

Smart Model for Classification and Orientation of Learners in a MOOC

<https://doi.org/10.3991/ijet.v17i05.28153>

Ilham Dhaioir^(✉), Mostafa Ezziyyani, Mohamed Khaldi
Abdelmalek Essaâdi University, Tanger, Morocco
dhaioir.ilham@gmail.com

Abstract—Distance education (E-Learning) is experiencing significant, rapid and continuous evolution all over the world, especially with the arrival of the covid-19 pandemic. MOOCs are considered as a personal learning process, which are addressed to a massive and varied number of learners. The problem of the free opening MOOCs puts us in front of a massive number of registrants, which means a large number of heterogeneous profiles, which makes the teacher's task more complicated, either in terms of follow-up or framing. As a solution to this problem, in this present work, we propose an approach that allows the classification and categorization of learner profiles via an intelligent and autonomous system developed on the basis of neural networks and in particular the self-organizing map (SOM). This approach which is based on the traceability of learners, allowed us to get homogenous groups in order to direct them towards MOOCs that meet their characteristics and needs. The tests carried out have shown that our approach is efficient in terms of classification and grouping of profiles, which allows us to manage a large number of learners either at the level of the choice of relevant contents or during the evaluation process.

Keywords—MOOC, profile, traceability, content, teacher, learner, learning objectives, classification, neural networks, self-organizing map

1 Introduction

Currently, MOOCs (Massive Open Online Courses), which can be defined as free courses distributed online to the general public with activities to be carried out via the Internet, are faced with a major problem, namely the increase in the drop-out rate due to the diversification of learners' profiles, which does not help the teacher to accompany them in their learning or to determine the pedagogical content and activities adapted to each profile. This leads the learner to abandon the MOOC and to look for another one, which causes a considerable loss of time and money, knowing that the creation of a MOOC requires a lot of effort, work and funding.

Defining the word profile, which is the focus of our work, we find that it represents a set of personal data information that characterizes the knowledge, skills, preferences and learning objectives of a learner. A profile can relate to a single person, or to a group of people with commonalities, such as members of a work group.

According to [1], a learner profile is what a student does in a learning context according to his or her learning style, which is related to his or her personality trait and the way he or she intervenes and engages in the course.

The use of learner profiles is one of the ways in which learners can improve their learning by addressing their problems and gaps.

For the teacher, it is also a way of facilitating the monitoring of the learner's learning progress [2] [3].

As for the learners themselves, research has shown the value of presenting them with information about the state of their knowledge in order to help them develop reflective skills and increase their motivation and responsibility for their learning [4].

The teacher who supervises a distance learning course has a very important and central role in the preparation of teaching resources (courses, activities, exercises, videos, etc.) and also in facilitating interactions with the learners.

But as MOOCs are open, free and address a large number of profiles that are heterogeneous in terms of: nationality, age, learning styles, background, skills, learning objectives, the teacher is unable to manage this massive number as well as to respond to their needs and learning objectives.

To solve this problem, the use of classification and grouping of learners according to their profiles in an intelligent way based on artificial neural networks in general and in particular the use of the self-organizing map is foreseen. This paradigm is widely used in the classification and clustering tasks we are interested in this work and it has undergone many improvements [5]. The integration of the SOM map in this work for the classification and clustering of learners will help us to control and direct learners with their profiles in an intelligent and autonomous way when they follow an online training.

Our work aims at improving the quality of online education and training by integrating our approach into MOOC platforms. The latter has become very important.

The success of a distance learning course is guaranteed by offering a course that adapts to the profiles of the learners, as well as by monitoring the learners in their learning process in order to have all the information that can be useful in characterizing the progress of the learners.

And to facilitate the task of the teachers it is necessary to ensure the operation of the classification of the profiles in homogeneous groups with the same needs and objectives as well as to orient them towards courses that meet their needs, objectives and learning capacities.

The process of classifying the profiles will then help the teacher to improve his pedagogical content and to better guide the learners in their learning to follow the MOOC to its end and to benefit from its content.

Also, it is preferable to integrate from time-to-time support tools that are often necessary [6] in a learning situation where learners want to acquire a skill or knowledge and are characterized by diversity in their: needs, objectives, skills, abilities and learning styles.

So, it is difficult to treat the diversity of each learner separately. Therefore, in this article, we propose our approach to develop a Smart Model that allows the classification

and orientation of profiles in order to improve their learning and meet their needs and objectives.

Firstly, we will identify the parameters on the basis of which we will classify the profiles from the information entered by the learners when filling in the registration form in the platform. Secondly, we will classify the profiles into homogeneous groups on the basis of these parameters.

