Proceedings of the Sixth IASTED International Conference
WEB-BASED EDUCATION
March 14–16, 2007, Chamonix, France
ISBN Hardcopy: 978-0-88986-650-8 / CD: 978-0-88986-651-5

EXTRACTION OF LOGICAL RULES TO DESCRIBE STUDENTS' LEARNING BEHAVIOR

Félix Castro & Àngela Nebot
Departament de LSI
Universitat Politècnica de Catalunya
Campus Nord
Jordi Girona Salgado 1-3
Barcelona 08034, Spain
fcastro@lsi.upc.edu, angela@lsi.upc.edu

Francisco Mugica
Instituto Latinoamericano de la
Comunicación Educativa (ILCE)
Calle del Puente 45
México D.F., México
fmugica@ilce.edu.mx

ABSTRACT. E-learning offers a new context for education where large amounts of information describing the continuum of the teaching-learning interactions are endlessly generated and ubiquitously available. But raw information by itself may be of no help to any of the e-learning actors. The use of Data Mining methods to extract knowledge from this information can, therefore, be an adequate approach to follow in order to use the obtained knowledge to fit the educational proposal to the students' needs and requirements. In this brief study we use an extension of Fuzzy Inductive Reasoning methodology to extract comprehensible, actionable and reasonable set of rules describing the students' learning behavior. The obtained rules can be used to improve the system understanding and to provide valuable information to tutors about the course performance. The extraction rules model presented in this research is applied to a real virtual campus graduate course of the Center of Studies in Communication and Educational Technologies (CECTE, as Spanish acronym).

KEY WORDS: Virtual campus; e-learning; Fuzzy inductive reasoning; Rules extraction.

1 Introduction

Any e-learning system is, by its own nature, likely to generate large amounts of information describing the continuum of the teaching-learning interactions almost in real time. All this information, gathered from diverse and usually heterogeneous sources, may be of no help by itself to any of the e-learning actors in its raw form. Actually, an excess of such information can become a liability for e-learning tutors and managers unless it is processed according to reasonable goals. Data Mining can provide the adequate tools for such processing, obtaining actionable patterns from large data repositories. The use of Data Mining methods to extract knowledge from the e-learning system available information can, therefore, be an adequate approach to follow in order to use the obtained knowledge to fit the educational proposal to the students' needs and requirements.

Virtual campus environments, such as the one that is subject of this case study, are fastly becoming a mainstream alternative to traditional distance higher education. The Internet medium they use to convey content, also allows the gathering of information on students' online behavior. Here, we focus on e-learning systems improvement through the analysis of the data generated by the virtual campus students, aiming to discover their system usage patterns. Most of this research uses Soft Computing techniques to analyze the available data. In turn, we can distinguish diverse Soft Computing-based approaches to e-learning process analysis, i.e. methods to classify students' based on their usage patterns on a web-based course [1, 2, 3], methods oriented towards system personalization [4]. For instance, a neural network model is proposed in [5] to recommend an adequate navigation strategy for the user. A methodology to improve the performance of developed courses through adaptation, using Evolutionary algorithms, is presented in [6]. And, finally, methods that allow automatic detection of atypical students' behavior such as the Bayesian predictive distribution model to detect irregular learning proposed in [7], and the Generative Topographic Mapping model to detect atypical behavior on the grouping structure of the users of a real virtual campus, presented in [8].

Research concerned with the analysis of data generated by the use of e-learning systems is still scarce, and there is a lack of standard methods and guidelines to address some open problems in distance education.

Extraction rules methods applied to e-learning have been investigated in the areas of learning recommendation systems [9, 10], learning material organization [11] and student learning assessments [12]. Association rule mining, inter-session and intra-session frequent pattern mining, were applied in [9, 10] to extract useful patterns that might help educators, educational managers and Web masters to evaluate and interpret on-line course activities. A similar approach can be found in [13], where contrast rules, defined as sets of conjunctive rules describing patterns of performance disparity between groups of students, were used.

Artificial neural networks for the prediction of the students' final marks were applied in [12] and the predictions made by the networks were interpreted using orthogonal search-based rule extraction (OSRE), in order to discover interesting behavioral rules in student usage information. In [14] an a priori algorithm for association rules was applied to capture relationships among URL references based on the navigational patterns of students.

