Indonesian Journal of Electrical Engineering and Computer Science
Vol. 26, No. 1, April 2022, pp. 482∼492
ISSN: 2502-4752, DOI: 10.11591/ijeecs.v26.i1.pp482-492

A cluster validity for optimal configuration of Kohonen maps in e-learning recommendation

Jamal Mawane[1], Abdelwahab Naji[2], Mohamed Ramdani[1]
[1]Department of Computer Science, Faculty of Sciences and Technics of Mohammedia, Hassan II University, Mohammedia, Morocco
[2]Signals, Distributed Systems and Artificial Intelligence (SSDIA) Laboratory, ENSET, Hassan II University, Mohammedia, Morocco

Article Info

Article history:

Received Nov 11, 2021
Revised Jan 20, 2022
Accepted Feb 7, 2022

Keywords:

Cluster validity
Coefficient of variation
Collaborative filtering
Homogeneity
Kohonen maps
Recommendation systems

ABSTRACT

This paper reviews the first block of our unsupervised deep collaborative recommendation (UDCF) system and proposes a platform whose goal is to try to find the adequate parameters of the Kohonen maps, to create homogeneous clusters in profile data and results, the homogeneity is verified thanks to the very low variance rate of the results obtained by the cluster population and a second criterion which is the high prediction rate of collaborative recommendation. Although the revision concerns only the clustering block, and the use of a symmetrical autoencoder without searching for its optimization, the result obtained (82.33%) for the optimal configurations with high homogeneity of the Kohonen map is equivalent to the optimized result of the UDCF and even better than the classical recommendation methods.

Corresponding Author:

Jamal Mawane
Department of Computer Science, Faculty of Sciences and Technics of Mohammedia, Hassan II University
Mohammedia, Morocco
Email: ja.mawane@gmail.com, jamal.mawane@etu.fstm.ac.ma

1. INTRODUCTION

The use of recommendation systems in E-learning platforms dates back to the 2000's [1]. Since that time, researchers' efforts have concentrated on improving these systems through finding adequate solutions to the different problems faced by these platforms. Recent work in the field of recommendation in e-learning platforms have focused on the development of recommendation methodologies that can achieve better performance compared to existing recommendation strategies. However, these studies failed to address two important shortcoming of previously developed recommendation systems [2]:

- A very low recommendation success rate
- A very long recommendation time

The fact that the success rate of recommendation systems is very low and the lengthy time these systems take to produce recommendations have been a stumbling block to the efficiency and effectiveness of these systems.

Our previous work,the UDCF recommendation approach [3], focused on solving the two classical recommendation problems, i.e., a cold start and change of learner status during the use of the e-learning platform, and succeeded in improving the prediction rate of the items to be recommended. This way, It has reached significantly better results than classical collaborative filtering methods. Yet, without guaranteeing the homogeneity of the groups formed, this approach remains less efficient and the recommendations are less relevant as they are tied to learners learning styles that are heterogeneous and are constantly evolving along time.

To optimize the relevance of recommendation systems, numerous studies have used different methods to select the relevant items for platforms' users [4], [5] such as classic data mining techniques that compute a similarity to each user or item, and use the top scores to recommend the selected items [6]. Nevertheless, the two major issues mentioned persisted and recent work has failed to address them.

Thus, In an attempt to improve the first clustering block of the UDCF recommendation approach and in order to overcome the low rate of the recommendation and shorten the length of time required, the present work proposes a platform that tries to find the adequate parameters of the Kohonen maps, to create homogeneous clusters in profile data and results in order to improve the efficiency of learning resources recommendation. A real dataset-based simulation was conducted to validate this model and visualize the behavior of the proposed system. The method, analysis, and the obtained results of this simulation will be presented in this paper. The next section will synthesize the related work from the literature before proceding to the introduction of the proposed platform in details

2. RELATED WORKS

Research in recommendation systems(RecSys) in general have adopted several strategies of recommendation. To overcome the classical problems of recommendation, research has increasingly used hybrid techniques, and ML algorithms. In the education field, the use of such techniques has helped address the diversity of problems related to the personalization of content offered to learners by e-learning platforms. In fact, recommender systems try to automatically identify items that could be of interest for a certain learner [7]. Unspervised techniques have frequently been used in many others fields (Business studies, Arts and entertainment research) using data mining tasks and visualization tools. However, there is a dirth of research that analyzes the use of clustering algorithms in recommender systems and their behavior in different aspects. In e-learning platforms, empirical research based on unsupervised techniques can be classified as follows:

2.1. Classical Clustering Recommendation Research

Nafea *et al.* [8] used attributes for efficient clustering of learners to address both cold start and scoring sparsity problems, based on k-means clustering algorithm and Felder Silverman learning style model [9]. Based on their findings, the predictive accuracy of the model for new items (learning objective (LO)) reached interesting level (85%), except that the small size of the dataset used in the tests poses a problem of scalability and generalization of the model on all e-learning platforms. Asadi *et al.* [10] used student characteristics to determine similar students and highlight dependencies between learners' course choices, based on clustering algoritms and fuzzy association rules. The results obtained allow for the simplification of course selection decision making, except that the cold start problem arises for any new course, besides, learners are required to follow a particular learning path predefined in the learning platform program. The prior choice of the number of clusters is a major drawback of clustering algorithms. However, the mastery of the K-means algorithm and its implementation pleads in its favor much more than other unsupervised algorithms such as hierarchical or fuzzy clustering algorithms. The other research used Elbow method to determine the optimal number of clusters [11], [12] . Despite the use of the elbow method to determine the optimal k value, any change in the dataset requires a new search for the optimal value.

