New Automatic Hybrid Approach for Tracking Learner Comprehension Progress in the LMS

https://doi.org/10.3991/ijim.v16i19.33733

Khalid Benabbes[1](✉), Brahim Hmedna[2], Khalid Housni[1], Ahmed Zellou[3], Ali El Mezouary[4]
[1] Faculty of Sciences, Ibn Tofail University, Kénitra, Morocco
[2] Faculty of Sciences, Ibn Zohr University, Agadir, Morocco
[3] ENSIAS, Mohammed V University, Rabat, Morocco
[4] EST, Ibn Zohr University, Agadir, Morocco
khalid.benabbes@uit.ac.ma

Abstract—Learning style is a significant learner-difference factor. Each learner has a preferred learning style and a different way of processing and understanding the novelty. In this paper, a new approach that automatically identifies learners learning styles based on their interaction with the Learning Management System (LMS) is introduced. To implement this approach, the traces of 920 enrolled learners in three agronomy courses were exploited using an unsupervised clustering method to group learners according to their degree of engagement. The decision tree classification algorithm relies on the decision rules construction, which is widely adopted to identify the accurate learning style. As the lack of good decision rules would lead to learning style misclassification, the Felder-Silverman Learning Style Model (FSLSM) is used as it is among the most adopted models in the technology of quality improvement process. The results of this research highlight that most learners prefer the global learning style.

Keywords—LMS, learning styles, FSLSM, dropout, decision tree, rules-based

1 Introduction

Learning style is the preferred mode of processing, thinking, and understanding information [1]. Each learner has a particular learning style that allows him to reach the optimal potential. A learning style is one of the various components representing the learner's model. The learner model identifies the learner's preferences and features to effectively assimilate the knowledge [2]. It is based on various aspects such as the order in which the training is presented and structured, the tutorials (assignments, notes, exams, training, exercises), and how the learning information is presented to the learners [3].

The course completion rate is defined as the number of certified learners divided by the total number of learners enrolled in the course [4]. It is low in LMSs, varies between 2 and 11% and typically less than 14%. Some researchers explained the issue of learner dropout by lack of learner engagement and motivation [3]. For other researchers, learners feel isolated from each other when they enrolled in a course [5]. For some scientists, the courses are not adjusted to the learners' orientations and preferences [6]. In that sense, the problem of detecting how learners, study and acquire new skills and knowledge has been raised [3], [4].

Since learners are diverse across their comprehension, some of them understand the complex concept step by step in a linear style with valid links between all aspects (sequential learners), while others ignore the relationships between the concept's aspects and aim to conclusions in a global style (global learners). Many solutions have been proposed based on the awareness of the learner's learning style to better help and satisfy his/her needs [1], [3], and the data collected can help to better understand learners' learning. In addition, the use of a reliable learning system that is appropriate to the learners' preferences is considered instead of providing the same materials to all learners in the same manner [3].

For this purpose, two approaches are used to identify the learning styles: the collaborative and the automatic approach [5]. The first one, a survey-based, is less precise as learners may randomly reply to queries tied to the learning style model with a lot of questions, which greatly contributes to decision tiredness [3]. The second approach is centered on gathering learners' traces generated through their interaction with the LMS platform and interpreting them to track learners' learning styles. This approach particularly has the advantage of manipulating real-time data to classify learners, thus it can track the learners' learning styles with high accuracy [5].

The proposed approach aims to identify automatically the learners' learning styles in relation to their understanding of training information through the learning system formed on the global/sequential dimension of FSLSM [5]. In that regard, we are proposing to use K-means clustering algorithm [4] implemented on learners' behaviors to group them according to their engagement level. Then the learners are classified using a decision tree classifier adjusted by a well-defined set of decision rules [6].

The manuscript is articulated as follows: First, we will introduce the basics on concepts, classification of learning styles and models, and we will provide detailed information about the approach adopted to identify learning styles. Then, we will present the methodology, including data preprocessing, feature extraction, normalization, clustering technique, and decision tree rules performed. The paper will end with the results, a conclusion and future works.

2 Background

In this section, we will first provide a general overview of the learning styles and discuss the most used models for their identification and classification. Then, we will describe the different approaches to identify learners' learning styles. These are mostly used by personalized adaptive solutions to provide an effective learning experience in line with the requirements and needs of the learner [39].

2.1 Learning styles concepts and classification

Learning style refers to the relatively stable indicator aspect of how learners perceive, interact and respond in a learning environment [7]. As all learners do not perceive a learning situation in the same way, each learner has a personal style for processing and organizing information [7]. The learning style refers also to the behavioral traits adopted by individuals in a specific context that make them different from others [8]. The term "learning style" was first used by Herb Thelen in 1954 [9]. Since then, it has gained enormous popularity among researchers, who have proposed several definitions accordingly [11], [26]. Thus, in the literature, there is no consensus on a single definition of the learning style [10].

In the same way, the evolution of recent technology makes it possible to design learning solutions adapted to learner preferences but still limited in their efficiency. Such systems require progressive improvements to make learners learn effectively during their classes following a preferred style and method [39]. As a consequence, more research has been conducted to address this issue and improve learning efficiency using learning style aspects [18]. There are numerous approaches to the learning styles of learners, that is why a variety of learning styles have been proposed.

Moreover, some learning style models developed in the past have been classified as shown in Table 1. Curry [25], for example, has established a classification of the main learning style models that she has named the "onion model", consisting of several layers. The outer layer, also called the instructional preference, refers to the conditions in which the individual wishes to learn, such as the learning environment and the teacher. The middle layer refers to the informational processing style, within which many strategies are used by the learner to process information. Yet, the inner layer refers to the cognitive personality style, it is defined as an underlying method of thinking that the learner uses. The internal layer is supposed to be the most stable in the model [11].

Table 1. Classification of learning styles models

	Classifications		
Features	**Curry [25]**	**Riding and Rayner [2]**	**Coffield et al. [12]**
Cognitive personality	×		
Cognitive skills		×	
Cognitive structure			×
Personality type			×
Learning processes	×	×	
Learning preferences			×
Learning orientations and strategies		×	×
Instructional performance	×	×	
Constitutionally based			×

However, the classification of Rayner & Riding [2], derived from the model proposed by Riding and Rayner (1997) is based on four levels. The model groups learners learning preferences into learning processes using experiential learning, orientation to study, instructional preferences, and development of learning strategies, which are related to the teaching and learning styles applied in the classroom [11]. The classification by Coffield et al. [12] identified the top 13 out of 71 different learning models, including their theoretical basis and similarities. The authors also constructed an independent study to investigate the relationship between the learning style and the student's learning capability [12].

