

Dynamic K : A Novel Satisfaction Mechanism for CAR-Based Classifiers

Raudel Hernández-León

Centro de Aplicaciones de Tecnologías de Avanzada (CENATAV),
7a # 21406 e/ 214 and 216, Rpto. Siboney, Playa, C.P. 12200, La Habana, Cuba
rhernandez@cenatav.co.cu

Abstract. In this paper, we propose a novel satisfaction mechanism, named “Dynamic K ”, which could be introduced in any Class Association Rules (CAR) based classifier, to determine the class of unseen transactions. Experiments over several datasets show that the new satisfaction mechanism has better performance than the main satisfaction mechanism reported (“Best Rule”, “Best K Rules” and “All Rules”). Additionally, the experiments show that “Dynamic K ” obtains the best results independent of the CAR-based classifier used.

Keywords: Supervised classification, Satisfaction mechanisms, Class association rules.

1 Introduction

The Classification Association Rule Mining (CARM) or associative classification, introduced in [2], integrates Classification Rule Mining (CRM) and Association Rule Mining (ARM). This integration involves mining a special subset of association rules, called Class Association Rules (CARs), using some quality measure (QM) to evaluate them. A classifier based on this approach usually consists of an ordered CAR list l , and a satisfaction mechanism for classifying unseen transactions using l [2,3,5]. Associative classification has been applied to many tasks including automatic error detection [19], detection of breast cancer [15], prediction of consumer behavior [17] and prediction of protein-protein interaction types [16], among others.

In associative classification, similar to ARM, a set of items $I = \{i_1, \dots, i_n\}$, a set of classes C , and a set of labeled transactions D , are given. Each transaction in D is represented by a set of items $X \subseteq I$ and a class $c \in C$. A lexicographic order among the items of I is assumed. The Support of an itemset $X \subseteq I$ is the fraction of transactions in D containing X . A CAR is an implication of the form $X \Rightarrow c$ where $X \subseteq I$ and $c \in C$. The most commonly used QM to evaluate CARs is the Confidence. The rule $X \Rightarrow c$ is held in D with certain Support s and Confidence α , where s is the fraction of transactions in D that contains $X \cup \{c\}$, and α is the probability of finding c in transactions that also contain X . A CAR $X \Rightarrow c$ covers a transaction t if $X \subseteq t$.

In general, CAR-based classifiers could be divided in two groups according to the strategy used for computing the set of CARs: (1) Two Stage classifiers - in a first stage, all CARs satisfying the Support and Confidence are mined and later, in a second stage, a classifier is built by selecting a small subset of CARs that fully covers the training set [2,3] and (2) Integrated classifiers - in these classifiers a small subset of CARs is directly generated [5,7,13].

Regardless of the strategy used for computing the set of CARs, in order to build the classifier we need to sort the CARs. In the literature, there are six main strategies for ordering CARs:

- a) CSA (Confidence - Support - Antecedent size): First, the rules are sorted in a descending order according to their Confidence. In case of ties, the tied CARs are sorted in a descending order according to their Support, and if the tie persist, CSA sorts the rules in ascending order according to the size of their rule antecedent [2].
- b) ACS (Antecedent size - Confidence - Support): This strategy is a variation of CSA, but it takes into account the size of the rule antecedent as first ordering criterion followed by Confidence and Support [7].
- c) SrQM (Specific rules - QM): First, the rules are sorted in a descending order according to the size of the CARs and in case of tie, the tied CARs are sorted in a descending order according to their quality measure [18].
- d) WRA (Weighted Relative Accuracy): The WRA rule ordering strategy assigns to each CAR a weight and then sorts the set of CARs in a descending order according to the assigned weights [12,14].
- e) LAP (Laplace Expected Error Estimate): LAP was introduced by Clark and Boswell [1] and it has been used to order CARs in CPAR classifier [5].
- f) χ^2 (Chi-Square): The χ^2 rule ordering strategy is a well known technique in statistics, which is used to determine whether two variables are independent or related. After computing an additive χ^2 value for each CAR, this value is used to sort the CARs in a descending order in the CMAR classifier [3].

Once the classifier has been built, we need to select a satisfaction mechanism for classifying unseen transactions. Three main satisfaction mechanisms have been reported in previous works [2,3,14]:

1. **Best Rule:** This mechanism selects the first (“best”) rule in the order that satisfies the transaction to be classified, and then the class associated to the selected rule is assigned to this transaction [2].
2. **Best K Rules:** This mechanism selects the best K rules (per each class) that satisfy the transaction to be classified and then the class is determined using these K rules, according to different criteria [14].
3. **All Rules:** This mechanism selects all rules that satisfy the transaction to be classified and use them to determine the class of the new transaction [3].

