Marcin S. Szczuka Daniel Howard
Dominik Ślęzak Haeng-kon Kim
Tai-hoon Kim Il-seok Ko
Geuk Lee Peter M.A. Sloot (Eds.)

LNAI 4413

Advances in Hybrid Information Technology

First International Conference, ICHIT 2006
Jeju Island, Korea, November 2006
Revised Selected Papers

Springer

Marcin S. Szczuka Daniel Howard
Dominik Ślęzak Haeng-kon Kim
Tai-hoon Kim Il-seok Ko Geuk Lee
Peter M.A. Sloot (Eds.)

Advances in Hybrid Information Technology

First International Conference, ICHIT 2006
Jeju Island, Korea, November 9-11, 2006
Revised Selected Papers

Springer

Series Editors

Jaime G. Carbonell, Carnegie Mellon University, Pittsburgh, PA, USA
Jörg Siekmann, University of Saarland, Saarbrücken, Germany

Volume Editors

Marcin S. Szczuka
Warsaw University, Poland, szczuka@mimuw.edu.pl

Daniel Howard
QinetiQ Group PLC, Malvern, UK, dhoward@taz.qinetiq.com

Dominik Ślęzak
Infobright Inc, Toronto, ON, Canada, slezak@infobright.com

Haeng-kon Kim
Daegu Catholic University, Gyeongbuk, Korea, hangkon@cu.ac.kr

Tai-hoon Kim
Hannam University, Daejeon, Korea, taihoonn@hannam.ac.kr

Il-seok Ko
Dongguk University, Korea, isko@dongguk.edu

Geuk Lee
Hannam University, Daejeon, Korea, leegeuk@ce.hannam.ac.kr

Peter M.A. Sloot
University of Amsterdam, The Netherlands, sloot@science.uva.nl

Library of Congress Control Number: 2007941600

CR Subject Classification (1998): I.2, H.4, H.3, J.1, H.5, K.6, K.4

LNCS Sublibrary: SL 7 – Artificial Intelligence

ISSN 0302-9743
ISBN-10 3-540-77367-3 Springer Berlin Heidelberg New York
ISBN-13 978-3-540-77367-2 Springer Berlin Heidelberg New York

Springer is a part of Springer Science+Business Media

springer.com

Printed in Germany

Typesetting: Camera-ready by author, data conversion by Scientific Publishing Services, Chennai, India
Printed on acid-free paper SPIN: 12206988 06/3180 5 4 3 2 1 0

Lecture Notes in Artificial Intelligence 4413

Edited by J. G. Carbonell and J. Siekmann

Subseries of Lecture Notes in Computer Science

Lecture Notes in Artificial Intelligence (LNAI)

Vol. 4874: J. Neves, M.F. Santos, J.M. Machado (Eds.), Progress in Artificial Intelligence. XVIII, 704 pages. 2007.

Vol. 4869: F. Botana, T. Recio (Eds.), Automated Deduction in Geometry. X, 213 pages. 2007.

Vol. 4850: M. Lungarella, F. Iida, J. Bongard, R. Pfeifer (Eds.), 50 Years of Artificial Intelligence. X, 399 pages. 2007.

Vol. 4845: N. Zhong, J. Liu, Y. Yao, J. Wu, S. Lu, K. Li (Eds.), Web Intelligence Meets Brain Informatics. XI, 516 pages. 2007.

Vol. 4830: M.A. Orgun, J. Thornton (Eds.), AI 2007: Advances in Artificial Intelligence. XIX, 841 pages. 2007.

Vol. 4828: M. Randall, H.A. Abbass, J. Wiles (Eds.), Progress in Artificial Life. XII, 402 pages. 2007.

Vol. 4827: A. Gelbukh, Á.F. Kuri Morales (Eds.), MICAI 2007: Advances in Artificial Intelligence. XXIV, 1234 pages. 2007.

