Tracking system based on traces analysis and SVM classification

Amali Said[#1], EL Faddouli Nour-eddine[*2], Mourchid Mohammed[#3]

[#1] *OMEGA Research Team, LERES Laboratory, Faculty FSJES, Moulay Ismaïl University in Meknes, Morocco*
[*2] *RIME Research Team, LRIE Laboratory, Mohammadia School of Engineers, Mohammed V University in Rabat Morocco*
[#3] *MIC Research Team, Laboratoire MISC, Faculty of Sciences, Ibn Tofail University in Kenitra, Morocco*

Abstract — *For several years, the Moroccan universities have put themselves in the digital technology with an increasingly sustained development of the e-learning. In this context, we have adopted a blended learning approach to support the face-to-face learning in the school of law and economics at Meknes. The success of such learning system is based on a good learners tracking by a tutor who needs accurate and aggregated information's to identify learners who need support to progress in their learning path. We propose in this paper a tracking system based on traces analysis and classification with SVM. This system is composed of several web services.*

Keywords — *Blended-learning; e-learning; trace analysis, learner's tracking; classification; method of binary SVM, web service.*

I. Introduction

Following the strategy and recommendations of the Moroccan Ministry of Higher Education that promotes Moroccan universities to implement e-Learning platforms to allow students to access educational resources available and the development of pilot projects of training in teaching with the possibility of teaching a part at a distance as indicated in the specifications of the national educational standards (CNPN), the introduction of e-learning in higher education aims to solve the overcrowding problem in universities and especially in open access institutions. A new form of pedagogy through blended learning [6] may be a solution to the problem of overstaffed students.

The mixed mode of learning (blended learning) is the joint use of e-Learning and traditional learning style often called "classroom based learning", in the same training scheme. A successful mix can make the most of each modality and maximize learning time. The aim is to ensure a level of training adapted to the specific needs of the learner. Thus, teachers can offer students learning and remote monitoring as well as a part in a classroom based training. Indeed, the immense capacity for dissemination over the internet must be accompanied by a teaching strategy in the sense that the teacher is no longer only the owner of knowledge; it is also a mediator of knowledge acquisition.

To follow up learners, the teacher needs more information on the learning path of each learner to make appropriate decisions according to the current situation of the learner (normal progress, difficulty, risk of abandonment ...).

Digital traces represent the information on the activities of the learners in the use of distance learning platform. They are generated by the platform due to interactions: learner - learner, learner - learning content and learner - tutor. The analysis of these traces is of obvious importance. It allows a proper follow-up of learners by helping to make decisions such as the customization of the learning path of learners, updating of educational policy, etc. .Most current LMS provide a large number of quantitative indicators on learning activities (number of connections, resources consulted, assessments on ratings, ...) based on collected evidence. However, the tutor must make an effort to classify learners into categories based on the progress and challenges in order to make the same decisions if needed for learners of the same class. To remedy this problem we propose to develop a dashboard containing relevant indicators to help tutor in their follow up of students via a decision -making system.

In the second section of this paper, we present the analysis of the needs of the development of a training project in blended learning in the School of Law and Economics, Meknes.

The third section will be devoted to the importance of trace analysis in the context of our project and the various indicators used in our study. The architecture of this system will be the fourth section.

The classification method of Support Vetcor Machines (SVM) is described in the fourth section. This method is used in our proposed system. This system is presented in the fifth section.

The taxonomy of indicators covered by our system is presented in the sixth section.

II. Motivations and Needs Analysis

For several years, considerable masses of graduates enroll in universities. They are 205 900 young people who have passed their baccalaureate in 2014, 6 % more than the previous year. The Ministry of Higher Education expects an increase of 5.84% per year [1]. The majority of new graduates will be hosted in the open-access faculties. The Schools of Law and Economics (FJESS) receive as usual the largest part. They are followed by those of letters and humanities. The Ministry is very concerned with this issue since the number of science graduates continues to decline in favor of literature graduates, which means more "mass education" in SLE (school of law and economics).

Figure 1 shows the evolution of the number of students in the SLE, Meknes in the last five years.