For a deeper and more detailed insight into the data mining methods applied to e-learning systems, the reader is referred to [15].

In this paper we present an extension of the Fuzzy Inductive Reasoning (FIR) methodology that allows the extraction of logical rules from a previous identified FIR model. The main goal of the FIR model is to forecast the final mark of the students, and to determine the most relevant features involved in this process. In the application at hand, the generated set of logical rules describes students' learning behavior patterns enrolled in a CECTE virtual course.

The remaining of the paper is organized as follows: section 2 presents the basics of the FIR methodology and the rules extraction method. A description of the data collected from the analyzed web-based course is provided in section 3. Results from the experiments are presented and discussed in section 4. Finally, section 5 summarizes the results obtained in this work with some conclusions.

2 The Fuzzy Inductive Reasoning methodology

The conceptualization of the Fuzzy Inductive Reasoning (FIR) methodology arises from the General Systems Theory proposed by Klir [16]. This modeling and qualitative simulation methodology is based on systems behavior rather that on structural knowledge. It is able to obtain good qualitative relations between the variables that compose the system and to infer the future behavior of that system. It also has the ability to describe systems that cannot easily be described by classical mathematics (e.g. differential equations), i.e. systems for which the underlying physical laws are not well understood. FIR consists of four main processes, namely: fuzzification, qualitative model identification, fuzzy forecast and defuzzification. Fig. 1 describes the structure of the FIR methodology.

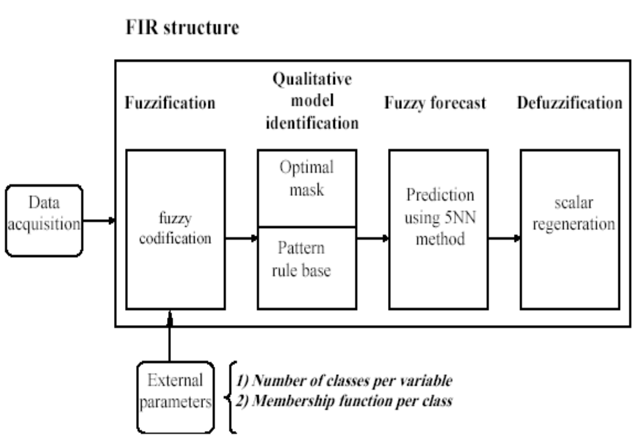

Fig. 1. Fuzzy Inductive Reasoning.

The fuzzification (discretization) process converts quantitative data stemming from the system into fuzzy data. The qualitative model identification process is responsible for finding causal and temporal relations between variables and therefore for obtaining the best model that represents the system. A FIR model is composed of a mask (model structure) and a pattern rule base (behavior matrix). Once the FIR model is available, the prediction system can take place using the FIR inference engine. This process is called fuzzy forecast. The FIR inference engine is a specialization of the k-nearest neighbor rule, commonly used in the pattern recognition field. Defuzzification is the inverse process of fuzzification. It allows converting the qualitative predicted output into quantitative values that can then be used as inputs to an external quantitative model.

In order to provide a better understanding of the rule extraction method proposed in this paper, we present a brief explanation of the pattern rule base determination process. The logical rule base is generated from the pattern rule base obtained by FIR methodology.

As previously stated, a FIR model is composed of a mask and a pattern rule base, an example of mask (that is used latter in Fig. 2) is:

t \ x	u_1	u_2	u_3	u_4	y_1
t - 2δt	-1	0	0	-2	0
t - δt	0	0	0	0	-3
t	0	-4	0	0	+1

Each negative element in the mask is called a m-input (mask input) and denotes a causal relation with the output. The process of finding the best mask that represents the system under study corresponds to the feature selection process in such a way that the negative elements in the mask are the relevant features selected and give us information of the temporal relation with the output.

