Proposal of Architecture And Application of Machine Learning (Ml) as A Strategy for the Reduction of University Desertion Levels Due to Academic Factors

Propuesta de arquitectura y construcción de aprendizaje automático (AA) como estrategia para la reducción de los niveles de deserción universitaria debido a factores académicos

Proposta de arquitetura e construção de Aprendizagem Automática (AA) como estratégia para a redução dos níveis de deserção universitária devido a fatores acadêmicos

José Ignacio Rodríguez Molano[1]
Leidy Daniela Forero Zea[2]
Yudy Fernanda Piñeros Reina[2]

Artículo de investigación. https://doi.org/10.16925/2357-6014.2019.03.06

[1] Facultad de Ingeniería. Universidad Distrital Francisco José de Caldas, Bogotá, Colombia
   ORCID: https://orcid.org/0000-0003-2581-277
   E-mail: jirodriguezr@udistrital.edu.co

[2] Facultad de Ingeniería. Universidad Distrital Francisco José de Caldas, Bogotá, Colombia
   ORCID: https://orcid.org/0000-0002-9995-6832
   E-mail: ldforeroz@correo.udistrital.edu.co

Abstract

Introduction: Machine Learning arises as one of the techniques of artificial intelligence, with the development of computer programs that, through algorithms, access data and use them to learn and predict results. Their application in education allows for the characterization of problems or difficulties in learning through the analysis of student performance.

Objective: Identification of applications of Machine Learning that can be applied to the educational field accompanied by a proposal of architecture for the application in an environment of personalized education.

Methodology: This article begins with the review of the literature on the characteristics of Machine Learning and academic desertion, with an emphasis on the Colombian case, the Hyper-personalization and its applicability to learning methodologies. Then, a proposal of architecture in a Machine Learning environment is generated in order to mitigate the academic desertion caused by academic factors. Finally, we propose mechanisms for evaluating the proposed architecture, with a subsequent synthesis and discussion of the results.

Conclusions: The construction of a Moodle architecture for the hyper-personalization of learning, is a global perspective of the representative factors proposed for the development of applications through Machine Learning. This could lead to a decrease in levels of university academic desertion because it facilitates the management of knowledge, information and adaptation through the analysis of scenarios.

Originality: The proposed architecture is shown as an application of machine learning in social cases such as academic desertion, allowing the inclusion of automatic learning models with the requirements of an educational environment.

Restrictions: The case for the application for the Hyper-personalization of learning uses an academic approach which can generate invalid results regarding desertion levels.

Keywords: Machine Learning, Academic Desertion, Hyper-personalization, Education.

Resumen

Introducción: El Machine Learning, surge como una de las técnicas de la inteligencia artificial, en la cual, a través de algoritmos, accede a los datos y los utiliza para aprender y predecir resultados. En cuanto su aplicación en la educación permite la caracterización de dificultades en el aprendizaje a través del análisis de su rendimiento.

Objetivo: Identificación de aplicaciones del Machine Learning aplicado al ámbito educativo que permitan la disminución de los niveles de deserción académica, a través de una propuesta de arquitectura para su aplicación en un entorno de educación personalizada.

Metodología: Se inicia con la revisión de la literatura sobre las características del aprendizaje automático, la deserción académica, con énfasis en el caso colombiano, la hiperpersonalización y su aplicabilidad a las metodologías de aprendizaje; generando a continuación una propuesta de arquitectura en un entorno de Aprendizaje Automático, con el fin de mitigar la deserción académica provocada por factores académicos. Finalmente, se proponen mecanismos de evaluación de la arquitectura propuesta, con una posterior síntesis y discusión de los resultados.

Conclusiones: La construcción de una arquitectura del Moodle de Hiperpersonalización del aprendizaje, es una perspectiva global de los factores representativos propuestos para el desarrollo de aplicaciones a través del Machine Learning, lo cual podría llevar a la disminución de los niveles de deserción académica universitaria, en el sentido en que se facilita la gestión del conocimiento, la información y la adaptación a través del análisis de escenarios.

Originalidad: La arquitectura propuesta se muestra como una aplicación del Machine Learning en casos de tipo social como la deserción académica, permitiendo la inclusión del modelado de aprendizaje automático con los requerimientos de un entorno educativo.

Limitación: El caso de las aplicaciones de Hiperpersonalización del aprendizaje, cuenta con un enfoque académico, lo cual puede generar resultados inválidos de los niveles de deserción, si este no es el factor principal o uno ampliamente significativo en el caso donde se desarrolle.

Palabras clave: Aprendizaje automático, Deserción académica, Hiperpersonalización, Educación.

Resumo

1. INTRODUCTION

The impact of technology on human life is increasing. The use of Artificial Intelligence and Big Data tools with which we interact daily, directly affect the way the world is perceived, our interpersonal relationships and even the methods of human learning.