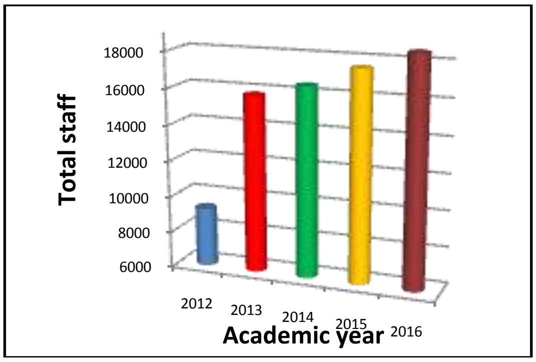


Fig 1: The evolution of the number of students in the SLE, Meknes

The number of enrolled students was 9260 in the academic 2011-2012. It rose to 18,521 in the academic year 2015-2016. It is an increase of 100%. We notice a rapid growth in the total number. Many problems related with the rise in number are caused such as low rates in supervision in some branches. Indeed, Table 1 gives the distribution of the number of students and the supervision rates in different departments:

TABLE 1
THE DISTRIBUTION OF THE NUMBER OF STUDENTS

Academic year 2015-2016			
Departements	***The number of students***	***Number of professors***	***Rates in supervision***
Economics and Management	8073	39	207 students /professor
Private Law	9 489	27	351 students /professor
Public Law	959	30	32 students /professor

Students in the Economics and Management Program and the Law in Arabic language Program are divided into 5 groups (with a rate of 1219 students / group) with a capacity of 500 students an amphitheater, which causes a significant massification (2.34 per student seat) that exceeds the national average (1.59 students per seat) [1].

With amphitheatres totally overworked and the difficulty of a classroom based learning, a new form of blended learning pedagogy (blended learning) could be a solution to the problem of overstaffed students.

A project of a joint training (classroom based learning-distance), as any e-learning project, must be conducted according to a project management methodology (See Figure 1): highly specific for identifying others [5] [9]:

- Feasibility difficulties (technical tools, availability of educational resources, ...)
- The modules selected for a Blended Learning training
- Parties of classroom based learning and parties at distance of each module,
- The overall goals and objectives of the parties provided for distance learning,
- Prerequisites,
- The number of beneficiaries of the training
- To conduct a joint pilot training project in the SLE of Meknes, we are limited to three modules. The first one is the basic training ; the second is a master; and the last is Professional BA (Professional license).
- We then identified the general objectives and specific objectives of the three modules and the prerequisites of each.
- In this study, we aimed students enrolled in different modules of the first semester for the academic year 2015-2016:
- Module "Mathematical Analysis" to register students in the basic branch "Economic Sciences and Management - EG1 " involved 500 students who are registered .
- Module "Databases" addressed 25 students in Professional BA (license professionnelle) (Computerized Management Organisations GIO) .
- Module "Information System" destined to 30 students in Research Master (Economics and International Management)

Our study therefore concerned a total population of 555 students, all levels of education included.

The developed learning system takes into account the functions reserved for students and teachers.

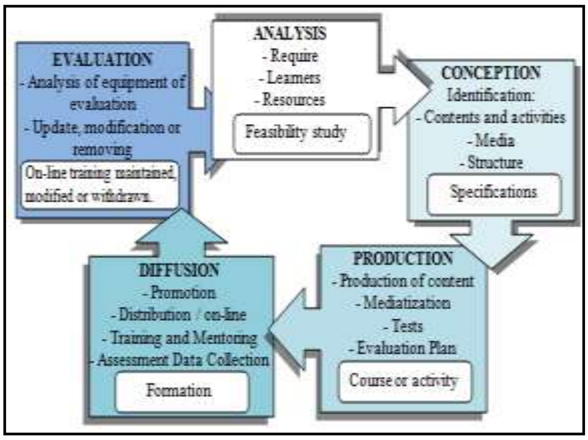


Fig 2 : process of implementation of an on-line training

III. TRACE ANALYSIS FOR MONITORING LEARNERS

Any learning platform is manipulated by different actors: tutors, learners, teachers, designers and administrators. Each user has a specific role in the learning process. The teacher defines the learning platform through its educational scenario, and uses a set of tools and resources offered by the platform. Subsequently, the tutor (he can be naturally a teacher) assists the students during their learning activity, to answer questions, raise interactions, and evaluate progress, etc. The learner uses the pedagogical scenario that indicates the route to follow to achieve the objectives of the training. He seeks, for example, to find the solution of an exercise, or understanding the content of a paragraph … etc. Tutors and learners are the principal actors during the learning activity progress.

