e-Learning Proposal Supported by Reasoning based on Instances of Learning Objects

Benjamin Maraza-Quispe[1], Olga Melina Alejandro-Oviedo[2], Walter Choquehuanca-Quispe[3]
Alejandra Hurtado-Mazeyra[4], Walter Fernandez-Gambarini[5]

Facultad de Ciencias de la Educación, Universidad Nacional de San Agustín, Arequipa-Perú

Abstract—In recent years, new research has appeared in the area of education, which has focused on the use of information technology and the Internet to promote online learning, breaking many barriers of traditional education such as space, time, quantity and coverage. However, we have found that these new proposals present problems such as linear access to content, patronized teaching structures, and non-flexible methods in the style of user learning. Therefore, we have proposed the use of an intelligent model of personalized learning management in a virtual simulation environment based on instances of learning objects, using a similarity function through the weighted multidimensional Euclidean distance. The results obtained by the proposed model show an efficiency of 99.5%; which is superior to other models such as Simple Logistic with 98.99% efficiency, Naive Bayes with 97.98% efficiency, Tree J48 with 96.98% efficiency, and Neural Networks with 94.97% efficiency. For this, we have designed and implemented the experimental platform MIGAP (Intelligent Model of Personalized Learning Management), which focuses on the assembly of mastery courses in Newtonian Mechanics. Additionally, the application of this model in other areas of knowledge will allow better identification of the best learning style of each student; with the objective of providing resources, activities and educational services that are flexible to the learning style of each student, improving the quality of current educational services.

Keywords—*Learning; management; intelligence; styles; instances; objects; reasoning; model; personalized*

I. INTRODUCTION

Artificial Intelligence (AI) in education is a highly researched field [1], which focuses primarily on the formulation and application of techniques for the development of systems that improve the teaching process through computer-assisted learning [2], with the goal of building more intelligent systems [3].

The term "intelligent" used in these systems is fundamentally determined by its capacity for continuous adaptation to the characteristics of learning and knowledge of different users [4].

For example, an article presents an approach that recognizes the relevant elements of the student profile seeking to meet their personal and academic needs by recovering reusable knowledge units with fine granularity stored in repositories of learning objects. The proposal is based on the technique of artificial intelligence, known as Case-Based Reasoning (CBR) [3].

For these reasons, we have decided to focus our research on developing dynamic methods for the search and identification of a student's best learning style. These methods are applied in real time, using Case-Based Reasoning CBR, through the similarity function, using the weighted multidimensional Euclidean distance.

Finally, the CBR will provide a method for personalizing learning according to the best learning strategy of each student. On the other hand, we have calculated the efficiency of our algorithm for the selection of learning styles via CBR, which has been compared with the results obtained by other learning styles selection algorithms such as Neural Networks, Naive Bayes, Tree J48 and Simple Logistic [5].

In this context, we have designed and implemented the MIGAP platform (Intelligent Model of Personalized Learning Management), with the aim of presenting learning contents, which can be adapted to the best learning style of each student according to the Honey-Alonso model.

The importance of our proposal is to adapt teaching to the specific needs of the student, giving flexibility and autonomy to the learning environment. For this we use artificial intelligence techniques such as Case Based Reasoning, whose efficiency is compared to other techniques or algorithms with RN, Naive Bayes, Tree J48 and Simple Logistic.

II. THEORETICAL FRAMEWORK

A. Artificial Intelligence in Education

In the field of Artificial Intelligence applied to education, we have found research focused on the development of systems for education, based on aspects of knowledge [6]. Figure 1 shows the main AI techniques applied to education.

B. Learning Styles

1) Definition: Learning styles can be defined as cognitive, affective and physiological traits, which serve as indicators of how students perceive interactions and respond to their learning environments. [7].