2.2 Learning styles model

Several models of learning style have been reported in the literature and they differ in terms of some topological variables [13]. Learning styles have been of interest to a variety of researchers, especially in pedagogy. Further studies have shown that learning styles can improve the learners' activity and knowledge level, making the learners more focused and getting the best results [38], [39]. Felder [40] has proved that learning difficulties are introduced when the learner is badly treated with the learning style. Moreover, there are various interpretations and models related to this topic. In this section, we will give an overview of the main models used (Figure 1).

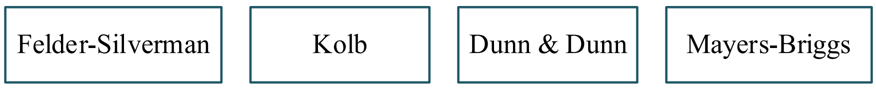

Fig. 1. The learning styles models

The Felder-Silverman Model (FSLSM) was proposed by Professor Richard M. Felder and psychologist Linda K. Silverman [13]. It is an approach within which learners perceive and acquire new concepts. This model is one of the most used models for identifying learning styles. It was designed to consider different learning styles. In that regard, four dimensions have been used and each dimension is based on two learning styles such as perception (sensory/intuitive), reception (visual/verbal), understanding (global/sequential) and processing information (active/reflective).

The Dunn Model (DM), developed by Rita Dunn and Kenneth Dunn [26] is one of the oldest models for identifying learning styles. The model considers learning style as the way a learner concentrates on, processes, and remembers new and hard skills. It proposes a series of variables to include in the learning process as factors causing differences between learners. These variables are divided into five categories: affective, sociological, environmental, psychological and physiological [14].

The Kolb Model (KM) presented by Dewey and Lewin [27] is based on the studies of the Swiss psychologist Piaget. Kolb is a four-step model containing concrete experience, reflective observation, abstract conceptualization, and active experimentation. The model starts by concrete experience (CE), a step that refers to how a learner actively experiences an activity. It is followed by a reflective observation (RO) that refers to which learner knowingly reflects on that experience. Then, an abstract conceptualization (AC), where a learner tries to use logic and ideas instead of sensing to understand problems and solve them. Lastly, active experimentation (AE) is when a learner attempts to organize and plan for testing a model or plan for future works [36].

The Mayers-Briggs Type Indicator (MBTI) is built on Jung's personality theory [28]. Bear in mind that the learning style of individuals reflects their psychological type, and indicates aspects of their personality and usual preferences. The MBTI is made up of four pairs of opposing ways to believing and reacting, with bipolar dimensions, known as dichotomy for each one of them. The four dimensions of learning styles are established in relation to (1) focus on the life (Extraverted/iNtroverted), (2) perception (Sensory/iNtuitive), (3) approach to judging by contract (Reflexive/Sentimental), and (4) focus on outside influences (Judgment/Perception) [36].

In this particular study, the FSLSM model was used because it is the most appropriate for the scientific questions raised. Each dimension is presented on a scale ranging from -11 to +11, which is more accurate to define the dominance of the learning style. All the dimensions are independent and distinctive, and constitute an index that can be used to better define learning styles [36].

2.3 Approaches to identifying learning styles

The process of identifying learning styles requires the use of one or more approaches. The most widely known ones are either collaborative or automatic approaches.

The collaborative approach is founded on the gathering of data from the survey, and provides a static learner profile model [14], [48]. The learner model is set up only once and remains unchanged throughout the learning process, which goes against the findings of researchers who believed that learning styles are rather unstable. The questions are often not clear to everyone and lead to wrong answers, which do not effectively reflect the nature of the student [41]. In addition, the long questions can demotivate the learners to complete all the questions asked. To overcome this weakness and auto-detect the learners' learning styles, the researchers suggested to automate the approach.

An automatic approach can be used to auto-detect the learning styles during the learner's interactivity with the learning platform through the data generated. However, it generates a lot of information, which may not be reliable and does not contain all the learner's desired preferences, and thus requires great carefulness to capture the global learner's profile [42]. Besides, the majority of studies opt for the collaborative approach (i.e., explicit), in contrast, only 25% of the studies use the automatic approach (i.e., implicit). The automatic approach can be either data-driven or rule-based [16].

Rule-based approach. The rule-based approach uses learners' traces to define their learning styles [17]. The model that determines learning styles is defined by using a combination of pre-defined rules to the learners' reactions within the learning system [18]. These pre-defined rules expressed as "if... then... else" statements, are applied to compute the preference level of the learning style. As shown in Figure 2, the nbr_sequential variable measures the number of times a learner accessed and used a learning resource, while mean_sequential variable represents the average number of times a learner accessed a learning resource. The two variables K_lower and K_superior vary for each learner, and are used to classify the sequential resources as low, medium or high.

```
If nbr_sequential < mean_sequential * K_lower:
    SEQUENTIAL_STYLE = Weak
Elif no_sequential < mean_sequential * K_superior:
    SEQUENTIAL_STYLE = Moderate
Else:
    SEQUENTIAL_STYLE = High
```

Fig. 2. Rules identifying learning style preferences (sequential dimension of FSLSM)

The advantage of this model is that it can be applied to the data coming from any training course, in spite of the difficulties encountered when trying to make an exhaustive list of rules to identify the learning styles.

Data-driven approach. This approach creates a model that looks like a learning style survey. It sets up the model from the learners' activities using classification algorithms such as (1) Decision Tree, (2) Bayesian method (Neural Network), (3) Reinforcement Method, (4) Markov Model and (5) Naïve Bayes. These classification methods use real data to classify learners and adapt to changes the learners' learning styles and update them according to their background [22].

In the following section, the Decision Tree Rules and the collaborative approaches were adopted to identify the learning styles, as one complements the other.

3 Methodology

This section outlines our methodological approach made up of five steps (Figure 3). It starts by extracting and preprocessing raw data in order to make each learner a composite of characteristics. Then the unsupervised clustering algorithm is applied to a group of learners according to their engagement level [31], [37]. Finally, the modified decision tree classification method, with well-defined rules, is introduced to identify the final categories of the understanding dimension (global vs. sequential) and the learning style preferences degree.

