

Research Article

A New Classification Approach Based on Multiple Classification Rules

Zhongmei Zhou

College of Computer Science, Minnan Normal University, Zhangzhou 363000, China

Correspondence should be addressed to Zhongmei Zhou; zzm@zju.edu.cn

Received 12 April 2014; Accepted 28 May 2014; Published 15 June 2014

Academic Editor: Jingjing Zhou

Copyright © 2014 Zhongmei Zhou. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

A good classifier can correctly predict new data for which the class label is unknown, so it is important to construct a high accuracy classifier. Hence, classification techniques are much useful in ubiquitous computing. Associative classification achieves higher classification accuracy than some traditional rule-based classification approaches. However, the approach also has two major deficiencies. First, it generates a very large number of association classification rules, especially when the minimum support is set to be low. It is difficult to select a high quality rule set for classification. Second, the accuracy of associative classification depends on the setting of the minimum support and the minimum confidence. In comparison with associative classification, some improved traditional rule-based classification approaches often produce a classification rule set that plays an important role in prediction. Thus, some improved traditional rule-based classification approaches not only achieve better efficiency than associative classification but also get higher accuracy. In this paper, we put forward a new classification approach called CMR (classification based on multiple classification rules). CMR combines the advantages of both associative classification and rule-based classification. Our experimental results show that CMR gets higher accuracy than some traditional rule-based classification methods.

1. Introduction

Classification is a pervasive data mining problem which has many applications, such as medical analysis, fraud detection, and network security [1]. A good classifier can correctly classify an object for which the class label is unknown. Therefore, building accurate and efficient classifier is one of the essential tasks of data mining. As a consequence, classification techniques are quite useful in ubiquitous computing [2–5]. The classification problem has been extensively studied by the research community. Various types of classification approaches have been proposed (e.g., KNN [6], Bayesian classifiers [7], decision trees [8], support vector machines [9], neural networks [10], and associative classifiers [11]). Classification is generally divided into two steps. First, we construct classification model based on the training dataset. Second, we use the model to predict new instances for which the class labels are unknown.

In recent years, associative classification has been investigated widely [11–19]. It integrates association rule mining algorithm and classification. Associative classification

induces a set of association classification rules from the training dataset which satisfies certain user-specified frequency and confidence. Then it selects a small set of high quality association classification rules and uses this rule set for prediction. The experimental results indicate in [11, 12] that associative classification gets higher accuracy than some traditional classification approaches such as decision trees [8]. In comparison with some traditional rule-based classification approaches, associative classification has two characteristics: (1) it generates a large number of association classification rules and (2) the measure support and confidence are used for evaluating the significance of association classification rules. However, associative classification has some weaknesses. First, it often generates a very large number of association classification rules in association rule mining, especially when the training dataset is large and dense. It takes great efforts to select a set of high quality classification rules from among them. Second, the accuracy of associative classification depends on the setting of the minimum support and the minimum confidence. Third, the efficiency of associative

classification is low when the minimum support is set to be low and the training dataset is large.

Traditional rule-based classification approaches also have been studied extensively [20–25]. Some traditional rule-based algorithms like FOIL [20], CN2 [21], and ELEM2 [22] discover a small set of high quality classification rules. They employ a sequential covering methodology and induce one rule at a time, and then they remove the positive instances that are covered by each new discovered rule. This rule induction process is done in a greedy fashion as it uses a heuristic function to select an attribute value to determine how each rule would be extended. They can achieve higher efficiency than associative classification. However, the accuracy of some traditional rule-based classification approaches may not be as high in some datasets. One of the reasons is that they usually generate a much small set of classification rules, especially when the training dataset is small. Some novel rule-based classification approaches have been proposed recently [22–25]. They can generate more classification rules than the one-rule-at-a-time algorithms and achieve higher accuracy. CPAR [23] keeps all close-to-the-best attribute values during the rule building process. By doing so, CPAR can select more attribute values and build several rules at one time. Thus, CPAR discovers more classification rules than FOIL. CPAR achieves higher average classification accuracy than associative classification algorithm CBA [11]. CATW algorithm [24] is different from PRM [23]. After an example is covered by a rule, CATW decreases both the tuple weight of the example and the weight of attribute values which the rule contains. As a result, CATW generates a much larger set of classification rules than FOIL. CATW achieves higher accuracy than CPAR and CMAR [12] in many cases. CMER [25] creates a seed set and a candidate set. It connects the seed set with the candidate set to generate several classification rules at a time. CMER generates a much larger set of classification rules than FOIL. The experimental results show that CMER achieves higher accuracy than FOIL.