Today, there is a great variety of technological advances aimed at the educational field; from the creation of specific-purpose applications such as language teaching or applications for memory training, to the development of specialized software in different learning methods for any of the educational levels [1]. For this reason, teachers and educational authorities are concerned not only about the content of teaching programs, but also about the way in which they are transmitted. [2]
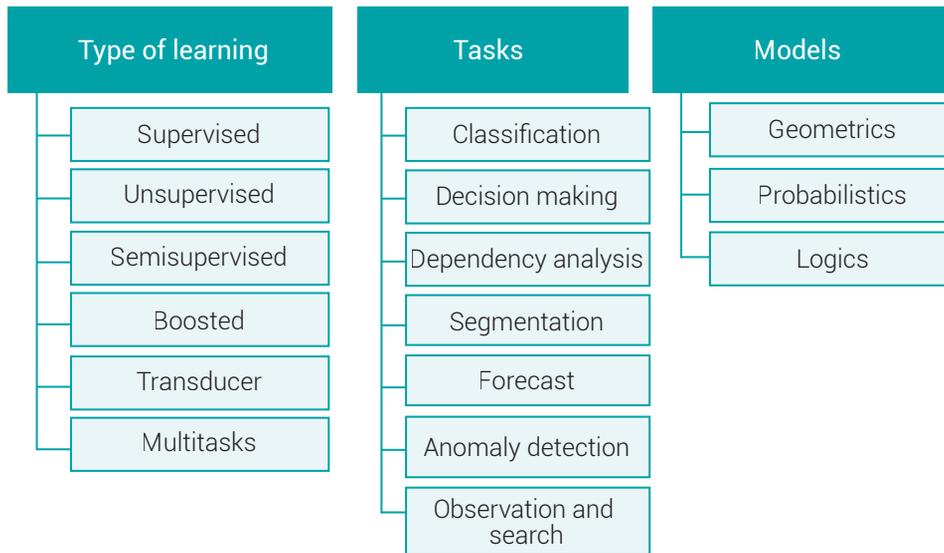
In education, there are three main lines of application in which important changes are perceived that will shape the new educational model and intelligent campuses: process automation, hyper-personalization of services and connected campuses [3]. The following article will approach the use of Machine Learning as a tool of Artificial Intelligence, from the concept of hyper-personalization linked to the improvement of quality in Higher Education, having the reduction of academic desertion levels as its focus.

2. LITERATURE REVIEW

2.1. MACHINE LEARNING

Machine Learning is a type of Artificial Intelligence, in which large volumes of data are accessed in order to interpret them, train the system, and predict new information using learning algorithms. These algorithms are classified according to several criteria: the type of learning, the tasks for which the algorithm is intended to be used, and the types of models used (Figure 1) [4] .

Classification is a task in which, given an individual of the system in question, it pretends to know what type of class it belongs to according to the learning of characteristics, patterns and behaviors that other individuals have previously taken, and for which there is a register [5]. Based on this historical information of individuals, summarized in a series of assigned variables –for example, age, gender, studies, etc.–, the classification algorithms will construct a model that allows a label for a given individual to be obtained as a result. Likewise, these algorithms are divided into two categories: binary classification and multiclass classification. The binary classification refers to those algorithms whose observation result must be catalogued as positive or negative. Therefore, the classification is carried out by means of a threshold with which the score generated for each iteration of the algorithm will be compared. On the other hand, in the multiclass classification, class prediction is made for the highest score [6].

Type of learning	Tasks	Models
Supervised	Classification	Geometrics
Unsupervised	Decision making	Probabilistics
Semisupervised	Dependency analysis	Logics
Boosted	Segmentation	
Transducer	Forecast	
Multitasks	Anomaly detection	
	Observation and search	

Figure 1. Classification of learning algorithms according to different criteria.
Source: own work

Binary classification algorithms

In this article, an emphasis is put on binary classification algorithms since the goal of the algorithm in the present proposal is to predict an "n profile" label for a student whose attributes allow his classification according to the detected patterns; both of personal characteristics provided and the results in the diagnostic tests. Therefore, a synthesis of the existing binary classification algorithms will be carried out below, which may later be evaluated in the generation of the Machine Learning model.

Averaged Perceptron. The inputs are classified into possible results considering the linear function to be then combined into a set of ponderations from which its characteristics can be derived [7].

Bayes Point Machine. This algorithm is a classification technique with a generality that defines that all characteristics are independent and not interrelated. Thus, the status of a specific characteristic within a class does not affect the status of another one [5].

Boosted Decision Tree. A reinforced decision tree is a conjoint learning method in which the second tree corrects the errors from the first tree, the third tree corrects the errors from the first and second trees and so on. The predictions are based in the complete set of trees that comprise the prediction [7].

Decision Forest. The conjoint methods are based on the general principle that, instead of trusting in a single model, a more generalized model can offer better results

by creating multiple models and combining them in some form. In general, the conjoint models offer better coverage and precision than the individual decision trees [7].

Decision Jungle. Decision jungles are a recent extension of decision forests and consist of a set of directed acyclical graphs (DAG).

Logistic Regression. Logistic regression is a statistical method that is used to predict the probability of a result and is especially popular for classification tasks. The algorithm predicts the probability of occurrence of an event by adjusting data to a logistic function.

Two-Class Neural Network. A neural network is a set of interconnected layers. The inputs are the first layers and are connected to an output layer through an acyclic graph made of edges and pondered nodes [5].

2.2. ACADEMIC DESERTION IN COLOMBIAN UNIVERSITIES

Academic desertion is defined as the condition of those persons enrolled in Higher Education who leave the institution for two or more consecutive periods at the time of study [3]. It is a public problem that, in Colombia and the world, affects the efforts made by societies and States to raise the level of training of human beings in order to improve competitiveness and successfully insert themselves into the knowledge society.

According to data obtained by 2015 by the Colombian System for the Prevention of Desertion in Higher Education Institutions -SPADIES-, in the report regarding desertion and graduation statistics of the Colombian Ministry of National Education and taking into account the desertion rate by cohort and by level of training, students who drop out of their undergraduate studies represent 41.60% of the total number of students entering an academic period at the national level and 45.25% at the district level. [8]