The various interactions with the learning platform (Moodle, ...) generate data traces on the user and his activities. They can be detected and possibly recorded, usually in platform logs files or in database. Examples include: the number of pages visited in a course, the results of benchmark tests , details on the choices and followed training courses, and time spent per module , etc. We also distinguish the traces of communication activities between users of the system: the number and content of messages exchanged between the tutors, administrators and learners during participation in discussion forums.

The objective of our trace analysis is to propose to the tutor a learner's classification tool according to the status of each one and the difficulties in the learning process. This tool will assist the tutor in learner's tracking. The classification will be done using the SVM classification method (Support Vetcor Machines) by variables such as:

- Results Evaluation,
- The number of consulted resources,
- The number of carried out activities,
- The length of the connections,
- The number of questions in the forum
- The number of responses in the forum,
- The frequency connections,
- ...

We will begin our study by proposing two classes:

- Learners in normal progress .
- Learners in difficulty with a risk of abandonment.

To establish this classification, we are going to use the method of binary SVM that will be described in following section.

IV. BINARY SVM

The SVM were introduced by Vladmir Vapnik at the early 90 [5]. They were used in their beginnings in the problems of classification and regression. Today, they are used in many fields of research and engineering such as medical diagnostics, biology, etc.

We will present in the following the binary case of the original form of SVM.

The binary SVM consist of a classification of two classes +1 and -1. The idea is to find an optimal hyperplane (right in the case of two dimensions) separating the above classes in the best way. The margin is the distance between the separation border and the nearest samples, called support vector. When data is linearly separable, we are talking about a machine vectors with hard margin; and when they are not or if they contain noisy data, we are talking about a machine of soft margin support vector.

In binary SVM hard margin, to the hyperplane separator is represented by equation (1):

$$H(x) = w^T x + b \qquad (1)$$

Where $x = (x_1, x_2, ..., x_n)$ is an input vector composed of the classification variables cited in section 3, and $w = (w_1, w_2, ..., w_n)$ is a weight vector. Depending on the value of H(x), we can determine the class as follows:

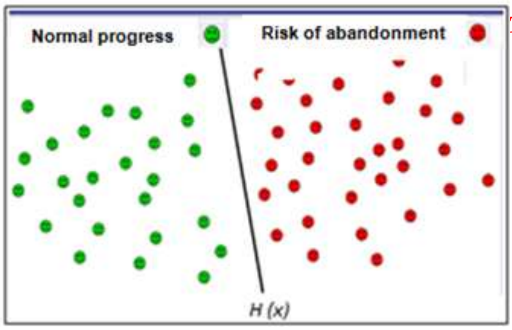
$$\text{Classe}(x) = \begin{cases} +1 \text{ si } H(x) > 0 \\ -1 \text{ si } H(x) < 0 \end{cases} \qquad (2)$$

Since both classes are linearly separable, no example satisfies the equation H(x) = 0. By asking β the minimum distance between the separator hyperplane and the nearest example, the above equation becomes:

Fig 3 : Regions separation by a hyperplan

In SVM, the separation boundary is chosen so as to maximize the margin (Figure 3). Solving this problem is to find the optimal separation border, from a set of training data used to train the SVM.

SVM are trained using a subset of the experimental data indicating the values of different variables (trace) and classification (+1 or -1)



conducted by an expert (tutor) . The remaining subset is used for the test to evaluate the performance of the classifier.

This will be part of a trace analysis system that allows calculation of the indicators and make the classification.

V. THE PROPOSED SYSTEM

The system we propose will consist of several modules as Web services (See Figure 4) each of which has a definite role. Thus, our system can be integrated in any LMS based on web services.

