This is previous version (preprint) of the article published in International Journal of Engineering Education, 35(5), p. 1316-1325, 2019. https://www.ijee.ie/contents/c350519.html

Comprehensive learning system based on the analysis of data and the recommendation of activities in a distance education environment

Villegas-Ch, William[1]; Palacios-Pacheco, Xavier[2]; Buenaño-Fernández, Diego[1]; Luján-Mora, Sergio[3]
[1] *Universidad de Las Américas, william.villegas@udla.edu.ec; diego.buenano@udla.edu.ec*
[2] *Universidad Internacional del Ecuador, xpalacio@uide.edu.ec*
[3] *Universidad de Alicante, sergio.lujan@ua.es*

Abstract

Traditional teaching, based on techniques in which students develop a passive function, has proven to be an inefficient method in the engineering learning process. Universities have been forced to improve their teaching methods and have found a partial solution in open source platforms; these platforms have allowed a greater collaboration between institutions that improve the contribution of technology to education. There are cases of collaboration between universities where their sole objective is to promote student learning and the automation of educational processes. The massification of this type of technological tools allows the use of systems and platforms commonly used in the business world. This adoption of open source tools has proven to be very effective in educational environments and has offered several benefits such as the reduction of costs and the constant updating of information systems. One of the frequent cases in which there are collaborative projects based on learning is the analysis of educational data that seek to detect students' deficiencies and to take actions before they abandon their studies. In this work, we propose the design of an integral learning system in which business intelligence, expert systems, learning management systems and different learning techniques converge. This integration seeks to create a system capable of recommending different activities that focus on the needs of students.

Keywords: Open source; expert system; engineering education; project based learning; data mining

1 Introduction

A model of distance education is part of current educational trends since it has advantages over a traditional modality where the student fulfills a classroom schedule that is already defined [1]. In the face-to-face model, traditional lectures encourage students to become passive individuals, with low levels of commitment, concentration, participation and motivation towards the subject [2]. Distance education allows the educational act through different methods, techniques, strategies and means in a situation in which students and teachers are physically separated or interact occasionally [3]. These conditions, if not accompanied by a plan to follow up and provide the student with resources that meet their needs can become a problem rather than a solution. In a model of distance education, integrating concepts such as personalized education, active learning and the recommendation of activities will contribute effectively to learning. This process gives the student time flexibility that facilitates the organization of personal time respecting family life and work obligations, in addition to allowing a personal pace of study.

In distance education, the student becomes the protagonist of their learning and it is responsibility of universities to turn this experience into a personalized education. Implementing these advantages presents technical and administrative difficulties, since talking about personalized education in an environment where there is no feedback from the student affects the measurement of learning. The solution is the creation of integrated learning systems that rely on information and communication technologies (ICT), and where the follow-up is continuous based on the information that the student generates. With the knowledge extracted about the students, the system will be able to recommend activities that suit each one of their needs. To guarantee and prioritize learning, the recommendation of activities must be within an active learning model. Active learning consists of the use of a set of more effective and interesting experimental methods, whereby students assume greater responsibility for their own education [4]. Active learning has many benefits for students, including a deeper understanding of the concepts of a certain subject and promoting the student's positive attitude towards learning and, consequently, a greater motivation towards the subject.

In this sense, there have been many proposals as to how to plan educational activities, the strategies that can be applied and the models that can be adopted. However, this process continues to be a difficult path, given the number of variables to consider when appropriating some of these methodological proposals. At present we can see the contributions of the open source tools to the enrichment of educational work [5]. Most research and proposals, especially in the field of artificial intelligence (AI) and education, are usually dedicated

This is previous version (preprint) of the article published in International Journal of Engineering Education, 35(5), p. 1316-1325, 2019. https://www.ijee.ie/contents/c350519.html

to the learning process and to the academic administration related to the management of information about students [6]. The AI provides tools and techniques that allow a knowledge-based system to face problems associated with decision making. In this work, the development of an expert system (ES) is sought; the ES, after evaluating different criteria or variables, will propose applicable activities to the student in a learning situation [7]. A comprehensive system, with systems based on knowledge or ES, comes to constitute permanent support for the student. They represent a response that is oriented to the efficient decision making on the teaching models and the didactic activities to be implemented in a learning management system (LMS). This condition makes it necessary to strengthen and potentiate the autonomy of the LMS with the aim that it works in its entirety to the ES and can recommend the exact activities to groups of students already defined.

Regarding the ideas revealed, ES in the educational field can be considered to have certain advantages, particularly those created for pedagogical and instructional purposes [8]. An ES can diagnose, debug and correct the development of student learning in a particular area of knowledge. In addition, the system determines the cognitive level of the student and helps him/her improve their weaknesses to reach a higher level of learning. The work is distributed in the following sections; Section 2 contains the theoretical foundation that contributes to the design and implementation of the system. Section 3 contains the method where the whole development is explained step by step. Section 4 presents the discussion based on the results obtained.