When evaluating areas of knowledge based on academic permanence indices (Table 1), at the university level, higher desertion rates are observed in areas related to mathematics and natural sciences, representing 11. 1%; agronomy, veterinary and related areas 10. 2%; and economy, administration, accounting and related areas 10. 1% [8]. However, the desertion rates by cohort for the same year show that the areas with the highest cumulative desertion rates (after 10 semesters) are engineering, architecture, urbanism and other related subjects with 55. 0%. [3]
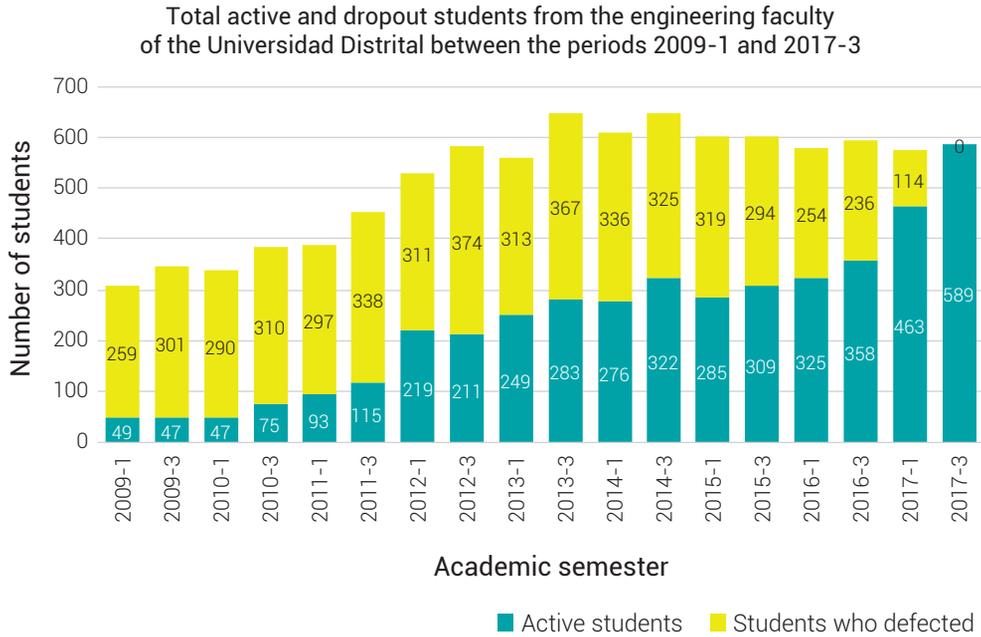
Table 1. Percentage of desertion by cohort by area of knowledge for the year 2015. Adapted by the authors from [3]. Cut-off date: April 2016

| Areas of knowledge	Semester									
	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
Agronomy and veterinary	21,1	30,1	35,4	39,2	41,9	44,8	46,8	47,8	48,5	49,7
Fine arts	20,4	29,1	34,7	39,1	42,7	46,2	48,4	50,2	51,4	52,8
Education Sciences	19,0	26,5	31,4	34,9	37,6	39,9	41,6	43,1	44,7	46,7
Health Sciences	17,8	24,1	28,3	31,6	34,2	36,2	37,8	39,3	40,4	41,9
Social and human sciences	16,8	23,7	27,8	30,9	33,0	34,8	36,2	37,7	39,1	41,9
Economy, adm., accounting	22,3	30,5	35,8	39,9	42,9	45,4	47,2	48,5	49,6	51,6
Engineering and architecture	22,7	32,1	38,3	42,7	46,0	48,9	50,9	52,5	53,6	55,0
Mathematics and natural sciences.	22,0	31,4	37,2	41,2	43,9	46,4	48,0	49,2	50,0	51,0

Source: own work

On the other hand, the indices of academic permanence in the Universidad Distrital Francisco José de Caldas, specifically in the faculty of engineering, are not encouraging. According to the latest report "Statistics of permanence, graduation and desertion of students in the Faculty of Engineering in undergraduate programs 2009-2017"; made by the Systems Advisory Office of the Universidad Distrital -OAS- of students of the Faculty who entered between the periods 2009-1 and 2017-3, 53.8% (or 5038 students in total) withdrew or lost their place as a student. Figure 2 is based upon data taken from the report of statistics of faculty permanence [9].

Within the motives of academic desertion, according to the results consigned in the format for Voluntary and Definitive Retirement from the Faculty of Engineering of the Universidad Distrital GA-FR-FI-01, between the semesters 2016-1 and 2017-1, the academic aspects are the third most recurrent motive for academic desertion of the faculty, with a representation of 14 % of the total of cases, preceded by economic motives that represent 29% and the change-of-career motive that represents 23 % of the total [2].

Total active and dropout students from the engineering faculty
of the Universidad Distrital between the periods 2009-1 and 2017-3

Figure 2. Ratio of active and deserted students between 2009-1 and 2017-3 for the
faculty of Engineering of the Universidad Distrital. Cut-off date: February 2018.
Source: adapted by authors from data provided by the Systems Advisory Office
of the Universidad Distrital Francisco José de Caldas [9]

Total desertion cases between 2016-1 and 2017-1
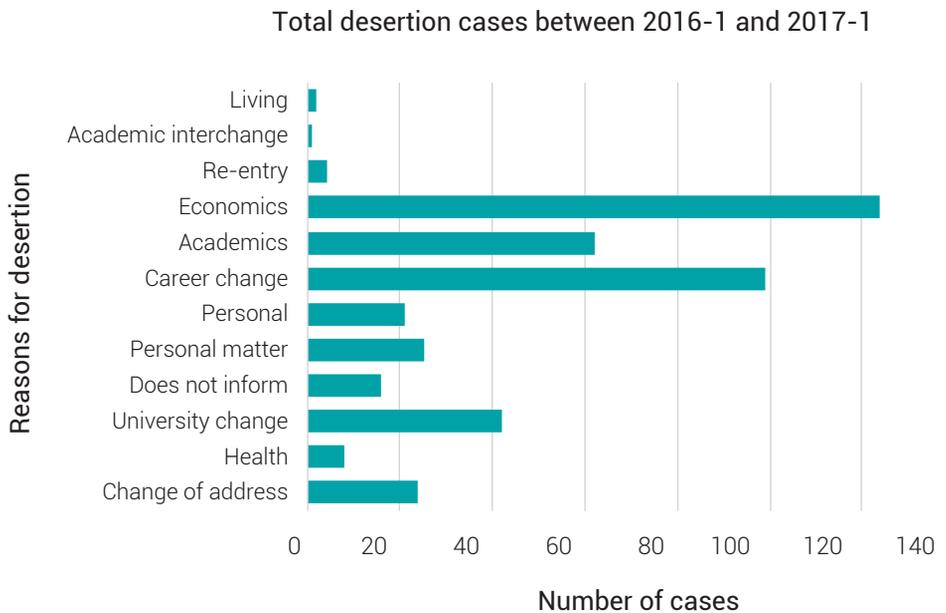
Figure 3. Total number of students who deserted the Faculty of Engineering
of the Universidad Distrital between the periods 2016-1 and 2017-1 along with their
given reasons for doing so.
Source: adapted by authors based on [2]

