

## Learning Algorithm of Kohonen Network With Selection Phase

Mohamed Ettaouil<sup>1</sup>, Essafi Abdelatifi<sup>2</sup>, Fatima Belhabib<sup>3</sup>, and Karim El moutaouakil<sup>4</sup>

<sup>1</sup>UFR: Scientific Computing and Computer sciences, Engineering sciences

Modelling and Scientific Computing Laboratory

university Sidi Mohamed Ben Abdelah

Box 2202, Fez

MOROCCO

<sup>1</sup>MohamedEttaouil@yahoo.fr

<sup>2</sup>essafiabdelatifi@hotmail.com

<sup>3</sup>F.BELHABIB@yahoo.fr

<sup>4</sup>karimmoutaouakil@yahoo.fr

*Abstract:* - The learning algorithm of Kohonen network realized away from any notion of class. So, the labelling phase is necessary to find the class associated to data. Generally, the size of the topological map is randomly chosen. Such choice effect the performance of the Kohonen algorithm. Consequently, the labelling phase becomes difficult to realize. To overcome this problem, we add to learning Kohonen algorithm a phase called selection stage. This phase is based on a function called selection function; to construct, we use a sub-set of the data set. In addition, we divided the topological map on two parts: The first contains the used neurons; while the second one is formed by the unused ones. It should be noted that the size of the first and the second parts are modified by the use of the selection function. To compare our method with the classical one, some experiments results are performed.

*Key-Words:* -Kohonen Network, Learning Kohonen, Neural architecture of optimization, Kohonen with Selection phase

### 1 Introduction

The Kohonen algorithm is an automatic classification method which is the origin of Self-Organizing Maps (SOM)[9]. This famous method falls within the framework of algorithms quantification vector and the method of k-means algorithm. More precisely, the SOM can be seen as an extension of the algorithm of pattern recognition and the automatic classification algorithms [5].

The latter algorithms detail the possible methodologies for introducing expertise that follows the learning unsupervised stage [1].

Because its interesting and specific quality, Kohonen is efficient tool in the important domain which is the unsupervised classification . Indeed this latter is considered as a useful tool in data mining because it accelerates the research of relevant information. In fact the access to this latter is in suitable time has become a daily need and

difficult task. This is due to the huge amount of information available in the website. An unsupervised classifier groups the similar information that refer to the same topic in same cluster and these which are dissimilar in the distinct ones. This avoid the search of the desired information in a lot of clusters, consequently an important time is economized. The Kohonen algorithm is unsupervised partitioned classifier i.e. it treat with unlabeled inputs and provides classes with no overlap. Beside its robustness i.e. its ability to resist to the noise, the Kohonen algorithm possesses other interesting properties. Indeed the self-organizing map is an unsupervised neural network which projects high-dimensional data onto a low-dimensional grid which called a topological map [9]. This projection must preserve the topology of inputs (more details of this point is given after in this paper). This lead to an organization and representation in map. This latter provides a



training, continuous-valued input vectors are presented either sequentially in time or in batch without specifying the desired output. This is called unsupervised learning. In addition, the weights will be organized such that topologically close nodes are sensitive to inputs that are physically similar. Output nodes will thus be ordered in a natural manner. This may be important in complex systems with many layers of processing because it can reduce lengths of inter-layer connections. After enough input vectors have been presented, weights will specify clusters or vector centers that sample the input space such that the point density function of the vector centers tend to approximate the probability density function of the input vectors [19].