2.2. Kohonen map Recommendation Research

Kohonen maps or self-organizing maps (SOM) are used in different fields of scientific [13]-[16]. On the other hand, a review of research in e-Learning recommendation platforms reveals that there are only a few studies in this area in which SOM is used. Tai *et al.* [17] proposes a recommendation system with association rule extraction procedure based on two modules. The first one is an affinity clustering module based on SOM, which performs clustering tasks. The second one is a significant pattern extraction module using association rule extraction for each homogeneous group. Mawane *et al.* [3] proposes a two-block deep network approach, based on kohonen maps for learner profiling and a recommendation block implemented by an auto-encoder, to predict the outcome of any item to be recommended to a particular learner, from the list of items best performed by fellow learners of the same group. Despite the improved recommendation reached; Mawane *et al.* [3] does not ensure groups homogeneity. Overall, these studies focused on correcting the shortcomings of previous recommendations, but they did not take into account the evaluation of the homogeneity of the obtained clusters and their validity. Incorporating a validity criterion can provide two aspects of perfection:

- Certifying and explaining the issued recommendation, and
- Proposing an evaluation criterion of the recommendation system

3. RESEARCH METHOD

This section is divided into two main parts. The first one concerns the theoretical background of our method. The second part explains the components and working concept of the recommender framework WebApp.

3.1. Theoretical background

The main objective of this study is to improve the existing UDCF approach and optimize the results of the first block of the unsupervised recommender system using Kohonen maps and clustering validity. This will be done through ensuring homogeneity and a low coefficient of variation. The following part provides an overview of these main theoretical components.

3.1.1. Kohonen maps

A Kohonen map or self-organized map (SOM) [18] was inspired from the human brain structure and based on a competitive unsupervised learning system, in order to reduce the input data dimension N, into a perceptual space d (usually two dimensions), by grouping in a neuron the similar individuals input [19]-[22]. Self-organizing maps are often used for unsupervised classification of data with topology preservation. To this end, each neuron represents a class, and each observation is assigned to the neuron based on the closest reference vector. The Kohonen map algorithm is the major element of the first block of UDCF system, so the revision seeks to customize these maps to obtain a better recommendation result, by adding a clustering validity step using learners' results.

3.1.2. Clustering validity

Cluster validity refers to the process of evaluating the results of a clustering algorithm to ensure the validity of the partition and to determine the extent to which the result is significant. Unsupervised learning presents a serious problem of clustering. For this reason, evaluating the validity of the results of the partition is an essential step to ensure the significance of the results. In general, methods should propose clusters whose members have a high degree of similarity and are well separated at the same time [23]. Research generally distinguishes between three approaches to measure cluster validity [24]:

- External validity: to evaluate the extent to which supplied class labels are externally aligned with cluster labels
- Internal Validity: to measure the goodness of a clustering structure
- Relative criteria: to compare two different clusters or clustering

Clustering, one of the most important unsupervised learning problems, is the task of dividing a set of objects into clusters such that objects within the same cluster are similar while objects in different clusters are distinct. The current work has focused on the internal criterion, since the objective is to obtain homogeneous clusters whose results are obtained by the members of the same cluster and have a very low coefficient of variation.

3.1.3. Homogeneity and coefficient of variation

Homogeneity is an index that measures the relationships within a group according to one or more criteria. A homogeneous group would be formed by a set of members coming from the same classification algorithm, having a latent link, on the characteristics that defines each member, or a notorious link as examples, a similar social and educational itinerary, or sharing the same ideologies, etc. One criterion for estimating variability (or homogeneity) within a group is the coefficient of variation. The coefficient of variation (or CV) is a dimensionless, scale-independent indicator of the homogeneity of the population relative to its mean. The can be defined mathematically as a ratio of the standard deviation of the distribution of a measure representing the population to the arithmetic mean of that measure.

$$CV = \frac{\mu}{\sigma} \quad (1)$$

The notion of CV is generally attributed to Karl Pearson. Pearson used the to assess the relative variability of data and emphasized that differences in relative variability indicate "inequality of mutual correlations". The coefficient of variation has applications in many fields of research, it can be, used:

- to measure sensitivity to risk in medical sciences and industry,
- to evaluate variability in agricultural experiments,
- to represent the reliability of tests,

- to evaluate the homogeneity of groups obtained in social sciences and education [25].

A coefficient of variation of less than 15% is considered to indicate that the population is homogeneous, otherwise it indicates that the values are relatively dispersed [26], [27]. In order to evaluate the average homogeneity of the clusters, the (1) is applied on the results obtained by the members of a cluster, which allows to evaluate the CV_{item} of each item and subsequently calculate the average of the $\overline{CV}_{items}$ in each cluster. Table 1 illustrates an example of the calculation of the average CV for a $cluster_i$.

$$CV = \frac{1}{p}\sum_{j=1}^{p}\frac{\mu_{item_1}}{\sigma_{item_1}} = \frac{1}{p}\sum_{j=1}^{p}VC_{item_1} \tag{2}$$

The evaluation of the homogeneity of the clusters will be done by calculating the average of the CV_{cl} according to (3).