2 Method

The process through which the integral system proposed for the recommendation of activities guarantees learning can be observed in Fig. 1. In the first phase, students are identified and classified by means of patterns that they share. The identification is carried out by a business intelligence (BI) subsystem that has the necessary capacity to perform the analysis of student data. Prior to this process, there is a whole set of tools that extract information from different data sources and turn it into knowledge. In the second phase, the resulting data are presented to the ES which, like a human expert, will identify each case and cross the information obtained from the BI with the information obtained from the interaction with the student. The analysis carried out by the ES based on its knowledge allows us to establish which activities are aligned to the needs of the students. In order to provide the system with greater possibilities and to obtain greater results with respect to learning, the activities it recommends are part of an active learning method. The ES is integrated into an additional module of the LMS, which also provides a system for continuous monitoring of learning. In the following sections, each of the stages of the integrated learning system is described in detail.

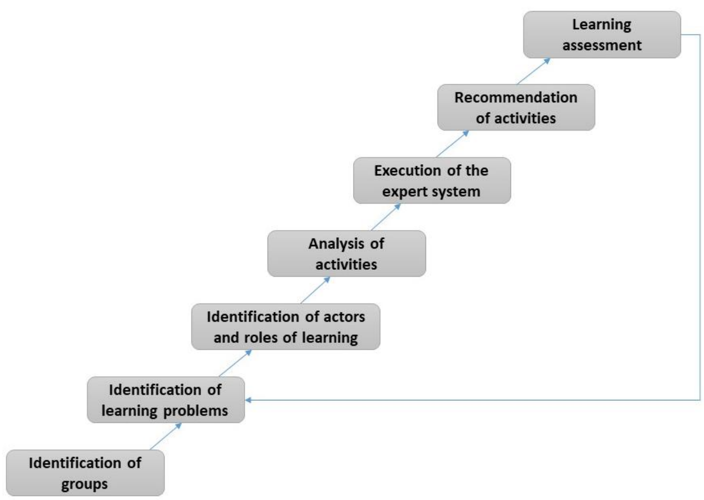

Fig. 1. Stages of the method of an activity recommendation system through an ES.

2.1 Identification of groups

In the identification of student groups, the BI module is used. This process applies data mining algorithms. The process is defined in context; however, the execution is very complex since a whole BI platform is used that considers all the stages for the extraction of knowledge of databases (KDD) [9]. The BI module keeps track of the students based on the data stored in the different repositories. The groups considered for the development of the system are the students that have a high percentage of repetition and those that have been the cause of a study of university desertion. The selection of the data is based on the study variables; these variables are chosen according to the Bean model [10]. The model identifies the academic, psychosocial and environmental variables that cause the dropout syndrome. In the academic variables, the student's performance and integration are considered; the common repositories that store this information are the systems of academic registration and academic monitoring. Psychosocial variables include objectives, alignment, peer interaction and interaction with teachers. This work, being focused on distance learning models, extracts information from the LMS; for this reason, the interaction with peers and the interactions of teachers in the LMS are modified in the Bean model. Environmental variables are parameterized according to funding, external social relations, transfer opportunities and interaction with the LMS and the use of ICT [11].

To facilitate the process of data selection, several questions are asked to help identify the fields that will be part of the variables, for example:

- Is there information available on the systems that can be used?
- Does this information help the analysis?
- Of all the available information, which one interests us?
- Are the details of all the available information interesting or just the detail of the information we need?

The data used belong to the university that participated in this work, specifically of a distance study modality. The careers included are engineering in information technology, engineering in communications networks and mechatronics engineering. The choice was mainly due to the availability of information provided by the university. Access was obtained from the data on the socioeconomic characteristics of the students that are collected in the student registration systems [12]. These data are important to carry out the academic follow-up of the students and detect their condition of defector, graduate or student active. Students who have similar learning patterns and who belong to the 2013 cohort are included in the analysis; a cohort being defined as the group of students who enroll in a university career in a given year. The advantage, unlike working with the entire population, is to identify a specific group of students who are in the same initial condition and face the same academic and social aspects. According to the data collected from the different systems, the total number of people registered in the distance education modality was 3,207 students across all the cohorts up until the year 2017. The sample considered for this study was the 2013 cohort composed of 208 students. The analysis process identified two groups of students; the sample responds to the following characteristics:

- Group 1. Students in a range of 35 - 44 years of age that have a grade lower than the average of the course considered in 6/10.
- Group 2. The low percentage of hours of significant interactivity.

2.2 Identification of learning problems

Table 1 identifies eight common problems within a distance education model that each group of students faces [13]. The data were obtained through the use of surveys carried out with the population included in the research.