Classifiers following the “Best Rule” mechanism could suffer biased classification or overfitting since the classification is based on only one rule. On the other hand, the “All Rules” mechanism includes rules with low ranking for classification and this could affect the accuracy of the classifier. The “Best K Rules”

mechanism has been the most used satisfaction mechanism for CAR-based classifiers, reporting the best results. However, the classification accuracy could be affected when most of the best K rules were obtained extending the same item, or when there is an imbalance among the numbers of CARs with high Confidence (or another QM) values, per each class, that cover the new transaction.

In this paper, we propose a novel satisfaction mechanism, named “Dynamic K ”, to determine the class of unseen transactions. “Dynamic K ” obtains better performance than the main satisfaction mechanism mentioned above and it could be introduced in any CAR-based classifier. This paper is organized as follows: The next section describes our proposal. In the third section the experimental results are shown. Finally, the conclusions are given in section four.

2 Our Proposal

As we mentioned above, the three satisfaction mechanisms reported have limitations that can affect the classification accuracy. In general, the “Best K Rules” mechanism has been the most widely used for CAR-based classifiers, reporting the best results [11]. However, using this mechanism could affect the classification accuracy. Ever more when most of the best K rules were obtained extending the same item, or when there is an imbalance among the numbers of CARs with high QM values, per each class, that cover the new transaction (see examples 1 and 2, respectively).

Example 1. Given the set of CARs shown in Tables 1(a) and 1(b). Suppose that you would like to classifier the transaction $\{i_1, i_2, i_3, i_4, i_5, i_6\}$ using the “Best K Rules” mechanism with $K = 5$, value commonly used in previous works. First, the CARs are sorted with SrQM rule ordering strategy (see Tables 1(c) and 1(d)). Then, for each class, the top five rules covering the transaction $\{i_1, i_2, i_3, i_4, i_5, i_6\}$ are selected. The average of those top five rules QM values (delimited by a line in Tables 1(c) and 1(d)), are 0.93 and 0.91, respectively; which means that c_1 would be assigned. Note that all antecedents of the selected CARs, belonging to class c_1 , are subsets of $\{i_1, i_2, i_3, i_4\}$ (almost all rules were obtained from extensions of $\{i_1\} \Rightarrow c_1$). On the other hand, the antecedents of the top five rules in class c_2 (see Table 1(d)) involve a higher number of different items, resulting more intuitive to assign the class c_2 .

Example 2. Given the set of CARs shown in the Tables 2(a) and 2(b). Suppose that you would like to classifier the transaction $\{i_1, i_2, i_3, i_4\}$ using the “Best K Rules” mechanism with $K = 5$. First, the CARs are sorted with SrQM rule ordering strategy (see Tables 2(c) and 2(d)). Then, for each class, the top five rules covering the transaction $\{i_1, i_2, i_3, i_4\}$ are selected. The average of those top five rules QM values, delimited by a line in Tables 2(c) and 2(d), are 0.90 and 0.92 respectively; which means that c_2 would be assigned. However, note if you only consider the top three CARs of each class, you would obtain the averages 0.95 and 0.92, respectively; which means that c_1 would be assigned. That happens because, the class c_2 has more rules with QM greater than 0.90

than class c_1 ; and the “Best K Rules” mechanism does not take into account the imbalance among the number of CARs with high QM values, for each class, covering the new transaction.

Table 1. Example of two set of rules ((a) and (b)) and the result of sort them using the SrQM rule ordering strategy ((c) and (d))