The neighbourhoods defined in the map can be chosen so varied, but generally they have rectangular form

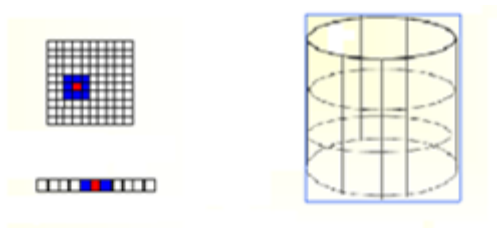


Fig 2 example of neighborhood in the kohonen

The advantage and disadvantage of this algorithm can be summarized as follows:

- Advantage
  - Visualization the of inputs and the Relationship between the classes
  - Projection of linear and nonlinear data while preserving the topology of the input space into a lower dimensional space: then it can be used as a means of dimension reduction and vector quantization.
  - Ability to solve problems with missing data
  - Auto-organisation of given inputs

➤ Disadvantage

The major disadvantage comes from the initialization phase in which the size of the map and the initial weight vector are randomly chosen.

Because of its effectiveness in resolution of unsupervised learning problems, the Self-Organizing Map (SOM) is used as a handy tool for feature analysis of high-dimensional data. This allow to a visualization in a low-dimensional space neuron layer . It also can be used in a speech recognizer as a vector quantization [20]. Kohonen has proposed an algorithm of self-organization planning space data on a discrete area of small size.

SOMs are the most commonly used strategy in Artificial Neural Networks for unsupervised learning [1]. During the training process the neurons tend to represent statistical properties of the input data, preserving the topology of the input space, even though it is unknown.

### 3. Related work

As is reported above the initial phase which is performed randomly has a great effect on the performance of the Kohonen algorithm. Indeed the choice of the initials weights and the neurons number has an impact into the convergence of the training methods and the stabilization of the network after the training stage [20]. Because of interesting quality of this algorithm many researches were proposed to overcome these drawbacks and improve its performance. Some of them use the techniques used in selection of relevant features; others use increasing network gas [1]. We find also some approaches which built the map in evolutionary way by adding some neurons and deleting others. More recently, a new approach was proposed to solve the problem of the optimization of the Kohonen network architecture [2]. This method consists modeling the investigated problem in term of a polynomial mix-integer problem with no linear-constraints [4], [7]. We recall that the of the architecture map signify to find the optimal number of hidden layers in the ANN, the number of neurons within each layer, and the good activation function, in order to maximize the performance of artificial neural networks [22]. Many variants and extensions have since been proposed, including the Visualization Induced SOM (ViSOM) [11]. In the literature, there are some context-aware SOM variants and typical examples are the SOAN (Self Organization with Adaptive Neighborhood Neural Network) and the Parameter Less PLSOM. Both use the current mapping error to adjust the internal parameters of the adaptation process. In the Time-Adaptive SOM (TASOM) [10],

Before concluding this section we present the following remark:

It is important to report, that although the simplicity of this algorithm its convergence to a stable state is still a difficult task: Indeed the algorithm must converge to the minimum of the following function:

$$D_{SOM}(x) = \sum_{i=1}^n \sum_{x/i=i_0(x)} \|x - C_i\|^2$$

Where  $i$  is close to  $i_0(x)$

In fact this function is not differentiable and has many local minima. So it presents mathematical difficulties.

#### 4. Kohonen Algorithm with a Selection Phase

Generally, the size of the topological map is randomly chosen. Such choice effect the performance of the Kohonen algorithm. Consequently, the labeling phase becomes difficult. To correct this arbitrary choice, we add to the learning Kohonen algorithm a selection phase which is based on a function. In this section, we give the philosophies of our method. Basing on this discussion, we propose a new algorithm which ameliorates the Kohonen one.

##### 4.1 The Proposed Method

In this part, we describe with some detail the main component of our approach.

It should be noted that this method consists integrating a new phase in the Kohonen algorithm. This later is called selection phase and has as role to facilitate the labelling phase.

##### Decomposition of the topological map

We call active neuron, the neuron of the topological map which participates in the competition phase. Our method consists decomposing the topological map into two groups of neurons : the first group is denoted  $P_A$  and is formed by the active neurons, the second group of neurons is denoted  $P_{NA}$ , and contains the non-active neurons. Several decompositions can be proposed, but the problem is to find the suitable one. In this context, we mean by a suitable decomposition the one whose the active neurons give a good representation of the investigated data. In this work, we construct iteratively this suitable decomposition. To this end, we begin by an arbitrary decomposition and we update it steep by steep using an adequate tool.[13][16].

