

A Soft Computing Decision Support Framework for e-Learning

by

Félix A. Castro Espinoza

fcastro@cs.upc.edu

A thesis submitted to the Computer Science Dept. of the Universitat Politècnica de Catalunya – BarcelonaTech in partial fulfilment of the requirement for the degree of Ph.D. in Artificial Intelligence

Artificial Intelligence PhD Program

Thesis by compendium of publications

Advisors:

Àngela Nebot Castells, PhD.

Universitat Politècnica de Catalunya - BarcelonaTech

Computer Science Department

Soft Computing Research Group

IDEAI Research Centre

angela@cs.upc.edu

Francisco Mugica Álvarez, PhD.

Universitat Politècnica de Catalunya - BarcelonaTech

Computer Science Department

Soft Computing Research Group

IDEAI Research Centre

fmugica@cs.upc.edu

This thesis is dedicated to my whole family,
especially my son

Acknowledgments

It is with much appreciation and gratitude that I want to thank to my advisors: Àngela and Paco who always maintained their post by my side from the inception to the completion of this research project. They gave me guidance throughout my research. Without his knowledge, patience, and support this dissertation would have not been possible. The knowledge and friendship I gained from them will definitely influence the rest of my life.

I am grateful to the anonymous referees for taking time to serve on the reviewing committee and overseeing this work. I deeply appreciate his support and valuable suggestions.

Abstract

Supported by technological development and its impact on everyday activities, e-Learning and b-Learning (Blended Learning) have experienced rapid growth mainly in higher education and training. Its inherent ability to break both physical and cultural distances, to disseminate knowledge and decrease the costs of the teaching-learning process allows it to reach anywhere and anyone. The educational community is divided as to its role in the future. It is believed that by 2019 half of the world's higher education courses will be delivered through e-Learning. While supporters say that this will be the educational mode of the future, its detractors point out that it is a fashion, that there are huge rates of abandonment and that their massification and potential low quality, will cause its fall, assigning it a major role of accompanying traditional education. There are, however, two interrelated features where there seems to be consensus. On the one hand, the enormous amount of information and evidence that Learning Management Systems (LMS) generate during the e-Learning process and which is the basis of the part of the process that can be automated. In contrast, there is the fundamental role of e-tutors and e-trainers who are guarantors of educational quality. These are continually overwhelmed by the need to provide timely and effective feedback to students, manage endless particular situations and casuistics that require decision making and process stored information. In this sense, the tools that e-Learning platforms currently provide to obtain reports and a certain level of follow-up are not sufficient or too adequate. It is in this point of convergence Information-Trainer, where the current developments of the LMS are centered and it is here where the proposed thesis tries to innovate.

This research proposes and develops a platform focused on decision support in e-Learning environments. Using soft computing and data mining techniques, it extracts knowledge from the data produced and stored by e-Learning systems, allowing the classification, analysis and generalization of the extracted knowledge. It includes tools to identify models of students' learning behavior and, from them, predict their future performance and enable trainers to provide adequate feedback. Likewise, students can self-assess, avoid those ineffective behavior patterns, and obtain real clues about how to improve their performance in the course, through appropriate routes and strategies based on the behavioral model of successful students. The methodological basis of the mentioned functionalities is the Fuzzy Inductive Reasoning (FIR), which is particularly useful in the modeling of dynamic systems. During the development of the research, the FIR methodology has been improved and empowered by the inclusion of several algorithms. First, an algorithm called CR-FIR, which allows determining the Causal Relevance that have the variables involved in the modeling of learning and assessment of students. In the present thesis, CR-FIR has been tested on a comprehensive set of classical test data, as well as real data sets, belonging to different areas of knowledge. Secondly, the detection of atypical behaviors in virtual campuses was approached using

the Generative Topographic Mapping (GTM) methodology, which is a probabilistic alternative to the well-known Self-Organizing Maps. GTM was used simultaneously for clustering, visualization and detection of atypical data.

The core of the platform has been the development of an algorithm for extracting linguistic rules in a language understandable to educational experts, which helps them to obtain patterns of student learning behavior. In order to achieve this functionality, the LR-FIR algorithm (Extraction of Linguistic Rules in FIR) was designed and developed as an extension of FIR that allows both to characterize general behavior and to identify interesting patterns. In the case of the application of the platform to several real e-Learning courses, the results obtained demonstrate its feasibility and originality. The teachers' perception about the usability of the tool is very good, and they consider that it could be a valuable resource to mitigate the time requirements of the trainer that the e-Learning courses demand. The identification of student behavior models and prediction processes have been validated as to their usefulness by expert trainers. LR-FIR has been applied and evaluated in a wide set of real problems, not all of them in the educational field, obtaining good results. The structure of the platform makes it possible to assume that its use is potentially valuable in those domains where knowledge management plays a preponderant role, or where decision-making processes are a key element, e.g. e-business, e-marketing, customer management, to mention just a few. The Soft Computing tools used and developed in this research: FIR, CR-FIR, LR-FIR and GTM, have been applied successfully in other real domains, such as music, medicine, weather behaviors, etc.

Resumen

Soportado por el desarrollo tecnológico y su impacto en las diferentes actividades cotidianas, el e-Learning (o aprendizaje electrónico) y el b-Learning (Blended Learning o aprendizaje mixto), han experimentado un crecimiento vertiginoso principalmente en la educación superior y la capacitación. Su habilidad inherente para romper distancias tanto físicas como culturales, para diseminar conocimiento y disminuir los costes del proceso enseñanza aprendizaje le permite llegar a cualquier sitio y a cualquier persona. La comunidad educativa se encuentra dividida en cuanto a su papel en el futuro. Se cree que para el año 2019 la mitad de los cursos de educación superior del mundo se impartirá a través del e-Learning. Mientras que los partidarios aseguran que ésta será la modalidad educativa del futuro, sus detractores señalan que es una moda, que hay enormes índices de abandono y que su masificación y potencial baja calidad, provocará su caída, reservándole un importante papel de acompañamiento a la educación tradicional. Hay, sin embargo, dos características interrelacionadas donde parece haber consenso. Por un lado, la enorme generación de información y evidencias que los sistemas de gestión del aprendizaje o LMS (Learning Management System) generan durante el proceso educativo electrónico y que son la base de la parte del proceso que se puede automatizar. En contraste, está el papel fundamental de los e-tutores y e-formadores que son los garantes de la calidad educativa. Éstos se ven continuamente desbordados por la necesidad de proporcionar retroalimentación oportuna y eficaz a los alumnos, gestionar un sin fin de situaciones particulares y casuísticas que requieren toma de decisiones y procesar la información almacenada. En este sentido, las herramientas que las plataformas de e-Learning proporcionan actualmente para obtener reportes y cierto nivel de seguimiento no son suficientes ni demasiado adecuadas. Es en este punto de convergencia Información-Formador, donde están centrados los actuales desarrollos de los LMS y es aquí donde la tesis que se propone pretende innovar.

La presente investigación propone y desarrolla una plataforma enfocada al apoyo en la toma de decisiones en ambientes e-Learning. Utilizando técnicas de Soft Computing y de minería de datos, extrae conocimiento de los datos producidos y almacenados por los sistemas e-Learning permitiendo clasificar, analizar y generalizar el conocimiento extraído. Incluye herramientas para identificar modelos del comportamiento de aprendizaje de los estudiantes y, a partir de ellos, predecir su desempeño futuro y permitir a los formadores proporcionar una retroalimentación adecuada. Así mismo, los estudiantes pueden autoevaluarse, evitar aquellos patrones de comportamiento poco efectivos y obtener pistas reales acerca de cómo mejorar su desempeño en el curso, mediante rutas y estrategias adecuadas a partir del modelo de comportamiento de los estudiantes exitosos. La base metodológica de las funcionalidades mencionadas es el

Razonamiento Inductivo Difuso (FIR, por sus siglas en inglés), que es particularmente útil en el modelado de sistemas dinámicos. Durante el desarrollo de la investigación, la metodología FIR ha sido mejorada y potenciada mediante la inclusión de varios algoritmos. En primer lugar un algoritmo denominado CR-FIR, que permite determinar la Relevancia Causal que tienen las variables involucradas en el modelado del aprendizaje y la evaluación de los estudiantes. En la presente tesis, CR-FIR se ha probado en un conjunto amplio de datos de prueba clásicos, así como conjuntos de datos reales, pertenecientes a diferentes áreas de conocimiento. En segundo lugar, la detección de comportamientos atípicos en campus virtuales se abordó mediante el enfoque de Mapeo Topográfico Generativo (GTM), que es una alternativa probabilística a los bien conocidos Mapas Auto-organizativos. GTM se utilizó simultáneamente para agrupamiento, visualización y detección de datos atípicos.

La parte medular de la plataforma ha sido el desarrollo de un algoritmo de extracción de reglas lingüísticas en un lenguaje entendible para los expertos educativos, que les ayude a obtener los patrones del comportamiento de aprendizaje de los estudiantes. Para lograr dicha funcionalidad, se diseñó y desarrolló el algoritmo LR-FIR, (extracción de Reglas Lingüísticas en FIR, por sus siglas en inglés) como una extensión de FIR que permite tanto caracterizar el comportamiento general, como identificar patrones interesantes. En el caso de la aplicación de la plataforma a varios cursos e-Learning reales, los resultados obtenidos demuestran su factibilidad y originalidad. La percepción de los profesores acerca de la usabilidad de la herramienta es muy buena, y consideran que podría ser un valioso recurso para mitigar los requerimientos de tiempo del formador que los cursos e-Learning exigen. La identificación de los modelos de comportamiento de los estudiantes y los procesos de predicción han sido validados en cuanto a su utilidad por los formadores expertos. LR-FIR se ha aplicado y evaluado en un amplio conjunto de problemas reales, no todos ellos del ámbito educativo, obteniendo buenos resultados. La estructura de la plataforma permite suponer que su utilización es potencialmente valiosa en aquellos dominios donde la administración del conocimiento juegue un papel preponderante, o donde los procesos de toma de decisiones sean una pieza clave, por ejemplo, e-business, e-marketing, administración de clientes, por mencionar sólo algunos. Las herramientas de Soft Computing utilizadas y desarrolladas en esta investigación: FIR, CR-FIR, LR-FIR y GTM, ha sido aplicadas con éxito en otros dominios reales, como música, medicina, comportamientos climáticos, etc.

Contents

Glossary of Terms.....	10
Chapter 1: Introduction.....	12
1.1 Motivation of the thesis	12
1.2 Methodology selection.....	13
1.3 Main goals of the thesis	15
1.3.1 Machine learning and soft computing goals	16
1.3.2 e-Learning goals	16
1.4 Contributions of the thesis and derived articles	16
1.5 List of articles derived from the thesis with their quality indexes	22
1.6 Organization of the document	26
Chapter 2: State of the Art	28
2.1 Data Mining in e-Learning from the Data Mining point of view.....	29
2.1.1 Where does Data Mining fit in e-Learning processes?	29
2.1.2 The classification problem in e-Learning	31
2.1.2.1 Fuzzy logic methods.....	31
2.1.2.2 Artificial Neural Networks and Evolutionary Computation	32
2.1.2.3 Graphs and Trees.....	34
2.1.2.4 Multi-agent systems	35
2.1.3 The clustering problem in e-Learning.....	36
2.1.4 The association rules mining problem in e-Learning.....	36
2.1.5 Other Data Mining problems in e-Learning	40
2.1.5.1 Prediction techniques	40
2.1.5.2 Visualization techniques	42
2.1.6 Other Data Mining methods applied in e-Learning	43
2.2 Data Mining in e-Learning from the e-Learning point of view	46
2.3 Discussion and opportunity for the use of Data Mining in e-Learning systems	50
2.3.1 e-Learning courseware optimization	53
2.3.2 Students' e-Learning experience improvement	53

2.3.3 Support tools for e-Learning tutors	54
2.4 Data Mining in e-Learning beyond academic publications: systems and research projects.....	55
2.5 Key e-Learning resources	55
2.6 Conclusions of the state of the art	63
Chapter 3: The Fuzzy Inductive Reasoning Methodology: FIR	64
3.1 Fuzzification.....	67
3.2 Qualitative modeling	69
3.3 Qualitative simulation	73
3.4 Defuzzification	76
Chapter 4: Conclusions and Future Research.....	77
4.1 Summary of results obtained	78
4.2 Future research	81
Chapter 5: Bibliography and References.....	84
Chapter 6: Copy of the Articles Derived from the Doctoral Thesis	95
6.1 Articles derived from <u>Objectives 1 and 6</u> : Perform a detailed review of the data mining techniques that have been applied to e-Learning problems and an analysis of the e-Learning problems to which data mining techniques have been applied.	96
6.2 Articles derived from <u>Objective 2</u> : Demonstrate the feasibility of FIR methodology for modeling and prediction of students' performance.	97
6.3 Articles derived from <u>Objective 3</u> : Design and development of an algorithm, in the context of FIR methodology, to provide a quantitative method for assessing the relative causal relevance of individual data features involved in the inferred system model for reducing the uncertainty during the forecasting stage and data understanding purposes.	98
6.4 Articles derived from <u>Objective 4</u> : Design and development of a rule extraction algorithm, in the context of FIR methodology, for data mining and knowledge discovery that enhances the system characterization and facilitates decision making.	99
6.5 Articles derived from <u>Objective 5</u> : Study of the use of Generative Topographic Mapping for the analysis of atypical student behavior.	100
6.6 Articles derived from <u>Objective 7</u> : Design and development of a framework to provide real time useful knowledge to e-Learning environments and improve the e-Learning experience.	101

Glossary of Terms

Classification: Is the task of assigning items in a collection to target categories or classes. The main objective in classification is to construct a model able to predict the target class for each item in the dataset. To accomplish the prediction it is necessary to build a model based on a training dataset, and the classification algorithm finds relationships between the values of the predictors and the values of the target. Classification models segment data by assigning it to previously-defined classes, which are specified in a target.

Clustering: Is the task of grouping a set of objects in such a way that objects in the same group (called *clusters*) are more similar to each other than objects from different clusters. By using clustering models we aim to discover and model the natural grouping structure of data. Clustering is mostly dealt with unsupervised learning techniques, where the class label information is not considered in the process of group definition.

Data Mining: Is a computational process of extracting useful knowledge from large dataset by combining statistical, machine learning, and pattern recognition techniques. The main goal is to discover and reveal interesting correlations, frequent patterns, associations, trends and anomalies, significant and/or causal structures that are embedded within a large and complex dataset.

Decision Support Systems (DSS): Is a computer-based application that collects, organizes and analyses data to facilitate quality decision-making for management, operations and planning and help people make decisions about problems that may be rapidly changing and not easily specified in advance. A DSS in e-Learning is a system that analyse the data generated during the interaction of the teaching-learning process and provide valuable information that support the decision-making activities to accomplish the e-Learning objectives.

Educational Data Mining (EDM): Is the application of data mining, machine learning and statistics techniques to data generated from educational environments. EDM uses mainstream data mining approaches to find and reveal embedded knowledge in educational data for educational purposes. The main concern of EDM is to develop methods that can explore the interesting information extracted from educational datasets, and employ those methods to better understand the students, and their learning behaviors.

e-Learning: Is a learning system based on formalised teaching but with the help of electronic resources, including the delivery of content via electronic media such as internet/intranet/extranet, audio or video, satellite broadcast, interactive TV, CD-ROM, and so on.

Learning Analytics (LA): Is the use of intelligent data analysis for collection, analysis, measurement and reporting of information about learners and their contexts, for purposes of discovering and understanding learning information and social connections to improve students' learning and teaching-learning environments. LA offers a method of gathering information on how students are interacting with learning resources, each other and their teachers.

Learning Management Systems (LMS): Is a software application or Web-based technology used to implement and evaluate a specific learning process. A LMS provides to e-teacher with a way to create and deliver content, monitor student participation, and assess their performance.

Machine Learning (ML): Is a field of computer science devoted in giving computers the ability to automatically learn and improve from experience without being explicitly programmed. ML focuses on the development of computer programs that can access data and use it learn for themselves.

Prediction: Is the task of making a statement about an uncertain event. Prediction involves the development of models that can accurately predict the outcome within the training data and perform well on independent test data. Prediction tries to explicitly model the relationship between inputs or independent variables and the outputs or dependent variables.

Soft Computing (SC): Is a set of computational techniques that combines a collection of methods for dealing with ambiguous situations like imprecision and uncertainty in order to exploit the tolerance for partial truth to achieve tractability, robustness and low cost solution. SC is often applied for defining approaches when the problem is not clear and its solution is unpredictable. SC techniques include mainly: fuzzy systems, neural networks, evolutionary computation and probabilistic reasoning.

Chapter 1: Introduction

This chapter provides the main motivations of this thesis, the reasons why the Fuzzy Inductive Reasoning (FIR) has been chosen as the central methodology to perform this research and the main goals and contributions of the thesis, grouped in machine learning/soft computing goals and e-Learning goals.

1.1 Motivation of the thesis

In the last two decades 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, 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. Internet takes this process of delocalization of the educative experience to a new realm, where the lack of face-to-face 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 face-to-face and distance learning teachers and learners.

This is e-Learning (also referred to as web-based education or 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.

e-Learning involves the use of electronic devices for learning, including the delivery of content via electronic media such as internet/intranet/extranet, audio or video, satellite broadcast, interactive TV, CD-ROM, and so on. e-Learning gives several advantages to students: cost effectiveness, timely content, and access flexibility (Lorenzetti, 2005; Zhang et al., 2010).

Over the last years, e-Learning systems, such as virtual campus environments, have gradually established themselves as a plausible alternative to, and a complement of, traditional distance education models. Initially, e-Learning was presented as the best solution to cover the needs and requirements of remote students, but also a helping tool in the teaching-learning process, reinforcing or replacing face-to-face education. However, usually a huge amount of time is required in the process of providing feedback to the virtual learners, resulting in an increasing demand of teachers and, therefore, of the educative costs, that not always can be fulfilled (Vellido et al., 2011; Abu-Naser et al., 2011; Zorrilla et al., 2010).

This is an important drawback that is addressed in this thesis with the goal of alleviating teachers' workload by providing a suitable way to assess the importance of each evaluation parameter in the learning process.

Furthermore, one of the most difficult and time consuming activities for teachers in distance education courses is the evaluation process, due to the fact that the reviewing process in this kind of courses is better accomplished through collaborative resources such as e-mail, discussion forums, chats, etc. As a result, this evaluation usually has to be done according to a large number of factors, whose influence in the final mark is not always well defined and/or understood. Moreover, it is hard to evaluate the structure of the course content and its effectiveness on the learning process as it develops (Vellido et al., 2011).

To deal with this limitation, it would be helpful to discover features that are highly relevant for the evaluation of groups (clusters) of students. In this way, it would be possible for teachers to provide feedback to students, on their learning activity, online and in real time in a more effective way than if we tried to do it individually. This option, using clustering methods, could reduce teachers' workload. This is also a goal of this thesis, which has been addressed by means of the Generative Topographic Mapping.

Therefore, on the one hand, it would be helpful to reduce the intrinsic system evaluation dimensionality, to identify factors that are highly relevant for the students' evaluation and that represent their learning performance, as it would become feasible for teachers to provide feedback to the students regarding their learning activities, online and in real time. On the other hand, it would be helpful to obtain objective feedback from learners in order to track the learning process and assess the online course structure effectiveness.

As commented before, the use of data mining methods to extract knowledge from the e-Learning system available information can be an adequate approach to follow, in order to use the obtained knowledge to fit the educational proposal to the students' needs and requirements. In fact, as stated by Zaïane, one of the most desired objectives in EDM is to provide an educative recommender system (Zaïane, 2002). This kind of systems suggests, in an intelligent way, actions or activities to students based on previous

decisions of others students with similar academic, demographic or personal characteristics.

Despite the existence of several research concerned with the mining of data generated by the use of e-Learning systems (Romero and Ventura, 2007; Castro et al., 2007; Romero and Ventura, 2010; Desmarais and Baker, 2012; Mohamad and Tasir, 2013; Peña-Ayala, 2014; Cheng et al., 2014; Drachsler et al., 2015; Hegazi and Abugroon, 2016), there are not enough standard methods and techniques to address some problems in e-Learning (Lockyer et al., 2013). This is the case, for instance, of methods that could detect similarities in learning behaviors and group students according to them. This strategy could be used, for instance, to provide teachers with an adaptive guidance tool to prevent student failure. Furthermore, the characterization of the students' online behavior would benefit from a method capable of determining the relevance of the features involved in the analysed data set in terms of this prediction.

This goal is also addressed in this thesis, providing a rule extraction algorithm that describes the analysed system using linguistic rules that are more comprehensive, readable, and which provide explanations (not only assumptions) that may be validated by domain experts, increasing confidence in the analysis. This knowledge could be used for real time student personalization guidance. The idea is to prevent students getting failing grades by forecasting students' performance in real time along the course duration. Also, the algorithm can provide valuable knowledge to the teachers in order to better understand the students' learning behavior patterns, and take into consideration this knowledge in the decision making processes.

This thesis tries to addresses these challenges by means of the development of new approaches to feature selection and rule extraction in the context of the Fuzzy Inductive Reasoning (FIR) soft computing methodology. The proposed algorithms increase FIR robustness and give to the methodology further capabilities to better deal with demanding real world systems. The use of the Generative Topographic Mapping (GTM) unsupervised machine learning algorithm is also studied for the characterization of atypical students' behavior. The diverse palette of data mining problems addressed in this research include data clustering and visualization, outlier management, feature selection, and rule extraction. Strong emphasis is placed on the interpretability of the results obtained through rule extraction, so that they can be feedback to the e-Learning system in a practical an efficient manner.

The methodological developments achieved in this research are included in the e-Learning decision support framework developed in this thesis. This framework aims to deal with the main e-Learning problems mentioned previously.

1.2 Methodology selection

As mentioned in the previous section, the FIR (Fuzzy Inductive Reasoning) methodology is the one chosen as the basis to perform the artificial intelligence research contributions addressed in this thesis due to some of its main advantages:

- The technique can be applied to any system available for testing and observation. The inductive reasoning is entirely based on patterns; therefore, it is not necessary to know the internal structure of the system under study. In this regard, the inductive reasoners are similar to neural networks.
- The methodology contains a model inherent validation mechanism within the simulation method. This mechanism prevents that the model reaches conclusions that cannot be justified on the basis of available data. In this regard, the inductive reasoners are similar to knowledge-based systems.
- Inductive reasoning operates in a qualitative form, as the knowledge-based reasoners do, and is able to provide information on the subset of causal and spatial relationships established between the variables used in the reasoning process, and can provide a justification of the prediction made on the basis of the qualitative states of the input variables selected.
- It has already been proven the great capacity of this methodology for model identification and system's simulation/prediction over dynamic and complex processes in different real world domains.
- FIR has achieved a high capacity to deal with uncertainty, very common in real world problems, especially in complex datasets.