$$CV(Clusters) = \frac{1}{n}\sum_{j=1}^{n}CV_{Cl_i}, \text{ where n} = \Omega(Clusters) \tag{3}$$

Table 1. Coefficient of variations' computation for a cluster (2)

$Learners$	$Cluster_i(Cl_i)Learningressources$						
	$Item_1$	$Item_2$	-	$Item_i$	-	$Item_p$
L_1	$Ev_1 1$					$Ev_1 p$
L_2	x	$Ev_2 2$	.	$Ev_2 i$	.	x
.	.	.	.	.	.	.
.	.	.	.	.	.	.
.	.	.	.	.	.	.
L_k	Ev_{k1}	Ev_{k2}	.	Ev_{ki}	.	Ev_{kp}
Coefficient of Variation (CV)	$CV_{item_1} = \frac{\mu_{item_1}}{\sigma_i tem_1}$	CV_{item_2}	.	CV_{item_i}	.	CV_{item_p}

3.2. Kohonen map parameters recommender framework

This part delineates the implementation of the first block of the UDCF approach, as well as the improvement modifications to be made to the kohonen map algorithm as proposed in this paper. The following sections explain the underpinnings of this revision starting with the input of the system.

3.2.1. Data collections and preprocessing

The objective of this study is to discover homogeneous clusters not only in terms of the profile characteristics of learners, but also at the level of their results. In general, the development of the clustering model must be independent of any collection of learning data, directly related to the aspects of the learners, and subsequently the validation of the model is verified through the results obtained. For this reason, the dataset is divided into two parts:

- Profile dataset: The characteristics of the learner's profile (age, interactivity, and level) retained in the e-learning platform for model prediction [9].
- Results dataset: The current learning results for model validation.

3.2.2. Kohonen map parameters recommender framework

The system needs to apply the correct map parameters to ensure that the model developed classifies the learners appropriately, based on a low coefficient of variation of learners results of each cluster, with the aim of getting a higher recommendation accuracy. Figure 1 gives an overview of how the framework functions.

To find the appropriate parameters values, and to ensure that our approach can be used in any e-learning platform, a webapp framework was developed to analyze the system's data, and to recommend optimized Kohonen map configurations with the lowest coefficient of variation and the highest recommendation accuracy to it. The Kohonen results validity webapp is designed around two main modules. Figure 2 gives an overview of these modules and their inputs and outputs.

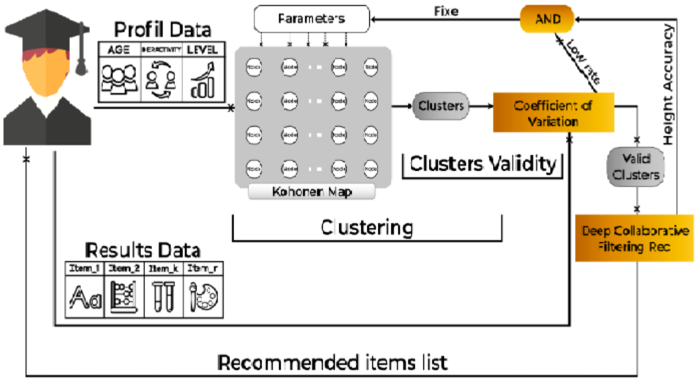

Figure 1. The architecture of the framework

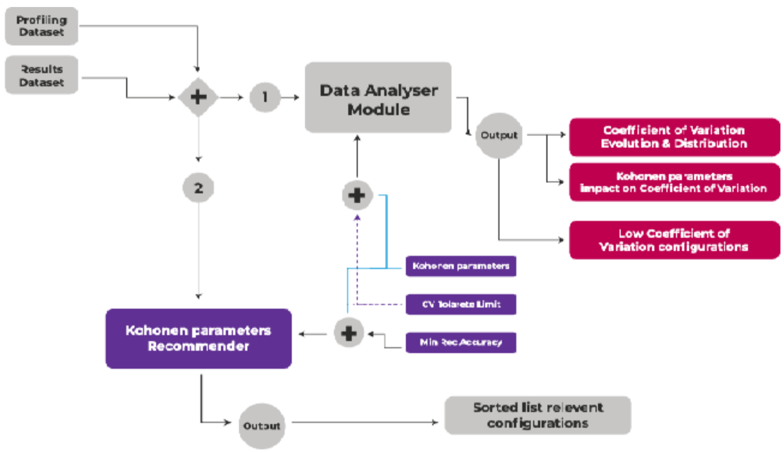

Figure 2. Kohonen results validity webapp modules description

Module 1: Data analyses. The 1^{st} module uses an adaptation of the kohonen algorithm, performing a blind or targeted search according to one or more parameters of the Kohonen maps, and an analysis of the clustering results and the average CV rate obtained. The evolution of the CV and the impact of any parameter of the kohonen maps on the average value of the obtained CV will be introduced in detail in the case study part.

For each combination of parameter values, Algorithm 1 executs the clustering by kohonen maps method, then the average CV is calculated on the obtained clusters, according to the process described in the definition part of the (3). The Algorithm 1 evaluates and compares the CV value to the variability bar to check for eligibility and to insert the combination of parameters in the list of possible configurations if it is below the bar.This is then sorted in ascending order of the average CV and truncated to the first K configurations to be displayed adjusted to the needs.