##### The selection function

The classical Kohonen algorithm consists decomposing the investigated data on tree sets which are : the learning set  $A$ , the test set  $T$  and the validation set  $V$ . Our approach consists integrating an additional set  $S$  called the selection set. Basing on this set and using the set of the active neurons  $P_A$  and the set of the non-active neurons  $P_{NA}$ , we build a function  $s$  called selection function as follows:

$$s(j) = \sum_{y \in S} \sum_{i \in P_A} e^{-\frac{\delta(i,j)}{\tau}} \|y - w^j\|^2 \quad (1)$$

The real  $\tau$  is the parameter which controls the neighborhood size. The value  $s(j)$  quantifies the quality of the neuron  $j$ . In fact, this value measures the adaptation of the neuron  $j$  to the constructed topology on the set  $P_A$  and the geometrical error associated to the neuron  $j$ . In this regard, if  $L$  is a sub-set of the neuron map, the neuron  $l$  which satisfy the equation  $s(l) = \text{Min}\{s(j)/j \in L\}$  is the most adapted neuron, in comparison with the neurons of the set  $L \setminus \{l\}$ , to the order constructed on the set of the active neurons  $P_A$ . Moreover, this neuron is associated with the smallest geometrical error for the current representation of the data. Basing on this discussion, we will use the function  $s$  to update the topological map decomposition.

##### The update of the topological map decomposition

Using the function 2, we integrate a new phase in the classical Kohonen algorithm. This phase is divided into two steps:

- In the first step, we active a neuron  $p$  from the set  $P_{NA}$  via the following equation :

$$p = \arg \min_{l \in P_{NA}} s(l) \quad (2)$$

This choice can be explained by the fact that the neuron  $p$  which satisfy the equation  $s(l) = \text{Min}\{s(r)/r \in P_{NA}\}$  is the most adapted neuron, in comparison with the neurons of the set  $P_{NA} \setminus \{p\}$ , to the order constructed on the set of the active neurons  $P_A$ . Moreover, this neuron is associated with the smallest geometrical error for the current representation of the data.

- In the second step, we put off the competition the neuron  $h$  of the set  $P_A$  which represents the smallest number of data, and which has participated in at least  $[N/2]$  competitions. Where  $[a]$  represents the integer part of the number  $a$ .

The data represented by the neuron  $h$  will be affected to the neuron  $q$  which fulfilled the following condition:

- 1)  $q$  is a neighbor of the neuron  $j$ ;
- 2)  $q$  represents a largest number of data.