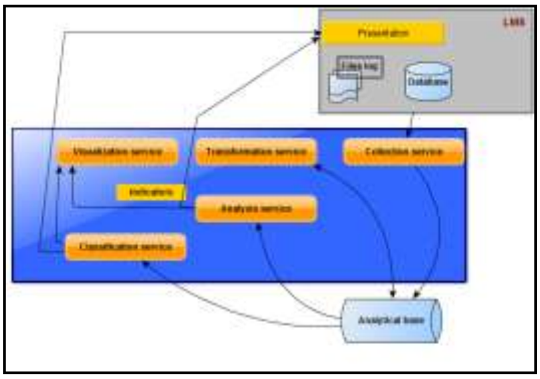


Fig 4: Services of the trace analysis system

- **Data Collection service:**

This service is responsible for extracting data from the learning system or from other data sources (surveys …) that may be available. The traces collected may affect learning or communication activities. The data flow between the LMS and the collection service is enabled by a data pattern indicating the data to be extracted.

The elementary information collected by the collecting service is determined by the predefined analysis indicators that are predefined by the objectives of the analysis.

- **Transformation Service:**

It performs the processing of data collected in a specific format used for storage and analysis

- **Analysis Service:**

The analysis of data collected and processed is made according to some indicators that can serve the participants of the system in their tasks and goals (tutor, learner, and teaching engineers ...). The operation of the actual analysis aims to measure predefined indicators and draw interpretations with the ability to propose recommendations for decision making and deliver the data to the service of these visualizations results

- **Visualization service:**

This service is used to present the analysis results in several forms depending on the nature of the indicators.

- **Classification service:**

It is used to implement the classification method by SVM. The result service will be communicated later to visualization service.

The analysis and classification results are also presented in a specific form to the presentation service of the LMS.

The calculated indicators can be of various types according to a taxonomy that we present in the following section.

VI. INDICATORS AND THEIR TAXONOMY

Learning indicators are data resulting from the transformation traces for interpret and analyze the interactions of learners in a distance learning environment [12].

Indicators according to [10] are variables that describe:

- The mode, process and quality of the cognitive system of a learning activity (cognitive indicators related to the learning process).
- Specifications and product quality of the interaction (cognitive indicators related to products resulting from the activity of learning).
- The method, process or quality collaboration (social indicator falling within cooperation and collaboration between learners).

Any indicator endowed the following attributes: name, definition, description, objectives, a domain of the possible values, a duration of validity, interpretation of the indicator according to context ...etc

The interpretation of the indicator depends on factors related to the learning context such as the environment and the learning activity (individual activity or collaborative activity)

Beyond the context, indicators are classified according to their type values (quantitative or qualitative), or according to the level of associate interpretation [10].

We can define three types of indicators related to the typology [11] depending on the dimension of the learning activity. The taxonomy of learning indicators includes three types of indicators:

- **The cognitive indicators:**

These indicators are related to productions and the process of learning. They reflect the level of knowledge acquisition (difficulty level, mistakes made ...) and inform on the using of learning resources and computer tools available on the learning platform (connection frequency ...).

These indicators are used by the tutor to help learners to organize their learning process according to their level as well as the difficulties encountered (adaptive learning).

- **Social indicators:**

The social indicator is used in a context of collaboration, and focuses on the interactions within a group of learners.

The social indicator focuses on the communication and cooperation within a group of learners performing a collective work in an online learning platform.

Learners working in groups, are brought at collaborate for a collective production through sending asynchronous messages in a discussion forum or synchronous messages by a chat tool.

The state of relations in a group of learners, involves a degree of homogeneity and cohesion in decision making during the execute of the learning scenario activities

The tutor measures the indicator of collaboration through the composition of several indicators. Among the social indicators we cite the indicator of the number of messages sent [14], the level of interaction of the group members[15] in a collaborative production.

- **Affective indicators:**

Affective indicator reflects the emotional state and the learner motivation. These indicators relate the personal way the learner in the interaction with learning content and with other learners.

We cite for example, motivational indicator that can be used to anticipate the learner's abandonment situation.

