Toward a new approach to improve the classification accuracy of the Kohonen's Self-Organizing Map during learning process.

Mostafa Ezziyyani[(1)], El Khatir HAIMOUDI[(2)], Hanane FAKHORI[(3)]
[(1)]Faculty of Sciences and Techniques, Mathematics and Application Laboratory, Tanger, Maroc
[(2)]Faculty Poly-disciplinary of Larache, Computer Sciences Department, Larache, Maroc
ezziyyani@gmail.com , helkhatir@gmail.com, hananef@gmail.com

Abstract

Kohonen's self-organization algorithm, known as "topologic maps algorithm", has been largely used in many applications for classification. However, few theoretical studies have been proposed to improve and optimize the learning process of classification and clustering for dynamic and scalable systems taking into account the evolution of multi-parameter objects. Our objective in this paper, is to provide a new approach to improve the accuracy and quality of the classification method based on the basic advantages of the Kohonen's self-organization algorithm and on new network functions to pre-eliminate the auto-detected of drawbacks and redundancy.

I. Introduction.

The self-organizing map of Kohonen (CAD), is a model of artificial neural networks (ANN) widely used in different domains, particularly the classification and clustering of multi-parameter objects. These intelligent systems are characterized by special abilities such as learning, adaptation and the possibility of visualization of Multiparameter objects with a reduced space. Other paradigms can beings used, such as retro-propagation network (Conter-propagation) and Back-propagation, both types unlike the SOM map based on supervised learning algorithms that requires advance preparation of desired outputs. Our objective is to reveal the ambiguities and obstacles that may limit the application of this paradigm in the different domains of human activities, and find possible solutions to eliminate them. To reach the desired result, we made a preliminary theoretical study of the learning process, and perform experimental tests. On the basis of analysis of results achieved, we fixed the nature and details of the problem and proposed a solution already practiced and justified theoretically can eliminate the obstacles detected.

II. The process of classification and clustering of multiparametrics objects with the self-organizing map of Kohonen.

The artificial neural networks develop rules for a well-defined resolution, which allow classifying the multi-parametric objects (events, situations, processes). Unlike classical classification methods, the RNA used to create a biological information model of the human and other beings brain.

The model of the Kohonen network uses the competitive learning method [3], this procedure distribute the objects of the learning multitude on clusters inherent to the regrouping of the input data [4]. During learning, the neurons are competing, and the network fixed the winners neurons for each group of similar input objects. The fixed neurons form the centers of the clusters. The metric used in this operation is the Euclidean distance between the synaptic weights vectors, and input objects vectors.

A. The structure of the Kohonen neural network.

The Kohonen network is composed of two layers: the input layer and output layer – precisely SOM. The elements of the map are dispersed in a space - usually one-dimensional or two-dimensional. The inputs data are presented as a matrix, the rows are the vectors of objects and the columns are the components of these objects.

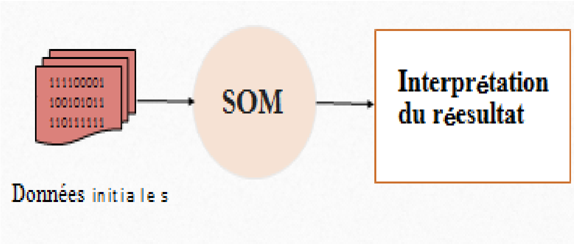

Fig.1.1. Structure of the Kohonen map.

The objects are presented to the network input one by one the idea of learning, is the application of successive rapprochement method, beginning with the random choice of the disposition of cluster centers, then the algorithm gradually improve them to perform the learning data clustering [5-6].

B. The Kohonen network learning algorithm.

The learning idea is application of successive rapprochement method, beginning with the random choice of the disposition of cluster centers, then the algorithm gradually improve them to perform the learning data clustering [5-6].

The learning procedure begins with the normalization of input data and synaptic weights to reduce the learning time [7]. This operation based on the following algebraic formula:

$$x_i = x_i / \sqrt{\sum_{j=0}^{n-1} x_j^2} \qquad (1.1)$$

Where: x_i - input object component or of vector of synaptic weights;

n - The number of variables in the vector x.

The main learning algorithm passes successively through a series of iterations. In each one the learning object vectors are presented successively to network input, and the desired output are absent. At the end of this procedure the topologically adjacent neurons, respond to similar input vectors.

