Learners’ Activity Indicators Prediction in e-Learning using Fuzzy Logic

Sanae CHEHBI¹, Chakir FRI³
PHD Student at Computer Science Department Faculty of Sciences, Moulay Ismail University of Meknes
Meknes, Morocco

Rachid ELOUAHBI²
Higher Education Professor Department of Computer Science, Faculty of Sciences, Moulay Ismail University of Meknes, Meknes, Morocco

Abstract—With the idea of introducing computer supports in education, Online Learning (named also e-learning) associated on one hand, the concept of network, therefore that of direct and concepts of communicating interaction, whether between the learner and the teacher (or tutor), or between the learners themselves and on the other hand exchanges and collaboration. Any activity in e-learning leaves recorded traces stored in a database system. Until now, data on student activity is stored as low-level information; however, the volume of this information is too large to be processed and interpreted by tutors, requiring data collection and preparation to give it meaning. In addition, according to the studies carried out in this direction, the tracking of learners must be guaranteed in all stages of e-learning process, to assist and help them when they encounter problems that they cannot solve. The lack of direct contact between the tutor and the learners can cause a lack of feedback of the learning activity; all these problems can lead to a high rate of abundance in e-learning. Our work aims to develop a model for predicting learner activity indicators using fuzzy logic without going through rigid calculations but based on consultation traces and skill assessment scores. Based on the traces collected from the Learning Management System (LMS) Moodle, it could give the tutor high level processing of the learning activity.

Keywords—e-Learning; tracking; Moodle; traces; activity indicators; fuzzy logic

I. INTRODUCTION

To give learners more choices of methods and means of learning, e-learning offers several advantages such as: accessibility, personalization, adaptability and profitability which qualify it as the best solution. Diversified forms of learning make it possible to increase the level of knowledge acquisition.

During the learning activity, learners leave data from which one can deduce interesting facts and describe well the learning process, called learning traces. These traces are recorded and stored in the LMS database system as low level information [1]. This information is abundant on interactions between students and teachers as well as on access to resources and the system. They can give an idea of how and when students perform their homework and tasks, their engagement in lessons, etc. In a context of e-learning and based on learning traces, teachers have a very partial view of the activity of learners and can on this basis, make judgments about the activity of the learner within the learning system [2].

However, the large volume of this information requires the collection of these traces, preparation and processing for this provided information to be meaningful and transformed into actionable knowledge, which is a difficult task [3]. As Peraya (2019) underlines, learning traces have significant potential for use, whether it is to predict certain learner behaviors, to visualize learning activities or to set up adaptive systems [4]. Authors have got the idea to develop a fuzzy logic system for learners’ activity indicators prediction, to assist tutor in tracking learners and give him a clear idea about the course of the learning activity and help him in decision making.

The paper is organized by following three sections. The first sections is to start by the collection and the transformation of traces from the Learning Management System (LMS) Moodle. The second one is to introduce fuzzy logic and its use and contribution in the context of activities’ indicators in e-learning LMS. The final sections is to presenting the implementation and simulation of the fuzzy activity indicators system.

II. TRACKING LEARNERS IN E-LEARNING

A. e-Learning

e-Learning offers several advantages, among which we cite: ease of access to information, flexibility and autonomy in time and space, variety of learning methods, personalization and individualization of content.

But, it has some limitations such as: the absence of human contact between learners and teacher, the need for greater passion and more rigorous work discipline.

B. Tracking Learners

Capturing the attention of learners is a major challenge whether in “classroom” or in “e-learning”. By using an LMS, the teacher does not interact directly with the learners to determine whether they have understood the course well or whether the course is suitable for their levels. This is the reason why researchers introduced tracking in LMS [5].

The crucial task of tutors is to reflect the evolution of the learners, to measure the quality of the training, to highlight the specific knowledge built by each one. This is why it is necessary to collect a set of information on the educational pathways of the learners.
Any activity carried out through software leaves traces recorded in a file, from which one can deduce interesting facts. Since training in Information Technology (IT) has existed, researchers have collected and analyzed the traces left by learners during an e-learning session.

