INT J COMPUT COMMUN, ISSN 1841-9836
7(5):933-944, December, 2012.

Mining Association Rules from Empirical Data in the Domain of Education

D. Radosav, E. Brtka, V. Brtka

Dragica Radosav, Eleonora Brtka,
Vladimir Brtka
University of Novi Sad
Technical Faculty "Mihajlo Pupin"
Serbia, 23000 Zrenjanin, Djure Djakovica bb
E-mail: radosav@tfzr.uns.ac.rs,
brtka@sbb.rs, vbrtka@tfzr.uns.ac.rs

Abstract:
The data mining techniques and their applications are widely recognized as powerful tools in various domains. In the domain of education there is a variety of data of various types that are collected. The important question is: Is it possible to process collected data with the data mining technique and what are main advantages of data mining and e-learning interaction? If an e-learning system accumulates a huge volume of data, then it is possible to deploy techniques and tools from the domain of data mining in order to gain valuable information. The research presented in this paper is conducted on real-life data that originates from the Balkan region. The software system Weka is used to generate association rules. The main result of this research is the assessment of the parameters that are associated with the opinion that computer skills will be helpful in the future, from the students point of view. This result is very important because it gives the exact insight to computer technology usage in the Balkans schools. Furthermore, some advantages of the usage of data mining techniques in the domain of education are determined.
Keywords: association rules, education, data mining.

1 Introduction

In the past few years e-learning techniques have significantly improved as the result of progress and increased use of the Internet. The "desktop" e-learning systems, in many cases, have been replaced by systems that operate using the Internet. Some of the web-based systems allow the determination of preferences for each participant in the process and adjustment of activities in accordance with the profile of participants [1]. Recently, techniques from the domain of data mining have been incorporated into systems for e-learning. The changes that are constantly taking place in terms of rapid technical and technological developments affect society as a whole. The educational system is experiencing changes in terms of modernization and globalization. However, the educational system that is "inert" is not suitable for rapid change and modernization. The educational processes in Serbia and some other countries in the region of western Balkans are changing so that the "reproduction" style of learning is replaced with the style that prefers "understanding" of the learning content and usage of the acquired knowledge. The theories of learning that are used are no longer associative and behavioral, but have become constructive and cognitive. Students are required to improve the style of self-training and their skills. In these processes the information capacity is an important factor in development, especially the Internet and the resources that are available through World Wide Web. The usage of the Internet by some Course Management System (CMS) is not unusual occurrence in Serbia, but cannot be said that such systems are widely present.

Efforts that have been invested in the integration of CMS and Data Mining (DM) systems are evident. This integration often means adding DM modules to the existing CMS [1, 2], but

it is possible to approach to CMS and DM system integration through serial connection [3, 4]. Serial connection, in this case, means collecting data with CMS, and then processing collected data by DM system. The results of DM analysis are fed back to CMS in order to improve their effectiveness.

However, this paper lists the basic DM techniques and some tools that allow the application of these techniques in domain of education, but CMS and their application is not the topic of this paper, although the importance of CMS is evident. This paper deals with some special DM techniques when applied to data collected in the domain of education. The application of DM techniques results in some rules or patterns that can be used as feedback to CMS. Rather than investigation of the connection between CMS and DM system, this paper deals with DM techniques when applied to data in the domain of education and gives some conclusions and remarks about the importance of inferred knowledge. Special contribution is the analysis of the results of DM technique when applied to real-life data collected from the region of Serbia and Bosnia and Herzegovina.

The paper is organized as follows: section two gives the short description of various DM techniques used in the domain of education. One of DM techniques is chosen to be used for the analysis of the real-life data. It is explained why this DM technique is the most suitable in this particular case. Section three contains data description, as well as the methodology description. Section four lists the results obtained by application of the DM technique, as well as the interpretation of these results. Finally, section five contains conclusions and remarks about applicability of DM techniques in the domain of education.

2 Previous Work and DM Techniques

The well-known CMS that are in use are: Blackboard, WebCT, ANGEL and Moodle. In previous papers DM techniques are used as a functional element (module) of CMS [2, 5]. The DM module can be an integral part of CMS but, in some cases, can be used separately. The field of data mining, like statistics, concerns itself with "learning from data" or "turning data into information" [6]. According to [5,7], DM can be defined as the intersection of the domain of statistics, computer science, artificial intelligence, machine learning, database management and data visualization. Data mining is the process of identifying valid, novel, potentially useful, and ultimately comprehensible and understandable patterns or models.