Vol. 4826: P. Perner, O. Salvetti (Eds.), Advances in Mass Data Analysis of Signals and Images in Medicine, Biotechnology and Chemistry. X, 183 pages. 2007.

Vol. 4819: T. Washio, Z.-H. Zhou, J.Z. Huang, X. Hu, J. Li, C. Xie, J. He, D. Zou, K.-C. Li, M.M. Freire (Eds.), Emerging Technologies in Knowledge Discovery and Data Mining. XIV, 675 pages. 2007.

Vol. 4798: Z. Zhang, J.H. Siekmann (Eds.), Knowledge Science and Engineering and Management. XVI, 669 pages. 2007.

Vol. 4795: F. Schilder, G. Katz, J. Pustejovsky (Eds.), Annotating, Extracting and Reasoning about Time and Events. VII, 141 pages. 2007.

Vol. 4790: N. Dershowitz, A. Voronkov (Eds.), Logic for Programming, Artificial Intelligence, and Reasoning. XIII, 562 pages. 2007.

Vol. 4788: D. Borrajo, L. Castillo, J.M. Corchado (Eds.), Current Topics in Artificial Intelligence. XI, 280 pages. 2007.

Vol. 4775: A. Esposito, M. Faundez-Zanuy, E. Keller, M. Marinaro (Eds.), Verbal and Nonverbal Communication Behaviours. XII, 325 pages. 2007.

Vol. 4772: H. Prade, V.S. Subrahmanian (Eds.), Scalable Uncertainty Management. X, 277 pages. 2007.

Vol. 4766: N. Maudet, S. Parsons, I. Rahwan (Eds.), Argumentation in Multi-Agent Systems. XII, 211 pages. 2007.

Vol. 4755: V. Corruble, M. Takeda, E. Suzuki (Eds.), Discovery Science. XI, 298 pages. 2007.

Vol. 4754: M. Hutter, R.A. Servedio, E. Takimoto (Eds.), Algorithmic Learning Theory. XI, 403 pages. 2007.

Vol. 4737: B. Berendt, A. Hotho, D. Mladenic, G. Semeraro (Eds.), From Web to Social Web: Discovering and Deploying User and Content Profiles. XI, 161 pages. 2007.

Vol. 4733: R. Basili, M.T. Pazienza (Eds.), AI*IA 2007: Artificial Intelligence and Human-Oriented Computing. XVII, 858 pages. 2007.

Vol. 4724: K. Mellouli (Ed.), Symbolic and Quantitative Approaches to Reasoning with Uncertainty. XV, 914 pages. 2007.

Vol. 4722: C. Pelachaud, J.-C. Martin, E. André, G. Chollet, K. Karpouzis, D. Pelé (Eds.), Intelligent Virtual Agents. XV, 425 pages. 2007.

Vol. 4720: B. Konev, F. Wolter (Eds.), Frontiers of Combining Systems. X, 283 pages. 2007.

Vol. 4702: J.N. Kok, J. Koronacki, R. Lopez de Mantaras, S. Matwin, D. Mladenič, A. Skowron (Eds.), Knowledge Discovery in Databases: PKDD 2007. XXIV, 640 pages. 2007.

Vol. 4701: J.N. Kok, J. Koronacki, R. Lopez de Mantaras, S. Matwin, D. Mladenič, A. Skowron (Eds.), Machine Learning: ECML 2007. XXII, 809 pages. 2007.

Vol. 4696: H.-D. Burkhard, G. Lindemann, R. Verbrugge, L.Z. Varga (Eds.), Multi-Agent Systems and Applications V. XIII, 350 pages. 2007.

Vol. 4694: B. Apolloni, R.J. Howlett, L. Jain (Eds.), Knowledge-Based Intelligent Information and Engineering Systems, Part III. XXIX, 1126 pages. 2007.

Vol. 4693: B. Apolloni, R.J. Howlett, L. Jain (Eds.), Knowledge-Based Intelligent Information and Engineering Systems, Part II. XXXII, 1380 pages. 2007.