(a)			(b)		
#	CAR	QM	#	CAR	QM
1	$\{i_1\} \Rightarrow c_1$	0.91	1	$\{i_2\} \Rightarrow c_2$	0.86
2	$\{i_1, i_2\} \Rightarrow c_1$	0.96	2	$\{i_2, i_3\} \Rightarrow c_2$	0.87
3	$\{i_1, i_2, i_3\} \Rightarrow c_1$	0.96	3	$\{i_2, i_3, i_4\} \Rightarrow c_2$	0.93
4	$\{i_1, i_2, i_3, i_4\} \Rightarrow c_1$	0.96	4	$\{i_2, i_4\} \Rightarrow c_2$	0.85
5	$\{i_1, i_3\} \Rightarrow c_1$	0.92	5	$\{i_2, i_4, i_5\} \Rightarrow c_2$	0.96
6	$\{i_2\} \Rightarrow c_1$	0.84	6	$\{i_3\} \Rightarrow c_2$	0.88
7	$\{i_2, i_3\} \Rightarrow c_1$	0.84	7	$\{i_3, i_5\} \Rightarrow c_2$	0.88
8	$\{i_2, i_3, i_4\} \Rightarrow c_1$	0.85	8	$\{i_3, i_5, i_6\} \Rightarrow c_2$	0.90
9	$\{i_3\} \Rightarrow c_1$	0.81	9	$\{i_4\} \Rightarrow c_2$	0.87
10	$\{i_3, i_4\} \Rightarrow c_1$	0.83	10	$\{i_4, i_5\} \Rightarrow c_2$	0.89

(c)			(d)		
#	CAR	QM	#	CAR	QM
1	$\{i_1, i_2, i_3, i_4\} \Rightarrow c_1$	0.96	1	$\{i_2, i_4, i_5\} \Rightarrow c_2$	0.96
2	$\{i_1, i_2, i_3\} \Rightarrow c_1$	0.96	2	$\{i_2, i_3, i_4\} \Rightarrow c_2$	0.93
3	$\{i_2, i_3, i_4\} \Rightarrow c_1$	0.85	3	$\{i_3, i_5, i_6\} \Rightarrow c_2$	0.90
4	$\{i_1, i_2\} \Rightarrow c_1$	0.96	4	$\{i_4, i_5\} \Rightarrow c_2$	0.89
5	$\{i_1, i_3\} \Rightarrow c_1$	0.92	5	$\{i_3, i_5\} \Rightarrow c_2$	0.88
6	$\{i_2, i_3\} \Rightarrow c_1$	0.84	6	$\{i_2, i_3\} \Rightarrow c_2$	0.87
7	$\{i_3, i_4\} \Rightarrow c_1$	0.83	7	$\{i_2, i_4\} \Rightarrow c_2$	0.85
8	$\{i_1\} \Rightarrow c_1$	0.91	8	$\{i_3\} \Rightarrow c_2$	0.88
9	$\{i_2\} \Rightarrow c_1$	0.84	9	$\{i_4\} \Rightarrow c_2$	0.87
10	$\{i_3\} \Rightarrow c_1$	0.81	10	$\{i_2\} \Rightarrow c_2$	0.86

Table 2. Example of two set of rules ((a) and (b)) and the result of sort them using the SrQM rule ordering strategy ((c) and (d))

(a)			(b)		
#	CAR	QM	#	CAR	QM
1	$\{i_1\} \Rightarrow c_1$	0.80	1	$\{i_1\} \Rightarrow c_2$	0.80
2	$\{i_1, i_2\} \Rightarrow c_1$	0.82	2	$\{i_2\} \Rightarrow c_2$	0.83
3	$\{i_1, i_2, i_3\} \Rightarrow c_1$	0.95	3	$\{i_2, i_3\} \Rightarrow c_2$	0.92
4	$\{i_2\} \Rightarrow c_1$	0.83	4	$\{i_2, i_3, i_4\} \Rightarrow c_2$	0.92
5	$\{i_2, i_3\} \Rightarrow c_1$	0.94	5	$\{i_3\} \Rightarrow c_2$	0.91
6	$\{i_3\} \Rightarrow c_1$	0.84	6	$\{i_3, i_4\} \Rightarrow c_2$	0.92
7	$\{i_3, i_4\} \Rightarrow c_1$	0.96	7	$\{i_4\} \Rightarrow c_2$	0.91

(c)			(d)		
#	CAR	QM	#	CAR	QM
1	$\{i_1, i_2, i_3\} \Rightarrow c_1$	0.95	1	$\{i_2, i_3, i_4\} \Rightarrow c_2$	0.92
2	$\{i_3, i_4\} \Rightarrow c_1$	0.96	2	$\{i_2, i_3\} \Rightarrow c_2$	0.92
3	$\{i_2, i_3\} \Rightarrow c_1$	0.94	3	$\{i_3, i_4\} \Rightarrow c_2$	0.92
4	$\{i_1, i_2\} \Rightarrow c_1$	0.82	4	$\{i_3\} \Rightarrow c_2$	0.91
5	$\{i_3\} \Rightarrow c_1$	0.84	5	$\{i_4\} \Rightarrow c_2$	0.91
6	$\{i_2\} \Rightarrow c_1$	0.83	6	$\{i_2\} \Rightarrow c_2$	0.83
7	$\{i_1\} \Rightarrow c_1$	0.80	7	$\{i_1\} \Rightarrow c_2$	0.80

