

M-SOM-ART: Growing Self Organizing Map for sequences clustering and classification

Farida Zehraoui and Younès Bennani¹

Abstract. This paper presents a new growing neural network for sequences clustering and classification. This network is a self organizing map (SOM), which has the properties of stability and plasticity. The stability concerns the preservation of previously learned knowledge and the plasticity concerns the adaptation to any change in the input environment. These properties are obtained using Adaptive Resonance Theory. In order to take into account the temporal information (the dynamics) and the correlation of the patterns contained in the sequences, the inputs of the map are modelled using their associated dynamic covariance matrices. This new model is inspired from the field of speaker recognition. We have modified a covariance matrix in order to represent a temporal order in the sequence. The experimentations show that our approach is better than some other temporal self organizing map for user’s Web navigation classification.

1 Introduction

This work reports on the problem of sequences classification and clustering using artificial neural networks. Our study is motivated by an application, which consists in predicting user behaviour in an e-commerce Web site. Past experiences are provided from the site users’ Web log files. This application is characterized by the big amount of data, the temporal aspect of these data, the presence of noise and the real time constraints (we must do the processing before the end of the user navigation in the site). In addition, priory field knowledge is not available.

In this paper we propose a new approach, which uses a growing self organising map for sequences clustering and classification.

Growing network models have no pre-defined structures, they are generated by successive additions (and possibly deletions) of elements. For these networks, suitable insertion strategies have to be defined as well as criteria to eventually stop the growth. Most of these networks have the plasticity property because they are adapted to the changes in the input environment, but they have not the stability property. When new inputs are presented to these networks, the old ones can be forgotten. The conflict between stability and plasticity is called the stability-plasticity dilemma. Most of existing algorithms are either stable but not capable of forming new clusters, or plastic but not stable. The Adaptive Resonance Theory ”ART” was specifically designed to overcome the stability-plasticity dilemma [4].

The majority of neural models deals with the learning of static mappings. However, real-world data are usually dynamic, so that a particular pattern cannot be assumed to be independent of its antecedents.

Processing these dynamic patterns differs fundamentally from processing static ones because the temporal order and correlation of the patterns being observed must be taken into account [5].

We present a new approach for sequences classification and clustering using self organizing map. In this approach, all the antecedents of each pattern are taken into account. The position and the “shape” of the cloud of points that represents the sequence are modelled at the input of the SOM network. The problem of the variable input length is overcome since all the inputs models have the same dimension. This model is inspired from methods developed for the speaker recognition task. We propose a new model, which introduces the dynamics in the covariance matrix associated to the input sequence. In our approach, the SOM is integrated in the ART paradigm. We obtain a growing self organising map, which can learn new knowledge without forgetting the previous learned ones.

The paper is organized as follows. Section 2 addresses the sequences processing using SOM and growing SOM. Section 3 details the proposed approach. Section 4 shows our experimental results. Section 5 presents the current state of our work and future work.

2 Related work

The Self Organizing Map (SOM) [15] is among the most used connectionist models for data clustering and visualization. The SOM models comprise an important class of competitive neural models. The main difference between the SOM and standard competitive networks is that the output neurons are arranged in specific geometrical forms.

Each neuron c in the SOM is associated with a weight vector $w_c = [w_{c_1}, w_{c_2}, \dots, w_{c_n}]^T \in \mathbb{R}^n$ with the same dimension as the input vector $\xi = [\xi_1, \xi_2, \dots, \xi_n]^T \in \mathbb{R}^n$. Through an unsupervised learning process, the output neurons become tuned and organized after several presentations of the data. The learning algorithm that leads to a self organization can be summarized in two steps:

- A winning or best-matching unit, denoted $s(\xi)$, of the map is found by using a distance or a similarity measure (Euclidian distance, for example) between the input and weight vectors:

$$s(\xi) = \arg \min_{c \in A} \text{dist}(\xi, w_r) \quad (1)$$

where A is the set of neurons and dist is the Euclidian distance

- The winner and its neighbours in the map have their weights $w_r(t)$ updated towards the current input ξ :

$$w_r(t+1) = w_r(t) + \epsilon(t)h_{rs}[\xi - w_r] \quad (2)$$

where $\epsilon(t)$ is the learning rate and h_{rs} is the neighbourhood function.

¹ LIPN, UMR 7030 du CNRS, Institut Galilée, Université Paris-Nord, Villeneuve, France. email: {fz, younes}@lipn.univ-paris13.fr

In this section, we present some work related to the use of SOM for sequences processing and others related to the stability and the plasticity properties.

Several efforts have been devoted in order to introduce the temporal information in the Kohonen map [17, 6, 13, 8, 7, 12, 14]. These maps allow taking into account the temporal and the space information of the data. A spatio-temporal sequence (or simply, temporal sequence) is a finite set of time ordered n-dimensional feature vectors. The temporal information added to the network can be external or internal.

In the external representation, when a finite sequence: $X(t) = x(1), x(2), \dots, x(p)$ of p vector input items is presented as a single input entity to SOM, each neuron in the map might be made to correspond to an operator that analyses $X(t)$.