The Colombian Ministry of National Education is the Colombian governmental entity in charge of formulating national education policy and promoting the development of a competitive and high-quality education. It assumes within its functions, the guarantee of opportunities for access to higher education so that Colombians can access it and benefit from its programs. It deems it necessary to promote the permanence of students in their professional training [10]. In order to carry out an adequate intervention on this phenomenon, it is essential to systematically monitor student desertion, in addition to studying it from different dimensions -higher education institution, academic program and mainly the student– [11]. Viana and Rullan emphasize that social and academic unification are a link to re-evaluate individual goals and responsibilities that ultimately influence decision making about whether or not to stay in school [12]. In this way, it is possible to identify with greater reliability the incident factors in academic desertion and therefore to take corrective and preventive action on them, thereby achieving a better impact, in this case, as a curricular project and as a University in the management of resources in order to comply with the aforementioned objective.

Previous studies have examined the risk of academic desertion, taking into account different attributes for the recognition of patterns of behavior or decisions. For example, Chung & Lee using characteristics of desertion as a function of attendance at the beginning of high school and through the use of the algorithm of random forests, with training data (80 %) equivalent to 132.572 cases, obtained a performance in their predictive model of school desertion with an accuracy of 95 %. [13] . On the other hand, Tan & Shao used two types of attributes, personal characteristics and academic development, of different students of a virtual course of the University of China as input attributes for the predictive model. This was developed with Artificial Neural Networks, Bayes and Decision Tree algorithms. The results obtained correlated well with expectations as the accuracy with which potential desertions were predicted in his experiment was relatively high [14]. In this way, the tendency to study social problems such as academic desertion is evident through the different tools offered by Artificial Intelligence, specifically Machine Learning. These tools are able to obtain predictions with high precision, thus giving them relevance in decision making.

2.3. TECHNOLOGY IN EDUCATION

Incorporating technology into education brings a number of benefits that help improve efficiency and productivity in the classroom, which through the rise of the Internet and mobile technology incorporate even more technological elements into

the educational environment; interactive whiteboards, virtual classrooms and an endless number of electronic resources to conduct research or do schoolwork, are some of the ways in which digital technology has been integrated into schools and universities [15].

In this sense, Burgos indicates that "technology applied to education is a set of theories and techniques that offer a process through which tools can be operated, changed, manipulated and controlled in a learning environment" [16]. However, the existence of technological tools implies the development of adaptive methods in teaching and learning methodologies.

However, the advantages of the availability of ICT (Information and Communication Technologies) in education include: firstly, students have the option of knowing about a given subject matter from different perspectives, where teachers become advisors or guides, since they should focus on giving students tools for intelligent search, critical analysis and the selection and application of knowledge; secondly, teachers have fewer opportunities to develop an obsolete program; and thirdly, students can carry out their learning processes in a more meaningful way, starting from their previous knowledge and experiences, because they have at their disposal a lot of information to choose from and the possibility of requesting and receiving, at any time, the advice of teachers and fellow students [17].

Today, many innovative institutions have focused, as Rob Curtin (Microsoft's global director of higher education) says, on using the intelligent cloud to scale highly personalized learning experiences that go beyond coverage and improve quality. On the other hand, the use of digital books and adaptive learning enables a personalized learning experience as teachers now have access to data about exactly what students consume and how they are dealing with the concepts they are trying to learn. Provided with this information, teachers can see if a student, or the class in general, is not capturing certain ideas or topics [18]. This information allows teachers to identify shortcomings in the transmission of knowledge and the level of reception transformed into learning for their students, and then to take action on the contents and teaching methodologies that are in accordance with the specific needs of their students. This is known as Hyper-personalization, which concludes in an improvement in the quality of education and a decrease in the level of academic desertion (academic base) of the institutions.

2.4. HYPER-PERSONALIZATION OF LEARNING

The learning process is different for each group, and even for each person; this is where the concept of adaptive learning arises, the development of which will facilitate the diversification of the standard educational system thanks to personalized training adapted to the knowledge, aptitudes and abilities of each student.

There is a diversity of computer projects aimed at researching the use of Machine Learning tools to improve automatic curricular design in novel educational technologies [19] such as: the Carnegie Mellon University in Pittsburgh with the EduBand project, which aims to create flexible course modules and use a platform that adapts the topics according to the student's learning, reinforcing the weakest topics and synthesizing those already mastered [20]; the MOOC (Massive Online) courses, which have promoted interactive, online and practical training in which the student can find support 24 hours a day [21]; or the development of the Mc Graw Hill company with Aleks, which is an adaptive, artificially intelligent learning system that provides students with an individualized learning experience tailored to their unique strengths and weaknesses [22].With these tools based on the development of Big Data and Artificial Intelligence, it is possible to identify when a student is at risk of dropping out of the course before it happens and act preventively with specific interventions by teachers or advisors.