In this paper, we propose a new classification approach called CMR (classification based on multiple classification rules). CMR combines the advantages of both associative classification and rule-based classification. It is distinguished from other rule induction methods in six aspects: (1) CMR uses multiple measures the minimum support and foil gain to select important attribute values. (2) CMR uses the minimum support and the minimum confidence to generate all classification rules with length one like associative classification. (3) CMR constructs a seed set and a candidate set. Both the seed set and the candidate set are selected from important attribute values. (4) CMR connects the seed set with the candidate set and then generates classification rules with length two and candidate patterns. (5) CMR adopts the minimum support and foil gain to build a new seed set using these candidate patterns. (6) CMR extends this new seed set and induces rules by adding the best attribute value one by one. The experimental results show that CMR gets higher accuracy than some traditional rule-based algorithms.

The outline of this paper is as follows. In Section 2, we comment on some related work. In Section 3, first we give some definitions. Second, we use an example to describe

the main ideas of inducing classification rules of CMR. Third, we develop the algorithm of CMR. Finally, we discuss how to predict class labels for unseen examples using the classification rule set generated by CMR. We report our experimental results in Section 4. We conclude the paper in Section 5.

2. Related Work

There are many classification approaches in classification domain. They are extremely different on inducing the classification rules from the training dataset and testing an object for which the class label is unknown. Our work presented in this paper is related to some existing researches of various classification methods. Therefore, we comment on some of these including the extraction of classification rules and testing strategies in the following.

(1) *Generating the Set of Classification Rules from the Training Dataset.* Decision tree method selects the best attribute with the highest information gain and then builds the decision tree as a classification rule set. It generates a small rule set, especially when the training dataset is small. Yin and Han [23] introduce three ways of generating the set of classification rules from the training dataset. They are FOIL, PRM, and CPAR. FOIL uses foil gain to select the best attribute value from the whole training dataset. And then it selects the best attribute value in the conditional database of the new selected attribute value. It adds it one by one to produce a classification rule. FOIL repeatedly searches for the current best rule and removes all the positive examples which are covered by the rule until all the positive examples in the dataset are removed. FOIL induces a small rule set as decision tree. PRM proposes an algorithm which modifies FOIL. PRM uses foil gain to select the best attribute value and adds attribute values one by one. However, after an example is covered by a rule, instead of removing it, its weight is decreased by multiplying a factor. PRM generates more rules than FOIL and each positive example is usually covered more than once. CPAR builds rules by adding attribute values one by one, which is similar to PRM. CPAR selects the best attribute value and keeps all close-to-the-best attribute values. CPAR connects the best attribute with the attribute values which are close to the best one. By doing so, CPAR selects more attribute values and builds several rules at one time. CATW [24] not only decreases the weight of an example by multiplying a factor, but also decreases the attribute values weight which the rules match after a rule is generated. ELEM2 takes into account the support of an attribute value and selects the most relevant attribute value for formulating rules. CMER [25] uses the support and foil gain to select several important attribute values to build the candidate set and the seed set. It connects the seed set with the candidate set and generates several classification rules at a time. CAEP [26] proposes a new measure growth rate for finding pattern. It produces emerge patterns with growth rate greater than or equal to the minimum threshold and then induces classification rule set with these patterns. In associative classification [11], association rule mining is used to generate candidate

rules, which includes all conjunctions of attribute values that meet the minimum support threshold. Then the measure confidence is used to generate the association classification rule set.