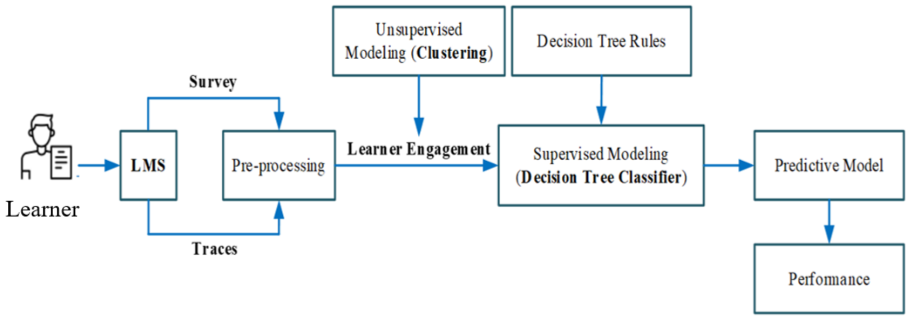

Fig. 3. The proposed approach for learning styles identification

3.1 Experimentation

The dataset used in this study was made available by the Hassan II Institute of Agronomy and Veterinary Medicine (IAV HASSAN II, Rabat, Morocco). The data collected was from three agronomy courses offered during winter sessions (2018 and 2019) through Moodle Learning Management System (LMS). The training courses lasted for 16 weeks and included video tutorials, quizzes, discussion forums and lesson docs. In this learning experiment study, 920 learners joined and produced some 136,534 events.

3.2 Data pre-processing

The pre-processing started by extracting the traces produced during 2018 and 2019 academic years from January 27 to May 30. The raw data collected may contain missing values, lack certain features, or sometimes meaningless. This type of data noise affects the performance of models, and it is, therefore, necessary to pre-process the data prior to inclusion in models. For this reason, the algorithm called Isolation Forest (Iforest) [43] with the same concept as Random Forest was used to isolate the anomalies from of the normal data. The benefits of this approach are: (1) it does not depend on density or distance measures to identify anomalies, (2) it properly maps a binary tree, (3) there are very few conditions required to dissociate outliers from normal cases; (4) it is linear in terms of time complexity and memory consumption, (5) it can be scaled to deal with high-volume datasets as well as higher dimensional problems. However, the results revealed that most of these anomalies are caused by a bug in the Moodle LMS that leads to recurrent sending of the same query, altering the learner's traces.

To organize the flow of events generated by each learner throughout 16 weeks learning, the events were reduced to a daily session, attached to a list of events filled in through a survey at the beginning of each day. The events feature allowed us to assign events to every week during which the course was delivered. The reconstruction process of each session is based on the time of the visit to the page limited to 1 hour, calculated from the difference between two timestamps. In the following subsection, we will discuss the extraction of the relevant features for the learning styles identification, the normalization step, and the dimensionality reduction.

Feature extraction. Learning style feature extraction is a method of searching the optimum model for the existing features to carry out a given task [8]. This method uses the feature engineering mechanism, known as the knowledge area, which serves to add features iteratively until it reaches the minimum number of features to create a features engine for the automatic learning algorithms operation [17]. To this end, the Moodle LMS traces were extracted for each learning style. Then each learner is presented as features vectors retrieved from the dataset, which we adapted to the learning style features of the understanding dimension of the FSLSM.

This operation is done in several steps. The actions of the clickstream are grouped by event_type attribute, which defines the types of events executed during the learning session. This aggregation allows us to calculate the total number of events performed. Using the time column, a new column called "time_duration" is generated, to measure the time of each learner's activities, such as the time spent at a learning resource. To make a session defined by learner's login and logout, several sessionization methods have been proposed [44]. In this particular study, we used a time-oriented method related to the session lifetime calculated using the difference between two timestamps that should not exceed 30 min otherwise, a new session is considered.

In Figure 4, we have synthesized the suggested features related to the FSLSM understanding dimension [19]. The simple fact of tracking the progress of pages by a learner may indicate the learner's global or continuous sequential progression [13]. Sequential learners are expected to frequently use the next and previous buttons to navigate through the course units [10]. Similarly, sequential learners use forums to read discussions, ask and answer questions in-depth [20]. However, global learners read the messages holistically, respond briefly to a given post. They also tend to go directly to a specific part of the page (seq_goto) and a portion of the video (seek_video) [21].

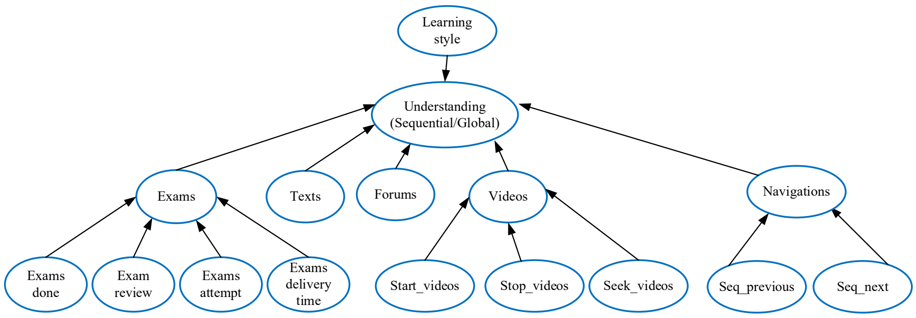

Fig. 4. Modeling learner's interaction with the Moodle LMS

Once the optimal features for identifying the learning styles are extracted, the data is normalized as discussed in the following section.

Normalization. To reduce the influence of the features processed by the learning algorithm on the output, the normalization is used to scale the features between zero (0) and one (1) [10]. In this sense, the MinMaxScaler function available in Sklearn [19] is used to execute the following equation:

$$x' = \frac{x - X_{min}}{X_{max} - X_{min}} \quad (1)$$

Where x' is the normalized value, x is the eventual value of X, X_{max} and X_{min} are respectively the maximum and minimum values of the feature X in the dataset.

Dimensionality reduction. The Principal Component Analysis (PCA) is performed to reduce the dimensionality of the baseline feature space to a 2D-space for each understanding dimension learning style without altering the links between the variables [17]. Therefore, each learning style is displayed in a bidirectional format. The output of PCA is processed to optimally cluster the learners using K-means.

In the following section, the next logical step of our methodology is introduced. It is divided into two parts specifically (1) the clustering of learners by engagement level and (2) the labeling of the dataset and its classification using the modified decision tree with rules-based approach (decision rules), as the most appropriate algorithm to predict the learning styles.

3.3 Unsupervised and supervised modeling

In this section, the unsupervised modeling process across the clustering [21] and the supervised modeling using decision rules of classification [16] are outlined to identify learning styles with high accuracy over the solutions proposed in the literature.

