Applying Data Mining Techniques to e-Learning Problems

Chapter · July 2007

DOI: 10.1007/978-3-540-71974-8_8

Félix Agustín Castro Espinoza
Autonomous University of Hidalgo

Alfredo Vellido
Universitat Politècnica de Catalunya

Àngela Nebot
Universitat Politècnica de Catalunya

Applying Data Mining Techniques to e-Learning Problems

Félix Castro[1, 2], Alfredo Vellido[1], Àngela Nebot[1], and Francisco Mugica[3]

[1] Dept. Llenguatges i Sistemes Informatics, LSI, Universitat Politècnica de Catalunya, Campus Nord, C. Jordi Girona 1-3, Barcelona 08034, España
{fcastro, avellido, angela}@lsi.upc.edu
[2] Centro de Investigación en Tecnologías de Información y Sistemas, CITIS, Universidad Autónoma del Estado de Hidalgo, Ciudad Universitaria, Carretera Pachuca-Tulancingo km. 4.5, Hidalgo, México
fcastro@lsi.upc.edu
[3] Instituto Latinoamericano de la Comunicación Educativa (ILCE), Calle del Puente 45, México D. F. 14380, México
fmugica@ilce.edu.mx

Abstract. This chapter aims to provide an up-to-date snapshot of the current state of research and applications of Data Mining methods in e-learning. The cross-fertilization of both areas is still in its infancy, and even academic references are scarce on the ground, although some leading education-related publications are already beginning to pay attention to this new field. In order to offer a reasonable organization of the available bibliographic information according to different criteria, firstly, and from the Data Mining practitioner point of view, references are organized according to the type of modelling techniques used, which include: Neural Networks, Genetic Algorithms, Clustering and Visualization Methods, Fuzzy Logic, Intelligent agents, and Inductive Reasoning, amongst others. From the same point of view, the information is organized according to the type of Data Mining problem dealt with: clustering, classification, prediction, etc. Finally, from the standpoint of the e-learning practitioner, we provide a taxonomy of e-learning problems to which Data Mining techniques have been applied, including, for instance: Students' classification based on their learning performance; detection of irregular learning behaviours; e-learning system navigation and interaction optimization; clustering according to similar e-learning system usage; and systems' adaptability to students' requirements and capacities.

1 Introduction

Within a decade, the Internet has become a pervasive medium that has changed completely, and perhaps irreversibly, the way information and knowledge are transmitted and shared throughout the world. The education community has not limited itself to the role of passive actor in this unfolding story, but it has been at the forefront of most of the changes.

Indeed, the Internet and the advance of telecommunication technologies allow us to share and manipulate information in nearly real time. This reality is determining the next generation of distance education tools. Distance education arose from traditional education in order to cover the necessities of remote students and/or help the teaching-learning process, reinforcing or replacing traditional education. The Internet takes this process of delocalization of the educative experience to a new realm, where the lack of presential intercourse is, at least partially, replaced by an increased level of technology-mediated interaction. Furthermore, telecommunications allow this interaction to take forms that were not available to traditional presential and distance learning teachers and learners.

This is e-learning (also referred to as web-based education and e-teaching), a new context for education where large amounts of information describing the continuum of the teaching-learning interactions are endlessly generated and ubiquitously available. This could be seen as a blessing: plenty of information readily available just a click away. But it could equally be seen as an exponentially growing nightmare, in which unstructured information chokes the educational system without providing any articulate knowledge to its actors.

Data Mining was born to tackle problems like this. As a field of research, it is almost contemporary to e-learning. It is, though, rather difficult to define. Not because of its intrinsic complexity, but because it has most of its roots in the ever-shifting world of business. At its most detailed, it can be understood not just as a collection of data analysis methods, but as a data analysis process that encompasses anything from data understanding, pre-processing and modelling to process evaluation and implementation [16]. It is nevertheless usual to pay preferential attention to the Data Mining methods themselves. These commonly bridge the fields of traditional statistics, pattern recognition and machine learning to provide analytical solutions to problems in areas as diverse as biomedicine, engineering, and business, to name just a few. An aspect that perhaps makes Data Mining unique is that it pays special attention to the compatibility of the modelling techniques with new Information Technologies (IT) and database technologies, usually focusing on large, heterogeneous and complex databases. E-learning databases often fit this description.

Therefore, Data Mining can be used to extract knowledge from e-learning systems through the analysis of the information available in the form of data generated by their users. In this case, the main objective becomes finding the patterns of system usage by teachers and students and, perhaps most importantly, discovering the students' learning behavior patterns.

This chapter aims to provide an as complete as possible review of the many applications of Data Mining to e-learning over the period 1999-2006; that is, a survey of the literature in this area up to date. We must acknowledge that this is not the first time a similar venture has been undertaken: a collection of papers that cover most of the important topics in the field was concurrently presented in [71].

The findings of the survey are organized from different points of view that might in turn match the different interests of its potential readers: The surveyed research can be seen as being displayed along two axes: Data Mining problems and methods, and e-learning applications. Section 2 presents the research along the axis of the Data Mining modelling techniques and methods, while section 3 presents the surveyed content along the e-learning applications axis. This organization of the surveyed content should allow readers to access the information in a more compact and self-contained way than that in [71].

A deeper analysis and discussion of the actual state of the research in the field is presented in section 4, highlighting its opportunities and limitations. Section 5 reports work on Data Mining in e-learning beyond academic publications. Finally, section 6 summarizes the findings and draws some conclusions.

Most of the information provided in this chapter takes the form of tables of publications. We consider this the best (or at least the most compact) way to organize it and ease, in a guided manner, the access to the main contents.

2 A survey of Data Mining in e-learning from the Data Mining point of view

As stated in the introduction, we aim to organize the findings of the survey in different ways that might correspond to the diverse readers' academic or professional backgrounds. In this section, we present the surveyed research according to the Data Mining problems (classification, clustering, etc.), techniques and methods (e.g., Neural Networks, Genetic Algorithms, Decision Trees, or Fuzzy Logic).

In fact, most of the existing research addresses problems of classification and clustering. For this reason, specific subsections will be devoted to them. But first, let us try to find a place for Data Mining in the world of e-learning.

2.1 Where does Data Mining fit in e-learning processes?

Some researchers have pointed out the close relation between the fields of Artificial Intelligence (AI) and Machine Learning (ML) -main sources of Data Mining techniques and methods- and education processes [4, 26, 30, 49, 79, 85].

In [4], the author establishes the research opportunities in AI and education on the basis of three models of educational processes: *models as scientific tool,* are used as a means for understanding and forecasting some aspect of an educational situation; *models as component*: corresponding to some characteristic of the teaching or learning process and used as a component of an educative artefact; and *models as basis for design of educational artefacts*: assisting the design of computer tools for education by providing design methodologies and system components, or by constraining the range of tools that might be available to learners.

In [49, 85], studies on how Data Mining techniques could successfully be incorporated to e-learning environments and how they could improve the learning tasks were carried out. In [85], data clustering was suggested as a means to promote group-based collaborative learning and to provide incremental student diagnosis.

A review of the possibilities of the application of Web Mining (Web usage mining and clustering) techniques to meet some of the current challenges in distance education was presented in [30]. The proposed approach could improve the effectiveness and efficiency of distance education in two ways: on the one hand, the discovery of aggregate and individual paths for students could help in the development of effective customized education, providing an indication of how to best organize the educator organization's courseware. On the other hand, virtual knowledge structure could be identified through Web Mining methods: The discovery of Association Rules could make it possible for Web-based distance tutors to identify knowledge patterns and reorganize the virtual course based on the patterns discovered.

An analysis on how ML techniques -again, a common source for Data Mining techniques- have been used to automate the construction and induction of student models, as well as the background knowledge necessary for student modelling, were presented in [79]. In this paper, the difficulty, appropriateness and potential of applying ML techniques to student modelling was commented.

2.2 The classification problem in e-learning

In classification problems, we usually aim to model the existing relationships (if any) between a set of multivariate data items and a certain set of outcomes for each of them in the form of class membership labels. Although plenty of classification methods that would fit in a Data Mining process exist, in what follows, we shall see that only a few techniques (or families of techniques) have been applied to e-learning.

2.2.1 Fuzzy logic methods

Fuzzy logic-based methods have only recently taken their first steps in the e-learning field [36, 39, 40, 81, 89].

In [81], a neuro-fuzzy model for the evaluation of students in an intelligent tutoring system (ITS) was presented. Fuzzy theory was used to measure and transform the interaction between the student and the ITS into linguistic terms. Then, Artificial Neural Networks were trained to realize fuzzy relations operated with the max-min composition. These fuzzy relations represent the estimation made by human tutors of the degree of association between an observed response and a student characteristic.

A fuzzy group-decision approach to assist users and domain experts in the evaluation of educational web sites was realized in the EWSE system, presented in [39]. In further work by Hwang and colleagues [36, 40], a fuzzy rules-based method for eliciting and integrating system management knowledge was proposed and served as the basis for the design of an intelligent management system for monitoring educational Web servers. This system is capable of predicting and handling possible failures of educational Web servers, improving their stability and reliability. It assists students' self-assessment and provides them with suggestions based on fuzzy reasoning techniques.

A two-phase fuzzy mining and learning algorithm was described in [89]. It integrates an association rule mining algorithm, called *Apriori*, with fuzzy set theory to find embedded information that could be fed back to teachers for refining or reorganizing the teaching materials and tests. In a second phase, it uses an inductive learning algorithm of the AQ family: AQR, to find the concept descriptions indicating the missing concepts during students' learning. The results of this phase could also be fed back to teachers for refining or reorganizing the learning path.

2.2.2 Artificial Neural Networks and Evolutionary Computation

Some research on the use of Artificial Neural Networks and Evolutionary Computation models to deal with e-learning topics can be found in [53, 55, 87].

A navigation support system based on an Artificial Neural Network (more precisely, a Multi-Layer Perceptron, or MLP) was put forward in [55] to decide on the appropriate navigation strategies. The Neural Network was used as a navigation strategy decision module in the system. Evaluation has validated the knowledge learned by the Neural Network and the level of effectiveness of the navigation strategy.

In [53, 87], evolutionary algorithms were used to evaluate the students' learning behaviour. A combination of multiple classifiers (CMC), for the classification of students and the prediction of their final grades, based on features extracted from logged data in an education web-based system, was described in [53]. The classification and prediction accuracies are improved through the weighting of the data feature vectors using a Genetic Algorithm. In [87] we find a random code generation and mutation process suggested as a method to examine the comprehension ability of students.

2.2.3 Graphs and Trees

Graph and/or tree theory was applied to e-learning in [9, 13, 14, 29, 42, 47, 48, 95, 97].

