Subgroup discovery in an e-learning usage study based on Moodle

C. J. Carmona, P. González, M. J. del Jesus
Department of Computer Science
University of Jaen. Jaen, Spain
{ccarmona|pglez|mjjesus}@ujaen.es

S. Ventura
Department of Computer Science and Numerical Analysis
University of Cordoba. Cordoba, Spain
sventura@uco.es

Abstract—**This paper presents an experimental study with several subgroup discovery algorithms using data from a web-based education system. The main objective of this contribution is to extract unusual subgroups to describe possible relationships between the use of the e-learning platform and marks obtained by the students. The results obtained by the best performing algorithm, NMEEF-SD, are also presented. Finally, the most representative results obtained by this algorithm are analised, in order to obtain knowledge that can allow teachers to take actions to improve student performance.**

Keywords-**Subgroup Discovery; Evolutionary Fuzzy Systems; NMEEF-SD; e-learning.**

I. INTRODUCTION

On-site education systems allow teachers to directly analyse the situation of the courses, using the interation with students to improve course content and thus achieve better grades for students. However, in learning management systems (LMS) this information can not be directly analised with respect to the use of the resources, activities, quizzes and so on, and teachers must rely on certain techniques like educational data mining.

Educational data mining is an emerging interdisciplinary research area that deals with the development of methods to explore data from educational contexts [1]. It is concerned with the development of mining methods to explore the unique types of data in educational environments and, using these methods, to better understand students and learning systems. A data mining algorithm can discover knowledge using different representation models and techniques from two different perspectives: predictive induction, whose objective is the discovery of knowledge for classification or prediction [2]; or descriptive induction, whose main objective is the extraction of interesting knowledge from data. In this area, attention can be drawn to the discovery of association rules following an unsupervised learning model [3], and other approaches to non-classificatory induction.

In this contribution, the Subgroup Discovery (SD) task [4] is employed. SD is a descriptive inductive learning area in which, given a set of data and a property of interest to the user, an attempt is made to locate subgroups which are statistically "most interesting" for the user. A subgroup is interesting if it has an unusual statistical distribution with respect to the property of interest. The objective is to discover interesting properties of subgroups by obtaining simple rules, which are highly significant and with high support.

An experimental study is performed with several SD algorithms to show the quality of the powerful NMEEF-SD algorithm in the e-learning usage study of the University of Cordoba. This system is based on the Moodle system which is one of the most used web-based e-learning systems. In addition, a study with the most representative subgroups obtained by NMEEF-SD is performed, as this algorithm obtains the best results in the study. This study, joined with an analysis from the point of view of the teacher, is also presented with the aim of improving the e-learning courses.

To do so, the paper is organised as follows: learning management systems are presented in section II. SD task and EFSs used for SD are presented in section III. In section IV the results obtained by different algorithms and the rules obtained by NMEEF-SD algorithm are analysed. Finally, conclusions are outlined.

II. LEARNING MANAGEMENT SYSTEMS

Nowadays, several systems have been developed for online education, most of them using web-based platforms. These web-based educational systems can be classified in different types such as adaptive and intelligent web based educational systems, particular web-based courses and learning management systems. In this paper, we focus on learning management systems (LMSs), also known under other different terms as course management systems or learning content management systems. LMSs are e-learning platforms offering different facilities for information sharing and communication between the participants in the e-learning process. These systems allow the distribution of information to students but also facilitates the task of the educators when producing content material, preparing assignments, managing distance classes, engaging in discussions and enabling collaborative learning with file storage areas, chats, forums or news services. There are both commercial LMSs (eg. Blackboard, Virtual-U, WebCT, or TopClass) and free LMSs (Moodle, Ilias, Claroline, or ATutor) [5].

It is common for LMSs using a relational database to store the students activities and usage information instead of using data log files. However, when the number of students is high it becomes complicated for a tutor to manage the information

stored, even using the reporting tools offered by some of the platforms. In this situation, the great amount of information makes it difficult to extract useful information to improve the learning process. Therefore, some researchers have recently proposed using data mining techniques in order to help the tutor in this task.