3. METHODOLOGY

For this proposal an analysis-synthesis methodology will be accomplished in which each part of the research is studied independently, starting with the analysis in which a separation of the concepts is carried out, studying each one of the specific details. For this process, the information available in academic databases was studied, performing a review of the state of literature about the characteristics of machine learning, and its applicability in practical cases. This is followed by an analysis of the information about academic desertion, with emphasis on the Colombian case, specifically in the engineering faculty of the Universidad Distrital, and some antecedents of investigations in which machine learning has been used in educational cases. The technology and the impact it can have by including it in educational and training subjects is also analyzed along with the concept of hyper-personalization, and the applicability of this concept in learning methodologies. Subsequently, these separate concepts are combined, establishing relationships between these elements in an architectural structure to be introduced into a Machine Learning environment as a tool

(in which the elements, authors and interested parties that make up the architecture were explicitly established and explained0, to strengthen knowledge and skills, in addition to predicting the risk of student failure in different subjects; − all of this with the aim of mitigating academic dropout caused by academic factors such as failure of subjects in the program and weighted average of grades, among others. Finally, an exploration of the results of the synthesis is carried out, establishing evaluation mechanisms for the proposed architecture and concluding with this process.

The methodology for the elaboration of the proposal of architecture is synthesized in Figure 4. Initially, requirements that determine the model were accounted for, making the core functionality of architecture proportional to the proper disposal of quality requirements. Such requests are generated given the influence of stakeholders, this procedure will use the 'use case' diagram (See Figure 5).
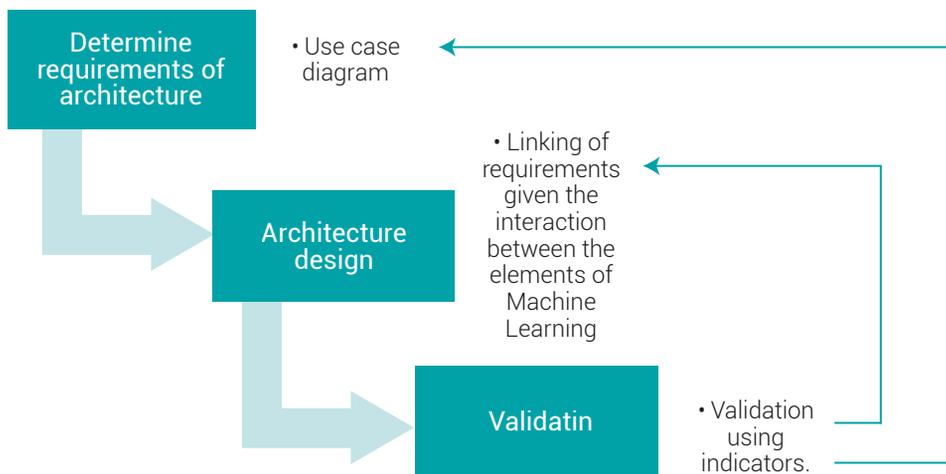
Figure 4. Methodology for architecture development.
Source: own work

Determination of the architectural requirements: creation of the model from requirements that will guide the design of architecture based on the attributes of expected quality.

Architecture design: This involves defining the structure and responsibilities of the components that will included in the architecture. For this purpose, a reference architecture recognized by academia or industry is chosen, and possible styles and patterns are discussed to support compliance with the requirements [23]. This is followed by the component assignment subprocess, whose objective is to define the main components that will be included in the architecture design, identifying the services

and interfaces that each component supports for itself, in addition to determining the dependencies between them [24].

Validation: The validity of the architecture is verified, taking into account the requirements initially defined, using qualitatively measurable characteristics for each of the users, with the objective of identifying possible deficiencies and weaknesses in the design so that they can be improved.

4. RESULTS

A Machine Learning (ML) architecture is proposed as a strategy to reduce the student's risk of dropping out due to academic factors; in this case based on a tool that reinforces knowledge and skills.

4.1. USE CASE DIAGRAM

Below is a diagram with the use cases that are considered important for the custom Moodle system in the context of Machine Learning:
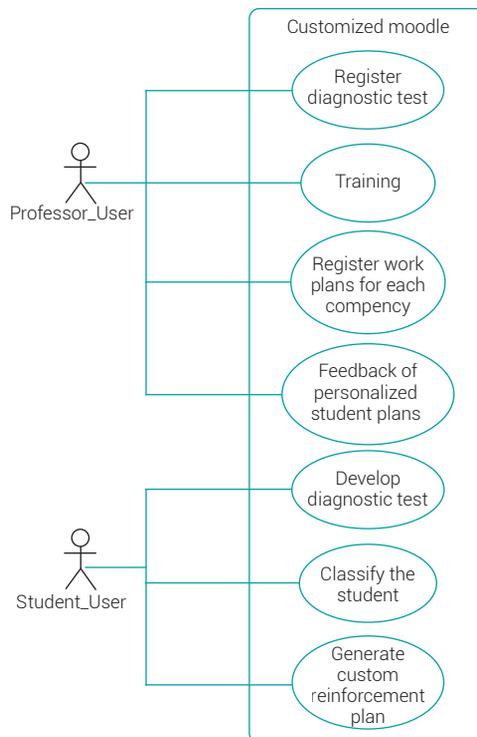
Figure 5. Use-case diagram for custom Moodle in context of Machine Learning.
Source: own work

As it is possible to observe in Figure 5, there are seven use-cases which are described below:

Register the diagnostic test

The professor enters the diagnostic test into the system that will generate input information for each student for the classification process.

Training: based on the possible results for the test, the professor will train the neural network using the possible labels for subsequent classification.

Register work plan for each competency

According to the possible scenarios resulting from the classification of the students, the teacher will enter certain help materials, procedures and other relevant resources to strengthen learning competencies, which will then be shown to the student who requires them.