Romero wondered if traces always used for the benefit of the learner? What happens when the traces can be used to justify a negative assessment of the learner’s engagement or to assess their participation in a team activity? [3].

1) Trace definition: Among all definitions given to what is a trace, we opt for "The trace carries meaning as any structured representation of inscriptions of knowledge can be (a classification, a reading) with the only difference that this structuring is not voluntary, declarative, but induced by the use of the inscriptions itself; as such, the trace is itself considered as an inscription of knowledge" [6].

C. Background and Related Works

Up to now, a large number of specific learners’ tracking tools have been developed to solve online educational problems. These tools are used in different educational environments: Learning and Management Systems (LMS), Intelligent Tutoring Systems (ITS), Adaptive and Intelligent Hypermedia Systems (AHIS), MOOC (Massive Open Online Course). Thus, it is impossible to review or compare our approach against others that already exist. To learn more, you will find a brief description of the specific tools for Moodle and others below.

Florian et al., focused on analyzing social perspectives to access Moodle tracking data. They tried to reuse the tracking data from the Moodle LMS so that they could model learners and groups. To create rich learning models they used data from Moodle diaries to gain insight into learner activities in a social context and to ensure support for learning. The authors implemented architecture to have a flexible and extensible interface to Moodle’s tracking data to transform the collected data into learning analytical information [7].

Conde and al., conducted a study that describes the different existing tools that facilitate the extraction and analysis of educational data for learning analysis. It focuses on two differentiated parts namely: a presentation of the different learning analysis tools that analyzed it, then a work that covers the results, including a comparison of the tools after analyzing the data sets of existing courses, this study summarized the results of the application of the different tools [3].

Iglesias-Pradas, Ruiz-de-Azcárate and Agudo-Peregrina explored the applicability of learning analysis for predicting the development of two transversal skills: teamwork and engagement [8]. They were based on the analysis of logs of Moodle interaction data for the benefit of master’s students. The results of the study call into question the appropriateness of the approach and show no relationship between online activity metrics and teamwork and engagement acquisition.

Djouad and Mille proposed to use the traces of the learning activity to develop an indicator management system based on traces (TB-IMS) independent of the platform. The approach adopted made it possible to create and reuse learning indicators independently of the source code of the learning platform. This work presents the theory and its implementation in a first TB-IMS which is illustrated through a real learning situation in Moodle [9].

Gamie and al., introduced a model to analyze and predict students’ performance based on two dimensions: teaching style, and e-learning activities. They collected data from educational settings within an academic institution. They analyzed data used to reveal knowledge and useful patterns from which critical decisions could be made [10].

Hernandez-Garcia et al., Proposed the design of a data extraction, transformation and loading (ETL) system of educational data from the LMS Moodle. Their design was based on indicators of teamwork at the individual and team level across four dimensions: communication, cooperation, coordination and monitoring / tracking. This design aims to transform data from teamwork activities, retrieved from Moodle, into useful information about teamwork behaviors [11].

Mazza et al., developed the GISMO tool to extract monitoring data from an online course, it queries the learning traces recorded by the Moodle platform and automatically generates various graphic representations that can be used by tutors. In order to provide visualizations of the course's behavioral, cognitive and social data, allowing constant monitoring of student activities, engagement and learning outcomes [12].

From what precedes, it is obvious that, if there are differences between all these researches in the axes treated and the methods of resolution proposed, they have as a common objective the improvement of learning in order to encourage tutors to make decisions geared towards improving the learning process.

D. Moodle Traces Collection

The during an e-learning session using an LMS, every user is supposed to leave traces of his activity recorded in specific files, from which it is sometimes possible to deduce some interesting facts. This is why the researchers worked on the collection, preparation, processing and analysis of the traces left during the use of the LMS, to ensure the follow-up of learners and to improve the systems put in place. They drew attention to the wealth of indicators extracted which possibly return relevant information on the progress of the learning activity.