Data mining techniques allow to identify patterns in the information related to the use of the platform which can be analysed not only to interpret the students' activity on the platform but also to get more objective feedback for instruction and more knowledge about how the students learn on the LMS [1]. In fact, some data mining techniques have already been used in e-learning problems. Thus, a grammar-based genetic programming with multi-objective optimization techniques was performed in [7] in order to provide a feedback to courseware authors. In [8] a clustering task was used to discover patterns reflecting user behavior. In addition, in [9] the sequencing capabilities of the SCORM standard to include the concept of recommended itinerary, by combining educators expertise with learned experience acquired by system usage analysis was presented.

Two different areas can be established with respect to the data mining techniques: predictive, in which the main objective is to find a model to classify new instances with respect to an interest variable, and descriptive, whose objective is to describe relationships between all the variables. However, there are some data mining techniques that are somehow halfway between prediction and description. This is the case with the subgroup discovery task, aimed at searching unusual relationships between data with respect to an interest variable. Actually, this technique was applied in [10] to search subgroups in an e-learning platform through evolutionary algorithms.

III. SUBGROUP DISCOVERY

The concept of SD was initially defined as [11]:

"In subgroup discovery, we assume we are given a so-called population of individuals (objects, customer, ...) and a property of those individuals we are interested in. The task of subgroup discovery is then to discover the subgroups of the population that are statistically "most interesting", i.e., are as large as possible and have the most unusual statistical (distributional) characteristics with respect to the property of interest."

The main goal in SD is to discover characteristics of the subgroups by constructing rules with high support and significance. As SD focusses its interest on partial relations instead of complete ones, small subgroups with interesting characteristics can be sufficient.

In SD, a rule R can be described as:

$$R : Cond \rightarrow T_v$$

where T_v or (target variable) is the property of interest that appears in the consequent part of the rule, and the antecedent

part of the rule, $Cond$, is a conjunction of features (attribute-value pairs) selected from the features describing the training instances [12].

A model corresponding to a subgroup in a problem with two classes (x and o) can be found in Fig. 1, where the rule defining the subgroup corresponds to class x. The model defined by the rule is very simple (represented in the figure as a circle) and therefore it is very interpretable. However, the model covers a high number of objects of class x, but not all of them, also including objects corresponding to the other class, o. This illustrates one of the main features of the SD task: it is normally preferred to obtain a simple model rather than a completely precise one.
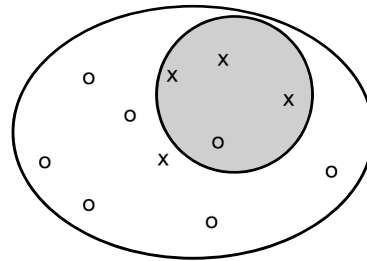

Fig. 1.   Representation of a SD rule for the class x

In the literature there is a wide number of algorithms presented for this task. According to the taxonomy presented in [13], these algorithms can be divided in three types: algorithms based on extensions of classification (such as CN2-SD algorithm [14]), based on extensions of association rules (such as Apriori-SD [15]) and based on evolutionary fuzzy systems (such as SDIGA [16], MESDIF [17] and NMEEF-SD [18]). This work is focused on the use of evolutionary fuzzy systems in a specific e-learning platform.

In the following subsections, the quality measures used in this work together the use of evolutionary fuzzy systems in SD are presented in section III-A and III-B, respectively.

A. Quality measures

In SD a wide number of quality measures defined can be found. These measures can be divided with respect to their properties such as interpretability, precision or generality among others. Due to the fact that a SD algorithm must aim good interpretability, interest and relation between sensibility and confidence, in this work the following quality measures are used:

- *Number of rules (n_r)*: It measures the number of induced rules.
- *Number of variables (n_v)*: It measures the number of variables of the antecedent. The number of variables for a set of rules is computed as the average of the variables for each rule of that set.
- *Significance*: This measure indicates the significance of a

finding, if measured by the likelihood ratio of a rule [4].