The basic techniques of data mining are [7]:

- *Classification* - examining the feature of a newly presented object and assigning it to a predefined set of classes.
- *Affinity grouping or association rules* - determining which things go together, also known as dependency modeling.
- *Clustering* - segmenting a population into a number of subgroups or clusters. Description and visualization - exploratory or visual data mining.

2.1 The Application of the DM

In practice, there are a lot of general and specific data mining tools [2]. The commercial mining tools are: DBMiner [8], SPSS Clementine: [9] and DB2 Intelligent Miner [10], etc. Some public domain mining tools are: Weka [11] and Keel [12]. There are also specific educational data mining tools such as: Mining tool [13] for association and pattern mining, MultiStar [14]

for association and classification, KAON [15] for clustering and text mining and CIECoF [16] for association rule mining. The application of these systems in the field of education includes the selection of suitable DM techniques. It is not unusual that multiple DM techniques are applied to the same data sample.

As in [2, 7] the main steps of application of DM techniques are:

1. Collection of the data. The CMS system is used and the collected data are stored in database. This step can be executed by a questionnaire or some other data collection technique instead of CMS usage.

2. Preprocessing the data. The data is "cleaned" and transformed into an appropriate format to be mined.

3. Application of suitable DM technique. The DM technique is applied to build the model that discovers new rules, patterns and knowledge. To execute this step, either a general or a specific data mining tool, or a commercial or a free data mining tool can be used.

4. The interpretation, evaluation and deployment of the results.

In particular, it is necessary to apply and elaborate in detail each of these steps depending on the data to be analyzed. Some of the systems for data mining that are used to analyze data from different domains are:

- **Rosetta** system, developed by researchers from the University of Warsaw and the University of Trondheim [17, 18]. Rosetta is capable of synthesis of the IF ... THEN rules by usage of the Rough sets theory. This system is based on classification, reduction of data and decision rules synthesis.

- **Weka** system was developed by researchers from the University of Waikato, New Zealand [19].

Rosetta system allows data to be loaded from MS Excel table; the format of the loaded data can also be CSV (Comma Separated Values). Rosetta system performs the extraction of the IF...THEN rules. The data can be collected by various methods; the format of the collected data does not have to be specially adapted to DM techniques implemented in Rosetta.

On the other hand, the Weka GUI Chooser provides a starting point for launching Wekas main GUI applications and supporting tools. The Weka system can be used to start the particular DM applications:

- Explorer - An environment for exploring data with Weka.

- Experimenter - An environment for performing experiments and conducting statistical tests between learning schemes.

- KnowledgeFlow - This application supports essentially the same functions as the Explorer but with a drag-and-drop interface. One advantage is that it supports incremental learning.

- SimpleCLI - Provides a simple command-line interface that allows direct execution of Weka commands for operating systems that do not provide their own command line interface.

System Weka allows (see Figure 1):