Therefore, it can be determined that each person has a "fingerprint" of their own learning. Which each person develops and powers following different types of educational strategies (some learn through reading, others through practices, others through group work, others through individual work), however, we all have in different percentages some feature of different learning styles. [8].
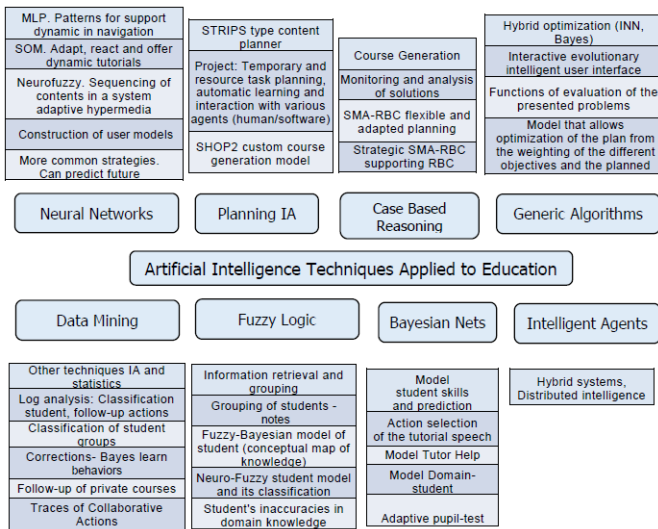
Fig. 1. Main AI Techniques Applied in Education.

2) Selected learning styles: To select a learning style show in Table I, it is necessary to first investigate what types of learning styles exist; we have found a model in which they use four styles of learning which are: [5].

- Active
- Theoretical
- Pragmatic
- Reflective

3) Artificial intelligence technique applied in the proposal: The technique of artificial intelligence that we have applied to the proposal is the Case-Based Reasoning, which first detects the student's learning style to determine the best learning strategy that best fits his learning style. CBR is the process of solving new problems based on solutions to previous problems. For example, an auto mechanic who repairs an engine because he remembered that another car had the same symptoms is using case-based reasoning [9]. An attorney who appeals to legal precedents to defend a case is using case-based reasoning.

Case-Based Reasoning: Case-Based Reasoning (CBR) is a set of concepts and techniques that address issues related to the representation of knowledge, reasoning and learning from experience [4]. The similarity is a concept that plays a fundamental role in CBR.

Case Definition: The case definition is also known as an instance, object or example. It can be defined as a piece of contextualized knowledge that represents a significant experience [10].

Stages of CBR: The main stages of CBR are four: Recovery, Reuse, Revision, and Retention. These four stages involve basic tasks such as case grouping and classification, case selection and generation, case learning and indexing, case similarity measurement, case recovery and inference, reasoning, adaptation rules, and data mining.

TABLE. I. CHARACTERISTICS OF EACH LEARNING STYLE [9]

Learning Styles	Main features
Active	Animator, Improviser, Discoverer, Risky.
Reflective	Weighted, Conscientious, Receptive, and Analytical.
Theoretical	Methodical, Logical, Objective, Critical, Structured
Pragmatic	Experiments, Practical, Direct, Efficacious, Realist

Life cycle of CBR: The life cycle of an RBC system for problem-solving consists of four states:

- Retrieve of similar cases from an experience base.
- Reuse of cases by copying or integrating solutions from recovered cases.
- Revise or adaptation of the recovered solutions to resolve the new problem.
- Retention of a new solution, once it has been confirmed or validated.

a) Functioning: Each case has predefined attributes. These are defined by a Name and a Datatype that can be String, MultiString, Float, Int, and Bool. The best case is found by the Euclidean distance, a technique that is similar to the Pythagorean Theorem in n dimensions. To find the best case, a similarity indicator (percentage) is used: [18].

$$M = \frac{100 \times (1 - D)}{\sqrt{\sum_{1}^{n} w}}$$

Where:

M = Best case percentage (between 0 and 100)

D = Distance

W = Weight

In addition, D is the distance between cases. It is a floating number between zero and one, which is calculated like this:

$$D = \sqrt{\sum_{1}^{n} Wn \times d(c - Cn)}$$

D = distance

W = Weight

d = Distance between the case "c" sought and the case Cn of the case base.