Input: This module automatically starts the analysis once the profiling data and the results are provided as input. The empirical analysis performs a blind search on a list of values by default, or a search on specific values of one or more hyperparameters of the Kohonen maps. The list of specific values for each parameter is presented in Table 2.

Table 2. Input parameters list of specific values and descriptions

Parameter	List possible values	Note
Nodes' number	$\lfloor[1, 5\sqrt{n}]\rfloor$	$5\sqrt{n}$ (Ref Kohonen 2) n number of learners
Square topology	from 2x2 to nxn	n=$\lfloor 5\sqrt{n}\rfloor$
Learning rate	$[0.1, 0.9]$	
neighborhood radius	$[\|1, Dim\|]$	Dim : The selected value of the topologie
Homogeneity limit	100%	

Algorithm 1: Input data analyser module

```
AnalyserModule(Map_Dimension = Map_dl,iterations = It_dl,init_Lr = Lr_dl,init_radius =
  Radius_dl,Homo_limit = 100%):                        // dl: parameter default list
List_config ⟵ [ ]
foreach Dim in Map_Dimension do
    foreach L_r in init_Lr do
        foreach radius in init_radius do
            foreach Iter in iterations do
                map = Map_Training(parameters)
                clusters = Clustering(map)
                CV = avg_Coef_Variation(clusters)          // references to Eq.3
                Draw_evolution(CV)
                Draw_distribution(CV)
                if CV < Homo_limit then
                    List_config ⟵ insert([parameters, CV])
Draw_params_impact(CV)
return List_config                    // Empty list if no configuration is valid

EndFunction
```

Result: At the end of the analysis, the 1^{st} module proposes K optimized configurations of the kohonen map parameters with the lowest average CV according to our cluster validity criterion (Table 3). Once the recommendation parameter is activated, the application switches to the second module. Module 2: Finding Kohonen parameters. The 2^{nd} module aims to recommend the configurations of the Kohonen maps giving a homogeneity ratio below the tolerance limit and the highest prediction accuracy of recommendation (Algorithm 2). Note also that a minimum accuracy rate limit can be applied to search for configurations whose recommendation rate is greater than this limit. In addition, at the end of the search, the 2^{nd} module displays a table summarizing the configurations to be recommended as well as the average CV, the accuracy, and the number of clusters among the nodes of the corresponding maps.

Table 3. K configurations with lowest CV (K=5)

5 Relevent config	Map_Dim	Iterations	Learning_rate	Neighbor_radius	Variation_Coef	Clusters_Number
1294	9	2000	0.900	9	0.0848	50
1702	10	2000	0.700	10	0.0916	56
1742	10	2000	0.600	10	0.0946	54
1915	10	3000	0.100	3	0.1009	76
1575	9	3000	0.100	7	0.1009	36

Algorithm 2: Kohonen map parameters recommendation

```
Rec_Module(Map_Dimension = Map_dl,iterations = It_dl,init_Lr = Lr_dl,init_radius =
  Radius_dl,Homo_limit = 100%,Rec_Acc):                // dl: parameter default list
Rec_config ⟵ [ ]
foreach Dim in Map_Dimension do
    foreach L_r in init_Lr do
        foreach radius in init_radius do
            foreach Iter in iterations do
                map = Map_Training(parameters)
                clusters = Clustering(map)
                CV = avg_Coef_Variation(clusters)          // references to Eq.(3)
                if CV < Homo_limit then
                    Accuracy = avg_Rec_predict(clusters)
                    if Accuracy ≥ Rec_Acc then
                        Rec_config ⟵ insert([parameters, CV, Accuracy, clusters])
return List_config                    // Empty list if no configuration is valid

EndFunction
```

Synthesis: The main objective of the revision, proposed in this paper, is to form clusters of learners of homogeneous profile validated by their results also proving the homogeneity of the clusters, and subsequently improving the recommendation within the formed clusters. In this paper, we test our webapp on the same dataset used as an experimental support for the deep unsupervised collaborative recommendation approach.

4. RESULTS AND DISCUSSION

In this section, Numerous tests are conducted to analyze the impact of each parameter of the kohonen self-organizing map algorithm on the homogeneity of the obtained clusters. To that end, throughout this part, different results of the empirical study, as well as the answer of some simulation of our framework will be exposed as we go along, taking as test base the dataset (OULAD) [28]. The objective is to determine the optimal parameters and to verify the performance of our proposed Framework (Revised UDCF) compared to the result of our previous UDCF approach and traditional collaborative filtering recommendation methods.

4.1. Dataset description

As described at the beginning of the previous section, student profiling is based on several facets, namely: personal data, level, and interactivity. The student profiling clustering validity is gauged by the degree of closeness of the results in the common items of each cluster. The synthesis of the StudentInfo and StudentVle tables allowed us to build the profiling dataset, while the StudentInfo, studentAssessment, Assessment, and Courses tables were used to build the result dataset [28].

4.2. The Kohonen map size impact test

From Figure 3, it can be seen that increasing the size of the Kohonen map considerably reduces the disparity of the results in the groups based on these maps. The objective in this case can be considered basic since increasing the size gives a significant similarity of results. Yet, increasing the size will create groups with a limited population size and subsequently will disfavor collaboration and recommendation. In this case, the optimization is to find the right size of the map giving the lowest possible coefficient of variation of the results. To do so, a simultaneous analysis of both the map size and the other parameters of the map is presented next.