An e-learning model for the personalization of courses, based both on the student's needs and capabilities and on the teacher's profile, was described in [9]. Personalized learning paths in the courses were modelled using graph theory. In [47, 48], Decision Trees (DT) as classification models were applied. A discussion of the implementation of the Distance Learning Algorithm (DLA), which uses Rough Set theory to find general decision rules, was presented by [47]: A DT was used to adequate the original algorithm to distance learning issues. On the basis of the obtained results, the instructor might consider the reorganization of the course materials. A system architecture for mining learners' online behaviour patterns was put forward in [13]. A framework for the integration of traditional Web log mining algorithms with pedagogical meanings of Web pages was presented. The approach is based on the definition of an e-learning system concept-hierarchy and the sequential patterns of the pages shown to users.

Also in [48], an automatic tool, based on the students' learning performance and communication preferences, for the generation and discovery of simple student models was described, with the ultimate goal of creating a personalized education environment. The approach was based on the PART algorithm, which produces rules from pruned partial DTs. In [97], a tool that can help trace deficiencies in students' understanding was presented. It resorts to a tree abstract data type (ADT), built from the concepts covered in a lab, lecture, or course. Once the tree ADT is created, each node can be associated with different entities such as student performance, class performance, or lab development. Using this tool, a teacher could help students by discovering concepts that needed additional coverage, while students might discover concepts for which they would need to spend additional working time.

A tool to perform a quantitative analysis based on students' learning performance was introduced in [14]. It proposes new courseware diagrams, combining tools provided by the theory of conceptual maps [63] and influence diagrams [75]. In [29, 42, 95], personalized Web-based learning systems were defined, applying Web usage mining techniques to personalized recommendation services. The approach is based on a Web page classification method, which uses attribute-oriented induction according to related domain knowledge shown by a concept hierarchy tree.

2.2.4 Association Rules

Association Rules for classification, applied to e-learning, have been investigated in the areas of learning recommendation systems [18, 98, 99], learning material organization [89], student learning assessments [38, 45, 52, 54, 69, 70], course adaptation to the students' behaviour [19, 35, 50], and evaluation of educational web sites [21].

Data Mining techniques such as Association Rule mining, and inter-session and intra-session frequent pattern mining, were applied in [98, 99] 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 [54], where contrast rules, defined as sets of conjunctive rules describing patterns of performance disparity between groups of students, were used. A computer-assisted approach to diagnosing student learning problems in science courses and offer students advice was presented in [38], based on the concept effect relationship (CER) model (a specification of the Association Rules technique).

A hypermedia learning environment with a tutorial component was described in [19]. It is called *Logiocando* and targets children of the fourth level of primary school (9-10 years old). It includes a tutor module, based on if-then rules, that emulates the teacher by providing suggestions on how and what to study. In [52] we find the description of a learning process assessment method that resorts to Association Rules, and the well-known ID3 DT learning method. A framework for the use of Web usage mining to support the validation of learning site designs was defined in [21], applying association and sequence techniques [80].

In [50], a framework for personalised e-learning based on aggregate usage profiles and a domain ontology were presented, and a combination of Semantic Web and Web mining methods was used. The *Apriori* algorithm for Association Rules was applied to capture relationships among URL references based on the navigational patterns of students. A test result feedback (TRF) model that analyzes the relationships between student learning time and the corresponding test results was introduced in [35]. The objective was twofold: on the one hand, developing a tool for supporting the tutor in reorganizing the course material; on the other, a personalization of the course tailored to the individual student needs. The approach was based in Association Rules mining.

A rule-based mechanism for the adaptive generation of problems in ITS in the context of web-based programming tutors was proposed in [45]. In [18], a web-based course recommendation system, used to provide students with suggestions when having trouble in choosing courses, was described. The approach integrates the *Apriori* algorithm with graph theory.

2.2.5 Multi-agent systems

Multi Agents Systems (MAS) for classification in e-learning have been proposed in [2, 28]. In [28] this takes the form of an adaptive interaction system based on three MAS: the *Interaction MAS* captures the user preferences applying some defined usability metrics (affect, efficiency, helpfulness, control and learnability). The *Learning MAS* shows the contents to the user according to the information collected by the Interaction MAS in the previous step; and the *Teaching MAS* offers recommendations to improve the virtual course. A multi-agent recommendation system, called InLix, was described in [2]; it suggests educational resources to students in a mobile learning platform. InLix combines content analysis and the development of students' virtual clusters. The model includes a process of classification and recommendation feedback in which the user agent learns from the student and adapts itself to the changes in user's interests. This provides the agent with the opportunity to be more accurate in future classification decisions and recommendation steps. Therefore, the more students use the system, the more the agent learns and more accurate its actions become.

2.3. The clustering problem in e-learning

Unlike in classification problems, in data grouping or clustering we are not interested in modelling a relation between a set of multivariate data items and a certain set of outcomes for each of them (being this in the form of class membership labels). Instead, we usually aim to discover and model the groups in which the data items are often clustered, according to some item similarity measure.

We find a first application of clustering methods in [37], where a network-based testing and diagnostic system was implemented. It entails a multiple-criteria test-sheet-generating problem and a dynamic programming approach to generate test sheets. The proposed approach employs fuzzy logic theory to determine the difficulty levels of test items according to the learning status and personal features of each student, and then applies an Artificial Neural Network model: Fuzzy Adaptive Resonance Theory (Fuzzy ART) [10] to cluster the test items into groups, as well as dynamic programming [22] for test sheet construction.

In [60, 61], an in-depth study describing the usability of Artificial Neural Networks and, more specifically, of Kohonen's Self-Organizing Maps (SOM) [43] for the evaluation of students in a tutorial supervisor (TS) system, as well as the ability of a fuzzy TS to adapt question difficulty in the evaluation process, was carried out. An investigation on how Data Mining techniques could be successfully incorporated to e-learning environments, and how this could improve the learning processes was presented in [85]. Here, data clustering is suggested as a means to promote group-based collaborative learning and to provide incremental student diagnosis.

In [86], user actions associated to students' Web usage were gathered and pre-processed as part of a Data Mining process. The Expectation-Maximization (EM) algorithm was then used to group the users into clusters according to their behaviours. These results could be used by teachers to provide specialized advice to students belonging to each cluster. The simplifying assumption that students belonging to each cluster should share web usage behaviour makes personalization strategies more scalable. The system administrators could also benefit from this acquired knowledge by adjusting the e-learning environment they manage according to it. The EM algorithm was also the method of choice in [82], where clustering was used to discover user behaviour patterns in collaborative activities in e-learning applications.

Some researchers [23, 31, 83] propose the use of clustering techniques to group similar course materials: An ontology-based tool, within a Web Semantics framework, was implemented in [83] with the goal of helping e-learning users to find and organize distributed courseware resources. An element of this tool was the implementation of the Bisection K-Means algorithm, used for the grouping of similar learning materials. Kohonen's well-known SOM algorithm was used in [23] to devise an intelligent searching tool to cluster similar learning material into classes, based on its semantic similarities. Clustering was proposed in [31] to group similar learning documents based on their topics and similarities. A Document Index Graph (DIG) for document representation was introduced, and some classical clustering algorithms (Hierarchical Agglomerative Clustering, Single Pass Clustering and k-NN) were implemented.

Different variants of the Generative Topographic Mapping (GTM) model, a probabilistic alternative to SOM, were used in [11, 12, 94] for the clustering and visualization of multivariate data concerning the behaviour of the students of a virtual course. More specifically, in [11, 94] a variant of GTM known to behave robustly in the presence of atypical data or outliers was used to successfully identify clusters of students with atypical learning behaviours. A different variant of GTM for feature relevance determination was used in [12] to rank the available data features according to their relevance for the definition of student clusters.

2.4 Other Data Mining problems in e-learning

As previously stated, most of the current research deals with problems of classification and clustering in e-learning environments. However, there are several applications that tackle other Data Mining problems such as prediction and visualization, which we review in this subsection.

2.4.1 Prediction techniques

Prediction is often also an interesting problem in e-learning, although it must be born in mind that it can easily overlap with classification and regression problems. The forecasting of students' behaviour and performance when using e-learning systems bears the potential of facilitating the improvement of virtual courses as well as e-learning environments in general.

A methodology to improve the performance of developed courses through adaptation was presented in [72, 73]. Course log-files stored in databases could be mined by teachers using evolutionary algorithms to discover important relationships and patterns, with the target of discovering relationships between students' knowledge levels, e-learning system usage times and students' scores.

A system for the automatic analysis of user actions in Web-based learning environments, which could be used to make predictions on future uses of the learning environment, was presented in [59]. It applies a C4.5 DT model for the analysis of the data; (Note that this reference could also have been included in the section reviewing classification methods).

Some studies apply regression methods for prediction [5, 27, 44]. In [27], a study that aimed to find the sources of error in the prediction of students' knowledge behaviour was carried out. Stepwise regression was applied to assess what metrics help to explain poor prediction of state exam scores. Linear regression was applied in [5] to predict whether the student's next response would be correct, and how long he or she would take to generate that response.

In [44], a set of experiments was conducted in order to predict the students' performance in e-learning courses, as well as to assess the relevance of the attributes involved. In this approach, several Data Mining methods were applied, including: Naïve Bayes, kNN, MLP Neural Network, C4.5, Logistic Regression and Support Vector Machines. With similar goals in mind, experiments applying the Fuzzy Inductive Reasoning (FIR) methodology to the prediction of the students' final marks in a course taken at a virtual campus were carried out in [62]. The relative relevance of specific features describing course online behaviour was also assessed. This work was extended in [25] using Artificial Neural Networks for the prediction of the students' final marks. In this work, the predictions made by the network were interpreted using Orthogonal Search-based Rule Extraction (OSRE) a novel rule extraction algorithm [24]. Rule extraction was also used in [72, 73] with the emphasis on the discovery of interesting prediction rules in student usage information, in order to use them to improve adaptive Web courses.

Graphical models and Bayesian methods have also been used in this context. For instance, an open learning platform for the development of intelligent Web-based educative systems, named MEDEA, was presented in [88]. Systems developed with MEDEA guide students in their learning process, and allow free navigation to better suit their learning needs. A Bayesian Network model lies at the core of MEDEA. In [3] an evaluation of students' attitudes and their relationship to students' performance in a tutoring system was implemented. Starting from a correlation analysis between variables, a Bayesian Network that inferred negative and positive students' attitudes was built. Finally, a Dynamic Bayes Net (DBN) was used in [15], for modelling students' knowledge behaviour and predict future performance in an ITS.