In this paper, in order to overcome the drawbacks of the existing satisfaction mechanisms, we propose the “Dynamic K ” mechanism. First, “Dynamic K ” sorts the CARs using the SrQM rule ordering strategy. Later, “Dynamic K ”

selects, for each class $c \in C$, the set of rules $X \Rightarrow c$ covering the new transaction t and satisfying the following conditions:

- $X \Rightarrow c$ is a maximal rule.
- for all $i \in I$, with i lexicographically greater than all items of X , $QM(X \cup \{i\} \Rightarrow c) < QM(X \Rightarrow c)$ holds.

Thereby we included more large rules with high QM values in the classification, avoiding redundancies and including more different items in the antecedents of the selected CARs. For example, Tables 3 and 4 show the rules of examples 1 and 2 that were selected for “Dynamic K ” and cover transactions $\{i_1, i_2, i_3, i_4, i_5, i_6\}$ and $\{i_1, i_2, i_3, i_4\}$, respectively;

Table 3. Rules of Example 1 that were selected by “Dynamic K ” and cover transaction $\{i_1, i_2, i_3, i_4, i_5, i_6\}$

(a)		(b)	
# CAR	QM	# CAR	QM
1 $\{i_1, i_2, i_3, i_4\} \Rightarrow c_1$	0.96	1 $\{i_2, i_4, i_5\} \Rightarrow c_2$	0.96
2 $\{i_2, i_3, i_4\} \Rightarrow c_1$	0.85	2 $\{i_2, i_3, i_4\} \Rightarrow c_2$	0.93
3 $\{i_1, i_3\} \Rightarrow c_1$	0.92	3 $\{i_3, i_5, i_6\} \Rightarrow c_2$	0.90
4 $\{i_3, i_4\} \Rightarrow c_1$	0.83	4 $\{i_4, i_5\} \Rightarrow c_2$	0.89
Average	0.89	Average	0.92

Table 4. Rules of Example 2 that were selected by “Dynamic K ” and cover transaction $\{i_1, i_2, i_3, i_4\}$

(a)		(b)	
# CAR	QM	# CAR	QM
1 $\{i_1, i_2, i_3\} \Rightarrow c_1$	0.95	1 $\{i_2, i_3, i_4\} \Rightarrow c_2$	0.92
2 $\{i_3, i_4\} \Rightarrow c_1$	0.96	2 $\{i_3, i_4\} \Rightarrow c_2$	0.92
3 $\{i_2, i_3\} \Rightarrow c_1$	0.94	3 $\{i_4\} \Rightarrow c_2$	0.91
		4 $\{i_1\} \Rightarrow c_2$	0.80
Average	0.95	Average	0.92

Let N_i be the set of maximal CARs of class c_i that were selected for “Dynamic K ” mechanism. After selecting all N_i (for $i = 1$ to $|C|$), “Dynamic K ” assigns the class c_j such that the QM average of all rules of N_j is greater than the QM average of the top $|N_j|$ rules of each N_i , with $i \neq j$ and $|N_i| \geq |N_j|$. In case of tie among classes with different number of CARs, the class with less number of CARs is preferred because the CARs are sorted in descendent order according to their sizes (SrQM rule ordering strategy); in case of tie among classes with equals number of CARs, the class with greater Support is selected, if the tie persist the class is selected randomly.

If in examples 1 and 2, we applied the “Dynamic K ” mechanism to classify the transactions $\{i_1, i_2, i_3, i_4, i_5, i_6\}$ and $\{i_1, i_2, i_3, i_4\}$, respectively, we will obtain the class c_2 in the first example (see Table 3) and the class c_1 in the second one (see Table 4).

The “Dynamic K ” mechanism does not have the drawbacks of the other existent mechanisms since:

- It selects the maximal rules with high QM values, avoiding redundancies and allowing the inclusion of more different items in the antecedents of the selected CARs, thereby CARs of low quality are not included for classifying.
- The result is not affected when there is an imbalance among the numbers of CARs with high QM values, for each class, that cover the new transaction, this happens because to classify a new transaction, “Dynamic K ” considers the average of the same amount of CARs.
- It considers all good quality CARs that cover the new transaction and not only the best one. Thereby, “Dynamic K ” does not fall on the mistake of assuming that the best rule is going to classify correctly all transactions that it covers.