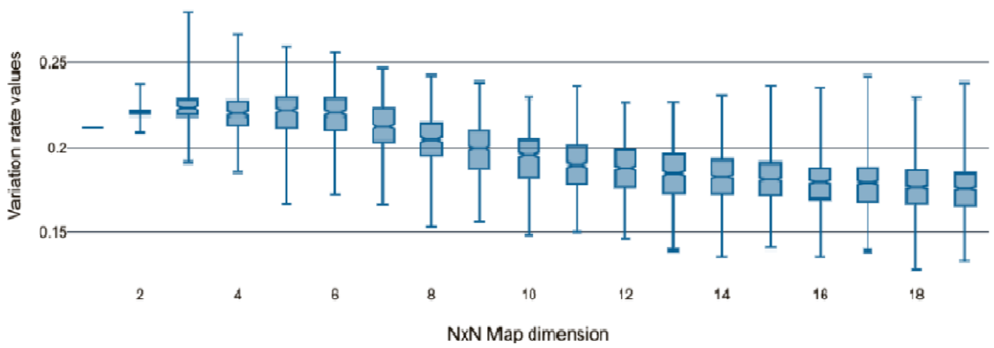

Figure 3. The impact of the maps size on the homogeneity of the clusters

4.3. The neighborhood impact test

Figure 4 shows the impact of the combination of map size and neighborhood radius on the average CV. The radius is by default limited to the map size, so Table 4 gives an overview of the lists of radiuses that allow to obtain a low value of the average CV of each size. The results presented in Figure 4 indicates that there is no significant impact of the radius neighborhood on the low rate of CV. However, we notice that the low average coefficient of variation rates are obtained on an interval of initial neighborhood radius which varies between both values $\left[\frac{dim}{2}, dim\right]$.

4.4. The Learning rate impact test

Figure 5 shows the impact of the combination of map size and initial learning rate on the average CV. We notice that for any learning rate value between 0.1 and 0.9, there is always a possibility to have at least one configuration giving homogeneous clusters, except that the result obtained does not allow to select a subset of the particular learning rate values.

4.5. The number of iterations impact test

In our empirical study, an overview of the impact of variation of the number of iterations is shown in the Table 5 and Figure 6 presents the minimum value for every map dimension. As shown in Table 5, the number of iterations equal to 2000 allows the different configurations of the Kohonen maps to have the lowest average rate of change or the one closest to the minimum value of the other numbers of iterations.

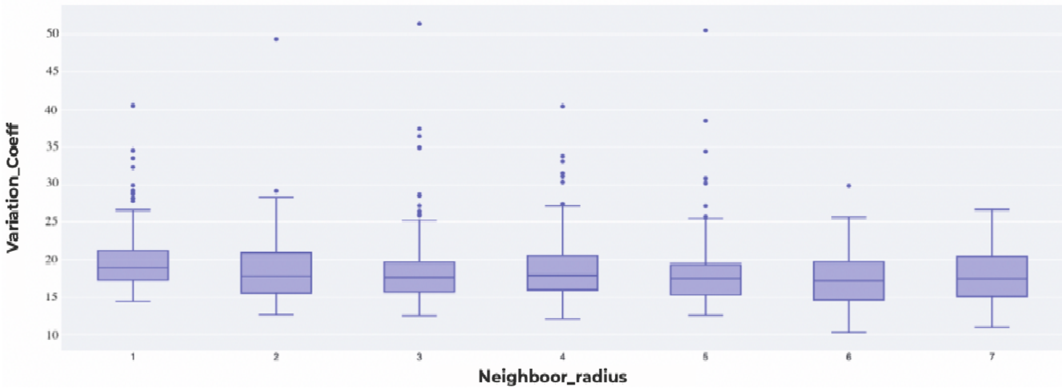

Figure 4. Neighborhood radius impact

Table 4. Optimal radiuses list

Map dimension	Initial neighboors radius 1	2	3	4	5	....	15	16	17	18	19
2	0.219	0.209				.					
3	0.197	0.193	0.192			.					
4	0.188	0.203	0.195	0.186		.					
5	0.188	0.209	0.193	0.194	0.167	.					
.	.	.	.	.	.	.	.	.	.	.	.
16	0.193	0.172	0.173	0.171	0.162	.	0.139	0.139			
17	0.188	0.180	0.159	0.145	0.153	.	0.145	0.140	0.146		
18	0.185	0.169	0.176	0.170	0.161	.	0.142	0.143	0.129	0.141	
19	0.184	0.167	0.161	0.168	0.145	.	0.144	0.133	0.144	0.141	0.136

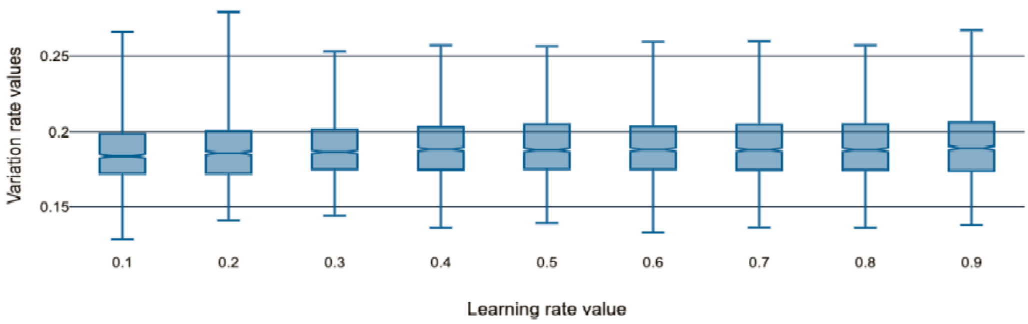

Figure 5. The impact of the initial learning rate on the homogeneity of the clusters