1.3 Main goals of the thesis

Although, as stated in the title, the current doctoral thesis has a strong educational application component, the theoretical/methodological developments are an essential part of it. All the applied research reported in this document has its foundations in novel theoretical developments that involve machine learning techniques and soft computing approaches. Each of such developments, though, is meant to have a practical application in the context of e-Learning. Therefore, the goals of the thesis are twofold: part of them are e-Learning problem-related goals, while another part are machine learning and soft computing related theoretical goals.

The main goal of this thesis is the development of novel algorithms in the area of machine learning and soft computing that can help to improve the e-Learning effectiveness by reducing teacher workload, supporting educative decision making and enhancing students' learning behavior, as well as the integration of these tools in an e-Learning decision support platform.

This main goal of the thesis can be divided in the following sub-objectives:

1.3.1 Machine learning and soft computing goals

1. Perform a detailed review of the data mining techniques that have been applied to e-Learning problems.
2. Demonstrate the feasibility of FIR methodology for modeling and prediction of students' performance.
3. Design and development of an algorithm, in the context of FIR methodology, to provide a quantitative method for assessing the relative causal relevance of individual data features involved in the inferred system model for reducing the uncertainty during the forecasting stage and data understanding purposes.
4. Design and development of a rule extraction algorithm, in the context of FIR methodology, for data mining and knowledge discovery that enhances the system characterization and facilitates decision making.
5. Study of the use of Generative Topographic Mapping for the analysis of atypical student behavior.

1.3.2 e-Learning goals

- 1 Perform an analysis of the e-Learning problems to which data mining techniques have been applied.
- 2 Design and development of a framework to provide real time useful knowledge to e-Learning environments and improve the e-Learning experience.
- 3 Help to prevent students getting failing grades by forecasting students' performance in real time along the course duration.
- 4 Alleviate teachers' workload by providing a suitable way to assess the importance of each evaluation parameter in the learning process.
- 5 Provide valuable knowledge to the teachers in order to better understand the students' learning behavior patterns, and take into consideration this knowledge in the decision making processes.

1.4 Contributions of the thesis and derived articles

In this section, a more detailed description of each goal is provided and the articles derived out of each objective are presented. A complete copy of the post-prints of all the articles can be found in chapter 6.

Objectives 1 and 6: Perform a detailed review of the data mining techniques that have been applied to e-Learning problems and an analysis of the e-Learning problems to which data mining techniques have been applied.

In this goal an up-to-date snapshot of the current state of research and applications of data mining methods in e-Learning is performed. An exhaustive search and analysis of papers, published from 1999 to 2016, which applied data mining approaches to improve the e-Learning environments has been done. The available bibliographic information has been examined according to different criteria, firstly, and from the data mining practitioner point of view, references are studied according to the type of modeling 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 investigated according to the type of data mining problem dealt with: clustering, classification, prediction, etc. A special attention has been given to the clustering educational data problem, leading this work to the following book chapter publication:

A. Vellido, F. Castro, A. Nebot. Clustering Educational Data. Handbook of Educational Data Mining, Chapter 6, pp. 75-92, 2011.

Finally, from the standpoint of the e-Learning practitioner, a taxonomy of e-Learning problems to which data mining techniques have been applied is provided, including, for instance: Students' classification based on their learning performance; detection of irregular learning behaviors; e-Learning system navigation and interaction optimization; clustering according to similar e-Learning system usage; and systems' adaptability to students' requirements and capacities. Part of this work was also published as a book chapter:

F. Castro, A. Vellido, A. Nebot, F. Mugica. Applying Data Mining Techniques to e-Learning Problems: a Survey and State of the Art. Evolution of Teaching and Learning Paradigms in Intelligent Environment. Studies in Computational Intelligence, 62, pp. 183-221, 2007.

An upgrade of the state of the art since these articles were published until now has been done, and the whole state of the art is presented in chapter 2 of this document.

Objective 2: Demonstrate the feasibility of FIR methodology for modeling and prediction of students' performance.

In order to study the feasibility of FIR methodology in the educational environment, a set of experiments where FIR has been applied to model and predict students' performance of different e-Learning real courses are carried out.

The first experiment has been performed using the “Didactic Planning” course of the Centre of Studies in Communication and Educational Technologies virtual campus (CECTE), and the goal is to predict the final mark of the students involved in the course. The CECTE is the part of the international organism known as ILCE (Latin-American Institute of the Educative Communication, in its original Spanish denomination) whose main goal is to offer postgraduate courses. The experiments carried out with the available data indicate that the final mark can be predicted with a low error and that the number of relevant features identified is small, reducing considerably the complexity of the evaluation process and minimizing the teachers’ workload. This work has been published on the *Conference on Web-based Education (IASTED)* and has been selected as finalist of the 5th International Competition of Ph.D. Students on research in web-based education area.

A. Nebot, F. Castro, F. Mugica, A. Vellido. Identification of fuzzy models to predict students performance in an e-Learning environment. Proceedings of the WBE’06: Fifth IASTED International Conference on Web-based Education, pp. 74-79, 2006.

A second set of experiments following the same goals has been performed using the CECTE “Introductory” course. In this case a dynamic assessment of students’ learning performance strategy based on FIR methodology is accomplished. The main goal is to help teachers to give feedback to potential failing students in time, i.e. students can still enhance their performance and pass the course by doing more exercises, studying harder, etc. Each model defines students’ learning behavior at one point in time. In this way it is possible to predict the performance of each student at different moments while course goes through. This work has been published on the *IEEE World Congress on Computational Intelligence*.

A. Nebot, F. Mugica, F. Castro, J. Acosta. Genetic Fuzzy System for Predictive and Decision Support Modelling in e-Learning. Proceedings of the WCCI’10: IEEE World Congress on Computational Intelligence, pp. 1804-1811, 2010.

A. Nebot, F. Mugica, F. Castro. Fuzzy Predictive Models to Help Teachers in e-Learning Courses. Proceedings of the WCCI’10: IEEE World Congress on Computational Intelligence, pp. 2077-2083, 18-23 July 2010.

Objective 3: Design and development of an algorithm, in the context of FIR methodology, to provide a quantitative method for assessing the relative causal relevance of individual data features involved in the inferred system model for reducing the uncertainty during the forecasting stage and data understanding purposes.

In this goal, the concept of causal relevance (CR) is introduced in the context of the FIR modeling and simulation methodology. The idea behind CR is to quantify how much influence each system variable has, from the spatial and temporal points of

view, on the prediction of the output. This research tries to improve the FIR inference engine by means of the CR concept, helping to reduce uncertainty during the forecasting stage. In this work four different CR measures are defined and implemented within the inference engine of the FIR methodology. The first two CR methods compute the relevancy of each feature by means of the quality of the optimal mask, obtained in the qualitative model identification step. The last two CR methods are based on the prediction error of a validation data set, not used in the model identification process.

Applications from different fields are studied in the light of the prediction process, and a comparison between the accuracy of the predictions obtained when using the classical inference engine and the CR option is performed. The results obtained from this research show that FIR predictions are more accurate and precise when the CR option is used, especially for systems where classical FIR forecasting performs rather poorly. This research has been published in the *International Journal of General Systems* and was nominated by members of the journal's Editorial Board community to receive the 2009 Best Paper Award (being considered among the best papers of that year).

A. Nebot, F. Mugica, F. Castro. Causal Relevance to Improve the Prediction Accuracy of Dynamical Systems using Inductive Reasoning, International Journal of General Systems, 38 (3), pp. 331-358, 2009.

The application of CR-FIR to educational data is published in the *Mexican International Conference on Artificial Intelligence*.

F. Castro, A. Nebot, F. Mugica. Causal Relevancy Approaches to Improve the Students' Prediction Performance in an e-Learning Environment. Proceedings of the MICAI'07: 6th Mexican International Conference on Artificial Intelligence, pp.: 342-351, 4-10 2007.

Objective 4: Design and development of a rule extraction algorithm, in the context of FIR methodology, for data mining and knowledge discovery that enhances the system characterization and facilitates decision making.

In this objective, a novel rule-extraction algorithm named LR-FIR (Linguistic Rules in FIR), which is able to extract linguistic rules from a FIR model is designed and developed. Due to the fact that LR-FIR was developed within the FIR methodology, the obtained rules could be considered as predictive rules and deal naturally with the uncertainty captured in the FIR models. The LR-FIR algorithm is developed with the goal to be a useful tool for decision makers. With this purpose in mind the ultimate goal of LR-FIR is to obtain interpretable, realistic and efficient behavioral rules, describing complex systems. LR-FIR foundations have been published in the *Applied Soft Computing* journal.

F. Castro, A. Nebot, F. Mugica. On the extraction of decision support rules from fuzzy predictive models, Applied Soft Computing, 11 (4), pp. 3463–3475, 2011

LR-FIR has been widely experimented in several domains resulting in a number of publications in conferences, book chapters and a journal, with the following goals:

- global warming decision support,

A. Nebot, B. Martínez, F. Castro, C. Gay. Fuzzy Rules Model for Global Warming Decision Support. Proceedings del VI Congreso Internacional de la Asociación Española de Climatología: Cambio Climático Regional y sus Impactos, pp. 565-576, 2008.

- brain tumour diagnosis and UCI benchmarks,

A. Nebot, F. Castro, A. Vellido, M. Julià-Sapé, C. Arús. Rule-based assistance to brain tumour diagnosis using LR-FIR. Lecture Notes in Artificial Intelligence, LNAI-5178, pp. 173-180, 2008.

A. Nebot, F. Castro, A. Vellido, M. Julià-Sapé, C. Arús. Rule-based assistance to brain tumour diagnosis using LR-FIR. Investigating Human Cancer with Computational Intelligence Techniques, Chapter 7, pp. 83-92, 2009.

The first article was published in LNAI as the proceedings of the 12th International Conference, KES 2008 Zagreb, Croatia, September 3-5, 2008. The second one is the same article that was selected from all the conference articles to be published in the book *Investigating Human Cancer with Computational Intelligence Techniques*, KES International.

- e-Learning,

F. Castro, A. Nebot, F. Mugica. Extraction of Logical Rules to Describe Students' Learning Behavior, Proceedings of the WBE'07: 6th IASTED International Conference on Web-based Education, pp. 164-169, 2007.

A. Vellido, F. Castro, T.A. Etchells, A. Nebot, F. Mugica. Data Mining of Virtual Campus Data: a Case Study on Clustering, Visualization, Outlier Detection, Classification, Feature Selection, and Rule Extraction. Evolution of Teaching and Learning Paradigms in Intelligent Environment. Studies in Computational Intelligence, 62, pp. 223-254, 2007.

A. Nebot, F. Mugica, F. Castro, J. Acosta. Genetic Learning of Fuzzy Parameters in Predictive and Decision Support Modelling, International Journal of Computational Intelligence Systems, 5 (2), pp. 387-402, 2012.

The rules extracted by LR-FIR capture the main behavior of each application, from the domain experts' point of view, demonstrating in this sense, the efficiency of the proposed algorithm.

Objective 5: Study of the use of Generative Topographic Mapping for the analysis of atypical student behavior

This objective offers a study of the use of the Generative Topographic Mapping (GTM) approach to the analysis of students that show a non-typical learning behavior.

The GTM model is both a probabilistic alternative to the well-known Self-Organizing Maps (SOM) (also referred to as Kohonen Maps) and a constrained mixture of Gaussians model. The GTM model is capable of detecting atypical usage behavior on the cluster structure of the users of a virtual campus, while neutralizing the negative impact of outliers on the clustering process. This model can simultaneously assess the relative relevance of individual variables on the cluster structure of the users. It simultaneously provides robust data clustering and visualization of the results, which become intuitively interpretable.

The GTM proposed approach has been applied to three e-Learning data sets, the “Didactic planning” and “Introductory” courses of the CECTE and the “Compilers I” subject of the Open University of Catalonia (UOC: Universitat Oberta de Catalunya in its original Catalan denomination). Experiments carried out on the available data indicate that atypical students’ behavior can be identified and interpreted in terms of those variables which are best at explaining and discriminating their different typologies. This research has been published as part of two book chapters and in three international conferences.

A. Vellido, F. Castro, T.A. Etchells, A. Nebot, F. Mugica. Data Mining of Virtual Campus Data: a Case Study on Clustering, Visualization, Outlier Detection, Classification, Feature Selection, and Rule Extraction. Evolution of Teaching and Learning Paradigms in Intelligent Environment. Studies in Computational Intelligence, 62, pp. 223-254, 2007.

The book chapter above is also appears in objective #4 because has contributions to both goals.

F. Castro, A. Vellido, A. Nebot, J. Minguillón. Detección de estudiantes con comportamiento atípico en entornos de aprendizaje e-Learning. TICs Aplicadas a la enseñanza/aprendizaje de la Ingenierías, pp. 23-30, 2006.

A. Vellido, F. Castro, A. Nebot, F. Mugica. Characterization of atypical virtual campus usage behavior through robust generative relevance analysis. Proceedings of the WBE'06: Fifth IASTED International Conference on Web-based Education, pp. 183-188, 2006.

F. Castro, A. Vellido, A. Nebot, J. Minguillón. Detecting atypical student behavior on an e-Learning system. Proceedings of the SINTICE'05: VI Congreso Nacional de Informática Educativa Simposio Nacional de Tecnologías de la Información y las Comunicaciones en la Educación, pp. 153-160, 14-16 September 2005.

F. Castro, A. Vellido, A. Nebot, J. Minguillón. Finding Relevant Features to Characterize Student Behavior on an e-Learning System. Proceedings of the FECS'05: 2005 International

Conference on Frontiers in Education: Computer Science and Computer Engineering, pp. 210-216, 20-23 June 2005.

Objective 7: Design and development of a framework to provide real time useful knowledge to e-Learning environments and improve the e-Learning experience.

With this goal in mind, in this thesis an e-Learning decision support framework (eL-DSF) that provides valuable knowledge to teachers and students has been developed. The eL-DSF presented is mainly based on the FIR methodology and the two key extensions described previously in objectives 3 and 4: CR-FIR and LR-FIR.

The framework offers the following knowledge: 1) gives a sets of rules describing the students' learning behavior; 2) provides a relative assessment of the features involved in the students' evaluation performance, i.e. detects and assess the most important topics involved in the course evaluation process; 3) groups the learning behavior of the students involved in online courses, in an incremental and dynamical way, with the ultimate goal to timely detect failing students, and properly provide them with a suitable and actionable feedback.

In this research the proposed framework is applied to the “Didactic Planning” and the “Introductory” courses of the CECTE. The application shows its usefulness, improving the course understanding and providing valuable knowledge to teachers about the course performance. The eL-DSF and its applications to the CECTE courses have been published in the Simultech and MSE international conferences and a journal paper is under preparation to be submitted in the *User Modeling And User-Adapted Interaction* journal.

A. Nebot, F. Mugica, F. Castro. A Framework to Provide Real Time Useful Knowledge in e-Learning Environments. Proceedings of the SIMULTECH'12: 2nd International Conference on Simulation and Modeling Methodologies, Technologies and Applications, pp. 103-108, 28-31 July 2012.

F. Castro, A. Nebot, F. Mugica. A Soft Computing Decision Support Framework to Improve the e-Learning Experience. Proceedings of the MSE'08: Modeling and Simulation Education – SpringSIM'08: Spring Simulation Multiconference, pp. 781-788, 14-17 April 2008

F. Castro, A. Nebot, F. Mugica. A decision support toolbox for reducing teachers' workload in e-Learning environments. To be presented in the Journal User Modeling and User-Adapted Interaction. Under preparation

1.5 List of articles derived from the thesis with their quality indexes

Publications are presented in four groups: JCR journals, chapters in books, international conferences and national conferences. For each of these groups the articles are listed by year of publication, from most recent to least recent.

JCR Journals

1. A. Nebot, F. Mugica, F. Castro, J. Acosta. Genetic Learning of Fuzzy Parameters in Predictive and Decision Support Modelling, **International Journal of Computational Intelligence Systems**, 5 (2), pp. 387-402, 2012. ISSN: 1875-6883, DOI: 10.2991/10.1080/18756891.2012.685328
Impact factor JCR 2010: **1.471**
Quartile and rank in the category:
Q2 COMPUTER SCIENCE, ARTIFICIAL INTELLIGENCE (49/108);
Q2 COMPUTER SCIENCE, INTERDISCIPLINARY APPLICATIONS (43/97)

Impact factor SCOPUS 2012, Cite Score: **1.16**
Quartile and rank in the category:
Q2 COMPUTER SCIENCE (54/166);
SNIP citations: 3 / GoogleScholar citations: 4
Scopus SNIP (*Source Normalized Impact per Paper*): 0.85
Scopus SJR (*SCImago Journal & Country Rank*): 0.56
2. F. Castro, A. Nebot, F. Mugica. On the extraction of decision support rules from fuzzy predictive models, **Applied Soft Computing**, 11 (4), pp. 3463–3475, 2011. ISSN: 1568-4946, DOI: <http://dx.doi.org/10.1016/j.asoc.2011.01.018>
Impact factor JCR 2011: **2.612**
Quartile and rank in the category:
Q1 COMPUTER SCIENCE, ARTIFICIAL INTELLIGENCE (13/111);
Q1 COMPUTER SCIENCE, INTERDISCIPLINARY APPLICATIONS (16/99)
Impact factor SCOPUS 2011, Cite Score: **4.65**
Quartile and rank in the category:
Q1 COMPUTER SCIENCE, SOFTWARE (13/337);
SNIP citations: 12; GoogleScholar citations: 14
3. A. Nebot, F. Mugica, F. Castro. Causal Relevance to Improve the Prediction Accuracy of Dynamical Systems using Inductive Reasoning, **International Journal of General Systems**, 38 (3), pp. 331-358, 2009. ISSN: 0308-1079, DOI: <http://dx.doi.org/10.1080/03081070802614247>
Impact factor JCR 2009: **0.611**
Quartile and rank in the category:
Q4 COMPUTER SCIENCE, THEORY AND METHODS (75/92);
Impact factor SCOPUS 2009, Cite Score: Scopus no tiene registros anteriores a 2011.
SNIP citations: 3; GoogleScholar citations: 3

This journal has been increasing its JCR level and is currently located in the first quartile (Q1) even though the number of researchers working in the area of general systems theory is small. It should be noted that it is the most prestigious journal in general systems theory.

On the other hand, this article was nominated for the 2009 Best Paper Award. This was notified by the editor with the following letter:

“I am pleased to inform you that your paper was nominated by members of the journal’s Editorial Board community to receive the 2009 Best Paper Award, which is given to the author(s) of the paper that receives the most nominations. Although your paper did not receive enough votes to win this award, we thought that you would be pleased to know that your paper received five (5) and was considered among the best papers of the entire volume. The Editor of the IJGS, Professor Klir, sends his most sincere congratulations and gratitude for this outstanding contribution to the journal. We encourage you to submit papers of this high quality to the International Journal of General Systems for publication in the future.”

Book Chapters

4. A. Vellido, F. Castro, A. Nebot. Clustering Educational Data. ***Handbook of Educational Data Mining***. Editors: Cristobal Romero, Sebastian Ventura, Mykola Pechenizkiy, Ryan Baker, ISBN: 978-1-4398-0457-5, Chapter 6, p.p. 75-92, 2011. GoogleScholar citations: 19
5. A. Nebot, F. Castro, A. Vellido, M. Julia-Sapé, C. Arús. Rule-based assistance to brain tumour diagnosis using LR-FIR. ***Investigating Human Cancer with Computational Intelligence Techniques***. Editorial: Future Technology Press, United Kingdom, ISBN: 978-0-9561516-0-5, Chapter 7, p.p. 83-92, 2009. SNIP citations: 3, GoogleScholar citations: 6
6. A. Nebot, F. Castro, A. Vellido, M. Julià-Sapé, C. Arús. Rule-based assistance to brain tumour diagnosis using LR-FIR. ***Lecture Notes in Artificial Intelligence*** (selected papers of the 12th International Conference on Knowledge-Based and Intelligent Information & Engineering Systems, Invited Sessions: Computational Intelligence in human cancer research). Editors: I. Lovrek, R.J. Howlett, and L.C. Jain / Springer-Verlag Berlin Heidelberg, LNAI- 5178, p.p. 173-180, 2008. SNIP citations: 3, GoogleScholar citations: 6
7. A. Vellido, F. Castro, T.A. Etchells, A. Nebot, F. Mugica. Data Mining of Virtual Campus Data: a Case Study on Clustering, Visualization, Outlier Detection, Classification, Feature Selection, and Rule Extraction. ***Evolution of Teaching and Learning Paradigms in Intelligent Environment. Studies in Computational Intelligence***. Editors: Jain, L.C., Tedman, R. and Tedman, D. (eds.), Springer-Verlag, Germany, ISBN: 978-3-540-71973-1, 62, p.p. 223-254, 2007. SNIP citations: 2, GoogleScholar citations: 2
8. F. Castro, A. Vellido, A. Nebot, F. Mugica. Applying Data Mining Techniques to e-Learning Problems: a Survey and State of the Art. ***Evolution of Teaching and Learning Paradigms in Intelligent Environment. Studies in Computational***

Intelligence. Editors: Jain, L.C., Tedman, R. and Tedman, D. (eds.), Springer-Verlag, Germany, ISBN: 978-3-540-71973-1, 62, p.p. 183-221, 2007. SNIP citations: 64, GoogleScholar citations: 242

9. F. Castro, A. Vellido, A. Nebot, J. Minguillón. Detección de estudiantes con comportamiento atípico en entornos de aprendizaje e-Learning. **TICs Aplicadas a la enseñanza/aprendizaje de la Ingeniería**. Editorial: IEEE, Sociedad de Educación: Capítulos Español, Portugués y Colombiano, ISBN 978-84-8158-375-5, p.p. 23-30, 2006

International Conferences

10. A. Nebot, F. Mugica, F. Castro. A Framework to Provide Real Time Useful Knowledge in e-Learning Environments. *Proceedings of the SIMULTECH'12: 2nd International Conference on Simulation and Modeling Methodologies, Technologies and Applications*, ISBN: 978-989-8565-20-4, pp. 103-108, 28-31 July 2012. GoogleScholar citations: 1
11. A. Nebot, F. Mugica, F. Castro, J. Acosta. Genetic Fuzzy System for Predictive and Decision Support Modelling in e-Learning. *Proceedings of the WCCI'10: IEEE World Congress on Computational Intelligence*, ISBN: 978-1-4244-6910-9, pp. 1804-1811, 18-23 July 2010. GoogleScholar citations: 2. This journal is considered remarkable by the UPC.
12. A. Nebot, F. Mugica, F. Castro. Fuzzy Predictive Models to Help Teachers in e-Learning Courses. *Proceedings of the WCCI'10: IEEE World Congress on Computational Intelligence*, ISBN: 978-1-4244-6910-9, pp. 2077-2083, 18-23 July 2010. SNIP citations: 1, GoogleScholar citations: 7. This journal is considered remarkable by the UPC.
13. F. Castro, A. Nebot, F. Mugica. A Soft Computing Decision Support Framework to Improve the e-Learning Experience. *Proceedings of the MSE'08: Modeling and Simulation Education – SpringSIM'08: Spring Simulation Multiconference*, Editorial: ACM Digital Library, pp. 781-788, 14-17 April 2008. SNIP citations: 1, GoogleScholar citations: 5
14. F. Castro, A. Nebot, F. Mugica. Causal Relevancy Approaches to Improve the Students' Prediction Performance in an e-Learning Environment. *Proceedings of the MICAI'07: 6th Mexican International Conference on Artificial Intelligence*, Editorial: IEEE Computer Society Highlights, p.p.: 342-351, 4-10 November 2007. SNIP citations: 1, GoogleScholar citations: 2
15. F. Castro, A. Nebot, F. Mugica. Extraction of Logical Rules to Describe Students' Learning Behavior, *Proceedings of the WBE'07: 6th IASTED International Conference on Web-based Education*, ISBN: 978-0-88986-651-5, pp. 164-169, 14-16 March 2007. SNIP citations: 6, GoogleScholar citations: 9

16. A. Vellido, F. Castro, A. Nebot, F. Mugica. Characterization of atypical virtual campus usage behavior through robust generative relevance analysis. *Proceedings of the WBE'06: Fifth IASTED International Conference on Web-based Education*, Acta press, pp. 183-188, 23-25 January 2006. SNIP citations: 3, GoogleScholar citations: 5
17. A. Nebot, F. Castro, F. Mugica, A. Vellido. Identification of fuzzy models to predict students performance in an e-Learning environment. *Proceedings of the WBE'06: Fifth IASTED International Conference on Web-based Education*, Acta press, pp. 74-79, 23-25 January 2006. SNIP citations: 13, GoogleScholar citations: 23

It should be noted that this article was selected as a finalist of “the 5th International Competition of Ph.D. Students on Research in Web-Based Education area”. The letter of nomination says: “Only 4 papers-finalists have been selected for this competition from more than 20 submissions. The WBE-2006 IPC members truly believe that papers-finalists demonstrate good theoretical background, innovative ideas and/or approaches, and sounding research findings and results”.