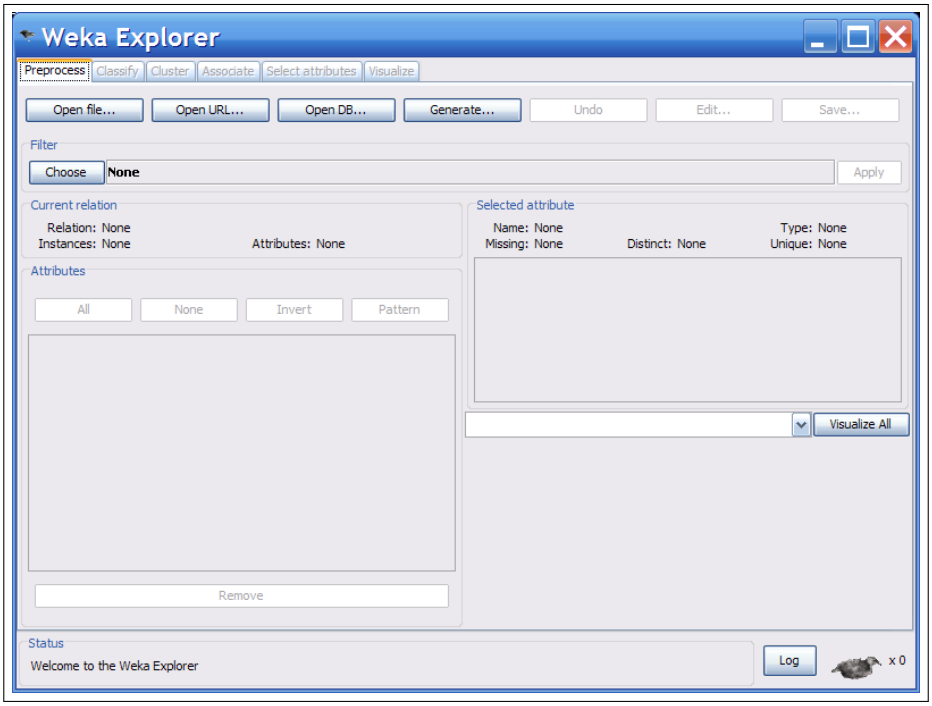

Figure 1: The main menu of Weka system

- **Classification** (*Classify*) - A classifier is a mapping from a (discrete or continuous) feature space X to a discrete set of labels Y [20]. Classification or discriminant analysis predicts class labels. This is supervised classification which provides a collection of labeled pre-classified patterns; the problem being to label a newly encountered, still unlabeled, pattern. In e-learning, classification has been used for: discovering potential student groups with similar characteristics and reactions to a specific pedagogical strategy [21]; predicting students performance and their final grade [22]; detecting students misuse or students playing around [23]; predicting the students performance, as well as assessing the relevance of the attributes involved [24]; grouping students as hint-driven or failure-driven and finding students common misconceptions [25]; identifying learners with little motivation and finding remedial actions in order to lower drop-out rates [26]; for predicting course success [27].

- **Clustering** (*Cluster*) - Clustering is a process of grouping objects into classes of similar objects [28]. It is an unsupervised classification or partitioning of patterns into groups or subsets (clusters) based on their locality and connectivity within an n-dimensional space. In e-learning, clustering has been used for: finding clusters of students with similar learning characteristics, and for promoting group-based collaborative learning, as well as for providing incremental learner diagnosis [29]; grouping students and personalized itineraries for courses based on learning objects [30]; grouping students in order to give them differentiated guiding according to their skills and other characteristics [31]; grouping tests and questions into related groups based on the data in the score matrix [32].

- **Association rule mining** (*Associate*) - Association rule mining discovers relationships among attributes in databases, producing if-then statements concerning attribute-values [33]. An association rule expresses a close correlation between items (attribute-value) in a database with values of support and confidence. The confidence of the rule is the percentage of transactions that contains the consequence in transactions that contain the

antecedent. The support of the rule is the percentage of transactions that contains both antecedent and consequence in all transactions in the database. Association rule mining has been applied to web-based educational systems for: building recommender agents that could recommend on-line learning activities or shortcuts [34]; diagnosing student learning problems and offering students advice [35]; guiding the learners activities automatically and recommending learning materials [36]; determining which learning materials are the most suitable to be recommended to the user [37]; identifying attributes characterizing patterns of performance disparity between various groups of students [38]; discovering interesting relationships from students usage information in order to provide feedback to the author of the course [39]; finding out relationships in learners behavior patterns [40]; finding students mistakes that often accompany each other [41]; guiding the search for the best fitting transfer models of student learning [42]; and optimizing the content of the e-learning portal by determining what most interests the user [43].

- **Selecting Attributes** (*Select attributes*) - Attribute selection involves searching through all possible combinations of attributes in the data to find which subset of attributes works best for prediction. To do this, two objects must be set up: an attribute evaluator and a search method. The evaluator determines what method is used to assign a value to each subset of attributes. The search method determines what style of search is performed.

- **Visualization** (*Visualize*) - Information visualization [44] is a branch of computer graphics and user interface which is concerned with the presentation of interactive or animated digital images so that users can understand data. These techniques facilitate analysis of large amounts of information by representing the data in some visual display. Weka visualization section allows visualizing 2D plots of the current relation.