To fix the winner's neurons, using the metric of the Euclidean distance [4] see formula below:

$$k : \|\mathbf{w}_k - \mathbf{x}\| \leq \|\mathbf{w}_o - \mathbf{x}\| \quad \forall o \quad (1.2)$$

Subsequently the algorithm performs a correction of synaptic weights to gradually minimize the distance between the neurons winners and input objects. For this correction using the following formula [4]:

$$w_{ij}(t+1) = w_{ij}(t) + \alpha_i(t) h(d,t) \cdot \left[y_i - w_{ij}(t) \right] \quad (1.3)$$

Where: y_i - the value of the output neuron i;

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

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

$$\alpha_i = \alpha_0 e_t^i \qquad (1.4)$$

Where: i - the iteration number;

t - the iteration rate.

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

$$h(d,t) = \begin{cases} 0, & d \geq \delta(t) \\ e^{\left(-\frac{d}{2\delta(t)}\right)}, & d < \delta(t) \end{cases} \qquad (1.5)$$

$$\delta(t) = \delta_0 e^{\left(-\frac{i}{\mu}\right)} \qquad (1.6)$$

Where: d- the distance between the winner neuron and an x neuron.

δ - Constant.

$$\mu = \frac{n}{log_{10}(\delta_0)} \qquad (1.7)$$

n – Iteration rate.

The learning process will be continued up to the stabilization of the self-organizing map.

C. The classification and clustering process with Kohonen's self-organizing map: The experiments results.

The test is performed to fix the impact of data normalization on the learning process, and the revelation of the restrictions which may limit the right application of the network in different domains.

The data used in the test are structured to form a matrix of a dimension of (8X3), whose lines present the input vector objects. Among these objects, there are two similar objects (Input 1 and 5), four objects having linear regularities between its components (Inputs 0, 4, 2 and 6) and two other normal objects (Inputs 3 and 7) see Tab.1.1.

Inputs N°	Initial data			data after normalization		
0	8	4	6	0,74	0,37	0.55
1	5	6	7	0,47	0,57	0,66
2	3	5	4	0,42	0.70	0.56
3	11	27	39	0,22	0,55	0,80
4	4	2	3	0,74	0,37	0.55
5	5	6	7	0,47	0,57	0,66
6	9	15	12	0,42	0,70	0.56
7	13	35	42	0,23	0,62	0,74

Tab.1.1 Input vectors Normalization (Matrix 8x3).

The result presented in Fig.2 show that the network designated a winner neuron for each pair of similar objects (1, 3, 5 and 7) and two different winner's neurons for both normal objects which seems normal. But for linear inputs objects (0, 2, 4 and 6) the network has fixed a single winner neuron for each pair that does not conform to our wishes.

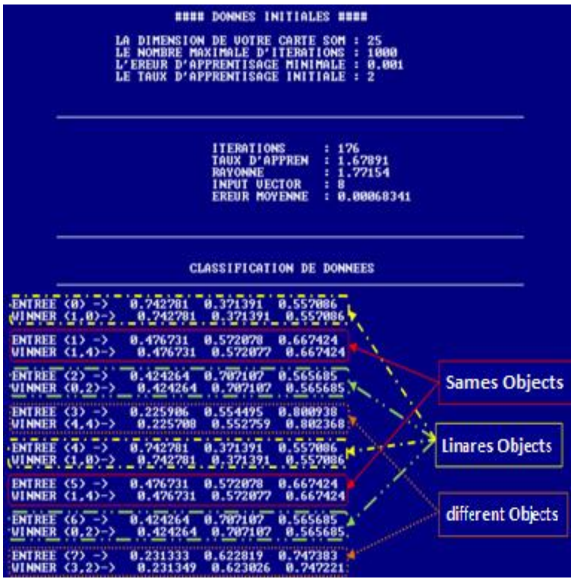

Fig.1.1 Learning results of the self-organizing map Matrix 8X3.

The theoretical analysis and experimental tests show that the possibilities of the card in the classification and Clustering task of multiparameter objects are limited by the loss of certain information about objects (the initial length of vectors). This loss is caused by the need of the input vectors normalization. So the practice shows that the speed and learning term depends on the size and complexity of the data multitude. Our perspective is to develop intelligent systems based on neural network models with competitive learning. These systems will be characterized with news capabilities to identify and classify correctly the linear or nearly linear objects.