18. F. Castro, A. Vellido, A. Nebot, J. Minguillón. Finding Relevant Features to Characterize Student Behavior on an e-Learning System. *Proceedings of the FECS'05: 2005 International Conference on Frontiers in Education: Computer Science and Computer Engineering*, ISBN: 1-932415-84-X, pp. 210-216, 20-23 June 2005. SNIP citations: 3, GoogleScholar citations: 10. This journal is considered remarkable by the UPC.

National Conferences

19. A. Nebot, B. Martínez, F. Castro, C. Gay. Fuzzy Rules Model for Global Warming Decision Support. *Proceedings del VI Congreso Internacional de la Asociación Española de Climatología: Cambio Climático Regional y sus Impactos*, ISBN: 978-84-612-6051-5, pp. 565-576, 8-11 October 2008.
20. F. Castro, A. Vellido, A. Nebot, J. Minguillón. Detecting atypical student behavior on an e-Learning system. *Proceedings of the SINTICE'05: VI Congreso Nacional de Informática Educativa Simposio Nacional de Tecnologías de la Información y las Comunicaciones en la Educación*, ISBN: 84-9732-437-4, pp. 153-160, 14-16 September 2005. GoogleScholar citations: 32

1.6 Organization of the document

This document is organized as follows. After this introduction, the state of the art is presented in Chapter 2. The state of the art chapter is organized along the axes of data mining techniques and methods applied to e-Learning, and e-Learning applications.

Therefore, we deal with data mining in e-Learning from both, the e-Learning point of view and from the data mining point of view.

Chapter 3 presents the main concepts of the Fuzzy Inductive Reasoning Methodology, FIR. Although this methodology is introduced and explained in several of the published papers, there is no article explaining it as a whole. For this reason we consider that it may be useful, to facilitate the reader's understanding, to add a chapter describing the origins of this methodology and explaining it in depth.

In Chapter 4 the main conclusions of this thesis are discussed and the future research is outlined. Chapter 5 contains the list of references and bibliography used in this research.

Finally, Chapter 6 contains the postprints of all the articles derived from this doctoral thesis grouped by the goals presented in the introduction chapter.

Chapter 2: State of the Art

This chapter tries to provide a complete review of the many applications of Data Mining to e-Learning over the period 1999-2017; 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. Several research efforts in this direction can be found in the literature (Croock et al., 2002; Romero and Ventura, 2006; Romero and Ventura, 2007; Rosmalen et al., 2005; Baker, 2010; Romero and Ventura, 2010; Desmarais and Baker, 2012; Mohamad and Tasir, 2013; Peña-Ayala, 2014; Cheng et al., 2014; Drachsler et al., 2015; Hegazi and Abugroon, 2016).

Over the last years, researchers and developers from the educational community started studying the use of analytic techniques for increasing comprehension into online students' behavior. Two areas under development oriented towards the inclusion and exploration of learning data capabilities are Educational Data Mining (EDM) and Learning Analytics (LA) (Papamitsiou and Economides, 2014). EDM refers to the application of data mining, machine learning and statistical techniques to data generated from educational environments to improve teaching-learning processes. LA means the use of intelligent data analysis for collection, analysis, measurement and reporting of information about learners in order to discovering and understanding learning information and social connections to improve students' learning and, in general, the teaching-learning environments.

In (Peña-Ayala, 2014), the author stated that the evolution of tools and their application is mainly oriented to facilitate the data processing, simplify feature selection stage, provide learning support to learners and enhance the collaboration between students and teacher. However, it is concluded that the majority of these tools are incipient and are in early stages, but growing their functionalities. In (Uskov and Uskov, 2008) the authors, by the analysis of a survey applied to leaders in Web-Based Education (WBE) area, identified a set of strategic issues that would benefit the WBE systems. Although the authors do not mention explicitly the use of Data Mining (DM) to achieve the found issues, these could be reach by using DM methods.

Cheng and his colleagues (Cheng et al., 2014) presented a bibliometric analysis of the e-Learning literature from 2000 to 2012. They identify six groups of papers describing e-Learning topics: continuing education, computer assisted training for professional development, computer assisted occupational health and safety education, computer assisted healthcare and nursing education, social media for informal learning and knowledge management in workplace e-Learning. In (Mohamad and Tasir, 2013) a review about how DM can tackle educational problems is presented and addresses the limitations of existing research. Drachsler and colleagues presented a state of the art on

recommender systems in educational area, specifically in technology enhanced learning (Drachler et al., 2015). The authors organize their findings in seven clusters based on terms of relevant contributions to the field.

An interesting work is (Hegazi and Abugroon, 2016), where an analysis of the state of the art on EDM in higher education is presented. The paper concludes that EDM discipline is start growing and new contributions are needed to be more useful, fully operative and available for all users.

The findings of the state of the art presented in this chapter are organized along two axes: Data Mining problems and methods, and e-Learning applications. Section 2.1 presents the research along the axis of the DM modeling techniques and methods, while section 2.2 presents the surveyed content along the e-Learning applications axis.

In this chapter, a deeper analysis and discussion of the actual state of the research in the field is presented in section 2.3, highlighting its opportunities and limitations. Section 2.4 reports work on DM in e-Learning beyond academic publications. In order to complete the state of the art presented in this chapter we synthesize, in section 2.5, the main key resources for the e-Learning community. Finally, section 2.6 summarizes the findings and draws some conclusions.

Most of the information provided in this chapter takes the form of tables of publications. It is 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.

2.1 Data Mining in e-Learning from the Data Mining point of view

In this section, the surveyed research according to the DM problems (classification, clustering, etc.) and techniques and methods (e.g., Neural Networks, Genetic Algorithms, Decision Trees, or Fuzzy Logic), is presented. 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 DM in the world of e-Learning.

2.1.1 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 DM techniques and methods, and education processes (Baker, 2000; Fasuga and Sarmanova, 2005; Ha et al., 2000; Margo, 2004; Sison and Shimura, 1998; Tang and McCalla, 2005; Velmurugan et al., 2016). In (Baker, 2000), 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 (Margo, 2004; Tang and McCalla, 2005; Barahate and Shelake, 2012; Sharma and Singh, 2013), studies on how DM techniques could successfully be incorporated to e-Learning environments and how they could improve the learning tasks were carried out. In (Tang and McCalla, 2005), data clustering was suggested as a means to promote group-based collaborative learning and to provide incremental student diagnosis. In (Barahate and Shelake, 2012) a review of DM techniques that have been applied in the educational field grouping them by task, is presented. A discussion of successful applications of DM on traditional face-to-face and web-based distance educational systems is introduced in (Sharma and Singh, 2013).

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 (Ha et al., 2000). 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 Machine Learning (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 modeling, were presented in (Sison and Shimura, 1998). In this paper, the difficulty, appropriateness and potential of applying ML techniques to student modeling were commented.

In (Velmurugan et al., 2016) the application of big data technologies and data analysis for capturing student's learning processes, and for supporting the strategic operational decisions at educational institutions, were presented and discussed. The authors stated that big data technologies could be applied for performance prediction, students' attrition risk detection, data visualization, intelligent feedback, course recommendation, student skill estimation and behavior detection; however, the paper approach is only a technological proposal and no experimental results are presented.

2.1.2 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 DM process exist, in what follows, we shall see that only a few techniques (or families of techniques) have been applied to e-Learning.

2.1.2.1 Fuzzy logic methods

Fuzzy logic (FL) based methods have taken their first steps in the e-Learning field at the beginning of this century (Hwang, 1999; Hwang et al., 2004a; Hwang et al., 2004b; Stathacopoulou and Grigoriadou, 1999; Tsai et al., 2001; Jurado et al., 2008; Jia et al., 2010; Yadav and Singh, 2011; Yadav et al., 2014; Jurado et al., 2012; Goel et al., 2012; Jeremic et al., 2012; Chrysafiadi and Virvou, 2013). In (Stathacopoulou and Grigoriadou, 1999), 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 (ANN) were trained to perform 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 (Hwang et al., 2004a). In further work by Hwang and colleagues (Hwang, 1999; Hwang et al., 2004b), 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 (Tsai et al., 2001). 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.

In (Jurado et al., 2008) FL was used to model students' assignments and to update the student model with their preferences. After that, students' evaluation is performed by means of ML techniques. Jia and colleagues applied fuzzy set theory to design an

adaptive learning system to support learners during the memorization processes (Jia et al., 2010). Goel and colleagues (Goel et al., 2012) used FL for students modeling. This model, based on the student system interaction, forecasts the error degree that the student will get when he/she tries to solve a new problem. Fuzzy models are used as well in (Jeremic et al., 2012), in order to model student's knowledge and cognitive characteristics by means of fuzzy rules, which are applied during the learning process in an intelligent tutoring system.

(Jurado et al., 2012) developed a method based on FL for evaluating programming algorithms. The method assess the structural part of an algorithm by means of obtaining a fuzzy ideal representation (algorithm) written by a teacher. The authors claim that the proposed method can be used in the implementation of intelligent tutoring systems for teaching/learning algorithm programming. Thus, the authors have proved that it is possible to use test cases and FL to assess the algorithms written by students and to give advice.

A new method for students' academic performance assessment based on both, FL techniques and expert systems, i.e. Fuzzy Expert System (FES), was introduced in (Yadav and Singh, 2011). An extension of the proposed FES is introduced in (Yadav et al., 2014) where a comparison between classical and new FES is performed. The obtained results have demonstrated the suitability of the new proposed approach.

A literature review on students modeling is presented in (Chrysafiadi and Virvou, 2013). The authors conclude that the use of fuzzy techniques and Bayesian networks have been increased in order to deal with the uncertainty associated to students learning.

2.1.2.2 Artificial Neural Networks and Evolutionary Computation

Some research on the use of Artificial Neural Networks (ANN) and Evolutionary Computation (EC) models to deal with e-Learning topics can be found in (Minaei-Bigdoli and Punch, 2003; Mitzue and Toshio, 2001; Traynor and Gibson, 2005; Kotsiantis et al., 2010; Ogor, 2007; Kardan et al., 2013; Fazlollahtabar and Mahdavi, 2009; Taylan and Karagozoglu, 2009; Lykourantzou et al., 2009; Vázquez-Barreiros et al., 2014).

A navigation support system based on an Artificial Neural Network (more precisely, a Multi-Layer Perceptron, or MLP) was put forward in (Mitzue and Toshio, 2001) 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 ANN and the level of effectiveness of the navigation strategy.

In (Minaei-Bigdoli and Punch, 2003; Traynor and Gibson, 2005), evolutionary algorithms were used to evaluate the students' learning behavior. A combination of multiple classifiers 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 (Minaei-Bigdoli and Punch, 2003). The classification and prediction accuracies are improved through the weighting of the data feature vectors using a Genetic Algorithm. In (Traynor and Gibson, 2005) a random code generation and mutation process suggested as a method to examine the comprehension ability of students can be found.

In (Vázquez-Barreiros et al., 2014) SoftLearn, a process mining platform that is able to discover students' learning paths from the logs of virtual learning environments, is presented. SoftLearn uses a genetic algorithm as a core to find the learning paths. In addition, SoftLearn includes a graphical interface that teachers can analyse to visualize learning paths as activity graphs and to access to the relevant data generated in the learning activities.

In (Kotsiantis et al., 2010) an ensemble of classifiers was introduced in order to improve the precision assessing of students enrolled in e-Learning environments. The proposed approach includes an ensemble that combines three online classifiers: Naive Bayes, 1-NN and WINNOW algorithms, using as a voting methodology. An ensemble of classifiers, that integrates Latent Semantic Analysis (LSA) and n-gram co-occurrence, is also adopted in (He et al., 2009). The proposed ensemble approach is applied to a summary assessment system for automatic grading of English summary writings.

In (Ogor, 2007) a methodology is developed in order to predict the students' final achievement status upon graduation by the derivation of performance prediction indicators mainly focusing on performance monitoring of students' continuous assessment and examination scores. Based on various DM techniques and the application of ML processes (Link Analysis, ANN, k-means, SOM, etc.), rules are derived that enable the classification of students in their predicted classes. Similarly, a combination of ML: Feed-Forward Neural Networks, support vector machines (SVM) and simplified fuzzy ARTMAP architecture, are applied in (Lykourentzou et al., 2009) to perform a dropout prediction method for students involved in e-Learning courses.

An ANN approach is introduced in (Kardan et al., 2013) for modeling the course offering problem. The goal is to identify the potential factors that affect student satisfaction concerning the online courses they select, modeling student course selection behavior and fitting a function to the training data using ANNs, and applying the obtained function to predict the final number registrations in every course after the drop and add period. In (Samigulina and Shayakhmetova, 2015) an information system of distance learning for people with impaired vision is introduced. The proposed system applied a combination of ANN, genetic algorithms, neuro-fuzzy models and FL for processing multidimensional data in real-time.

In (Fazlollahtabar and Mahdavi, 2009) a neuro-fuzzy approach is proposed based on an evolutionary technique to obtain an optimal learning pathway for both instructor and learner enrolled in distance learning courses based on their profile. The neuro-fuzzy

approach allows the diagnostic model to imitate the instructor in inferring students' characteristics, and provides the intelligent learning environment with reasoning capabilities. Similarly, in (Taylan and Karagozoglu, 2009) a neuro-fuzzy approach was developed for prediction of student's academic performance.

2.1.2.3 Graphs and Trees

Graph and/or tree theory was applied to e-Learning in (Carchiolo et al., 2003; Chang and Wang, 2001; Chang et al., 2003; Grieser et al., 2002; Jantke et al., 2004; Liang et al., 2000; Liccheli et al., 2004; Wang et al., 2002; Yoo et al., 2006; Ozpolat and Akar, 2009; Shannaq et al., 2010; Baradwaj and Pal, 2011; Dejaeger et al., 2012; Sen et al., 2012; Parmar et al., 2015; Kolo et al., 2015; Dwivedi and Rawat, 2017).

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 (Carchiolo et al., 2003). Personalized learning paths in the courses were modelled using graph theory. In (Liang et al., 2000; Liccheli et al., 2004), 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 (Liang et al., 2000): 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 behavior patterns was put forward in (Chang and Wang, 2001). 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 (Liccheli et al., 2004), 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 (Yoo et al., 2006), 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 (Chang et al., 2003). It proposes new courseware diagrams, combining tools provided by the theory of conceptual maps (Novak and Gowin, 1984) and influence diagrams (Schachter, 1986). In (Grieser et al., 2002; Jantke et al., 2004; Wang et al.,

2002), 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.

In (Ozpolat and Akar, 2009) an automatic learner model approach, based on personality factors like learning styles, behavioral factors like user's browsing history and knowledge factors like user's prior knowledge, is presented. They use NBTree classification in conjunction with Binary Relevance classifier. At the beginning, the learner interest is collected using generic queries. Then, the learner profile is constructed using a conversion unit based on keyword mapping. The learner model is built by processing the learner profile over a clustering unit and then using a decision unit. A random tree model was applied in (Parmar et al., 2015) for the prediction of students' performance in distributed environment. In (Shannaq et al., 2010) an educational rule generation process based on decision trees is applied to predict the students' loyalty, i.e. the number of enrolled students in educational courses.

A decision tree approach for predicting students' academic performance is introduced in (Kolo et al., 2015). The authors used the SPSS program for building the predictive decision tree; the dataset used for the experiments was obtained from a questionnaire and includes personal and academic data, such as finance level, motivation level, gender and grades obtained in previous courses. A decision tree is used too in (Baradwaj and Pal, 2011), to predict the students' performance at the end of the semester examination. The authors consider that this information can help to identify the dropouts and the students who need special attention and allow the teachers to provide appropriate advising.

The applicability of different DM techniques to identify the main clues of student satisfaction in two business education institutions is presented in (Dejaeger et al., 2012). The introduced approach includes the selection of two different decision tree learners and cumulative logistic regression. A case study for the forecasting of secondary education is developed in (Sen et al., 2012). In this study, four popular classification methods are used, and compared to each other, i.e. ANN, SVM, decision trees and logistic regression.

In (Dwivedi and Rawat, 2017) the authors introduced an approach for the recommendation of the best combination of courses to the students. The proposed architecture is based on the Apriori algorithm in order to provide association rule mining. The association rules discovered are matched with real life choices of courses.

2.1.2.4 Multi-agent systems

Multi Agents Systems (MAS) for classification in e-Learning have been proposed in (Andronico et al., 2003; Fernández et al., 2003; Schiaffino et al., 2008). In (Fernández et al., 2003) 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 (Andronico et al., 2003); 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.

In (Schiaffino et al., 2008) the eTeacher, an intelligent agent that delivers personalized assistance to e-Learning students is introduced. eTeacher detects student's behavior while he/she is on online courses and builds the student's profile. This profile contains the student's learning style and information about his/her performance, such as exercises done, topics studied and exam results.

2.1.3 The clustering problem in e-Learning

Unlike in classification problems, in data grouping or clustering we are not interested in modeling 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 (Hwang, 2003), 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) (Carpenter et al., 1991) to cluster the test items into groups, as well as dynamic programming (Dreyfus and Law, 1977) for test sheet construction.

In (Mullier, 2003; Mullier et al., 2001), an in-depth study describing the usability of ANN and, more specifically, of Kohonen's Self-Organizing Maps (SOM) (Kohonen, 2000) 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 (Tang and McCalla, 2005). Here, data clustering is suggested as a means to promote group-based collaborative learning and to provide incremental student diagnosis.

In (Kato et al., 2016) the authors applied the k-means clustering algorithm in order to examine the relationship between the students' programming behaviors and their programming modes during programming exercises. The algorithm was applied over the chronological records of the compilation and execution of individual students. As a result, the authors have found that there is a correlation between the programming activities and the time needed for solving the problems.

In (Teng et al., 2004), 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 behaviors. Teachers could use these results in order to provide specialized advice to students belonging to each cluster. The simplifying assumption that students belonging to each cluster should share web usage behavior 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 (Talavera and Gaudioso, 2004), where clustering was used to discover user behavior patterns in collaborative activities in e-Learning applications.

Some researchers (Drigas and Vrettaros, 2004; Hammouda and Kamel, 2005; Tane et al., 2004) propose the use of clustering techniques to group similar course materials: An ontology-based tool, within a Web Semantics framework, was implemented in (Tane et al., 2004) 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 (Drigas and Vrettaros, 2004) to devise an intelligent searching tool to cluster similar learning material into classes, based on its semantic similarities. Clustering was proposed in (Hammouda and Kamel, 2005) 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.

A model to optimize the order of presentation of educational contents in the Moodle e-Learning platform is introduced in (Franco-Lugo et al., 2016). The proposed model firstly group students according to a student model, then for each group, an Ant Colony Optimization algorithm is used for organize the sequencing of the educational content that better adapts to its learning characteristics. In (Aher and Lobo, 2013) the authors propose a clustering and association rule mining approach to develop a course recommendation system.

Different variants of the Generative Topographic Mapping (GTM) model, a probabilistic alternative to SOM, were used in (Castro et al., 2005b; Castro et al., 2005a; Vellido et al., 2006) for the clustering and visualization of multivariate data concerning the behavior of the students of a virtual course. More specifically, in (Castro et al., 2005b; Vellido et al., 2006) 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 behaviors. A different variant of GTM for feature relevance determination was used in (Castro et al., 2005a) to rank the available data features according to their relevance for the definition of student clusters.

The application of different fuzzy clustering techniques (FCM and KFCM) to find learners' profiles is introduced in (Hogo, 2010). The paper presents a hybridization approach of artificial intelligence techniques and statistical tools to evaluate students' profiles and adapt the e-Learning systems.