2 Related work

In order to analyze and manage the massive number of registrants in a MOOC, several avenues of research are considered to allow the appropriation of a MOOC by different users, but all these approaches have shown limitations to overcome or offer very general solutions in terms of solving the problem of profiles in a MOOC system.

Starting with the work of [7] which proposed a method to manage the profiles of learners following a distance learning course by associating each learner profile with a stereotype that represents a general assessment of the set of characteristics deduced from them, with the aim of proposing activities appropriate to their profiles.

The problem reported is that this approach [7] does not take into account all the characteristics of the learners.

Therefore, other authors have used a set of systems and software to find a solution to the problem of managing a set of heterogeneous profiles.

Citing then first the ViSMod system adopted by [8] which is represented as a Bayesian network that allows us the visualization by the learner and the teacher of a model of the learner coming from another computer system but that just allows the reuse of a limited number of profiles, which excludes the number of profiles coming from HITEs, as well as most of the paper and pencil profiles created by teachers. Also, other problems are pointed out by this approach, namely that the intervention of the teacher is rare, and his practices are not considered.

As for [9] who proposed the PERLEA project (Reused Student Profiles of External Software Explored in Depth) which aims at using and processing profiles already known and existing in the classroom by a system that allows teachers to follow learners in their evolution as well as the management of their profiles through the results obtained during assessments. This approach has shown that the use of the software in the classroom is not really insufficient because it is used by a few teachers and it only takes into consideration a part of the curriculum, with a limited number of profiles that are processed.

As for the approach of [10] which envisages enriching learner profiles from heterogeneous applications by taking into account just the heterogeneous profiles by proposing a model of the learner profile integrating IMS-LIP and taking into account some cognitive and metacognitive elements.

The main limitation of this approach lies in its lack of flexibility because in the case of a change of information, the software will require a total modification and the work to be redone.

The approach of [11] opted for the use of an e-learning system based on a priori All algorithm with the aim of meeting the needs of the different profiles that follow an online course by offering a set of recommendations based on collaborative filtering to meet the needs, and learning objectives of the learners.

As for [12] proposed a recommendation model of threads facilitating the search of information by learners in discussion forums in which he could participate easily by giving a small list to each learner of topics that could interest him.

The recommendation takes into account three types of modelling: content modelling of forum discussions based on word analysis, preference modelling which is deduced from the history of participation in the forum and finally modelling of the learner's social connection which supports his interactions with peers.

[13] have proposed a recommendation system based on a Python programming algorithm that uses the collaborative filtering method to recommend to each participant the content of courses that suit their learning preferences. Participants are asked to rate the courses by giving a score from 1 to 5 which is then saved in a CSV file. Based on this rating, a prediction function is calculated to determine the degree of appreciation of each participant.

Both approaches [12] and [13] have two major limitations. On the one hand, the recommendation is aimed, in particular, at forum participants who represent only a limited minority between 5% and 10% of the learners of a MOOC.

On the other hand, the proposed recommendation does not verify whether the pedagogical courses made available to the learners have been interpreted by them.

Their successful referral system proposes to offer learners in need of help a list of relevant learners that are made available to them. Once learners are in need of help, they select learners who can offer them help. In turn the targeted learners have the choice to send a private message to report them either favored or ignored and also there is the possibility to open a chat window with them.

3 The application of neural networks in classification tasks

Neural networks are highly recommended for creating intelligent and autonomous classification systems based on machine learning.

Artificial neural networks are a class of systems inspired by the functioning of biological nervous systems that mimic their ability to learn from observations of any new situation [14]. These systems represent a set of strongly connected elements called formal neurons capable of producing a weighted output based on the inputs and weights of synaptic connections (see Figure 1).

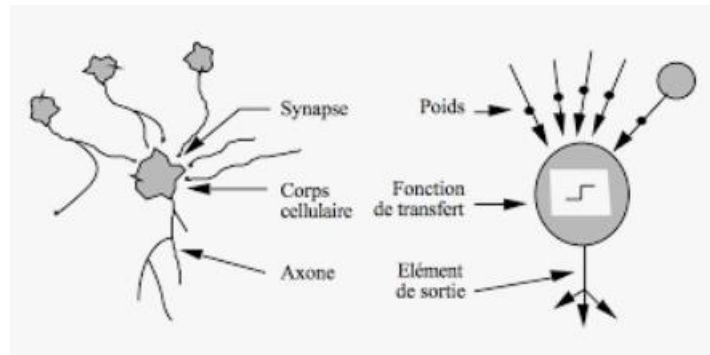


Fig. 1. The biological neuron and the formal neuron

Artificial neural networks differ according to two main criteria: the structure and the learning algorithm.