II. The possibilities of the principal component analysis method.

The principal component analysis (PCA) is a way to identify regularities in the data, and to express data so as to highlight their similarities and differences. The main advantage of the PCA is that once these regularities are found in the data, it can eliminate them and compress the data by reducing the number of dimensions, without losing too much information. This technique can beings used to filter the data and make it suitable with the Kohonen map.

A. The algorithm of the principal component analysis method.

The principal component analysis is a way to identify patterns in the data, and expressing data so as to highlight their similarities and differences. The main advantage of the (PCA) is that once these patterns are found in the data, it can compress data by reducing the number of dimensions, without losing too much information. This technique is the basis of our method to achieve our objectives referred to in this work.

The initial data is organized as an array, whose rows are the individuals (objects) and columns present variables (Components) see figure1.

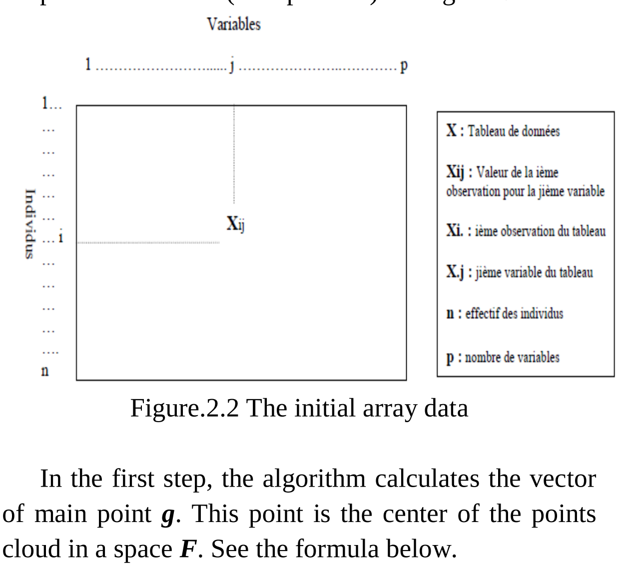

Figure.2.2 The initial array data

In the first step, the algorithm calculates the vector of main point ***g***. This point is the center of the points cloud in a space ***F***. See the formula below.

$$g^t = (\bar{x}^1, \dots, \bar{x}^p) \qquad (2.1)$$

$$\text{Où : } \bar{x}^j = \frac{1}{n}\sum_{i=1}^{n} x_i^j$$

At the base of the vector ***g*** is calculated the data centered matrix, which is written in terms of ***X*** of the following way:

$$Y = X - 1g^t \qquad (2.2)$$

Where: g^t is the transposed of *g*.

The term centered signifies that the means of the variables $\bar{y}^j$ are zero.

The centered data matrix ***Y*** is used in this step for calculating the variance-covariance matrix ***V***, which is written according to the following way of ***Y***:

$$V = \frac{1}{n} Y^t Y \qquad (2.3)$$

Where: Y^t is the transposed of ***Y***.

The V matrix is presented as follows:

$$V = \begin{pmatrix} s_1 & . & . & s_{1p} \\ s_{21} & . & . & . \\ . & . & . & . \\ s_{p1} & . & . & s_p \end{pmatrix}$$

Where: S_{kl} is the covariance of the variables *k* and *l*, and S_k is the variance of the variable *k*.

In the last step and to develop the correlation matrix ***R***, we calculate the two diagonal matrices $\boldsymbol{D_{1/S^2}}$ ***et*** $\boldsymbol{D_{1/S}}$ as a function of ***V*** for the following way.

$$D_{1/S^2} = \frac{1}{Diag(V)}$$

2.4

$$D_{1/S} = \frac{1}{Diag(D_{1/S^2})}$$

2.5

The correlation matrix ***R*** comprises the coefficients of linear correlation between ***p*** variables taken two by two. It sums and shows the structure of linear dependencies between ***p*** variables. It is symmetrical and her diagonal is composed of 1. ***R*** is written according to the following way of V:

$$R = D_{1/s} V D_{1/s}$$

2.6

Where: $D_{1/s}$ is the diagonal matrix containing $\frac{1}{s_1}, \dots\dots\dots, \frac{1}{s_p}$ on its diagonal.