Vol. 4692: B. Apolloni, R.J. Howlett, L. Jain (Eds.), Knowledge-Based Intelligent Information and Engineering Systems, Part I. LV, 882 pages. 2007.

Vol. 4687: P. Petta, J.P. Müller, M. Klusch, M. Georgeff (Eds.), Multiagent System Technologies. X, 207 pages. 2007.

Vol. 4682: D.-S. Huang, L. Heutte, M. Loog (Eds.), Advanced Intelligent Computing Theories and Applications. XXVII, 1373 pages. 2007.

Vol. 4676: M. Klusch, K.V. Hindriks, M.P. Papazoglou, L. Sterling (Eds.), Cooperative Information Agents XI. XI, 361 pages. 2007.

Vol. 4667: J. Hertzberg, M. Beetz, R. Englert (Eds.), KI 2007: Advances in Artificial Intelligence. IX, 516 pages. 2007.

Vol. 4660: S. Džeroski, L. Todorovski (Eds.), Computational Discovery of Scientific Knowledge. X, 327 pages. 2007.

Vol. 4659: V. Mařík, V. Vyatkin, A.W. Colombo (Eds.), Holonic and Multi-Agent Systems for Manufacturing. VIII, 456 pages. 2007.

Vol. 4651: F. Azevedo, P. Barahona, F. Fages, F. Rossi (Eds.), Recent Advances in Constraints. VIII, 185 pages. 2007.

Vol. 4648: F. Almeida e Costa, L.M. Rocha, E. Costa, I. Harvey, A. Coutinho (Eds.), Advances in Artificial Life. XVIII, 1215 pages. 2007.

Vol. 4635: B. Kokinov, D.C. Richardson, T.R. Roth-Berghofer, L. Vieu (Eds.), Modeling and Using Context. XIV, 574 pages. 2007.

Vol. 4632: R. Alhajj, H. Gao, X. Li, J. Li, O.R. Zaïane (Eds.), Advanced Data Mining and Applications. XV, 634 pages. 2007.

Vol. 4629: V. Matoušek, P. Mautner (Eds.), Text, Speech and Dialogue. XVII, 663 pages. 2007.

Vol. 4626: R.O. Weber, M.M. Richter (Eds.), Case-Based Reasoning Research and Development. XIII, 534 pages. 2007.

Vol. 4617: V. Torra, Y. Narukawa, Y. Yoshida (Eds.), Modeling Decisions for Artificial Intelligence. XII, 502 pages. 2007.

Vol. 4612: I. Miguel, W. Ruml (Eds.), Abstraction, Reformulation, and Approximation. XI, 418 pages. 2007.

Vol. 4604: U. Priss, S. Polovina, R. Hill (Eds.), Conceptual Structures: Knowledge Architectures for Smart Applications. XII, 514 pages. 2007.

Vol. 4603: F. Pfenning (Ed.), Automated Deduction – CADE-21. XII, 522 pages. 2007.

Vol. 4597: P. Perner (Ed.), Advances in Data Mining. XI, 353 pages. 2007.

Vol. 4594: R. Bellazzi, A. Abu-Hanna, J. Hunter (Eds.), Artificial Intelligence in Medicine. XVI, 509 pages. 2007.

Vol. 4585: M. Kryszkiewicz, J.F. Peters, H. Rybinski, A. Skowron (Eds.), Rough Sets and Intelligent Systems Paradigms. XIX, 836 pages. 2007.

Vol. 4578: F. Masulli, S. Mitra, G. Pasi (Eds.), Applications of Fuzzy Sets Theory. XVIII, 693 pages. 2007.

Vol. 4573: M. Kauers, M. Kerber, R. Miner, W. Windsteiger (Eds.), Towards Mechanized Mathematical Assistants. XIII, 407 pages. 2007.

Vol. 4571: P. Perner (Ed.), Machine Learning and Data Mining in Pattern Recognition. XIV, 913 pages. 2007.