Clustering. In this part, the interest is to group learners by their level of engagement [33]. Through the analysis of the related features, the behaviors, the traces and the frequency by which each learner carries out a set of activities when they interact with the LMS (Figure 4). Although several profiles and taxonomies of learners have been recommended [32] - [34], we opted for the clustering classification proposed by [35] where Active learners are fully engaged and use all available resources on the LMS. Whereas Passive learners just browse the content without participating actively in the different discussions and activities. Drop-ins are active learners who follow only a part of the course. Observers can explore the content of the LMS, but don't have a specific learning goal, and see the LMS as a source of educational resources and not as a course or training in its own right. No-shows sign up for the course but never sign in, they constitute the most vital category. For this particular group, the clustering technique have been used to automatically define the different groups of learners that have similar features in order to keep the most active learners. These learners generate a huge number of traces that are enough to train the learner model with high accuracy. Each learner is shown as a stream of pre-processed features.

To overcome the problem of estimating the exact number of clusters (K), considered as one of the main challenges of K-means, we opted for the elbow method as the best existing solution to estimate the ideal number of clusters [21]. Then, the changes of the clustering cost in relation to the number of clusters are plotted. If the value drops rapidly to reach a certain point and then decreases its rate of change (Figure 5), this point corresponds to the right number of clusters [18]. The cost function is expressed by the following equation:

$$\text{Cost} = \text{Log}\left(\frac{1}{n}\sum_{i=1}^{K}\sum_{x\in \text{Cluster}_i}|x - \text{Centoid}_i|^2\right) \quad (2)$$

Therefore, the number of clusters chosen to be five (Figure 5). The learners are then divided into clusters according to their features and their engagement levels. The study conducted proved that K-means++ is the most suitable algorithm for selecting the initial centroid of clusters. This is an enhanced version that overcomes and corrects the K-means errors [20].

Based on these results, the traces of no-shows and observers have been excluded to bring the dataset down to 750 learners and 9,508 traces.

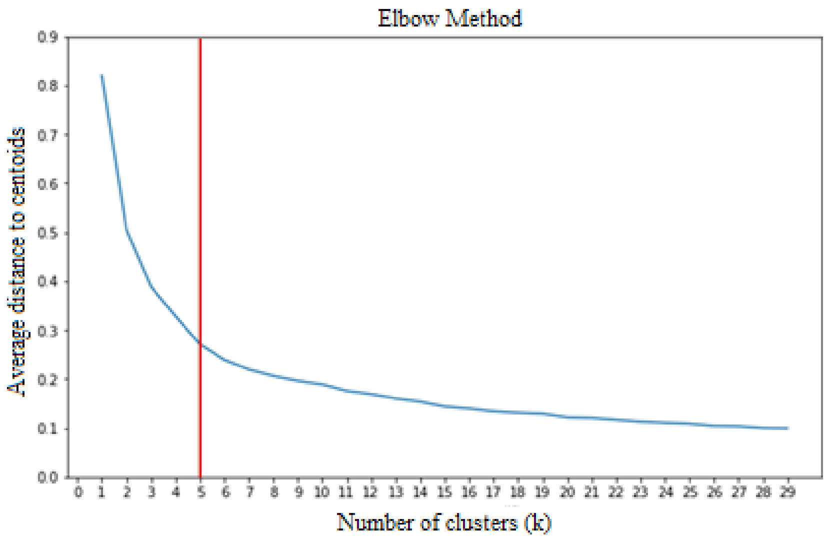

Fig. 5. Understanding the learning style (elbow point occurs at number of clusters 5)

Decision tree rules. The traces selected from learners' interactions are used as input to the Decision Tree Classifier setup, with relevant rules [15] that improve the accuracy of learners' style classification according to the FSLSM understanding dimension. The rules used to label the dataset are presented as categories in Table 2. The final values are calculated and used to set the final categories of the understanding dimension (global vs. sequential). The features gained are given as input to the decision tree classifier. This process aims to enhance the precision and accuracy of the classification of each learner's learning styles based on FSLSM.

Table 2. Decision tree rules for features of understanding dimension [15]

Course resources	Features	Rules	Weight
Forum	No participation	No participation	0

	Only view discussions	Number of views	1
	View & add a short post	Number of post and view	2
	View & add long post	Number of post and view	3
	Add a short response for discussion	Number of post	2
	Add a long response for discussion	Number of post	3
Exam	Exam delivery Time	Consume > 70% of time	4
		Consume 60% - 70% of time	3
		Consume 50% - 60% of time	2
		Consume 40% - 50% of time	1
		Consume < 40% of time	0
	Exam done	Check > 70% exams	4
		Check 60% - 70% exams	3
		Check 50% - 60% exams	2
		Check 40% - 50% exams	1
		Check < 40% exams	0
	Exam attempt	The try > 70% exams	4
		The try 50% - 70% exams	3
		The try 30% - 50% exams	2
		The try 10% - 30% exams	1
		The try < 10 exams	0
	Exam review	Do > 20% review	4
		Do 15% - 20% review	3
		Do 10% - 15% review	2
		Do 5% - 10% review	1
		Do < 5% review	0
Text (word)	View Course	Number of times viewed	1
	No view	No overview	0
Videos	Start_video	Number of times forward	2
	Stop_video	Number of times turned back.	2
	Seek_video	Number of times seeking a video	1

A set of rules proposed by Sheeba and Krishnan [15] (Table 2) was used to derive the categories of the understanding dimension according to the final calculated value: (1) if the value is between 0 and 1, the understanding dimension is Strong Global, (2) if the value is between 2 and 3, the understanding dimension is Medium Global, (3) if the value is between 4 and 5, the understanding dimension is Balanced, (4) if the value is between 6 and 7, the understanding dimension is Medium Sequential, (5) if the value is between 7 and 8, it means that the understanding dimension is Strong Sequential.

The decision tree classification method is used to train the learning style prediction model. It is made up of a root, decision nodes and branches [16]. The nodes are labeled with text attributes and the branches with attribute values. The classification involves two steps, the training and the testing (Figure 6). The training step builds the tree recursively from a dataset until it integrates all records into the same class. The testing step is processed to assess the accuracy of the trained tree [23].