In [90, 91], a tool for the automatic detection of atypical behaviours on the students' use of the e-learning system was defined. It resorts to a Bayesian predictive distribution model to detect irregular learning processes on the basis of the students' response time. Note that some models for the detection of atypical student behaviour were also referenced in the section reviewing clustering applications [11, 94].

2.4.2 Visualization techniques

One of the most important phases of a Data Mining process (and one that is usually neglected) is that of data exploration through visualization methods.

Visualization was understood in [68] in the context of Social Network Analysis adapted to collaborative distance-learning, where the cohesion of small learning groups is measured. The cohesion is computed in several ways in order to highlight isolated people, active sub-groups and various roles of the members in the group communication structure. Note the links between this goal and that of atypical student behaviour described in previous sections. The method allows the display of global properties both at individual level and at group level, as well as to efficiently assist the virtual tutor in following the collaboration patterns within the group.

An educational Data Mining tool is presented in [57, 58] that shows, in a hierarchical and partially ordered fashion, the students' interaction with the e-learning environment and their virtual tutors. The tool provides case analysis and visualizes the results in an event tree, exploiting MySQL databases to obtain tutorial events.

One main limitation to the analysis of high-dimensional multivariate data is the difficulty of representing those data faithfully in an intuitive visual way. Latent methods (of which Principal Component Analysis, or PCA, is perhaps the most widely known) allow such representation. One such latent method was used in [11, 12, 94] to display high-dimensional student behaviour data in a 2-dimensional representation. This type of visualization helps detecting the characteristics of the data distributions and their grouping or cluster structure.

2.5 Other Data Mining methods applied in e-learning

Not all Data Mining in e-learning concerns advanced AI or ML methods: traditional statistics are also used in [1, 32, 74, 77], as well as Semantic Web technologies [34], ontologies [46], Case-Based Reasoning [33] and/or theoretical modern didactical approaches [6, 7, 41, 96].

Although it could have been included in the section devoted to classification, Naïve Bayes, the model used in [78, 84], also fits in the description of general statistical method. An approach to automate the classification process of Web learning resources was developed in [78]. The model organizes and labels learning resources according to a concept hierarchy extracted from the extended ontology of the ACM Computing Curricula 2001 for Computer Science. In [84], a method to construct personalized courseware was proposed. It consists of the building of a personalized Web tutor tree using the Naïve algorithm, for mining both the context and the structure of the courseware.

Statistical methods were applied in [8, 56, 64]. In [64], the goals were the discovery and extraction of knowledge from an e-learning database to support the analysis of student learning processes, as well as the evaluation of the effectiveness and usability of Web-based courses. Three Web Mining-based evaluation criteria were considered: *session statistics*, *session patterns* and *time series of session data*. In the first, basic statistics about sessions, such as average session, length in time or in number of content requests were gathered. In *session patterns*, the learning processes were extracted from navigation and request behaviour. Finally, in the *time series of session data*, the evolution of session statistics and session patterns over a period of time was analyzed. All methods were applied to Web log entries. In [8], a personalized learning environment applying different symmetric and asymmetric distance measures between the students' profiles and their interests was proposed. In [56], tools for the analysis of student activity were developed to provide decision makers and course developers with an understanding of the e-learners needs. Some statistical analyses of the learner's activities were performed.

An experiment combining a MAS and self-regulation strategies to allow flexible and incremental design, and to provide a more realistic social context for interactions between students and the teachable agent, were presented in [6]. In [41], a model called Learning Response Dynamics that analyzes learning systems through the concepts of learning dynamics, energy, speed, force, and acceleration, was described. In [7, 96], the problems of developing versatile adaptive and intelligent learning systems that could be used in the context of practical Web-based education were discussed. One such system: ELM-ART was developed; it supports learning programming in LISP, and provides adaptive navigation support, course sequencing, individualized diagnosis of student solutions, and example-based problem-solving support.

MAS have also been applied to e-learning beyond classification problems. In [76], one called IDEAL was designed to support student-centred, self-paced, and highly interactive learning. The analysis was carried out on the students' learning-related profile, which includes learning style and background knowledge in selecting, organizing, and presenting the learning material to support active learning. IDEAL supports personalized interaction between the students and the learning system and enables adaptive course delivery of educational contents. The student learning behaviour (student model) is inferred from the performance data using a Bayesian Belief Network model. In [66, 67], a MAS called Cooperative Intelligent Distance Learning Environments (CIDLE) was described. It extracts knowledge from domain knowledge and students' behaviour during a learning discussion. It therefore infers the learners' behaviour and adapts to them the presentation of course material in order to improve their success rate in answering questions. In [51], software agents were proposed as an alternative for data extraction from e-learning environments, in order to organize them in intelligent ways. The approach includes pedagogical agents to monitor and evaluate Web-based learning tools, from the educational intentions point of view.

In [33], a Case-Based Reasoning system was developed to offer navigational guidance to the student. It is based on past user's interaction logs and it includes a model describing learning sessions.

A system that evaluates the students' performance in Web based e-learning was presented in [65]. Its functioning is controlled by an expert system using "neurules": a hybrid concept that integrates symbolic rules and neural computing. Internally, each "neurule" is represented and considered as an Adaline neuron.

Finally, in [17], Social Network Analysis was proposed as a method to evaluate the relationships between communication styles, social networks, and learning performance in a computer-supported collaborative learning (CSCL) community. The students' learning performance was measured by their final grades in the second semester of the CSCL course and was calculated through a combination of final exam score, group assignment evaluation, and peer-evaluation.

3 A survey of Data Mining in e-learning from the e-learning point of view

In this section, we present the surveyed research according to the e-learning problems to which the Data Mining methods are applied.

As mentioned in the introduction, and to avoid unnecessary redundancies, we now present in Tables 1 to 5 a survey of the available literature according to the different e-learning topics addressed in it. All tables include, column-wise, the following information: bibliographic reference, Data Mining problem addressed (DM objective), Data Mining technique used (DM technique), e-learning actors involved, and type of publication: Journal (J), International Conference (C), or Book Chapter (B).

Each of these tables summarizes, in turn, the references on one of the following e-learning subjects:

1. Applications dealing with the assessment of students' learning performance.
2. Applications that provide course adaptation and learning recommendations based on the students' learning behavior.
3. Approaches dealing with the evaluation of learning material and educational web-based courses.
4. Applications that involve feedback to both teachers and students of e-learning courses, based on the students' learning behavior.
5. Developments for the detection of atypical students' learning behavior.

Table 1. Research works that perform students' learning assessment.

Reference	DM objective	DM approach	e-learning actor	Type of publication
[56]	Statistical analysis	Basic statistical methods	Student and Staff	J
[36]	Classification	Fuzzy reasoning	Student	J
[37]	Clustering	Clustering, dynamic programming and fuzzy logic theory	Student and Teacher	J
[14]	Classification	Conceptual maps	Student and teacher	J
[1]	Statistical analysis	Metadata analysis	Student and Teacher	C
[38]	Classification	Concept effect relationship (CER) model	Teacher	J
[74]	Statistical analysis	Basic statistical methods	Student and Teacher	C
[32]	Statistical analysis	Metadata analysis	Student and Teacher	C
[52]	Classification	ID3	Teacher	C
[97]	Classification and visualization	ADT Tree	Student and Teacher	C
[64]	Classification	Basic statistical methods	Teacher	C
[68]	Visualization and clustering	Social Network Analysis	Teacher	C
[17]	Classification	Social Network Analysis	Teacher	J
[87]	Classification	Code generation and mutation.	Teacher	C

[81]	Classification	Neuro-fuzzy model	Teacher	C
[65]	Classification	Expert systems and Neural computing	Teacher	C
[53]	Classification	Combination of: k-NN, MLP and Decision Tree	Teacher	C
[54]	Classification	Contrast rules	Teacher	C
[69, 70]	Classification	Association Rules	Teacher	C; C
[13]	Classification	Association Rules	Teacher	C
[60, 61]	Clustering	SOM	Teacher	J; C
[35]	Classification	Association Rules	Student and Teacher	C
[45]	Classification	Association Rules	Teacher	C
[77]	Statistical analysis	Basic statistical methods	Student, Teacher and Staff	J
[85]	Clustering	Navigation path clustering ad hoc algorithm	Teacher	C
[48]	Classification	Decision tree-based rule extraction	Teacher	C
[59]	Prediction	Decision tree	Teacher	C
[3]	Prediction	Bayesian Network	Teacher	B
[44]	Classification and Prediction	Naïve Bayes, kNN, MLP-ANN, C4.5, Logistic Regression and SVM	Teacher	J
[5]	Prediction	Linear regression	Teacher	C
[27]	Prediction	Regression	Teacher	C
[57]	Visualization	SQL queries	Teacher	C
[58]	Visualization	SQL queries	Teacher	C
[62]	Prediction	FIR	Teacher	C
[25]	Prediction	FIR and OSRE	Teacher	C
[82]	Clustering	EM algorithm	Teacher	C
[15]	Prediction	Dynamic Bayes Net	Teacher	C

Although an important deal of research effort has been devoted to improve the students' e-learning experience (see Tables 2 and, partially, 4), even more has focused assisting online tutors' tasks, including the analysis and assessment of the students' performance and the evaluation of course materials (see Tables 1, 3 and 5, as well as, partially, 3.4).

The assessment of students is the e-learning issue most commonly tackled by means of Data Mining methods. This is probably due to the fact that such assessment is closer to the evaluation methods available in the traditional presential education. One of the e-learning topics with the least results obtained in this survey is the analysis of the atypical students' learning behaviour. This is probably due to the inherently difficult problem of successfully establishing when the learning behaviour of a student is atypical or not.

Table 2. Research works that offer course adaptation based on students' learning behaviour.

Reference	DM objective	DM approach	e-learning actor	Type of publication
[29]	Classification	Consistency Queries (CQ) inductive inference machine	Student	C
[42]	Classification	Consistency Queries (CQ) inductive inference machine	Student	C
[93]	Prediction	Software agents	Student	C
[84]	Prediction	Ad hoc naïve algorithm for tutor tree	Student	C
[28]	Classification	Multi-agent systems	Student	C
[9]	Classification	Graph theory	Student	C
[19]	Classification	IF-THEN rules	Student	
[2]	Classification	Multi-agent systems	Student	
[50]	Classification	Apriori algorithm	Student	C
[8]	Classification	Distance measures	Student	C
[95]	Classification	Association Rules	Student	C
[35]	Classification	Association Rules	Student and Teacher	C
[55]	Classification	Neural Network	Student	J
[48]	Classification	Decision Tree-based rule extraction	Teacher	C
[72, 73]	Prediction	Prediction rules	Student	C; J
[33]	Classification	Case-based reasoning	Student	C
[31]	Clustering	HAC, Single-Pass and k-NN	Student	B
[47]	Classification	Rough set theory and decision trees	Student and Teacher	C
[66, 67]	Prediction	Multi-agent systems and ID3	Teacher	C; C
[76]	Prediction	Bayesian Network	Student	J
[88]	Prediction	Bayesian Network	Student	C

Table 3. Data Mining applications providing an evaluation of the learning material.