In (Dogan and Camurcu, 2010) the k-means and fuzzy c-means algorithms are used to clustering and visualizing the concept-level scores in order to provide meaningful and nontrivial insights into the workings of a course. Such information is useful for the teacher to discover which concepts are difficult for students and which are not. Similarly, in (Rajibussalim, 2010) a k-means algorithm is used to cluster students based on their similar behavior when using an educational Web-based system. Once the clustering model is available, the main objective is to classify students with respect their final exam mark. The classification task is performed by using the J48 tree-based algorithm. (Maull et al., 2010) introduces a Web-based curriculum-planning tool that is designed to help teachers to review curricular objectives, locate relevant supplementary digital resources, and develop differentiated instructional plans that connect their curricular goals and digital materials with classroom activities and assessments. To accomplish the digital resources organization K-means and E-M clustering algorithms are used.

2.1.4 The association rules mining problem in e-Learning

Association Rules Mining (ARM) for descriptive behavior in data, applied to e-Learning, have mainly been investigated in the areas of learning recommendation systems (Chu et al., 2003; Zaïane, 2002; Zaïane and Luo, 2001; García et al., 2011), learning material organization (Tsai et al., 2001), student learning assessments (Hwang et al., 2003; Kumar, 2005; Matsui and Okamoto, 2003; Minaei-Bigdoli et al., 2004; Resende and Pires, 2001; Resende and Pires, 2002; Buldu and Ucgun, 2010; Weng, 2011), course adaptation to the students' behavior (Costabile et al., 2003; Hsu et al., 2003; Markellou et al., 2005), evaluation of educational web sites (Dos Santos et al., 2003), mining rare association rules (Romero et al., 2010) and subgroup discovery (Carmona et al., 2010; Romero et al., 2009, Herrera et al., 2011).

Data Mining techniques such as Association Rule mining, and inter-session and intra-session frequent pattern mining, were applied in (Zaïane, 2002; Zaïane and Luo, 2001) 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 (Minaei-Bigdoli et al., 2004), 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 (Hwang et al., 2003), 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 (Costabile et al., 2003). 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 (Matsui and Okamoto, 2003) 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 (Dos Santos et al., 2003), applying association and sequence techniques (Srivastava et al., 2000).

In (Markellou et al., 2005), 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 analyses the relationships between student learning time and the corresponding test results was introduced in (Hsu et al., 2003). 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 (Kumar, 2005). In (Chu et al., 2003), 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.

In (García et al., 2011) a collaborative educational data mining tool based on association rule mining and collaborative filtering was introduced, with the goal of making recommendations to instructors about how to improve e-Learning courses, and allowing teachers with similar course profiles to share and score the discovered information. Similarly, Weng introduced a new algorithm based on the Apriori approach to mine fuzzy specific rare itemsets from quantitative data. Then, fuzzy association rules can be generated from these fuzzy specific rare itemsets. The patterns are useful to discover learning problems (Weng, 2011). The Apriori algorithm is also applied in (Buldu and Ucgun, 2010), to characterize the relationship between the students' behavior and the courses enrolled into.

One of the first works related to the extraction process of rare association rules from e-Learning data is the one described in (Romero et al., 2010). Different association rule

mining approaches were studied and compared from the point of view of providing relevant knowledge about non-frequent patterns when gathering student usage data from a Moodle system. In this work the authors explored how some specific algorithms, such as Apriori-Inverse and Apriori-Rare, are better at discovering rare-association rules than other non-specific algorithms, such as Apriori-Frequent and Apriori-Infrequent.

Other interesting research related to e-Learning descriptive tasks are (Carmona et al., 2010) and (Romero et al., 2009), which were focused on the application of subgroup discovery techniques to e-Learning data. The main objective was to obtain knowledge from the usage data and to use it to improve the marks obtained by the students.

A first approach of the subgroup discovery iterative genetic algorithm (SDIGA), based on accuracy, coverage, and significance, applied to this problem was used in (Romero et al., 2009), where a comparison study was presented. In (Carmona et al., 2010), a different version of SDIGA was proposed based on unusualness, support, and confidence. This new approach was compared with classical techniques and evolutionary subgroup discovery algorithms. An interesting and complete overview on subgroup discovery can be found in (Herrera et al., 2011), where the different areas where this methodology has been applied are reviewed.

2.1.5 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 DM problems such as prediction and visualization, which we review in this subsection.

2.1.5.1 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' behavior 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 (Romero et al., 2004; Romero et al., 2003). 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 (Muehlenbrock, 2005). 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 (Beck and Woolf, 2000; Feng et al., 2005; Kotsiantis et al., 2004). In (Feng et al., 2005), a study that aimed to find the sources of error in the prediction of students' knowledge behavior 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 (Beck and Woolf, 2000) to predict whether the student's next response would be correct, and how long he or she would take to generate that response. Logistic regression is used in (Burgos et al., 2017) in order to predict whether or not a student will drop out a course. Based on the resulting prediction model, the authors created an action plan to reduce the dropout rate in e-Learning courses.

In (Kotsiantis et al., 2004), 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, k-NN, MLP Neural Network, C4.5, Logistic Regression and Support Vector Machines (SVM). SVM is also adopted in (Wang et al., 2009) to provide personalized learning resource recommendation in educational environments. 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 (Nebot et al., 2006). The relative relevance of specific features describing course online behavior was also assessed. This work was extended in (Etchells et al., 2006) 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 (Etchells and Lisboa, 2006). Rule extraction was also used in (Romero et al., 2004; Romero et al., 2003) 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 (Trella et al., 2005). 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 (Arroyo et al., 2004) 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. A Dynamic Bayes Net (DBN) was used in

(Chang et al., 2006), for modeling students' knowledge behavior and predict future performance in an ITS.

In (Inventado et al., 2010) the authors applied a combination of Bayesian networks and ML techniques with the aim to detect students' reactions while using an intelligent tutoring system and to provide customizable feedback to each student. Baker and colleagues (Baker et al. 2010) study the Contextual-Guess-and-Slip variant on Bayesian Knowledge Tracing to classical four-parameter Bayesian Knowledge Tracing and the Individual Difference Weights in investigating how well each model variant predicts student performance, not only within the intelligent tutoring system, but on paper post-tests outside of the system.

In (Ueno, 2003a; Ueno, 2003b), a tool for the automatic detection of atypical behaviors 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 behavior were also referenced in the section reviewing clustering applications (Castro et al., 2005b; Vellido et al., 2006).

2.1.5.2 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 (Reffay and Chanier, 2003) 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 behavior 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.

A review of the application of Social Network Analysis (SNA) as a useful tool for improving e-Learning systems is presented in (Cela et al., 2014). This work assesses the evidence of using SNA as a way to understand and improve e-Learning systems and suggests directions for future research. SNA, particularly when combined with content analysis, can provide a detailed understanding of the nature and type of interactions between members of the network, allowing the optimisation of course design, composition of learner groups and identification of learners in danger of dropping out.

In (Zaki, 2016) the PageRank algorithm was applied to identify homogenous groups of students, and select the best peer leaders in classrooms who can disseminate the

educational knowledge in an efficient way. In the proposed approach, the graph knowledge and social interactions are used as a way to create natural groups of students.

An educational Data Mining tool is presented in (Mostow et al., 2005; Mostow et al., 2006) 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 (Castro et al., 2005b; Castro et al., 2005a; Vellido et al., 2006) to display high-dimensional student behavior data in a 2-dimensional representation. This type of visualization helps detecting the characteristics of the data distributions and their grouping or cluster structure.

2.1.6 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 (Abe et al., 2003; Hasegawa and Ochimizu, 2005; Seki et al., 2005; Sheard et al., 2003; Jeong and Biswas, 2008), as well as Semantic Web technologies (Holohan et al., 2005), ontologies (Leidig, 2001), Case-Based Reasoning (Heraud et al., 2004), novel framework technologies (Le Ru et al., 2015), Grid-aware technology (Caballé et al., 2005), argumentation theories (Wang, 2014) and/or theoretical modern didactical approaches (Biswas et al., 2004; Brusilovsky, 2001; Hwang et al., 2004c; Weber and Brusilovsky, 2001).

Although it could have been included in the section devoted to classification, Naïve Bayes, the model used in (Singh, 2004; Tang et al., 2000), also fits in the description of general statistical method. An approach to automate the classification process of Web learning resources was developed in (Singh, 2004). 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 (Tang et al., 2000), 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 (Carbonaro, 2003; Monk, 2005; Pahl and Donnellan, 2002; Vialardi et al., 2011). In (Pahl and Donnellan, 2002), 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 behavior. Finally, in the *time series of session data*, the evolution of session statistics and session patterns over a period of time was analysed. All methods were applied to Web log entries. In (Carbonaro, 2003), a personalized learning environment applying different symmetric and asymmetric distance measures between the students' profiles and their interests was proposed. In (Monk, 2005), 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.

In (Vialardi et al., 2011) the authors propose a data mining approach to guide students through the enrolment process based on academic performance. In this research the CRISP-DM methodology is applied to develop a recommender system for the enrolment process. A conceptual model to examine the influences of perceived individual and social learning support on employees' acceptance of competency-based e-Learning systems is introduced in (Cheng et al., 2011). In (Caballé et al., 2005) a grid technology is applied for providing effective feedback to online learning groups. For the proposed approach, the event log files from the Basic Support for Collaborative Work (BSCW) is utilized and exploited. The BSCW is a shared workspace system used in Open University of Catalonia, which enables collaboration over the Web by supporting document upload, group management and event service, to name just a few features.

A Hidden Markov model was applied in (Jeong and Biswas, 2008) with the ultimate goal of building students' behavior models from data collected in the log files. 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 (Biswas et al., 2004). In (Hwang et al., 2004c), a model called Learning Response Dynamics that analyses learning systems through the concepts of learning dynamics, energy, speed, force, and acceleration, was described. In (Brusilovsky, 2001; Weber and Brusilovsky, 2001), 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.

A Web-based two-tier diagnostic assessment and Web-based dynamic assessment is used in (Wang, 2014) to develop an assessment-centered e-Learning system, named GPAM-WATA. This system provides a personalized dynamic assessment that generates customized evaluation based on a pre-test and personalized e-Learning material. From

this study it is found that the personalized dynamic assessment is significantly more effective in facilitating student learning achievement and improvement of misconceptions, especially for students with low-level prior knowledge.

In (Le Ru et al., 2015) the authors propose a Web-based system to help students, teachers and academic staff to assess automatically programming assignments. The proposed technologies include Laravel Framework, LAMP, design patterns and twitter bootstrap.

MAS have also been applied to e-Learning beyond classification problems. In (Shang et al., 2001), 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 behavior (student model) is inferred from the performance data using a Bayesian Belief Network model. In (Razek et al., 2002; Razek et al., 2003), a MAS called Cooperative Intelligent Distance Learning Environments (CIDLE) was described. It extracts knowledge from domain knowledge and students' behavior during a learning discussion. It therefore infers the learners' behavior and adapts to them the presentation of course material in order to improve their success rate in answering questions. In (Markham et al., 2003), 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 (Heraud et al., 2004), 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 (Prentzas et al., 2002). 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. The main goal of the research presented in (AlAjmi, et al., 2012) is to inspect the impact of a number of e-Learning actions on the students' learning development. The results show that involvement in virtual classroom sessions has the most considerable impact on the students' final scoring or grade.

Finally, in (Cho et al., 2007), 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.

2.2 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.

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 2.1. Research works that perform students' learning assessment

Reference	DM objective	DM approach	e-Learning actor	Type of publication
(Monk, 2005)	Statistical analysis	Basic Statistical Methods	Student and Staff	J
(Hwang, 1999)	Classification	Fuzzy Reasoning	Student	J
(Hwang, 2003)	Clustering	Clustering, Dynamic Programming and Fuzzy Logic Theory	Student and Teacher	J
(Chang et al., 2003)	Classification	Conceptual Maps	Student and teacher	J
(Abe et al., 2003)	Statistical analysis	Metadata Analysis	Student and Teacher	C
(Hwang et al., 2003)	Descriptive behavior in data	Concept Effect Relationship (CER) Model	Teacher	J
(Seki et al., 2005)	Statistical analysis	Basic Statistical Methods	Student and Teacher	C

(Hasegawa and Ochimizu, 2005)	Statistical analysis	Metadata Analysis	Student and Teacher	C
(Matsui and Okamoto, 2003)	Descriptive behavior in data	ID3	Teacher	C
(Yoo et al., 2006)	Classification and Visualization	ADT Tree	Student and Teacher	C
(Pahl and Donnellan, 2002)	Classification	Basic Statistical Methods	Teacher	C
(Reffay and Chanier, 2003)	Visualization and Clustering	Social Network Analysis	Teacher	C
(Cho et al., 2007)	Classification	Social Network Analysis	Teacher	J
(Traynor and Gibson, 2005)	Classification	Code Generation and Mutation.	Teacher	C
(Stathacopoulou and Grigoriadou, 1999)	Classification	Neuro-Fuzzy Model	Teacher	C
(Prentzas et al., 2002)	Classification	Expert Systems and Neural Computing	Teacher	C
(Minaei-Bigdoli and Punch, 2003)	Classification	Combination of: k-NN, MLP and Decision Tree	Teacher	C
(Minaei-Bigdoli et al., 2004)	Descriptive behavior in data	Contrast Rules	Teacher	C
(Resende and Pires, 2001; Resende and Pires, 2002)	Descriptive behavior in data	Association Rules	Teacher	C; C
(Chang and Wang, 2001)	Classification	Association Rules	Teacher	C
(Mullier, 2003; Mullier et al., 2001)	Clustering	SOM	Teacher	J; C
(Hsu et al., 2003)	Descriptive behavior in data	Association Rules	Student and Teacher	C
(Kumar, 2005)	Descriptive behavior in data	Association Rules	Teacher	C
(Sheard et al., 2003)	Statistical analysis	Basic Statistical Methods	Student, Teacher and Staff	J
(Tang and McCalla, 2005)	Clustering	Navigation Path Clustering ad-hoc Algorithm	Teacher	C
(Liccheli et al., 2004)	Classification	Decision Tree-based Rule Extraction	Teacher	C
(Muehlenbrock, 2005)	Prediction	Decision Tree	Teacher	C
(Arroyo et al., 2004)	Prediction	Bayesian Network	Teacher	B
(Kotsiantis et al., 2004)	Classification and Prediction	Naïve Bayes, k-NN, MLP-ANN, C4.5, Logistic Regression and SVM	Teacher	J
(Beck and Woolf, 2000)	Prediction	Linear Regression	Teacher	C
(Feng et al., 2005)	Prediction	Regression	Teacher	C
(Mostow et al., 2005)	Visualization	SQL Queries	Teacher	C
(Mostow et al., 2006)	Visualization	SQL Queries	Teacher	C
(Nebot et al., 2006)	Prediction	FIR	Teacher	C
(Etchells et al., 2006)	Prediction	FIR and OSRE	Teacher	C
(Talavera and Gaudioso, 2004)	Clustering	EM Algorithm	Teacher	C
(Chang et al., 2006)	Prediction	Dynamic Bayes Net	Teacher	C
(Ogor, 2007)	Classification	Neural Network and k-Means	Teacher	C
(Jurado et al., 2008)	Classification	Fuzzy Models	Student	C
(Jeong and Biswas, 2008)	Classification	Hidden Markov Models	Teacher	C
(Taylan and Karagozoglu,	Prediction	Neuro-Fuzzy Model	Teacher	J

2009)				
(Lykourantzou et al., 2009)	Classification	NN, SVM and Fuzzy ARTMAP	Teacher	J
(Ozpolat and Akar, 2009)	Classification	NBTree	Teacher	J
(Kotsiantis et al., 2010)	Classification	Ensemble of Classifiers	Teacher	J
(Hogo, 2010)	Clustering	Fuzzy Clustering Techniques	Teacher	J
(Rajibussalim, 2010)	Clustering and Classification	K-Means and J48 Algorithm	Teacher	C
(Baker et al., 2010)	Classification	Bayesian Knowledge Tracing	Teacher	C
(Yadav and Singh, 2011)	Classification	Fuzzy Expert Systems	Teacher	J
(Baradwaj and Pal, 2011)	Classification	Decision Tree	Teacher and Student	J
(Dejaeger et al., 2012)	Classification	M5 and CART	Teacher	J
(Sen et al., 2012)	Classification	C5 Decision Tree, ANN, SVM and Logistic Regression	Teacher	J
(Goel et al., 2012)	Prediction	Fuzzy Logic	Teacher	C
(Jeremic et al., 2012)	Classification	Fuzzy Models	Teacher	J
(Jurado et al., 2012)	Classification	Fuzzy Logic	Student	J
(Kardan et al., 2013)	Prediction	Neural Network	Teacher	J
(Wang, 2014)	Classification	Argumentation Theory	Teacher and Student	J
(Yadav et al., 2014)	Classification	Fuzzy Expert Systems	Teacher	C
(Le Ru et al., 2015)	Classification	Design Patterns and Novel Software Technologies	Student, Teacher and Staff	B
(Parmar et al., 2015)	Prediction	Random Tree	Teacher	B
(Kolo et al., 2015)	Classification	Decision Tree	Teacher	J
(Burgos et al., 2017)	Classification	Logistic Regression	Teacher and Staff	J
(Carmona et al., 2010)	Descriptive behavior in data; Subgroup Discovery	Association Rules; Evolutionary Algorithms	Teacher	C
(Romero et al., 2009)	Descriptive behavior in data; Subgroup Discovery	Association Rules; Evolutionary Algorithms	Teacher	J

Although an important deal of research effort has been devoted to improve the students' e-Learning experience (see Tables 2.2 and, partially, 2.4), 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 2.1, 2.3 and 2.5, as well as, partially, 2.4).

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

Table 2.2. Research works that offer course adaptation based on students' learning behavior

Reference	DM objective	DM approach	e-Learning actor	Type of publication
(Grieser et al., 2002)	Classification	Consistency Queries (CQ) Inductive Inference Machine	Student	C
(Jantke et al., 2004)	Classification	Consistency Queries (CQ) Inductive Inference Machine	Student	C
(Van Rosmalen et al., 2005)	Prediction	Software Agents	Student	C
(Tang et al., 2000)	Prediction	Ad hoc Naïve algorithm for Tutor Tree	Student	C
(Fernández et al., 2003)	Classification	Multi-Agent Systems	Student	C
(Carchiolo et al., 2003)	Classification	Graph Theory	Student	C
(Costabile et al., 2003)	Descriptive behavior in data	IF-THEN Rules	Student	
(Andronico et al., 2003)	Classification	Multi-Agent Systems	Student	
(Markellou et al., 2005)	Descriptive behavior in data	Apriori Algorithm	Student	C
(Carbonaro, 2003)	Classification	Distance Measures	Student	C
(Wang et al., 2002)	Classification	Association Rules	Student	C
(Hsu et al., 2003)	Descriptive behavior in data	Association Rules	Student and Teacher	C
(Mitzue and Toshio, 2001)	Classification	Neural Network	Student	J
(Liccheli et al., 2004)	Classification	Decision Tree-based Rule Extraction	Teacher	C
(Romero et al., 2004; Romero et al., 2003)	Prediction	Prediction Rules	Student	C; J
(Heraud et al., 2004)	Classification	Case-Based Reasoning	Student	C
(Hammouda and Kamel, 2005)	Clustering	HAC, Single-Pass and k-NN	Student	B
(Liang et al., 2000)	Classification	Rough Set Theory and Decision Trees	Student and Teacher	C
(Razek et al., 2002; Razek et al., 2003)	Prediction	Multi-Agent Systems and ID3	Teacher	C; C
(Shang et al., 2001)	Prediction	Bayesian Network	Student	J
(Trella et al., 2005)	Prediction	Bayesian Network	Student	C
(Fazlollahtabar and Mahdavi, 2009)	Classification	Neuro-Fuzzy Model	Teacher and Student	J
(Jia et al., 2010)	Classification	Fuzzy Theory	Student	B
(Franco-Lugo et al., 2016)	Classification	Ant Colony Optimization	Student	B

Table 2.3. Data mining applications providing an evaluation of the learning material

Reference	DM objective	DM approach	e-Learning actor	Type of publication
(Zaïane, 2002; Zaïane and Luo, 2001)	Descriptive behavior in data	Software Agents and Association Rules	Student	C; C

(Tsai et al., 2001)	Descriptive behavior in data	Association Rules (integrating Apriori Algorithm), Fuzzy Set Theory and Inductive Learning (AQR algorithm)	Teacher	C
(Hwang et al., 2004a)	Group Decision methods	Group Decision Method, Grey System and Fuzzy Theory	Student, Teacher and Staff	J
(Hwang et al., 2004b)	Classification and prediction	Fuzzy Rules	Student, Teacher and Staff	C
(Singh, 2004)	Classification	Naïve Bayes	Teacher	C
(Pahl and Donnellan, 2002)	Classification	Basic Statistical Methods	Teacher	C
(Dos Santos et al., 2003)	Descriptive behavior in data	Web Usage Mining: Association and Sequence	Teacher	C
(Sheard et al., 2003)	Statistical analysis	Basic Statistical Methods	Student, Teacher and Staff	J
(Tane et al., 2004)	Clustering and Visualization	Bisection k-Means	Teacher	C
(Drigas and Vrettaros, 2004)	Clustering	SOM	Teacher	J
(García et al., 2011)	Descriptive behavior in data	Association Rule Mining	Teacher	J

Table 2.4. 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
(Zaïane, 2002; Zaïane and Luo, 2001)	Descriptive behavior in data	Software Agents and Association Rules	Student	C; C
(Hwang, 1999)	Classification	Fuzzy Reasoning	Student	J
(Abe et al., 2003)	Statistical analysis	Metadata Analysis	Student and Teacher	C
(Hasegawa and Ochimizu, 2005)	Statistical analysis	Metadata Analysis	Student and Teacher	C
(Yoo et al., 2006)	Classification	ADT Tree	Student and Teacher	C
(Hsu et al., 2003)	Descriptive behavior in data	Association Rules	Student and Teacher	C
(Chu et al., 2003)	Descriptive behavior in data	Apriori Algorithm	Student	C
(Liang et al., 2000)	Classification	Rough Set Theory and Decision Trees	Student and Teacher	C
(Teng et al., 2004)	Clustering	EM Algorithm	Teacher	C
(Arroyo et al., 2004)	Prediction	Bayesian Network	Teacher	B
(Beck and Woolf, 2000)	Prediction	Linear Regression	Teacher	C
(Feng et al., 2005)	Prediction	Regression	Teacher	C

(Caballé et al., 2005)	Clustering	Grid Technology	Student	C
(Etchells et al., 2006)	Prediction	FIR and OSRE	Teacher	C
(Nebot et al., 2006)	Prediction	FIR	Teacher	C
(Castro et al., 2005b)	Clustering	GTM	Teacher	C
(Chang et al., 2006)	Prediction	Dynamic Bayes Net	Teacher	C
(Schiaffino et al., 2008)	Classification	Intelligent Agents	Student	J
(Wang et al., 2009)	Classification	Support Vector Machine	Student	C
(Inventado et al., 2010)	Classification	Bayesian Networks	Student	C
(Buldu and Ucgun, 2010)	Descriptive behavior in data	Apriori Algorithm	Teacher and Student	J
(Maull et al., 2010)	Clustering	K-Means and E-M	Teacher	C
(Shannaq et al., 2010)	Classification	Rule Extraction from Decision Trees	Teacher and Staff	J
(Dogan and Camurcu, 2010)	Clustering	K-Means and Fuzzy C-Means	Teacher	J
(Weng, 2011)	Descriptive behavior in data	Fuzzy Apriori Algorithm	Teacher and Student	J
(Vialardi et al., 2011)	Classification	C4.5, KNN and Naïve Bayes	Student	J
(Cheng et al., 2011)	Clustering	Structural Equation Modeling	Student	J
(Cela et al., 2014)	Visualization	Social Network Analysis	Teacher and Student	J
(Aher and Lobo, 2013)	Clustering and Classification	K-Means and Apriori Algorithm	Student	J
(Vázquez-Barreiros et al., 2014)	Classification	Genetic Algorithms	Teacher	C
(Samigulina and Shayakhmetova, 2015)	Clustering and Classification	Neural Networks, Genetic Algorithms and Fuzzy Logic	Student	B
(Velmurugan et al., 2016)	Classification	Big Data Technologies	Teacher and Staff	J
(Kato et al., 2016)	Clustering	K-Means	Teacher and Student	B
(Zaki, 2016)	Clustering	PageRank	Teacher and Student	B
(Dwivedi and Rawat, 2017)	Classification	Apriori Algorithm	Student	J

Table 2.5. Data Mining applications for the detection of atypical learning behaviors

Reference	DM objective	DM approach	e-Learning actor	Type of publication
(Ueno, 2003a; Ueno, 2003b)	Outliers detection	Bayesian Predictive Distribution Model	Teacher	C; C
(Castro et al., 2005a)	Outliers detection	GTM	Teacher	C
(Vellido et al., 2006)	Outliers detection	GTM	Teacher	C
(Weng, 2011)	Descriptive behavior in data	Fuzzy Apriori Algorithm	Teacher and Student	J

(Romero et al., 2010)	and Outliers detection Descriptive behavior in data	Association Rules (Apriori-Inverse; Apriori-Rare)	Teacher and Student	C
-----------------------	--	---	------------------------	---

2.3 Discussion and opportunity for the use of Data Mining in e-Learning systems

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

2.3.1 e-Learning courseware optimization

The possibility of tracking user behavior in virtual e-Learning environments makes possible the mining of the resulting databases. 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 2.3, could be used. Classification problems are dominant in this area, although prediction and clustering are also present.