Vol. 4570: H.G. Okuno, M. Ali (Eds.), New Trends in Applied Artificial Intelligence. XXI, 1194 pages. 2007.

Vol. 4565: D.D. Schmorrow, L.M. Reeves (Eds.), Foundations of Augmented Cognition. XIX, 450 pages. 2007.

Vol. 4562: D. Harris (Ed.), Engineering Psychology and Cognitive Ergonomics. XXIII, 879 pages. 2007.

Vol. 4548: N. Olivetti (Ed.), Automated Reasoning with Analytic Tableaux and Related Methods. X, 245 pages. 2007.

Vol. 4539: N.H. Bshouty, C. Gentile (Eds.), Learning Theory. XII, 634 pages. 2007.

Vol. 4529: P. Melin, O. Castillo, L.T. Aguilar, J. Kacprzyk, W. Pedrycz (Eds.), Foundations of Fuzzy Logic and Soft Computing. XIX, 830 pages. 2007.

Vol. 4520: M.V. Butz, O. Sigaud, G. Pezzulo, G. Baldassarre (Eds.), Anticipatory Behavior in Adaptive Learning Systems. X, 379 pages. 2007.

Vol. 4511: C. Conati, K. McCoy, G. Paliouras (Eds.), User Modeling 2007. XVI, 487 pages. 2007.

Vol. 4509: Z. Kobti, D. Wu (Eds.), Advances in Artificial Intelligence. XII, 552 pages. 2007.

Vol. 4496: N.T. Nguyen, A. Grzech, R.J. Howlett, L.C. Jain (Eds.), Agent and Multi-Agent Systems: Technologies and Applications. XXI, 1046 pages. 2007.

Vol. 4483: C. Baral, G. Brewka, J. Schlipf (Eds.), Logic Programming and Nonmonotonic Reasoning. IX, 327 pages. 2007.

Vol. 4482: A. An, J. Stefanowski, S. Ramanna, C.J. Butz, W. Pedrycz, G. Wang (Eds.), Rough Sets, Fuzzy Sets, Data Mining and Granular Computing. XIV, 585 pages. 2007.

Vol. 4481: J. Yao, P. Lingras, W.-Z. Wu, M.S. Szczuka, N.J. Cercone, D. Ślęzak (Eds.), Rough Sets and Knowledge Technology. XIV, 576 pages. 2007.

Vol. 4476: V. Gorodetsky, C. Zhang, V.A. Skormin, L. Cao (Eds.), Autonomous Intelligent Systems: Multi-Agents and Data Mining. XIII, 323 pages. 2007.

Vol. 4460: S. Aguzzoli, A. Ciabattoni, B. Gerla, C. Manara, V. Marra (Eds.), Algebraic and Proof-theoretic Aspects of Non-classical Logics. VIII, 309 pages. 2007.

Vol. 4457: G.M.P. O'Hare, A. Ricci, M.J. O'Grady, O. Dikenelli (Eds.), Engineering Societies in the Agents World VII. XI, 401 pages. 2007.

Vol. 4456: Y. Wang, Y.-m. Cheung, H. Liu (Eds.), Computational Intelligence and Security. XXIII, 1118 pages. 2007.

Vol. 4455: S. Muggleton, R. Otero, A. Tamaddoni-Nezhad (Eds.), Inductive Logic Programming. XII, 456 pages. 2007.

Vol. 4452: M. Fasli, O. Shehory (Eds.), Agent-Mediated Electronic Commerce. VIII, 249 pages. 2007.

Vol. 4451: T.S. Huang, A. Nijholt, M. Pantic, A. Pentland (Eds.), Artifical Intelligence for Human Computing. XVI, 359 pages. 2007.

Vol. 4442: L. Antunes, K. Takadama (Eds.), Multi-Agent-Based Simulation VII. X, 189 pages. 2007.