Reference	DM objective	DM approach	e-learning actor	Type of publication
[98, 99]	Classification	Software agents and Association Rules	Student	C; C
[89]	Classification	Association Rules (integrating Apriori algorithm), fuzzy set theory and inductive learning (AQR algorithm)	Teacher	C
[39]	Group Decision methods	Group decision method, grey system and fuzzy theory	Student, Teacher and Staff	J
[40]	Classification and prediction	Fuzzy rules	Student, Teacher and Staff	C
[78]	Classification	Naïve Bayes	Teacher	C
[64]	Classification	Basic statistical methods	Teacher	C
[21]	Classification	Web usage mining: association and sequence	Teacher	C
[77]	Statistical analysis	Basic statistical methods	Student, Teacher and Staff	J
[83]	Clustering and Visualization	Bisection K-Means	Teacher	C
[23]	Clustering	SOM	Teacher	J

Table 4. Data Mining applications providing feedback to e-learning actors (students, tutors and educational managers).

Reference	DM objective	DM approach	e-learning actor	Type of publication
[98, 99]	Classification	Software agents and Association Rules	Student	C; C
[36]	Classification	Fuzzy reasoning	Student	J
[1]	Statistical analysis	Metadata analysis	Student and Teacher	C
[32]	Statistical analysis	Metadata analysis	Student and Teacher	C
[97]	Classification	ADT Tree	Student and Teacher	C
[35]	Classification	Association Rules	Student and Teacher	C
[18]	Classification	Apriori algorithm	Student	C
[47]	Classification	Rough set theory and decision trees	Student and Teacher	C
[86]	Clustering	EM algorithm	Teacher	C

[3]	Prediction	Bayesian Network	Teacher	B
[5]	Prediction	Linear regression	Teacher	C
[27]	Prediction	Regression	Teacher	C
[25]	Prediction	FIR and OSRE	Teacher	C
[62]	Prediction	FIR	Teacher	C
[11]	Clustering	GTM	Teacher	C
[15]	Prediction	Dynamic Bayes Net	Teacher	C

Table 5. Data Mining applications for the detection of atypical learning behaviours.

Reference	DM objective	DM approach	e-learning actor	Type of publication
[90, 91]	Outliers detection	Bayesian predictive distribution model	Teacher	C; C
[12]	Outliers detection	GTM	Teacher	C
[94]	Outliers detection	GTM	Teacher	C

4 Discussion and opportunity for the use of Data Mining in e-learning systems

In this section, we analyze in some more detail the current state of the research in Data Mining applied to e-learning, highlighting its future perspectives and opportunities, as well as its limitations. On the basis of the research papers surveyed in this chapter, we can roughly characterize the aforementioned opportunities as follows:

4.1 E-learning courseware optimization

The possibility of tracking user behaviour in virtual e-learning environments makes possible the mining of the resulting data bases. This opens new possibilities for the pedagogical and instructional designers who create and organize the learning contents.

In order to improve the content and organization of the resources of virtual courses, Data Mining methods concerned with the evaluation of learning materials, such as those summarized in Table 3, could be used. Classification problems are dominant in this area, although prediction and clustering are also present.

Some of the publications reported in Table 1 could also indirectly be used to improve the course resources. If the students' evaluation was unsatisfactory, it could hint to the fact that the course resources and learning materials are inadequate.

The Data Mining methods applied to evaluate the learning material in an e-learning course, summarized in Table 3, include: Association Rules techniques, Fuzzy theory and clustering techniques, amongst others. We think that a sensible starting point for the development of course material evaluation is the exploration of Web usage models, applying Association Rules to explore the relationships between the usability of the course materials and the students' learning performance, on the basis of the information gathered from the interaction between the user and the learning environment.

4.2 Students' e-learning experience improvement

One of the most important goals in e-learning, and one of its major challenges, is the improvement of the e-learning experience of the students enrolled in a virtual course. As seen in Tables 1, 2 and 4, several publications have addressed self-evaluation, learning strategies recommendation, users' course adaptation based on the student's profile and necessities. Diverse Data Mining models have been applied to these problems, including Association Rules, Fuzzy Theory, Neural Networks, Decision Trees and traditional statistical analysis.

Applying Data Mining (text Mining or Web Mining) techniques to analyze Web logs, in order to discover useful navigation patterns, or deduce hypotheses that can be used to improve web applications, is the main idea behind Web usage mining. Web usage mining can be used for many different purposes and applications such as user profiling and Web page personalization, server performance enhancement, Web site structure improvement, etc. [80].

Clustering and visualization methods could also enhance the e-learning experience, due to the capacity of the former to group similar actors based on their similarities and the ability of the later to describe and explore these groups intuitively. If it was possible to cluster similar student behaviours on the basis of students' interaction with the learning environment, the tutor could provide scalable feedback and learning recommendation to learners.

Combinations of Data Mining methods have demonstrated their potential in web-based environments, such as the combination of multiple classifiers and genetic algorithms described in [53] and the neuro-fuzzy models put forward in [81].

4.3 Support tools for e-learning tutors

The provision of a set of automatic, or semiautomatic, tools for virtual tutors that allowed them to get objective feedback from students' learning behaviour in order to track their learning process, has been an important line of research on Data Mining for e-learning, as can be deduced from the information summarized in tables 1, 4 and 5. Based on the publications surveyed, the experimental tools developed with this goal in mind could be roughly grouped into:

1. Tools to evaluate the students' learning performance (Table 1).
2. Tools that allow performing an evaluation of the learning materials (Table 3).
3. Tools that provide feedback to the tutors based on the students' learning behavior (Tables 4-5).

Diverse Data Mining methods have been applied to assess the students' learning performance, including: Clustering, Decision Trees, Social Network Analysis, Neural Networks, Fuzzy methods and Association Rules. In fact, this is perhaps the e-learning topic with more significant research advances in the field of applications we are surveying.

One of the most difficult and time-consuming activities for teachers in distance education courses is the evaluation process, due to the fact that, in this type of course, the review process is better accomplished through collaborative resources such as e-mail, discussion forums, chats, etc. As a result, this evaluation has usually to be carried out according to a large number of parameters, whose influence in the final mark is not always well defined and/or understood. Therefore, it would be helpful to discover features that are highly relevant for students' evaluation. In this way, it would be possible for teachers to provide feedback to students regarding their learning activities online and in real time. In this sense, GTM [12, 94] with feature relevance determination and FIR [25, 62] methodologies, have been applied.

From the virtual teacher standpoint, valuable information could be obtain from the e-mail or discussion forum resources; however there is still a lack of automated tools with this purpose, probably due to the difficulty of analyzing the learning behaviour from the aforementioned sources. Such tool would entail the use of Text Mining (or Web Mining) techniques. Natural Language Processing (NLP) techniques would be of potential interest to tackle this problem in e-learning, due their ability to automatically extract useful information that would be difficult, or almost impossible to obtain, through other techniques. Unfortunately, NLP techniques have not been applied extensively in e-learning. Some exceptions can be found in [23, 31], where NLP and clustering models were proposed for grouping similar learning materials based on their topics and semantic similarities.

Another almost unexplored research path in Data Mining for e-learning, which, in the authors' opinion, bears a great potential, is that of the application of methods for the explicit analysis of time series. That is despite the fact that much of the information that could be gathered from e-learning systems usage takes precisely this form.

5 Data Mining in e-learning beyond academic publications: systems and research projects

Beyond academic publications, Data Mining methods have been integrated into software platforms implemented in real e-learning systems. A general review of these types of systems: WebCT, Blackboard, TopClass, Ingenium Docent, etc. [20, 92], commonly used in universities and higher education, showed two main types of platforms: The first type takes a course as the building block, while the second takes the organisation as a whole. The former (e.g. WebCT, TopClass) normally does not make a distinction between teacher and author (course-developer). This way, such systems allow the teacher much flexibility but also assume that the teacher will create course materials. The latter (e.g. Ingenium, Docent), have clearly defined and distinct roles. Content can be developed outside the system.

Table 6. E-learning projects in which Data Mining techniques are used.

Project name	DM techniques applied	e-Learning Topic	University or institution	URL of the project
LON-CAPA	k-NN, MLP, Decision Trees, Association Rules, Multiple Classifiers, Genetic Algorithms and K-means	Assessment system and feedback to e-learning actor, Feature selection and clustering of students performance	Michigan State University, USA	www.lon-capa.org/
ATutor	Statistical analysis	Assessment system and student behaviour tracking	University of Toronto, Canada	www.atutor.ca/
LExIKON	Consistency queries (CQ) inductive inference	Course adaptation to the students' navigational behaviour	German Research Center for Artificial Intelligence, Technische Universität Darmstadt, and others, Germany	http://lexikon.dfki.de/
aLFanet	Software Agents, Machine Learning, Association Rules	Course adaptation to the students' navigational behaviour	Universidad Nacional de Educación a Distancia and Open University of the Netherlands. Spain Portugal, Germany and Netherlands	http://alfanet.ia.uned.es/alfanet
AHA!	Prediction Rules	Course adaptation to the students' navigational behaviour	Eindhoven University of Technology and Cordoba University. Netherlands and Spain	http://aha.win.tue.nl
WebCT	Statistical Analysis	Assessment system and student behaviour tracking	WebCT	www.webct.com/
Blackboard	Statistical Analysis	Assessment system and student behaviour tracking	Blackboard	www.blackboard.com/us/index.aspx

All these systems claim to be innovative and stress the importance of content but, unfortunately, they hardly provide any information about which didactical methods and models they implement; it is therefore difficult to assess them. As far as adaptation is an integral part of the systems, it would require extensive customisation. Most of the surveyed systems do support collaborative learning tasks; however they do not allow the use of any specific scenario. They allow collaboration but merely provide the basic tools for its implementation [93].

Several large research projects have dealt with the integration of Data Mining methods in e-learning (see Table 6). The ALFANET project consists of an e-learning platform that provides individuals with interactive, adaptive and personalized learning through the Internet. ALFANET includes a component to provide support to the interpretation and presentation of dynamic adaptive questionnaires and their evaluation at run-time, based on the student preferences and profile. The adaptation component applies ML techniques, Association Rules, and Multi-Agent architectures to provide online real-time recommendations and advice to learners based on previous users' interactions, the course structure, the contents characterization and the questionnaires' results.