VII. CONCLUSION

In this paper, we presented the architecture and functionality of our traces analysis system that integrates several web services for collecting, transformation, analysis and visualization. We have also integrated a learner's classification service into two classes (normal progress, risk of abandonment) using the SVM method.

We envisage as a perspective, the definition of key indicators for the analysis and the crude traces we must collect.

We also intend to define recommendations according to the interpretation of the indicators in the analysis step. We will also generalize our classification method for several learners' classes.

REFERENCES

[1] L'ECONOMISTE, UNIVERSITÉ: FINI LES PROBLÈMES DE SOUS-CAPACITÉ?, Édition N° 4340 du 2014/08/18

[2] Youssef Ait ouguengay , Nour-eddine El faddouli, "An Interaction Analysis System for e-Learning : Toward a web SOA tool for the improvement of learning" worldcomp Int'l Conference. e-Learning 2011 Las Vegas USA pages 30-35

[3] Mohammed CHEKOUR, Mohammed AL ACHHAB, Mohamed LAAFOU et Badr Eddine EL MOHAJIR, Contribution à l'intégration de l'apprentissage mixte dans le système éducatif marocain, Revue internationale des technologies en pédagogie universitaire 2014

[4] Marie Prat, E-learning, réussir un projet : pédagogie, méthodes et outils de conception, déploiement, évaluation. Edition Eni 2015

[5] J.A.K. Suykens , J. Vandewalle, Least Squares Support Vector Machine Classifiers, Neural Processing Letters, volume 9, 1999

[6] Antoine Cornuéjols, Une nouvelle méthode d'apprentissage :Les SVM. Séparateurs à vaste marge, BULLETIN DE L'AFIA, numéro 51, 2002

[7] Nicolas Henriet, Des ateliers TICE de l'INP Toulouse à l'édition de Ressources Pédagogiques Numériques ouvertes : un cercle vertueux, PédagoTICE 2013

[8] Youssef Ait ouguengay , Nour-eddine El faddouli, "An Interaction Analysis System for e-Learning : Toward a web SOA tool for the improvement of learning" worldcomp Int'l Conference. e-Learning 2011 Las Vegas USA pages 30-35

[9] Lucie audit, "Guide de formation et de soutien aux enseignants et formateurs en formation à distance" Réseau d'enseignement francophone à distance du Canada (http://www.refad.ca) 2013.

[10] Dimitracopoulou A. Martinez Mones A, Dimitriadis Y., Morch A, Ludvigsen S., Harrer A.,Hoppe U., Barros B., Verdejo F., Hulshof C., de Jong T., Fessakis G., Petrou A., Lund K., Baker M., Jermann P., Dillenbourg P., Kollias V., Vosniadou S. "State of the Art on Interaction Analysis: « Interaction Analysis Indicators » ,2004,délivrable 26.1 du projet ICALT du réseau européen Kaleidoscope.

[11] Dimitracopoulou A. Kollias V., Harrer A., Martinez A., Petrou A., Dimitriadis Y. Antonio J., Bollen L., Wichmann A. State of the Art on Interaction Analysis,2005, Délivrable 31.1 du projet IA Interaction Analysis' supporting Teachers & Students' Self-regulation.

[12] B.Nabila, Analyse des traces de navigation des apprenants dans un environnement de formation dans une perspective de détection automatique des styles d'apprentissage, rapport de thèse, 2011.

[13] REFFAY, C., & CHANIER, T. How social network analysis can help to measure cohesion in collaborative distance-learning. Paper presented at Proceeding of Computer Supported Collaborative Learning conference (CSCL'2003) Computer Supported Collaborative Learning CSCL2003, Bergen Norvège, 2003.

[14] Dimitracopoulou A., Bruillard E. Enrichir les interfaces de forums par la visualisation d'analyses automatiques des interactions. Revue STICEF, Vol. 13, 2006.

[15] Lavoie, N., Levesque, J. Y., & Marin, J. Les interactions lors d'activités d'écriture collaborative au premier cycle du primaire: la contribution de l'étayage de l'enseignant. La Recherche en Education, 2011, no 6.