The process to obtain the pattern rule base from the mask is illustrated in Fig. 2. The mask can be used to 'flatten' dynamic relationships into pseudo-static relationships. The left-hand side of Fig. 2 shows an excerpt of the matrix that stores the class values of the training data set. It shows the numerical rather than the symbolic class values. In the example shown in Fig. 2, all the variables were discretized into three classes, with the exception of variable *y1* that was discretized into two classes. The dashed box symbolizes the mask that is shifted downwards along the class value matrix. The round shaded 'holes' in the mask denote the positions of the m-inputs, whereas the square shaded 'hole' indicates the position of the m-output. The class values are read out from the class value matrix through the 'holes' of the mask, and are placed next to each other in the behavior matrix that is shown on the right-hand side of Fig. 2 and that contains the set of pattern rules. Each pattern rule has associated three values, the class value (shown in Fig. 2) and the membership and side values (do not shown in Fig. 2). Notice that the number of pattern rules obtained by FIR corresponds almost to the number of training data available. Therefore, the size of the pattern rule base impedes their comprehension and understanding by the educational actors. For a deeper and more detailed insight into the FIR methodology, the reader is referred to [17].

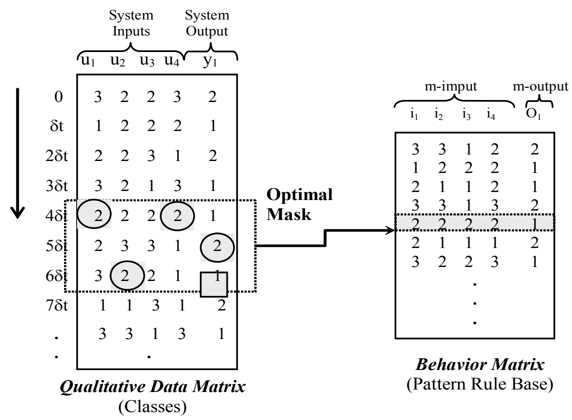

Fig. 2. FIR qualitative model identification process.

2.1 Rules extraction using FIR

Starting from the description of FIR in last section, we now explain how rule extraction can be implemented as part of this methodology. Figure 3 shows in a schematic way the main phases of the proposed algorithm described next.

The proposed method is an iterative process that compact the pattern rule base obtained by FIR. On the one hand, we aim to obtain interpretable, realistic and efficient rules, describing students' learning behavior. On the other hand, we want to compact the pattern rule base to speed up the prediction process. In order to get a set of logical rules congruent with the pattern rules previously identified by FIR, the proposed algorithm is based on its initial discretization. The model can be summarized as a set of ordered steps:

1. *Basic compactation*. This is an iterative step that evaluates, one at a time, all the rules in a pattern rule base. The pattern rule base, R, is compacted on the basis of the "knowledge" obtained by FIR. A subset of rules R_c can be compacted in the form of a single rule r_c, when all premises P but one (P_a), as well as the consequence C share the same values. Premises, in this context, represent the input features, whereas consequence is the output feature in a rule. If the subset contains all legal values LV_a of P_a, all these rules can be replaced by a single rule, r_c , that has a value of -1 in the premise P_a. When more than one -1 value, P_{ni}, is present in a compacted rule r_c, it is compulsory to evaluate the existence of conflicts by expanding all P_{ni} to all their legal values LV_a, and comparing the resultant rules Xr with the original rules R. If conflicts, Cf, exist, the compacted rule r_c is rejected, and otherwise accepted. In the latter case, the previous subset, R_c is replaced by the compacted one r_c. Conflicts occur when one or more extended rules, Xr have the same values in all its premises, P, but different values in the consequence C.

2. *Improved compactation*. Whereas the previous step only structures the available knowledge and represents it in a more compact form, the improved compactation step extends the knowledge base R to cases that have not been previously used to build the model: R_b. Thus, whereas step 1 leads to a compacted data base that only contains knowledge, the enhanced algorithm contains undisputed knowledge and uncontested belief. Two options are studied: In the first one, using the compacted rule base R' obtained in step 1, all input features P (premises) are visited once more in all the rules r that have nonnegative vales (not compacted), and their values are replaced by -1. An expansion to all possible full sets of rules Xr and their comparison with the original rules R are carried out. If no conflicts Cf are found, the compacted rule, r_c, is accepted, and otherwise, rejected. The second option is an extension of the basic compactation, where a consistent and reasonable minimal ratio, MR, of the legal values LV_a should be present in the candidate subset R_c, in order to compact it in the form of a single rule r_c. This latter option seems sensible because, although a reasonable ratio was used to compact R_c in a single rule r_c, the assumed beliefs are minimal and do not compromise the model previously identified by FIR. Instead, in option 1, beliefs are assumed to be consistent with the original rules; nevertheless, this could compromise the agreement with model identified, specially when the training data are poor and do not describe well all possible behaviors.