The AHA! project was initially developed to support an on-line course to add adaptation to hypermedia courses at the Eindhoven University of Technology. AHA! is currently in its 3.0 version. One of its most important features is the adaptation of the presentation and navigation system of a course on the basis of the level of knowledge of a particular student. AHA! applies specific prediction rules to achieve the adaptation goals.

The LearningOnline Network with a Computer Assisted Personalized Approach (LON-CAPA) is an integrated system for online learning and assessment. It consists of a learning content authoring and management system that allows new and existing content to be shared and re-used within and across institutions; a course management system; and an individualized homework and automatic grading system. In LON-CAPA some Data Mining methods, such as k-NN, MLP Neural Networks, Decision Trees, Association Rules, Combinations of Multiple Classifiers, Genetic Algorithms and K-means, are employed to analyze individual access paths though the material interaction behaviour.

LExIKON is a research and development project with an innovative approach to knowledge extraction from the Internet. The underlying learning mechanisms invoke inductive inference of text patterns as well as inductive inference of elementary formal systems. A specific inductive inference method called consistency queries (CQ) was designed and applied to this purpose.

ATutor is an Open Source Web-based LCMS designed with accessibility and adaptability features. ATutor has also adopted the IMS/SCORM Content Packaging specifications, allowing content developers to create reusable content that can be swapped between different e-learning systems. In ATutor, the tutors can assign partial credit for certain answers and can view grades, by student, and for all students on all tests, even can get reports showing the number of times, the time, date, and the frequency with which each student accessed course content.

WebCT is a commercial e-learning suite providing a Course Management system and an e-learning platform. In WebCT, the tutors can create self-assessments and the system automatically scores multiple choice, matching, calculated, jumbled sentence, fill-in-the-blank, true-false and short answers type questions, and can display instructor-created feedback and links to relevant course material. The tutors can monitor students' activities in the e-learning system and get different reports about the tracking data of their students.

Blackboard is another commercial e-learning suite that allows tutors to create e-learning courses and develop custom learning paths for group or individual students, providing tools that facilitate the interaction, communication and collaboration between all actors. The system provides data analysis for surveys and test item, and the results can be exported for further analysis. The report includes the number of times and dates on which each student accessed course contents, discussion forums and assignments.

6 Conclusions

The pervasiveness of the Internet has enabled online distance education to become far more mainstream than it used to be, and that has happened in a surprisingly short time. E-learning course offerings are now plentiful, and many new e-learning platforms and systems have been developed and implemented with varying degrees of success. These systems generate an exponentially increasing amount of data, and much of this information has the potential to become new knowledge to improve all instances of e-learning. Data Mining processes should enable the extraction of this knowledge.

It is still early days for the integration of Data Mining in e-learning systems and not many real and fully operative implementations are available. Nevertheless, a good deal of academic research in this area has been published over the last few years. From the point of view of the Data Mining problems dealt with in the surveyed works, we have seen that these are dominated by research on classification and clustering. This is somehow unsurprising, given the variety and wide availability of Data Mining methods, techniques and software tools for both of them. From the e-learning problems viewpoint, most work deals with students' learning assessment, learning materials and course evaluation, and course adaptation based on students' learning behaviour.

In this chapter we have presented a general and up-to-date survey on Data Mining application in e-learning, as reported in the academic literature. Although we aimed to make it as complete as possible, we may have failed to find and identify some papers, journals and conferences that should have been included. The authors apologise in advance for any such errors that may have occurred. We hope this chapter becomes useful not only for Data Mining practitioners and e-learning system managers and developers, but also even for members and users, teachers and learners, of the e-learning community at large.

7 Key e-learning resources

In this section, we synthesize, in a self-contained manner, some key resources for the e-learning community. Once again, the information is provided in the form of tables and includes: International journals and conferences specialized on e-learning; main e-learning discussion forums; main e-learning organizations; e-learning repositories; e-learning standards; key e-learning research papers, books and book chapters; and open source e-learning software.

The last years have witnessed the appearance of a rapidly increasing number of scholarly publications either devoted to e-learning or including e-learning within their scope, as well as the organization of specialised conferences in the field. Table 7 summarizes this information.

Table 7. International journals and conferences specialized on e-learning or including it within their scope and main topics. Conference edition corresponds to that held on 2006.

Scientific Journal	International Conference
ACM Journal of Educational Resources in Computing (JERIC), ACM	IASTED International Conference on Web-Based Education (WBE), on its 5th edition
Computers & Education, Elsevier	IEEE International Conference on Advanced Learning Technologies (ICALT), on its 6th edition
Education and Information Technologies, Springer-Verlag	International Conference of the Association for Learning Technology, (ALT-C), on its 13th edition
European Journal of Open and Distance Learning (EURODL), European Distance and e-Learning Network (online only)	International Conference on Artificial Intelligence in Education (International AIED Society), on its 12th edition
E-Learning, Symposium Journals (online only)	International Conference on Computers in Education, (ICCE), on its 14th edition
IEEE Transactions on Education, IEEE Education Society	International Conference on Engineering Education, (ICEE), on its 9th edition
International Journal of Artificial Intelligence in Education, International AIED Society	International Conference on Intelligent Tutoring Systems, (ITS), on its 8th edition
International Journal on e-Learning (IJEL), AACE	International Conference on Interactive Computer Aided Learning, (ICL), on its 9th edition
Journal of Educational Multimedia and Hypermedia (JEMH), AACE	International Conference on Web-based Learning, (ICWL), on its 5th edition
Journal of Information Technology Education (JITE), Informing Science Institute	MERLOT International Conference (MIC), on its 6th edition
Journal of Interactive Learning Research, (JILR), AACE	Society for Information Technology & Teacher Education, (SITE), on its 18th edition
Journal of Online Teaching and Learning (JOLT), MERLOT (online only)	World Conference on Educational Multimedia, Hypermedia and Telecommunications, (ED-MEDIA), on its 18th edition
User Modelling and User-Adapted Interaction (UMUAI), Springer-Verlag	World Conference on E-Learning in Corporate, Government, Health, & Higher Education (E-Learn), on its 11th edition

In Table 8, the main discussion forums concerning e-learning topics are listed, together with their corresponding URLs. Furthermore, many institutions delivering e-learning courses provide discussion forums to improve the interaction between their students and tutors.

Table 8. e-Learning discussion forums.

Name	URL of the forum
eLearning Forum eCommunity	http://elf.collabhost.com/logon.do
eLearning Forum vPortal	http://elearningforum.vportal.net/
The Common Room - eLearning Discussion Forum	http://bbs.odeluce.stir.ac.uk/index.php
Support Insight e-learning discussion forums (numbers 3, 13, 22)	http://www.supportinsight.com/snitz/default.asp
ASTD E-Learning Discussion Board	http://community.astd.org/eve/ubb.x/a/frm/f/6401041
VTU eLearning Center- Discussion Forum	http://forum.vtu.ac.in/index.php

In Table 9, the most important e-learning organizations, societies and interest groups are presented.

Table 9. e-Learning organizations.

Name	URL of the organization
The eLearning Guild	http://www.elearningguild.com
Learning Economics Group	http://www.learningeconomics.org
Greater Arizona eLearning Association	http://www.gazel.org
New England Learning Association	http://www.nelearning.org
International Association for Distance Learning	http://www.iadl.org.uk/associations.htm
Consortium of College Testing Centers	http://www.ncta-testing.org/cctc/
Sloan Consortium	http://www.sloan-c.org/
Masie Center e-Learning Consortium	http://www.masie.com/masie/default.cfm?page=default
IMS Global Learning Consortium	http://www.imsglobal.org/
Association of Learning Technology (ALT)	http://www.alt.ac.uk/
British Learning Association	http://www.british-learning.com/
European Institute for E-Learning (EIfEL)	http://www.eife-l.org/
eLearning Alliance	http://www.elearningalliance.org
eLearning Network	http://www.elearningnetwork.org/
Learning Federation	http://www.learningfederation.org/

In order to fast-track the access to what we consider the most successful experiments applying Data Mining techniques to e-learning problems, and the most interesting published information in the field, Table 10 shortlists some key research papers and, whereas Table 11 lists some main books and book chapters.

Table 10. Key e-learning research papers.

Paper
Alpaslan F.N. and Jain L.C., "Virtual Al Classroom: A Proposal", Proc. 1st International Workshop on Hybrid Intelligent Systems (HIS-2001) in Advances in Soft Computing, 2002, Springer, Germany, pp.485-495.
Jain, L.C., "Knowledge-Based engineering: An Innovative Teaching Approach," Proceedings of the Proceedings of the Eighth Turkish Symposium on Artificial Intelligence and Neural Networks, June 1999, pp. 15-19.
Rowland, J.G. and Jain, L.C., "Artificial Intelligence Languages in Engineering Education," Proceedings of PRCEE, Adelaide, 1992, pp.201-206.
Fasuga, R., Sarmanova, J.: Usage of Artificial Intelligence in Education Process. In: International Conference for Engineering Education & Research, ICEER2005. Tainan, Taiwan (2005).
Kotsiantis, S.B., Pierrakeas, C.J., Pintelas, P.E.: Predicting Students’ Performance in Distance Learning Using Machine Learning Techniques. Applied Artificial Intelligence 18(5) (2004) 411-426.
Margo, H.: Data Mining in the e-Learning Domain. Computers & Education 42(3) (2004) 267-287.
Matsui, T., Okamoto, T.: Knowledge Discovery from Learning History Data and its Effective Use for Learning Process Assessment Under the e-Learning Environment. In: Crawford, C., et al. (eds.): Society for Information Technology and Teacher Education International Conference. (2003) 3141-3144.
Minaei-Bidgoli, B., Punch, W.F.: Using Genetic Algorithms for Data Mining Optimization in an Educational Web-based System. In: Cantu, P.E., et al. (eds.): Genetic and Evolutionary Computation Conference, GECCO 2003. (2003) 2252-2263.
Monk, D.: Using Data Mining for e-Learning Decision Making. The Electronic Journal of e-Learning 3 (2005) 41-54.
Nebot, A., Castro, F., Vellido, A., Mugica, F.: Identification of Fuzzy Models to Predict Students Performance in an e-Learning Environment. In: Uskov, V. (ed.): The Fifth IASTED International Conference on Web-Based Education, WBE 2006. Puerto Vallarta, Mexico (2006) 74-79.
Pahl, C., Donnellan, D.: Data Mining Technology for the Evaluation of Web-based Teaching and Learning Systems. In: World Conference on e-Learning in Corp., Govt., Health., & Higher Education. (2002) 747-752.
Sison, R., Shimura, M.: Student Modelling and Machine Learning. International Journal of Artificial Intelligence in Education 9 (1998) 128-158.
Tang, C., Lau, R.W., Li, Q., Yin, H., Li, T., Kilis, D.: Personalized Courseware Construction Based on Web Data Mining. In: The First international Conference on Web information Systems Engineering, WISE’00. IEEE Computer Society. June 19 - 20, Washington, USA (2000) 204-211.
Vellido, A., Castro, F., Nebot, A., Mugica, F.: Characterization of Atypical Virtual Campus Usage Behavior Through Robust Generative Relevance Analysis. In: Uskov, V. (ed.): The 5th IASTED International Conference on Web-Based Education, WBE 2006. Puerto Vallarta, Mexico (2006) 183-188.
Zaïane, O.R., Luo, J.: Towards Evaluating Learners’ Behavior in a Web-based Distance Learning Environment. In: IEEE International Conference on Advanced Learning Technologies, ICALT’01. August 6-8, Madison, WI (2001) 357-360.