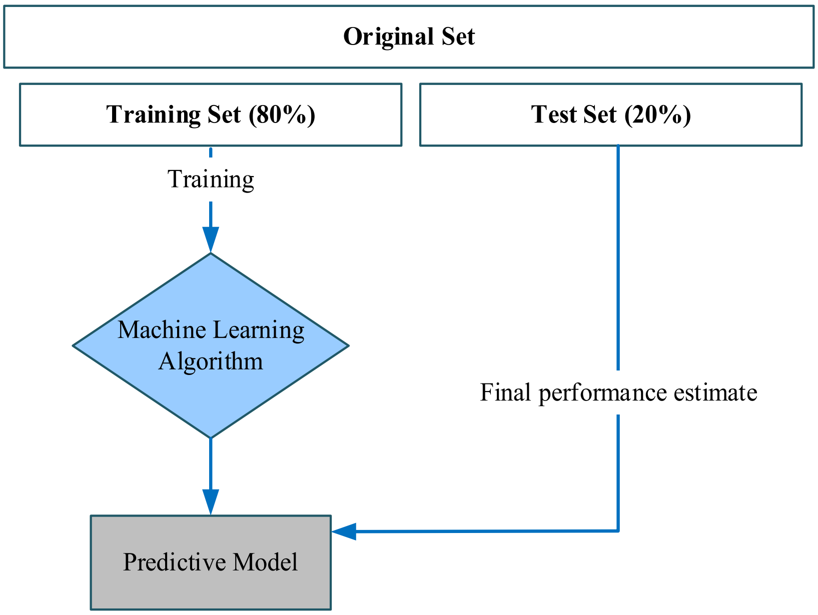

Fig. 6. Overview of the training and test set

Finally, the learners' learning style is predicted using their traces (Figure 4), through the decision tree classifier model strengthened by the rules shown in Table 2 as the rule-based decision tree classifier model.

4 Results & discussions

The features (9,508 traces) related to the understanding dimension of the FSLSM [22] were gathered from 750 learners enrolled in the three agronomy courses available online on Moodle LMS [18]. A data pre-processing [17] was conducted to clean and extract features from the learners' traces. The selected data were normalized in the interval [0, 1], and the K-means++ clustering algorithm [19] was applied to automatically define the different learner groups with the same engagement level [34], and thus selecting the most active ones. Then, the higher decision rules are applied to label the dataset, classify the learners' learning styles and quantify the degree of dominance of each learning style. According to the distribution of learning styles related to the understanding dimension shown in Figure 7, the largest number of learners shows a moderate global learning style with 23 learners certified and 239 learners failed. These learners achieved worse results than the sequential learners. As shown in table 2, this is because most learners use videos and PowerPoint and go directly to a specific unit (seq_goto) or to a specific part of the video (seek_video), thus spending less time analyzing the textual documents viewed. They also used little the forums, as they rarely or never contributed to the discussions. Therefore, we found more variance in global dimension. This is in line with the findings of [23], which showed that global learners have a high risk of dropping out of the training.

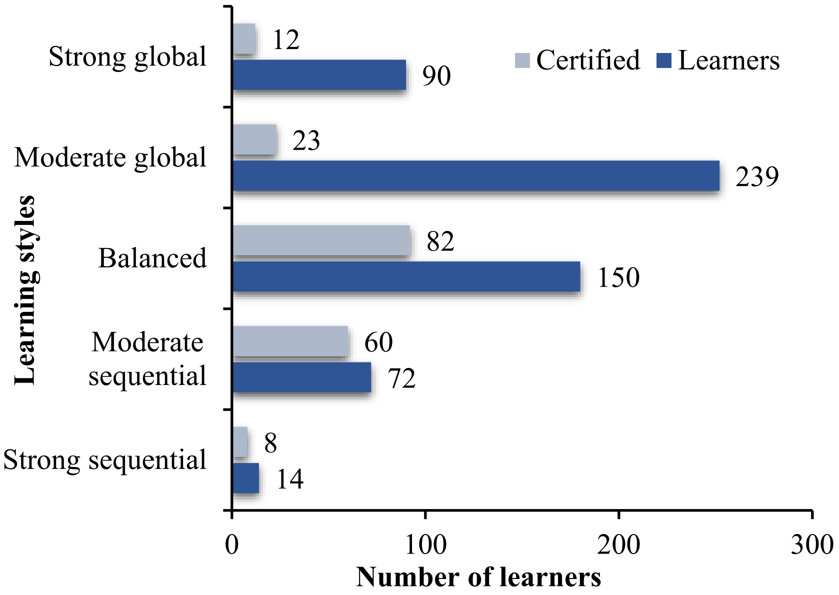

Fig. 7. Distribution of learning styles

The main evaluation method is the confusion matrix (Figure 8), which confronts the predicted and the observed classes as a table of frequencies. These measures are used to illustrate the accuracy of the classification model produced, which confirms the performance of our evaluated system.

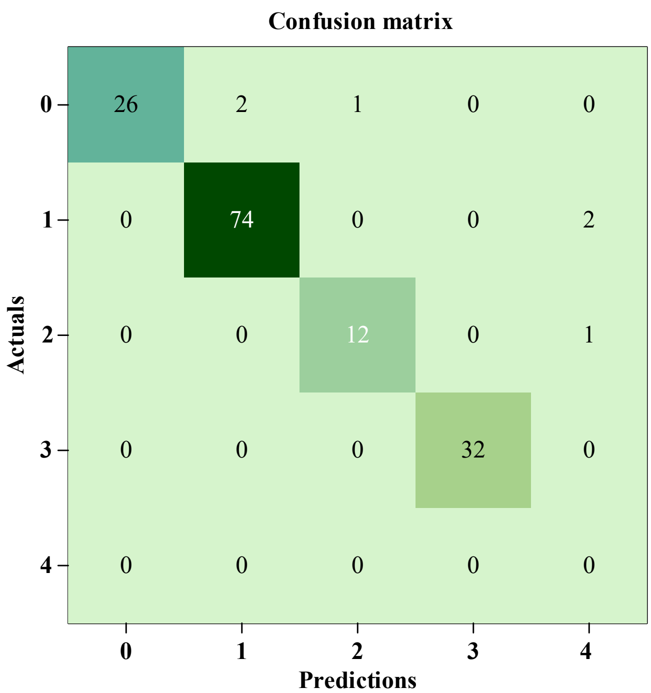

Fig. 8. Confusion matrix (Actuals vs. Predictions)

The performance achieved (Figure 9) from the experiments carried out is also analyzed using the following accuracy classification equation:

$$\text{Accuracy} = \frac{\sum_{i=1}^{5} \text{TP}_i}{\sum_{i=1}^{5} \text{TP}_i + \text{FN}_i} \tag{3}$$

Where accuracy is the ratio of correct predictions made by the model, TP (true positives) is the number of true positives or the number of samples, and FN (false negatives) is the number of false negatives or the total number of misclassified samples [29].