According to the first criterion there are several paradigms that can be classified according to the type of connection between neurons on neural networks: monolayer, multilayer, forward propagation, recurrent and others, the Figure 2 shows some paradigms.

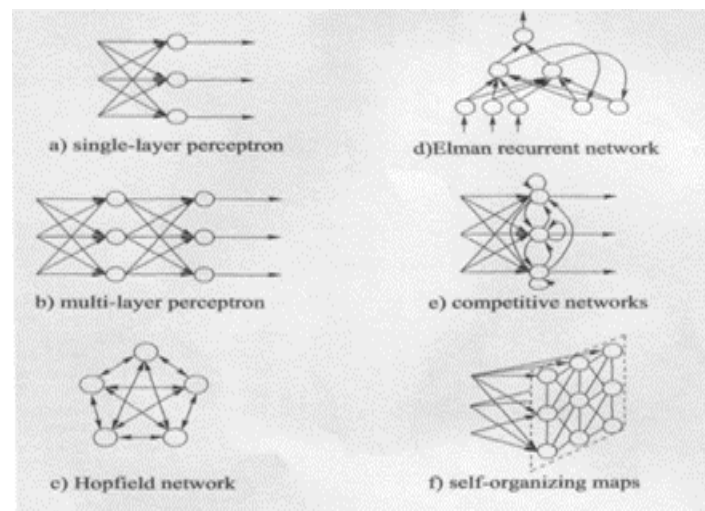


Fig. 2. Examples of RNA paradigms

Regarding the second criterion, there are two learning methods: supervised learning and unsupervised learning [15]. The supervised learning method is used to classify objects with a label (a class). The unsupervised learning method is used to solve classification and clustering tasks. One of the neural network paradigms that uses this method is the SOM map, which has been used in several classification tasks and has undergone

several improvements and evolutions to increase the relevance of the classification and the learning speed [16], [17]. (See Figure 3).

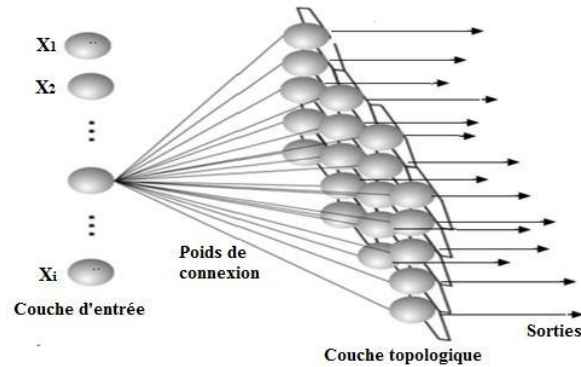


Fig. 3. Korhonen's self-organizing map

The SOM map consists of an input and an output tick Figure 3, The elements of the map are scattered in a space - usually one or two dimensional (see Figure 3).

The input data is presented as a matrix, the rows are the vectors of the objects and the columns are the components of these objects. This paradigm uses competitive learning, in which it tries to distribute the training set into groups (clusters), which are specific to the input data [16], [17]. This type of neural network processes only the input vectors X and thus implements the "unsupervised" learning procedure.

All neurons in the output layer K combined with all neurons in the input layer i connect with synaptic weight coefficients w_{io} . When the input vector x is supplied to the input of the neural network, only one output is activated ("Winner"). After efficient learning of the network, all input vectors with the same cluster will have a winner. Then, there are two methods to determine the winner and to establish appropriate learning rules: the scalar product method and the Euclidean distance method [14].

The learning procedure starts with the normalization of the input data and the synaptic weights to an even length to reduce the learning time [16], [17]. This operation is based on the following algebraic formula:

$$x_i = x_i / \sqrt{\sum_{j=0}^{n-1} x_j^2} \quad (1)$$

Where:

x_i - input object component or synaptic weight vector.

n - the number of variables in the vector x.

The main learning algorithm successively goes through a series of iterations, and at each iteration, the learning object vectors are presented to the input of the network, with no desired output. At the end of this procedure, neurons that are topologically adjacent respond to similar input vectors.