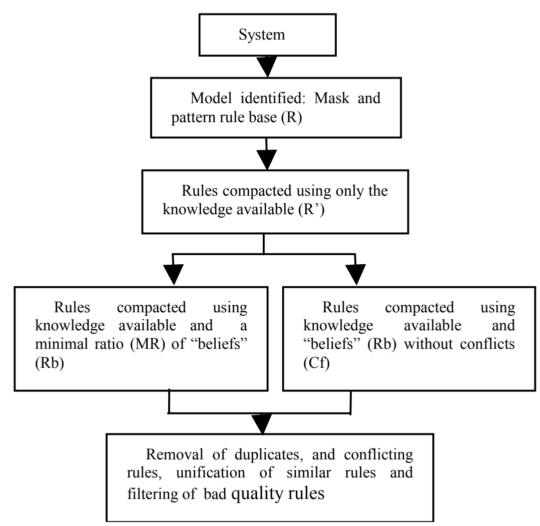

Fig. 3. Ordered steps of the rule extraction method.

The obtained set of rules is subjected to a number of refinement steps: removal of duplicate rules and conflicting rules; unification of similar rules; evaluation of the obtained rules and removal of rules with low specificity. Specificity is a standard metric, described latter, that assess the quality of the obtained rules. For space limitations the refinements steps can not be explained in detail in this paper.

3 Data from the CECTE virtual campus

The Center of Studies in Communication and Educational Technologies (CECTE: Spanish acronym) is a partially virtual campus, offering postgraduate courses and continuous education (graduate, workshops and specific courses) to Latin-American students. The CECTE is part of the international Latin-American Institute of Educative Communication (ILCE: Spanish acronym), whose main goal is to offer postgraduate courses.

Table 1. Data features collected for the experiment.

Feature	Alias	Description
Age of the student	AGE	Age of the student.
Area of expertise	EXP	Area of expertise of the student (mathematics, chemistry, Mexican history, etc.).
Gender	G	Student's gender.
Level of studies	STD	Level of studies (graduate, master, Ph.D., etc.).
Position of the student	POS	Position of the student as a teacher in his/her school.
Percentage of the activities performed by the student	ACT	Percentage of the activities performed by the student with respect to the total activities of the course.
Percentage of session assistance	ASS	Percentage of student's session assistance with respect to the total number of sessions of the course.
Average mark of the e-mail	MAIL	Average mark obtained by the student in the activities sent by e-mail.
Average mark of the co-evaluation	COEV	Average mark of the co-evaluation performed by the student of the class plan of other students.
Average mark of the forum participation	F	Average mark of the student's forum participation (referring to topics related to the course).
Average mark of the forum class plan	FCP	Average mark of the forum class plan (referring only to topics related to the class plan exclusively).
Average mark of the final class plan	FC	Average mark obtained by the student in his/her final class plan.
Average mark of the initial class plan	IC	Average mark obtained by the student in his/her initial class plan.
Average mark of the experience report	ER	Average mark obtained by the student in the experience report. In this report the student evaluates his/her learning process and describes the main concepts learned.
Average mark of the work in the branch	BR	Average mark of the work (activities) performed in the branch.
Final mark	MARK	Final mark obtained by the student in the course.

The most demanded CECTE courses follow a hybrid, semi-presential model, in which students take courses online (WCECTE) but also attend weekly TV sessions through the National System of Educative Television (EDUSAT). Through WCECTE, students can access the course materials and communicate and interact with each other through an e-mail system and a discussion forum. The environment also includes an agenda, a news system, virtual classrooms, a digital library, interactive tutorials, and other related tools.

The tutor is a very relevant actor, as he or she interacts directly with students, assigning learning activities, answering doubts, opening topics in discussion forums, evaluating the activities performed by learners, and verifying that the teaching-learning process be adequate, taking advantage of all the tools provided by WCECTE.