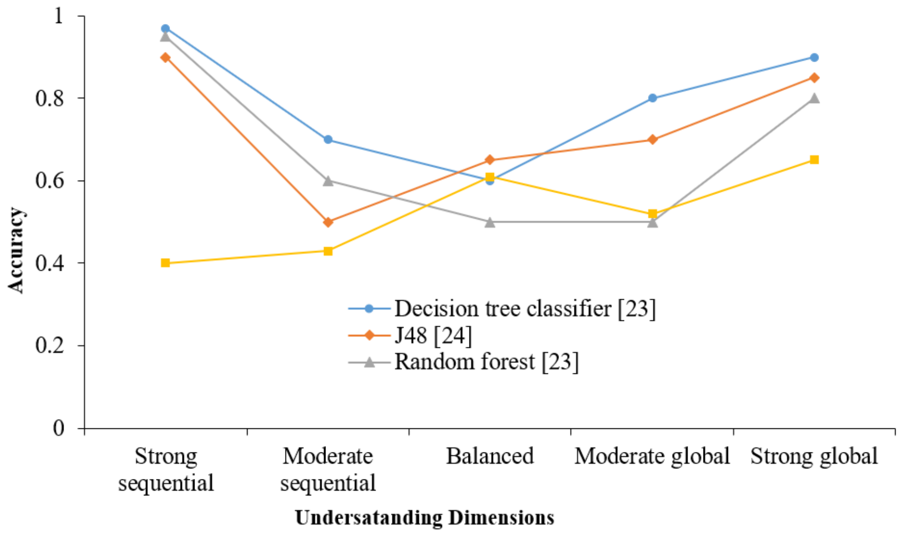

Fig. 9. Comparison of precision for understanding dimensions of FSLSM

It is evident that the decision tree rules method has an accuracy of up to 95%, which is a high value, followed by J48 and Random Forest. The fuzzy logic model is the algorithm with the lowest performance that proves that the proposed solution can classify and predict learners' learning styles with high accuracy, under the conditions that learners are induced to effectively interact with the groups in the Moodle LMS and that the learning styles are well aligned with their preferences and engagement following the FSLSM understanding dimension.

In line with our findings, it seems that the Felder Silverman model appears in the vast majority of pre-existing research [46]. Our study is also consistent with the literature [16], [23], [45]. The researchers' focus on Felder Silverman compared to other models is due to its effectiveness and the easiness of the survey adapted to personalized learning environments [40]. Our study is in concordance with the research literature in higher education compared to low percentage targeted on secondary and primary education levels [38]. However, the research has focused on learners only, neglecting the role of the instructor who should be making suggestions, adapting presentations, and evaluating the impact of learning systems [47].

The proposed model can help instructors better identify the needs and preferences of their learners and tailor instruction in a format that matches their learning style, rather than providing similar content to everyone. According to similar studies, the use of learning style understanding dimension in adaptive learning systems, has a positive impact on academic achievement, level of learner satisfaction and makes it easier to manipulate personalized courses [44].

Finally, since the proposed model can be adapted to the changing of the learners' learning style, it is crucial to implement the approach in an LMS or a MOOC. This will provide instructions and activities suitable to the learners' learning style, by linking both the data gathered through the survey and the traces generated while interacting with the learning environment.

5 Conclusion and future work

In this paper, we aimed at identifying the sequential/global learning styles of the learners through their interaction with the Moodle LMS platform, using collaborative and automatic methods. The experiments were conducted on a dataset extracted from the Moodle LMS platform available at the Hassan II Institute of Agronomy and Veterinary Medicine, collected from three agronomy courses offered during 2018 and 2019 winter sessions). We carried out two experiments: first, an unsupervised clustering technique to group learners according to their engagement on the platform, by understanding their behaviors through the analysis of their activities. In the second experiment, we evaluated four machine learning algorithms: rule-based decision tree, random forest, fuzzy logical model and J48. The results achieved indicate that the decision tree rule approach is effective in detecting and predicting the learning style with high accuracy in the understanding dimension. The decision tree approach, made of precise decision rules, was used to label the engaged groups in the datasets and accurately classify the learning style based on the understanding dimension of the FSLM. The results of our research showed that the majority of learners subscribed to the Moodle LMS have a global learning style preference mode. The proposed solution can be used to tailor courses provided to the learner's learning style understanding type and help teachers frame educational resources according to learners' learning styles.

The future research is to build a secure adaptive cross-platform system, by including the proposed classifier into MOOC for automatic learning style identification.

6 Acknowledgment

I am grateful to the Hassan II Institute of Agronomy and Veterinary Medicine for its extensive support. I also wish to express my thanks to Mustapha NAIMI, Professor at the Hassan II Institute of Agronomy and Veterinary Medicine and Mr. Abderrahman El HADDI, CTO and Founder of Enduradata Corporation, for their valuable comments and suggestions.

7 References

[1] M. Barry and A. Egan, "An adult learner's learning style should inform but not limit educational choices," Int Rev Educ, vol. 64, no. 1, pp. 31-42, Feb. 2018. https://doi.org/10.1007/s11159-017-9694-6

[2] J. Bernard, T.-W. Chang, E. Popescu, and S. Graf, "Learning style Identifier: Improving the precision of learning style identification through computational intelligence algorithms," Expert Systems with Applications, vol. 75, pp. 94-108, Jun. 2017. https://doi.org/10.1016/j.eswa.2017.01.021