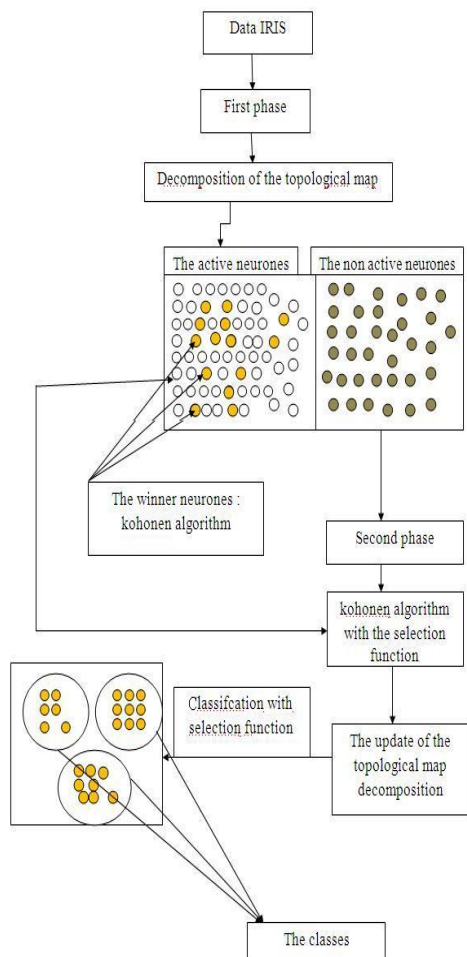


Fig.3 The classification diagram

The diagram shows the different steps of our method

**Remarks**

The choice of the selection set has great impact on the algorithm performance and on its complexity. The selection of neuron must reflect the most information of the population space. Since the selection function is defined on  $S$ , a big size of this set increases the complexity of the system. So, a suitable choice is required [17][15].

- The dimension of the investigated inputs is typically huge. The neural network of kohonen is commonly used as an efficient mean to reduce this dimension. In the beginning no information about the class is available. So the choice of map size and the active neuron is made randomly. Hence the performance will be affected and some problem will arise in the labeling phase. The update of the set  $P_A$  and  $P_{NA}$  can be considered as a correction of initialization phase and also as a way for searching a suitable architecture map. Thereby the performance of the system will be improved.
- The proposed method modelizes a behavior used naturally by the human. Indeed, this latter when he wants accomplishes a task., at first he choose randomly a set of persons. After, this set is updated by liberating some persons and adding others basing on their competency. This behavior is also used by certain animals as ants [15].

**3.2 A Training Algorithm with a Selection Phase**

Basing on the notions of selection set  $S$ , selection function  $s$ , active part  $P_A$  and non active part  $P_{NA}$  of the topological map, we integrate in the Kohonen algorithm. A new phase called a selection phase. The procees of our method can be summarized as follows:

- The initiation phase: In this phase we initialize the training set  $P_A$ , the inactive set  $P_{NA}$ , the test set  $T$  and the selection set  $S$ .
- The classical phase: After the initialization phase the classical Kohonen algorithm is conducted on the set  $P_A$ .
- Selection phase: In this stage, using the selection function  $S$ , a neuron from  $P_{NA}$  is activated. This neuron will participate in the following competitions.
- The update of training set  $P_A$ : The neuron selected in the previous operation is added to the set  $P_A$ .
- The reduction of set  $P_A$ : We put of  $P_A$  the neuron which has won smallest number of data. Before leaving the process, the neurons must participate in a sufficiently number of competitions. The data represented by this neuron are affected to suitable a neuron.

The proposed method can be translated by the following algorithm:

**Algorithm with selection phase:**

- /\*Convergence parameters\*/
1. Fix  $t_{max}$  and  $\epsilon$ ;
  - /\*Initial weights and initial partition\*/
  2. Initialization phase:
    - 1.1. Initialize randomly the weights vectors  
 $w_1, \dots, w_N$ ;
    - 1.2. Divide the topological map into two subsets  $P_A$  and  $P_{NA}$ .
  - /\*Applying the classical Kohonen algorithm for N iteration\*/
  3. For  $t = 1$  to  $N$ 
    - 3.1. Allocation phase for the neurons of the set  $P_A$ ;
    - 3.2. Minimization phase for the neurons of the set  $P_A$ ;
  - /\*From the N+1 iteration on, the selection phase is integrated\*/
  4. Selection phase:
    - 4.1 Do while  $t \leq t_{max}$  and  $\max_{i \in P_A(t)} \|w^i(t-1) - w^i(t)\| < \epsilon$
    - 4.2. Allocation phase for  $P_A$ ;
    - 4.3. Competition phase for  $P_A$ ;

/\*The selection phase\*/

4.4 Calculate the following quantity for every neuron  $j$  from the set  $P_{NA}$  :

/\*Selection function\*/

$$s(j) = \sum_{y \in P_A} \sum_{i \in P_A} e^{-\frac{\delta(i,j)}{\tau}} \|y - w^j\|^2 \quad (3)$$

4.4.1. Determine from  $P_{NA}$  the neurons  $p$  which will participate in the future competition using the following equation:

/\*Selection phase criteria\*/

4.4.2. Determine the neurons  $h$  of the set  $P_A$  which represents the smallest number of data, and which have participated in at least  $[N/2]$  competitions.