Rosetta system and Weka are particularly suitable for data analysis in the field of education because they offer a selection of DM techniques and are relatively easy to use.

3 Methodology

The data sample, in form of MS Excel document, consists of a total of 256 instances (students). It is important to mention that the data are not collected with the aim to be analyzed by DM techniques. The survey was conducted on students from the territory of the Republic of Serbia and the territory of Bosnia and Herzegovina. The computer technology that is used in this region is mostly out of date but is sufficient for elementary usage in education. Description of the data, presented in Table 1, shows the names of attributes and their possible values. Attribute names and associated values comply with the form of survey.

Various techniques can be used on this small data set: First of all, there are statistical techniques, for example students distribution that is used when estimating the mean of a normally distributed population when the data set is small; then there are techniques for inferring decision rules (based on Pawlaks rough sets theory, decision trees, etc.), even neural networks can be trained on small data set [45]. In addition, the data sample which is described in Table 1 can be analyzed by various DM techniques or DM systems. The association rule mining is adopted as the most suitable DM technique.

As defined by Agrawal et al. [33] the problem of association rule mining is defined as:

Let $U = \{u_1, u_2, ..., u_m\}$ be a discrete universe, a finite set of objects. Let $A = \{a_1, a_2, ..., a_n\}$ be a finite set of attributes with binary values. Each object of universe U is described by attributes a_i, $i = 1, 2, ..., n$ thus generating a data set. An associative rule is defined as an implication of the form $X \Rightarrow Y$ where $X, Y \in A$ and $X \cap Y \neq \oslash$. The set of attributes X is

Table 1: Description of used data, attribute names and their possible values

Name	Value, number of objects and distribution	Description of the attribute
A1	1. yes, 210 , 82.03125% 2. no, 46, 17.96875%	Does the student has a computer at home
A2	1. do not know, 31, 12.109375% 2. other, 7, 2.734375% 3. PII, 7, 2.734376% 4. PIII, 60, 23.4375% 5. PIV, 144, 56.25% 6. laptop, 7, 2.734376%	What type of computer do the student have
A3	1. yes, 123, 48.046875% 2. no, 133, 51.953125%	Does the student use the Internet
A4	1. 1 hour per day, 98, 38.28125% 2. 2 hours per day, 85, 33.203125% 3. 3 hours a day, 23, 8.984375% 4. 4 hours a day, 24, 9.375% 5. 5 hours a day, 26, 10.15625%	How many hours per day does the student use the computer
A5	1. 0 hours a day, 67, 26.171875% 2. 1 hour per day, 74, 28.90625% 3. 2 hours per day, 64, 25.0% 4. 3 hours a day, 18, 7.03125% 5. 4 hours a day, 12, 4.6875% 6. 5 hours a day, 21, 8.203125%	How many hours per day the student use the Internet
A6	1. yes, 89, 34.765625% 2. no, 167, 65.234375%	Does the student use e-mail
A7	1. web sites on Serbian, 142, 55.46875% 2. web sites on other lenguages, 114, 44.53125%	What web sites does the student visit most frequently
A8	1. educational, 27, 10.546875% 2. entertainment, 116, 45.3125% 3. other, 113, 44.140625%	What kind of web sites is the most visited by student
A9	1. yes, 185, 72.265625% 2. no, 71, 27.734375%	Does the student use his/her home computer for learning
A10	1. educational, 42, 16.40625% 2. film, 18, 7.03125% 3. music, 130, 50.78125% 4. other, 66, 25.78225%	What type of media does the student use most frequently at home
A11	1. yes, 159, 62.109375% 2. no, 5, 1.953125% 3. do not know, 92, 35.9375%	Does the student want more educational computer software to be used in school in order to improve teaching
A12	1. yes, 192, 75.0% 2. no, 30, 11.71875% 3. do not know, 34, 13.28125%	Does the student want to review learning materials used in school, by Distant Learning System at home
A13	1. games, 78, 30.46875% 2. educational, 11, 4.296875% 3. games, educational, 167, 65.234375%	What type of software is most frequently used by student
A14	1. yes, 221, 86.328125% 2. no, 35, 13.671875%	Does the student like the subject of informatics
A15	1. yes, 219, 85.546875% 2. no, 37, 14.453125%	Does the student believe that he/she gained enough knowledge to work independently on a computer
A16	1. independently, 82, 32.03125% 2. from others, 36, 14.0625% 3. at school, 138, 53.90625%	From who or where has the student learned most about computer usage
A17	1. yes, 235, 91.796875% 2. no, 21, 8.203125%	Does the student think that his/her computer skills will help him/her in the future