Vol. 4441: C. Müller (Ed.), Speaker Classification II. X, 309 pages. 2007.

Vol. 4438: L. Maicher, A. Sigel, L.M. Garshol (Eds.), Leveraging the Semantics of Topic Maps. X, 257 pages. 2007.

Vol. 4434: G. Lakemeyer, E. Sklar, D.G. Sorrenti, T. Takahashi (Eds.), RoboCup 2006: Robot Soccer World Cup X. XIII, 566 pages. 2007.

Vol. 4429: R. Lu, J.H. Siekmann, C. Ullrich (Eds.), Cognitive Systems. X, 161 pages. 2007.

Vol. 4428: S. Edelkamp, A. Lomuscio (Eds.), Model Checking and Artificial Intelligence. IX, 185 pages. 2007.

¥671.00元

Preface

As information technologies (IT) become specialized and fragmented, it is easy to lose sight that many topics in IT have common threads and because of this, advances in one sub-discipline may transmit to another. The presentation of results between different sub-disciplines of IT encourages this interchange for the advancement of IT as a whole. Of particular interest is the hybrid approach or combining ideas from one discipline with those of another to achieve a result that is more significant than the sum of the individual parts. Through this hybrid philosophy, a new or common principle can be discovered which has the propensity to propagate throughout this multifaceted discipline.

This volume comprises the selection of extended versions of papers that were presented in their shortened form at the 2006 International Conference on Hybrid Information Technology (http://www.sersc.org/ICHIT2006/). Sixty-four papers out of the 235 that were published in ICHIT 2006 electronic proceedings were deemed suitable for inclusion in this volume, in a selection that was guided by technical quality and relevance to the balance of topics in hybrid information technology. The conference reflected a change in the thinking of scientists and practitioners, who now tend to join their efforts within multidisciplinary projects. As a consequence, the readers may observe that many papers might conceivably be classified into more than one chapter, given their interdisciplinary scope. The contributions in this monograph are clustered into six chapters: Data Analysis, Modeling, and Learning (11 papers); Imaging, Speech, and Complex Data (11 papers); Applications of Artificial Intelligence (11 papers); Hybrid, Smart, and Ubiquitous Systems (11 papers); Hardware and Software Engineering (9 papers); as well as Networking and Telecommunications (11 papers).

We would like to acknowledge the great effort of all in the ICHIT 2006 International Advisory Board and members of the International Program Committee of ICHIT 2006, as well as all organizations and individuals who supported the idea of publishing these advances in hybrid information technology, including SERSC (http://www.sersc.org/) and Springer. We strongly believe in the need for continuing this undertaking in the future, in the form of a conference, journal, or book series. In this respect we welcome any feedback.

October 2007

Marcin Szczuka
Daniel Howard
Dominik Ślęzak
Haeng-kon Kim
Tai-hoon Kim
Il-seok Ko
Geuk Lee
Peter Sloot

Table of Contents

Data Analysis, Modelling, and Learning

Imaging, Speech, and Complex Data

Applications of Artificial Intelligence

Hybrid, Smart, and Ubiquitous Systems

Hardware and Software Engineering

Networking and Telecommunications

Taking Class Importance into Account

José-Luis Polo, Fernando Berzal, and Juan-Carlos Cubero

Department of Computer Science and Artificial Intelligence,
University of Granada, 18071 Granada, Spain
{jlpolo,fberzal,JC.Cubero}@decsai.ugr.es

Abstract. In many classification problems, some classes are more important than others from the users' perspective. In this paper, we introduce a novel approach, *weighted classification*, to address this issue by modeling class importance through weights in the [0,1] interval. We also propose novel metrics to evaluate the performance of classifiers in a weighted classification context. In addition, we make some modifications to the ART[1] classification model [1] in order to deal with weighted classification.