Table 1. Identification of relevant problems in a distance education modality

Problems identified in a distance education modality	**Group 1. Age between 35-44 years**	**Group 2. Low percentage of hours of interactivity**
Absence of significant interactivity	X	XXX
Use of materials not designed for online education	XX	X
The lack of evaluations and exercises	X	XXX
The use of irrelevant content	X	XXX
Complex navigation	XXX	X
Instructions for insufficient or unclear use	XXX	X

Criterion of maximum impact XXX

This is previous version (preprint) of the article published in International Journal of Engineering Education, 35(5), p. 1316-1325, 2019. https://www.ijee.ie/contents/c350519.html

Criterion of medium impact XX
Criterion of minimum impact X

Overcoming the problems found in each of the groups is a topic of analysis on which several areas of the university have worked. When activities are suggested, these are adjusted to the needs of each student, considering the time they spend on academic activities, what they want and should learn, and the process of evaluating learning. The evaluation of the learning process is given by the results, processes and conditions that can be represented by questions about any learning situation [14]. For example, what do we want to learn? How do we learn that? What conditions favor this learning?

2.3 Identification of the actors and their roles in learning

It is important to clearly establish the actors and their roles in education: the student, who is the owner of their learning, and the teacher, who guides the student in how to get to the learning. Considering that the student owns their learning, they fall back on what they want; for example, there are students who simply set their goal to pass a certain course and others have a real desire to generate knowledge [15]. Here the teacher enters into action because the task falls on them to propose strategies that help to define what they want the student to learn. Once have defined what to learn, they have to look for techniques that enable doing it and that are within the learning processes. Finally, we must meet the conditions for the student to learn. In any learning scenario, teachers can vary conditions to meet the learning objective. Modifying the conditions is to modify the tactics and the scenarios of learning and teaching. Next, the roles in learning are presented according to their participation guidelines.

Participation of students:

- They move from a passive listening role to active involvement in learning (readings, discussions, reflections, etc.).
- They are involved in higher-order thought processes such as analysis, synthesis and evaluation.
- They learn through dialogue and through interaction with the content and development of competencies.
- Students receive immediate feedback from the teacher and their classmates.

Participation of teachers:

- They design activities according to their discipline and at the current curricular moment needs of their students.
- They adapt the learning activity to the possibilities and needs of the group.
- They facilitate the process of the activity taking care of the extension and the depth of knowledge that is approached.
- Feedback in a timely manner on the performance of the group and individual students.

2.4 Analysis of activities

Established the groups and the participation of each actor in the learning is analyzed the activities that contribute to the learning considering that they must be designed to be presented in an LMS. Under these considerations, the following activities are proposed that can be classified as projects based on learning [16]:

- The “One Minute Paper”: This is a highly effective technique to monitor the progress of the student, both in their understanding of the subject and in reacting to the course material. It can be generated by reading a resource within the Moodle activities.
- Reading Contests are a way to motivate students to read the assigned material. Active learning depends on students reviewing the resources available in each module of the LMS. A reading test can also be used as an effective measure of student comprehension of the reading material so that their level of reading sophistication can be measured.
- Puzzles/Paradoxes are one of the most useful means of eliciting insights from students about a topic. It presents a paradox or a puzzle related to the concept in question and guides them towards a solution. By forcing students to work without authority, it increases the likelihood that they will be able to critically evaluate the theories that are presented.

- Discussion students are asked to pair up and answer a question. This can be easily combined with other techniques such as questions and answers.
- Conceptual maps are a way to illustrate the connections that exist between the terms or concepts covered in the course material. Students construct conceptual maps by connecting individual terms by lines that indicate the relationship between each set of terms connected.

When moving from a normal learning scenario to active learning, one must consider that this focuses on the student by promoting their participation and continuous reflection through activities that promote dialogue, collaboration, development and knowledge construction, as well as skills and attitudes. The activities must be motivating and challenging, and capable of promoting active adaptation to solving problems.

2.5 Execution of the expert system

The ES is in charge of interacting directly with the student; the communication protocol is through questions that will allow solving problems much faster than a human expert. The ES is based on a deterministic methodology that aligns with the exposed needs as well as having a large number of success stories [17]. For the development of ES, the stages proposed by Weiss and Kulikowski [18] are considered; each stage must be properly analyzed and developed since this directly influences the quality resulting from the ES.

a) System tools

For the design of the expert system and the necessary components for the recommendation of learning activities, open source tools are used. These tools are used based on costs, availability of information and the learning curve that is very low. For the design of the expert system, the Scrum methodology is used, which allows continuous adaptation to the different circumstances in the development of the system. The development language used is prolog that allows a logical programming and becomes the processing core. Prolog is based on two elements that must necessarily be identified to meet the expectations of the research. These elements are the atoms that define in a generic way an object of the environment that we want to represent and the predicates that are responsible for specifying the characteristics of the objects in the environment and the relationship between them [19]. For example, the logical representation for the case of students with learning problems. (Student_learning_problems (X) has (X, bad grades): Student_learning_problems (X)). In this rule generalizes the fact that any object that is a student with learning problems, will have bad grades. In development, it must be borne in mind that the fact that an object has poor grades is not a sufficient condition for it to be a student with learning problems.