At each iteration the winning neuron for each input object is determined, using the Euclidean distance metric [14], see formula below:

$$k : \|w_k - x\| \leq \|w_o - x\| \quad \forall o \quad (2)$$

After each determination of a winner, the map performs the adjustment of its synaptic weights to progressively minimize the distance between this winning neuron and the input object. For this operation, the following formula is used [14]:

$$w_{ij}(t + 1) = w_{ij}(t) + \alpha_i(t)h(d, t) \cdot [y_i - w_{ij}(t)] \quad (3)$$

Where:

y_i - the value of output neuron i .

$w_{ij}(t)$ and $w_{ij}(t + 1)$ - the synaptic weights in iterations t and $(t + 1)$.

$\alpha_i(t)$ - learning rate, this coefficient can have a value between 0 and 1, and is calculated using the following equation:

$$\alpha_i = \alpha_0 e^{-i} \quad (4)$$

Where:

i - The iteration number.

t - The iteration rate.

$h(d, t)$ - neighborhood function, it is written according to the formula below :

$$h(d, t) = \begin{cases} 0, & d \geq \delta(t) \\ e^{-\frac{d}{2\delta(t)}}, & d < \delta(t) \end{cases} \quad (5)$$

$$\delta(t) = \delta_0 e^{-\frac{t}{\mu}} \quad (6)$$

Where:

d - the distance between the winning neuron and a neuron x .

δ - Constant.

$$\mu = \frac{n}{\log_{10}(\delta_0)} \quad (7)$$

n - the Iteration step.

The learning process will continue until the stabilization of the self-organizing map, when the input objects are distributed on groups (Clusters), and for each cluster a winning neuron that plays the role of the center is determined, as well as a distance to determine the membership of the objects to this group.

Following our study, we found that this paradigm supports our work, because it is able to classify and group profiles into homogeneous groups according to their: Ages, background, skills, preferences, needs and learning objectives.

Therefore, we proposed our approach based on the use of the self-organizing map (SOM) to create an intelligent classification and clustering system to determine the membership of learners to profiles. Our system is in the form of a software model that can be used in MOOC platforms to improve the management of learners and other resources.

4 The proposed intelligent approach to learner classification and orientation

The resolution of the limitations of the above-mentioned authors' work is guaranteed by the creation of an intelligent system of classification and grouping of similar profiles in order to direct them to courses that meet their profiles. To this end, and after a thorough study in different fields of distance learning, systems and means of artificial intelligence and in particular neural networks, we have developed an intelligent approach based on the SOM map paradigm which is able to analyze and structure all the data that characterize the profiles, leading to the creation of clusters that respond to the groups of learners' profiles. This operation will allow us to orientate the learners towards contents that are appropriate for their profiles. In Figure 4 we present the schema of our approach.

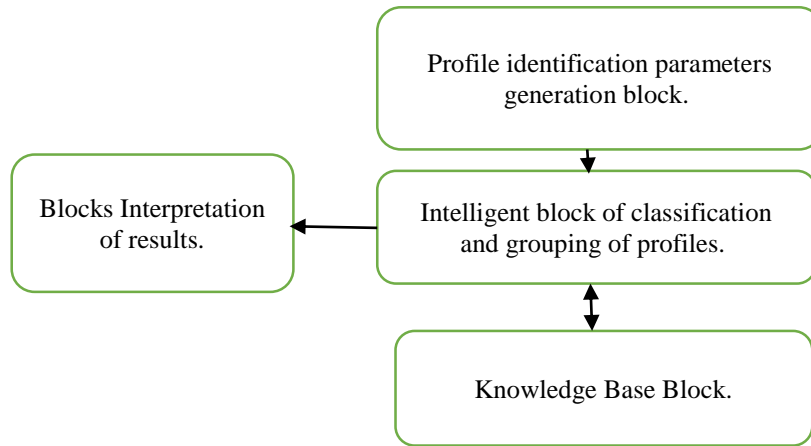


Fig. 4. Schematic of the intelligent profile grouping and orientation system

The diagram shows that the proposed system is composed of four blocks. The first block is intended for the generation of the profile parameters, which are the data collected during the enrolment of the learners, and which play the role of a learning base for the SOM card. The latter is the main core of the second block, which we have named "Intelligent block of classification and grouping of profiles". This block contains a pre-processing component that performs the normalization of training data and the elimination of correlation between data, especially the linear dependency between training objects [16], [17].

The pre-processed data is passed to the input of the SOM map in order to carry out a learning phase which ends with the definition of the classes and their clusters, this result will be stored in the form of a knowledge base, which is formed by the values of the connection weights and represented in Figure 4 by the third block named "Knowledge Base Block".