Table 2: Simple example of data set

Object	The student has a computer at home (A1)	The student uses the Internet (A2)	The student thinks that his/her computer skills will help him/her in the future (A3)
1.	yes	yes	no
2.	no	yes	yes
3.	no	no	no
4.	yes	yes	yes
5.	no	yes	no

called antecedent (left-hand-side or LHS) of the rule; the set of attributes Y is called consequent (right-hand-side or RHS) of the rule.

There are many rules of the form $X \Rightarrow Y$, but to select interesting rules from the set of all possible rules, various measures of significance can be used; the best-known are minimum thresholds on support and confidence. The support $supp(X)$ is defined as the proportion of objects in the data set which contains the attributes from X. The confidence of a rule is defined as:

$$conf(X \Rightarrow Y) = \frac{supp(X \cup Y)}{supp(X)}$$

The previous concepts are explained in next simple example. For a data set given in Table 2 it is possible to infer some association rules, as well as the confidence and support parameters.

For $X = \{A1, A2, A3\}$ $supp(X) = \frac{1}{5} = 0.2$ because there is one object (number four) for which there is a "yes" value for every attribute.

For example, the confidence of the rule $X \Rightarrow Y$, where $X = \{A1, A2\}$ and $Y = \{A3\}$ is:

$$conf(X \Rightarrow Y) = \frac{supp(X \cup Y)}{supp(X)} = \frac{0.2}{0.4} = 0.5$$

The previous rule $X \Rightarrow Y$ is interpreted as follows: The student who has a computer at home and uses the Internet is associated with the opinion that his/her computer skills will help him/her in the future.

There are many algorithms for association rule computation, but so called apriori algorithm [46] is the best-known algorithm to mine association rules. It is based on breadth-first search strategy [47]. In the data set described by Table 2 attribute selection is more complicated by the fact that some attribute values are not binary so this involves more extensive search.

The association rule generation is chosen to be performed due to multiple reasons:

1. This is an exact method and therefore excludes any subjective influence while analyzing data set.

2. The result is presented in a readable and easy-to-understand form.

3. It is expected that number of generated association rules would not be high (data set contains 256 instances) due to computation of confidence for each rule and selection of rules with the highest confidence.

4. It is expected that association rules have a great value when inferred from data set in education domain because association rules can be treated as a hypothesis.

The experiment was conducted on data set by software system Weka in order to generate association rules. After loading, the data are ready for pre-processing and application of DM techniques. Weka system requires that the attributes with numerical values do not participate in the association rule mining, so they are ignored. Association rules generated by Weka system (apriori algorithm is used) are shown in Table 3.

Table 3: Association rules generated by apriori algorithm

Rule	IF	THEN	Rule confidence
1	A11=yes, 159	A17=yes, 154	0.97
2	A9=yes AND A14=yes, 159	A17=yes, 154	0.97
3	A13=games, educational, 167	A17=yes, 159	0.95
4	A9=yes AND A15=yes, 166	A17=yes, 158	0.95
5	A12=yes AND A14=yes, 169	A17=yes, 160	0.95
6	A9=yes, 185	A17=yes, 174	0.94
7	A14=yes AND A15=yes, 195	A17=yes, 183	0.94
8	A12=yes, 192	A17=yes, 180	0.94
9	A6=no, 167	A17=yes, 156	0.93
10	A14=yes, 221	A17=yes, 206	0.93

There are 10 rules generated. The IF part of every rule is followed by support measure, as is the case with the THEN part of each rule. The confidence for each rule is given in the separate column.

4 Results

The Analysis of generated association rules provides insight into the dependence of the monitored parameters. Each rule is accompanied by a factor of confidence that takes a value in the range [0, 1]. Ten association rules have been generated, see Table 3.