4.4.3. Replacing the set  $P_A$  by the set:  $P_A = P_A \cup \{p \text{ fulfilled the selection criteria}\} \setminus \{h\}$ ;

4.4.3. Affect the data represented by the neuron  $h$  to the neuron  $q$  which fulfilled the following condition:

- a)  $q$  is a neighbor of the neuron  $h$ ;
- b)  $q$  represents a largest number of data;

4.4.3. Return to the steep 4.1;

In this algorithm, the positive real  $\epsilon$  represents the threshold. The concept of neighbourhoods can be introduced between  $l$  and  $k$  by the positive and symmetrical functions  $\beta_{l,k}$  that defined as follows:

$$\beta_{l,k} = \exp\left(\frac{\|r_l - r_k\|}{(\sigma(t))^2}\right) \quad (4)$$

Where  $\|r_l - r_k\| \approx \|w^l - w^k\|$  and the vector  $r_i$  represents the coordinates of the weights and  $\sigma$  functions are decreasing.

The main important parameter in our method is the selection set. In this context, the choice of this latter must be doing carefully taking into a count the learning quality and a small complexity of the proposed method. In addition, it is possible to active or to disable many neurons in each iteration. Finally, various versions of this technique can be considered by varying either the number of those neurone or the initial distribution of the topological map or by choosing another selection function.

**5. Computational experiments**

In order to assess the advantages of our method, we apply the proposed algorithm to a widely used dataset: Iris dataset for classification [12]. It consists of three target classes: Iris Setosa, Iris Virginica and Iris Versicolor. Each species contains 50 data samples. Each sample has four real-valued features: sepal length, sepal width, and petal length and petal width. Before training, data are normalized [1]. The half of the data samples are used for training, 75 items for testing. The selection set is formed by 10 items from each class. We aim by this choice to built a map basing on the tree kind of data and to avoid the complexity of the proposed algorithm .

In order to study the effect of the choice of those elements on the results, several experiences are carried out varying in each of them the number of inputs of each class. To cluster data IRIS, in the beginning we initialize the map by a random number of neurons. Then this number varies during the during the conduct of the algorithm. This variation is controlled by the objective function defined above. Indeed the unnecessary units are dropped from the map and replaced by the others which can improve the performance.

The tables 1 and 3 presents the results provided by the classical kohonen algorithm and the ours. In order to compare those algorithms they are applied to the same dataset. The performance of the algorithms is evaluated by measuring the quality of the classification using the misclassified error.. This later is defined as a percentage of the incorrectly classified object.

Table 1: Numerical results for clustering the testing data with algorithm SOM method

	ND	MSC	CC	A.TS %
Setosa	75	0	75	100

Virginica	75	0	75	100
Versicolor	75	8	67	89.33
Total	225	8	217	96.44

- ND: is number of data set
- MSD :Misclassified for testing set
- CC: is correctly classified
- A.TS: Accuracy for training set
- IT: Number for iteration

The results presented in this table shows that using the classical kohonen algorithm, 8 inputs among 225 are wrongly classified. We also note that all these misclassified data belong to the same class . In order to investigate the effect of the number of iteration on quality of classification, we carry out the following experience:

In the following experience, we perform a number of iterations, after each twenty one, we measure the quality of the resulted classification. The results of this experience are stored in table 2.