The development of the interface that interacts with the user is developed in PHP, which is an open source and is widely used in web development environments. It is important to consider that the whole system is in a testing process for which it is accessed via intranet by a sample of the student population [20]. PHP provides the best solution for the environment mentioned by its ease of use, as well as its orientation to the development of dynamic web applications. These tools are easily integrated into the LMS and the BI platform, since the LMS is Moodle with a database in MySQL and for its part the BI platform is developed in Pentaho in its community version [21]. These characteristics make the whole system is based on open source tools acquiring the advantages already mentioned.

b) Problem statement

The objective of identifying the problem for the execution of the ES is to extract general knowledge about the university, its components, characteristics, structures and processes. These parameters contribute to the design of the ES, as well as a global vision of the current state of the university with respect to its environment, its shortcomings, its resources and relationships. Within this context, a university has education as its main activity; this concept integrates a number of variables to consider, as you cannot establish a general method and expect similar results in all students. The variation of learning in students depends on economic as well as psychosocial factors and determining a strategic plan that allows for taking action measures requires much technical and administrative effort. The creation of an ES that is in charge of the analysis of these factors and recommends activities that adjust to each student proposes to solve this problem.

Specifically, the problems that are intensive in knowledge are characterized because they are based on facts and rules of a domain, as well as on the experience of the people and organizations that make it up. The sets of problems that accompany universities according to knowledge engineering are:

This is previous version (preprint) of the article published in International Journal of Engineering Education, 35(5), p. 1316-1325, 2019. https://www.ijee.ie/contents/c350519.html

- Learning problems (Very High).
- Lack of feedback from the student (High).
- Acceptance of the educational model is low (Medium).
- There is no desire to learn in the student (Low), etc.

c) Find human experts

In this phase, several expert teachers work in different areas of engineering. The objective is to establish the questions that the system must ask the students and that they serve as a guide to obtain an accurate conclusion. The design of activities is another task that this group of teachers must perform focusing on active learning. Another working group is made up of experts in educational psychology who seek methods through which the student is interested in generating knowledge and not in obtaining a grade. A third group is made up of experts in the use of open source LMS, specifically Moodle. The objective of this group of experts is to create the activities proposed by the experts in the subject areas and translate them into the platform considering the methodologies proposed by the second group of experts.

d) Expert system design

The design must be flexible and gather characteristics similar to human behavior. This includes the ability to acquire knowledge, the reliability that allows trusting the results, the knowledge domain that provides solidity in a process and the ability to solve problems. The design is done based on a Bayesian theorem, which is an algorithm that is based on probabilistic theory and combines the Bayes theorem with the expressivity of the directed graphs [22]. The directed graphs allow the representation of a causal model by means of the graphic representation of the dependencies between the random variables and the causal influences [23]. The design is constituted by the Bayes theorem (1), whereby given two variables "x" and "y", such that P (x)> 0 for all x and P (y)> 0 for all Y is fulfilled:

Equation 1. Bayes theorem

$$P(x|y) = \frac{P(x)\ \cdot P(y|x)}{\sum_{x^1} P(x^1) \cdot P(y|x^1)}$$

For example, several tests similar to the following process have been carried out for development. For a course with 50 students (23, 18, 9), divided into groups defined by their performance (high, medium, low), it is suggested that they perform some action to improve their learning.

- 30% Games, 40% Quick Tests, 25% Forums and 5% Puzzle.
- 25% Games, 35% Quick Tests, 30% Forums and 10% Puzzle.
- 20% Games, 50% Quick Tests, 10% Forums and 20% Puzzle.

The environment analyzed is as follows: a) is the recommended activity the third?; b) if we find out what the forums are, what would it be?

Solution:

a) A priori solution: P (x) = 9/50 = 0.18 = 18%
b) To answer this request, we describe the data in Table 2 where x = activities | y = groups of students.

Table 2. Identification of activities and groups of students through Bayes

P(x|y)	**x1**	**x2**	**x3**
Yc	0.30	0.25	0.20
Yi	0.40	0.35	0.50
Yh	0.25	0.30	0.10
Ye	0.05	0.10	0.20

In the equation (2) the data of the table is processed and the percentage with which the results are processed in the ES is obtained.

Equation 2. Identification of activities and groups through Bayes theorem

This is previous version (preprint) of the article published in International Journal of Engineering Education, 35(5), p. 1316-1325, 2019. https://www.ijee.ie/contents/c350519.html

$$P * (x^3) = P(x^3|y^e) = \frac{P(x^3) \cdot P(y^e|x^3)}{\sum_x P(x) \cdot P(y^e|x)} = \frac{0.18*0.20}{0.46*0.05+0.36*0.10+0.18*0.20} = 0.379 = 37.9\%$$

The design of the ES, which is based on Bayesian networks, constructs a decision tree that allows combining the judgment of the experts with the available data and making the inference between any set of variables [24]. The architecture of an ES is organized according to three main elements, which are the basis of knowledge, the inference engine and the basis of facts; however, to ensure the dialogue between the machine and the human, additional components are required [25]. The components that act additionally are the human component and the knowledge acquisition systems, etc.