$$Sign(R) = 2 \cdot \sum_{k=1}^{n_c} n(T_{v_k} \cdot Cond) \cdot log \frac{n(T_{v_k} \cdot Cond)}{n(T_{v_k}) \cdot p(Cond)}$$
(1)

where n_c is the number of values of the target variable, $p(Cond)$, computed as $n(Cond)/n_s$, is used as a normalised factor, $n(Cond)$ is the number of examples which satisfy the conditions determined by the antecedent part of the rule, $n(T_v \cdot Cond)$ is the number of examples which satisfy the conditions and also belong to the value for the target variable in the rule and $n(T_v)$ are all the examples of the target variable.

It must be noted that although each rule is for a specific T_v, the significance measures the novelty in the distribution impartially, for all the values.

- *Unusualness*: This measure is defined as the weighted relative accuracy of a rule [19]. It can be computed as:

$$Unus(R) = \frac{n(Cond)}{n_s} \left(\frac{n(T_v \cdot Cond)}{n(Cond)} - \frac{n(T_v)}{n_s} \right)$$
(2)

where n_s is the number of total examples.

The unusualness of a rule can be described as the balance between the coverage of the rule $p(Cond_i)$ and its accuracy gain $p(T_v \cdot Cond) - p(T_v)$. This quality measure is derived from *novelty*.

- *Sensitivity*: This measure is the the proportion of actual matches that have been classified correctly [4]. It can be computed as:

$$Sens(R) = TPr = \frac{TP}{Pos} = \frac{n(T_v \cdot Cond)}{n(T_v)}$$
(3)

Sensitivity combines precision and generality related to the target variable.

- *Confidence*: It measures the relative frequency of examples satisfying the complete rule among those satisfying only the antecedent. This can be computed with different expressions, e.g. [20]:

$$Conf(R) = \frac{n(T_v \cdot Cond)}{n(Cond)}$$
(4)

This quality measure can also be found as *accuracy* in the specialised bibliography. In this paper an expression adapted for fuzzy rules is used, Fuzzy Confidence:

$$FCnf(R_i) = \frac{\sum_{E^k \in E / E^k \in T_{v_j}} APC(E^k, R_i)}{\sum_{E^k \in E} APC(E^k, R_i)}$$
(5)

where APC (Antecedent Part Compatibility) is the degree of compatibility between an example and the antecedent part of a fuzzy rule and $\{E^k = (e_1^k, e_2^k, \ldots, e_{n_v}^k)/k = 1, \ldots, n_s\}$ is a set of examples.

A description of all quality measures presented for SD can be analysed in [13].

B. Evolutionary fuzzy systems applied to subgroup discovery

EFSs are essentially fuzzy systems enhanced by a learning process based on an evolutionary algorithm [21], [22]. Currently, EFSs are being applied to a wide range of real-world problems. The research related to this area is growing, and a number of open problems and future directions can be found in [23], [24].

Evolutionary algorithms [25] were not specifically designed for learning, but they offer a set of advantages for knowledge extraction and specifically for rule induction processes:

- They tend to cope well with attribute interaction because they usually evaluate a rule as a whole via a fitness function rather than evaluating the impact of adding/removing one condition to/from a rule.
- They have the ability to scour a search space thoroughly and to handle a fitness function adapted to the problem to be solved.
- In addition, the genetic search performs implicit backtracking in its search of the rule space, thereby allowing it to find complex interactions that other non-backtracking searches would miss.
- An additional advantage over other conventional rule-learning algorithms is that the search is carried out among a set of competing candidate rules or sets of rules.

On the other hand, fuzzy systems are one of the most important areas for the application of the fuzzy set theory [26]. Fuzzy sets correspond to linguistic labels which are defined by means of their corresponding membership functions. These can be specified by the user or defined through uniform partitions. In addition, fuzzy systems handle with continuous variables without the necessity to discretise them before to apply the algorithms. In general, this discretisation could lead to a loss of information in the data.