In this paper, we propose a new classification approach called CMR. First, CMR uses the support and confidence to generate classification rules with length one. Second, CMR connects the seed set with the candidate set and generates rules with length two. Third, CMR builds new seed set and then induces rules by connecting the new seed set with the attribute values which are the best in the conditional databases. Finally, CMR removes the examples which are covered by all the rules generated and iterates this process. CMR induces many rules at a time.

(2) *Classifying New Objects.* When predicting the class label, CBA uses the best rule whose body matches the example. CMAR uses multiple association classification rules and weighted Chi-square to measure the strength of group rule set under both conditional support and class distribution. CPAR selects the best k rules for each class and compares the average expected accuracy of each class. CAEP sums the contributions of the individual EPs. ELEM2 gives a decision score for each class that the matched rules indicate. CMR selects the best k rules for each class and compares the average decision score.

3. CMR: Classification Based on Multiple Classification Rules

In this part, we first give some definitions. Second, we use an example to describe the process of rule mining of CMR. Third we develop the algorithm of CMR. Finally, we give the measure of the significance of the classification rules and introduce how to predict new examples.

3.1. *Definitions.* Let $T = P \cup N$ be a training dataset with m distinguished attributes (A_1, A_2, \dots, A_m) . Let $C = \{C_1, C_2\}$ be the class set of T . We introduce some definitions as follows.

Definition 1 (foil gain). Suppose that v is an attribute value. There are $|P|$ positive examples and $|N|$ negative examples in T . The foil gain of v is defined as follows:

$$\text{fgain}(v) = |P_1| \left(\log \left(\frac{|P_1|}{(|P_1| + |N_1|)} \right) - \log \left(\frac{|P|}{(|P| + |N|)} \right) \right), \quad (1)$$

where $|p_1|$ is the number of positive examples which contain attribute value v and $|n_1|$ is the number of negative examples which contain attribute value v .

Definition 2 (support). The support of pattern X is defined as follows:

$$\text{sup}(X) = \left(\frac{\text{count}(X)}{|T|} \right) \times 100\%, \quad (2)$$

where $\text{count}(X)$ is the number of examples in T which contain pattern X and $|T|$ is the number of examples in dataset T .

Definition 3 (confidence). The confidence of pattern X is defined as follows:

$$\text{conf}(X) = \left(\frac{\text{count}(X_c)}{\text{count}(X)} \right) \times 100\%, \quad (3)$$

where $\text{count}(X_c)$ is the number of examples which contain pattern X and have a class C .

3.2. *Inducing Rules.* We first generate the rules with length one. Then we construct a candidate set and a seed set. Third, we connect the seed set with the candidate set and generate the rules with length 2. Fourth, we build a new seed set and generate the classification rules based on the new seed set. Finally, we remove the examples which are covered by the just-found rules and iterate the process. The following example shows the detailed process of inducing rules of CMR.

Example 4. The training dataset T is shown in Table 1. We suppose that the attribute J is the decision attribute and others are the condition attributes. In the training dataset T , we suppose that all examples which have the class ($J = 89$) are positive examples and all examples which have the class ($J = 90$) are negative examples. Let the minimum support be 20% and let the minimum confidence be 60%. We induce classification rules for the class ($J = 89$).

First, CMR generates the rules with length one. CMR adopts the minimum support and the minimum confidence as the measurement. We calculate the support and the confidence of all the attribute values in the positive examples. If the support of an attribute value p is greater than the minimum support, and if the confidence of p is greater than the minimum confidence, then p is selected as the rules with length one. We have rules with length one as shown in Table 2.

Second, CMR constructs the candidate set and the seed set. If the foil gain of an attribute value p in the positive examples is greater than the zero, then p is selected as an element of the candidate set. We use the average foil gain of all elements in the candidate set as the threshold of the minimum foil gain. If the foil gain of an attribute value p in the candidate set is greater than the minimum foil gain, then p is selected as an element of the seed set. We have the candidate set as shown in Table 3. In this example, the seed set consists of one element. It is the best attribute value in the candidate set. Attribute value 88 is the element of the seed set with the foil gain 1.150 and support 70%.