To collect traces we have conducted a study during confinement in a secondary school with students in their final year. In the Fig. 1, the connection log files from the LMS. It contains the session information such as: connection time, the event context and name, the component, the description, the origin of the connection and the IP address.

When consulting the participation of learners in the course, we get a table that contains the number of consultations by the number of learners of a given resource and the last access to the resource, as shown in Fig. 2.
The LMS Moodle store a large number of information about teacher student interactions about access to resources and system, it produce many traces as low-level information. This information describes how and when students perform their missions, tasks and course engagement, but it is difficult for tutors to support learners in their educational course, because the information big volume it is difficult for tutors to deal with, which requires data collection and treatment in order to deliver meaningful information.

In this part, we place the traces cited in the literature and which appear among the factors influencing the learning process. We propose a model to analyze and predict students’ learning performance based on three indicators; session indicator, productivity indicator and interactivity indicator. We collect data from Moodle and analyze it in order to be used to reveal knowledge and useful information from which decisions could be made by tutors in time.

These traces are: the degree of interest expressed by students to e-learning in general, the frequency of Internet connection per week, the frequency of consultation of the LMS per week, the average learning time, the average number of course units visited during a learning session, the degree of use of the chat and / or forum, the number of messages sent by chat and / or forum during a learning session, learners’ productions and the score obtained in the course unit validation test.

We decide to predict the first activity indicator named session indicator based on the frequency of consultation of the courses on the platform per week, the average learning time, the average number of course units visited during a learning session, the degree of use of the chat and / or forum, the number of messages sent by chat and / or forum during a learning session.

In this part, we place the traces cited in the literature and which appear among the factors influencing the learning process. We propose a model to analyze and predict students’ learning performance based on three indicators; session indicator, productivity indicator and interactivity indicator. We collect data from Moodle and analyze it in order to be used to reveal knowledge and useful information from which decisions could be made by tutors in time.
III. ESTABLISH A FUZZY LOGIC SYSTEM TO CALCULATE ACTIVITY INDICATORS

Among the known techniques of artificial intelligence, we have chosen fuzzy logic because it is an approach based on human reasoning rather than on rigid calculations which is need in our study case. The fuzzy logic reasoning method is more intuitive and closer to human reason and thinking unlike classic logic. It allows a better understanding and interpretation of natural phenomena, thanks to the inference rules and the membership functions of fuzzy sets. These phenomena are often imprecise and difficult to model, which is applicable in our case where we measure the performance of learning activities in e-learning LMS.

A. Introduction to Fuzzy Logic

Nowadays, fuzzy logic is an important research area on which many scientists focus. Technological benefits are already available. The theoretical foundations of fuzzy logic were formulated in 1965 by Professor Lotfi A. Zadeh, of the University of Berkeley in California [13]. Fuzzy logic is based on fuzzy set theory, which is a generalization of classical set theory. By introducing the notion of degree in the verification of a condition, thus allowing a condition to be in a state other than true or false, fuzzy logic confers a very appreciable flexibility to the reasoning using it, making possible to take into account imprecisions and uncertainties [14].

1) Fuzzy sets and subsets: One of the interests of fuzzy logic to formalize human reasoning is that the rules are stated in natural language.

Definition: Let X be a set. A fuzzy subset A of X is characterized by a membership function \( \mu : X \rightarrow [0,1] \).

Note: this membership function is the equivalent of the characteristic function of a classic set [14].

According to the usual practices of the literature, the terms fuzzy subsets and fuzzy sets are used interchangeably. Classical sets are also called net sets, as opposed to fuzzy, and similarly classical logic is also called Boolean or Binary logic.

The Fig. 3 shows graphically the difference between a classic set and a fuzzy set.

2) Membership function: In our case, we take example of the first activity indicator, which is the session indicator; we will have to redefine membership functions for each fuzzy subset of each of our four variables:

Input 1: the frequency of consultation of the course. Subsets: low, medium and high.

Input 2: the duration of the learning session. Subsets: low, medium and high.

Input 3: the number of course units studied. Subsets: low, medium and high.