e) Acquisition of knowledge

In this phase, the basic knowledge in the subject is provided and the knowledge engineers transfer this knowledge to a language that the ES understands. This stage requires great dedication and maximum effort due to the different levels of knowledge of those involved and the different experiences they have [26]. For the development of the ES, the technical knowledge of experts in the area of computer tools, networks, computer audit and robotics, among others, is used. These areas of expertise are considered when contributing to engineering careers related to ICT. The experts in teaching crystallize the activities given by the experts with a psychological tinge that contributes to the active learning of the students. Once the experts are selected, and they agree to give their knowledge, they start playing the role of the "Knowledge Engineer". As such, they are in charge of extracting the knowledge from the expert and giving it an appropriate representation, either in the form of rules or another type of representation, thus forming the knowledge base of the expert system.

f) Knowledge base

The knowledge base is the phase where specialists are responsible for providing knowledge engineers with an orderly and structured base and a set of well-defined and explained relationships. This structured way of thinking requires that human experts rethink, reorganize and restructure the knowledge base and, as a result, the specialist becomes a better connoisseur of their own field of expertise [27]. However, we must differentiate between data and knowledge; knowledge refers to statements of general validity such as rules, probability distributions, etc. This is stored in the knowledge base and the data is stored in the working memory. There are three ways of representing knowledge: production rules, semantic rules and frames.

The most commonly used form of representation is production rules, also called rules of inferences [28]. Most ESs are based on this type of representation. For this reason, the design for this work considers this option for success cases in other areas, as well as the amount of existing information that contributes to the development.

The rules of production are of the type:

- IF Premise THEN Conclusion (YES A THEN B)

Where both the premise and the conclusion are no more than a chain of events connected by "Y" or "O", in general it would be:

- IF Done1 AND/OR Done2 AND/OR ... DoneN THEN Done1 AND/OR ... DoneN

The facts are statements that serve to represent concepts, data, objects, etc. The set of facts that describe the problem is the basis of facts.

g) The inference engine

The inference engine is the supervisor; it is responsible for extracting the conclusions based on the symbolic data by applying the rules that govern the system in which it works; thus, a change in the rules will result in different conclusions, which is why it uses data that are facts or evidence, and knowledge that is the set of rules stored in the knowledge base to obtain new conclusions or facts [29].

The function of the inference engine is to execute actions to solve the problem from an initial set of facts and, eventually, through an interaction with the user. The execution can lead to the deduction of new facts. This process is done through rules that model the general knowledge of the domain and constitute the knowledge base, long-term memory and implications. For example, if x is 45 years old and studies, "then" x is a university

This is previous version (preprint) of the article published in International Journal of Engineering Education, 35(5), p. 1316-1325, 2019. https://www.ijee.ie/contents/c350519.html

student. The ES will process all the information through the aforementioned stages and reach a conclusion. The percentage of the uncertainty with which it reaches the conclusion will depend on the allocation of activities. Table 3 details the percentages of uncertainties with which the system will work.

Table 3. Percentages of ES uncertainty

% Acceptance	
%	Conclusion
0 – 50	Not feasible
50 – 70	Feasible with uncertainty
70 – 90	Feasible
90 – 100	Sure

The questions that the ES gives to the students are designed in natural language and look for the interaction in a simple and direct way. Below are several questions used to interact with the student; these seek to know if the student is giving a true answer [30]:

- Hello “X”, how old are you?
- Is the time you dedicate to the course sufficient?
- Are the contents of the course easy to understand?
- Do practical activities strengthen your learning?
- Do you want to develop a case applied to business reality?

For example, asking if the student is on a computer tools course allows the ES to validate the information it receives from the student. The validation of information allows it to reach conclusions more accurately. This process is possible by integrating the ES with systems that manage student information, thus allowing advance knowledge of several student data. The ES takes this information and contrasts it with the information received in order to verify that it is the person who claims to be.

The answers in most cases are given by the following four options: Yes; It seems to me that yes; I don't know; I don’t think so; No

2.6 Recommendation of activities

The process of recommending activities is based on the three following stages: a) the profile of the user is loaded directly from the data that the ES finds in the LMS, however, the student's name and age are asked as a question of courtesy; b) it extracts information from the available activities and adjusts this to the needs of each student; c) it initiates the recommender with the ES to compare the user's information with the information stored in the knowledge base and recommends an activity. Fig. 2 presents the entire process of the recommender system [31].

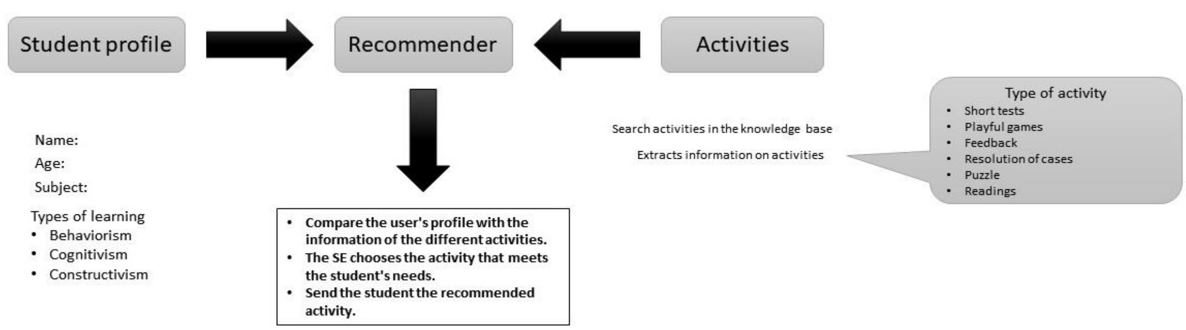

Fig. 2. Process of the recommender system.