Some of the publications reported in Table 2.1 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 2.3, 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.

2.3.2 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 2.1, 2.2 and 2.4, 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 analyse 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. (Srivastava et al., 2000).

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 behaviors 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 (Minaei-Bigdoli and Punch, 2003) and the neuro-fuzzy models put forward in (Stathacopoulou and Grigoriadou, 1999).

2.3.3 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 behavior 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 2.1, 2.4 and 2.5. 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 2.1).
2. Tools that allow performing an evaluation of the learning materials (Table 2.3).
3. Tools that provide feedback to the tutors based on the students' learning behavior (Tables 2.4-2.5).

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 (Castro et al., 2005a; Vellido et al., 2006) with feature

relevance determination and FIR (Etchells et al., 2006; Nebot et al., 2006) methodologies have been applied.

From the virtual teacher standpoint, valuable information could be obtained 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 analysing the learning behavior 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 (Drigas and Vrettaros, 2004; Hammouda and Kamel, 2005), 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 DM for e-Learning, which, in our 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.

2.4 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: Blackboard, TopClass, Ingenium Docent, etc. (Croock et al., 2002; Van der Klink et al., 2002), 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 (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.

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 (Van Rosmalen et al., 2005).

Several large research projects have dealt with the integration of Data Mining methods in e-Learning (Table 2.6). 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 Learning Online 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 analyse individual access paths through the material interaction behavior.

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.

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.

Table 2.6. 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
aLFanet	Software Agents, Machine Learning, Association Rules	Course adaptation to the students' navigational behavior	Universidad Nacional de Educación a Distancia and Open University of the Netherlands. Spain Portugal, Germany and Netherlands	http://adenu.ia.uned.es/alfanet/ (1/2/2018)
AHA!	Prediction Rules	Course adaptation to the students' navigational behavior	Eindhoven University of Technology and Cordoba University. Netherlands and Spain	http://aha.win.tue.nl (1/2/2018)
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/ (1/2/2018)
ATutor	Statistical analysis	Assessment system and student behavior tracking	University of Toronto, Canada	www.atutor.ca/ (1/2/2018)
LEXIKON	Consistency queries (CQ) inductive inference	Course adaptation to the students' navigational behavior	German Research Center for Artificial Intelligence, Technische Universität Darmstadt, and others, Germany	http://lexikon.dfki.de/ (1/2/2018)
Blackboard	Statistical Analysis	Assessment system and student behavior tracking	Blackboard	http://www.blackboard.com/index.html (1/2/2018)
ELM-ART	Intelligent program analysis, concept-based hyperspace organization, adaptive navigation support	Course adaptation to the students' navigational behavior	University of Education Freiburg, Freiburg im Breisgau, Germany and University of Pittsburgh, Pittsburgh, PA, USA	http://art2.ph-freiburg.de/Lisp-Course (1/2/2018)
Moodle Learning Analytics	Statistical Analysis and Machine Learning	Learning Analytics: Student behavior tracking	Moodle	https://docs.moodle.org/34/en/Analytics (1/2/2018)

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.

ELM-ART is a Web-based Intelligent Educational system that offers a creative combination of two different paradigms - Intelligent Tutoring and Adaptive Hypermedia technologies. ELM-ART provides all learning material online in the form of an adaptive interactive textbook. Using a combination of an overlay model and an episodic student model, ELM-ART provides adaptive navigation support, course sequencing, individualized diagnosis of student solutions, and example-based problem-solving

support. Furthermore, in ELM-ART students can answer questions in the text sections, explore program examples by running them with different parameters, interactively solve problems and receive system feedback. The textbook should also be intelligent in several senses, i.e., using several AI technologies to support readers – the intelligent program analysis, concept-based hyperspace organization with a set of domain concepts behind each content page, intelligent links between examples and problems, and, finally, adaptive navigation support that could guide students to the material that is most appropriate for the current knowledge (Weber and Brusilovsky, 2016).

Moodle is one of the most known and used open source LMS. Since the version 3.4 Moodle has included a plug-in module to perform learning analytics, by machine learning algorithms, that provide predictions of learner success, and diagnosis about the learners behaviors. One of the most interesting capabilities of the plug-in is the report of Student at risk of dropping out. This functionality detects students who are at risk of non-completion of a Moodle course, based on low student engagement. In this model, the definition of "dropping out" is "no student activity in the last quarter of the course" (MOODLE, 2018).

2.5 Key e-Learning resources

In this section, we synthesize, in a self-contained manner, some key resources for the e-Learning community. 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; key e-Learning books and book chapters; and open source e-Learning software.

Table 2.7. International journals and conferences specialized on e-Learning or including it within their scope and main topics

Scientific Journal	International Conference (the edition corresponds to that held on 2017)
The ACM Transactions on Computing Education (TOCE), ACM	The 4 th International KES Conference on Smart Education and E-Learning (KES-SEEL)
Computers & Education, Elsevier	IEEE International Conference on Advanced Learning Technologies (ICALT), on its 17 th edition
Education and Information Technologies, Springer-Verlag	International Conference of the Association for Learning Technology, (ALT-C), on its 24 th edition
European Journal of Open, Distance and e-Learning (EURODL), European Distance and e-Learning Network (online only)	International Conference on Artificial Intelligence in Education (International AIED Society), on its 18 th edition
Electronic Journal of e-Learning (EJEL), Academic Conferences and Publishing International Limited	World Conference on E-Learning (E-Learn), on its 22 nd edition
Journal of Computer Assisted Learning, Blackwell Publishing Inc.	The 28 th Annual Conference of the Society for Information Technology and Teacher Education (SITE)

International Journal of Computer-Supported Collaborative Learning, Springer-Verlag	Global Conference on Learning and Technology (Global Learn), on its 6 th edition
Online Learning Journal, Online Learning Consortium	World Conference on Educational Media & Technology (EdMedia), on its 29 th edition
IEEE Transactions on Learning Technologies, IEEE Computer Society and IEEE Education Society	International Conference on E-Education, E-Business, E-Management and E-Learning (IC4E), on its 8 th edition
Interactive Learning Environments, Swets & Zeitlinger	Annual Hawaii International Conference on Education (HICE Education), on its 15 th edition
International Journal of Innovation and Learning, Inderscience Publishers	International Conference on Education and New Learning Technologies (EDULEARN), on its 8 th edition
E-Learning, Symposium Journals Ltd. (online only)	International Conference on Computers in Education, (ICCE), on its 25 th edition
IEEE Transactions on Education, IEEE Education Society	International Conference on Engineering Education, (ICEE), on its 19 th edition
International Journal of Artificial Intelligence in Education, Springer-Verlag	International Conference on Intelligent Tutoring Systems, (ITS), on its 13 th edition
International Journal on e-Learning (IJEL), AACE	International Conference on Interactive Collaborative Learning, (ICL), on its 20 th edition
Journal of Educational Multimedia and Hypermedia (JEMH), AACE	International Conference on Web-based Learning, (ICWL), on its 16 th edition
Journal of Information Technology Education (JITE), Informing Science Institute	International Conference on Web-based Learning (ICWL), on its 16 th edition
Journal of Interactive Learning Research, (JILR), AACE	IEEE International Conference on Engineering Education (ICEED), on its 9 th edition
Journal of Online Teaching and Learning (JOLT), MERLOT (online only)	IEEE Frontiers in Education Conference (FIE), on its 47 th edition
User Modeling and User-Adapted Interaction (UMUAI), Springer-Verlag	The European Conference on e-Learning (ECEL), on its 16 th edition
	The IEEE Global Engineering Education Conference (EDUCON), on its 9 th edition

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 2.7 summarizes this information.

In Table 2.8, 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 2.8. e-Learning discussion forums

Name	URL of the forum
Adapt Learning Steering Group	https://www.adaptlearning.org/index.php/forums/ (30/1/2018)
Ako Aotearoa eLearning Forum	https://ako.aotearoa.ac.nz/communities/elearning-forum

	(30/1/2018)
eLearning Forum	http://elearningforum.com/ (30/1/2018)
eLearning Forum for NATO'S Leading Conference	http://www.act.nato.int/e-Learning-forum (30/1/2018)
eLearning Learning Forum	https://www.elearninglearning.com/forum/ (30/1/2018)
The ELearning Guild forum	https://www.elearningguild.com/ (30/1/2018)
London Global University Forum	https://www.ucl.ac.uk/isd/services/learning-teaching/elearning-staff/other-elearning-tools/communication/forums (30/1/2018)
E-Learning Forum from EPNoSL	http://www.schoolleadership.eu/portal/article/e-Learning-forum (30/1/2018)
Met E-Learning Forum	http://www.met-elearning.org/moodle/mod/forum/ (30/1/2018)
eLearning Developers Forum	http://learningandskillsgroup.ning.com/group/elearningdevelopersforum (30/1/2018)

In Table 2.9, the most important e-Learning organizations, societies and interest groups are presented.

Table 2.9. e-Learning organizations

Name	URL of the organization
The eLearning Guild	http://www.elearningguild.com (1/2/2018)
GAZEL, Greater Arizona eLearning Association	http://www.gazel.org (1/2/2018)
New England Learning Association	http://www.nelearning.org/about/default.asp (1/2/2018)
International Association for Distance Learning	http://www.iadl.org.uk/ (1/2/2018)
Online Learning Consortium	https://onlinelearningconsortium.org/ (1/2/2018)
Masie Center e-Learning Consortium	https://masie.com/MASIE-Information/masieconsortium.html (1/2/2018)
IMS Global Learning Consortium	http://www.imsglobal.org/ (1/2/2018)
Association of Learning Technology (ALT)	https://www.alt.ac.uk/ (1/2/2018)
British Learning Association	http://www.bircham.net/bla-the-british-learning-association.html (1/2/2018)
European Institute for E-Learning (EifEL)	https://www.openeducationeuropa.eu/en/directory/European-Institute-for-E-Learning---EifEL (1/2/2018)
eLearning Alliance	http://www.elearningalliance.org (1/2/2018)
eLearning Network	https://www.eln.com.pk/ (1/2/2018)
National Digital Learning Resources Network from Australian Schools	http://www.ndlrn.edu.au/ (1/2/2018)
Cabot Learning Federation	http://www.cabotlearningfederation.net/ (1/2/2018)
European Distance and E-Learning Network	http://www.eden-online.org/ (1/2/2018)
IEEE Education Society	http://ieeeducationsociety.org/ (1/2/2018)

Table 2.10 lists some main books and book chapters.

Table 2.10. 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-Verlag, 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).

Allen's, M.: *Designing Successful e-Learning*. John Wiley & Sons (2007).

Horton, W.: *E-Learning by Design*. John Wiley & Sons (2011).

Neves-Silva, R., Tsihrintzis, G.A., Uskov, V., Howlett, R.J., Jain, L.C.: *Smart Digital Futures 2014*. IOS Press (2014).

Elkins, D., Pinder, D.: *E-Learning Fundamentals: A Practical Guide*. Assn. for Talent Development (2015).

McQuiggan, S., McQuiggan, J., Sabourin, J., Kosturko, L.: *Mobile Learning: A Handbook for Developers, Educators, and Learners*. John Wiley & Sons (2015).

Uskov, V.L., Howlett, L.J., Jain, L.C. (Editors): *Smart Education and e-Learning*. Smart Innovation, Systems and Technologies 59. Springer-Verlag (2016).

An important issue for the development of e-Learning environments is the existence and availability of open source software. In Table 2.11, the most popular, open source learning management systems are presented.

Table 2.11. Open source e-Learning software

Name	URL of the open source software
ATutor	http://www.atutor.ca/ (1/2/2018)
Claroline	http://www.claroline.net/ (1/2/2018)
Docebo LMS	https://www.docebo.com/ (1/2/2018)
Opigno LMS	https://www.opigno.org (1/2/2018)
ILIAS	http://www.ilias.de/ (1/2/2018)
LAMS	http://lamsfoundation.org/ (1/2/2018)
LRN	http://www.dotlrn.org/ (1/2/2018)
Manhattan Virtual Classroom	http://manhattan.sourceforge.net/ (1/2/2018)
Moodle	http://moodle.com/ (1/2/2018)
OLAT - Online Learning and Training	http://www.olat.org/ (1/2/2018)
Sakai	http://www.sakaiproject.org/

	(1/2/2018)
Eliademy	https://eliademy.com/ (1/2/2018)
Forma LMS	https://www.formalms.org/ (1/2/2018)
Dokeos	https://www.dokeos.com/ (1/2/2018)
Opigno	https://www.opigno.org/en (1/2/2018)
OLAT	http://www.olat.org/ (1/2/2018)
Canvas	http://gocanvas.stanford.edu/ (1/2/2018)
Chamilo	https://chamilo.org/ (1/2/2018)
Ganesha LMS	http://ganesha.fr/ (1/2/2018)
efront	https://www.efrontlearning.com/ (1/2/2018)
Metastudy	http://metacoon.net/ (1/2/2018)
Totara	https://www.totarlms.com/ (1/2/2018)
Kornukopia LMS	http://kornukopia.com/ (1/2/2018)
LON-CAPA	http://www.lon-capa.org/ (1/2/2018)

2.6 Conclusions of the state of the art

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.

Although there are several research efforts including data mining in e-Learning 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 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 behavior.

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 and/or omissions that may have occurred.

Chapter 3: The Fuzzy Inductive Reasoning Methodology: FIR

The inductive reasoning methodology had originally been developed by George Klir as a tool for general systems analysis to study the conceptual modes of behavior of systems. The inductive reasoning set of methods and algorithms forms a subset of Klir's General Systems Problem Solver (GSPS) framework, facets of which have been described in numerous of his publications starting in the late sixties (Klir, 1969; Klir, 1985; Klir and Folger, 1988). In order to go deeply into the GSPS approach let us start by giving a proper definition. The general systems problem solver, GSPS, can be defined as *a conceptual framework through which types of systems problems are defined together with methodological tools for solving problems of these types.*

The GSPS methodology distinguishes between different types of systems characterized by different levels of abstraction. A hierarchical classification by epistemological levels of the fundamental system types found in different disciplines research areas was developed by Klir. Figure 3.1 shows a diagram of this classification.

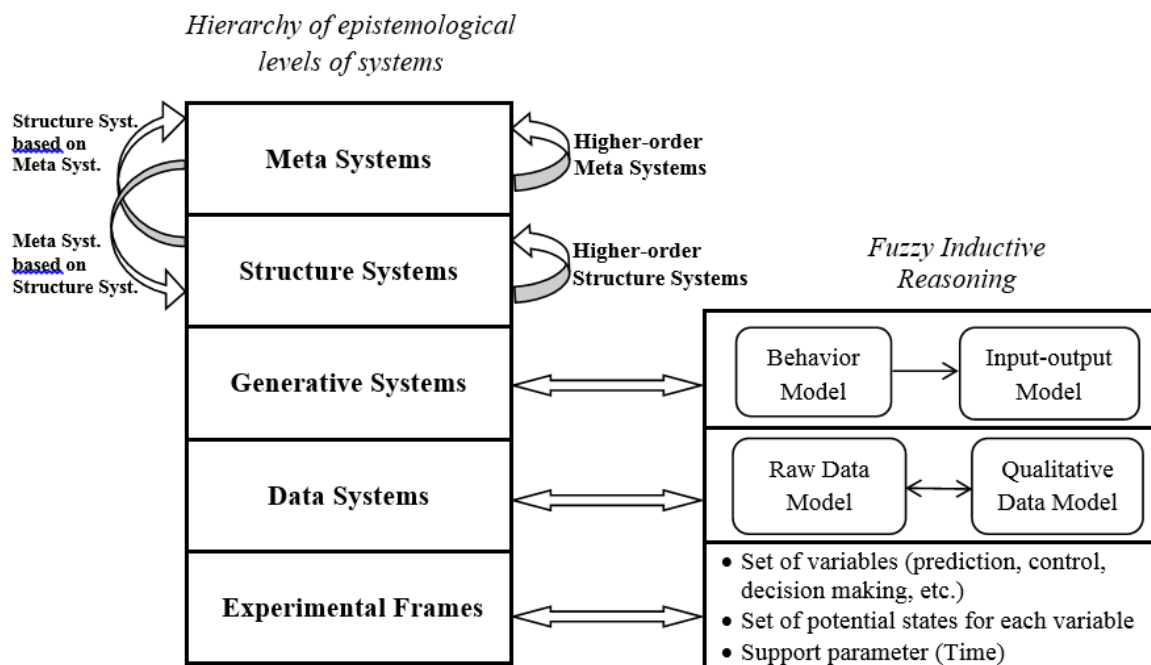


Figure 3.1 Hierarchy of Epistemological Levels of Systems and its relation with the Inductive Reasoning Methodology

The left side of figure 3.1 represents the skeleton of the GSPS taxonomy being such hierarchy vital to the development of any organized package of methodological tools for system problem solving. In our own terminology, a *system* is the physical entity from which mathematical descriptions of varying abstraction, so called *models*, can be derived.

The GSPS methodology distinguishes between an infinity of abstraction levels. The most abstract model is the *source or base model*. It simply encodes knowledge about which facets of the real system are to be captured in the mathematical description. For most practical purposes this knowledge consists in a declaration of a *set of variables* to be contained in the model. Source models can be classified by various criteria through which methodologically significant special properties of the variables are distinguished. One of such criteria is the classification of the variables in input and output variables. Other classifications within this level are the distinctions between crisp and fuzzy variables and continuous and discrete variables. A higher level model entails all knowledge of the corresponding model at any lower level and contains additional knowledge which is not available at the lower levels Therefore the source model is included in all the models of higher levels.

The next higher, i.e. more refined or less abstract level, along Klir's epistemological hierarchy is the *data model*. In order to climb the epistemological hierarchy ladder from the level of the source model to that of the data model it is necessary to supplement the source model by data. Depending on the problem data can be obtained by observation or measurement. The result is a bunch of data streams, or trajectories, which, at this point, still do not contain a description of any logical or causal relationship connecting these data streams to each other. Their only known relationship so far is their common time stamp. Figure 3.1 describes the correspondence between GSPS epistemological levels and the inductive reasoning methodology. In our own terminology we shall call this the *raw data* model because the data have not yet been processed in any way. The raw data model in our implementation of the methodology is represented by a real-valued matrix, whereby each column denotes one variable trajectory, i.e. the recording of the values of one variable as a function of time, whereas each row denotes one data record, i.e. a collection of the values of all variables with identical time stamp.

In order to proceed to higher levels along the epistemological hierarchy ladder, it will prove useful to preprocess these data. In our implementation of the methodology, the raw data model will be preprocessed into a *qualitative data model*, whereby each raw (quantitative) data value is being replaced by a qualitative triplet. The details of this process will be explained in due course. Since the raw data model and the qualitative data model contain exactly the same information, GSPS does not distinguish between the two. Both are located at the same epistemological hierarchy level. Since the transformation from the raw (quantitative) data model to its qualitative counterpart does not add any information to the model, Klir places them at the same hierarchy level.

Climbing up the ladder one rung further, we end up with the *generative or behavior model*. The behavior model adds logical or causal relationships to subsets of the variables. Whereas before we did not know anything about the causal relationship between the recorded variables, they might even stem from entirely different physical objects for that matter, at the new level, this is no longer possible. Now, we know which set of variables we must consult to infer knowledge about one or several other variables. All so-called *input/output models* are located at this hierarchical level. Therefore, this level involves knowledge of some “support-invariant relational characteristics” of the variables involved, that may be exact (deterministic) or approximate (stochastic, fuzzy).