Table 5. CV iterations' impact

Map dimension iterations	2x2	3x3	4x4	5x5	..	15x15	16x16	17x17	18x18	19x19	Mean value
1000	0.212	0.192	0.199	0.187	..	0.155	0.155	0.154	0.157	0.153	0.174
2000	0.209	0.192	0.200	0.167	..	0.141	0.139	0.139	0.129	0.133	0.161
3000	0.213	0.193	0.195	0.196	..	0.146	0.136	0.143	0.142	0.144	0.168

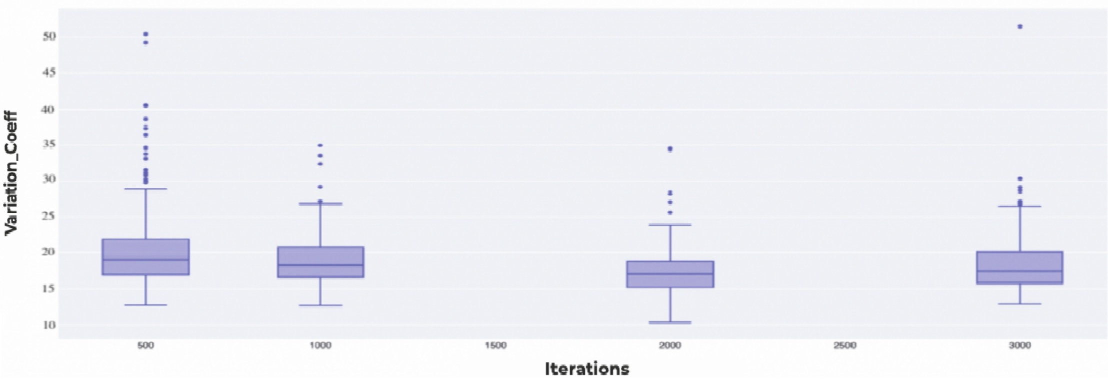

Figure 6. Iterations impact

4.6. Recommendation of Kohonen map configurations

In order to test the recommendation rate of selected configurations of homogeneous population Kohonen maps, a symmetric autoencoder of reduced depth, as shown in Figure 7, is used to evaluate the accuracy rate of unsupervised deep collaborative filtering (UDCF) recommendation for e-learning platforms. Thus, the average of the recommendation rates obtained, by applying such a configuration on all the clusters of the map (Figure 8), is retained as the recommendation score of the block. The result of the recommendation block obtained from the dataset is presented in Table 6.

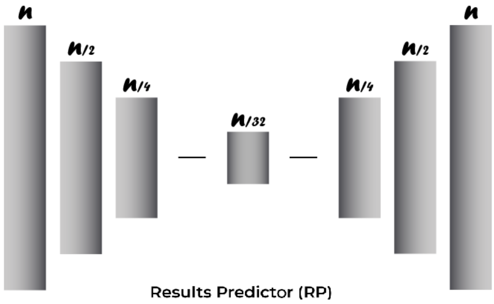

Figure 7. Results predictor (RP) autoencoder' configuration

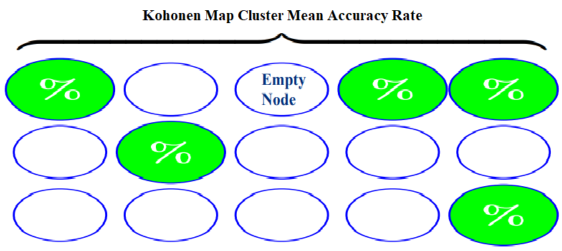

Figure 8. Recommendation configuration

Table 6. Recommendation results (valid configurations)

No.	Square dimension value	Iterations number	Initial learning rate	Initial neighboors radius	Coefficient of variation	Rec accuracy	Number of clusters
1	19	2000	0.900	19	5.868	82.38	164
2	20	2000	0.900	19	6.790	82.43	185
3	20	2000	0.700	20	6.918	82.51	163
4	17	2000	0.900	17	7.860	82.43	131
5	17	2000	0.800	17	7.220	82.54	128

4.7. Discussion

The result obtained, by revising the first block of UDCF approach, and comparing it to the previous result, as well as to the classical recommendation algorithms [3], is exposed in the table of results (Table 7). The revised UDCF accuracy is slightly lower than the first UDCF despite the use of a simple, non-deep and symmetric configuration of the autoencoder network of the second block of the approach (not subject of this revision). This result is very stimulating for the second phase of the revision, in order to recommend optimized configurations of the kohonen maps not only to obtain high homogeneity clusters but also to reach a high recommendation rate.