Kangas [13] has proposed two temporal variants of the SOM to represent sequential features of the input data, which are applied to phoneme recognition tasks. In the first model, called SOM with exponentially weighted decay, the input pattern of SOM is a function of former inputs. The second model uses an external delay line model: the last T items of the input sequence are concatenated and presented to the network. This procedure requires high computational efforts, increasing the training time. However, it produces very good results in recognition tasks. Kohonen [14] has proposed the "Hypermap architecture", where different time windows centred on a particular time step t are used to construct two types of network input vectors: the context vector and the pattern vector. The context vector is first used to select a subset of nodes in the network. The best-matching neuron is then chosen on the basis of the pattern vector from this subset. This approach is used in phoneme recognition, biological sequences processing and speech recognition.

In the internal representation, the temporal information is integrated in the network at the neuron level or at the connection level.

Chappell and Taylor [6] have proposed a "Temporal Kohonen Map (TKM)" for sequences classification. The TKM maintains the activation history of each neuron by means of a variable called the leaky integrator potential. This approach succeeded in classifying a word in the same position within a set of sentences having different contexts. Euliano and Principe [8] have proposed the SOTPAR model based on the biologically inspired diffusion of activation through time over the neurons in the map and temporal decay of activation. These couplings are based on the propagation of activation waves starting at each winning neuron and decaying as time goes by.

Several growing self organizing maps are proposed in the literature. Some models like the growing cell structure (GCS) model [9], the growing neural gas (GNG) model [11] and the growing grid (GG) method [10] have several common properties. Their network structure is a graph of nodes and edges connecting nodes (hyper tetrahedrons structure for the GCS, hyper-rectangular structure for the GG and no explicit constraints on the GNG graph). At each adaptation step local error information is accumulated at the winning unit. The accumulated error information is used to determine (after a fixed number of adaptation steps) where to insert new units in the network. Error variables are locally re-distributed and another adaptation steps are performed. Some units and edges are also deleted. Other insertion criteria can be used in growing self organizing maps. For example, in the incremental model (GSOM) [2], instead of using accumulated error information, authors always insert units near the center of the current topology.

These models can learn new knowledge but don't "assure" the preservation of the old ones. In order to build a system which have the stability and plasticity properties, Carpenter and Grossberg [4] have proposed the adaptive resonance theory "ART". An ART system con-

sists of two subsystems: an attentional subsystem and an orienting subsystem. The orienting subsystem selects among processing units, candidated by the attentional subsystem for being resonant, those that match external input. To select these units, the orienting subsystem employs a vigilance test combined with a mismatch reset condition and search process.

The first and the most basic architecture of ART model is the ART1 network [4]. It was designed to learn and recognize binary input patterns. In this model, when the best-matching template, selected according to a match function, does not satisfy the vigilance test (using an activation function), a search process is required to detect the resonance domain.

In order to overcome ART1 limitations, A. Baraldi and E. Alpaydm [1] have proposed a new class of ART processing schemes called Simplified ART (SART) capable of processing real-valued multidimensional patterns. In this model, the match and the activation functions are chosen such that a new processing unit can be immediately allocated to match the input pattern when the vigilance test is not satisfied.

3 Our Approach

We propose a growing self organizing map "M-SOM-ART", for sequences classification and clustering, using Simplified ART paradigm and inspired from the field of speaker recognition (see figure 1).

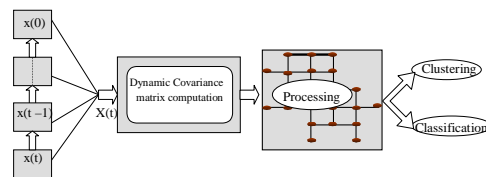


Figure 1. M-SOM-ART architecture

A sequence X is defined as a finite set of ordered n-dimensional feature vectors $x(i) \in R^n$ ($1 \leq i \leq p_X$). In our approach, the temporal information is represented externally: before using the SOM map, the input sequence is modelled using its associated dynamic covariance matrix [19] $COV_X \in R^n \times R^n$ defined as follows:

$$COV_X = \frac{1}{p_X} [x(1)x(1)^T + \sum_{i=2}^{p_X} (x(i) - \bar{x}(i))(x(i) - \bar{x}(i))^T] \quad (3)$$

where $\bar{x}(t) = \frac{1}{t} \sum_{i=1}^t x(i)$ ($t \geq 2$) is the dynamic mean vector associated to $x(t) \in R^n$ in the sequence and computed using the precedents and the current vectors $\{x(i)\}$, ($1 \leq i \leq t$).

This model allows representing the position (because the mean vector is introduced in the computation of the covariance matrix) and the shape of the cloud of points representing the sequence. We have introduced a dynamic mean vector in order to take into account the order of the vectors in the sequence. All sequences models have the same dimensions.

In order to take into account the modelling of inputs as matrices, we have generalized the Kohonen map, which has vectors as inputs.