Table 11. Key e-learning books and books chapters (in chronological order).

Books and Book Chapters
Tedman, D. and Jain, L.C., An Introduction to Innovative Teaching and Learning. In: Jain L.C. (Editor) *Innovative Teaching and Learning: Knowledge-based Paradigms*, Springer, pp. 1-30, Chapter 1 (2000).
Horton, W.: Evaluating E-Learning. ASTD E-Learning Series, Pearson (2001).
Jain, L.C., Howlett, R.J., Ichalkaranje, N., and Tonfoni, G.(Editors), Virtual Environments for Teaching and Learning, World Scientific, Singapore (2002).
Bersin, J. The Blended Learning Handbook: Best Practices, Proven Methodologies, and Lessons Learned. John Wiley & Sons (2004).
Ghaoui, C., Jain, M., Bannore, V. and Jain, L.C. (Editors), Knowledge-Based Virtual Education: User-centred Paradigms, Springer-Verlag (2005).
Rosenberg, M.L.: Beyond E-Learning: Approaches and Technologies to Enhance Organizational Knowledge, Learning, and Performance. John Wiley & Sons (2006).
Romero, C., Ventura, S. (Editors): Data Mining in e-Learning. WIT Press (2006).
Hammouda, K., Kamel, M.: Data Mining in e-Learning. In: Pierre, S. (Editor): *e-Learning Networked Environments and Architectures: A Knowledge Processing Perspective*. Springer-Verlag, (2006).

The deployment of an e-learning solution is a usually difficult process. Table 12 lists the best-known information repositories that may help to facilitate this task. In Table 13, we present a number of repositories containing learning objects ready to use in the development of e-learning courses. Learning objects are any digital resource that can be reused for learning or training, and constitute a valuable resource for e-learning course development.

Table 12. e-Learning information repositories.

Name	URL of the document repository
Distributed eLearning Repositories	http://www.markcarey.com/elearning/distributed-elearning-repositories.html
Educational Resource Information Center – ERIC	http://www.eric.ed.gov/
Distance Learning Database	http://icdllit.open.ac.uk/
Database of Research on International Education	http://cunningham.acer.edu.au/dbtw-wpd/textbase/ndrie/ndrie.html
Education-line	http://www.leeds.ac.uk/educol/
Association for the Advancement of Computing in Education - AACE Digital Library	http://www.aace.org/
e-Learning Centre	http://www.e-learningcentre.co.uk
E-Learning Resources	http://www.grayharriman.com
elearnspace	http://www.elearnspace.com
E-Learning Knowledge Base	http://ekb.mwr.biz
ASTD e-Learning Community	http://www.astd.org/astd/Resources/elearning_community/elearning_home.htm

Table 13. Learning objects repositories.

Name	URL of the learning object repository
Multimedia Educational Resource for Online Learning and Teaching – MERLOT	www.merlot.org
Wisconsin Online Resource Center	http://www.wisc-online.com/
FERL: Learning Object Technology	http://ferl.becta.org.uk/display.cfm?page=307
ARIADNE - European Knowledge Pool System	http://www.ariadne-eu.org/
Campus Alberta Repository of Educational Objects – CAREO	http://careo.netera.ca/
Fathom Archive	http://www.fathom.com/
Lydia Global Repository	http://www.lydialearn.com/devwelcomepage.cfm
Scottish Electronic Staff Development Library	http://www.sesdl.scotcit.ac.uk:8082/main.html
Learning Resource Catalogue – LRC	http://www.lrc3.unsw.edu.au
Cooperative Learning Object Exchange - CLOE	http://cloe.on.ca/
Smete Digital Library	http://www.smete.org
Education Network Australia – EdNA	http://www.edna.edu.au/edna/go
MIT OpenCourseWare	http://ocw.mit.edu/index.html
Interactive Dialogue with Educators from Across the State- IDEAS	http://ideas.wisconsin.edu/
National Learning Network: Materials	http://www.nln.ac.uk/Materials
Global Campus	http://www.csulb.edu/~gcampus/

e-Learning systems require standards for their design and deployment. Educative organizations can also save time and resources, as well as guarantee their continuity, by adhering to a reliable and well-known set of standards. Some of them can be found in Table 14.

Table 14. e-Learning standards.

Name	URL of the specifications
ADL SCORM	http://www.adlnet.org/index.cfm?fuseaction=SCORMDown
AICC (CMI Guidelines)	http://www.aicc.org
IEEE LTCS: LOM	http://ltsc.ieee.org/
MELD IT standards for healthcare education	http://meld.medbiq.org/meld_library/standards/index.htm
IMS	http://www.imsglobal.org/
Learning Systems Architecture Lab	http://meld.medbiq.org/primers/e-learning_standards_pasini.htm

An important issue for the development of e-learning environments is the existence and availability of open source software. In Table 15, the most popular, open source learning management systems are presented. Some of the software presented in this table has already been cited in section 5.

Table 15. Open source e-learning software.

Name	URL of the open source software
ATutor	http://www.atutor.ca/
Brihaspati	http://home.iitk.ac.in/~ynsingh/tool/brihaspati.shtml
Claroline	http://www.claroline.net/
COSE	http://www.staffs.ac.uk/COSE/
CourseWork	http://getcoursework.stanford.edu/
Didactor	http://www.didactor.nl/
Docebo LMS	http://www.docebo.org/doceboCms/
Drupal	http://drupal.org/
Fle3 Learning Environment	http://fle3.uiah.fi/
ILIAS	http://www.ilias.de/ios/index-e.html
LAMS	http://lamsfoundation.org/
.LRN	http://www.dotlrn.org/
Mambo	http://www.mamboserver.com/
Manhattan Virtual Classroom	http://manhattan.sourceforge.net/
Moodle	http://moodle.com/
MySource Matrix	http://www.squiz.co.uk/mysource_matrix
OLAT - Online Learning and Training	http://www.olat.org
Open Source Portfolio (OSP) Initiative	http://www.osportfolio.org/
Sakai	http://www.sakaiproject.org/
Wordcircle CMS	http://wordcircle.org/

Acknowledgments

Félix Castro is a research fellow within the PROMEP program of the Mexican Secretary of Public Education. Alfredo Vellido is a research fellow within the Ramón y Cajal program of the Spanish Ministry of Education and Science.

References

1. Abe, H., Hasegawa, S., Ochimizu, K.: A Learning Management System with Navigation Supports. In: The International Conference on Computers in Education, ICCE 2003. Hong Kong (2003) 509-513.
2. Andronico, A., Carbonaro, A., Casadei, G., Colazzo, L., Molinari, A., Ronchetti, M.: Integrating a Multi-agent Recommendation System into a Mobile Learning Management System. In: Workshop on Artificial Intelligence in Mobile System 2003, AIMS 2003. October 12-15, Seattle, USA (2003).

3. Arroyo, I., Murray, T., Woolf, B., Beal, C.: Inferring Unobservable Learning Variables from Students' Help Seeking Behavior. Lecture Notes in Computer Science, Vol. 3220. Springer-Verlag, Berlin Heidelberg New York (2004) 782-784.
4. Baker, M.: The Roles of Models in Artificial Intelligence and Education Research: A Prospective View. International Journal of Artificial Intelligence in Education 11 (2000) 122-143.
5. Beck, J.E., Woolf, B.P.: High-Level Student Modeling with Machine Learning. In: Gauthier, G., et al. (eds.): Intelligent Tutoring Systems, ITS 2000. Lecture Notes in Computer Science, Vol. 1839. Springer-Verlag, Berlin Heidelberg New York (2000) 584-593.
6. Biswas, G., Leelawong, K., Belynne, K., Viswanath, K.: Developing Learning by Teaching Environments that Support Self-Regulated Learning. In: The 7th International Conference on Intelligent Tutoring Systems. Maceió, Brazil (2004).
7. Brusilovsky, P.: Adaptive Hypermedia. User Modelling and User Adapted Interaction, Ten Year Anniversary Issue 11(1-2) (2001) 87-110.
8. Carbonaro, A.: Personalization Mechanisms for Active Learning in a Distance Learning System. In: International Conference on Simulation and Multimedia in Engineering Education, ICSEE'03. Florida, USA (2003).
9. Carchiolo, V., Longheu, A., Malgeri, M., Mangioni, G.: Courses Personalization in an e-Learning Environment. In: The 3rd IEEE International Conference on Advanced Learning Technologies, ICALT'03. July 9-11 (2003) 252-253.
10. Carpenter, G., Grossberg, S., Rosen, D.B.: Fuzzy ART: Fast Stable Learning and Categorization of Analog Patterns by an Adaptive Resonance System. Neural Networks 4 (1991) 759–771.
11. Castro, F., Vellido, A., Nebot, A., Minguillón, J.: Finding Relevant Features to Characterize Student Behavior on an e-Learning System. In: Hamid, R.A. (ed.): Proceedings of the International Conference on Frontiers in Education: Computer Science and Computer Engineering, FECS'05. Las Vegas, USA (2005) 210-216.
12. Castro, F., Vellido, A., Nebot, A., Minguillón, J.: Detecting Atypical Student Behaviour on an e-Learning System. In: VI Congreso Nacional de Informática Educativa, Simposio Nacional de Tecnologías de la Información y las Comunicaciones en la Educación, SINTICE'2005. September 14-16, Granada, Spain (2005) 153-160.
13. Chang, C., Wang, K.: Discover Sequential Patterns of Learning Concepts for Behavioral Diagnosis by Interpreting Web Page Contents. In: Kommers, P., Richards, G. (eds.): World Conference on Educational Multimedia, Hypermedia and Telecommunications. Chesapeake, VA (2001) 251-256.
14. Chang, F.C.I., Hung, L.P., Shih, T.K.: A New Courseware for Quantitative Measurement of Distance Learning Courses. Journal of Information Science and Engineering 19 (2003) 989-1014.
15. Chang, K., Beck, J., Mostow, J., Corbett, A.: A Bayes Net Toolkit for Student Modeling in Intelligent Tutoring Systems. In: Ikeda, M., et al. (eds.): 8th International Conference on Intelligent Tutoring Systems, ITS2006. LNCS Vol. 4053. Springer-Verlag, Berlin Heidelberg New York (2006) 104-113.
16. Chapman, P., Clinton, J., Kerber, R., Khabaza, T., Reinartz, T., Shearer, C., Wirth, R.: CRIPS-DM 1.0 Step by Step Data Mining Guide. CRISP-DM Consortium (2000).
17. Cho, H., Gay, G., Davidson, B., Ingraffea, A.: Social Networks, Communication Styles and Learning Performance in a CSCL Community. Computers & Education, In Press (2006).
18. Chu, K., Chang, M., Hsia, Y.: Designing a Course Recommendation System on Web based on the Students' Course Selection Records. In: World Conference on Educational Multimedia, Hypermedia and Telecommunications (2003) 14-21.
19. Costabile, M.F., De Angeli, A., Roselli, T., Lanzilotti, R., Plantamura, P.: Evaluating the Educational Impact of a Tutoring Hypermedia for Children. Information Technology in Childhood Education Annual (2003) 289-308.