The knowledge base will be used to determine the membership profiles of new registrants, which are represented by their data generated by the system during their registration. These data play the role of stimuli in the operation phase of the system, which form adequate responses for them.

The last block named "Results Interpretation Block", it is used for the interpretation of the results of the classification and clustering, and it can be presented in the form of two-dimensional topographic maps which was used in our work where we used a two-dimensional map with a coloring system. In the map each winning neuron responds to an appropriate profile which will take a unique sticker represented on the map in the form of a circle identified by their index: row number and column number. Once you click on a neuron you get the necessary information about the appropriate profile.

5 Application of the proposed intelligent system for classification and grouping of learners

In this part of the work, we will test the application of the proposed intelligent system on a database composed of the traceability parameters of a group of learners registered in a MOOC. Our objective is to evaluate our approach in order to reveal its advantages, as well as its limitations. The system has been realized in the form of a software developed in Java. In order to collect the information of each learner, we will opt for the use of any trace left by the registrants, when filling in the registration form on the MOOC platform. With these traces, we will elaborate the set of data necessary for learning. In order to ensure the classification of learners, we will take into consideration three traceability parameters that will allow us to classify learners into homogeneous groups expressing the same learning needs and expectations.

These parameters will be integrated into the registration form that the learner must fill in before starting a distance learning course. With the help of the proposed system and on the basis of these parameters, information on all the possible profiles of the learners will be obtained, which will then play the role of the group profiles to which the newly registered learners will belong and automatically they will be directed to the groups that meet their needs, interests and characteristics. The Table 1 shows the parameters used for the classification and grouping of profiles.

Table 1. The main parameters

Parameters	Signification	Code
Language	English	6391
	French	6392
	Spanish	6393
Level	Advanced	1
	Medium	2
	Low	3
Speciality	Mathematics	25
	Computer Science	27
	Biological sciences	57

5.1 Experimental tests and results

Despite the progress made in the field of technology and information, it is now possible to collect, store and manage large masses of data in an autonomous and intelligent way, especially when the volume of information is large.

For this reason, we have chosen to develop our system based on the self-organizing map as an efficient way of classifying and grouping patterns in an intelligent way.

The use of the SOM map foresees a pre-processing of training data, in this operation the algorithms used try to eliminate the correlation in this type of data to fight against the linear dependence between the objects which prevents the SOM map to perform the classification in an efficient way and increases the training time [14]. After normalization to an even length, the data is passed to the classification block, so that the Kohonen neural network learns until it stabilizes. The values of the connection weights form the knowledge base of the classification and the grouping of the learners into homogeneous profile groups.

The result is used by the interpretation block in text mode or with the help of topographic maps.

The learning process starts with the successive presentation of objects representing the learners. These objects are in the form of vectors which consist of the parameters that have been entered by the learners. And the set of vectors constitutes a learning matrix M (see Figure 5).

$$M = \begin{pmatrix} x_{11} & x_{12} & x_{13} \\ x_{21} & x_{22} & x_{23} \\ \dots & \dots & \dots \\ x_{k1} & x_{k2} & x_{k3} \\ \dots & \dots & \dots \\ x_{m1} & x_{m2} & x_{m3} \end{pmatrix}$$

Fig. 5. Example of training data matrix

The main learning algorithm successively goes through a series of iterations, and at each iteration the learning object vectors are presented to the input of SOM, with no desired output. At the end of this procedure, neurons that are topologically adjacent respond to similar input vectors.

At each iteration the winning neuron for each input object is determined, using the Euclidean distance metric [16], see the formula below:

$$k : \|w_k - x\| \leq \|w_o - x\| \quad \forall o \tag{2}$$

After each determination of a winner, the map performs the adjustment of its synaptic weights to progressively minimize the distance between this winner neuron and the input object. For this operation the following formula is used [14]:

$$w_{ij}(t + 1) = w_{ij}(t) + \alpha_i(t)h(d, t) \cdot [y_i - w_{ij}(t)] \tag{3}$$

Where: y_i - the value of output neuron i ;

$W_{ij}(t)$ and $w_{ij}(t+1)$ - the synaptic weights in iterations t and $(t+1)$.

$\alpha_i(t)$ - learning rate, this coefficient can have a value between 0 and 1, and is calculated using the following equation:

$$\alpha_i = \alpha_0 e^{-i} \tag{4}$$

Where: i - The iteration number;

t - The iteration rate.