This floating value *D* means:

- 0 means exact accuracy (identical case).
- 1 means maximum distance (Not a similar case at all).

The weight W (weight n) is the weight for the attribute "n". It is an Integer $> = 0$, default = 5.

In addition, if an attribute for a case is not entered; the case is disqualified and does not enter the search.

b) Description of the proposed solution: A case-based system has three main components: a user interface, an

inference engine, and a case database. The case database contains descriptions of previously solved problems in the form of features (predictors and targets). Each case can describe a particular episode or a generalization of a set of related episodes.

The inference engine is the reasoning machine of the system, which compares the problem inserted with those stored in the case database and as a result infers a response, which has the greatest degree of similarity to the desired response. The user interface allows communication between the system and the user, giving the possibility of interacting with the case database, being able to raise new problems and consult the inferred results [11].

To solve the problems with the use of this technique, a similar case to the new one is recovered and the solution of the recovered problem is proposed as a potential solution for the new problem. This is derived from an adaptation process in which the old solution is adapted to the new situation. In this system, we have defined a series of steps and components that interact in a cycle of reasoning. Then from a new problem, the system recovers cases similar to the one entered, which go through an adaptation process with the aim of obtaining a response according to the proposed situation. Finally, if necessary and after its revision, the system decides whether or not to learn the given solution. All that we have mentioned is considered the cycle of reasoning based on cases as shown in figure 2.
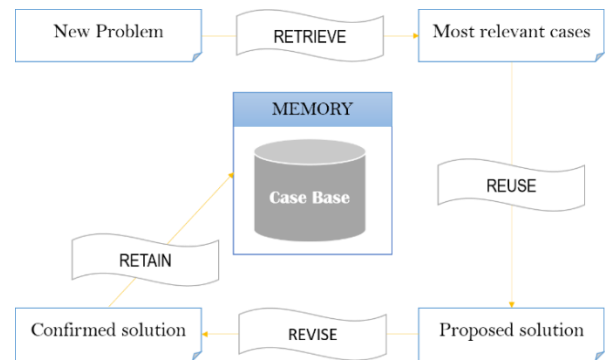

Fig. 2. Life Cycle of CBR [18].

III. CASE STUDY

A. Architecture of the Proposed Model

Fig. 3 describes the proposed architecture, which is based on the general structure of an intelligent learning management system which considers student learning styles, integrating Case-Based Reasoning for the selection of teaching-learning strategies and Neural Networks for the identification of learning styles [12]. The architecture proposes innovations in the representation of the tutor module and the knowledge module. In particular, the tutor module incorporates the CBR technique, which will choose the contents considering the teaching strategies that support the learning styles of the student.
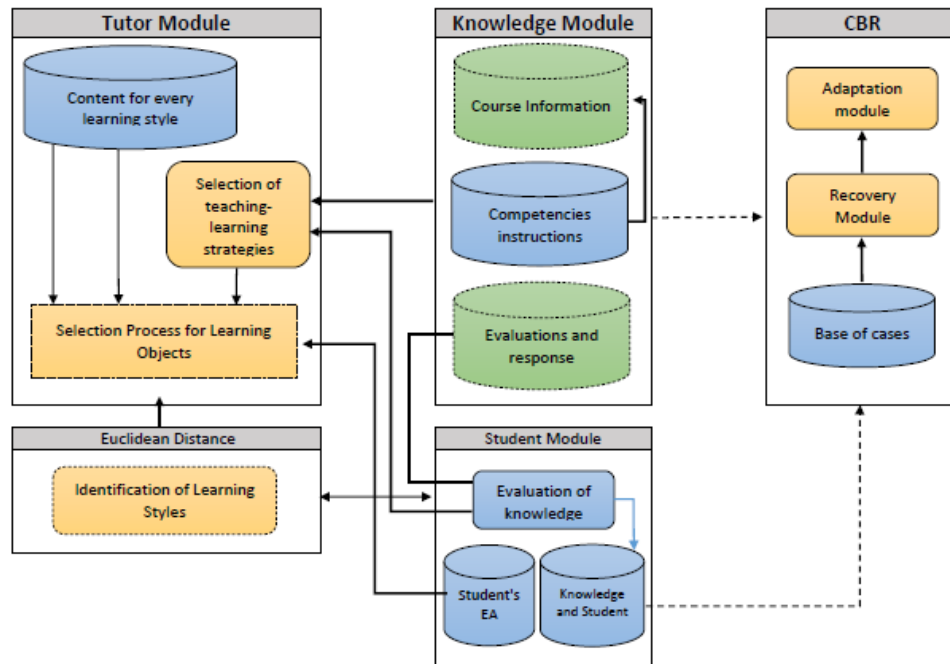