Third, CMR connects the seed set with the candidate set to produce patterns. If the confidence of a pattern X is greater than the minimum confidence, then $X \rightarrow C$ is a classification rule which has length two. If the support of a pattern X is greater than the minimum support, and the foil gain of X is greater than the minimum foil gain, then X is selected as an element of the new seed set. The new seed set consists of only one element. It is the pattern (88, 11).

Fourth, we generate rules by extending pattern (88, 11). CMR selects an attribute value with the best foil gain from

TABLE 1: The training data set.

	A	B	C	D	E	F	G	H	I	J
X1	1	9	19	30	55	60	72	79	82	89
X2	5	9	11	32	55	59	76	80	88	90
X3	3	7	14	35	53	60	73	81	88	89
X4	3	10	17	32	55	59	76	80	83	90
X5	3	9	11	35	55	58	72	81	88	89
X6	3	9	14	33	52	57	75	80	85	90
X7	4	9	11	36	54	59	73	80	88	89
X8	4	9	19	33	52	56	75	80	85	90
X9	3	9	11	37	54	60	73	81	88	89
X10	3	10	14	33	55	60	75	79	82	90

TABLE 2: The rule set with length one.

Attribute	Value	Support	Confidence
B	9	40%	66.7%
E	54	20%	66.7%
F	60	30%	75%

TABLE 3: The candidate set.

Attribute	Value	FGain	Support
I	88	1.150	70%
C	11	0.546	50%
D	35	0.546	50%
G	73	0.546	40%
H	81	0.546	50%

the conditional database of pattern (88, 11). In this example, attribute value 81 has the best foil gain. CMR forms the new pattern (88, 11, 81). CMR computes confidence of pattern (88, 11, 81). If the confidence of (88, 11, 81) is greater than the minimum confidence, then a classification rule is induced. Otherwise it continues finding the best attribute value from the conditional of pattern (88, 11, 81). After the rule which contains pattern (88, 11) is produced, CMR removes it from the new seed set and generates rules for other patterns in the new seed set until there are no patterns in the new seed set.

Finally, CMR removes the examples that are covered by rules produced and iterates the process.

3.3. The Algorithm of Inducing Rules of CMR. Algorithm 1 is the algorithm of inducing rules of CMR.

3.4. Predication Using Classification Rule Set. When we predict an example, we use all rules that satisfy the example. Three cases are possible for matching an example with a set of rules. There may be only one match (i.e., the example matches only one rule), more than one match (i.e., the example matches more than one rule), or no match (i.e., the example does not match any rules). If the matched rules do not agree on the class labels, we give a decision score for each class. Thus, we need to evaluate every rule to determine its

TABLE 4: Accuracy of CMR with different the minimum supports.

Size	3%	5%	7%
100	0.997456	0.996729	0.996729
200	0.993825	0.995822	0.995822
300	0.995639	0.996548	0.996548
400	0.996911	0.998365	0.998365
500	0.998911	0.998911	0.997821
600	0.993822	0.996002	0.995276
700	0.998910	0.999273	0.999273
800	0.998184	0.998184	0.998184
900	0.998910	0.999273	0.999273
1000	0.998910	0.999273	0.999273
Average	0.997148	0.997838	0.997656

predication power. The significance of a rule is defined as follows.

For rule $r : X \rightarrow C$, we use the expected accuracy to estimate the significance of rule r . The expected accuracy of rule r is given by (4) (denoted as SIG):

$$\text{SIG}(r) = \frac{(n_c + 1)}{(n + k)}, \quad (4)$$

where k is the number of classes, n is the total number of examples which contain pattern X , and n_c is the total number of examples which contain pattern X and have the class label C .

For a testing example, we select the best k rules which are matched by the example. If all the best k rules have the same class label, then the testing example is classified into this class. If the matched rules are not in the same class, CMR computes a decision score for each class. If the class label C has the number m rules which match the example, the decision score of the class label C is the average SIG of m rules. CMR classifies the example into the class with the highest decision score.