$h(d, t)$ - neighborhood function, it is written according to the formula below:

$$h(d, t) = \begin{cases} 0, & d \geq \delta(t) \\ e^{-\frac{d}{2\delta(t)}}, & d < \delta(t) \end{cases} \tag{5}$$

$$\delta(t) = \delta_0 e^{-\frac{t}{\mu}} \tag{6}$$

Where: d - the distance between the winning neuron and a neuron x .

δ - Constant.

$$\mu = \frac{n}{\log_{10}(\delta_0)} \tag{7}$$

n - the Iteration step.

The learning process will continue until the stabilization of the self-organizing map, when the input objects are distributed on groups (Clusters), and for each cluster a winning neuron that plays the role of the center is determined, as well as a distance to determine the membership of the objects to this group.

Following our study, we found that this paradigm supports our work, as it is able to classify and group the profiles into homogeneous groups according to their: Ages, background, skills, preferences, needs and learning objectives.

For this reason, we proposed our approach, which is based on the use of the self-organizing map (SOM) to create an intelligent classification and grouping system in order to determine which learners belong to the appropriate profiles. Our system is in the form of a software model that can be used in MOOC platforms to improve the management of learners and other resources.

In the tests, which we carried out, we used a database of 36 objects, some of which are shown in the Table 2:

Table 2. Parameters of learner objects

Objects	Languages	Levels	Specialty
1	1	1	10
2	1	1	11
3	1	1	12
4	1	1	13
13	2	1	10
14	2	1	11
15	2	1	12

16	2	1	13
25	3	1	10
29	3	2	10
30	3	2	11
31	3	2	12
32	3	2	13
33	3	3	10
34	3	3	11
35	3	3	12
36	3	3	13

The software developed relies on graphical interfaces to allow the user to enter the initial parameters and display the results. The initial data and the results of the tests carried out are shown in the figures below.

In the initial phase the size of the SOM map is specified by indicating the number of neurons whose power gives us the total number of neurons that will be drawn on the map.

After drawing the map, whose neurons are presented as rectangles, we then move on to the second phase for the input of the following parameters: iteration rate, average error and learning rate or neighborhood. Figure 6 shows an example of a 20X20 neuron map with the initial training parameters.

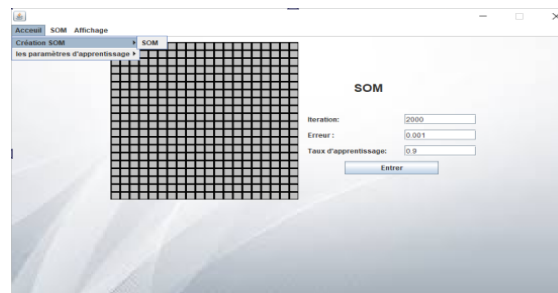


Fig. 6. Window for the input of initial parameters

As soon as the data and the learning parameters are entered, the learning phase is started and continues until the stop conditions are met. The training results can be interpreted in two ways.

The first way illustrates the result in textual form by displaying the iteration rate, the final average error and the final learning rate (neighborhood), as well as the winning neurons with their position and their connection weight values, followed by their equivalence objects. Figure 7 shows the textual results.

```
Affichage
***** Resultats Finaux *****
L'iteration maximale (Donnée entrée):      500
L'iteration finale:                        500
Le taux d'apprentissage initial (Donnée entrée) :  0.7
Le taux d'apprentissage final:             0.14045136867561533
Erreur Moyenne (Donnée entrée):           1.0E-4
Erreur Moyenne:                           9.729206565686991E-4
*****
Entree no: 1
Numero du neurone:72
Coordonnées: (1,7)
Les poids:
0.09545126035889703
0.09545127254806085
0.9908278167272079
```

Fig. 7. Textual results window

The second way will allow us to interpret the results visually in the form of a two-dimensional topographic map. The coloring system used characterizes each winning neuron with a specific color, and on each neuron we mention the number of corresponding input objects. Figure 8 shows an example of the SOM map after training.

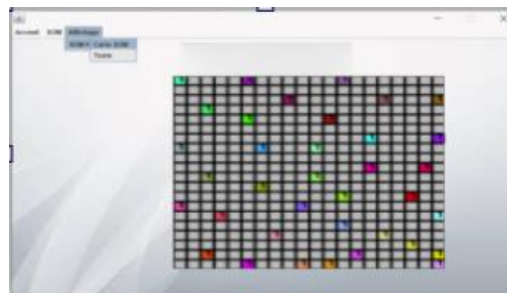


Fig. 8. SOM map viewer window

6 Test results and discussion

The objective of the tests carried out is to find out the capabilities of the proposed approach if it can learn to classify learners into well-defined profiles, in order to use them later for the grouping and orientation of new learners according to their needs and learning objectives in the appropriate courses. The study of the results will allow us to define the ambiguities of the system and to set up the next perspectives. The data used in this work was prepared based on a preliminary study carried out by a group of teachers, in which they set up the number of profiles and the parameters that constitute the object components. The multitude of learning objects consists of 36 objects, each of which is composed of three parameters (see Table 2).