After calculating the matrix ***R*** using the iterated power method, to search the Eigenvectors that constitute the rows of the resulting matrix *M* of input objects.

$$M = \begin{bmatrix} x_{11} & \cdots & x_{1l} \\ \vdots & \ddots & \vdots \\ x_{n1} & \cdots & x_{nl} \end{bmatrix}$$

III. Conclusions.

The results of this work show that the Kohonen network exceeds other means and methods used for classification and clustering multi-parameter objects, through its characteristics derived the animal brain, and the special abilities such as learning, adaptation and the possibility of visualization of multi-parameter objects with a reduced space.. This phenomenon can beings used to perform intelligent systems and apply them in various fields, but these opportunities can beings limited by restrictions detected. So these obstacles can beings removed by the preliminary data treatment at the base of the principal components analysis method, which filters and removes regularities in the data.

References

[1]. McCulloch W.S and W. Pitts, A logical calculus of the ideas immanent in nervous activity, Bull. Math. Biophys., 5, 115-133 (1943).

[2]. Ripley B.D. Pattern recognition and neural networks. – Cambridge: Cambridge University Press, 1996. – P403.

[3]. Kohonen T. Self-organizing maps (2nd edition). – Springer Verlag, 1997. – P 448.

[4]. Krose B., van der Smagt P. An introduction to neural networks. – The University of Amsterdam. – 1996. – P 135.

[5]. Anil K., Jain, Jianchang Mao, K.M. Mohiuddin. Introduction to Artificial Neural Networks // Open systems. - 1997. - №4. - P 16-24.

[6]. EZHOV A.A., SHUMSKY S.A., Neurocomputing et son application dans l'économie et les affaires: Proc. Allocation, 1998, P 222

[7]. P D. Wasserman. Neural Computing: Theory and Practice. ANZA Research, Inc. 1989. P 64.

[8]. Jérôme Pagès, Analyse factorielle multiple avec R, EDP sciences, Paris, 2013, 253 p.

[9]. Catherine Bolley. Analyse numerique. _Ecole d'ingenieur. Nantes, France. 2012, p.97.

[10]. Michel Volle, Analyse des données, Economica, 4e édition, 1997.

[11]. R.Christian ;S.A. Yvan Mathématiques et technologie. Springer Science+ Bisness 2008. P 431.

[12]. M. Mc Cord Nelson, W.T. Illingworth A practical guide to neural nets, Addison-Wesley Publishing ompany 1991.

[13]. E. Davalo, P. Naïm, Des Réseaux de neurones, Edition Eyrolles 1993.

[14]. A.Deweze, L'accès en ligne aux bases documentaire, Collection MASSON 1983.

[15]. J. A. Farrel, A. N. Michel Associative memory via artificial neural networks IEEE control system magazine 1990.

[16]. J. A. Freeman, D. Skapma, Neural Networks, Algorithms, and programming technique, Addison-Wesley publishing company 1992

[17]. C.Guinchat, Y.Skouse, Guide pratique des techniques documentaire, Vol 1, 2, EDICEF 1989.

[18]. D. O. Hebb The organization of behavior J. Wiley & Sons NY, 1949 [HER 94] : J Herault, C Jutten, réseaux neurones et traitement de signal, Edition HERMES 1994.

[19]. C. L. Jimenez Guaim, Opération d'accès par le contenue à une base de document textuelle, Thèse de doctorat INPG 1989.

[20]. Kohonen, Self-Organized Formation of topologically correct features maps, Biological Cybernetics 1988.

[21]. T. Kohonen, Self-Organising Maps, Edition Springer, second édition 1997.

[22]. C. Leloup, Moteur d'indexation et de recherche, Edition Eyrolles, 1998.

[23]. Xia lin, D Soergel et G Marchionini, A self-organising semantic map for information retrieval, Int AMC/SIGIR Conference on R&D information retrieval, pp 262-269, 1991

[24]. P. Martin Réseaux de neurones artificiels Application à la reconnaissance optique de partitions musicales Thèse de doctorat université de J. F. Grenoble I 1992.

[25]. P.Poincot, Création d'une carte bibliographique à l'aide des cartes autoorganisatrice de Kohonen, congres INFOSID'97, Toulouse, 10-13 Juin 1997, pp 625-641.

[26]. G. N. Reddy, Artificial neural networks Theory, Computer simulation programs and Applications Dept. of Electrical Engineering Lamar University Beaumont, Texas, USA Nov. 1995.