20.Croock, M., Mofers, F., Van Veen, M., Van Rosmalen, P., Brouns, F., Boticario, J., Barrera, C., Santos, O., Ayala, A., Gaudioso, E., Hernández, F., Arana, C., Trueba, I.: State-of-the-Art. ALFanet/IST-2001-33288 Deliverable D12, Open Universiteit Nederland (2002). URL: http://learningnetworks.org/
21.Dos Santos, M.L., Becker, K.: Distance Education: a Web Usage Mining Case Study for the Evaluation of Learning Sites. In: The 3rd IEEE International Conference on Advanced Learning Technologies, ICALT'03. IEEE Computer Society. Athens Greece (2003) 360-361.
22.Dreyfus, S.E., Law, A.M.: The Art and Theory of Dynamic Programming. Academic Press Inc. New York (1977).
23.Drigas, A., Vrettaros, J.: An Intelligent Tool for Building e-Learning Contend-Material Using Natural Language in Digital Libraries. WSEAS Transactions on Information Science and Applications 5(1) (2004) 1197-1205.
24.Etchells, T.A., Lisboa, P.J.G.: Orthogonal Search-based Rule Extraction (OSRE) Method for Trained Neural Networks: A Practical and Efficient Approach. IEEE Transactions on Neural Networks 17(2) (2006) 374-384.
25.Etchells, T.A., Nebot, A., Vellido, A., Lisboa, P.J.G., Mugica, F.: Learning What is Important: Feature Selection and Rule Extraction in a Virtual Course. In: The 14th European Symposium on Artificial Neural Networks, ESANN 2006. Bruges, Belgium (2006) 401-406.
26.Fasuga, R., Sarmanova, J.: Usage of Artificial Intelligence in Education Process. In: International Conference for Engineering Education & Research, ICEER2005. Tainan, Taiwan (2005).
27.Feng, M., Heffernan, N., Koedinger, K.: Looking for Sources of Error in Predicting Student's Knowledge. In: The Twentieth National Conference on Artificial Intelligence by the American Association for Artificial Intelligence, AAAI'05, Workshop on Educational Data Mining. July 9-13, Pittsburgh, Pennsylvania (2005) 54-61.
28.Fernández, C.A., López, J.V., Montero, F., González, P.: Adaptive Interaction Multi-agent Systems in e-Learning/e-Teaching on the Web. In: Third International Conference on Web Engineering, ICWE 2003. July 14-18, Oviedo, Spain (2003) 144-153.
29.Grieser, G., Klaus, P.J., Lange, S.: Consistency Queries in Information Extraction. In: The 13th International Conference on Algorithmic Learning Theory. Lecture Notes in Artificial Intelligence Vol. 2533. Springer-Verlag, Berlin Heidelberg New York (2002) 173-187.
30.Ha, S.H., Bae, S.M., Park, S.C.: Web Mining for Distance Education. In: IEEE International Conference on Management of Innovation and Technology, ICMIT'00. (2000) 715-719.
31.Hammouda, K., Kamel, M.: Data Mining in e-Learning. In: Pierre, S. (ed.): e-Learning Networked Environments and Architectures: A Knowledge Processing Perspective. Springer-Verlag, Berlin Heidelberg New York (2005).
32.Hasegawa, S., Ochimizu, K.: A Learning Management System Based on the Life Cycle Management Model of e-Learning Courseware. In: Goodyear, P., et al. (eds.): The Fifth IEEE International Conference on Advanced Learning Technologies, ICALT'05. Los Angeles, CA (2005) 35-37.
33.Heraud, J., France, L., Mille, A.: Pixed: an ITS that Guides Students with the Help of Learners' Interaction Log. In: International Conference on Intelligent Tutoring Systems, Workshop Analyzing Student Tutor Interaction Logs to Improve Educational Outcomes. Maceio, Brazil (2004) 57-64.
34.Holohan, E., Melia, M., McMullen, D., Pahl, C.: Adaptive e-Learning Content Generation Based on Semantic Web Technology. In: Workshop on Applications of Semantic Web Technologies for e-Learning, SW-EL@ AIED'05. July 18, Amsterdam, The Netherlands (2005).
35.Hsu, H.H., Chen, C.H., Tai, W.P.: Towards Error-Free and Personalized Web-Based Courses. In: The 17th International Conference on Advanced Information Networking and Applications, AINA'03. March 27-29, Xian, China (2003) 99-104.

36. Hwang, G.J.: A Knowledge-Based System as an Intelligent Learning Advisor on Computer Networks. J. Systems, Man, and Cybernetics 2 (1999) 153-158.
37. Hwang, G.J.: A Test-Sheet-Generating Algorithm for Multiple Assessment Requirements. IEEE Transactions on Education 46(3) (2003) 329-337.
38. Hwang, G.J., Hsiao, C.L., Tseng, C.R.: A Computer-Assisted Approach to Diagnosing Student Learning Problems in Science Courses. Journal of Information Science and Engineering 19 (2003) 229-248.
39. Hwang, G.J., Huang, T.C.K., Tseng, C.R.: A Group-Decision Approach for Evaluating Educational Web Sites. Computers & Education 42 (2004) 65-86.
40. Hwang, G.J., Judy, C.R., Wu, C.H., Li, C.M., Hwang, G.H.: Development of an Intelligent Management System for Monitoring Educational Web Servers. In: 10th Pacific Asia Conference on Information Systems, PACIS 2004. (2004) 2334-2340.
41. Hwang-Wu, Y., Chang, C.B., Chen, G.J.: The Relationship of Learning Traits, Motivation and Performance-Learning Response Dynamics. Computers & Education 42 (2004) 267-287.
42. Jantke, K.P., Grieser, G., Lange, S.: Adaptation to the Learners' Needs and Desires by Induction and Negotiation of Hypotheses. In: Auer, M.E., Auer U. (eds.): International Conference on Interactive Computer Aided Learning, ICL 2004. Villach, Austria (2004).
43. Kohonen, T.: Self-Organizing Maps. 3rd edition, Springer-Verlag, Berlin (2000).
44. Kotsiantis, S.B., Pierrakeas, C.J., Pintelas, P.E.: Predicting Students' Performance in Distance Learning Using Machine Learning Techniques. Applied Artificial Intelligence 18(5) (2004) 411-426.
45. Kumar, A.: Rule-Based Adaptive Problem Generation in Programming Tutors and its Evaluation. In: 12th International Conference on Artificial Intelligence in Education. July 18-22, Amsterdam (2005) 36-44.
46. Leidig, T.: L3-Towards an Open Learning Environment. ACM Journal of Educational Resources in Computing, JERIC. ACM Press 1(1) (2001).
47. Liang, A., Ziarco, W., Maguire, B.: The Application of a Distance Learning Algorithm in Web-Based Course Delivery. In: Ziarko, W., Yao, Y. (eds.): Second International Conference on Rough Sets and Current Trends in Computing. Lecture Notes in Computer Science. Springer-Verlag, Berlin Heidelberg New York (2000) 338-345.
48. Licchelli, O., Basile, T.M., Di Mauro, N., Esposito, F.: Machine Learning Approaches for Inducing Student Models. In: 17th International Conference on Innovations in Applied Artificial Intelligence, IEA/AIE 2004. LNAI Vol. 3029. Springer-Verlag, Berlin Heidelberg New York (2004) 935-944.
49. Margo, H.: Data Mining in the e-Learning Domain. Computers & Education 42(3) (2004) 267-287.
50. Markellou, P., Mousourouli, I., Spiros, S., Tsakalidis, A.: Using Semantic Web Mining Technologies for Personalized e-Learning Experiences. In: Uskov, V. (ed.): The Fourth IASTED Conference on Web-based Education, WBE-2005. Grindelwald, Switzerland (2005).
51. Markham, S., Ceddia, J., Sheard, J., Burvill, C., Weir, J., Field, B., et al.: Applying Agent Technology to Evaluation Tasks in e-Learning Environments. In: Proceedings of the Exploring Educational Technologies Conference. Monash University, Melbourne, Australia (2003) 31-37.
52. Matsui, T., Okamoto, T.: Knowledge Discovery from Learning History Data and its Effective Use for Learning Process Assessment Under the e-Learning Environment. In: Crawford, C., et al. (eds.): Society for Information Technology and Teacher Education International Conference. (2003) 3141-3144.
53. Minaei-Bidgoli, B., Punch, W.F.: Using Genetic Algorithms for Data Mining Optimization in an Educational Web-based System. In: Cantu, P.E., et al. (eds.): Genetic and Evolutionary Computation Conference, GECCO 2003. (2003) 2252-2263.