Table 7. Revised unsupervised deep e-learning RS compared to relevant methods (Evaluation)

RecSys methods	Accuracy (100%)
Restricted boltzmann machine (RBM)	77.45
User-item-based collaboratif filtering	77.45
Revised UDCF	82.32
UDCF	83

5. CONCLUSION

An adaptive and personalized e-learning system aims to support the learners' individual needs and interests to facilitate the learning process. The researchers in this area have tried a variety of techniques to achieve this goal. However, ensuring relevant recommendations for a wide range of varied profiles and learning preferences has been a daunting challenge. In this paper we set out to revise and improve the results of our previous UDCF model through using clustering validity and Kohonen maps configuration, in order to increase the rate of cluster homogeneity and achieve a low coefficient of variation rate. The proposed platform has reached a high rate of accuracy of the model testifying to the relevance of the approach implemented. Thus, this proposed model can help meet the relevance of recommended items by classifying learners into homogeneous groups before generating recommendations. The revised UDCF recommendation system in e-learning platforms provides a novel addition to the field of hybrid recommendation systems in E-learning by providing a more personalized learning resource to each Learning from relevant resources filtered from the other members of the same homogenic cluster. The revised UDCF WebApp will find the relevant configurations of kohonen map parameters to get homogenous groups of learners based on their profiling data and assessment once the learning process is initiated by the learner. Our future research will focus first on a revision of the learning item recommendation block (block 2), in order to enable Webapp to recommend complete optimal configurations of unsupervised deep e-learning collaborative recommendation platform approach for any e-learning platform. Secondly, we will reconsider the learner's contribution both as an implicit actor of the collaborative recommendation, and as an explicit actor of the selectivity of the most relevant learning.

REFERENCES

[1] O. R. Zaiane, "Building a recommender agent for e-learning systems," *International Conference on Computers in Education, 2002. Proceedings.*, 2002, pp. 55-59 vol.1, doi: 10.1109/CIE.2002.1185862.

[2] F. Ricci, L. Rokach, and B. Shapira, "Recommender systems: introduction and challenges," in *Recommender systems handbook*, Springer, 2015, pp. 1-34, doi: 10.1007/978-1-4899-7637-6_1.

[3] J. Mawane, A. Naji, and M. Ramdani, "Unsupervised deep collaborative filtering recommender system for e-learning platforms," in *Smart Applications and Data Analysis.SADASC 2020*, Springer, Cham., 2020, pp. 146-161, doi: 10.1007/978-3-030-45183-7_11.

[4] S. Souabi, A. Retbi, M. K. Idrissi, and S. Bennani, "Towards an evolution of e-learning recommendation systems: From 2000 to nowadays," *International Journal of Emerging Technologies in Learning*, vol. 16, no. 6, pp. 286-298, 2021, doi: 10.3991/ijet.v16i06.18159.

[5] S. S. Kundu, D. Sarkar, P. Jana, and D. K. Kole, "Personalization in education using recommendation system: An overview," *Computational Intelligence in Digital Pedagogy*, vol. 197, pp. 85-111, 2021, doi: 10.1007/978-981-15-8744-3_5.

[6] J. Wang, H. Xie, F. L. Wang, L.-K. Lee, and O. T. S. Au, "Top-N personalized recommendation with graph neural networks in MOOCs," *Computers and Education: Artificial Intelligence*, vol. 2, p. 100010, 2021, doi: 10.1016/j.caeai.2021.100010.

[7] M. Balabanović and Y. Shoham, "Fab: content-based, collaborative recommendation," *Communications of the ACM*, vol. 40, no. 3, pp. 66-72, 1997, doi: 10.1145/245108.245124.

[8] S. M. Nafea, F. Siewe, and Y. He, "A novel algorithm for course learning object recommendation based on student learning styles," *2019 International Conference on Innovative Trends in Computer Engineering (ITCE)*, 2019, pp. 192-201, doi: 10.1109/ITCE.2019.8646355.

[9] R. M. Felder and L. K. Silverman, "Learning and teaching styles in engineering education," *Engineering education*, vol. 78, no. 7, pp. 674-681, 1988.

[10] S. Asadi, S. Jafari, and Z. Shokrollahi, "Developing a course recommender by combining clustering and fuzzy association rules," *Journal of AI and Data mining*, vol. 7, no. 2, pp. 249-262, 2019, doi: 10.22044/JADM.2018.6260.1739.

[11] B. A. Jaafar, M. T. Gaata, and M. N. Jasim, "Home appliances recommendation system based on weather information using combined modified k-means and elbow algorithms," *Indonesian Journal of Electrical Engineering and Computer Science (IJEECS)*, vol. 19, no. 3, pp. 1635-1642, 2020, doi: 10.11591/ijeecs.v19.i3.pp1635-1642.

[12] J. Mawane, A. Naji, and M. Ramdani, "Clustering collaborative filtering approach for Diftari e-learning platform' recommendation system," in *SITA'18: Proceedings of the 12th International Conference on Intelligent Systems: Theories and Applications*, 2018, pp. 1-6, doi: 10.1145/3289402.3289535.

[13] D. Galvan, L. Effting, H. Cremasco, and C. A. Conte-Junior, "The spread of the COVID-19 outbreak in Brazil: An overview by Kohonen self-organizing map networks," *Medicina*, vol. 57, no. 3, 2021, doi: 10.3390/medicina57030235.

[14] E. C. Djamal, D. P. Gustiawan, and D. Djajasasmita, "Significant variables extraction of post-stroke eeg signal using wavelet and SOM Kohonen," *TELKOMNIKA (Telecommunication, Computing, Electronics and Control)*, vol. 17, no. 3, pp. 1149-1158, 2019, doi: 10.12928/telkomnika.v17i3.11776.