4. Experimental Results

All experiments are performed on mushroom characteristic dataset. The number of the testing dataset is set to be 500

Input: Training data set $T = P \cup N$ (P and N are the sets of all positive and negative examples, respectively), the minimum support, the minimum confidence.

Output: A set of classification rule

- (1) the rule set $R \leftarrow \emptyset$, the candidate set $cs \leftarrow \emptyset$, the seed set $ss \leftarrow \emptyset$, the frequent pattern $FP \leftarrow \emptyset$, the conditional positive example $P' \leftarrow \emptyset$, the conditional negative example $N' \leftarrow \emptyset$
- (2) while ($|P| > 0$) do
- (3) compute the support and the confidence of each attribute value p in P
- (4) if ($\text{sup}(p) > \text{minsup} \ \&\& \ \text{conf}(p) > \text{minconf}$)
- (5) $R \leftarrow R \cup \{p\}$
- (6) else if ($\text{sup}(p) > \text{minsup}$)
- (7) $FP \leftarrow FP \cup \{p\}$
- (8) end if
- (9) compute the foil gain of each attribute value p in FP
- (10) if ($\text{fgain}(p) > 0$)
- $cs \leftarrow cs \cup \{p\}$
- end if
- (11) compute the average foil gain of all attribute values in cs
- if ($\text{fgain}(p) > \text{average foil gain}$)
- $ss \leftarrow ss \cup \{p\}$
- end if
- (12) connect the seed set with the candidate set to generate patterns
- (13) compute the support and confidence of each pattern X
- (14) if ($\text{conf}(X) > \text{minconf}$)
- (15) $R \leftarrow R \cup \{X\}$
- (16) else if ($\text{sup}(X) > \text{minsup}$ and $\text{fgain}(X) > \text{the average foil gain}$)
- $ss \leftarrow ss \cup \{X\}$
- (17) else delete X
- (18) end if
- (19) For each pattern X in ss
- (20) if (the example t in positive example set P contains X)
- (21) $P' \leftarrow P' \cup \{t\}$
- (22) end if
- (23) if (the example t in negative example set N contains X)
- (24) $N' \leftarrow N' \cup \{t\}$
- (25) end if
- (26) while ($|P'| > 0$)
- (27) find the best attribute value p according to P' and N'
- (28) append p to r
- (29) remove all examples from P' that do not satisfying r
- (30) end while
- (31) end
- (32) remove all examples from positive example set P that satisfy the rule set R
- (33) end while

ALGORITHM 1

in all experiments. We select them from 0–5500 in turn in mushroom dataset. We select the best 3 rules for prediction.

In Table 4, we choose the size of the training dataset from 100 to 1000 in turn. We select the training dataset randomly in mushroom dataset. The minimum support is varied from 3% to 7%. From Table 4, we can see that CMR has different accuracy in different support. The average accuracy of CMR is the highest when the minimum support is 5%. However, the accuracy of CMR is not varied obviously by the minimum support.

Table 5 and Figure 1 show the accuracy of FOIL, CMR, and CMER, respectively. In Table 5 and Figure 1, the minimum support is set to be 5%. The minimum confidence is set to be 100%. From Table 5 and Figure 1, we can conclude that (1) the accuracy of CMER is higher than Foil, no matter how large the training date set is, (2) when the training dataset is small, the accuracy of CMR is much higher than Foil, (3) CMR achieves higher accuracy than CMER in many cases, and (4) CMR achieves higher average classification accuracy than CMER.

TABLE 5: Accuracy of FOIL, CMER and CMR.