Climbing the epistemological hierarchy ladder even further, we reach the rung of the *structure models*. Most deductively derived (classic) differential equation models are located at that level. Here, the causal relationships of the former behavioral models are concretized to explicit structural relationships between variables, i.e., formulae replacing mere tabulations. Finally, GSPS defines infinitely many rung of so-called *meta-models* that are not further qualified in the GSPS architecture. In some publications, the first meta model level is characterized by variable structure models, i.e. by models that abruptly change their behavior as a consequence of a discrete event taking place (Uyttenhove, 1979). However, since the FIR methodology does not deal at all with these higher elevated rung of the GSPS methodology, there is no need to explore their properties any further in this thesis.

The FIR methodology, a subset of the GSPS methodology, is located entirely at the hierarchical levels of the source, data and behavioral models. It deals with transformations within each of data and behavior levels, and with transitions between the two levels.

In the late seventies, a Ph.D. student of George Klir’s, Hugo Uyttenhove, went about to implement a subset of the GSPS methodology under the name Systems Approach Problem Solver, abbreviated as SAPS (Uyttenhove, 1979). The limited computer science tools available at that time did not lend themselves to a sufficiently flexible implementation of the GSPS concepts and consequently, SAPS could never be used for anything but mere toy problems. In the mid eighties, Cellier and his students went about to reimplement SAPS as a CTRL-C function library. The new implementation was called SAPS-II (Cellier and Yandell, 1987). Fuzzy measures were introduced into the GSPS methodology in the late eighties (Li and Cellier, 1990). More recently, new efficient and extended implementation of SAPS-II, called FIR, has been developed and demonstrated to be an effective tool for qualitatively studying the behavior of highly complex non-linear real systems (de Alborno, 1996; Mugica, 1995; Nebot, 1994; Nebot et al., 2003; Nebot et al., 2009; Nebot and Mugica, 2012). FIR is currently available as a Matlab toolbox.

The FIR methodology is composed of four basic functions: *fuzzification* (fuzzy recoding), *qualitative modeling* (fuzzy optimization), *qualitative simulation* (fuzzy forecasting) and *defuzzification* (fuzzy regeneration) as shown in Figure 3.2.

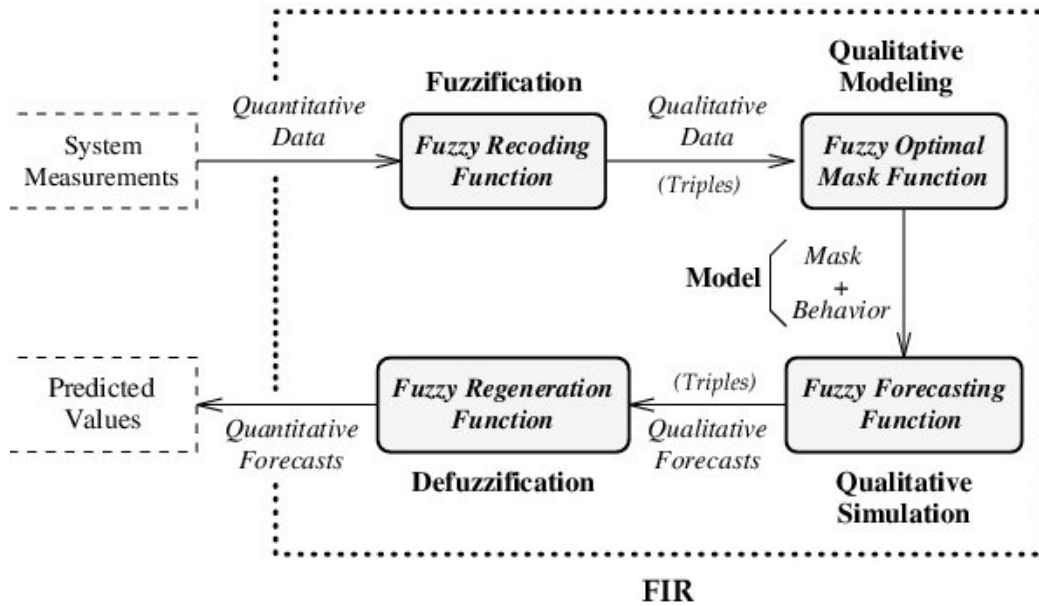


Figure 3.2 FIR main functions

The *fuzzification* module describes a transformation within the data model level, namely from the quantitative (raw) data model to its qualitative counterpart. The *qualitative modeling* module describes the step up the ladder from the data model to the behavioral model. This is accomplished by *induction*. The term induction is synonymous with climbing up the epistemological ladder, while deduction means descending it. The *qualitative simulation* module denotes the transition back down the ladder to the previous level, and the defuzzification module performs another transformation at the data model level.

3.1 Fuzzification

A transformation from quantitative values into qualitative triplets is very useful for the purpose of inductive modeling. Any data-fitting algorithm (and this is what inductive

modeling is all about) invariably involves some sort of optimization procedure. Thus, inductive modeling applied to the original quantitative, i.e. real-valued, variables involves a search across an n-dimensional continuous space. Such a search is invariably very time-consuming. By converting the quantitative values to qualitative triplets (explained later), the search is simplified dramatically, since the search space gets reduced to the n-dimensional discrete search space of the class values. Using this approach, the class values are used for a fairly coarse optimization, whereas the fuzzy membership values are then used for the fine interpolation between neighboring class values, once the optimal class value has been found. In the FIR methodology, the fuzzification process is accomplished by means of the fuzzy recoding function. Recoding denotes the process of converting a quantitative variable to a qualitative variable. In most transformations from a quantitative to a qualitative space, some information is lost in the process. Obviously, a temperature value of 30°C contains more information than the value hot. The FIR recoding technique avoids this problem. Figure 3.3 shows an example of fuzzy recoding of the variable temperature.

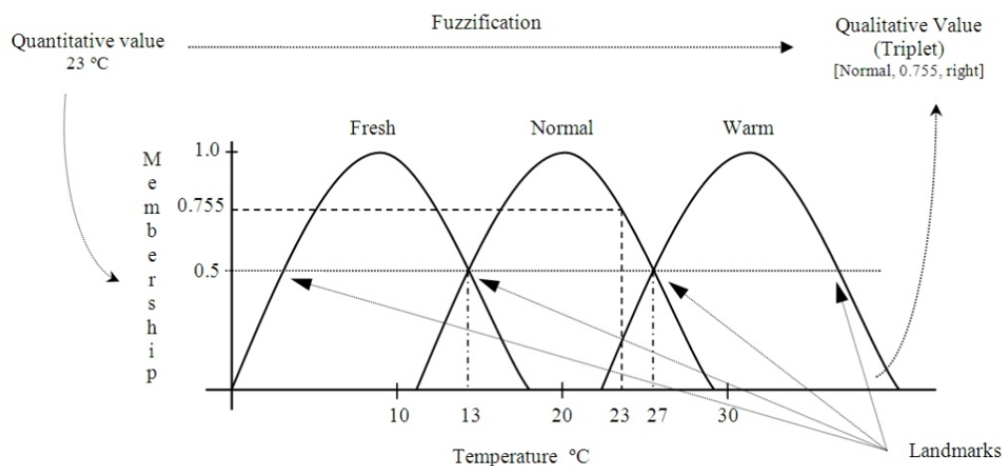


Figure 3.3 Fuzzy recoding of a temperature value of 23°C

In this example, the temperature has been discretized into three classes: fresh, normal and warm using a bell-shaped or Gaussian membership function. For instance, a quantitative temperature value of 23 degrees Centigrade is recoded into a qualitative class value of *normal*, with a fuzzy membership function value of 0.755 and a side function value of *right* (since 23 is to the right of the maximum of the bell-shaped membership function that characterizes the class *normal*). Thus, a single quantitative value is recoded into a qualitative triplet.

Any temperature with a quantitative value between 13 and 27 will be recoded into the qualitative class value *normal*. The fuzzy membership function denotes the value of the bell-shaped fuzzy membership curve that is associated with the selected class, read out at

the point of the quantitative value. It is always a value between 0.5 and 1. Other fuzzification techniques make use of the tails of the membership functions to resolve ambiguity. They assign multiple class values and multiple membership values to a single quantitative value. Our own dialect of fuzzy logic handles the ambiguity issue differently. The side value matrix is introduced as a third piece of information to eliminate the ambiguity, and the tails of the membership functions can thus be ignored.

Evidently, no information is lost in the process of FIR fuzzification. The qualitative triplet contains exactly the same information as the original quantitative value, and it is thus possible to regenerate the quantitative value from the qualitative triplet precisely, i.e. without any error or uncertainty, at any point in time.

By now, the quantitative trajectory behavior has been recoded into a qualitative episodal behavior. In FIR, the episodal behavior is stored in the *qualitative data matrices*. It consists of three matrices of identical size, one containing the class values, the second storing the membership information, and the third recording the side values. Each column represents one of the observed variables and each row denotes one time point, i.e. one recording of all variables, or one recorded state. The class values are in the set of legal levels that each variable can assume. They are all positive integers, as FIR uses integers in place of symbolic values to represent qualitative levels.

3.2 Qualitative modeling

In the FIR methodology, the fuzzy modeling process is performed by means of the fuzzy optimal mask function. It optimizes the predictiveness of the model by performing a search in the discrete space of the class values. The details of how this is accomplished are presented in this section.

How does the episodal behavior support the identification of a qualitative model of a given system for the purpose of forecasting its future behavior for any given input stream?

In the process of modeling, it is desired to discover finite automata relations among the recoded variables that make the resulting state transition matrices as deterministic as possible. If such a relationship is found for every output variable, the behavior of the system can be forecast by iterating through the state transition matrices. The more deterministic the state transition matrices are, the higher is the likelihood that the future system behavior will be predicted correctly.

A possible relation among the qualitative variables for this example could be of the form presented in equation 3.1.

$$y(t) = \tilde{f}(u_1(t - 2\delta t), y(t - 2\delta t), u_2(t - \delta t), u_1(t)) \quad (3.1)$$

where \tilde{f} denotes a qualitative relationship. Notice that \tilde{f} does not stand for any (known or unknown) explicit formula relating the input arguments to the output argument, but only represents a generic causality relationship that, in the case of the FIR methodology, will be encoded in the form of a tabulation of likely input/output patterns, i.e. a state transition table. In FIR, Equation 3.1 is represented by the matrix shown in Figure 3.4.

x	u ₁	u ₂	y
t			
t - 2δt	-1	0	-2
t - δt	0	-3	0
t	-4	0	+1

Figure 3.4 Example of a mask (model structure)

The negative elements in this matrix are referred to as m-inputs, mask inputs. M-inputs denote input arguments of the qualitative functional relationship. They can be either inputs or outputs of the subsystem to be modeled, and they can have different time stamps. The above example contains four m-inputs. The sequence in which they are enumerated is immaterial. They are usually enumerated from left to right and top to bottom. The single positive value denotes the m-output. The terms m-input and m-output are used in order to avoid a potential confusion with the inputs and outputs of the plant. In the above example, the first m-input corresponds to the input variable u_1 two sampling intervals back, whereas the second m-input refers to the output variable y two sampling interval into the past, etc.

In the FIR methodology, such a representation is called a mask. A mask denotes a dynamic relationship among qualitative variables. A mask has the same number of columns as the episodic behavior to which it should be applied, and it has a certain number of rows, the depth of the mask.

How is a mask found that, within the framework of all allowable masks, represents the most deterministic state transition matrix? This mask will optimize the predictiveness of the model.

In FIR, the concept of a *mask candidate matrix* has been introduced. A mask candidate matrix is the ensemble of all possible masks from which the best is chosen by either a mechanism of exhaustive search of exponential complexity, or by one of various suboptimal search strategies of polynomial complexity, as described in (Jerez and Nebot, 1997). The mask candidate matrix contains -1 elements, where the mask has a potential m-input, a +1 element where the mask has its m-output, and 0 elements to denote

forbidden connections. A good mask candidate matrix to determine a predictive model for variable y in the example of Figure 3.4 might be the one presented in Figure 3.5.

	x	u ₁	u ₂	y
t				
t - 2δt		-1	-1	-1
t - δt		-1	-1	-1
t		-1	-1	+1

Figure 3.5 Example of a mask candidate matrix

Based on the candidate matrix, all the possible masks are obtained and each of the possible masks is compared to the others with respect to its potential merit. The optimality of the mask is evaluated with respect to the maximization of its forecasting power. Lets focus on the computation of the goodness of a specific mask.

The Shannon entropy measure is used to determine the uncertainty associated with forecasting a particular output state given any legal input state. The Shannon entropy relative to one input state is calculated as:

$$H_i = \sum_{o} p(o|i) \cdot \log_2 p(o|i) \quad (3.2)$$

where $p(o|i)$ is the conditional probability of a certain m -output state o to occur, given that the m -input state i has already occurred. It denotes the quotient of the observed frequency of a particular state divided by the highest possible frequency of that state. The overall entropy of the mask is then computed as the weighted sum of the entropy over all input states:

$$H_m = - \sum_{i} p(i) \cdot H_i \quad (3.3)$$

where $p(i)$ is the probability of that input state to occur. The highest possible entropy H_{\max} is obtained when all probabilities are equal, and zero entropy corresponds to totally deterministic relationships. A normalized overall entropy reduction H_r is defined as:

$$H_r = 1.0 - \left(\frac{H_m}{H_{\max}} \right) \quad (3.4)$$

H_r is a real number in the range between 0.0 and 1.0, where high values indicate an improved forecasting power. The masks with highest entropy reduction values generate forecasts with the smallest amounts of uncertainty.

One problem still remains. The size of the input/output matrices increases as the complexity of the mask grows and consequently the number of legal states of the model grows quickly. Since the total number of observed states remains constant the frequency of observation of each state shrinks rapidly and so does the predictiveness of the model. With increasing complexity, H_r simply keeps growing. Very soon, a situation is encountered where every state that has ever been observed precisely once. This obviously leads to a totally deterministic state transition matrix and H_r assumes a value of 1.0. Yet the predictiveness of the model will be dismal since in all likelihood already the next predicted state has never before been observed and that means the end of forecasting. Therefore, this consideration must be included in the overall quality measure

From a statistical point of view, every state should be observed at least five times (Law and Kelton, 1990). Therefore, an observation ratio, O_r , is introduced as an additional contributor to the overall quality measure:

$$O_r = \frac{5 \cdot n_{5x} + 4 \cdot n_{4x} + 3 \cdot n_{3x} + 2 \cdot n_{2x} + n_{1x}}{5 \cdot n_{leg}} \quad (3.5)$$

where: n_{leg} is the number of legal m -input states, n_{1x} is the number of m -input states observed only once, n_{2x} is the number of m -inputs states observed twice, and so on.

The idea behind the O_r measure is to penalize those states that have been observed less than five times. If every m -input state has been observed at least five times, O_r is equal to 1.0. If no m -input state has been observed at all (no data are available), O_r is equal to 0.0.

The overall quality of a mask, Q , is then defined as the product of its uncertainty reduction measure, H_r , and its observation ratio, O_r :

$$Q = H_r \cdot O_r \quad (3.6)$$

Once the best mask has been identified, it can be applied to the qualitative data matrices that were previously obtained in the recoding process, resulting in a fuzzy pattern rule base that, in FIR terminology, is called the behavior matrix. How is the pattern rule base obtained from the mask? This process is illustrated in Fig. 3.6.

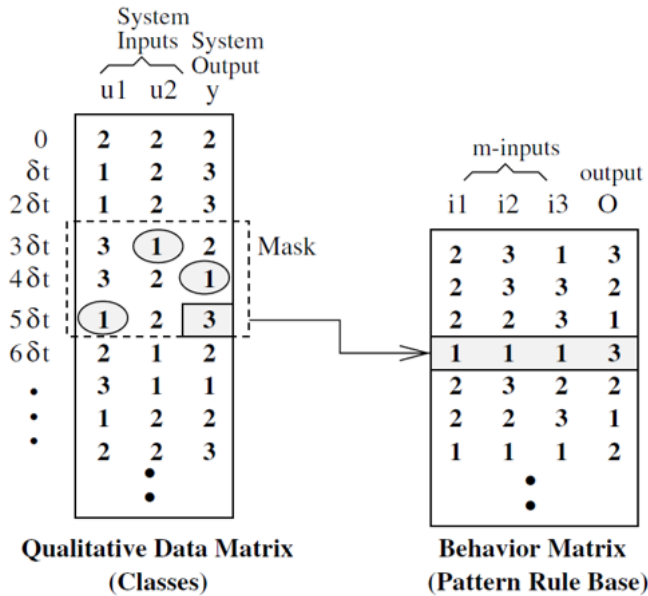


Figure 3.6 FIR pattern rule base construction

The mask can be used to ‘flatten’ dynamic relationships into pseudo-static relationships. The left side of Figure 3.6 shows an excerpt of the qualitative data matrix that stores the class values. The dashed box symbolizes the mask that is shifted downwards along the class value matrix. The round shaded ‘holes’ in the mask denote the positions of the m -inputs, whereas the square shaded ‘hole’ indicates the position of the m -output. The class values are read out from the class value matrix through the ‘holes’ of the mask, and are placed next to each other in the behavior matrix that is shown on the right side of Fig. 3.6. Here, each row represents one position of the mask along the class value matrix. It is lined up with the bottom row of the mask. Each row of the behavior matrix represents one pseudo-static qualitative state or qualitative rule (also called pattern rule). For example, the shaded rule of Figure 3.6 can be read as follows: “If the first m -input, i_1 , has a value of ‘1’ (corresponding to ‘low’), and the second and third m -inputs, i_2 and i_3 , have also values of ‘1’ (corresponding to ‘low’) then the output, o , assumes a value of ‘3’ (corresponding to ‘high’). The qualitative pattern rules can be invoked during qualitative simulation to predict new qualitative outputs.

Note that a FIR model is composed by the mask and the behavior matrix. The mask represents the structure of the model, whereas the behavior matrix is the associated pattern rule base.

3.3 Qualitative simulation

The FIR inference engine is based on a variant of the k -nearest neighbor rule. The k NN pattern matching algorithm is the core of the FIR inferencing process. The forecast of the

output variable is obtained as a weighted average of the potential conclusions that result from firing the k rules, whose antecedents best match the actual state. The prediction procedure is presented in the diagram of Figure 3.7 for an example containing three inputs and one output and using a k value of 5.

The optimal mask is placed on top of the qualitative data matrix in such a way that the m -output matches with the first element to be predicted. The values of the m -inputs are read out from the mask, and the behavior matrix (pattern rule base) is used to determine the future value of the m -output, which can then be copied back into the qualitative data matrix. The mask is then shifted further down by one position to predict the next output value. This process is repeated until all desired values have been forecast. The qualitative simulation process predicts an entire qualitative triplet, from which a quantitative variable can be obtained whenever needed.

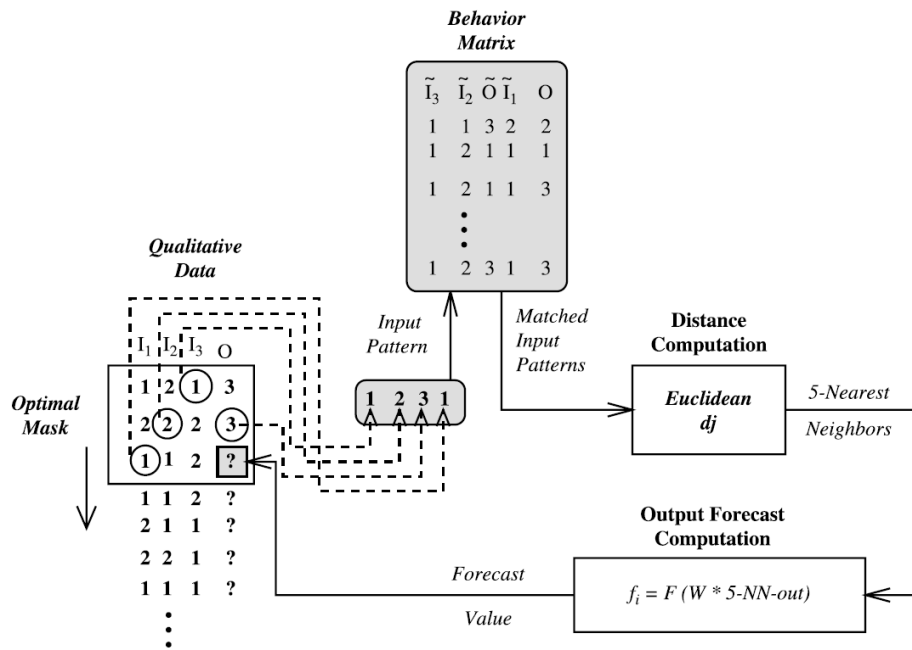


Figure 3.7 FIR forecasting process diagram.

The prediction process works as follows. The values of the m -inputs are read out from the mask and concatenated from the right to form the input pattern (input state) associated to the output value to be predicted. The pattern rule base is used to obtain the previous (historical) input states that match the new input pattern from the behavior matrix.

In fuzzy forecasting the membership and side functions of the new input state are compared with those of all previous recordings of the same input state contained in the class behavior matrix. For this purpose a normalization (pseudo-regenerated) function is computed for every element of the new input state, by means of equation 3.7.

$$psr_i = Class_i + Side_i \cdot (1.0 - Memb_i) \quad (3.7)$$

Irrespective of the original values of the input variable, psr_i assumes values in the range [1.0, 1.5] for the lowest class, [1.5, 2.5] for the next higher class, etc.

The psr_i values are quantitative variables that can be used to represent the relative magnitude of a particular qualitative triplet. However, they are not regenerations of the original quantitative signals. They are normalized variables. Consequently different pri_i signals can be compared to each other or can be summed up without weighing them relative to each other something that would not be meaningful using the original or regenerated signals.

The normalization function is computed for every input variable of the new input state, and the normalized psr_i values are then concatenated to form the $psrIN$ vector of equation 3.8.

$$psrIN = [psr_1, psr_2, psr_3, \dots, psr_N] \quad (3.8)$$

The $L2$ norms of the differences between the $psrIN$ vector representing the new input state and the $psrPR_j$ vectors representing all previous recordings of the same input state are computed using the well-known Euclidean distance measure described in equation 3.9.

$$d_j = \sqrt{\sum_{i=1}^N (psrIN_i - psrPR_{j,i})^2} \quad (3.9)$$

Notice that the sub-index i stands for the different m-inputs of a given input pattern whereas the sub-index j stands for the different historical input patterns found in the pattern rule base that match the new input pattern in terms of class values. The contribution of each neighbor to the estimation of the prediction of the new output state is a function of its proximity. This is expressed by giving a distance-weight to each neighbour.