Output: the session indicator. Subsets: low, medium and high.

In a reference set X, a fuzzy subset A of this frame of reference is characterized by a membership function \( \mu \) of A, which associates with each element x of X, the degree \( \mu A (x) \), between 0 and 1, for which x belongs to A. The value of the membership function at element x represents the "grade of membership" of x in A.

This function is the extension of the characteristic function of a classical subset. It can be represented as a Triangular, Gaussian, Trapezoidal or parabolic function. For the sake of clarity and to facilitate calculations, we will only use the first two forms.

We present a comparison between a characteristic function of a classical set and a membership function of a fuzzy set, in Fig. 4.

Fuzzy logic is applied in many fields (industrial processes, commerce, chemical industry, etc.) where it gives very satisfactory results. In our case, it is applied to the prediction of indicators of learner activity.

The fuzzy logic system is built by choosing independent variables that describe well the dependent variable. Fuzzy sets in human language are used to describe a variable instead of using the numeric value. The membership function determines the degree of certainty that each variable belongs to a fuzzy set. The inference engine is responsible for applying each of the inference rules. They represent the knowledge that we have of the system due to human expertise. Each rule will generate an exit command. Then comes the step allowing merging different commands generated by the inference engine to give only one output command and to transform this linguistic output variable into numerical data.
B. Fuzzy Logic applied to Predict Activity Indicators

1) Data sources and simulation software: For the realization of this model we used data collected from the Learning Management System Moodle, this information on the secondary students in their last year as well as on their history of login files and course participation reports. We used Matlab for modeling the fuzzy system. We capture all input membership functions and the output function. The inference rules are also captured.

2) Session indicator fuzzy inference system: The first activity indicator named Session indicator is based on the frequency of consultation of the courses on the platform per week “freqCons”, the average learning time “dureeSess”, the average number of course units visited during a learning session “nbrUnit”; it is represented in Fig. 5.

The inputs variables of the Session indicator fuzzy system are the freqCons, dureeSess and the nbrUnit. After the process of Fuzzification and Defuzzification we obtain the output variable: “IndicSess”.

3) Productivity indicator fuzzy inference system: The second indicator is the productivity indicator, it is the result of the number of learners’ productions and the score obtained in the course unit validation test; it is represented in the Fig. 6.

The inputs variables of the productivity indicator fuzzy system are the production and the score. After the process of Fuzzification and Defuzzification we obtain the output variable: “IndicProd”.

4) Interactivity indicator fuzzy inference system: The last indicator is the interactivity indicator, it is the result of the degree of use of the chat and / or forum, the number of messages sent by chat and / or forum during a learning session; it is represented in the Fig. 7.

The inputs variables of the productivity indicator fuzzy system are the nbrMsg and the ChatForum. After the process of Fuzzification and Defuzzification we obtain the output variable: “IndicInterac”.

5) Fuzzification and defuzzification: The purpose of the fuzzification step is to transform a digital data into a linguistic variable; the inference mechanism used is based on the Mamdani model [13]. In the second step, we move on to define the linguistic variables and their membership functions. We used two types of membership functions for inputs and outputs because each variable has its own characteristics and can be modeled differently from another. The last step is the defuzzification. In the second step, we generated a set of commands in the form of linguistic variables (one command per rule). The purpose of defuzzification is to merge these controls and transform the resulting parameters into numeric data.

The aim of this paper is to calculate learners’ activity indicators in e-learning using a fuzzy logic system. By modifying the entries for this system the indicators vary accordingly.
IV. IMPLEMENTATION OF ACTIVITY INDICATORS FUZZY LOGIC MOTIVATION SYSTEM

The use of linguistic variables makes it possible to introduce great flexibility into the characterization of fuzzy, vague or imprecise descriptions, and to avoid the artificially rigid boundaries of standard descriptions (statistical or probabilistic descriptions). The fuzzy model therefore makes it possible to be closer to reality.

When two descriptions are close, or when they can be deduced from each other, this leads to modifiers such as "very", "more or less", which constitute a kind of standardization of the linguistic expression used to modulate a description.