Fig. 3. General Architecture of the Proposed Model.

The knowledge module is influenced by the teaching strategies of the student's learning styles. These teaching strategies will be the link of the learning objects through the teaching-learning strategies applied to the design of the contents of the area [13].

The following is a description of the modifications made to the modules of the general architecture of the Intelligent Tutorial System (ITS).

a) Description of the model: The description of the model incorporates the classic architecture of an Intelligent Tutor System, a process of selection of learning objects (contents), influenced by the teaching strategies of the student's learning styles. [14].

The modifications made to the modules of the general ITS architecture are described.

The tutor module incorporates the teaching-learning strategies considered in the design of the themes of the different courses, as well as the redefinition of the teaching strategies according to the learning style of the student. It also incorporates a process to adapt the content to be presented:

- Identify learning styles through a test of learning styles [5].

- Select the topics to be shown to the student, linking their learning style with the teaching strategies used in the creation of the topics and thus promote their learning.

- In the knowledge module, a database is added that will store the competences of the subject. As well as the use of some metadata in the contents of the course to characterize the competences to be developed.

- The interface module will display the learning objects chosen by the tutor module selection process.

- The Case-Based Reasoning module is added, which is an approach that addresses new problems by reference to similar problems solved in the past. So similar problems have similar solutions.

A database of 199 students according to their learning styles is entered into the Case-Based Reasoning mechanism. This, prior to a case indexing process, which recovers cases using as a measure of similarity the Euclidean distance in n dimensions. Once the evaluation process is over, the winner is reviewed, returning the personalized content according to the learning style entered, if this case is significant it is retained; as can be seen in figure 4.

b) Customization of content according to the student's learning style: The experimentation was carried out in the Virtual Innovation Center, with a sample of 199 students who interacted with the MIGAP platform. In order to determine the predominant learning style, a questionnaire about the preferences of teaching strategies was applied. [15].
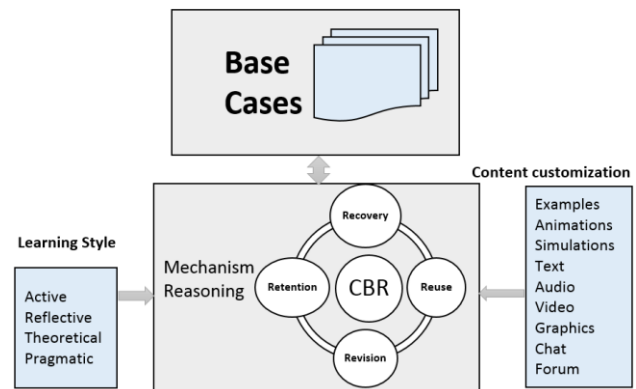

Fig. 4. CBR Proposal to Customize Content.

We also analyzed the frequencies of the learning styles detected in each of the students of each course, to know if these influenced the performance of the students. Figure 5 shows the graph corresponding to the learning styles, where a preponderance of reflexive and pragmatic learning styles can be observed. [16]

The following is a statistical summary of the learning styles detected in the 199 students, which were used as a sample to form the case base.