54. Minaei-Bidgoli, B., Tan, P.N., Punch, W.F.: Mining Interesting Contrast Rules for a Web-based Educational System. In: The 2004 International Conference on Machine Learning and Applications, ICMLA'04. Louisville, KY (2004).
55. Mizue, K., Toshio, O.: N3: Neural Network Navigation Support-Knowledge-Navigation in Hyperspace: The Sub-symbolic Approach. Journal of Educational Multimedia and Hypermedia 10(1) (2001) 85-103.
56. Monk, D.: Using Data Mining for e-Learning Decision Making. The Electronic Journal of e-Learning 3 (2005) 41-54.
57. Mostow, J., Beck, J., Cen, H., Cuneo, A., Gouvea, E., Heiner, C.: An Educational Data Mining Tool to Browse Tutor-Student Interactions: Time will tell!. In: Proceedings of the Workshop on Educational Data Mining 2005. (2005) 15-22.
58. Mostow, J., Beck, J.: Some Useful Tactics to Modify, Map and Mine Data from Intelligent Tutors. Natural Language Engineering 12 (2006) 195-208.
59. Muehlenbrock, M.: Automatic Action Analysis in an Interactive Learning Environment. In: 12th International Conference on Artificial Intelligence in Education, AIED 2005. July 18, Amsterdam, The Netherlands (2005) 73-80.
60. Mullier, D.: A Tutorial Supervisor for Automatic Assessment in Educational Systems. International Journal on e-Learning 2(1) (2003) 37-49.
61. Mullier, D., Moore, D., Hobbs, D.: A Neural-Network System for Automatically Assessing Students. In: Kommers, P., Richards, G. (eds.): World Conference on Educational Multimedia, Hypermedia and Telecommunications. (2001) 1366-1371.
62. Nebot, A., Castro, F., Vellido, A., Mugica, F.: Identification of Fuzzy Models to Predict Students Performance in an e-Learning Environment. In: Uskov, V. (ed.): The Fifth IASTED International Conference on Web-Based Education, WBE 2006. Puerto Vallarta, Mexico (2006) 74-79.
63. Novak, J.D., Gowin, B.C.: Concept Mapping for Meaningful Learning. In: International Conference on Learning. Cambridge University Press (1984) 36-37.
64. Pahl, C., Donnellan, D.: Data Mining Technology for the Evaluation of Web-based Teaching and Learning Systems. In: World Conference on e-Learning in Corp., Govt., Health., & Higher Education. (2002) 747-752.
65. Prentzas, J., Hatzilygeroudis, I., Garofalakis, J.: A Web-based Intelligent Tutoring System Using Hybrid Rules as its Representational Basis. In: Cerri, S.A., et al. (eds.): Intelligent Tutoring Systems, ITS 2002. LNCS Vol. 2363. Springer-Verlag, Berlin Heidelberg New York (2002)119-128.
66. Razek, M., Frasson, C., Kaltenbach, M.: A Confidence Agent: Toward More Effective Intelligent Distance Learning Environments. In: International Conference on Machine Learning and Applications, ICMLA'02. Las Vegas, USA (2002) 187-193.
67. Razek, M., Frasson, C., Kaltenbach, M.: A Context-Based Information Agent for Supporting Intelligent Distance learning Environments. In: The Twelfth International World Wide Web Conference, WWW 2003. May 20-24, Budapest, Hungary (2003).
68. Reffay, C., Chanier, T.: How Social Network Analysis Can Help to Measure Cohesion in Collaborative Distance-Learning. In: International Conference on Computer Supported Collaborative Learning. Kluwer Academic Publishers, Bergen (2003) 343-352.
69. Resende, S.D., Pires, V.M.T.: An Ongoing Assessment Model for Distance Learning. In: Hamza, M.H. (ed.): Fifth IASTED International Conference Internet and Multimedia Systems and Applications. Acta Press (2001) 17-21.
70. Resende, S.D., Pires, V.M.T.: Using Data Warehouse and Data Mining Resources for Ongoing Assessment of Distance Learning. In: IEEE International Conference on Advanced Learning Technologies, ICALT 2002. (2002).
71. Romero, C., Ventura, S.: Data Mining in e-Learning. WIT Press (2006).

72. Romero, C., Ventura, S., De Bra, P.: Knowledge Discovery with Genetic Programming for Providing Feedback to Courseware. User Modeling and User-Adapted Interaction 14(5) (2004) 425-464.
73. Romero, C., Ventura, S., De Bra, P., De Castro, C.: Discovering Prediction Rules in AHA! Courses. In: User Modelling Conference. June 2003, Johnstown, Pennsylvania (2003) 35-44.
74. Seki, K., Tsukahara, W., Okamoto, T.: System Development and Practice of e-Learning in Graduate School. In: The Fifth IEEE International Conference on Advanced Learning Technologies, ICALT'05. IEEE Computer Society, Los Angeles, CA (2005) 740-744.
75. Shachter, R.D.: Evaluating Influence Diagrams. Operating Research 34 (1986) 871-882.
76. Shang, Y., Shi, H., Chen, S.S.: An Intelligent Distributed Environment for Active Learning. ACM Journal of Educational Resources in Computing 1(2) (2001).
77. Sheard, J., Ceddia, J., Hurst, G.: Inferring Student Learning Behaviour from Website Interactions: A Usage Analysis. Education and Information Technologies 8(3) (2003) 245-266.
78. Singh, S.P.: Hierarchical Classification of Learning Resources Through Supervised Learning. In: World Conference on e-Learning in Corp., Govt., Health., & Higher Education. (2004) 178-183.
79. Sison, R., Shimura, M.: Student Modelling and Machine Learning. International Journal of Artificial Intelligence in Education 9 (1998) 128-158.
80. Srivastava, J., Cooley, R., Deshpande, M., Tan, P.: Web Usage Mining: Discovery and Applications of Usage Patterns from Web Data. ACM SIGKDD Explorations. 1(2) (2000) 12-23.
81. Stathacopoulou, G.D., Grigoriadou, M.: Neural Network-Based Fuzzy Modeling of the Student in Intelligent Tutoring Systems. In: International Joint Conference on Neural Networks. Washington (1999) 3517-3521.
82. Talavera, L., Gaudioso, E.: Mining Student Data to Characterize Similar Behavior Groups in Unstructured Collaboration Spaces. In: Workshop in Artificial Intelligence in Computer Supported Collaborative Learning in conjuntion with 16th European Conference on Artificial Intelligence, ECAI'2003. Valencia, Spain (2004) 17-22.
83. Tane, J., Schmitz, C., Stumme, G.: Semantic Resource Management for the Web: An e-Learning Application. In: Fieldman, S., Uretsky, M. (eds.): The 13th World Wide Web Conference 2004, WWW2004. ACM Press, New York (2004) 1-10.
84. Tang, C., Lau, R.W., Li, Q., Yin, H., Li, T., Kilis, D.: Personalized Courseware Construction Based on Web Data Mining. In: The First international Conference on Web information Systems Engineering, WISE'00. IEEE Computer Society. June 19 - 20, Washington, USA (2000) 204-211.
85. Tang, T.Y., McCalla, G.: Smart Recommendation for an Evolving e-Learning System: Architecture and Experiment. International Journal on e-Learning 4(1) (2005) 105-129.
86. Teng, C., Lin, C., Cheng, S., Heh, J.: Analyzing User Behavior Distribution on e-Learning Platform with Techniques of Clustering. In: Society for Information Technology and Teacher Education International Conference. (2004) 3052-3058.
87. Traynor, D., Gibson, J.P.: Synthesis and Analysis of Automatic Assessment Methods in CS1. In: The 36th SIGCSE Technical Symposium on Computer Science Education 2005, SIGCSE'05. ACM Press. February 23-27, St. Louis Missouri, USA (2005) 495-499.
88. Trella, M., Carmona, C., Conejo, R.: MEDEA: an Open Service-Based Learning Platform for Developing Intelligent Educational Systems for the Web. In: 12th International Conference on Artificial Intelligence in Education. July 18-22, Amsterdam (2005) 27-34.
89. Tsai, C.J., Tseng, S.S., Lin, C.Y.: A Two-Phase Fuzzy Mining and Learning Algorithm for Adaptive Learning Environment. In: Alexandrov, V.N., et al. (eds.): International Conference on Computational Science, ICCS 2001. LNCS Vol. 2074. Springer-Verlag, Berlin Heidelberg New York (2001) 429-438.

90. Ueno, M.: LMS with Irregular Learning Processes Detection System. In: World Conference on e-Learning in Corp., Govt., Health, & Higher Education. (2003) 2486-2493.
91. Ueno, M.: On-Line Statistical Outlier Detection of Irregular Learning Processes for e-Learning. In: World Conference on Educational Multimedia, Hypermedia and Telecommunications. (2003) 227-234.
92. Van der Klink, M., Boon, J., Rusman, E., Rodrigo, M., Fuentes, C., Arana, C., Barrera, C., Hoke, I., Franco, M.: Initial Market Study, ALFanet/IST-2001-33288 Deliverable D72. Open Universiteit Nederland (2002). URL: http://learningnetworks.org/downloads/alfanet-d72-initialmarket -studies.pdf
93. Van Rosmalen, P., Brouns, F., Tattersall, C., Vogten, H., van Bruggen, J., Sloep, P., Koper, R.: Towards an Open Framework for Adaptive, Agent-Supported e-Learning. International Journal Continuing Engineering Education and Lifelong Learning 15(3–6) (2005) 261-275.
94. Vellido, A., Castro, F., Nebot, A., Mugica, F.: Characterization of Atypical Virtual Campus Usage Behavior Through Robust Generative Relevance Analysis. In: Uskov, V. (ed.): The 5th IASTED International Conference on Web-Based Education, WBE 2006. Puerto Vallarta, Mexico (2006) 183-188.
95. Wang, D., Bao, Y., Yu, G., Wang, G.: Using Page Classification and Association Rule Mining for Personalized Recommendation in Distance Learning. In: Fong, J., et al. (eds.): International Conference on Web Based Learning, ICWL 2002. LNCS Vol. 2436. Springer-Verlag, Berlin Heidelberg New York (2002) 363-374.
96. Weber, G., Brusilovsky, P.: ELM-ART: An Adaptive Versatile System for Web-based Instruction. International Journal of Artificial Intelligence in Education 12(4) (2001) 352-384.
97. Yoo, J., Yoo, S., Lance, C., Hankins, J.: Student Progress Monitoring Tool Using Treeview. In: The 37th Technical Symposium on Computer Science Education, SIGCSE'06. ACM Press. March 1-5, Houston, USA (2006) 373-377.
98. Zaïane, O.R.: Building a Recommender Agent for e-Learning Systems. In: The International Conference on Computers in Education, ICCE'02. (2002) 55-59.
99. Zaïane, O.R., Luo, J.: Towards Evaluating Learners' Behavior in a Web-based Distance Learning Environment. In: IEEE International Conference on Advanced Learning Technologies, ICALT'01. August 6-8, Madison, WI (2001) 357-360.