By association rule 1 attribute A11 (Does the student want more educational computer software to be used in school in order to improve learning) is associated with attribute A17 (Does the student think that his/her computer skills will help him/her in the future). The factor of confidence for this rule is 0.97, that rule is guaranteed to a great extent. If the value of attribute A11 is "yes" then, by this association rule, the value of attribute A17 is also "yes". A possible conclusion which can be drawn is that most of the students think that usage of computer technology in school generates skills and knowledge that could be used in the future. This students opinion is a good indicator of the importance of the usage of computer technology and educational software in teaching process. It is evident that technology usage in future is closely related to computer technology software and methods which are used at present. Other association rules can be interpreted in analogous way.

The rule 2 can be interpreted as follows. Two facts: The fact that student uses home computer for learning and the fact that student likes the subjects of computer science, are associated with opinion that computer skills will be helpful in the future. Rule number 3 associates the type of software that is most frequently used by student (games, educational software) with the opinion that computer skills will be helpful in the future.

The opinion that computer skills will be helpful in the future is mostly associated with: believing that student gained enough knowledge to work independently on a computer, the usage of the computer for learning and aspiration to use distant learning system at home to review learning materials previously used in school. Attribute A14 (The student likes the subject of computer science) is most frequently associated with opinion that computer skills will be helpful in the future.

However, rule number nine associate the students that are not satisfied with the usage of e-mail with the opinion that computer skills will be helpful in the future. In fact, this may be in accordance with rule number eight: The usage of distant learning system at home to review learning materials previously used in school is associated with the opinion that computer skills will be helpful in the future. So, in students opinion, distant learning system makes e-mail service obsolete in a way.

5 Conclusions and Future Works

The application of association rule mining allows automatic generation of hypotheses and factors of confidence that are related to them. Considering rule number one (see Table 3) the hypothesis is: If a student wants more educational computer software to be used in school in order to improve learning then the student thinks that his/her computer skills will help him/her in the future. Other rules may also be interpreted as a hypothesis. Obviously the factor of confidence has a great impact on confirmation of hypotheses. At the Technical Faculty "Mihajlo Pupin" in Zrenjanin, Serbia, extensive research is underway, investigating the possibilities of applying DM techniques to data from the domain of education, extracted from a survey conducted in the wider Balkan region. The fact that it is not obligatory that the data are collected with the aim to be analyzed by DM techniques, offers an excellent chance to assess real possibilities in the actual practice. In future, this leads to identification of the advantages of DM techniques over standard statistical techniques. So far, there have been identified the following advantages of application of the DM techniques to data from the domain of education:

- It is possible to extract the "special" cases of the statistical rarity that are often discarded as "noise". The discovery of such "isolated" cases is conducted by automated IF ... THEN rule synthesis. As in [48] the association data mining based on a uniform support misses some patterns of low support. Although it is possible to exploit support constrains, which specifies some minimum supports, we opted to use automated IF ... THEN rules synthesis by Rosetta system or similar.

- Clustering technique allows the exact definition of average student and special groups (clusters) of students, so that further action of e-learning can be implemented while respecting the characteristics of each cluster.

- Analysis of generated association rules provides insight into the dependence of the monitored parameters.

- Ranking of attributes (parameters) in order of importance of influence on a selected attribute allows the rejection of the parameters of lesser importance.

- Data visualization provides a figurative description of the data, so we can actually see patterns which may exist.

The usage of Weka system, or a similar system, is of great help because it generates association rules (hypotheses) automatically. Furthermore, Weka system generates association rules for which the factor of confidence is high. This method can be used in any case of data, but Weka system requires that attributes with numerical values do not participate in generating the association rule mining, so they are ignored. This is not a great handicap because linguistic terms are frequently used in queries and surveys. The final conclusion is: the usage of Weka system in order to generated association rules automatically is of great help because the hypotheses of

little importance are avoided.

Future work will be practical and it will refer to the selection of a particular CMS and integration with DM system.

Bibliography

[1] C. Chih-Ming, Personalized E-learning system with self-regulated assisted mechanisms for promoting learning performance", *An Int. J. of Expert Systems with Applications* 36: 8816-8829, 2009.

[2] C. Romero, S. Ventura, E. Garcia, Data mining in course management system: Moodle case study and tutorial, *An Int. J. of Computers and Education* 51, pp. 368-384, 2008.