For a better understanding of the process carried out by the ES to recommend one or more activities, Table 5 is presented. Here, the weights are established through which the system can perform the calculations to reach a certain conclusion. This process is a basic example of the internal functioning of the ES, whereby

This is previous version (preprint) of the article published in International Journal of Engineering Education, 35(5), p. 1316-1325, 2019. https://www.ijee.ie/contents/c350519.html

five unique values of probability are handled; these are: 1 = yes, 0.75 = I think so, 0.50 = I do not know, 0.25 = I do not think so, and 0 = no.

If universities want to identify recommended student tasks, but they have little available time to develop activities and for students to learn to solve problems of a specific topic, the system eliminates activities that do not meet these criteria and recommends activities with greater weight. For students who meet the indicated conditions after carrying out the process and calculation, the system will recommend play activities. With a lower weight, the system recommends roundtable activities.

Table 5. Activity recommendation process

Has	Reading control	Rapid test	Playful games	Development of cases	Puzzle	Discussion through forums	Conceptual maps	Roundtables
Little time available to develop activities	0	1	0.75	0	0	0.5	0	0.75
Capture the idea with ease	1	0	1	0.75	0.5	1	1	0
Generates knowledge when solving problems	0	0	0.75	1	1	0	0.75	0.25
Discussion on specific topics	0	0	0	0.5	0	1	0.75	1
Constant review	1	0	0.75	0	0	0.25	0	0
I understand what I do	0	0.25	0.75	1	0.75	0	1	0
I master what I teach to another	0	0	0.25	0.75	0	1	0.5	1
I understand what I talk about with another	0	0	0	1	0	1	0.75	1
Recognizes different appearances and circumstances	0.75	0.25	0.75	1	1	0	0.75	0.5

The functional model is more sophisticated as it includes rules that help to deduce answers from the previous ones, allowing solutions to be found more quickly and efficiently. The system asks very generic questions that allow a significant reduction in the recommendation of activities.

3 Results

Through the use of the BI platform, a group of 10 students of the engineering degree in information technologies has been selected. The filters of group is the low percentage of hours of interactivity in the LMS platform and the low qualifications of the partial I. To this group the ES recommended the activities of case and puzzle developments with the purpose that they generate knowledge by solving problems. The activities have been planned within the LMS with the objective that encourage the use the platform and that the interaction time is effective generating the desired knowledge in the student. Table 6 presents the interaction data of the students and the grades obtained. Each partial consists of eight weeks therefore they contain information on the performance of students before applying the recommendation system. This information is compared to the

This is previous version (preprint) of the article published in International Journal of Engineering Education, 35(5), p. 1316-1325, 2019. https://www.ijee.ie/contents/c350519.html

second eight weeks that are part of the same academic period. The second part contains the academic data of the system with the functioning of the recommendation system.

Table 6. Partial I without ES of activities vs Partial II with ES of activities

	Partial I without ES of activities		Partial II with ES of activities	
Student	**Rating / 10**	**Interaction time hrs/month**	**Rating / 10**	**Interaction time hrs/month**
ID2431	3.4	2.36	7.6	5.01
ID5246	1.1	0.35	4.3	4.37
ID4876	2.7	4.21	5	6.35
ID7548	3.2	1.52	4.0	2
ID6584	2.9	2.65	7.1	6.35
ID8954	3.3	3.12	6.5	7.32
ID5842	2.2	0.59	8.6	12.25
ID6547	3.1	3.24	1	0.25
ID9854	4	2.54	6.1	4.54
ID8774	5	1.51	2	2.35

The BI platform identified students with low interaction with the platform and low scores in the first part. The activities developed in this time are shared by the teacher without there being any difference between students. With this knowledge, it is contrasted with the data of the second part where, the recommender system has followed the students assigning tasks that motivate the use of the platform. These activities are, the resolution of practical cases on the corresponding topics in each week, as well as the resolution of puzzles on these subjects. The differences are visible since 80% of the sample has improved the times of interaction with the platform and the ratings. The data presented is general, since the system is barely functioning and it is expected that in the period 2020-20 the first cut will be made, which will allow a deeper analysis regarding the efficiency of the activities in each case.

4 Discussion

The problem that is presented to the designers of educational environments is the number of variables that interact with each other and which must be taken into account for an effective design that includes an integral learning system. The strategies of instruction embedded in a highly effective design consist of the adequate interaction between the elements that intervene in the instruction scenario in order to optimal achievement of educational objectives. The proposed system assists designers and tutors in identifying the best activities that are determined by the scenario and defined by the characteristics of the learner and the learning context. The ES to identify the best instructional strategies receives the following scenario data as input: the characteristics of the domain of knowledge to be transmitted, the profile of the student who will act as a user, a description of the technological environment that will act as a facilitator of learning and the description of skills that the individual is expected to achieve as a result of the instruction. Once all these data are known, the ES identifies the pattern or combination of patterns that best suits the situation presented by the user, thus recommending the activities to be used by the designer.