Two novel evaluation topics, not often used in e-learning environments, were incorporated in the course: co-evaluation and experience report. In co-evaluation, the advisor grades how well the student evaluates the class plans of his/her course mates. The experience report is a student description of his/her perception of the course. It can be viewed as a self-evaluation of the student's own learning process.

For the experiments in this study, a set of 722 students, enrolled in the "Didactic Planning" graduate course, was selected. The course is addressed to second term high school teachers offering specialized subjects, namely Mathematics, Chemistry, Mexican History, Computer Science, English, as well as Reading and Writing, and Ethics and Values workshops. The students are meant to perform a set of activities throughout the course with the main purpose of learning new methods and strategies for planning the classes that they teach. This is the reason why these activities are centred on the so-called "class plan". A class plan is a document where a set of strategies are suggested to develop a teaching-learning session, taking into account different factors that appear in the educational process, such as students' characteristics, teaching style, teachers' experience, etc. The data features available for this study are detailed on Table 1.

4 Experiments: Results and discussion

In a previous work [18], the FIR methodology was used in the same application described here. The goal of this research was twofold. On the one hand, the identification of FIR models capable of predicting student's performance. On the other hand, the determination of data features with the highest relevance from the student's performance point of view.

The best mask obtained in this work was the following:

t \x	AGE	EXP	G	STD	POS	ACT	ASS	MAIL	COEV	F	FCP	FC	I C	E R	B R	MARK
T	0	0	0	0	0	0	0	0	-1	0	0	0	-2	-3	0	+1

Meaning that, from all the features involved (see Table 1), three of them i.e. COEV, IC and ER, were the more relevant when the goal was to predict student's learning performance. The results presented in [18] were very successful in the prediction of student's performance (96% of accuracy). However, the pattern rule base obtained had a size of 720 fuzzy pattern rules, a number extremely high to become useful for the easy understanding of student's learning performance.

In this paper, the interpretability of the prediction results is improved by their description in terms of simple and actionable rules. This is accomplished through the application of the new extension of FIR that allows the extraction of logical rules from the already available pattern rule base. The obtained logical rules are more comprehensive, readable and provide explanations (not only assumptions) that may be validated by domain experts, increasing confidence in the analysis.

The experimental results obtained using the rule extraction algorithm described in section 2 are presented in Table 2. The first column of this table presents the logical rules obtained by the extraction process. The second and third columns show the specificity, sensitivity and positive predictive value (PPV) measures for the training and test data sets, respectively. Specificity is defined as one minus the ratio of the number of out-of-class data records that the rule identifies to the total number of out-of-class data. Sensitivity is the ratio of the number of in-class data that the rule identifies to the total number of in-class data. PPV is the ratio of the number of in-class data that the rule identifies to the total number of data the rule identifies.

The number of logical rules extracted from the set of 720 pattern rules is 6. Therefore, a huge reduction has been obtained. A set of 6 rules gives comprehensible explanation to the educative actors making easiest the understanding of students learning behavior.

From table 2 it can be seen that the specificity and the PPV measures have reasonable values for each specific rule as well as for the whole set of rules (last row of Table 2). However, the sensitivity has very low values in 4 of the 6 logical rules extracted. Only the sensitivity of rules 2 and 6 (in bold in Table 2) have reasonable values, showing a very common pattern in the analyzed data.

A high sensitivity value implies a very general rule, i.e. a high number of students fit in that rule. A small sensitivity value denotes a very specific rule, i.e. the rule describes a small set of students. However, it is possible that this set of students is not represented with any other rule. Therefore, we consider important to keep the rules with low sensitivity but high specificity in such a way that all the students (each one with its own characteristics) are represented in the rule base. In that way we are assuring better predictions of the students' performance. A rule with a high specificity value indicates that when the premises (the input features) are fulfilled the output is also fulfilled. We think that this is the most valuable measure when the goal is to obtain good predictions of the output value, i.e. students mark in our case.