Absolute weights are computed using one of two formulae. If none of the five smallest distance functions, d_j is exactly equal to zero, we use the equation 3.10.

$$w_{absj} = \frac{(d_{max}^2 - d_j^2)}{d_{max} \cdot d_j} \quad (3.10)$$

where the index j loops over the k closest neighbours, and $d_a \leq d_b$; $a < b$; $d_{max} = d_k$. Obviously, the above formula will not work if any of the d_j values is zero, since this leads to a singularity. In this situation, equation 3.11 is being used instead.

$$w_{absj} = \begin{cases} 0.0; & d_j \neq 0.0 \\ 1.0; & d_j = 0.0 \end{cases} \quad (3.11)$$

The idea behind these formulae is that if one of the previous observations leads to a very small distance function its weight should dominate the computation yet if all distance functions are equally large we should make use of an arithmetic mean between the previous distance functions.

Using the sum of the k absolute weights, $s_w = \sum_{j=1}^{j=k} w_{absj}$, it is possible to compute relative weights (distance-weights), as shown in equation 3.12.

$$w_{relj} = \frac{w_{absj}}{s_w} \quad (3.12)$$

The relative weights are numbers between 0.0 and 1.0, and their sum always equals 1.0. Thus, the relative weights can be interpreted as percentages. Using this idea, the new output state values in the normalized space can be computed as a weighted sum of the output states of the previously observed k nearest neighbours.

FIR implements other formulae to compute the distance and the weights, however the ones presented here are the ones defined by default.

3.4 Defuzzification

Regeneration is the inverse function of recoding. It converts qualitative triples into quantitative values. As has been mentioned earlier, no information is lost in the process of fuzzification. The qualitative triple contains exactly the same information as the original quantitative value, and it is thus possible to recover a unique quantitative value from the qualitative triple.

Chapter 4: Conclusions and Future Research

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.

Although educational data mining has not accomplished greater success as compared to other domains such as e-commerce, up-to-date research results evidences that educational institutions would be able to apply data mining to discover knowledge and improve the learning effectiveness for students as well as enhance their experiences.

e-Learning environments provide a database that stores all the system's information: personal or students profile, academic results, users' interaction data, collaborative resources activity, etc. Although some platforms offer some reporting tools, it becomes hard for a tutor or teacher to extract useful information when there are a great number of students enrolled in the virtual course. Most of the real e-Learning systems do not exploit the available data, except for simple summaries or statistics. This kind of data by itself may be of no help to any of the e-Learning actors. This information is especially unintelligible and very hard to transform to a source of useful knowledge that supports decision-making processes. The use of data mining methods to extract knowledge from the e-Learning system can be an adequate approach to follow, in order to use the obtained knowledge to fit the educational proposal to the students' needs and requirements.

However, 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 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 behavior.

An important problem not solved yet is that a huge amount of time is required for both, the students' assessment process, and for providing feedback to the virtual learners, resulting in an increasing demand of teachers and, therefore, of the educative costs, that not always can be fulfilled. This teacher involvement may limit the e-Learning experience, and actually, increase the teacher's workload. Therefore, it is very difficult

and time consuming for teachers to thoroughly track and assess all the activities performed by all students.

The aim of the work developed in this dissertation was to address some of these difficulties, of e-Learning in general and of data mining applied to e-Learning in particular, and to come up with some data mining methodological developments that would allow to improve the knowledge extraction and, as a result, improve the e-Learning experience.

4.1 Summary of results obtained

The results obtained in this doctoral thesis address several of the problems that are characteristic of e-Learning system. In a first step, an analysis of the feasibility of FIR methodology for modeling and prediction of students' performance was developed. As described before, the FIR methodology offers a model-based approach to predicting either univariate or multi-variate time series. A FIR model is a qualitative, non-parametric, shallow model based on fuzzy logic. FIR was applied to two different courses: the *Didactic Planning* and the *Introductory* courses, both of the Centre of Studies in Communication and Educational Technologies virtual campus (CECTE), in Mexico. The results obtained in this work indicated that FIR is a good alternative to predict accurately the final mark of the students. On the other hand, FIR allowed to determine the more relevant features, reducing considerably system's complexity.

The introductory course was also used to study the dynamic assessment of students' learning performance. The aim of this study was to dynamically forecast students' learning performance in order to help them to prevent possible introductory course failing. The knowledge derived from the assessments is useful for teachers to give feedback to predicted failing students during the course, in such a way that the students can react and perform better the rest of the course. Therefore, it is of major importance to obtain reliable predictions early in time in order to give feedback to the students as soon as possible, encouraging students to work harder and pass the course.

Second, a methodology for assessing the relative Causal Relevance (CR) of individual data features involved in the inferred system model for reducing the uncertainty during the forecasting stage and data understanding purposes was developed.

As has been explained in Chapter 3, the FIR qualitative modeling process has the task of identifying the most important relations (temporal and spatial) between variables that represent the system in an optimal manner. The *mask* captures the result of this task, containing the more relevant relations between the system's variables. Therefore, a feature selection is performed in FIR methodology by means of the qualitative modeling function. However, a more accurate and detailed study can be done after this first stage. The question now is: Do all of the variables that appear in the mask affect the prediction

of the output to the same degree? Is it justifiable to treat all m-inputs as equals? It is quite obvious that the answer of the previous two questions depends on the particular application at hand. However in general terms, the answer should be no. Usually, some m-inputs are more relevant than others from the causality point of view. It thus becomes an interesting question to know these relevancies and to quantify them somehow. The CR concept addresses this issue by quantifying the influence of each m-input with respect to the output by using different metrics. The first one of the proposed metrics takes into account the quality of the mask, and studies the *direct* and *indirect* influence of a specific m-input to the prediction performance. The second one of the proposed metrics takes into account the prediction error, and studies, also, the *direct* and *indirect* approaches.

Four different applications were used in this thesis to study the validity of the CR concept, one in the biology area, two in the medical field, and the fourth a simple linear system. The results obtained with these applications show that the CR helped to improve the prediction of all the applications studied. In most cases, the improvement was quite significant. Afterwards, The CR-FIR was applied to educational data, in concrete to the *Introductory* course of the CECTE virtual campus. The new approaches help to improve the understanding of the educative process by describing how much influence each system feature has on the predicted output. The results obtained from the experiments performed, prove the feasibility of using the causal relevancy approaches proposed in this study, especially when the prediction results obtained by means of the FIR inference engine (without causal relevancy) are not good enough.

The advantages of using the CR-FIR in e-Learning environments would be fully experienced when dealing with data sets of much higher dimensionality than the one used in this study.

Third, in this thesis a novel rule extraction algorithm LR-FIR (linguistic rules in FIR), was developed based on fuzzy logic, that is able to derive linguistic rules from a FIR model. The LR-FIR functioning is similar to those used in Boolean algebra. However the premises and consequences of rules are not necessarily binary in nature, hence the algorithm is able to deal with multi-valued logic, and accept partial do-not-care conditions. Due to the fact that LR-FIR was developed within the FIR methodology, the obtained rules could be considered as predictive rules and deal naturally with the uncertainty captured in the FIR models.

The LR-FIR algorithm proposed is developed with the goal to be a useful tool for decision makers. With this purpose in mind the rules extracted by LR-FIR describe in a very intuitive and actionable way the system behavior. LR-FIR extracts predictive rules of the type IF-THEN and allows to represent multi-valued logic functions, i.e. neither inputs nor outputs are restricted to binary logic. Instead to avoid overlapping rules, these rules are treated in the compaction and unification steps where rules sharing contiguous input spaces in a feature and the same values in the remaining features are unified in a

unique rule. In this way LR-FIR represents as much accurately as possible the system behavior while preserving the main goal, i.e. the simplicity of the resulting rule base.

LR-FIR was evaluated using five data-sets from different domains: e-Learning, global change temperature, brain tumour diagnosis, and two of the most used classical UCI data-sets: IRIS and Pima Indian Diabetes. The rules extracted by LR-FIR capture the main behavior of each application, from the domain experts' point of view, demonstrating in this sense, the efficiency of the proposed algorithm.

Fourth, a study of the use of the Generative Topographic Mapping (GTM) approach to the analysis of students that show a non-typical learning behavior is also performed in this thesis. The presence of atypical observations, or outliers, in a data set can distort the results obtained from their analysis. Therefore, the data analyst would benefit from models that behave robustly in the presence of outliers. A constrained mixture of t distributions: the t-GTM, has been studied in this dissertation. It simultaneously provides robust data clustering and visualization of the results, which become intuitively interpretable. It also effectively neutralizes the negative effects of outliers. In the current study, data obtained from real virtual campuses e-Learning students' experiences have been analyzed, and data outliers corresponding to students' atypical online behaviors have been identified and characterized, illustrating the tGTM model capabilities. Two different courses of real virtual campuses were modeled, i.e. *Compilers I* course of the Open University of Catalonia (UOC) in Barcelona and *Didactic Planning* course of the Centre of Studies in Communication and Educational Technologies virtual campus (CECTE) in Mexico.

The experimental results have shown that useful knowledge can be extracted from the tGTM combination of outlier detection and data clustering and visualization. This knowledge can be used for real time student personalized guidance, and to help teachers to find patterns of student behavior.

Finally, in this thesis a framework to provide real time useful knowledge to e-Learning environments and improve the e-Learning experience was developed.

The main goal of the framework is to alleviate the virtual tutors' workload and to provide an effective and valuable feedback to learners. To deal with these objectives the framework offers tools to discover relevant learning behavior patterns from students' interaction with the educational materials. The knowledge obtained can be used by teachers to design courses more effectively and detect students with learning difficulties. The knowledge extracted can also be helpful for the students to know their own learning performance and therefore use more efficiently the educational resources. The soft computing methodologies (FIR, LR-FIR, CR-FIR and tGTM) that are the data mining core of the framework are able to offer valuable knowledge to both, teachers and students that can be used to enhanced course performance and that opens new possibilities for the pedagogical and instructional designers, who create and organize the learning contents.

In this thesis the framework developed was tested with the *Didactic Planning* course of the CECTE. Much more evaluations are needed to adjust the framework and enhance all the issues that for sure will arise during its use.

Summarizing, the major contributions of this doctoral thesis are the following:

- A survey of the data mining techniques that have been applied to e-Learning problems and of the e-Learning problems to which data mining techniques have been applied.
- An analysis of the feasibility of FIR methodology for modeling and prediction of students' performance.
- An algorithm, CR-FIR, for assessing the relative causal relevance of individual data features involved in the inferred system model for reducing the uncertainty during the forecasting stage and data understanding purposes.
- An algorithm, LR-FIR, for rule extraction, in the context of FIR methodology, for data mining and knowledge discovery that enhances the system characterization and facilitates decision making.
- A study of the use of Generative Topographic Mapping for the analysis of atypical student behavior.
- A framework to provide real time useful knowledge to e-Learning environments and improve the e-Learning experience.
 - Help to prevent students getting failing grades by forecasting students' performance in real time along the course duration.
 - Alleviate teachers' workload by providing a suitable way to assess the importance of each evaluation parameter in the learning process.
 - Provide valuable knowledge to the teachers in order to better understand the students' learning behavior patterns, and take into consideration this knowledge in the decision making processes

4.2 Future research

Although we think that we are going in the right direction, in the sense that the work done in this doctoral thesis is a significant step towards the improvement of e-Learning systems and that the framework proposed can help teachers and students providing useful knowledge, there is still a long way to go.

The first think that should be performed in the near future is to validate more extensively, in the context of the e-Learning environment, the two algorithms developed in this dissertation, i.e. CR-FIR and LR-FIR. As explained before, both were proved broadly using different kind of applications, but due to limited access to educational data, only three e-Learning sets of data from real virtual campuses were used in this research.

It is necessary to go a step further and apply both methodologies to e-Learning data sets with much more variables involved and a larger number of registers.

At the moment, the e-Learning framework presented is functioning only on the CECTE intranet; however it is necessary in the near future to implement several plug-ins to allow the connection with the most known e-Learning platforms, in such a way that all the educative institutions interested in using this framework can do it.

We hope to gain further insights into how the system works in practical everyday usage, and use this feedback to improve the framework. User opinions should be added to the project. The development methodology applied in the proposed framework, handles to see bugs and application problems during the development stage. Therefore, it is easy to perform changes or improvements.

The possibility of tracking user behavior in virtual campus e-Learning environments makes possible the web mining of the resulting databases. This opens new possibilities for the pedagogical and instructional designers who create and organize the learning contents. One of the most interesting ones is the personalization of the e-Learning process.

On the other hand, some security considerations should be taken into account in a future version of the framework. Right now, the implementation of the toolbox servlets makes it possible to execute all code on the host computer. MATLAB should run in a secure shell and access to the servlet should be restricted to users on a password basis in order to compensate for the security hazards. Moreover, in order to provide more secure access processes we propose to include Remote Method Invocation (RMI). To afford a standard and more open way to access, we are planning to offer the platform as a web service.

Another interesting topic for additional research is the development of a specific Moodle data mining tool, to be used by on-line instructors, which will eliminate the need for Content Management Systems (CMS) administrators to help the instructors to pre-process or to apply data mining techniques. It should have an intuitive and user-friendly interface and automatically pre-processes Moodle data, making it easier to configure and execute due to its parameter-free data mining algorithms. This tool should be integrated into the Moodle environment itself as another Moodle author tool such as Graphical Interactive Student Monitoring (GISMO). In this way, instructors can both create/maintain courses and carry out all data mining processing in the same interface. Likewise, they can directly apply feedback and results obtained by data mining into Moodle courses.

An almost unexplored and relevant research path in DM for e-Learning is that of the application of methods for the explicit analysis of time series. That would be really

appropriate since much of the information that could be gathered from e-Learning systems usage takes precisely this form.

We hope that, with this thesis, we have provided a significant contribution to the fields of data-mining and e-Learning, and that our results will prove to be useful for many other researchers dealing with learning analytics and knowledge extraction in e-Learning environments.

Chapter 5: Bibliography and References

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.

Abu-Naser, S., Al-Masri, Abu-Sultan, Y., Zaqout, I.: A Prototype Decision Support System for Optimizing the Effectiveness of Elearning in Educational Institutions. In: International Journal of Data Mining & Knowledge Management Process (IJDKP) 1(4) (2011) .

Aher, S.B., Lobo, L.M.R.J.: Combination of Machine Learning Algorithms for Recommendation of Courses in E-Learning System based on Historical Data. Knowledge-Based Systems 51 (2013) 1-14.

AlAjmi, M.A., Khan, S., Zamani, A.S.: Using Instructive Data Mining Methods to Revise the Impact of Virtual Classroom in E-Learning. International Journal of Advanced Science and Technology 45 (2012) 125-134.

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, AIMS. October 12-15, Seattle, USA (2003).

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.

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.

Baker, R.S.J., Corbett, A.T., Gowda, S.M., Wagner, A.Z., MacLaren, B.A., Kauffman, L.R., Mitchell, A.P., Giguere, S.: Contextual Slip and Prediction of Student Performance after Use of an Intelligent Tutor. In: De Bra et al., (eds.): International Conference on User Modeling, Adaptation, and Personalization, UMAP2010. Lecture Notes in Computer Science, Vol. 6075. Springer-Verlag, Berlin, Heidelberg (2010) 52-63.

Baradwaj, B.K., Pal, S.: Mining Educational Data to Analyze Students' Performance. International of Advanced Computer Science and Applications 2(6) (2011) 63-69.

Barahate, S.R., Shelake, V.M.: A Survey and Future Vision of Data Mining in Educational Field. In: Second International Conference on Advanced Computing and Communication Technologies (2012) 96-100.

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.

Biswas, G., Leelawong, K., Belyne, 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).

Brusilovsky, P.: Adaptive Hypermedia. User Modelling and User Adapted Interaction, Ten Year Anniversary Issue 11(1-2) (2001) 87-110.

Buldu, A., Üçgün, K.: Data Mining Application on Students' Data. Procedia-Social and Behavioral Sciences 2(2) (2010) 5251-5259.

Burgos, C., Campanario, M.L., De la Peña, D., Lara, J.A., Lizcano, D. Martínez, M.A.: Data Mining for Modeling Students' Performance: A Tutoring Action Plan to Prevent Academic Dropout. Computers and Electrical Engineering, In Press, (2017).

- Caballé, S., Paniagua, C., Xhafa, F., Daradoumis, T.: Grid-Aware Implementation for Providing Effective Feedback to On-Line Learning Groups. In: Meersman, R. et al. (eds.): OTM Workshops 2005, LNCS 3762, Springer-Verlag (2005) 274-283.
- 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).
- 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.
- 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.
- 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.
- Castro, F., Vellido, A., Nebot, A., Minguillón, J.: Detecting Atypical Student Behavior 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.
- Castro, F., Vellido, A., Nebot, A., Mugica, F.: Applying Data Mining Techniques to e-Learning Problems: a Survey and State of the Art. Evolution of Teaching and Learning Paradigms in Intelligent Environment. *Studies in Computational Intelligence*, 62, (2007) 183-221.
- 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.
- 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.
- 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.
- Cheng, B., Wang, M., Yang, S.J.H., Peng, J.: Acceptance of Competency-based Workplace e-Learning Systems: Effects of Individual and Peer Learning Support. *Computers & Education* 57(1) (2011) 1317-1333.
- Cheng, B., Wang, M., Morch, A.I., Chen, N.S., Kinshuk, Spector, J.M.: Research on e-Learning in the Workplace 2000-2012: A Bibliometric Analysis of the Literature. *Educational Research Review* 11 (2014) 56-72.
- Cho, H., Gay, G., Davidson, B., Ingraffea, A.: Social Networks, Communication Styles and Learning Performance in a CSCL Community. *Computers & Education* 49(2) (2007) 309-329.
- Chrysafiadi, K., Virvou, M.: Student Modeling Approaches: A Literature Review for the Last Decade. *Expert Systems with Applications* 40 (2013) 4715-4729.
- 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.
- Cela, K.L., Sicilia, M.A., Sanchez, S.: Social Network Analysis in e-Learning Environments: A Preliminary Systematic Review. *Educ. Psych. Rev. Springer Science-Business Media, New York* (2014) 219-246.

Cellier, F.E., Yandell, D.W.: SAPS-II: A New Implementation of the Systems Approach Problem Solver. *International Journal of General Systems* 13 (1987) 307-322.

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.

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/>

de Albornoz, A.: Inductive Reasoning and Reconstruction Analysis: Two Complementary Tools for Qualitative Fault Monitoring of Large-scale Systems. Ph.D. Thesis, Technical University of Catalonia, Barcelona, Spain (1996).

Dejaeger, K., Verbeke, W., Martens, D., Baesens, B.: Data Mining Techniques for Software Effort Estimation: A Comparative Study. *IEEE Transactions on Software Engineering* 38(2) (2012) 375-397.

Desmarais, M.C., Baker, R.S.J.: A Review of Recent Advances in Learner and Skill Modeling in Intelligent Learning Environments. *User Model User-Adap Inter* 22 (2012) 9-38.

Dogan, B., Camurcu, A.Y.: Visual Clustering of Multidimensional Educational Data From an Intelligent Tutoring System. *Computer Applications in Engineering Education* 18(2) (2010) 375-382.

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.

Drachsler, H., Verbert, K., Santos, O.C., Manouselis, N.: Panorama of Recommender Systems to Support Learning. Ricci, F. et al. (eds.): *Recommender Systems Handbook*, Springer-Verlag (2015) 421-451.

Dreyfus, S.E., Law, A.M.: *The Art and Theory of Dynamic Programming*. Academic Press Inc. New York (1977).

Drigas, A., Vrettaros, J.: An Intelligent Tool for Building e-Learning Content-Material Using Natural Language in Digital Libraries. *WSEAS Transactions on Information Science and Applications* 5(1) (2004) 1197-1205.

Dwivedi, S.K., Rawat, B.: An Architecture for Recommendation of Courses in e-Learning. *IJ. Information Technology and Computer Science* (4) (2017) 39-47.

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.

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.

Fasuga, R., Sarmanova, J.: Usage of Artificial Intelligence in Education Process. In: *International Conference for Engineering Education & Research, ICEER2005*. Tainan, Taiwan (2005).

Fazlollahtabar, H., Mahdavi, I.: User/Tutor Optimal Learning Path in e-Learning Using Comprehensive Neuro-Fuzzy Approach. *Educational Research Review* 4(2) (2009) 142-155.

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.

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.

- Franco-Lugo, M.J., Von Lücken, C., Ramírez, E.E.: Sequencing Educational Contents Using Clustering and Ant Colony Algorithms. In: V.L. Uskov et al. (eds.): Smart Education and e-Learning 2016, Smart Innovation, Systems and Technologies 59, (2016) 375-385.
- García, E., Romero, C., Ventura, S., de Castro, C.: A Collaborative Educational Association Rule Mining Tool. *The Internet and Higher Education* 14(2) (2011) 77-88.
- Goel, G., Lallé, S., Luengo, V.: Fuzzy Logic Representation for Student Modelling Case Study on Geometry. In Proceedings of the 11th international conference on intelligent tutoring systems. Chania, Greece (2012) 428-433.
- 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.
- 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.
- 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).
- 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.
- He, Y., Hui S.C., Quan, T.T.: Automatic Summary Assessment for Intelligent Tutoring Systems. *Computers & Education* 53 (2009) 890-899.
- Hegazi, M.O., Abugroon, M.A.: The State of the Art on Educational Data Mining in Higher Education. *International Journal of Computer Trends and Technology (IJCTT)* 31(1) (2016).
- 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.
- Herrera, F., Carmona, C.J., González, P. et al. *Knowl Inf Syst* (2011) 29: 495. <https://doi.org/recursos.biblioteca.upc.edu/10.1007/s10115-010-0356-2>
- Hogo, M.A.: Evaluation of e-Learning Systems Based on Fuzzy Clustering Models and Statistical Tools. *Expert Systems with Applications* 37(10) (2010) 6891-6903.
- 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).
- 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.
- Hwang, G.J.: A Knowledge-Based System as an Intelligent Learning Advisor on Computer Networks. *J. Systems, Man, and Cybernetics* 2 (1999) 153-158.
- Hwang, G.J.: A Test-Sheet-Generating Algorithm for Multiple Assessment Requirements. *IEEE Transactions on Education* 46(3) (2003) 329-337.
- 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.
- 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.