With respect to the representation, in the specialised literature two approaches are followed in order to encode rules within a population of individuals [22]: "chromosome=rule" and "chromosome=set of rules". In processes aimed to the extraction of rules for the SD task, the "Chromosome = Rule" approach is more suited because the objective is to find a reduced set of rules in which the quality of each rule is evaluated independently from the rest, and it is not necessary to evaluate the set of rules jointly.

The evolutionary fuzzy system for SD developed so far are implemented in the software tool KEEL [28], [29] and summarised below:

- SDIGA [16] is an evolutionary fuzzy rule induction system. It uses as quality measures for the SD task adaptations of the measurements used in the association rules induction algorithms, confidence and support, and can also use other measures such as interest, significance, sensitivity or unusualness. The algorithm evaluates the quality of the rules by means of a weighted average of the measures selected. An analysis of different combinations of quality measures can be observed in [27]. SDIGA uses linguistic rules [17] as description language to specify the

subgroups.

- MESDIF [17] is a multi-objective genetic algorithm for the extraction of fuzzy rules which describe subgroups. The algorithm extracts a variable number of different rules expressing information on a single value of the target variable. The search is based on the multi-objective SPEA2 [30] approach, and so applies the concepts of elitism in the rule selection (using a secondary or elite population) and the search for optimal solutions in the Pareto front. It can use several quality measures at a time to evaluate the rules obtained, like confidence, support, significance or unusualness.

- NMEEF-SD [18] is an evolutionary fuzzy system whose objective is to extract descriptive fuzzy and/or crisp rules for the SD task, depending on the type of variables present in the problem. NMEEF-SD has a multi-objective approach whose search strategy is based on NSGA-II [31], which is based on a non-dominated sorting approach, and on the use of elitism. This algorithm uses specific operators to promote the extraction of simple, interpretable and high quality rules. It allows a number of quality measures to be used both for the selection and the evaluation of rules within the evolutionary process, including confidence, support, sensitivity, significance and unusualness.

IV. E-Learning case study: Data on the use of Moodle at the University of Cordoba

This experimental study is divided into two sections. First, a experimental study of the usage data of Moodle e-learning system at the University of Cordoba with several SD algorithms is performed, where the problem analysed is described and the experimental results obtained in the execution of different SD algorithms are shown. Then, a study of the usage data with NMEEF-SD algorithm is shown, where several of the rules obtained are analysed from the point of view of the teacher with the aim of improving the courses content.

A. Experimental study with the usage data at the University of Cordoba

Moodle system [32] is one of the most used web-based e-learning systems. In addition, Moodle is an alternative to proprietary commercial online learning solutions, is distributed free under open source licensing and has been installed at universities and institutions all over the world.

Moodle system contains a great deal of detailed information on course content, users, usage, etc., stored in a relational data base keeping detailed logs of all the activities performed by the students. We can use these logs in order to determine which students have been active in the course, what they did, when, or if everyone has done a certain task or spent a required amount of time online within certain activities [33].

In this work, available information corresponding to 5 different courses of the University of Cordoba, involving a total of 293 students, is used. In this experimentation, courses with high student participation have been selected to obtain

more general results. Furthermore, there is no a minimum amount of students to obtain any rule. This information has been preprocessed for obtaining a summary table with the most important information related to our objective. Table I shows this summary including the activities completed by each student in an e-learning course. The information obtained has been exported to a text file using the structure of the KEEL platform files [28], [29] because the SD algorithms used are implemented for the KEEL data mining platform. With respect to the mark obtained of the students, they have discretised into different values: fail, pass, good and excellent. In the experimentation could have been used numerical values but it is more representative using these values in order to codify them as the rule consequent.