Feedback from personalized student plans

The professor will receive a consolidated classification of each student, along with the generated work plan and related additional information.

Develop diagnostic test

Students will be able to do the diagnostic test related to their knowledge of the subject, in addition to other qualifying aspects according to variables such as gender, restart conditions of the subject, grades of previous subjects, among other factors.

Classify student

Based on the results of the previous use-case, students will be classified according to the competencies they have not yet developed, along with a prediction of the risk of reprobation of the subject, information that will be shown to the professor in the case of the use of a feedback work plan.

Generate custom reinforcement plan

Students will be able to check the suggestions of the road map according to their condition in such a way that it allows them, through this plan of academic reinforcement, to overcome their weaknesses and difficulties in the subject and at the same time to develop the competences for which it was designed.

4.2. ARCHITECTURE

The following figure (Figure 6) shows a proposed architecture for the custom Moodle system in the context of Machine Learning, in which four layers are defined that seek

the integration of the different elements and actors. The following describes the components of the architecture considering the use-cases of the system, which has two users (teacher and student), an API Rest, a communication layer, a Machine Learning modeling layer, and a visualization layer.
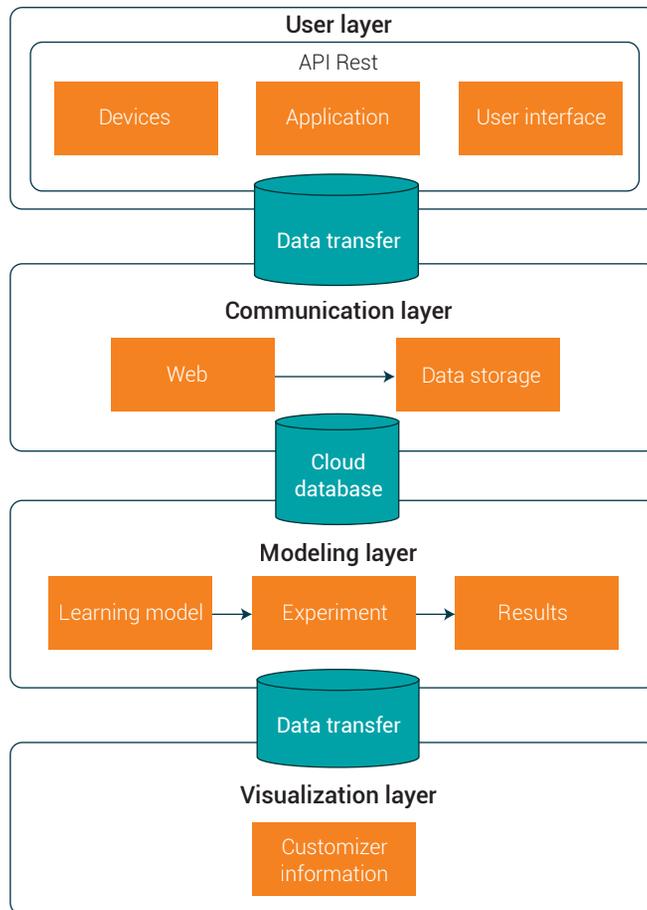
Figure 6 . Architecture proposal for personalized Moodle in the context of Machine Learning.
Source: authors based on [25]

User layer

This layer contains the user interface between the systems to obtain data or generate operations on that data in all required formats. The developed application must be entered through a user account, authenticating and authorizing the entry to ensure the integrity of the information contained, thus managing user access to the application, including activities of information management, service management and technical management.

Communication layer

In this layer is the data storage infrastructure on the web for the subsequent use of data and the generation of an integration between the information received and that provided to the other layers of analysis.

Modeling layer

This is where the platform with an adapted environment for the use of Machine Learning (e.g. Azure Machine Learning, Amazon ML, etc.) intervenes. The model is presented according to the information provided by the databases contained in the cloud, the learning and training experiment, and finally the classification and prediction results based on the input information.

Visualization layer

In this layer are all the webpages that are developed towards the user; from requests for information, errors, and of course the display of personalized information generated from the learning results of the previous layer.

For a proper evaluation of the proposed architecture, we suggest a survey using 2 criteria to understand the opinion that users will have regarding platforms with a Hyper-personalized approach.

Evaluation indicators for proposed architecture layers

For an adequate evaluation of the proposed architecture for Hyper-personalization, a survey is suggested for each actor involved regarding the use of the platform, where the priority level is the same for each of them:

Student

Criterion 1: Ease of use

  a) Access to the platform: User and password management

  b) Test application: Compressible and manageable interface for test and evaluation solution.

  c) Understanding the results: The reading of the test results is understandable.

$$Ease\ of\ use = \frac{Number\ of\ approved\ factors}{Total\ factors}\ (Ec.1)$$

All aspects (3/3) must be met for it to be considered an easy-to-use platform.

Criterion 2: Personalization's level of sessions
a) Relevance of the themes: Reinforced subjects are consistent with subjects with which the student has academic difficulties.
b) Reinforcement method: The reinforcement methods proposed by the professor facilitate and improve the understanding of the subject.
c) Timely attention: Time and quality in the resolution of doubts is appropriate

$$Personalization's\ level\ of\ sessions = \frac{Number\ of\ approved\ factors}{Total\ factors} \quad (Ec.\,2)$$

Professor

Criterion 1: Ease of use
a) Access to the platform: User and password management
b) Test evaluation: Compressible and manageable interface for scoring tests and evaluations.
c) Understanding of the results: The visualization of the classification of the results of the students is simple and identifies the deficiencies in the evaluated thematic.

$$Ease\ of\ use = \frac{Number\ of\ approved\ factors}{Total\ factors} \quad (Ec.\,3)$$

Criterion 2: Personalization's level of sessions
a) Relevance of the themes: The classification of the reinforced topics is directed to the students who, according to the results, require it.
b) Reinforcement method: The platform provides tools to improve the reinforcement methods proposed by the professor.
c) Timely attention: The software has a filter of frequently asked questions and facilitates the task of the professor in resolving doubts in a timely and effective way.

$$Personalization's\ level\ of\ sessions = \frac{Number\ of\ approved\ factors}{Total\ factors} \quad (Ec.\,4)$$

Qualitative evaluation of the criteria:

Table 2. Qualitative evaluation of performance measurement criteria of the proposed architecture.