A. Simulation Environment

MATLAB is an easy and efficient programming environment; it constitutes an interactive and user-friendly system of numerical calculation and graphic visualization. Aimed at engineers and scientists, it is a widely used tool, in universities and in the industrial world. Matlab integrates hundreds of mathematical and numerical analysis functions (matrix calculation, signal processing, image processing, graphic visualizations, etc.).

B. Fuzzy Logic Indicators Simulation

The purpose of this research is to predict the learning activity indicators of learners in e-learning. We got the idea to use a fuzzy expert system for the first time in learning activity indicators, considering that these indicators can be expressed using human reasoning language. Suppose that it can be evaluated according to a scale of appreciation, which to be easily usable by the person who evaluates the item, must be linguistic. Indeed, it is more natural to express an assessment in the form "the learner is rather producer" than to give a grade, whatever the scale used.

The fuzzy expert systems of the three activity indicators calculated based on traces collected from the LMS Moodle. By modifying the entries for these systems the indicators varies accordingly.

The membership functions of each evaluation are represented by a triangle, a Gaussian curve or a trapezoid, whose base or large base, respectively, covers part of the symbolic scale (in our case, [0.4]). The equation (1) of the Triangular membership functions of the freqCons and the nbrUnit variables:

$$\mu_A(x) = \begin{cases} x-x_1 & \text{si } x \in [x_1, x_2] \\ x_2-x_1 & \text{si } x \in [x_2, x_3] \\ x_3-x_2 & \text{si } x \in [x_3, x_4] \end{cases}$$

The Triangular membership functions of the variables of the fuzzy activity indicators are represented in Fig. 8 and 9.

The equation (2) of the Gaussian membership functions of the dureSess and the IndicSess variables:

$$\mu_A(x) = \begin{cases} e^{-(x-c)^2/2\sigma_g^2} & \text{si } x < c \\ e^{-2(x-c)^2/\sigma_d^2} & \text{si } x \geq c \end{cases}$$

The Gaussian membership function of the variable dureSess and the output variable of the first activity indicator are represented in Fig. 10 and 11.
Fig. 10. The Membership Function of the Third Variable DureSess.

Fig. 11. The Membership Function of the Output Variable IndicSess.

The inference rules of the Session Indicator fuzzy model Fig. 12.

The 3D surface simulations of the Session indicator according to the input variables are presented in the Fig. 13, 14 and 15.

Concerning the membership functions of the inputs variable of the productivity indicator, we used the Triangular membership function for the variable score and the Gaussian one for the variable production and the output variable IndicProd.

The 3D surface simulations of the productivity indicator in function of the inputs variables are presented in Fig. 16.

Fig. 12. The Inference Rules of the Session Indicator.

Fig. 13. 3D Surface Simulation of Session Indicator in Function of the Variables DureSess and FreqCons.

Fig. 14. 3D Surface Simulation of Session Indicator in Function of Variables NbrUnit and FreqCons.

Fig. 15. 3D Surface Simulation of Session Indicator in Function of NbrUnit and DureSess Variables.

The inputs and outputs variables for the Interactivity indicator are presented with a Triangular and Gaussian membership function. The Fig. 17 presents the 3D simulation of the Interactivity indicator.
learning activity, learners leave traces stored and recorded in dedicated files of the LMS. Until now, this data is stored at low level information with a large volume requiring processing. The task of collection, treatment and analyzing this data to extract meaningful information is very heavy. In addition, virtual mode does not offer the same working conditions as face-to-face mode.

To remedy this big problem, we decided to set up a tracking system for learning activity, by collecting information from the LMS Moodle which are called traces, after we passed to the filtering of this information to keep only the most relevant which better describe the course of the learner's activity. In the next step we chose to elaborate the activity indicators prediction system using fuzzy logic. These activity indicators give a clear idea to tutors of how the course is learned based on meaningful information about learners, it could give tutors a high level processing of the learning activity.