[3] E. Brtka, D. Radosav, V. Brtka, The Data Mining module as a part of the e-learning system, (In Serbian), *In Proc. of InfoTech Conference*, Vrnjacka Banja, Serbia, 2009.

[4] E. Brtka, The data mining analysis approach in pedagogical research, Master Thesis, (In Serbian), Technical Faculty "Mihajlo Pupin", Zrenjanin, Serbia, 2009.

[5] E. Gaudioso, L. Talavera, Data mining to suport tutoring in virtual learning communities: Experiences and challenges. *In C. Romero and S. Ventura (Eds.), Data mining in e-learning*, Southampton UK, Wit Press, pp. 207-226, 2006.

[6] K. Diego, Data Mining and Statistics: What is the Connection?, *The Data Administration Newsletter*, LLC - www.TDAN.com

[7] D. Hand, H. Mannila, P. Smyth, *Principles of Data Mining*, MIT Press, Cambridge, MA. ISBN 0-262-08290-X. OCLC 226126187, 2001.

[8] DBMiner (2007), http://www.dbminer.com

[9] Clementine (2007), http://www.spss.com/clementine/

[10] Miner (2007), http://www-306.ibm.com/software/data/iminer/

[11] Weka (2007), http://www.cs.waikato.ac.nz/ml/weka/

[12] Keel (2007), http://www.keel.es/

[13] O. Zaiane, J. Luo, Web usage mining for a better web-based learning environment, *In Proc. of conf. on advanced technology for education*, Banff, Alberta, pp. 60-64, 2001.

[14] D. Silva, M. Vieira, Using data warehouse and data mining resources for ongoing assessment in distance learning, *In IEEE int. conf. on advanced learning technologies*, Kazan, Russia, pp. 40-45, 2002.

[15] J. Tane, C. Schmitz, G. Stumme, Semantic resource management for the web: An elearning application, *In Proc. of the WWW conference*, New York, USA, pp. 1-10, 2004.

[16] E. Garcia, C. Romero, S. Ventura, C. Castro, Using rules discovery for the continuous improvement of e-learning courses, *In Int. conf. intelligent data engineering and automated learning*, Burgos, Spain, pp. 887-895, 2006.

[17] Z. Pawlak, A. Skowron, Rudiments of rough sets, *An Int. J. of Information Sciences* 177:3-27, 2007.

[18] A. Ohrn, *Discernibility and Rough Sets in Medicine: Tools and Applications*, PhD thesis, Department of Computer and Information Science, Norwegian University of Science and Technology, Trondheim, Norway, 1999.

[19] I. H. Witten, E. Frank, *Data Mining: Practical machine learning tools and techniques*, 2nd Edition, Morgan Kaufman, San Francisco, 2005.

[20] R. O. Duda, P. E. Hart, D. G. Stork, *Pattern classification*, Wiley Interscience, 2000.

[21] G. Chen, C. Liu, K. Ou, B. Liu, Discovering decision knowledge from web log portfolio for managing classroom processes by applying decision tree and data cube technology, *Journal of Educational Computing Research* 23(3):305-332, 2000.

[22] B. Minaei-Bidgoli, W. Punch, Using genetic algorithms for data mining optimization in an educational web-based system, *In Genetic and evolutionary computation conference*, Chicago, USA, pp. 2252-2263, 2003.

[23] R. Baker, A. Corbett, K. Koedinger, Detecting student misuse of intelligent tutoring systems, *In Intelligent tutoring systems*, Alagoas, Brazil, pp. 531-540, 2004.

[24] S. B. Kotsiantis, C. J. Pierrakeas, P. E. Pintelas, Predicting students performance in distance learning using machine learning techniques", *Applied Artificial Intelligence* 18(5):411-426, 2004.

[25] M. V. Yudelson, O. Medvedeva, E. Legowski, M. Castine, D. Jukic, C. Rebecca, Mining student learning data to develop high level pedagogic strategy in a medical ITS, *In Proceedings of AAAI workshop on educational data mining*, Boston, pp. 1-8, 2006.

[26] M. Cocea, S. Weibelzahl, Can log files analysis estimate learners level of motivation? *In Proceedings of the workshop week Lernen - Wissensentdeckung - Adaptivitat*, Hildesheim, pp. 32-35, 2006.