Affirmative answers/ Total answers	Criteria level
3/3	High
2/3	Medium
1/3	Low

Source: own work

Comparison of academic results

$$Level\ of\ Improvement\ in\ Results(NMR) = \frac{Results\ after\ the\ process}{Results\ before\ the\ process}(Ec.5)$$

Level of improvement in results
If NMR >1 then an improvement in the student's academic performance is concluded.
If NMR <1 then a decrease in the student's academic performance is concluded.
If NMR=1 there was no difference between the academic results before and after the process is concluded.

It is important to clarify that these weights are quantitative, strictly based on the grades obtained by the student, and do not take into account the psychological, social or economic factors that may affect the final result.

Academic desertion level

$$Academic\ desertion\ level(NDA)$$
$$= \frac{Real\ number\ of\ students\ who\ drop\ out\ at\ the\ end\ of\ the\ period}{Expected\ number\ of\ students\ who\ drop\ out\ at\ the\ end\ of\ the\ period}(Ec.6)$$

If NDA <1 then an improvement in the results of the level of academic decision-making is concluded.
If NDA >1 then a deterioration in the results of the level of academic decision making is concluded.
If NDA =1 it is concluded that there was no difference between the level of theoretical and actual academic decision making.

It is important to clarify that these weights are quantitative, strictly based on the academic factor. The expected number of students who desert can be obtained from statistical measurements, however, Machine Learning provides support tools in this regard.

5. DISCUSSION AND CONCLUSIONS

In conclusion, the development of an architecture such as the one presented in this document, whose objective is the Hyper-personalization of education, allows for great advances in the academic field from the perspective in which the learning process depends on each individual in question. Therefore, the characterization and analysis of the information specific to Artificial Intelligence allows for an improvement in the academic level of higher education, assuming that new trends in technological tools are included in the development of education.

Although the article presents evaluation indexes to be considered during the development of the architecture, it should be clarified that the judgment of each developer is independent, and this allows him to propose or eliminate criteria that he considers relevant for the subsequent evaluation. For academic aspects, the rate of failure of subjects is one of the most repetitive causes for academic university desertion. The reduction of this rate through the use of technological tools such as the one presented in this article is a potential solution to mitigate this collateral effect of abandoning higher education.

The completion of university courses is of global concern, so it is necessary to have alternative methods to assess the causes and mitigate the levels of student desertion, allowing innovative solutions and demonstrating that the use of Artificial Intelligence contributes to the improvement of social problems.

It should be noted that although the insertion of technology into education has generated a positive impact on improving academic outcomes and facilitating knowledge and information management, these tools are based on the teacher's strategies and methods of teaching. It is the teacher who must contemplate the types of intelligence and how their students manage to retain the information. This is a foundation for programming tools that "learn" based on different scenarios. All this without leaving behind the work of the student and that their responsibility and commitment is in the same fundamental measure for their formation. The construction of these applications should take into account, among other things, indicators that measure and compare the student's commitment to their formative process and academic

results; which can also be a way to validate whether the strategies of education and personalization of information are being effective.

On the other hand, the proposed architecture for the personalized Moodle in the context of Machine Learning (Figure 6), allows for the contemplation, in a global way, of the information and factors necessary for the construction of this type of digital tool. Now, focusing on the Modeling Layer section, where the use of Machine Learning (ML) is concentrated, it should be borne in mind that ML software development companies have different approaches, be it commercial, scientific, research, financial, among others. For this reason, it is important to know the programs that allow the developer a concrete answer based on his objective.

6. REFERENCES

[1]   E. Himmel, "Modelo de análisis de la deserción estudiantil en la educación superior," *Calidad en la Educación*, no. 17, pp. 91–108, May 2018. [Online]. doi: http://dx.doi.org/10.31619/cale-du.n17.409

[2]   M. Sánchez, L. Cruz, and R. Ferro, "Modelo de aproximación al comportamiento de la de-serción voluntaria universitaria en pregrados de Ingeniería periodo 2015-2018," *Ingeniería solidaria*, vol. 14, no. 26, pp. 1–27, 2018. [Online]. doi: https://doi.org/10.16925/in.v14i26.2452

[3]   L. A. Melo-Becerra, J. E. Ramos-Forero, and P. Oswaldo Hernández-Santamaría, "La educa-ción superior en Colombia: situación actual y análisis de eficiencia," *Desarrollo y Sociedad*, pp. 59–111, 2017. [Online]. doi: 10.13043/DYS.78.2

[4]   I. G. Maglogiannis, *Emerging artificial intelligence applications in computer enginee-ring: real word AI systems with applications in eHealth, HCI, information retrieval and pervasive technologies*. IOS Press, 2007. [Online]. Available: https://books.google.es/books?id=vLiTXDHr_sYC&lpg=PA3&dq=classification%20machine%20learning%20al-gorithms&lr&hl=es&pg=PA3#v=onepage&q=classification%20machine%20learning%20algo-rithms&f=false

[5]   D. Sisodia and D. S. Sisodia, "Prediction of Diabetes using Classification Algorithms," *Procedia Computer Science*, vol. 132, pp. 1578–1585, Jan. 2018. [Online]. doi: https://doi.org/10.1016/j.procs.2018.05.122