TABLE I
DESCRIPTION OF THE ATTRIBUTES EMPLOYED FOR EACH STUDENT

Name	Description	Type
course	Identification of the course	Discrete
nAssigment	Number of assignments completed	Continuous
nAssigmentP	Number of assignments passed	Continuous
nAssigmentF	Number of assignments failed	Continuous
nQuizz	Number of quizzes completed	Continuous
nQuizzP	Number of quizzes passed	Continuous
nQuizzF	Number of quizzes failed	Continuous
nMessagesC	Number of messages sent to the chat	Continuous
nMessagesT	Number of messages sent to the teacher	Continuous
nMessagesF	Number of messages sent to the forum	Continuous
nRead	Number of forum messages read	Continuous

In this problem, SD is used in order to analyse the possible relation between the usage of complementary activities of a course and the final marks obtained by the students. This is conducted using different algorithms. The final mark is used as the variable to characterise, using the different marks to divide the data into classes and codifying them as values of the consequent of the rules. Therefore, the final purpose is to present the results to the teacher in form of rules in order to allow the use of this knowledge in the decision making concerning the complementary activities of the course. For example, the teacher can decide to promote the use of some type of activities to improve marks, or on the contrary eliminate some activities because they are associated with low marks.

The study has been performed using the evolutionary algorithms SDIGA, MESDIF and NMEEF-SD. Moreover, to present a complete study the most used classical SD algorithms like CN2-SD and Apriori-SD are employed.

For the classical deterministic algorithms Apriori-SD and CN2-SD a set of runs has been performed varying one of their parameters each time. In the case of Apriori-SD different minimum confidence values (0.6, 0.7, 0.8 and 0.9) have been used and for CN2-SD the weight parameter (0.9, 0.7, 0.5 and additive –add–). However, for the evolutionary algorithms have been used different minimum confidence vales (0.2, 0.4, 0.6 and 0.8) for SDIGA and NMEEF-SD, while MESDIF uses different size for the elite population (3,4 and 5 individuals). In addition, evolutionary algorithms are executed with different

Algorithm	Disc/LL_s	Param	$\mathbf{n}_r$	$\mathbf{n}_v$	SIGN	UNUS	SENS	CONF
Apriori-SD	Fayyad	0.8	5.00	2.51	12.924	0.046	0.601	0.802
CN2-SD	Fayyad	0.5	16.60	3.28	8.736	0.041	0.661	0.420
SDIGA	3	0.6	5.00	3.10	6.376	0.039	0.555	0.421
MESDIF	3	5	20.00	4.29	4.635	0.027	0.499	0.346
NMEEF-SD	3	0.8	15.40	4.38	5.799	0.104	0.716	0.831

TABLE III
RULES MORE REPRESENTATIVE OBTAINED BY NMEEF-SD

$\mathbf{n}_r$	Rule
R_1	IF (Course=29) AND (nMessagesT=0) THEN *Fail*
	Sign:25.700 - Unus:0.110 - Sens:0.631 - Conf:0.856
R_2	IF (nAssigment=Low) AND (nQuizz=Low) THEN *Fail*
	Sign:25.829 - Unus:0.107 - Sens:0.765 - Conf:0.836
R_3	IF (nQuizzP=Low) THEN *Fail*
	Sign:6.913 - Unus:0.075 - Sens:0.955 - Conf:0.703
R_4	IF (nAssigment=Normal) AND (nMessagesT=0) THEN *Pass*
	Sign:1.423 - Unus:0.023 - Sens:0.789 - Conf:0.241
R_5	IF (nAssigment=Normal) THEN *Pass*
	Sign:1.141 - Unus:0.018 - Sens:0.824 - Conf:0.231
R_6	IF (nRead=Low) THEN *Pass*
	Sign:0.973 - Unus:0.001 - Sens:0.965 - Conf:0.194
R_7	IF (nQuizzP=High) THEN *Good*
	Sign:29.912 - Unus:0.079 - Sens:0.772 - Conf:0.655
R_8	IF (Course=110) THEN *Good*
	Sign:25.536 - Unus:0.081 - Sens:0.750 - Conf:0.532
R_9	IF (nAssigment=High) AND (nRead=Low) THEN *Excellent*
	Sign:25.309 - Unus:0.005 - Sens:0.750 - Conf:0.076
R_{10}	IF (nQuizzP=High) THEN *Excellent*
	Sign:29.912 - Unus:0.006 - Sens:0.417 - Conf:0.081

number of linguistic labels for the representation of the continuous variables: 3 and 5 linguistic labels. Finally, evolutionary algorithms are executed five times for each experiment because they are non-deterministic algorithms. All the parameters of the evolutionary process for SDIGA, MESDIF and NMEEF-SD are used by definition.