Fig. 5 shows that 37 students have the active learning style, 59 students have the reflective learning style, 44 students have the theoretical learning style and 59 students have the pragmatic learning style.
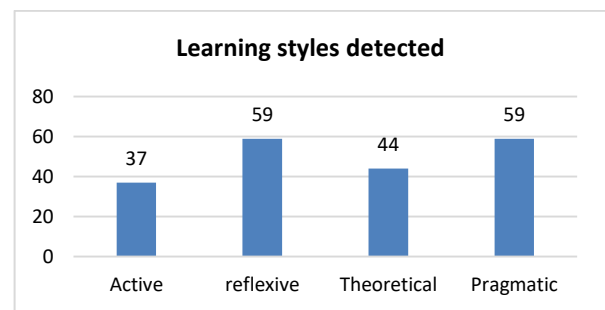

Fig. 5. Learning Styles Detected.

IV. EXPERIMENTATION AND RESULTS

To carry out the evaluation of the proposal with other algorithms, we have used the cross validation technique that is a way to evaluate the results of a statistical analysis. That consists in repeating and calculating the arithmetic average obtained from the evaluation measures on different partitions [17].

It is used in environments where the main objective is a prediction, we have selected it since we want to estimate the accuracy of a model. It is a technique widely used in artificial intelligence projects to validate generated models.

The following graphs show a statistical analysis developed with the different Artificial Intelligence techniques used in the proposal. [10]

Table II shows that our proposal has a precision with a weighted average of 0.99 %. This is due to the fact that we were able to correctly classify a large number of positive instances with an average of 0.99% and classified few negative elements with an average of 0.007 %.

Table III shows that the Simple Logistic algorithm turned out to be the second best in the list; this is due to the fact that although I was able to correctly classify a large number of positive instances, I was also able to classify erroneously negative instances with an average of 0.003.

In the confusion matrix applying the naive Bayesian classifier (Probabilistic classifier) Table IV, based on the Bayes Theorem, we can observe that there is no good classification in learning styles since this Bayesian classifier uses some additional simplified hypotheses and because of these simplifications, which are usually summarized in the hypothesis of independence between predictor variables, is called Naive.

The Table IV shows that the Tree J48 algorithm turned out to be the fourth best classifier since it presents an accuracy with a weighted average of 0.97. This is due to the fact that it was able to classify correctly a positive instance with an average of 0.97 and classified few negative elements with an average of 0.01. We can observe that the TP Rate and Recall measures clearly and unequivocally reflect the number of successes in the classification of the positive instances.

The Table VI shows that the Neuronal Network was the fifth best classifier since it presents an accuracy with a weighted average of 0.95. This is due to the fact that it was able to classify correctly a positive instance with an average of 0.95 and classified few negative elements with an average of 0.018.

TABLE. II. CONFUSION MATRIX APPLYING CBR AND NEURAL NETWORKS

	Active	Reflective	Theoretical	Pragmatic	Facts	Accuracy	Errors
Active	37	0	0	0	37	37	0
Reflective	0	58	1	0	59	58	1
Theoretical	0	0	44	0	44	44	0
Pragmatic	0	0	0	59	59	59	0
						198	1
						99.50%	0.50%

TABLE. III. CONFUSION MATRIX APPLYING SIMPLE LOGISTIC ALGORITHM

	Active	Reflective	Theoretical	Pragmatic	Facts	Accuracy	Errors
Active	37	0	0	0	37	37	0
Reflective	0	57	2	0	59	57	2
Theoretical	0	0	44	0	44	44	0
Pragmatic	0	0	0	59	59	59	0
						197	2
						98.99%	1.005%

TABLE. IV. CONFUSION MATRIX APPLYING NAIVE BAYES ALGORITHM

	Active	Reflective	Theoretical	Pragmatic	Facts	Accuracy	Errors
Active	35	0	0	2	37	35	2
Reflective	0	58	1	0	59	58	1
Theoretical	0	0	44	0	44	44	0
Pragmatic	1	0	0	58	59	58	1
						195	4
						97.98%	2.01%