[6] S. B. Kotsiantis, I. D. Zaharakis, and P. E. Pintelas, "Machine learning: a review of classification and combining techniques," *Artif Intell Rev*, vol. 26, pp. 159–190, 2006. [Online]. doi: 10.1007/s10462-007-9052-3

[7] M. Copeland, J. Soh, A. Puca, M. Manning, and D. Gollob, "Microsoft Azure Machine Learning," *Microsoft Azure*, 2015. [Online]. Available: https://link.springer.com/content/pdf/bfm%3A978-1-4842-1043-7%2F1.pdf

[8] Ministerio de Educación, "Estádisticas de deserción y graduación," pp. 1–4, 2016. [Online]. Available: https://www.mineducacion.gov.co/sistemasdeinformacion/1735/articles-357549_recurso_5.pdf

[9] Universidad Distrital, "Estadística de la permanencia, graduación y deserción de los estudiantes en la Facultad de Ingeniería en programas de pregrado 2009-2017," pp. 13–19,47–56, 2018. [Online]. Available: http://www1.udistrital.edu.co:8080/documents/11171/e49da6d0-2402-43b0-a554-3136d196e8fa

[10] S. P. Barragán Moreno and L. González Támara, "Acercamiento a la deserción estudiantil desde la integración social y académica," *Revista de la Educación Superior*, vol. 46, no. 183, pp. 63–86, Jul. 2017. [Online]. doi: https://doi.org/10.1016/j.resu.2017.05.004

[11] Ministerio de Educación Nacional, *Deserción estudiantil en la educación superior colombiana*. pp. 33–134, 2009. [Online]. Available: https://www.mineducacion.gov.co/sistemasdeinformacion/1735/articles-254702_libro_desercion.pdf

[12] A. Viana, N., Rullán, "Reflections about school dropout in Finland and Puerto Rico," *Education Policy Analysis Archives*, vol. 18, pp. 1–33, 2010. [Online]. doi: http://dx.doi.org/10.14507/epaa.v18n4.2010

[13] J. Y. Chung and S. Lee, "Dropout early warning systems for high school students using machine learning," *Children and Youth Services Review*, pp. 346–353, 2018. [Online]. doi: https://doi.org/10.1016/j.childyouth.2018.11.030

[14] M. Tan and P. Shao, "Prediction of student dropout in E-learning program through the use of machine learning method," *International Journal of Emerging Technologies in Learning*, vol. 10, no. 1, pp. 11–17, 2015. [Online]. doi: http://dx.doi.org/10.3991/ijet.v10i1.4189

[15] P. Marqués, *Las Tecnologías de la Información y las Comunicaciones en el sistema universitario español. Conferencia de Rectores de las Universidades Españolas (CRUE)* pp. 41–81, 2008. [Online]. Available: https://ddd.uab.cat/pub/dim/16993748n0/16993748n0a6.pdf

[16] A. Lozano and J. Burgos, "El reto de la radio interactiva y la tutoría virtual," in *Tecnología Educativa en un Modelo de Educación a Distancia Centrado en la Persona*, 1st ed., México: Limusa, 2015, pp. 241–276. [Online]. Available: https://revistas.ucr.ac.cr/index.php/rlm/article/download/19692/19771/

[17] R. Canales and P. Marquès, "Factores de buenas prácticas educativas con apoyo de las TIC Análisis de su presencia en tres centros educativos," *Educar*, pp. 115–133, 2007. [Online]. Available: https://ddd.uab.cat/pub/educar/0211819Xn39/0211819Xn39p115.pdf

[18]   Revista Semana, "Uso de la tecnología en la educación," *Uso de la tecnología en la educación*, 2017. [Entrevista R. C.]. [Online]. Available: https://www.semana.com/educacion/articulo/uso-de-la-tecnologia-en-la-educacion/539903

[19] M. I. Jordan and T. M. Mitchell, "Machine learning: Trends, perspectives, and prospects.," *Science (New York, N.Y.)*, vol. 349, no. 6245, pp. 255–60, Jul. 2015. [Online]. Available: https://www.cs.cmu.edu/~tom/pubs/Science-ML-2015.pdf

[20] INRIA, "Applying machine learning to education," *INRIA (The French National Institute for Computer Science and Applied Mathematics)*, 2016. [Online]. Available: https://www.inria.fr/en/centre/lille/news/applying-machine-learning-to-education

[21] K. S. Hone and G. R. El Said, "Exploring the factors affecting MOOC retention: A survey study," *Computers and Education*, pp. 157–168, 2016. [Online]. doi: https://doi.org/10.1016/j.compedu.2016.03.016

[22] P. L. Smith and C. L. Dillon, "Comparing distance learning and classroom learning: Conceptual considerations," *American Journal of Distance Education*, vol. 13, no. 2, pp. 6–23, Sep. 2009. [Online]. doi: https://doi.org/10.1080/08923649909527020

[23] D. Garlan et al., *Documenting Software Architectures:Views and Beyond*, Second Edi. Pearson, pp. 12–96, 2010. [Online]. Available: http://ebooks.bharathuniv.ac.in/gdlc1/gdlc1/Computer%20Science%20Books/20110711documenting-software-architectures-views-and-beyond-2nd-edition.pdf

[24] L. Bass et al., *Software Architecture in Practice Third Edition*. pp. 78–97, 2013. [Online]. Available: https://jegadeesansite.files.wordpress.com/2018/01/sei-series-in-software-en-gineering-len-bass-paul-clements-rick-kazman-software-architecture-in-practice-addi-son-wesley-professional-2012.pdf

[25] Microsoft Azure, "Azure Machine Learning Studio Documentation - Tutorials, API Reference | Microsoft Docs," *docs.microsoft.com*, 2017. [Online]. Available: https://docs.microsoft.com/en-us/azure/machine-learning/studio/.