Table II shows (between all the possible values obtained with the different configuration of parameters) only the best results for each SD algorithm. Therefore, discretise or linguistic labels ($Disc/LL_s$), parameter of the algorithm ($Param$) and the quality measures n_r, n_v, $SIGN$, $UNUS$, $SENS$ and $CONF$ are presented. This table represents fuzzy confidence for the evolutionary algorithms and confidence for classical, i.e. $CONF$ represents crisp and fuzzy confidence to avoid confusion. Results showed in this table are the average of the set of rules obtained for each algorithm so the name of the quality measures are represented with uppercase.

As can be observed in table II, the best performance is obtained by NMEEF-SD algorithm. NMEEF-SD not only obtains the best results with respect to the sensibility-confidence relationship, but also with respect to the unusualness measure. In this sense, NMEEF-SD obtains subgroups with more than 70% of the examples covered and with 83% of examples correctly described. These values show the quality of the subgroups obtained by NMEEF-SD. NMEEF-SD does not obtains the best results with respect to the interpretability (number of rules and number of variables per rule), but the differences are not relevant.

Taking into account the results in table II, next subsection

analyses the rules obtained by NMEEF-SD, with the aim of bringing new knowledge to the teachers in order to enable them to act to improve the results of their students.

B. e-learning usage study with the NMEEF-SD algorithm

NMEEF-SD algorithm returns a comprehensive set of subgroups employing a low number of variables with the highest unusualness. Therefore, in this paper an analysis with the most representative subgroups is performed. In table III the best subgroups for each value of the target variable are presented.

From the results in table III, and for each one of the marks analysed in this experimentation, some key ideas can be highlighted:

- *Fail*. A set of subgroups with high values in all the quality measures analysed can be observed in this group, where generic rules with low number of variables obtain high values of unusualness and confidence. The information to highlight for the students with the mark *fail* would be a low participation of them, and a low interest for the professors to perform quizzes.
- *Pass*. In this group are obtained subgroups which cover the majority of the students but with confidence very low. This group is very difficult to analyse due to the low instances that the dataset contains for this target value.
- *Good*. In this target value good results are obtained with considerable values in all the quality measures analysed. It would be interesting to note the subgroup R_8 where it could indicate to the professor the indifference with respect to the the assignment performed for the student in the platform for the course. In this way, the subgroups indicates him that he should review the relationships between the course and the activities planned. In addition, there is another subgroup (R_7) with excellent results in all the quality measures for courses with high number of quizzes passed.
- *Excellent*. In this target value occurs a similar situation like *pass* value, i.e. is a minority class. However, in this set of subgroups there is a high significance with good values of sensitivity. Furthermore, the same rule (R_{10}) that appears in the target value *good* is obtained, though as can be observed in the table the confidence is lower. A remarkable new information obtained is that students obtain excellent results if the number of forum messages read is low and the number of assignments completed is high. Of course, teachers want their students to use the forum (as it is a valuable tool to improve the students'

skills) but perhaps they need to make an effort to educate their students in its proper use.

V. Conclusions

In this contribution, results obtained by NMEEF-SD algorithm from the usage data of Moodle at the University of Cordoba have been presented. This algorithm obtains the best results with important differences with respect to the remaining algorithms: SDIGA, MESDIF, CN2-SD and Apriori-SD. In addition, the set of most representative subgroups obtained by this algorithm is presented to the experts.

The quality of this algorithm is based on an evolutionary process –a multi-objective evolutionary algorithm–, the use of unusualness and sensibility as objectives of the evolutionary process (that provide an excellent balance in the results obtained for all the quality measures analysed) and the use of fuzzy logic for the continuous variables which avoids to perform a previous discretisation of the data.