TABLE. V. CONFUSION MATRIX APPLYING TREE J48 ALGORITHM

	Active	Reflective	Theoretical	Pragmatic	Facts	Accuracy	Errors
Active	37	0	0	0	37	37	0
Reflective	0	54	4	1	59	54	5
Theoretical	0	1	43	0	44	43	1
Pragmatic	0	0	0	59	59	59	0
						193	6
						96.48%	3.01%

TABLE. VI.   CONFUSION MATRIX APPLYING NEURAL NETWORKS

	Assets	Reflective	Theoretical	Pragmatic	Facts	Accuracy	Errors
Active	35	0	0	2	37	35	2
Reflective	0	55	4	0	59	55	4
Theoretical	0	4	40	0	44	40	4
Pragmatic	0	0	0	59	59	59	0
						189	10
						94.97%	5.02%

As a conclusion of the analysis developed, it can be observed that the greatest number of successes in the classification corresponds to the proposed technique of Case-Based Reasoning with 99.50% success and 0.5% error, as opposed to the use of the other techniques used that have a percentage of successes below the proposal. [18].

Fig. 6 shows that the greatest number of cases correctly classified 198, corresponds to the CBR with only one case incorrectly classified, with an average absolute error of 0.0279. After the comparisons with other classification algorithms were made, Simple Logistic followed second, Naive Bayes third, Tree J48 fourth, and Artificial Neural Networks fifth.

The Fig.7 shows the general average grades obtained by the students of each course, where it is observed that the average grades obtained by the students who took the personalized course are 60.5 points, while the average grades of the students who took the non-customized course are 39.5 points, which shows a very considerable difference of 21 points.
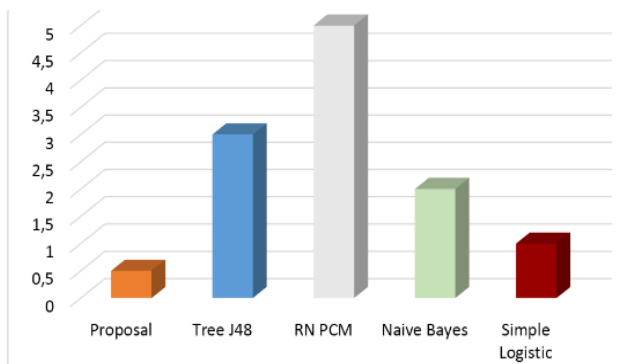

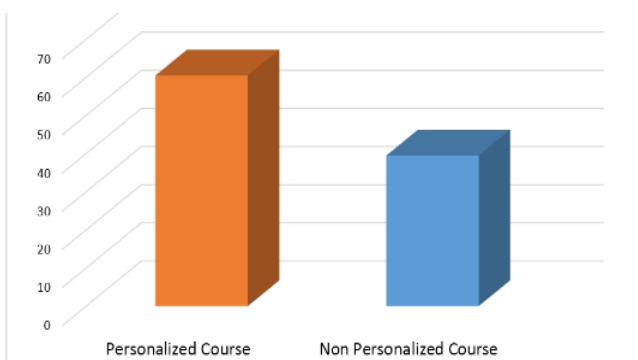Fig. 6.   Comparison of CBR Error Rate with other Techniques.

Fig. 7.   Overall Grade Point Average Per Course.

V.   CONCLUSIONS

- It has been possible to propose an Intelligent Personalized Learning Management Model based on instances of learning objects whose results show that the proposed model has an efficiency of 99.5%; above the models: Simple Logistic with 98.99%, Naive Bayes with 97.98%, Tree J48 with 96.98%, and Neural Networks with 94.97% success.

- A demonstration prototype can be seen at http://benjaminmaraza.my-place.us/ after proposing, experimenting, analysing and evaluating the Case-Based Reasoning, acceptable and significant behaviour is appreciated in the personalization of contents according to the learning style of the students.

- It has been possible to identify the problems in traditional systems as well as the theoretical foundations, determining the aspects that reduce the quality of learning management systems and proposing improvement strategies based on theory.