[3] B. Hmedna, A. El Mezouary, and O. Baz, "A predictive model for the identification of learning styles in MOOC environments," Cluster Comput, vol. 23, no. 2, pp. 1303-1328, Jun. 2020. https://doi.org/10.1007/s10586-019-02992-4
[4] L. Akharraz, A. El Mezouary, and Z. Mahani, "To context-aware learner modeling based on ontology," in 2018 IEEE Global Engineering Education Conference (EDUCON), Tenerife, Apr. 2018, pp. 1326-1334. https://doi.org/10.1109/EDUCON.2018.8363383
[5] S. Aghaei, M. J. Azizi, and P. Vayanos, "Learning Optimal and Fair Decision Trees for Non-Discriminative Decision-Making," AAAI, vol. 33, pp. 1418-1426, Jul. 2019. https://doi.org/10.1609/aaai.v33i01.33011418
[6] K. Adil, E. Amrani, and O. Lahcen, "Modeling and Implementing Ontology for Managing Learners' Profiles," ijacsa, vol. 8, no. 8, 2017. https://doi.org/10.14569/IJACSA.2017.080819
[7] R. N. Aljabali and N. Ahmad, "A Review on Adopting Personalized Gamified Experience in the Learning Context," in 2018 IEEE Conference on e-Learning, e-Management and e-Services (IC3e), Langkawi Island, Malaysia, Nov. 2018, pp. 61-66. https://doi.org/10.1109/IC3e.2018.8632635
[8] M. Barry and A. Egan, "An adult learner's learning style should inform but not limit educational choices," Int Rev Educ, vol. 64, no. 1, Art. no. 1, Feb. 2018. https://doi.org/10.1007/s11159-017-9694-6
[9] V. Thongchotchat, K. Sato, and H. Suto, "Recommender System Utilizing Learning Style: Systematic Literature Review," in 2021 6th International Conference on Business and Industrial Research (ICBIR), Bangkok, Thailand, May 2021, pp. 184-187. https://doi.org/10.1109/ICBIR52339.2021.9465832
[10] A. Agarwal, D. S. Mishra, and S. V. Kolekar, "Knowledge-based recommendation system using semantic web rules based on Learning styles for MOOCs," Cogent Engineering, vol. 9, no. 1, p. 2022568, Dec. 2022. https://doi.org/10.1080/23311916.2021.2022568
[11] J. Bicans and J. Grundspenkis, "Student Learning Style Extraction from On-Campus Learning Context Data," Procedia Computer Science, vol. 104, pp. 272-278, 2017. https://doi.org/10.1016/j.procs.2017.01.135
[12] A. Hidayat and V. G. Utomo, "An Architecture of Adaptive Online Module System Based on Felder-Silverman Learning Style Model," presented at the International Conference on Online and Blended Learning 2019 (ICOBL 2019), Yogyakarta, Indonesia, 2020. https://doi.org/10.2991/assehr.k.200521.015
[13] S. Majid and S. Chaudhary, "An exploratory etudy of preferred learning styles of visually impaired children," p. 16.
[14] B. Hmedna, A. E. Mezouary, and O. Baz, "How Does Learners' Prefer to Process Information in MOOCs? A Data-driven Study," Procedia Computer Science, vol. 148, pp. 371-379, 2019. https://doi.org/10.1016/j.procs.2019.01.045
[15] T. Sheeba and R. Krishnan, "Prediction of student learning style using modified decision tree algorithm in e-learning system," in Proceedings of the 2018 International Conference on Data Science and Information Technology - DSIT '18, Singapore, Singapore, 2018, pp. 85-90. https://doi.org/10.1145/3239283.3239319
[16] N. B. H. Ahmad and S. M. Shamsuddin, "A comparative analysis of mining techniques for automatic detection of student's learning style," in 2010 10th International Conference on Intelligent Systems Design and Applications, Cairo, Egypt, Nov. 2010, pp. 877-882. https://doi.org/10.1109/ISDA.2010.5687150
[17] L. Leka and A. Kika, "Implementation of Moodle plugins to support student's learning styles," vol. 7, no. 2, p. 4, 2021.

[18] M. Raleiras, A. H. Nabizadeh, and F. A. Costa, "Automatic learning styles prediction: a survey of the State-of-the-Art (2006–2021)," J. Comput. Educ., Feb. 2022. https://doi.org/10.1007/s40692-021-00215-7
[19] B. Hmedna, A. El Mezouary, and O. Baz, "A predictive model for the identification of learning styles in MOOC environments," Cluster Comput, vol. 23, no. 2, pp. 1303-1328, Jun. 2020. https://doi.org/10.1007/s10586-019-02992-4
[20] B. Fernando, T. Tommasi, and T. Tuytelaars, "Joint cross-domain classification and subspace learning for unsupervised adaptation," Pattern Recognition Letters, vol. 65, pp. 60-66, Nov. 2015. https://doi.org/10.1016/j.patrec.2015.07.009
[21] C. Li and H. Zhou, "Enhancing the Efficiency of Massive Online Learning by Integrating Intelligent Analysis into MOOCs with an Application to Education of Sustainability," Sustainability, vol. 10, no. 2, p. 468, Feb. 2018. https://doi.org/10.3390/su10020468
[22] T. Parshakova, J.-M. Andreoli, and M. Dymetman, "Global Autoregressive Models for Data-Efficient Sequence Learning," 2019. https://doi.org/10.18653/v1/K19-1084
[23] S. Sahu and B. M. Mehtre, "Network intrusion detection system using J48 Decision Tree," in 2015 International Conference on Advances in Computing, Communications and Informatics (ICACCI), Kochi, India, Aug. 2015, pp. 2023-2026. https://doi.org/10.1109/ICACCI.2015.7275914
[24] A. A. Nadiri, S. Shokri, F. T.-C. Tsai, and A. Asghari Moghaddam, "Prediction of effluent quality parameters of a wastewater treatment plant using a supervised committee fuzzy logic model," Journal of Cleaner Production, vol. 180, pp. 539-549, Apr. 2018. https://doi.org/10.1016/j.jclepro.2018.01.139
[25] Curry, L. (1983). An organization of learning styles theory and constructs.
[26] Dunn, R. S., Dunn, K. J., & Price, G. E. (1979). Learning style inventory. Price Systems Lawrence, KS
[27] Kolb, D. (1984). Experiential learning:turning experience into learning. New Jersey: Prentice Hall.
[28] Myers, I. B., McCaulley, M. H., Quenk, N. L., & Hammer, A. L. (1998). MBTI manual: A guide to the development and use of the Myers-Briggs Type Indicator (Vol. 3). Consulting Psychologists Press Palo Alto, CA.
[29] N. Idrissi and A. Zellou, "A systematic literature review of sparsity issues in recommender systems," Soc. Netw. Anal. Min., vol. 10, no. 1, p. 15, Dec. 2020. https://doi.org/10.1007/s13278-020-0626-2
[30] N. Idrissi, A. Zellou, O. Hourrane, Z. Bakkoury, and E. H. Benlahmar, "Addressing Cold Start Challenges In Recommender Systems: Towards A New Hybrid Approach," in 2019 International Conference on Smart Applications, Communications and Networking (SmartNets), Sharm El Sheik, Egypt, Dec. 2019, pp. 1–6. https://doi.org/10.1109/SmartNets48225.2019.9069801
[31] M. Khalil and M. Ebner, "Clustering patterns of engagement in Massive Open Online Courses (MOOCs): the use of learning analytics to reveal student categories," J Comput High Educ, vol. 29, no. 1, pp. 114–132, Apr. 2017. https://doi.org/10.1007/s12528-016-9126-9
[32] R. Ferguson and D. Clow, "Examining engagement: analysing learner subpopulations in massive open online courses (MOOCs)," in Proceedings of the Fifth International Conference on Learning Analytics And Knowledge, Poughkeepsie New York, Mar. 2015, pp. 51–58. https://doi.org/10.1145/2723576.2723606
[33] R. F. Kizilcec, C. Piech, and E. Schneider, "Deconstructing disengagement: analyzing learner subpopulations in massive open online courses," in Proceedings of the Third