1 Introduction

Classification is an extensively studied problem in Machine Learning research. Despite this, many classification problems exhibit specific features that render most classification models ineffective. Several models have been proposed to deal with class attribute peculiarities:

- In *imbalanced classification* [7] [8] problems, some classes are supported by a very low number of examples. Unfortunately, traditional classification models tend to ignore such classes, no matter what their importance is.
- *Cost-sensitive classification* [4] [5] [6] models take into account misclassification costs. These models are useful when the cost of a false positive is not the same for every class.
- *Subgroup discovery* [9]: In this case, there is only a class that is important for the expert. The aim of subgroup discovery is finding the most interesting subgroups of examples according to statistical criteria.

In this paper, we focus on a class attribute feature that is ignored by traditional classification models: the relative importance of each class.

2 The Weighted Classification Problem

Each class in a classification problem may have a different degree of importance. In some situations, the user might be interested in achieving the maximum possible accuracy for specific classes while keeping the classification model complexity

[1] ART in this paper stands for Association Rule Tree, not to be confused with Adaptive Resonance Theory - a paradigm commonly associated with Artificial Neural Networks.

M.S. Szczuka et al. (Eds.): ICHIT 2006, LNAI 4413, pp. 1–10, 2007.

to a minimum, even at the cost of lower accuracy for less important classes. She might also desire a minimum model complexity while preserving a reasonable accuracy level, even for the most important classes. Moreover, different users could attach different importance degrees to each class depending on their personal goals even for the same problem.

Therefore, we need classification inducers that take class importance into account when building classification models. In order to represent the relative importance of each class, we can resort to *relative weights* w_i for each class. For the sake of simplicity, we will assume that the weights w_i are floating-point values between 0 and 1. These values can always be normalized.

In this work, we focus on getting classification models as simple as possible for the important classes without penalizing classification accuracy. If we were only interested in classification accuracy, existing models could have been used. In particular, we could have used a cost-sensitive model [4] by defining a cost matrix, which would have reflected the relative importance of each class. However, classifier complexity is also a fundamental issue in supervised learning, since complexity is closely related to interpretability. A classifier might be useless from a practical point of view if it reaches a good accuracy level but is too complex to be understood by the decision makers who need a rationale behind their decisions.

In particular, *weighted classification models* (that is, classification models built by taking class weights into account) can be useful in situations such as the ones described by the following examples from the UCI Machine Learning repository [10]:

- **Extreme classes problems:** In some problems, experts could be specially interested in properly classifying 'extreme' classes, i.e. classes whose importance is paramount in the decision making process. For example, when dealing with the CAR data set, we could be interested in getting clear rules for good cars in order to recommend them (and for bad ones in order to avoid them).

  It should be noted that cost-sensitive classification could have also been used in this problem, since the expert might be interested in not mistaking a bad car for a good one, regardless of the required classifier complexity.
- **Two-class problems:** In binary classification problems, it is relatively common for the proper description of one class to be much more important than the other's for providing the rationale behind a given decision. For example, in the ADULT data set, where the class attribute is personal income, with values >50K and <=50K, a tax inspector might be more interested in people who earn more money in order to perform a financial investigation.
- **Classes and ontologies:** When the classes in a classification problem can be organized somehow, we can also resort to importance degrees in order to focus on related classes that might be specially relevant for the user. For example, a use hierarchy can be defined for the 6-class GLASS data set: three kinds of glass are used to make windows (one for vehicle windows, two for building windows), while the other three have other applications (containers,

tableware, and headlamps). If we were interested in identifying glass from a broken window, we could assign high importance degrees to all the kinds of glass used to make windows.