In Table 2 we found rules with very low sensitivity but high specificity values. This is for example the case of rule 4 with a sensitivity of 0.07 and a specificity of 1. If we analyze this rule carefully we find out that it represents 4 students of the 61 that did not pass the course, i.e. that have a mark value between 0 and 4.9. These 4 students are the only ones that obtained very low evaluations on the three features involved, i.e. COEV, ER and IC. Contrarily, the rule 6 has relatively high values of specificity and sensitivity measures. In that case, we obtain a sensitivity of 0.81 because this rule represents 314 students of 384 that have the mark in the same range (high evaluation). It is interesting to notice that the distribution of the data available in this study is unbalanced, i.e. 61 students did not pass the course, 104 passed the course with a low grade and 384 students obtained high grades. Based on the results obtained and on our experience we think that in those applications where the available data are more balanced than in the current application the extraction algorithm will obtain rules with higher sensitivity values.

The learning behavior rules obtained were analyzed and validated by educative experts of CECTE. They agreed that the obtained results were intuitive, realistic, and mostly consistent with their own perception of the CECTE course students' learning behavior. When the students obtain high evaluations in two relevant variables determined by FIR, i.e. COEV and ER, they obtain a high degree in the course (rule 6 in Table 2). When the COEV grade is not very high the student can still obtain a good final mark if he/she has a high evaluation of the initial class plan (rule 5 in Table 2). Few students obtain very bad degrees, only 61, but there are different circumstances of students' bad results. This is the reason way three rules are necessary to cover these students behaviour (rules 2, 3 and 4 in Table 2). Only one rule (rule 1 in Table 2) is obtained to define those students that passed the course with low degrees.

Table 2. Results obtained using the new rule extraction method for both training and test data sets. Spec stands for Specificity, Sens for Sensitivity and PPV for Positive Predictive Value.

RULE	TRAIN			TEST		
	Spec	Sens	PPV	Spec	Sens	PPV
IF 0<=IC<=5.1 AND 4.9<=COEV<=10 THEN 0<=MARK<=7.9	0.98	0.29	0.86	0.98	0.36	0.90
IF 5.1<=IC<=10 AND 0<=ER<=8.1 THEN 0<=MARK<=4.9	**0.79**	**0.67**	**0.28**	**0.78**	**0.54**	**0.17**
IF 0<=COEV<=4.90 AND 8.1<=ER<=10 THEN 0<=MARK<= 4.9	1.00	0.11	0.78	0.97	0.08	0.20
IF 0<=IC<=5.1 AND 0<=COEV<=4.90 AND 0<=ER<=8.1 THEN 0<=MARK<= 4.9	1.00	0.07	1.00	1.00	0.08	1.00
IF 5.1<=IC<=10 AND 4.90<=COEV<=7.9 THEN 7.9<=MARK<=10	0.95	0.03	0.58	0.87	0.04	0.42
IF 7.90<=COEV<=10 AND 8.1<=ER<=10 THEN 7.90<=MARK<=10	**0.75**	**0.81**	**0.88**	**0.83**	**0.81**	**0.91**
TOTAL RULES	**0.93**	**0.38**	**0.72**	**0.92**	**0.39**	**0.69**

5 Conclusion

The possibility of tracking user behavior in virtual campus e-learning environments makes the web mining of the resulting databases possible. This opens new possibilities for the pedagogical and instructional designers who create and organize the learning contents. One of the most interesting options is the personalization of the e-learning process. The characterization of the students' online behavior would benefit from a method capable of determining the relevance of the features involved in the analyzed data set in terms of the students' mark prediction. In this study, we have presented a new rule extraction method based on FIR methodology. This method has been used on a data set obtained from a virtual campus real e-learning experience.

The logical rules obtained are easily understandable by experts in an educative domain and they may expose problems with the data itself. This knowledge could be used for real time student personalization guidance, and to help teachers in finding patterns of students' behavior. For this knowledge to have an intuitive and useful form, results have been described in terms of a set of logical rules describing the diverse levels of the students' performance.

The experimental results have shown that the extraction rules method presented in this paper was able to obtain comprehensible, actionable and realistic logical rules describing students' learning behavior patterns.

Acknowledgment

Félix Castro is a research fellow within the PROMEP program of the Mexican Secretary of Public Education.

References

[1] B. Minaei-Bidgoli and W. F. Punch, Using Genetic Algorithms for Data Mining Optimization in an Educational Web-based System: *Genetic and Evolutionary Computation Conference*, 2003.