Size	FOIL	CMER	CMR
100	0.988009	0.995457	0.996729
200	0.990371	0.996548	0.995822
300	0.988375	0.995275	0.996548
400	0.988014	0.997820	0.998365
500	0.992917	0.998182	0.998911
600	0.990919	0.997456	0.996002
700	0.996367	0.998183	0.999273
800	0.998545	1.000000	0.998184
900	0.996004	0.998183	0.999273
1000	0.996367	0.998183	0.999273
Average	0.992589	0.997529	0.997838

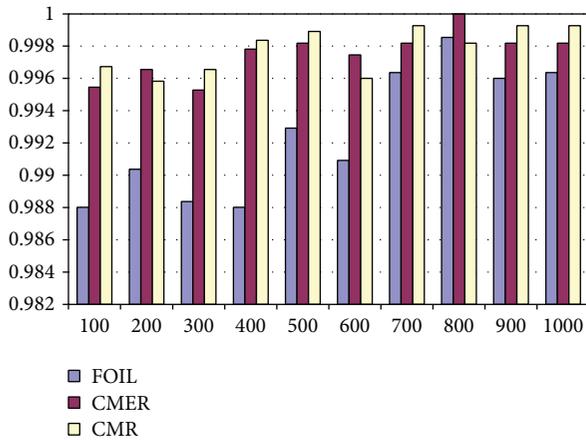


FIGURE 1: The accuracy of FOIL, CMER, and CMR.

5. Conclusions

Accuracy and efficiency are crucial factors in classification tasks in data mining. Associative classification gets higher accuracy than some traditional rule-based classification approaches in some cases. However, it generates a large number of association classification rules. Therefore, the efficiency of associative classification is not high when the minimum support is set to be low and the training dataset is large. In comparison with associative classification, one of the reasons that traditional rule-based classification methods cannot achieve high accuracy is that they often generate a few classification rules. In this paper, a new classification approach called CMR is proposed. CMR combines the advantages of both associative classification and rule-based classification. It induces many rules at a time. As a result, CMR generates much more classification rules than many other traditional rule-based classification methods, especially when the training dataset is small. Our experimental results show that the techniques developed in this paper are feasible. Our experimental results also show that CMR achieves high accuracy.

Conflict of Interests

The author declares that there is no conflict of interests regarding the publication of this paper.

Acknowledgments

This work is funded by the China NFS Program (no. 61170129) and by the Fujian province NSF Program (no. 2013J01259).

References

- [1] J. Han, K. Micheline, and J. Pei, *Data Mining: Concepts and Techniques*, The Morgan Kaufmann, 2011.
- [2] A. Pareek and M. Gupta, "Review of data mining techniques in cloud computing database," *International Journal of Advanced Computer Research*, vol. 2, no. 2, pp. 52–55, 2012.
- [3] R. S. Petre, "Data mining in cloud computing," *Database Systems Journal*, vol. 3, no. 3, pp. 67–71, 2012.
- [4] S. Adapa, M. K. Srinivas, and A. V. R. K. Harsha Vardhan Varma, "A study on cloud computing data mining," *International Journal of Innovative Research in Computer and Communication Engineering*, vol. 1, no. 5, pp. 1232–1237, 2013.
- [5] J. Wang, S. Wu, G. Li, and C. Wen, "An effective semi-supervised clustering framework integrating pairwise constraints and attribute preferences," *Computing and Informatics*, vol. 31, no. 3, pp. 597–612, 2012.
- [6] M. Zhang and Z. Zhou, "A k-nearest neighbor based algorithm for multi-label classification," in *Proceedings of the 1st IEEE International Conference on Granular Computing (GrC'05)*, vol. 2, pp. 718–721, Beijing, China, 2005.
- [7] B. Elena, C. Luca, and G. Paolo, "EnBay: a novel pattern-based Bayesian classifier," *IEEE Transactions on Knowledge and Data Engineering*, vol. 25, no. 12, pp. 2780–2795, 2013.
- [8] J. Quinlan, *C4.5: Programs for Machine Learning*, The Morgan Kaufmann, 1993.
- [9] Y. Li, T. K. James, and Z. Zhou, "Cost-sensitive semi-supervised support vector machine," in *Proceedings of the 24th AAAI Conference on Artificial Intelligence (AAAI '10)*, pp. 500–505, Atlanta, Ga, USA, July 2010.
- [10] Z. Zhou and X. Liu, "Training cost-sensitive neural networks with methods addressing the class imbalance problem," *IEEE Transactions on Knowledge and Data Engineering*, vol. 18, no. 1, pp. 63–77, 2006.
- [11] B. Liu, W. Hsu, and Y. Ma, "Integrating classification and association rule mining," in *Proceedings of the 4th International Conference on Knowledge Discovery and Data Mining (KDD '98)*, pp. 80–86, New York, NY, USA, 1998.
- [12] W. Li, J. Han, and J. Pei, "CMAR: accurate and efficient classification based on multiple class-association rules," in *Proceedings of the 1st IEEE International Conference on Data Mining (ICDM '01)*, pp. 369–376, San Jose, Calif, USA, December 2001.
- [13] G. Chen, H. Liu, L. Yu, Q. Wei, and X. Zhang, "A new approach to classification based on association rule mining," *Decision Support Systems*, vol. 42, no. 2, pp. 674–689, 2006.
- [14] F. A. Thabtah and P. I. Cowling, "A greedy classification algorithm based on association rule," *Applied Soft Computing Journal*, vol. 7, no. 3, pp. 1102–1111, 2007.