After having almost wiped out all the studies done on the level of learner motivation, we have found that it plays a key role in preserving learners' persistence, which also guarantees the continuity of training.

Once we got to this stage, we created a motivation prediction system using fuzzy logic, since motivation is a rough concept that cannot be measured. This decision-making system will help the tutor to have a clear view of the e-learning process so that they can act accordingly.

ACKNOWLEDGMENT

We would like to warmly thank the members of our research team and the journal scientific committee that will evaluate our paper.

REFERENCES

[1] A. F. Agudo-Peregrina, S. Iglesias-Pradas, M. Á. Conde-González, and Á. Hernández-García, “Can we predict success from log data in VLEs? Classification of interactions for learning analytics and their relation with performance in VLE-supported F2F and online learning”. Computers in human behavior, 2014, vol. 31, pp. 542–550.

[2] M. Romero, “Analyser les apprentissages à partir des traces : Des opportunités aux enjeux éthiques”. Distances et médiations des savoirs, 2019, vol. 26.

[3] M. Á. Conde, Á. Hernández-García, F. J. García-Peñalvo, and M. L. Sén-Echaluce, “Exploring Student Interactions: Learning Analytics Tools for Student Tracking”. In: Zaphiris P., Ioannou A. (eds) Learning and Collaboration Technologies. LCT 2015. Lecture Notes in Computer Science, 2015, 9192, 50–61, Springer, Cham.

[4] D. Peraya, “Les Learning Analytics en question”. Distances et médiations des savoirs, 2019, vol. 25.

[5] S. Chehbi, R. Elouahbi, and F. El Khoukhi, “Collection and transformation of Moodle traces”. In 2016 4th IEEE International Colloquium on Information Science and Technology (CISiT), pp. 570–574.

[6] J. Laflaquière, J., L. S. Setoutti, Y. Prié, and A. Mille, “Traces et inscriptions de connaissance”, 18es Journées Francophones d'Ingénierie des Connaissances, 2007, Grenoble, France.

[7] B. Florian, C. Glahn, H. Drachsler, M. Specht, and R. Fabregat Gesa, “Activity-Based Learner-Models for Learner Monitoring and Recommendations in Moodle”. In: Kloos C.D., Gillet D., Crespo García R.M., Wild F., Wolpers M. (eds) Towards Ubiquitous Learning. EC-TEL 2011. Lecture Notes in Computer Science, 6964. Springer, Berlin, Heidelberg.
[8] S. Iglesias-Pradas, C. Ruiz-de-Azcárate, Á. F. Agudo-Peregrina, “Assessing the suitability of student interactions from Moodle data logs as predictors of cross-curricular competencies”. Computers in Human Behavior, 2015, vol. 47, pp. 81–89.

[9] T. Djoud, and A. Mille, “Observing and understanding an on-line learning activity: A model-based approach for activity indicator engineering”. Technology, Knowledge and Learning, 2018, vol. 23(1), pp. 41–64.

[10] E. A. Gamie, M. S. A. El-Seoud, M. A. Salama, and W. Hussein, “Pedagogical and Elearning Logs Analyses to Enhance Students’ Performance”. In Proceedings of the 7th International Conference on Software and Information Engineering, 2018, pp. 116–120.

[11] Á. Hernández-García, E. Acquila-Natale, S. Iglesias-Pradas, and J. Chaparro-Peláez, “Design of an Extraction, Transform and Load Process for Calculation of Teamwork Indicators in Moodle”. In LASI-SPAIN, 2018, pp. 62–73.

[12] R. Mazza, M. Bettoni, M. Faré, and L. Mazzola, “Moclog—monitoring online courses with log data”, 2012.

[13] L. A. Zadeh, “Soft computing and fuzzy logic”. In Fuzzy Sets, Fuzzy Logic, and Fuzzy Systems. Selected Papers by Lotfi a Zadeh, 1996, pp. 796–804.

[14] F. Dernoncourt, and E. Métais, “La Logique Floue: le raisonnement humain au cœur du système décisionnel?”. Memory NFE211 engineering decision systems Paris, February 2011.