- 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.
- 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.
- Inventado, P.S., Legaspi, R., Bui, T.D., Suarez, M.: Predicting Student's Appraisal of Feedback in an ITS Using Previous Affective States and Continuous Affect Labels from EEG Data. In: Proceedings of the 18th International Conference on Computers in Education, Putrajaya, Malaysia (2010) 71-75.
- Jeremic, Z., Jovanovic, J., Gasevic, D.: Student Modeling and Assessment in Intelligent Tutoring of Software Patterns. *Expert Systems with Applications* 39(1) (2012) 210-222.
- 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).
- Jia, B., Zhong, S., Zheng, T., Liu, Z.: The Study and Design of Adaptive Learning System Based on Fuzzy Set Theory. In: Cheok, Z.A.D. et al., (eds.): Transactions on edutainment IV. Berlin, Heidelberg: Springer-Verlag (2010) 1-11.
- Jeong, H., Biswas, G.: Mining Student Behavior Models in Learning-by-Teaching Environments. In: The 1st International Conference on Educational Data Mining, Montreal, Québec, Canada, June 20-21 (2008) 127-136.
- Jerez, A., Nebot, A.: Genetic Algorithms Versus Classical Search Techniques for Identification of Fuzzy Models. In: Proceedings of the 1997 European Congress on Intelligent Techniques and Soft Computing, Aachen, Germany (1997) 769-773.
- Jurado, F., Santos, O.C., Redondo, M.A., Boticario, J.G., Ortega, M.: Providing Dynamic Instructional Adaptation in Programming Learning. In: Proceedings of the 3rd international Workshop on Hybrid Artificial Intelligence Systems. Burgos, Spain (2008) 329-336.
- Jurado, F., Redondo, M.A., Ortega, M.: Using Fuzzy Logic Applied to Software Metrics and Test Cases to Assess Programming Assignments and Give Advice. *Journal of Network and Computer Applications* 35 (2012) 695-712.
- Kardan, A.A., Sadeghi, H., Ghidary, S.S., Reza, M., Sani, F.: Prediction of Student Course Selection in Online Higher Education Institutes Using Neural Network. *Computers & Education* 65 (2013) 1-11.
- Kato, T., Kambayashi, Y., Kodama, Y.: Data Mining of Students' Behaviors in Programming Exercises. In: V.L. Uskov et al. (eds.): Smart Education and e-Learning 2016, Smart Innovation, Systems and Technologies 59 (2016) 121-133.
- Klir, G.J.: An Approach to General System Theory. Van Nostrand Reinhold Company, New York (1969).
- Klir, G. J.: Architecture of Systems Problem Solving. New York, Plenum Press (1985).
- Klir, G.J., Folger, T.A.: Fuzzy Sets, Uncertainty and Information. Englewood Cliffs, NJ: Prentice-Hall (1988).
- Kohonen, T.: Self-Organizing Maps. 3rd edition, Springer-Verlag, Berlin (2000).
- Kolo, D.K., Adepoju, S.A., Alhassan, J.K.: A Decision Tree Approach for Predicting Students Academic Performance. *International Journal of Education and Management Engineering (IJEME)* 5(5) (2015) 12-19.
- 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.
- Kotsiantis, S., Patriarcheas, K., Xenos, M.: A Combinational Incremental Ensemble of Classifiers as a Technique for Predicting Students' Performance in Distance Education. *Knowledge-Based Systems* 23 (2010) 529-535.

- 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.
- Leidig, T.: L3-Towards an Open Learning Environment. *ACM Journal of Educational Resources in Computing, JERIC*. ACM Press 1(1) (2001).
- Le Ru, Y., Aron, M., Gerval, J.P., Napoleon, T.: Tests Generation Oriented Web-Based Automatic Assessment of Programming Assignments. In: V.L. Uskov et al. (eds.): *Smart Education and e-Learning 2016, Smart Innovation, Systems and Technologies 41* (2015) 117-127.
- Li, D., Cellier, F.E.: Fuzzy Measures in Inductive Reasoning. *Proceedings of the 1990 Winter Simulation Conference*. New Orleans (1990) 527-538.
- 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.
- Liccheli, 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.
- Lockyer, L., Heathcote, E., Dawson, S.: Informing Pedagogical Action: Aligning Learning Analytics With Learning Design. *American Behavioral Scientist* 57(10) (2013) 1439-1459.
- Lorenzetti, J. P.: How e-learning is changing higher education: a new look. *Distance Education Report* 22 (2005) 4-7.
- Lykourantzou, I., Giannoukos, I., Nikolopoulos, V., Mpardis, G., Loumos, V.: Dropout Prediction in e-Learning Courses through the Combination of Machine Learning Techniques. *Computers & Education* 53(3) (2009) 950-965.
- Margo, H.: Data Mining in the e-Learning Domain. *Computers & Education* 42(3) (2004) 267-287.
- 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.
- 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).
- 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.
- Maull, K. E., Saldivar, M. G., & Sumner, T.: Online Curriculum Planning Behavior of Teachers. In: *The 3rd International Conference on Educational Data Mining*, Pittsburgh, PA, USA, June 11-13 (2010) 121-130.
- 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, GECCO2003* (2003) 2252-2263.
- 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).
- 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.
- Mohamad, S.K., Tasir, Z.: Educational Data Mining: A Review. *Procedia - Social and Behavioral Sciences* 97 (2013) 320-324.

- Monk, D.: Using Data Mining for e-Learning Decision Making. *The Electronic Journal of e-Learning* 3 (2005) 41-54.
- Moodle.: Learning Analytics. (2018) Retrieved from: <https://docs.moodle.org/34/en/Analytics>, 30/172018 (1/2/2018)
- 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.
- Mostow, J., Beck, J.: Some Useful Tactics to Modify, Map and Mine Data from Intelligent Tutors. *Natural Language Engineering* 12 (2006) 195-208.
- 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.
- Mugica, F.: *Diseño Sistemático de Controladores Difusos Usando Razonamiento Inductivo*. Ph.D. thesis, Technical University of Catalonia, Barcelona, Spain (1995).
- Mullier, D.: A Tutorial Supervisor for Automatic Assessment in Educational Systems. *International Journal on e-Learning* 2(1) (2003) 37-49.
- 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.
- Nebot, A.: *Qualitative Modeling and Simulation of Biomedical Systems Using Fuzzy Inductive Reasoning*. Ph.D. thesis, Technical University of Catalonia, Barcelona, Spain (1994).
- Nebot, A., Mugica, F., Cellier, F.E., Vallverdú, M.: Modeling and Simulation of the Central Nervous System Control with Generic Fuzzy Models. *Simulation: Transactions of the Society for Modeling and Simulation* 79 (2003) 648-669.
- 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.
- Nebot, A., Cellier, F.E., Carvajal, R., Mugica, F.: Fuzzy Inductive Reasoning for Variable Selection Analysis and Modeling of Biological Systems. *International Journal of General Systems* 38 (2009) 793-811.
- Nebot, A., Mugica, F.: Fuzzy Inductive Reasoning: a Consolidated Approach to Data-Driven Construction of Complex Dynamical Systems. *International Journal of General Systems* 41(7) (2012) 645-665. Invited survey paper.
- Novak, J.D., Gowin, B.C.: *Concept Mapping for Meaningful Learning*. In: *International Conference on Learning*. Cambridge University Press (1984) 36-37.
- Papamitsiou, Z., Economides, A.: Learning Analytics and Educational Data Mining in Practice: A Systematic Literature Review of Empirical Evidence. *Educational Technology & Society* 17(4) (2014) 49-64.
- Peña-Ayala, A.: Educational Data Mining: A Survey and a Data Mining-based Analysis. *Expert Systems with Applications* 41 (2014) 1432-1462.
- Ogor, E.M.: Student Academic Performance Monitoring and Evaluation Using Data Mining Techniques. *Conference: Electronics, Robotics and Automotive Mechanics Conference (CERMA 2007)* 25-28 Sept (2007) 354-359.
- Özpolat, E., Akar, E.: Automatic Detection of Learning Styles for an e-Learning System. *Computers & Education* 53(2) (2009) 355-367.

- 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.
- Parmar, N., Robinson, K., Salter, M.: Library and e-Learning: Further Adventures with Augmented Reality, In: Needham, G, Ally, M. (eds.): M-Libraries 5: From Devices to People. Facet Publishing (2015) 183-199.
- 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.
- Rajibussalim, R.: Mining Students' Interaction Data from a System that Support Learning by Reflection. The 3rd International Conference on Educational Data Mining, Pittsburgh, PA, USA, June 11-13, (2010) 249-256.
- 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.
- 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).
- 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.
- 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.
- 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).
- 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.
- 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.
- Romero, C., Ventura, S.: Data Mining in e-Learning. *WIT* (2006).
- Romero, C., Ventura, S.: Educational Data Mining: A Survey from 1995 to 2005. *Expert Systems with Applications* 33(1) (2007) 135-146.
- Romero C., González P., Ventura S., del Jesus MJ., Herrera F.: Evolutionary algorithm for subgroup discovery in e-learning: a practical application using Moodle data. *Expert Syst Appl* 36:1632–164 (2009).
- Romero, C, Ventura, S.: Educational Data Mining: A Review of the State of the Art. *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews)*, 40(6) (2010) 601-618.
- Romero, C., Romero, J.R., Luna, J.M., Ventura, S.: Mining rare association rules from e-learning data. In: The 3rd International Conference on Educational Data Mining (2010) 171-180.
- Carmona, C.J., González, P., del-Jesus, M.J., Romero, C., Ventura, S.: Evolutionary algorithms for subgroup discovery applied to e-learning data. In: IEEE EDUCON Education Engineering 2010 -The Future of Global Learning Engineering Education (2010) 983-990.
- Rosmalen, P.V., Brouns, F., Tattersall, C., Vogten, H., Bruggen, J.V., Sloep, P., Koper, R.: Towards an Open Framework for Adaptive Agent-Supported e-Learning. In: *International Journal Continuing Engineering Education and Lifelong Learning* 15(3-6) (2005) 261-275.

- Samigulina, G., Shayakhmetova, A.: The Information System of Distance Learning for People with Impaired Vision on the Basis of Artificial Intelligence Approaches. In: Uskov, V.L., et al. (eds.): Smart Education and e-Learning 2016, Smart Innovation, Systems and Technologies 41 (2015) 255-263.
- Shachter, R.D.: Evaluating Influence Diagrams. *Operating Research* 34 (1986) 871-882.
- Sheard, J., Ceddia, J., Hurst, G.: Inferring Student Learning Behavior from Website Interactions: A Usage Analysis. *Education and Information Technologies* 8(3) (2003) 245-266.
- Schiaffino, S., Garcia, P., Amandi, A.: eTeacher: Providing Personalized Assistance to e-Learning Students. *Computers & Education* 51 (2008) 1744-1754.
- Sen, B., Ucar, E., Denle, D.: Predicting and Analyzing Secondary Education Placement-Test Scores: A Data Mining Approach. *Expert Systems with Applications* 39(10) (2012) 9468-9476.
- 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.
- Shang, Y., Shi, H., Chen, S.S.: An Intelligent Distributed Environment for Active Learning. *ACM Journal of Educational Resources in Computing* 1(2) (2001).
- Shannaq, B., Rafael, Y., Alexandro, V.: Student Relationship in Higher Education Using Data Mining Techniques. *Global Journal of Computer Science and Technology* 10(11) (2010) 54-59.
- Sharma, R., Singh, H.: Data Mining in Education Sector. *International Journal of Electronics & Data Communication* 2(1) (2013) 4-8.
- 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.
- Sison, R., Shimura, M.: Student Modelling and Machine Learning. *International Journal of Artificial Intelligence in Education* 9 (1998) 128-158.
- 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.
- 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.
- 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 conjunction with 16th European Conference on Artificial Intelligence, ECAI'2003. Valencia, Spain (2004) 17-22.
- 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.
- 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.
- 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.
- Taylan, O., Karagözoglu, B.: An Adaptive Neuro-Fuzzy Model for Prediction of Student's Academic Performance. *Computers & Industrial Engineering* 57 (2009) 732-741.
- 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.

- 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.
- 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.
- 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.
- Ueno, M.: LMS with Irregular Learning Processes Detection System. In: World Conference on e-Learning in Corp., Govt., Health, & Higher Education (2003) 2486-2493.
- 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.
- Uskov, V., Uskov, A.: Web-Based Education: Strategic Issues for 2008-2015. *Advanced Technology for Learning* 5(1) (2008) 1-11.
- Uyttenhove, H.J.: SAPS: System Approach Problem Solver. Ph.D. thesis, SUNY, Binghamton, NY (1979).
- 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>
- 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.
- Vázquez-Barreiros, B., Lama, M., Mucientes, Vidal, J.C.: SoftLearn: A Process Mining Platform for the Discovery of Learning Paths. In: Proceedings of the IEEE 14th International Conference on Advanced Learning Technologies (ICALT), Athens, Greece (2014) 373-375.
- 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.
- Velmurugan, K.R., Saravanan M.R. Certain Investigations on Big Data Approaches in Education and Learning Analytics. *The IIOAB Journal* 7(9) (2016) 457-462.
- Vellido, A., Castro, F., Nebot, A.: Clustering Educational Data. In: *Handbook of Educational Data Mining*, Chapter 6 (2011) 75-92.
- Vialardi, C., Chue, J., Peche, J.P., Alvarado, G., Vinatea, B., Estrella, J., Ortigoza, A.: A Data Mining Approach to Guide Students through the Enrollment Process based on Academic Performance. *User Model User-Adap. Inter.* 21 (2011) 217-248.
- 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.
- Wang, T.H.: Developing an Assessment-centered e-Learning System for Improving Student Learning Effectiveness. *Computers & Education* 73 (2014) 189-203.
- 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.
- Weber, G., Brusilovsky, P.: ELM-ART - An Interactive and Intelligent Web-Based Electronic Textbook. *International Journal of Artificial Intelligence in Education* 26 (1) (2016) 72-81.

- Weng, C.H.: Mining Fuzzy Specific Rare Itemsets for Education Data. *Knowledge-Based Systems* 24(5) (2011) 697–708.
- X. Wang, Y. Yang, X. Wen: Study on Blended Learning Approach for English Teaching. In: *IEEE International Conference on Systems, Man and Cybernetics* (2009) 4641 – 4644.
- Yadav, R.S., Singh, V.P.: Modeling Academic Performance Evaluation Using Soft Computing Techniques: A Fuzzy Logic Approach. *International Journal on Computer Science and Engineering (IJCSE)* 3(2) (2011) 676-686.
- Yadav, R.S., Soni, A.K., Pal, S.: A Study of Academic Performance Evaluation Using Fuzzy Logic Techniques. In: *The 2014 International Conference on Computing for Sustainable Global Development (INDIACom)* (2014) 48-53.
- 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
- Zaïane, O.R.: Building a Recommender Agent for e-Learning Systems. In: *The International Conference on Computers in Education, ICCE'02* (2002) 55-59.
- 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.
- Zaki, N.: PageRank Algorithm to Improve the Peer-Led Team Learning Pedagogical Approach. In: V.L. Uskov et al. (Eds.): *Smart Education and e-Learning 2016, Smart Innovation, Systems and Technologies* 59 (2016) 225-234.
- Zhang, L., Zhang, X., Duan, Y., FU, Z., Wang, Y.: Evaluation of Learning Performance of E-Learning in China: A Methodology Based on Change of Internal Mental Model of Learners. *TOJET: The Turkish Online Journal of Educational Technology* 9(1) (2010) 70-82.
- Zorrilla, M., García, D., Álvarez, E.: A Decision Support System to improve e-Learning Environments. *International Workshop on Business intelligence and the WEB* (2010).

Chapter 6: Copy of the Articles Derived from the Doctoral Thesis

6.1 Articles derived from Objectives 1 and 6: Perform a detailed review of the data mining techniques that have been applied to e-Learning problems and an analysis of the e-Learning problems to which data mining techniques have been applied.

A. Vellido, F. Castro, A. Nebot. Clustering Educational Data. Handbook of Educational Data Mining, Chapter 6, pp. 75-92, 2011.

F. Castro, A. Vellido, A. Nebot, F. Mugica. Applying Data Mining Techniques to e-Learning Problems: a Survey and State of the Art. Evolution of Teaching and Learning Paradigms in Intelligent Environment. Studies in Computational Intelligence, 62, pp. 183-221, 2007.

6.2 Articles derived from Objective 2: Demonstrate the feasibility of FIR methodology for modeling and prediction of students' performance.

A. Nebot, F. Castro, F. Mugica, A. Vellido. Identification of fuzzy models to predict students performance in an e-Learning environment. Proceedings of the WBE'06: Fifth IASTED International Conference on Web-based Education, pp. 74-79, 2006.

A. Nebot, F. Mugica, F. Castro, J. Acosta. Genetic Fuzzy System for Predictive and Decision Support Modelling in e-Learning. Proceedings of the WCCI'10: IEEE World Congress on Computational Intelligence, pp. 1804-1811, 2010.

A. Nebot, F. Mugica, F. Castro. Fuzzy Predictive Models to Help Teachers in e-Learning Courses. Proceedings of the WCCI'10: IEEE World Congress on Computational Intelligence, pp. 2077-2083, 18-23 July 2010.

6.3 Articles derived from Objective 3: Design and development of an algorithm, in the context of FIR methodology, to provide a quantitative method for assessing the relative causal relevance of individual data features involved in the inferred system model for reducing the uncertainty during the forecasting stage and data understanding purposes.

A. Nebot, F. Mugica, F. Castro. Causal Relevance to Improve the Prediction Accuracy of Dynamical Systems using Inductive Reasoning, International Journal of General Systems, 38 (3), pp. 331-358, 2009.

F. Castro, A. Nebot, F. Mugica. Causal Relevancy Approaches to Improve the Students' Prediction Performance in an e-Learning Environment. Proceedings of the MICAI'07: 6th Mexican International Conference on Artificial Intelligence, pp.: 342-351, 4-10, 2007.

6.4 Articles derived from Objective 4: Design and development of a rule extraction algorithm, in the context of FIR methodology, for data mining and knowledge discovery that enhances the system characterization and facilitates decision making.

F. Castro, A. Nebot, F. Mugica. On the extraction of decision support rules from fuzzy predictive models, Applied Soft Computing, 11 (4), pp. 3463–3475, 2011.

A. Nebot, B. Martínez, F. Castro, C. Gay. Fuzzy Rules Model for Global Warming Decision Support. Proceedings del VI Congreso Internacional de la Asociación Española de Climatología: Cambio Climático Regional y sus Impactos, pp. 565-576, 2008.

A. Nebot, F. Castro, A. Vellido, M. Julià-Sapé, C. Arús. Rule-based assistance to brain tumour diagnosis using LR-FIR. Lecture Notes in Artificial Intelligence, LNAI-5178, pp. 173-180, 2008.

A. Nebot, F. Castro, A. Vellido, M. Julià-Sapé, C. Arús. Rule-based assistance to brain tumour diagnosis using LR-FIR. Investigating Human Cancer with Computational Intelligence Techniques, Chapter 7, pp. 83-92, 2009.

The previous two articles have the same content because the first was published in LNAI as the proceedings of the 12th International Conference, KES 2008 Zagreb, Croatia, September 3-5, 2008, and it was selected from all the conference articles to be published in the book *Investigating Human Cancer with Computational Intelligence Techniques*, KES International. Therefore, we only include the second one here.

F. Castro, A. Nebot, F. Mugica. Extraction of Logical Rules to Describe Students' Learning Behavior, Proceedings of the WBE'07: 6th IASTED International Conference on Web-based Education, pp. 164-169, 2007.

A. Vellido, F. Castro, T.A. Etchells, A. Nebot, F. Mugica. Data Mining of Virtual Campus Data: a Case Study on Clustering, Visualization, Outlier Detection, Classification, Feature Selection, and Rule Extraction. Evolution of Teaching and Learning Paradigms in Intelligent Environment. Studies in Computational Intelligence, 62, pp. 223-254, 2007.

A. Nebot, F. Mugica, F. Castro, J. Acosta. Genetic Learning of Fuzzy Parameters in Predictive and Decision Support Modelling, International Journal of Computational Intelligence Systems, 5 (2), pp. 387-402, 2012.

6.5 Articles derived from Objective 5: Study of the use of Generative Topographic Mapping for the analysis of atypical student behavior.

A. Vellido, F. Castro, T.A. Etchells, A. Nebot, F. Mugica. Data Mining of Virtual Campus Data: a Case Study on Clustering, Visualization, Outlier Detection, Classification, Feature Selection, and Rule Extraction. Evolution of Teaching and Learning Paradigms in Intelligent Environment. Studies in Computational Intelligence, 62, pp. 223-254, 2007.

The book chapter above also appears in objective #4 because has contributions to both goals. Therefore, the full text has been included already in the previous section (6.4).

F. Castro, A. Vellido, A. Nebot, J. Minguillón. Detección de estudiantes con comportamiento atípico en entornos de aprendizaje e-Learning. TICs Aplicadas a la enseñanza/aprendizaje de la Ingenierías, pp. 23-30, 2006.

A. Vellido, F. Castro, A. Nebot, F. Mugica. Characterization of atypical virtual campus usage behavior through robust generative relevance analysis. Proceedings of the WBE'06: Fifth IASTED International Conference on Web-based Education, pp. 183-188, 2006.

F. Castro, A. Vellido, A. Nebot, J. Minguillón. Detecting atypical student behavior on an e-Learning system. Proceedings of the SINTICE'05: VI Congreso Nacional de Informática Educativa Simposio Nacional de Tecnologías de la Información y las Comunicaciones en la Educación, pp. 153-160, 14-16 September 2005.

F. Castro, A. Vellido, A. Nebot, J. Minguillón. Finding Relevant Features to Characterize Student Behavior on an e-Learning System. Proceedings of the FECS'05: 2005 International Conference on Frontiers in Education: Computer Science and Computer Engineering, pp. 210-216, 20-23 June 2005.

6.6 Articles derived from Objective 7: Design and development of a framework to provide real time useful knowledge to e-Learning environments and improve the e-Learning experience.

A. Nebot, F. Mugica, F. Castro. A Framework to Provide Real Time Useful Knowledge in e-Learning Environments. Proceedings of the SIMULTECH'12: 2nd International Conference on Simulation and Modeling Methodologies, Technologies and Applications, pp. 103-108, 28-31 July 2012.

F. Castro, A. Nebot, F. Mugica. A Soft Computing Decision Support Framework to Improve the e-Learning Experience. Proceedings of the MSE'08: Modeling and Simulation Education – SpringSIM'08: Spring Simulation Multiconference, pp. 781-788, 14-17 April 2008.