[2] D. Mullier, A Tutorial Supervisor for Automatic Assessment in Educational Systems: *International Journal on E-Learning*, 2003.

[3] G. D. Stathacopoulou and M. Grigoriadou, Neural Network-Based Fuzzy Modeling of the Student in Intelligent Tutoring Systems: *International Joint Conference on Neural Networks*, Washington, July 1999.

[4] P. Brusilovsky, Adaptive hypermedia: *User Modeling and User Adapted Interaction*, Ten Year Anniversary Issue, 11 (1/2), 2001, 87-110.

[5] K. Mizue and O. Toshio, N3: Neural Network Navigation Support–Knowledge-Navigation in Hyperspace: *The Sub-symbolic Approach: Journal of Educational Multimedia and Hypermedia* 10(1), 2001, 85-103.

[6] C. Romero, S. Ventura and P. De Bra, Knowledge Discovery with Genetic Programming for Providing Feedback to Courseware: *User Modeling and User-Adapted Interaction*, 14(5), 2004, 425-464.

[7] M. Ueno, LMS with irregular learning processes detection system: *World Conference on E-Learning in Corp., Govt., Health, & Higher Ed.*, 2003(1), 2486-2493.

[8] F. Castro, A. Vellido, A. Nebot and J. Minguillón, Detecting Atypical Student Behaviour on an e-learning System: *VI Congreso Nacional de Informática Educativa, Simposio Nacional de Tecnologías de la Información y las Comunicaciones en la Educación (SINTICE'2005)*. Granada, España, September 14-16, 2005.

[9] O. R. Zaïane, Building a Recommender Agent for e-Learning Systems: *The International Conference on Computers in Education (ICCE'02)*, 2002, 55-59.

[10] O. R. Zaïane and J. Luo, Towards Evaluating Learners' Behavior in a Web-based Distance Learning Environment: *IEEE International Conference on Advanced Learning Technologies (ICALT'01)*. August 6-8, Madison, WI, 2001, 357-360.

[11] C. J. Tsai, S. S. Tseng and C. Y. Lin, A Two-Phase Fuzzy Mining and Learning Algorithm for Adaptive Learning Environment: *International Conference on Computational Science (ICCS 2001)*. LNCS Vol. 2074. Springer-Verlag, Berlin Heidelberg New York, 2001, 429-438.

[12] T. A. Etchells, A. Nebot, A. Vellido, P. J. G. Lisboa and F. Mugica, Learning What is Important: Feature Selection and Rule Extraction in a Virtual Course. *The 14th European Symposium on Artificial Neural Networks (ESANN 2006)*. Bruges, Belgium, 2006, 401-406.

[13] B.Minaei-Bidgoli, P. N. Tan and W. F. Punch, Mining Interesting Contrast Rules for a Web-based Educational System: *The 2004 International Conference on Machine Learning and Applications (ICMLA'04)*. Louisville, KY, 2004.

[14] P. Markellou, I. Mousourouli, S. Spiros and A. Tsakalidis, Using Semantic Web Mining Technologies for Personalized e-Learning Experiences: *The Fourth IASTED Conference on Web-based Education (WBE 2005)*. Grindelwald, Switzerland, 2005.

[15] F. Castro, A. Vellido, A. Nebot and F. Mugica, Applying Data Mining Techniques to e-Learning Problems: a Survey and State of the Art: *Evolution of Technology and Pedagogy*. Springer-Verlag, Germany, accepted for publication.

[16] G. Klir, *Architecture of systems problem solving,* New York: Plenum Press, 1985.

[17] A. Nebot, Qualitative Modeling and Simulation of Biomedical Systems Using Fuzzy Inductive Reasoning: Ph.D. thesis, Dept. Llenguatges i Sistemes Informàtics, Universitat Politècnica de Catalunya, Barcelona, Spain, 1994.

[18] A. Nebot, F. Castro, A. Vellido, and F. Mugica, Identification of Fuzzy Models to Predict Students Performance in an e-Learning Environment: *The Fifth IASTED International Conference on Web-Based Education (WBE 2006)*. Puerto Vallarta, Mexico, 2006, 74-79.