Table 2: The variation of the misclassified error

It	A.TS%
20	60
40	73.33
60	73.33
80	76
100	76
120	80
140	84
160	86,66
180	96,33
200	96,33
220	96,88
240	96,88

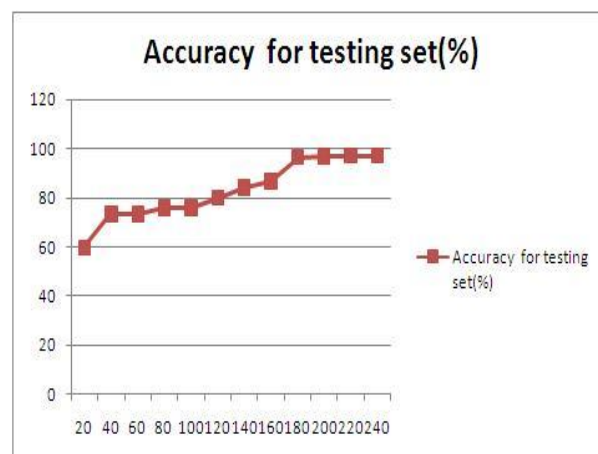


Fig.4 Curve presenting the variation of the misclassification error according to the number of iteration.

Both of table 2 and the fig. 3 shows that after a growth phase the performance of the system stagnates in certain value.

Table 3: Numerical results for clustering the testing data with algorithm SOM select method

	Nbr. Data	Mi-sclassified	Cr. Class	A.TS%	It
Setosa	75	0	75	100	100
Virginica	75	2	73	97.33	100
Versicolor	75	5	70	93.33	100
Total	225	7	220	96.88	100

- Nbr.Data: is number of data set
- Misclassified for testing set
- Cr.Class: is correctly classified
- A.TS: Accuracy for training set

This table shows that our method outperform the classical one. Indeed the first provides a score better than that obtained by the second one. Therefore this score is realized using a reduced number of neurons. Thus our method can also be used as a mean of optimization the size map.

We also compare the proposed method to some others classifier systems. The table 4 presents the results obtained by those algorithms.

Table 4: Comparison for iris data classification

	<i>It</i>	<i>M.T</i>	<i>M.TS</i>	<i>A.T%</i>	<i>A.TS%</i>
EBP	500	3	2	96	97.3
EBP	800	2	1	97.3	98.6
RBF	85	4	4	94.6	94.6
RBF	111	4	2	96	97.3
Proposed method	150	4	4	94.6	94.6

- *It*: number for iteration
- *M.T*: Misclassified for training set
- *M.TS*: Misclassified for testing set
- *A.T*: Accuracy for training set
- *A.TS*: Accuracy for testing set

From the Table 2 and 4, we see that the proposed method gives satisfying results, in comparisons with the best ones as RBF and EBP. In this context, we report that this result can be improved by a suitable adjustment of the parameters used in the algorithm., Tanks to it importance , any improvement of Kohonen algorithm will be well appreciated. The obtained amelioration can be explained as follows :

- The first term of the objective function of our proposed model controls the geometric error on a given classification.
- The second term of the objective function keeps only the neurons that represent the phase select of neuron.

In addition, the number of hidden neurons must be decided before training in both EBP and RBF neural networks. Different number of hidden neurons results in different training time and training accuracy. It is still a difficult task to determine the number of hidden neurons in advance. The experiments indicated that clustering by the proposed method is computationally effective approach.

In the next section, there is the optimization of Kohonen map and resutats obtained.

The following diagram shows the steps of select automatically the neuron for each iteration.