Within the set of results obtained, both specific subgroups for concrete courses and generic ones can be found. Based on the subgroups obtained, teachers can make decisions about course activities to improve the performance of their students because the subgroups show the strengths and weakness of the activities and properties of their courses with respect to the students' marks.

References

[1] C. Romero and S. Ventura, "Educational Data Mining: A Review of the State-of-the-Art," *IEEE Tansactions on Systems, Man and Cybernetics, part C: Applications and Reviews*, vol. 40, no. 6, pp. 601–618, 2010.

[2] D. Michie, D. J. Spiegelhalter, and C. C. Tayloy, *Machine Learning*. Ellis Horwood, 1994.

[3] R. Agrawal, T. Imieliski, and A. Swami, "Mining association rules between sets of items in large databases," in *Proceedings of the 1993 ACM SIGMOD International Conference on Management of Data*. ACM Press, 1993, pp. 207–216.

[4] W. Kloesgen, "Explora: A Multipattern and Multistrategy Discovery Assistant," in *Advances in Knowledge Discovery and Data Mining*. American Association for Artificial Intelligence, 1996, pp. 249–271.

[5] M. Paulsen, *Online Education and Learning Management Systems*, Bekkestua, Ed. NKI Forlaget, 2003.

[6] C. Romero and S. Ventura, "Educational data mining: a survey from 1995 to 2005," *Experty Systems with Applications*, vol. 33, no. 1, pp. 135–146, 2007.

[7] C. Romero, S. Ventura, and P. D. Bra, "Knowledge discovery with genetic programming for providing feedback to courseware author," *User Modeling and User-Adapted Interaction: The Journal of Personalization Research*, vol. 14, no. 5, pp. 425–464, 2004.

[8] L. Talavera and E. Gaudioso, "Mining student data to characterize similar behavior groups in unstructured collaboration spaces," in *Proc. 16th European conference on artificial intelligence*, 2004, pp. 17–23.

[9] E. Mor and J. Minguillon, "E-learning personalization based on itineraries and long-term navigational behavior," in *Proc. of the 13th international world wide web conference*, 2004, pp. 264–265.

[10] C. Romero, P. González, S. Ventura, M. J. del Jesus, and F. Herrera, "Evolutionary algorithm for subgroup discovery in e-learning: A practical application using Moodle data," *Expert Systems with Applications*, vol. 36, pp. 1632–1644, 2009.

[11] S. Wrobel, *Inductive logic programming for knowledge discovery in databases*. Springer, 2001, ch. Relational Data Mining, pp. 74–101.

[12] N. Lavrac, B. Cestnik, D. Gamberger, and P. A. Flach, "Decision Support Through Subgroup Discovery: Three Case Studies and the Lessons Learned," *Machine Learning*, vol. 57, no. 1-2, pp. 115–143, 2004.

[13] F. Herrera, C. J. Carmona, P. González, and M. J. del Jesus, "An overview on Subgroup Discovery: Foundations and Applications," *Knowledge and Information Systems*, vol. In press, 2011.

[14] N. Lavrac, B. Kavsek, P. A. Flach, and L. Todorovski, "Subgroup Discovery with CN2-SD," *Journal of Machine Learning Research*, vol. 5, pp. 153–188, 2004.

[15] B. Kavsek and N. Lavrac, "APRIORI-SD: Adapting association rule learning to subgroup discovery," *Applied Artificial Intelligence*, vol. 20, pp. 543–583, 2006.

[16] M. J. del Jesus, P. González, F. Herrera, and M. Mesonero, "Evolutionary Fuzzy Rule Induction Process for Subgroup Discovery: A case study in marketing," *IEEE Transactions on Fuzzy Systems*, vol. 15, no. 4, pp. 578–592, 2007.

[17] M. J. del Jesus, P. González, and F. Herrera, "Multiobjective Genetic Algorithm for Extracting Subgroup Discovery Fuzzy Rules," in *Proceedings of the IEEE Symposium on Computational Intelligence in Multicriteria Decision Making*. IEEE Press, 2007, pp. 50–57.