5 Conclusions

The integration of an ES with an LMS that interacts directly with the student guarantees the constant accompaniment of the expert with the student during the learning process. The result of this integration brings more benefits and not only to the student but also to the teacher. With this application, the teacher better uses his time devoted to the construction of knowledge through the design of new activities. For the teacher, the ES becomes a teaching assistant that facilitates learning and assesses the learning of each activity proposed by the teacher. This contribution optimizes the time of dedication of the teacher in the design of interactive resources that accompanied by the recommendation of activities by the ES greatly improves the learning of a student. An important point that allows measuring the effectiveness of the ES are the percentages of desertion that in the short time of execution of the system has been able to reduce.

Technically the solution must go through a time of maturity during which the results in the students must be analyzed. However, the process so far is satisfactory considering that its design is scalable and that its construction is based on the use of an open source language. These features will allow its evolution at a low cost, and it can also include the community that uses this type of language for its continuous improvement.

The next step to follow in this work is to increase the number of variables that the ES manages to improve their learning about how students learn in the proposed ecosystem. The obligation of those in charge of administering the ES is to evaluate the effectiveness of the proposed activities by means of data analysis systems. The validity of the results obtained highlights the role of experts who contributed with their knowledge to deduce the relevant variables of each subject and its importance in the learning. The result of the use of the ES allows recommending a variety of activities that contribute to learning using a multi-criteria concept.

References

[1] R. M. Bernard, P. C. Abrami, Y. Lou, E. Borokhovski, A. Wade, L. Wozney, et al. How does distance education compare with classroom instruction? A meta-analysis of the empirical literature. *Review of educational research*, **74(**3), pp. 379-439, 2004.

[2] J. H. Mcmillan, and S. Schumacher. Research in Education: Evidence-Based Inquiry, MyEducationLab Series, *Pearson*, 7, pp. 528, 2010.

[3] M. G. Moore and G. Kearsley. *Distance education: A systems view of online learning*, Cengage Learning, pp. 45-293, 2011.

[4] M. Simonson, S. Smaldino and S. M. Zvacek. *Teaching and learning at a distance: Foundations of distance education.* IAP, pp. 162-323, 2014.

[5] D. E. Drummond. Open sourcing education for Data Engineering and Data Science, *in Education Conference (FIE)*, pp. 1-1, 2016.

[6] B. Settles. Active learning. Synthesis Lectures on Artificial Intelligence and Machine Learning, **6**(1), pp. 1-114, 2012.

[7] R. O. Tandogan, and A. Orhan. The Effects of Problem-Based Active Learning in Science Education on Students' Academic Achievement, Attitude and Concept Learning, *Online Submission, Eurasia Journal of Mathematics, Science & Technology Education*, **3**(1), pp. 71-81, 2007.

[8] R. J. Spiro, B. C. Bruce and W. F. Brewer. (eds), *Theoretical Issues in Reading Comprehension*, vol. 11, 1ra edn, London: Routledge, 2017.

[9] W. Villegas-Ch, S. Luján-Mora and D. Buenaño-Fernández, Data mining toolkit for extraction of knowledge from LMS. In *Proceedings of the international Conference on Education Technology and Computers*. ACM, pp. 31-35, 2017.

[10] J. P. Bean. The application of a model of turnover in work organizations to the student attrition process. *The review of higher education*, **6**(2), pp. 129-148, 1983.

[11] S. M. Weiss, and C. A. Kulikowski. *Computer systems that learn: classification and prediction methods from statistics, neural nets, machine learning, and expert systems.* Morgan Kaufmann, 1991.

[12] W. Villegas-Ch and S. Luján.Mora. Analysis of data mining techniques applied to LMS for personalized education. In *World Engineering Education Conference, IEEE*, pp. 85-89, 2017.

[13] J. R. Savery. Overview of problem-based learning: Definitions and distinctions, *Essential readings in problem-based learning: Exploring and extending the legacy of Howard S. Barrows*, **9**, pp. 5-15, 2015.

[14] R. E. Mayer, *Learning strategies: An overview.* In Learning and study strategies, pp. 11-22, 1988 Academic Press.

[15] E. Wenger. *Artificial intelligence and tutoring systems: computational and cognitive approaches to the communication of knowledge*, Morgan Kaufmann, 2014.

[16] J. L. Jensen, T. A. Kummer, and P. D.D. M Godoy. Improvements from a flipped classroom may simply be the fruits of active learning. *CBE-Life Sciences Education*, **14**(1), pp. ar5, 2015.

[17] D. J. Power, R Sharda, and F. Burstein. *Decision support systems.* John Wiley & Sons, Ltd, 2015.