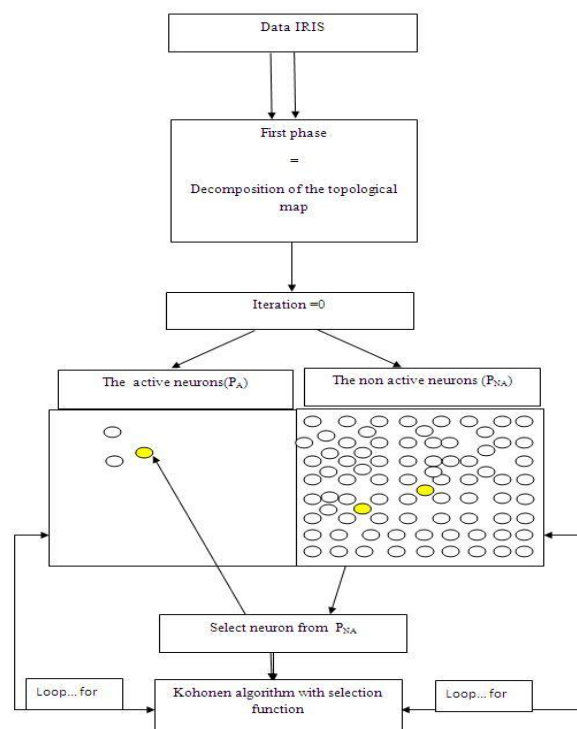


Fig. 5: schema presenting the automatic variation of the initial number of initial neuron.

Table 5: optimization of neural architecture

<i>It</i>	<i>N.T</i>	<i>N.N.O</i>	<i>P.R</i>
100	20	15	0.25
100	30	25	0.16
100	35	25	0.28
100	40	35	0.12
100	42	35	0.16
100	45	37	0.17
100	48	37	0.22



100	50	37	0.26
100	55	44	0.20
100	60	48	0.20
100	62	49	0.20
100	70	56	0.20
100	75	61	0.18
100	80	66	0.17
100	85	71	0.16
100	90	76	0.15
100	100	86	0.14

- It: number for iteration
- N.T: Number Total of neurons
- N.N.O: Number of Neurons of optimized map
- P.R Percentage of reduction

Table 5 presents the reduction of the the Kohonen map. It shows that only with 100 iterations, the proposed method lead to a reduction of the card.

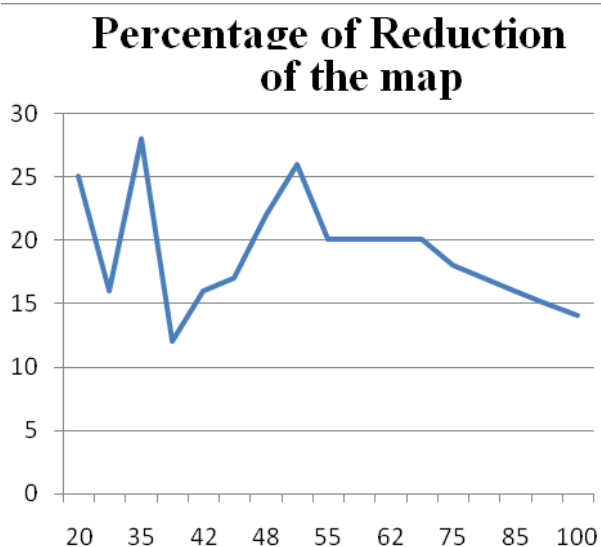


Fig. 6: Curve presents the variation of the reduction of map the size according to the number of the activated neurons .

The proposed method completes the Kohonen learning method. In fact, our approach realizes two tasks at the same time: the learning task and the classification of the objet one which consists minimizing the size of the map.

## 6. Conclusions

In this paper, we have proposed an amelioration of the classical Kohonen network and optimization its architecture map. To this end , an additional phase was integrated in the learning algorithm of kohonen network. This new phase is based on a function called a selection function. This latter permit to correct the initial choice of the Kohonen network an conserve at the same time the notion of neighborhood defined on the observation set. So the additional phase allows to remove useless neurons at each step of learning process. Using the data set Iris which is widely used in the clustering area. We have shown that our method outperform the classical one. Because it efficacy, we will use the proposed method in the image compression and speech processing. Finally we report that the result can be more improved by choosing a more suitable selection function.