3 Experimental Results

In this section, we present the results of our experimental comparison among “Dynamic K ” and the main satisfaction mechanisms reported in the literature: “Best rule” [2], “Best K rules” [14] and “All rules” [3].

For the experiment showed in Table 5, the four satisfaction mechanisms were implemented inside the CAR-IC classifier [18], using the Confidence threshold set to 0.5 and the Support threshold set to 0.01, as it was reported in other works [3,5,7,14,20]. All our experiments were done using ten-fold cross-validation reporting the average over the ten folds. All the tests were performed on a PC with an Intel Core 2 Duo at 1.86 GHz CPU with 1 GB DDR2 RAM. Similar to other works [2,3,6,13,20], we used several datasets, specifically 20. The chosen datasets were originally taken from the UCI Machine Learning Repository [10], and their numerical attributes were discretized by Frans Coenen using the LUCS-KDD [4] discretized/normalized CARM Data Library.

For the experiment showed in Table 6, the codes of CBA, CMAR, CPAR and TFPC were downloaded from the Frans Coenen’s homepage (<http://www.csc.liv.ac.uk/~frans>) and the codes of DDPMine and CAR-IC was provided by their authors. For CBA, CMAR, CPAR, TFPC and CAR-IC classifiers we used the Confidence threshold set to 0.5 and the Support threshold set to 0.01, as their authors suggested.

In Table 5, the results show that “Dynamic K ” yields an average accuracy higher than all other evaluated mechanisms, having a difference of 0.92% with respect to the mechanism in the second place (“Best K rules” with K set to 5, the same value used in other works [5,11,12,14]). Additionally, “Dynamic K ” wins in 16 of the 20 datasets and ties in the other four. In Table 6, we show that “Dynamic K ” obtains the best results independent of the CAR-based classifier used, being CAR-IC the most benefited with the new mechanism (see rows 3 and 4 of Table 6).

Finally, in order to determine if the results shown in Table 5 are statistically significant, we performed a pairwise comparison between “Dynamic K ” and the other satisfaction mechanisms. Each cell in Table 7 contains the number of

Table 5. Classification accuracy of CAR-IC using the different satisfaction mechanisms

Dataset	Best rule	All rules	Best K rules	Dynamic K
adult	82.14	81.09	82.61	82.85
anneal	91.77	90.95	92.73	93.26
breast	84.45	83.52	90.03	90.46
connect4	55.98	55.01	56.02	57.24
dermatology	78.45	77.22	80.16	83.93
ecoli	82.04	80.46	82.06	82.16
flare	86.00	85.38	85.98	86.45
glass	68.10	67.29	68.95	71.12
heart	53.23	52.14	54.35	56.48
hepatitis	84.54	83.66	84.62	84.62
horseColic	82.48	81.75	82.47	84.54
ionosphere	84.06	83.02	86.10	86.24
iris	96.07	95.98	96.67	97.91
led7	72.70	71.43	73.02	73.02
letRecog	73.17	72.50	73.14	75.23
mushroom	98.51	97.86	98.54	98.54
pageBlocks	91.85	91.13	92.26	92.59
penDigits	77.83	76.42	81.93	82.78
pima	75.35	74.59	76.01	76.01
waveform	73.14	72.24	74.39	75.06
Average	79.59	78.68	80.60	81.52

Table 6. Average accuracy of other classifiers over the tested datasets, for different satisfaction mechanisms

Dataset	CBA	CMAR	CPAR	TFCP	DDPMine	CAR-IC
All rules	70.03	75.28	74.39	73.08	77.81	78.68
Best rule	71.81	76.49	75.64	74.21	78.88	79.59
Best K rules	72.71	77.99	76.86	75.76	79.72	80.60
Dynamic K	73.28	78.53	77.32	76.21	80.16	81.52

Table 7. Pairwise comparison between “Dynamic K ” and the other satisfaction mechanism, using CAR-IC classifier. Each cell shows the number of times “Dynamic K ” Win/Lose with respect to the corresponding mechanism over the 20 selected datasets.

	All rules	Best rule	Best K rules	Dynamic K
All rules		0/6	0/14	0/19
Best rule	6/0		0/5	0/10
Best K rules	14/0	5/0		0/7
Dynamic K	19/0	10/0	7/0	

datasets where “Dynamic K ” significantly Win/Lose to each other mechanism. We detected ties using a one-tailed T-Test [8] with significance level of 0.05. The results in the pairwise comparison reveal that the “Dynamic K ” mechanism beats all other evaluated mechanisms over most of the tested datasets.