[15] F. Ciampi, V. Cillo, and F. Fiano, "Combining Kohonen maps and prior payment behavior for small enterprise default prediction," *Small Business Economics*, vol. 54, pp. 1007-1039, 2018, doi: 10.1007/s11187-018-0117-2.
[16] M. Ha, C. Gan, C. Nguyen, and P. Anthony, "Self-organising (Kohonen) maps for the Vietnam banking industry," *Journal of Risk and Financial Management*, vol. 14, no. 10, p. 485, 2021, doi: 10.3390/jrfm14100485.
[17] D. W.-S. Tai, H.-J. Wu, and P.-H. Li, "Effective e-learning recommendation system based on self-organizing maps and association mining," *The Electronic Library*, vol. 26, no. 3, pp. 329-344, Jan. 2008, doi: 10.1108/02640470810879482.
[18] T. Kohonen, "Essentials of the self-organizing map," *Neural Networks*, vol. 37, pp. 52-65, 2013, doi: 10.1016/j.neunet.2012.09.018.
[19] X. Wang and K. K. Paliwal, "Feature extraction and dimensionality reduction algorithms and their applications in vowel recognition," *Pattern Recognition*, vol. 36, no. 10, pp. 2429-2439, Oct. 2003, doi: 10.1016/S0031-3203(03)00044-X.
[20] S. Ayesha, M. K. Hanif, and R. Talib, "Overview and comparative study of dimensionality reduction techniques for high dimensional data," *Information Fusion*, vol. 59, pp. 44-58, Jul. 2020, doi: 10.1016/j.inffus.2020.01.005.
[21] P. Ray, S. S. Reddy, and T. Banerjee, "Various dimension reduction techniques for high dimensional data analysis: a review," *Artificial Intelligence Review*, vol. 54, no. 5, pp. 3473-3515, Jun. 2021, doi: 10.1007/s10462-020-09928-0.
[22] B. Venkatesh and J. Anuradha, "A review of feature selection and its methods," *Cybernetics and Information Technologies*, vol. 19, no. 1, pp. 3-26, Mar. 2019, doi: 10.2478/cait-2019-0001.
[23] Y. Liu, Z. Li, H. Xiong, X. Gao, and J. Wu, "Understanding of internal clustering validation measures," in *2010 IEEE International Conference on Data Mining*, 2010, pp. 911-916, doi: 10.1109/ICDM.2010.35.
[24] K. Koutroumbas and S. Theodoridis, *Pattern recognition*, 4th ed. Burlington, MA London: Academic Press, 2009.
[25] C. Santos and C. Dias, "Note on the coefficient of variation properties," *Brazilian Electronic Journal of Mathematics*, vol. 2, no. 4, pp. 101-111, 2021.
[26] C. Simard, *Méthodes quantitatives. Version étudiante*, 6th ed. Modulo, 2018.
[27] G. Baillargeon, *Methodes statistiques avec applications dans differents secteurs de lentreprise*, vol. 1, Canada: Trois-Rivieres, 1984.
[28] J. Kuzilek, M. Hlosta, and Z. Zdrahal, "Open university learning analytics dataset," *Scientific data*, vol. 4, no. 1, pp. 1-8, 2017, doi: 10.1038/sdata.2017.171.

BIOGRAPHIES OF AUTHORS

Jamal Mawane received his Master's degree in Science and Technology from the Faculty of Science and Technology of Gueliz of the Cadi Ayyad University of Marrakech and his Master's degree in Computer Engineering specializing in Distributed Information Systems from the Hassan 2 University of Casablanca in 2012. He is presently employed as a computer science instructor in two-year extremely selective programs preparing students for competitive tests to the "Grandes Ecoles" system. His study focuses on E-Learning platform recommendation algorithms. He may be reached at the email address ja.mawane@gmail.com.

Abdelwahab Naji received a Ph.D. degree in computer science from the Faculty of Science And techniques, Hassan II University, Casablanca, Morocco, in 2013, He is currently a Professor of Computer science and machine learning at ENSET of Moahammedia. His current research interests include machine learning algorithms, metaheuristic optimization methods, intelligent education systems, and robot navigation. He may be reached at the email address abdelwahab.naji@gmail.com.

Prof. Mohamed Ramdani is a Director of Intelligence Machine Laboratory (LIM), FST Mohammedia-University Hassan II of Casablanca, and President of Moroccan Association of Intelligent Systems. Received his Ph.D. in Fuzzy Machine Learning in 1994 at the University Paris VI, France. He became Assistant Professor (1995), in 2001 he obtained his HDR in perceptual computation at University Paris VI, France, and in 2005 he became Professor at the University Hassan II of Casablanca, Morocco. For the periods 1996-1998 and 2003-2005, he held the position of head of the Computer Science department at the Faculty of Sciences and Technologies of Mohammedia. Between 2008 and 2014, he was Pedagogical Director of the engineering department "Software Engineering and Systems Integration" (ILIS) within the Faculty of Sciences and Technologies of Mohammedia. Since 2006, he is the Director of the Computer Science Lab at the University Hassan II of Casablanca. In May 2016, he is President of the Moroccan Intelligent Systems Association (AMSI). From 2017 to 2019, he is the director of the project "Connected Objects and Bigdata"(BDOC) at the University Hassan II of Casablanca. His research interests include Explanation in Machine Learning, perceptual computation with Fuzzy Logic, and Big data mining. He has directed several Ph.D. theses in these fields and published several articles in many indexed journals. He may be reached at the email address ramdani@fstm.ac.ma.