## References

- [1] M. Ayache, M. Khalil and F. Tranquart. "Artificial Neural Network for Transfer Function Placental Development: DCT and DWT Approach". IJCSI International Journal of Computer Science Issues, Vol. 8, Issue 5, No 3, September 2011.
- [2] G. Dryfus, J.M.Martinez, M.Samuelides, M.B.Gordan, "Reseaux de neurons Mthodologie et applications". EYRLLES, SC 924, 2000.
- [3] M. Ettaouil, and Y.Ghanou, and K. Elmoutaouakil, M. Lazaar. "A New Architecture Optimization Model for the Kohonen Networks and Clustering", Journal of Advanced Research in Computer Science (JARCS), Volume 3, Issue 1, 2011,pp. 14 - 32.
- [4] M. Ettaouil, and Y. Ghanou, and K. El Moutaouakil, M. Lazaar, "Image Medical Compression by A new Architecture Optimization Model for the Kohonen Networks". International Journal of Computer

- Theory and Engineering vol. 3, no. 2, 2011, pp. 204-210.
- [5] M. Ettaouil, and K. Elmoutaouakil, and Y.Ghanou. "The continuous hopfield networks (CHN) for the placement of the electronic circuits problem", WSEAS Transactions on Computer, Volume 8 Issue 12, December 2009.
- [6] M. ETTAOUI and M. LAZAAR, "Improved Self-Organizing Maps and Speech Compression", International Journal of Computer Science Issues (IJCSI), Volume 9, Issue 2, No 1, pp. 197-205, 2012.
- [7] B.W. Forgey, Cluster analysis of multivariate data :efficiency versus interpretability of classification,Biometrics, 1965, 21 : 768.
- [8] A. Ghosh, and S.K. Pal, Object Background classification using Hopfield type neural networks. International Journal of Patten Recognition and artificial Intelligence, 1992, pp. 989-1008.
- [9] S. Ghorbel, M. Ben Jemaa and M. Chtourou. "Object-based Video compression using neural networks". IJCSI International Journal of Computer Science Issues, Vol. 8, Issue 4, No 1, July 2011.
- [10] J.J. Hopfield, Neurons with graded response have collective computational properties like those of two-states neurons. proceedings of the National academy of sciences of the USA 81 ,1984,pp. 3088-3092.
- [11] C. C. Hsu, Generalizing Self-Organizing Map for Categorical Data.IEEE Transactions on neural networks, Vol. 17, No. 2, 2006.,pp. 294-304.
- [12] S .Marsland, S.U .Nehmzow,, and Shapiro, J., "A self-organizing network that grows when required", in Neural Networks, Volume 15, 2002, Issue 8-9, pp.1041-1058.
- [13] N. Nasrabadi and Y. Feng, "Vector quantization of images based upon the Kohonen self-organizing feature maps", in IEEE Int. Conf. Neural Networks, San Diego, CA, vol. 1, pp. 101-108, 1988.
- [14] T. Kohonen. Self Organizing Maps. Springer, 3e edition, 2001.
- [15] H. Shah-Hosseini, and R. Safabakhsh. TASOM: The Time Adaptive Self-Organizing Map. The International Conference on Information Technology: Coding and Computing (ITCC'00), 2000, 422.
- [16] H. Yin. "ViSOMA Novel Method for Multivariate Data Projection and Structure Visualization". IEEE Transactions on Neural Networks, Vol 13, 2002, 1,pp. 237-243.
- [17] [www.ics.uci.edu/mllearn/MLRepository.html](http://www.ics.uci.edu/mllearn/MLRepository.html).
- [18] V. Selvi, R.Umarani " Comparative analysis of ant colony and Particle Swarm optimization techniques " International journal of Computer Application Volume 5-No.4, August 2010,pp.0975-8887
- [19] D. Wang, 'Fast Constructive-Covertng Algorithm for neural networks and its implement in classification', Applied Soft Computing 8 ,2008, pp. 166-173.
- [20] D. Wang, N.S. Chaudhari, 'A constructive unsupervised learning algorithm for Boolean neural networks based on multi-level geometrical expansion', Neurocomputing 57C, 2004, pp.455-461.