4 Conclusions

In this paper, a new satisfaction mechanism, called “Dynamic K ”, was presented. “Dynamic K ” overcomes the drawbacks of the existing satisfaction mechanisms

and it could be introduced in any CAR-based classifier. Experimental results show that “Dynamic K ” obtains a classification accuracy higher than all other evaluated mechanisms, independent of the CAR-based classifier used.

References

1. Clark, P., Boswell, R.: Rule Induction with CN2: Some Recent Improvements. In: Proc. of European Working Session on Learning, pp. 151–163 (1991)
2. Liu, B., Hsu, W., Ma, Y.: Integrating classification and association rule mining. In: Proc. of the KDD, pp. 80–86 (1998)
3. Li, W., Han, J., Pei, J.: CMAR: Accurate and efficient classification based on multiple class-association rules. In: Proc. of the ICDM, pp. 369–376 (2001)
4. Coenen, F.: The LUCS-KDD discretised/normalised ARM and CARM Data Library (2003), <http://www.csc.liv.ac.uk/~frans/KDD/Software/LUCS-KDD-DN>
5. Yin, X., Han, J.: CPAR: Classification based on Predictive Association Rules. In: Proc. of the SIAM International Conference on Data Mining, pp. 331–335 (2003)
6. Wang, J., Karypis, G.: HARMONY: Efficiently mining the best rules for classification. In: Proc. of SDM, pp. 205–216 (2005)
7. Coenen, F., Leng, P., Zhang, L.: Threshold Tuning for Improved Classification Association Rule Mining. In: Ho, T.-B., Cheung, D., Liu, H. (eds.) PAKDD 2005. LNCS (LNAI), vol. 3518, pp. 216–225. Springer, Heidelberg (2005)
8. Demšar, J.: Statistical Comparisons of Classifiers over Multiple Data Sets. *J. Mach. Learn. Res.* 7, 1–30 (2006)
9. Steinbach, M., Kumar, V.: Generalizing the notion of confidence. *Knowl. Inf. Syst.* 12(3), 279–299 (2007)
10. Asuncion, A., Newman, D.J.: UCI Machine Learning Repository (2007), <http://www.ics.uci.edu/~mllearn/{MLR}epository.html>
11. Wang, Y.J., Xin, Q., Coenen, F.: A Novel Rule Weighting Approach in Classification Association Rule Mining. In: International Conference on Data Mining Workshops, pp. 271–276 (2007)
12. Wang, Y.J., Xin, Q., Coenen, F.: A Novel Rule Ordering Approach in Classification Association Rule Mining. In: Perner, P. (ed.) MLDM 2007. LNCS (LNAI), vol. 4571, pp. 339–348. Springer, Heidelberg (2007)
13. Cheng, H., Yan, X., Han, J., Philip, S.Y.: Direct Discriminative Pattern Mining for Effective Classification. In: Proc. of the ICDE, pp. 169–178 (2008)
14. Wang, Y.J., Xin, Q., Coenen, F.: Hybrid Rule Ordering in Classification Association Rule Mining. *Trans. MLDM* 1(1), 1–15 (2008)
15. Karabatak, M., Ince, M.C.: An expert system for detection of breast cancer based on association rules and neural network. *Expert Syst. Appl.* 36, 3465–3469 (2009)
16. Park, S.H., Reyes, J.A., Gilbert, D.R., Kim, J.W., Kim, S.: Prediction of protein-protein interaction types using association rule based classification. *BMC Bioinformatics* 10(1) (2009)
17. Bae, J.K., Kim, J.: Integration of heterogeneous models to predict consumer behavior. *Expert Syst. Appl.* 37, 1821–1826 (2010)
18. Hernández, R., Carrasco, J.A., Martínez, F.J., Hernández, J.: Classifying using Specific Rules with High Confidence. In: Proc. of the MICA, pp. 75–80 (2010)
19. Malik, W.A., Unwin, A.: Automated error detection using association rules. *Intelligent Data Analysis* 15(5), 749–761 (2011)
20. Hernández, R., Carrasco, J.A., Martínez, F.J., Hernández, J.: CAR-NF: A Classifier based on Specific Rules with High Netconf. *Intelligent Data Analysis* 16(1), 49–68 (2012)