.

4 The Comprehensive Decis ion of Multi Classifier System Process

The multi classifier system is constituted by many single classifiers. Its decision process is made from two parts: (1) single classifier which composes the multi classifier system made their own decision. (2) according to certain tactics of decision integration, conclude the results of all single classifier.

4.1 The De cision Pr ocessing Based on Single Classifier

Algorithm 5: Decision Processing Based on Single Classifier

Input： （1）the series W_i that visited by users is one of the samples that being tested.（2）the rule set R_k.

Output：predict result r_k

Start：

(1) For all the rules in R_k, R_{ki}

(2) If Wi fits the antecedent then of the rule R_{ki}

(3) $R=R\bigcup\{R_{ki}\}$

(4) End If

(5) End For

(6)If $R=\phi$ then

(7) $r_k=$ {Error Mark }

(8)Else

(9) calculate the support degree of all the rules in R，choose the rule $R_{\max}$ which is of largest

support degree，and take the seccedent of the rule $R_{\max}$ as predict result r_k, output.

(11)End If

End.

4.2 The Decision Tactics of the Multi Classifier System

In the decision system which based on multi classifier, one reduction set can only imply a part of the features of the sample, because every classifier is built on different reduction sets. Therefore, different classifiers have different recognition abilities towards various characters. The decision process which based on multi classifier can be taken as a process of decision integration of multi classifiers. For the details of methods and principles of decision integration, please refer to Literature [11], [12], and [13]. We concluded following methods according to the involved application features:

(1) Voting.

According to the principle that the minority is subordinate to the majority, choose the result that most probable as the final output from all classifiers.

(2) The Integration Technology Based on Evidence Theory

The core of this method is to calculate the belief functions that every classifier acts towards different results, and then calculate a whole belief function by using the compositional rule, according to which make final decision. Literature [8] applies this method.

(3) Decision Matrix Method

Calculate the decision matrix of every single classifier, and calculate their class conditional probabilities, according to which make a comprehensive judgment. For example, in this application, the decision matrix that corresponds to classifier is a 10×10 matrix. Let classifier be C_k，and its decision matrix $M_k=\left[r_{i,j}\right]_{10\times10}$, r_{ij} is the value i which is in the testing sample set, while decision result is the sample number of j. Therefore, when the output result of classifier C_k is i, the probability that actual result is the class condition of j is：$P_k(i\mid j)=\frac{r_{ij}}{R_i}$. R_i is the number of the samples that value is i in the training samples. The whole class conditional probability of multi classifier is：$P(i\mid j)=\prod_{k=1}^{\|c\|}\left(P_k(i\mid j)\right)$. $\|c\|$ is the number of classifiers in the multi classifier constitution. Finally, make decision according to the value of $P(i\mid j)$.

5 Design &Analyze the Experiment

There are three recognition results: (1) correct recognition; (2) error recognition; (3) refusal recognition. To a selected testing sample set R, we use P^+ to represent the number of samples that are correctly predicted, while P^- represents the number of samples that are erroneously predicted, and use recognition precision and recognition accuracy rate to represent the features of classifier:

$$\text{prediction precision}=\frac{P^+}{P^++P^-},$$

$$\text{prediction accuracy}=\frac{P^+}{\|R\|}$$ $\|R\|$ is the total number of testing samples.

In the experiment, we take the recognition of 10 digital characters and choose 100 samples per character in training stage. There are 1000 samples in number, while testing set chooses 50 samples per character, so 500 samples in total.

First, we get 5 reduction sets of the sample set by processing it by the method that mentioned in the 4th part, and then extract rule sets from the 5 sample sets one by one. Every rule set equals to a classifier.

In the stage of decision, we do the decision integration with the three methods that mentioned in 5.2,

① uses voting method，that is, the result would be output as final result when the 5 classifiers generate 3 approvals. ② applies the method in Literature [8]. ③ uses the method of decision matrix that mentioned in previous part, we choose 20 samples per character as decision matrix of the testing sample constitution classifier. While the final class conditional probability is changed to:

$$P(i \mid j) = \prod_{k=1}^{\|c\|} \left(P_k(i \mid j) + 0.001 \right)$$

As experiment result 1 implies:

Table 1. Experiment Results

	C1	C2	C3	C4	C5	①	②	③
Correct Recognition	450	439	443	436	432	420	465	472
Error Recognition	40	51	50	46	53	8	35	28
Refusal Recognition	10	10	7	18	15	72	0	0

Note：C1—C5 are 5 separate rule sets. ①，②，③：is the classifiers that built by the three methods that mentioned in part 5.2.

We can get the following conclusions according to the experiment results in the diagram above: （1）the correct recognition rate of the multi rule set integration method applied by methods ② and ③ is obviously superior to single rule sets（C1 to C5）. （2）method ① fits serial multi-level classifier for it has the highest recognition rate. In practical subjects, we do the recognition by using method ① as pre-classifier and the method of neural network as secondary classifier, the whole recognition rate can up to 70%.

6 Conclusion & Prospect

Character recognition is an important application of the model recognition theory, there is no recognition method that fits all kinds of situation, therefore, it is an efficient way to improve recognition rate by using multi classifier to make comprehensive decision. This paper first introduces the way to build multi rough set, then put the idea that multi classifier decision integration to the character recognition which based on rough set theory, and proves its feasibility by experiments. Next, we will focus our research on how to combine the advantage of all kinds of classifier so that they can display maximum efficiency.

Note: This project is supported by the Science Fund Object of Jiangxi Province (2007GZS1871).

References

[1] Pawlak, Z.: Rough Sets - Theoretical Aspects of Reasoning about Data. KluwerAcademic Publishers, 1991

[2] Zheng Wenxu:The theory and method of Rough set.Science Press,[2001]

[3] LiuQing:Roug set and Rough set reasoning.Science press,[2001]

[4] Ban ZhaoQi:Pattern Recongition.Tsinghua press.[2000]

[5] Meng-Chieh, L.: Software System For Intelligent Data Processing And Discovering Based On The Fuzzy-Rough Sets Theory. Thesis, San Diego State University, 1995

[6] Hu Ke yun, Lu Yu chang, Shi Chun yi : Advances in rough set theory and its applications. Beijing:Tsinghua University (Science & Technology) , 2001,(4) : 64～ 68

[7] Wong S KM , ZiarkoW : On optimal decision rules in decision tables〔J〕. Bulletin of Polish Academy of Science, 1985, 33: 693～696

[8] Ban ZhaoQi:Pattern Recongition. Tsinghua press .[2000]:47~54

[9] Zdzislaw Pawlak : The Rough Set View on Bayes 'Theorem. N.R. Pal and M. Sugeno (Eds.): AFSS 2002,LNAI 2275, pp. 106–116, 2002.

[10] Hu X : Knowledge discovery in database .an attribute oriented rough set approach . University of Regina,Canada, 1995.