[18] C. J. Carmona, P. González, M. J. del Jesus, and F. Herrera, "NMEEF-SD: Non-dominated Multi-objective Evolutionary algorithm for Extracting Fuzzy rules in Subgroup Discovery," *IEEE Transactions on Fuzzy Systems*, vol. 18, no. 5, pp. 958–970, 2010.

[19] N. Lavrac, P. A. Flach, and B. Zupan, "Rule Evaluation Measures: A Unifying View," in *Proceedings of the 9th International Workshop on ILP*, ser. LNCS, vol. 1634. Springer, 1999, pp. 174–185.

[20] R. Agrawal, H. Mannila, R. Srikant, H. Toivonen, and A. Verkamo, "Fast discovery of association rules," in *Advances in Knowledge Discovery and data mining*, U. Fayyad, G. Piatetsky-Shapiro, P. Smyth, and R. Uthurusamy, Eds. AAAI Press, 1996, pp. 307–328.

[21] O. Cordón, F. A. C. Gomide, F. Herrera, F. Hoffmann, and L. Magdalena, "Ten years of genetic fuzzy systems. Current framework and new trends," *Fuzzy Sets and Systems*, vol. 14, pp. 5–31, 2004.

[22] F. Herrera, "Genetic fuzzy systems: taxomy, current research trends and prospects," *Evolutionary Intelligence*, vol. 1, pp. 27–46, 2008.

[23] J. Casillas and B. Carse, "Special issue on Genetic Fuzzy Systems: Recent Developments and Future Directions," *Soft Computing*, vol. 13, no. 5, pp. 417–418, 2009.

[24] Y. Nojima, R. Alcalá, H. Ishibuchi, and F. Herrera, "Special Issue on Evolutionary Fuzzy Systems," *Soft-Computing*, vol. In Press, 2011.

[25] J. H. Holland, "Adaptation in natural and artificial systems," *University of Michigan Press*, 1975.

[26] L. A. Zadeh, "The concept of a linguistic variable and its applications to approximate reasoning. Parts I, II, III," *Information Science*, vol. 8-9, pp. 199–249,301–357,43–80, 1975.

[27] C. J. Carmona, P. González, M. J. del Jesus, and F. Herrera, "An Analysis of Evolutionary Algorithms with Different Types of Fuzzy Rules in Subgroup Discovery," in *Proceedings of the IEEE International Conference on Fuzzy Systems*, 2009, pp. 1706–1711.

[28] J. Alcalá-Fdez, L. Sánchez, S. García, M. del Jesus, S. Ventura, J. Garrell, J. Otero, C. Romero, J. Bacardit, V. Rivas, J. Fernández, and F. Herrera, "KEEL: A Software Tool to Assess Evolutionary Algorithms for Data Mining Problems," *Soft Computing*, vol. 13, no. 3, pp. 307–318, 2009.

[29] J. Alcalá-Fdez, A. Fernández, J. Luengo, J. Derrac, S. García, L. Sánchez, and F. Herrera, "KEEL Data-Mining Software Tool: Data Set Repository, Integration of Algorithms and Experimental Analysis Framework," *Journal of Multiple-Valued Logic and Soft Computing*, vol. 17, no. 2-3, pp. 255–287, 2011.

[30] E. Zitzler, M. Laumanns, and L. Thiele, "SPEA2: Improving the strength pareto evolutionary algorithm for multiobjective optimization," in *International Congress on Evolutionary Methods for Design Optimization and Control with Apps to Industrial Problems*, 2002, pp. 95–100.

[31] K. Deb, A. Pratap, S. Agrawal, and T. Meyarivan, "A fast and elitist multiobjective genetic algorithm: NSGA-II," *IEEE Transactions Evolutionary Computation*, vol. 6, no. 2, pp. 182–197, 2002.

[32] M. Flate, "Online education and learning management systems. Global e-learning in a Scandinavian perspective, NKI Gorlaget, Oslo," 2003.

[33] C. Romero, S. Ventura, and E. Salcines, "Data mining in course management systems: Moodle case study and tutorial," *Computer & Education*, vol. 51, no. 1, pp. 368–384, 2008.