- A model has been developed that facilitates teaching-learning activities in e-Learning systems.

- The tests with this prototype show that the use of this e-Learning technology would directly affect the educational quality of the region. Allowing to optimize some elements of the learning process that are still traditional in our environment.

- It has been possible to Model the student in the Student Modeller Module attributing flexibility capabilities and personalization attributes provided by a Case Based Reasoning system based on the Euclidean distance in a N dimensional space.

- As future work is recommended the implementation of much more elaborate tests to give robustness to this proposal, since the tests are in the maturation phase and there is much research for the development of indicators to measure quality, flexibility, customization in an LMS system since it has to do with a neurological issue: Education.

REFERENCES

[1]   J. Carbonell. (2014). "An Artificial Intelligence Approach to Computer assisted instruction" .CAI, vol. Vol 4, N° 6.

[2]   W. Murray. (2013). "Control for intelligent tutoring systems: A comparison of blackboard architecture and discourse management networks". Research Report R-6267, FMC Corp.

[3]   Moreno, R., Joyanes, L., Giraldo, L. M., Duque, n. d. y Tabares, V. (2018). Modelo para personalización de actividades educativas

aprovechando la técnica de Razonamiento basado en Casos (RBC). Campus Virtuales, Vol. IV, num. 1, pp. 118-127.

[4]   E. Wenger. (2015). "Artificial intelligence and tutoring systems".

[5]   A. C. y. D. J. Honey, P. (2015). "Learning styles: diagnostic and improvement procedures." 2nd ed. Bilbao: Mensajero Editions, 1994. Pages. 104-116.

[6]   R. Vicari; D. Ovalle. (2010). "ALLEGRO: Ambiente Multi-Agente de Apoyo a la Enseñanza-Aprendizaje utilizando Planificación Instruccional y Razonamiento Basado en Casos (CBR)". In Proceedings of XIII Congreso Iberoamericano de Educación superior en Computaci1th Latin-American Conference on Informatics CLEI, Medellin, Colombia.

[7]   L. Zadeh. (2013). "Foreword of foundations of soft case-based reasoning". Ediciones Berkely, CA.

[8]   G. P. H. y. C. M. Alonso, Domingo J. (1994). "Learning styles: diagnostic procedures and improvement." 2nd ed. Mensajero, D.L. 1994. ISBN: 84-271-1914-3. Bilbao, Spain.

[9]   B. Maraza. (2015). "Intelligent Model of Personalized Learning Management in a Virtual Simulation Environment Based on Instances of Learning Objects". Doctoral Thesis. National University of San Agustín. Peru.

[10]   B. Collis. (2015). "Tele-learning in a digital world".

[11]   G. Verdejo, F. Davies, (1997). "Trends for higher education and training".

[12]   B. Expensive. (2013). "Estrategia de formación planificada de docentes," Ph.D. Thesis, Universidad de San Buenaventura, Universidad de San Buenaventura, Colombia.

[13]   M. Urretavizcay-Loinaz and I. Fernandez de Castro. (2012). "Artificial intelligence and educacia ovierview". In Conference on Advanced Information Systems Engineering www.upgrade-cepis.org/issues/2012/5/upgrade_vIII-j.html.

[14]   D. Cruz, R. Alexandre. (2013). "Technical Report Department of Information Technology and Automatics." University of Salamanca: Adison-Wesley Publishing Company, Inc.

[15]   R. Dunn and K. Dunn. (2014). "Dunn and dunn learning style inventory department of educational and community programs". Queens College, City University of New York.

[16]   Hunt, P. (2017). "Community iswhat i think everyone is talking about, remedial and special education". In ACM.

[17]   C. Peña. (2014). "Intelligent agents to improve adaptivity in a web-based learning environment". Universidad de Girona. España.

[18]   Sankar K Pal and Simon K. Shiu. (2014). Foundation of soft Case-Based Reasoning. Indian Statistical Institute, Hong Kong Polytechnic University. By Jhon Wiley & Sons, Inc 2014.