- [15] Y. Liu, Y. Jiang, X. Liu, and S.-L. Yang, "CSMC: a combination strategy for multi-class classification based on multiple association rules," *Knowledge-Based Systems*, vol. 21, no. 8, pp. 786–793, 2008.
- [16] S. Dua, H. Singh, and H. W. Thompson, "Associative classification of mammograms using weighted rules," *Expert Systems with Applications*, vol. 36, no. 5, pp. 9250–9259, 2009.
- [17] J. Song, Z. Ma, and Y. Xu, "DRAC: a direct rule mining approach for associative classification," in *Proceedings of the International Conference on Artificial Intelligence and Computational Intelligence (AICI '10)*, pp. 150–154, Sanya, China, October 2010.
- [18] M. Lin, T. Li, and S.-C. Hsueh, "Improving classification accuracy of associative classifiers by using k-conflict-rule preservation," in *Proceedings of the 7th International Conference on Ubiquitous Information Management and Communication (ICUIMC '13)*, ACM, Kota Kinabalu, Malaysia, January 2013.
- [19] W. Wang and Z. Zhou, "A review of associative classification approaches," *Transaction on IoT and Cloud Computing*, vol. 2, no. 2, pp. 31–42, 2014.
- [20] J. R. Quinlan and R. M. Cameron-Jones, "FOIL: a midterm report," in *Proceedings of the European Conference Machine Learning*, pp. 3–20, Vienna, Austria, 1993.
- [21] P. Clark and R. Boswell, "Rule induction with CN2: some recent improvements," in *Proceedings of the European Working Session on Learning (EWSL '91)*, pp. 151–163, Porto, Portugal, March 1991.
- [22] A. An, "Learning classification rules from data," *Computers & Mathematics with Applications*, vol. 45, no. 4-5, pp. 737–748, 2003.
- [23] X. Yin and J. Han, "CPAR: classification based on predictive association rules," in *Proceedings of the Society for Industrial and Applied Mathematics International Conference on Data Mining*, San Francisco, Calif, USA, May 2003.
- [24] Y. Zheng, Z. Huang, and T. He, "Classification based on both attribute value weight and tuple weight under the cloud computing," *Mathematical Problems in Engineering*, vol. 2013, Article ID 436368, 7 pages, 2013.
- [25] X. Wang, Z. Zhou, and G. Pan, "CMER: classification based on multiple excellent rules," *Journal of Theoretical and Applied Information Technology*, vol. 48, no. 1, pp. 661–665, 2013.
- [26] G. Dong, X. Zhang, L. Wong, and J. Li, "CAEP: classification by aggregating emerging patterns," in *Discovery Science*, vol. 1721 of *Lecture Notes in Computer Science*, pp. 30–42, Springer, Berlin, Germany, 1999.



Hindawi

Submit your manuscripts at
<http://www.hindawi.com>