International Conference on Learning Analytics and Knowledge - LAK '13, Leuven, Belgium, 2013, p. 170. https://doi.org/10.1145/2460296.2460330
[34] Hill, P. (2013). Combining MOOC student patterns graphic with Stanford Analysis. Http://mfeldstein.Com/combining-mooc-student-patterns-graphic-stanfordanalysis
[35] P. E. Ramírez-Correa, F. J. Rondan-Cataluña, J. Arenas-Gaitán, and J. L. Alfaro-Perez, "Moderating effect of learning styles on a learning management system's success," Telematics and Informatics, vol. 34, no. 1, pp. 272–286, Feb. 2017. https://doi.org/10.1016/j.tele.2016.04.006
[36] A. E. mezouary, B. Hmedna, "An Unsupervised Method for Discovering How Does Learners Progress toward Understanding in MOOCs," IJITEE, vol. 10, no. 5, pp. 40-49, Mar. 2021. https://doi.org/10.35940/ijitee.E8673.0310521
[37] Dahal, N., Manandhar, N., Luitel, L., Luitel, B., Pant, B., & Shrestha, I., "ICT tools for remote teaching and learning mathematics: A proposal for autonomy and engagements," Adv Mobile Learn Educ Res, vol. 2, no. 1, pp. 289–296, 2022. https://doi.org/10.25082/AMLER.2022.01.013
[38] Kikilias, P., Papachristos, D., Alafodimos, N., Kalogiannakis, M. & Papadakis, St., "An Educational Model for Asynchronous E-Learning. A Case Study in a Higher Technology Education," In D. Guralnick (ed.) Proceedings of the International Conference on E-Learning in the Workplace (ICELW-09), 10-12 June 2009, New York: Kaleidoscope Learning.
[39] I. Katsaris, N. Vidakis, "Adaptive e-learning systems through learning styles: A review of the literature," Adv Mobile Learn Educ Res, vol. 1, no. 2, pp. 124–145, 2021. https://doi.org/10.25082/AMLER.2021.02.007
[40] R. M. Felder et L. K. Silverman, « Learning and teaching styles in engineering education », Engineering education, vol. 78, no 7, p. 674-681, 1988.
[41] Bendahmane, M., Falaki, B. El, & Benattou, M. (2019). Paper-Toward a Personalized learning Path through a Services-Oriented Approach. International Journal of Emerging Technologies in Learning (IJET), 14(15), 52-66. https://doi.org/10.3991/ijet.v14i15.10951
[42] Almeida, T. O., VIsta, B., & De Magalhaes Netto, J. F. (2019). Adaptive Educational Resource Model to Promote Robotic Teaching in STEM Courses. Proceedings - Frontiers in Education Conference, FIE, 2019-Octob. https://doi.org/10.1109/FIE43999.2019.9028417
[43] Liu, F.T., Ting, K.M., Zhou, Zhi-Hua: Isolation forest. In: 2008 Eighth IEEE International Conference on Data Mining, pages 413 – 422. IEEE, (2008). https://doi.org/10.1109/ICDM.2008.17
[44] Sensuse, D. I., Hasani, L. M., & Bagustari, B. (2020). Personalization Strategies Based on Felder- Silverman Learning Styles and Its Impact on Learning: A Literature Review, 293-298. https://doi.org/10.1109/IC2IE50715.2020.9274670
[45] Truong, H. M. (2016). Computers in Human Behavior Integrating learning styles and adaptive e-learning system: Current developments , problems and opportunities. Computers in Human Behavior, 55, 1185-1193. https://doi.org/10.1016/j.chb.2015.02.014
[46] G. Akçapinar, M.-R. A. Chen, R. Majumdar, B. Flanagan, and H. Ogata, "Exploring student approaches to learning through sequence analysis of reading logs," in Proceedings of the Tenth International Conference on Learning Analytics & Knowledge, Frankfurt Germany, Mar. 2020, pp. 106–111. https://doi.org/10.1145/3375462.3375492
[47] J Aguayo, J. M., Valdes, J., Cordoba, V. H., Nájera, M., Vázquez, F. R., Muñoz, E., & García Lirios, C, "Digital activism in students of a university in central Mexico in the COVID-19 era," Adv Mobile Learn Educ Res, vol. 2, no. 1, pp. 297–307, 2022. https://doi.org/10.25082/AMLER.2022.01.014

[48] Can, Y., & Bardakci, S, "Teachers' opinions on (urgent) distance education activities during the pandemic period," Adv Mobile Learn Educ Res, vol. 2, no. 2, pp. 351–374, 2022. https://doi.org/10.25082/AMLER.2022.02.005

8 Authors

Khalid Benabbes is a PhD student at the MISC Laboratory, Faculty of Sciences, Ibn Tofail University, Kénitra, Morocco. He is currently a software engineer at Hassan II Institute of Agronomy & Veterinary Medicine in Rabat. He holds an engineering degree in Computer Sciences from ENSA, Agadir. His research interests include MOOC, Recommender systems, Machine Learning, and Data Science.

Brahim Hmedna Received his PhD in computer sciences from the IMIS Laboratory, Faculty of Sciences, University of Ibn Zohr, Agadir, Morocco. He is currently a assistant professor at the National School of Business and Management in Agadir. His research interests include MOOC, learning styles, Machine Learning and Data science (E-mail: brahim.hmedna@edu.uiz.ac.ma).

Khalid Housni received his PhD in computer sciences from the University of Ibn Zohr Agadir, Morocco. In 2014, he joined the Department of Computer Science, Faculty of Sciences, Ibn Tofail University, Kénitra, Morocco. His research interests include image/video processing and networks reliability (E-mail: housni.khalid@uit.ac.ma).

Ahmed Zellou received his PhD in Computer Sciences from SPM Research Team, ENSIAS, Mohammedia School of Engineering, Rabat, Morocco. He is currently a professor at the National School of Computer Science and Systems Analysis (ENSIAS). His research interests include Information Integration, Hybrid, Semantic and Fuzzy, Content Integration, Semantic P2P, Cloud Integration (E-mail: ahmed.zellou@um5.ac.ma).

Ali El Mezouary received his PhD in computer sciences from the IRF-SIC Laboratory, EST, University of Ibn Zohr, Agadir, Morocco. He is a full professor at ESTA, University of Ibn Zohr. His research interests include Web Semantic and Adaptive Learning Systems (E-mail: elmezouaryali@gmail.com).

Article submitted 2022-07-01. Resubmitted 2022-08-24. Final acceptance 2022-08-29. Final version published as submitted by the authors.