[27] W. Hamalainen, M. Vinni, Comparison of machine learning methods for intelligent tutoring systems, *In Proceedings of the eighth international conference in intelligent tutoring systems*, Taiwan, pp. 525-534, 2006.

[28] A. K. Jain, M. N. Murty, P. J. Flynn, Data clustering: A review, *ACM Computing Surveys* 31(3):264-323, 1999.

[29] T. Tang, G. McCalla, Smart recommendation for an evolving e-learning system, *International Journal on E-Learning* 4(1):105-129, 2005.

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

[31] W. Hamalainen, J. Suhonen, E. Sutinen, H. Toivonen, "Data mining in personalizing distance education courses", *In World conference on open learning and distance education*, Hong Kong, pp. 1-11, 2004.

[32] J. Spacco, T. Winters, T. Payne, T. Inferring use cases from unit testing, *In AAAI workshop on educational data mining*, New York, pp. 1-7, 2006.

[33] R. Agrawal, T. Imielinski, A. Swami, Mining association rules between sets of items in large databases, *In Proc.of the ACM SIGMOD international conference on management of data*, Washington DC, USA, pp. 1-22, 1993.

[34] O. Zaiane, Building a recommender agent for e-learning systems, *In Proc.of the int. conference in education*, Auckland, New Zealand, pp. 55-59, 2002.

[35] G. J. Hwang, C. L. Hsiao, C. R. Tseng, A computer-assisted approach to diagnosing student learning problems in science courses, *Journal of Information Science and Engineering* 19: 229-248, 2003.

[36] J. Lu, Personalized e-learning material recommender system, *In International conference on information technology for application*, Utah, USA, pp. 374-379, 2004.

[37] P. Markellou, I. Mousourouli, S. Spiros, A. Tsakalidis, Using semantic web mining technologies for personalized e-learning experiences, *In Proc. of the web-based education*, Grindelwald, Switzerland, pp. 461-826, 2005.

[38] B. Minaei-Bidgoli, P. Tan, W. Punch, Mining interesting contrast rules for a web-based educational system, *In Int. conf.on machine learning applications*, Los Angeles, California, pp. 1-8, 2004.

[39] C. Romero, S. Ventura, 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* 14(5):425-464, 2004.

[40] P. Yu, C. Own, L. Lin, On learning behavior analysis of web based interactive environment, *In Proc. of the implementing curricular change in engineering education*, Oslo, Norway, pp. 1-10, 2001.

[41] A. Merceron, K. Yacef,Mining student data captured from a web-based tutoring tool: Initial exploration and results, *Journal of Interactive Learning Research* 15(4):319-346, 2004.

[42] J. Freyberger, N. Heffernan, C. Ruiz, Using association rules to guide a search for best fitting transfer models of student learning, *In Workshop on analyzing studenttutor interactions logs to improve educational outcomes at ITS conference*, Alagoas, Brazil, pp. 1-10, 2004.

[43] A. A. Ramli, Web usage mining using apriori algorithm: UUM learning care portal case, *In Int. conf. on knowledge management*, Malaysia, pp. 1-19, 2005.

[44] R. Spence, *Information visualization*, Addison-Wesley, 2001.

[45] R. Andonie, Extreme Data Mining: Inference from Small Datasets, *INT J COMPUT COMMUN*, ISSN 1841-9836, Vol. 5(3):280-291, 2010.

[46] R. Agrawal, R. Srikant, Fast algorithms for mining association rules in large databases, *In Jorge B. Bocca, Matthias Jarke, and Carlo Zaniolo (eds.), Proc. of the 20th International Conference on Very Large Data Bases*, VLDB, Santiago, Chile, pp. 487-499, 1994.

[47] G. Luger, W. Stubblefield, Artificial Intelligence - structures and strategies for complex problem solving, University of New Mexico, Albuquerque, The Benjamin/Cummings Publishing Company Inc, 1993.

[48] M. Pater, D.E. Popescu, Multi-Level Database Mining Using AFOPT Data Structure and Adaptive Support Constrains, *INT J COMPUT COMMUN*, ISSN 1841-9836, 3(S):437-441, 2008.