In Table 3 we present the results of the learning tests, these results will be interpreted also in the form of a topographic map (Figure 9 and Figure 10), with the possibility of

displaying all the information of the selected winning neuron, and the appropriate input object.

Table 3. Learning results

	Test1	Test2
Number of objects.	36	36
Size of the card.	10X10	20X20
Total iteration.	1500	2000
Medium error.	0.0001	0.001
Initial learning rate.	0.7	0.9
Number of the winning neurons.	36	36
Final medium error.	9,30E-04	9,57E-04
Final learning rates.	0.1401	0.1507

The data in Table 3 shows that the system has well classified the objects in the training set for both tests, where the number of winning neurons is 36 (36 profiles). The average error value is less than 0.00001, which shows that the input objects in the realization space were well matched to the winning neurons. These neurons present the centers of clusters, the membership of which will be determined by the learning rate or neighborhood parameter (0.14 for the first test and 0.15 for the second test).



Fig. 9. The topographic map test1

The two topographic maps in both Figures 9 and 10, show that the winning neurons were well dispersed over the entire map surface. This shows that the proposed system has structured the training data well, which gives an efficient organization of the objects in the realization space and facilitates their classification.



Fig. 10.The topographic map test2

In analyzing the results interpreted by the two maps, it was found that dimension can play an important role in the notion of neighborhood where winning neurons may have other neurons as neighbors depending on the learning rate calculated by the system.

From a comparative point of view, other works had shown and appeared to play an important role in the clustering and classification of learner profiles based on traces left by learners in mostly forums, chat...

But the problem is that most of these works are based on the traces collected during and at the end of the MOOC to determine the characteristics of each learner profile, which sometimes generates a high number of in our approach, however, the operation of generating traces that will allow us to group and classify learners' profiles into homogeneous groups is carried out before the start of the MOOC, based on the learners' profiles.

On the other hand, in our approach the operation of generating traces that will allow us to group and classify the profiles of the learners in homogeneous groups is carried out before the beginning of the MOOC, based on the parameters of traceability of the learners that are integrated in the registration form of the learners in the platform which will automatically facilitate the teachers in the follow-up, the accompaniment and the supervision of the learners since the latter will work with groups of learners having the same characteristics and learning needs. Other research works have pointed out limitations in either taking into consideration all the characteristics of the learners in particular or in the inability to manage a massive number of learners which results in the consideration of a limited number of learners as in the case of the works of [7], [8] and [9].

In contrast to our approach, all the characteristics of the learners' profiles are taken into account and new registrants will automatically be classified and grouped into groups with the same characteristics of their profiles.

7 Conclusion and perspectives

The world is currently aware of the spread of the Corona pandemic, which has necessitated several strict precautions, and this has negatively affected all areas of life, including the field of education in all its types and levels. The horrific spread of this epidemic forced us to find quick and appropriate solutions, so distance learning was the perfect solution for this epidemic.

In this work we have discussed the problems we face when using online courses and training. In this work we have addressed the problems we face when using online courses and training. When using new information technologies as a teaching medium, we were faced with a large number of learners and MOOCs offered on the Internet, which makes the task of managing this type of education very complicated and requires adequate solutions. Our proposed approach is based on the use of the SOM card as an artificial intelligence tool to develop an autonomous system capable of learning, classifying and grouping learners to orient them according to their learning needs and objectives. In order to test our method, the learning data were selected on the basis of a study which allowed us to form 36 profile classes to which the new learners will be oriented according to their needs and objectives. The results of the tests carried out showed that our approach met our expectations, especially the fight against the problem of managing learner profiles and content. The use of artificial intelligence and in particular the SOM card has enabled our system to be able to:

- Learn from previous situations.
- Use a knowledge base for further classification and clustering.
- Visualizing multi-parametric data in a two-dimensional space.
- Keeping the topographic relationships between the data on the map for better analysis.

The work with the results we have obtained has allowed us to form an idea about the next works to realize a system able to group between the digital processing and semantic processing of data.