[18] D. Heckerman and B. N. Nathwani. Toward Normative Expert Systems Part II, *Methods of Information in medicine*, **31**, 2016.

[19] C. Dunchev, F. Guidi, C. C. Sacerdoti, E. Tassi, *ELPI: Fast, Embeddable, λProlog Interpreter*, in M. Davis, A. Fehnker, A. McIver, A. Voronkov (eds), *Logic for Programming, Artificial Intelligence, and Reasoning*, vol 9450, Springer, Berlin, Heidelberg, p. 460-468, 2015.

[20] B. Wang, P. S. L. Liu and X. F. Wang, Research on the Application of Flipped Classroom Model Based on MOOC in the Course PHP Dynamic Website Development. In *International Symposium on Educational Technology*, pp. 200-203, 2018.

[21] V. Vargas, A. Syed, A. Mohammad and M. Halgamuge, Pentaho and Jaspersoft: a comparative study

of business intelligence open source tools processing big data to evaluate performances. In *International Journal of Advanced Computer Science and Applications*, **10**(14569), pp. 1-10, 2016

[22] G. F. Cooper and E. Herskovits. A Bayesian method for the induction of probabilistic networks from data. *Machine learning*, **9**(4), pp. 309-347, 1992.

[23] D. E. Heckerman, E. J. Horvitz, and B. N. Nathwani. Toward normative expert systems part i. *Methods of information in medicine*, **31**, 2016.

[24] A. Okutan and O. T. Yıldız. Software defect prediction using Bayesian networks. *Empirical Software Engineering*, **19**(1), pp. 154-181, 2014.

[25] M. Monsion, B. Bergeon, A. Khaddad and M. Bansard. An expert system for industrial process identification, In *Artificial Intelligence in Real-Time Control*, pp. 85-89, 1988.

[26] J. H. Boose. A knowledge acquisition program for expert systems based on personal construct psychology, *International Journal of Man-Machine Studies*, **23**(5), pp. 495-525, 1985.

[27] J. F. Sowa. (ed). Principles of semantic networks: Explorations in the representation of knowledge. *Morgan Kaufmann*, 2014.

[28] H. Liu, A. Gegov and M. Cocea. Rule-based systems: a granular computing perspective, *Granular Computing*, **1**(4), pp. 259-274, 2016.

[29] J. Pearl. Probabilistic reasoning in intelligent systems: networks of plausible inference, *Elsevier*, 2014.

[30] U. Lindqvist and P. A. Porras. Detecting computer and network misuse through the production-based expert system toolset (P-BEST), In *Security and Privacy, Proceedings of the Symposium on. IEEE*, pp. 146-161, 1999.

[31] G. Adomavicius and A. Tuzhilin. Context-aware recommender systems. In *Recommender systems handbook. Springer, MA*, pp. 191-226, 2015.

Biographies of the authors:

William Villegas-Ch is professor in the area of Electronics and Information Networks of the University of the Americas. He obtained his masters in communications networks and is an engineer in systems with mention in robotics in artificial intelligence. His main research topics include web applications, accessibility and web usability, data mining, e-learning. He has taught courses in “Ofimatics”, “Operating Systems”, “Network Security”, “Network Quality Service” and “E-bussines” at the International University of Ecuador and at the Universidad de Las Américas.

Xavier Ivan Palacios Pacheco is Engineer in Computer Systems and Computer Science, Master in Connectivity and Telecommunications Networks, both degrees obtained at the National Polytechnic School. He has been teaching since 2000 at the International University of Ecuador, in face-to-face and distance learning, in Commercial Engineering, Computer Engineering and Multimedia, Engineering in Risk Management and Emergencies. He has held the position of Director of the Systems Department of the International University of Ecuador since 2000. He is currently developing research on learning analytics and data mining in the academic field.

Diego Buenaño-Fernández is an Assistant Professor of Computer Systems and Computer Science at the University of the Americas (Quito, Ecuador). Engineer in computer systems and computer science (1999) by the National Polytechnic School (Quito, Ecuador). Coordinator of Quality Assurance of the Faculty of Engineering and Agricultural Sciences of the Universidad de Las Américas. PhD student in the PhD program in computer science at the University of Alicante (Spain). His research line is related to Data Mining in educational environments and Learning Analytics. At the level of undergraduate professor of the subjects of “Operating Systems” and “Electronic Business” and at the graduate level, professor of “Workshops titling”.

Sergio Luján-Mora received the Ph.D. degree in Computer Engineering from the Department of Software and Computing Systems at the University of Alicante, Alicante, Spain, in 2005. He received the Computer Science and Engineering degree at the University of Alicante in 1998. He is currently a senior lecturer of the Department of Software and Computing Systems at the University of Alicante. His main research interests include web applications, web development, and web accessibility and usability. In recent years, he has focused on e-learning, Massive Open Online Courses (MOOCs), Open Educational Resources (OERs), and the accessibility of video games. He is the author of several books and many published artEicles in various conferences (ER, UML, DOLAP) and high-impact journals (DKE, JCIS, JDBM, JECR, JIS, JWE, IJEE, UAIS).