After the definition of weighted classification and the study of potential application areas, we face the problem of evaluating weighted classification models. In this paper, we propose two metrics that take into account class importance. They might be helpful when evaluating the accuracy and complexity of weighted classifiers:

- Classifier accuracy is the main goal of any classification system. In weighted classification problems, we recommend the use of the following **weighted accuracy** measure:

$$wAcc = \sum_{i=1}^{\#classes} w_i \cdot acc(i) \tag{1}$$

  where *acc(i)* is the average accuracy for the *i*-th class, w_i is the weight for the *i*-th class, and *#classes* is the number of classes.
- We also propose an analogous **weighted complexity** measure for evaluating classifier complexity, which is closely related to its understandability and interpretability:

$$wOpacity = \sum_{i=1}^{\#classes} w_i \cdot opacity(i) \tag{2}$$

  where *opacity(i)* is the value of the complexity measure for the *i*-th class. For instance, *opacity(i)* might represent the average depth for nodes belonging to the *i*-th class in a decision tree. In this case, the complexity measure for *i*-th class can be defined as follows:

$$opacity(i) = depth(i) = \frac{\sum_{x \in class(i)} level(x)}{freq(i)} \tag{3}$$

  where $level(x)$ is the depth of the leave corresponding to the example x, $freq(i)$ is the number of examples belonging to *i*-th class, and $class(i)$ is the set of examples belonging to the *i*-th class.

A weighted classifier should be evaluated according to these measures. An optimal classifier would optimize all of them at the same time, although this multi-objective optimization is not always possible, so we will usually have to achieve a trade-off between accuracy and complexity.

3 Adapting ART for Weighted Classification

In this paper, we show how standard classification models can be adapted for dealing with class weights. In particular, we focus on the ART classification model [1]. ART, which stands for Association Rule Tree, is a Separate and Conquer algorithm that is suitable for Data Mining applications because it makes use of efficient association rule mining techniques.

The special kind of decision list ART obtains can be considered as a degenerate, polythetic decision tree. The ART algorithm outline is shown in Figure 2. Unlike traditional TDIDT algorithms, ART branches the decision tree by simultaneously using several attributes.

Internally, ART makes use of association rules in order to find good descriptions of class values. When evaluating candidate rules, the classical confidence measure used in association rule mining is employed to rank the discovered rules, even though alternative criteria might be used [2].

Once ART discovers potentially useful classification rules, they are grouped according to the attributes in their antecedents, as shown in Figure 1. A rule selection mechanism is also necessary for choosing one of the resulting rule groups. The chosen group is used to branch the decision tree and the whole process is repeated for the remaining examples.

$$\begin{gathered}\{A_3\&B_2 \rightarrow C_1,\ B_4\&C_2 \rightarrow C_2,\ A_4\&B_1 \rightarrow C_2,\ B_0\&C_1 \rightarrow C_1,\ A_2\&C_2 \rightarrow C_1\} \\ \Downarrow \\ \{A_3\&B_2 \rightarrow C_1,\ A_4\&B_1 \rightarrow C_2\} \\ \{B_4\&C_2 \rightarrow C_2,\ B_0\&C_1 \rightarrow C_1\} \\ \{A_2\&C_2 \rightarrow C_1\}\end{gathered}$$

Fig. 1. Grouping rules with compatible antecedents

In the following sections, we propose some modifications to the rule evaluation and rule selection criteria used by the ART algorithm in order to deal with weighted classification problems.

3.1 Weighted Selection Criterion

As we have mentioned before, a selection criterion is needed when alternative sets of rules are considered good enough to branch the tree. In ART, the criterion is based on the support of the rules belonging to the group. The best set of rules is the set that covers the maximum number of examples. However, this approach does not take the weights of the rules into account. We propose a modified criterion, which we call weighted coverage:

$$weightedCoverage(RuleSet) = \sum_{r \in RuleSet} support(r) \cdot w(r)$$

where $w(r)$ is the weight of the class in the consequent of the rule r and $support(r)$ is the number of examples supporting the rule r. In some sense, this is similar to the idea used in boosting algorithms such as AdaBoost [3].

By using weighted coverage, the set of rules that cover a larger number of more important classes is preferred over other sets. Since such a set will be selected as soon as possible, its level in the tree will tend to be lower and, therefore, the classifier opacity is expected to be reduced.