8 References

- [1] Waldeck, J. H. (2007). Answer the Question: Student Perceptions of Personalized Education and the Construct's Relationship to Learning Outcomes. *Communication Education*, 56(4), 409-432. <https://doi.org/10.1080/03634520701400090>
- [2] Grandbastien M. et Labat J.-M., Introduction, Environnements informatiques pour l'apprentissage humain (Traité IC2, série Cognition et traitement de l'information), Grandbastien M. et Labat J.- M., pp. 17-24, 2006b.
- [3] Hibbi Fatima-Zohra, Abdoun Otman, Haimoudi El Khatir, Smart Tutoring System: A Predictive Personalized Feedback in a Pedagogical Sequence, October 2021, *International Journal of Emerging Technologies in Learning (iJET)* 16(20). <https://doi.org/10.3991/ijet.v16i20.24783>
- [4] Bull S., Dimitrova V. et McCall G., Open Learner Models: Research Questions, Preface of Special Issue of the *International Journal of Artificial Intelligence in Education (IJAIED)*, vol. 17 n°2, pp. 83- 87, 2007a.
- [5] Ben Krose et Patrick van der Smagt, *An Introduction to Neural Networks* [archive], 8e éd., 1996.
- [6] Czarkowski, M., & Kay, J. (2003). How to give the user a sense of control over the personalization of adaptive hypertext? Paper presented at the Workshop on Adaptive Hypermedia and Adaptive Web-Based Systems, User Modeling Session.
- [7] Zapata-Rivera, J. D., Greer, J., « Interacting with Inspectable Bayesian Student Models », *IJAIED*, vol. 14; 2004, p. 1-37.

- [8] Stéphanie Jean-Daubias. Exploitation de profils d'apprenants. Environnements Informatiques pour l'Apprentissage Humain, 2003, Strasbourg, France. pp.535-538 (hal-01575481).
- [9] Stéphanie Jean-Daubias, Carole Eyssautier-Bavay, Marie Lefevre, Modèles et outils pour rendre possible la réutilisation informatique de profils d'apprenants hétérogènes, Revue STICEF, Volume 16, 2009. <https://doi.org/10.3406/stice.2009.996>
- [10] Ramandalahy, M. T., Vidal, P., Huet, N., Broisin, J. (2009). Partage et réutilisation d'un profil ouvert de l'apprenant. EIAH'2009. Le Mans, France, p. 85-92.
- [11] Yang D., Piergallini M., Howley I., Rose C. (2014). Forum thread recommendation for massive open online courses. In Educational data mining 2014, p. 257–260.
- [12] D. F., Sinclair J. (2015). Collaborative filtering recommendation system: a framework in massive open online courses. INTED2015 Proceedings, p. 1249–1257.
- [13] Labarthe H., Bouchet F., Bachelet R., Yacef K. (2016). Does a peer recommender foster students' engagement in moocs? In 9th international conference on educational data mining, p. 418–423.
- [14] Haimoudi E K, Ikram Issarti, Ali Daanoun, The Particularities of the Counter Propagation Neural Network Application in Pattern Recognition Tasks, April 2018 Conference: International Conference on Advanced Information Technology, Services and Systems. https://doi.org/10.1007/978-3-319-69137-4_42
- [15] Jodouin (J.F.) 1994, Les réseaux de neurones, principes et définitions, Hermès, Paris.
- [16] Sara, B., Otman, A., Haimoudi, E.K., New learning approach for unsupervised neural networks model with application to agriculture field, International Journal of Advanced Computer Science and Applications, 11(5), pp. 360-369. <https://doi.org/10.14569/IJACSA.2020.0110548>
- [17] Haimoudi E K, FAKHOURI H, CHERRAT L, Ezziyyani M, towards a New Approach to Improve the Classification Accuracy of the Kohonen's Self-Organizing Map During Learning Process, March 2016 International Journal of Advanced Computer Science and Applications 7(3). <https://doi.org/10.14569/IJACSA.2016.070333>

9 Authors

Ilham Dhaoui is a PhD candidate at the Department of Computer Science at Abdelmalk Essaâdi University. She had her engineering in pedagogy and multimedia. Her research interest includes learning analytics, e-learning.

Mostafa Ezziyyani is a professor of computer science at Abdelmalek Essaâdi University, Tanger, Morocco (email: mezziyyani@uae.ac.ma).

Mohamed Khaldi is a professor in the department of computer science in Abdelmalek Essaâdi University, Tetouan, Morocco (email: m.khaldi1@uae.ac.ma).

Article submitted 2021-11-08. Resubmitted 2021-12-18. Final acceptance 2021-12-20. Final version published as submitted by the authors.