The distance between a covariance matrix

$$COV_X = (x_{ij}), 1 \leq i, j \leq n$$

and neuron weights

$$W_c = (w_{ij}^c), 1 \leq i, j \leq n$$

is the Frobenius matricial distance (fd) given by:

$$fd(COV_X, W_c) = [tr(COV_X - W_c)^T(COV_X - W_c)]^{1/2} = [\sum_i \sum_j (x_{ij} - w_{ij}^c)^2]^{1/2} \quad (4)$$

where $tr(M)$ is the trace of the matrix M .

This distance is close to the Euclidian distance between vectors.

The M-SOM-ART is first initialised with one neuron then new neurons are added to the map following the SART paradigm using the vigilance test to select the winner. If the vigilance constraint is satisfied, the map is updated; else a neuron is added to the map. Instead of using an error measure to add the neuron, like in most growing self organizing maps, the closest neuron to the input in the perimeter of the map is determined and a new neuron is added in its neighbourhood (with respecting the square neighbourhood structure of the map). The corresponding connections are also added. At the last time of the learning process, the neurons are labelled using the labels of the inputs that have activated them.

In this approach, the proposed SOM map performs clustering like in SOM. The classification is done by labelling the neurons of the map: at the last time of learning phase, each neuron is labelled using the majority vote on the input labels that activate them. If one neuron is not activated, the majority vote is done on the labels of its closer neighbours.

The learning algorithm is described in Algorithm 3.1.

Table 1 shows M-SOM-ART properties compared to other SOM and SART models properties. The symbol '+' means that the model or some of its variants have the property, '-' means that it does not have the property.

	SOM-ART	SART	SOM	Growing SOM
Stability	+	+	-	-
Plasticity	+	+	-	+
Classification	+	-	+	+
Clustering	+	+	+	+
Visualization	+	-	+	+
Sequences processing	+	-	+	-

Table 1. Comparison of M-SOM-ART and other neural networks properties.

4 Experiments and results

We have performed several experiments on log files of an e-commerce Web site, where approximately 3000 navigations are registered every day. More precisely, the behaviour of each site user is described by the information about the succession of pages that he has visited. This succession represents the temporal aspect of the data. Since the log file contains noise, different processing are done before the use of M-SOM-ART (see Figure 2).

Algorithm 3.1 M-SOM-ART learning algorithm.

1. Initialise the set A of neurons to contain one unit c_1 .
Initialise the time parameter t : $t = 0$.
Initialise the connection set C , $C \subset A \times A$ to the empty set: $C = \emptyset$.
2. Present an input signal $X = (x(i))$ ($1 \leq i \leq p_X$), $x(i) \in R^n$
3. Model the input X by its associated dynamic covariance matrix:

$$COV_X = \frac{1}{p_X} [x(1)x(1)^T + \sum_{i=2}^{p_X} (x(i) - \bar{x}(i))(x(i) - \bar{x}(i))^T]$$

Initialise the reference vector W_{c_1} of the first unit c_1 to the dynamic covariance matrix of the first input signal.

4. Determine the winner $s(X) \in A$ by:

$$s(X) = \arg \min_{c \in A} dist(COV_X, W_r)$$

where $dist$ is the Frobenius distance.

5. Subject the winner $s(X)$ to the vigilance test:
- **If** $(\frac{1}{1+dist(cov_x, s(X))}) \geq \rho$: $\rho \in [0, 1]$ is the vigilance parameter, adapt each unit c_r according to:

$$\Delta W_{c_r} = \epsilon(t) h_{r,s} [COV_X - W_{c_r}]$$

$\epsilon(t)$ is the learning rate, σ is the standard deviation of the Gaussian, σ_i and $h_{r,s}$ is the neighbourhood function.

- **else** add a new unit c_r in the neighbourhood of the closest neuron to the input in the perimeter of the map, initialise its reference vector W_{c_r} by: $W_{c_r} = COV_X$.

6. Increase the time parameter t : $t = t + 1$
 7. If $(t \leq t_{max})$, continue with step 2.
If $(t = t_{max})$ label each unit using a majority vote on the labels of the inputs that have activated these units.
-

The data are first filtered in order to remove a part of noise. Then, they are coded using quasi-behavioural matrices [18]. The principle of this coding is to calculate for each page its frequency of precedence and succession over all the other pages and to regroup these frequencies in one matrix. Data are then processed using SOM network in order to obtain an evolution space. The navigation pages are then represented by the weights of the map neurons. The navigation is represented by the succession of these vectors (corresponding to the succession of pages). Navigation lengths are variable.

The goal of the experiments is to cluster the navigations and then to classify the site user in one of the two classes {buyer, non-buyer}, this is done using M-SOM-ART.

We can view our approach as a hierarchical SOM. The first SOM provides an evolution space and the second one performs the classification and the clustering of the sequences.

In order to evaluate our approach (M-SOM-ART), we have used the following metrics:

- **Total result:** represents the ratio of the number of correct classifications to the number of all classifications ($\frac{\text{number_correct_classifications}}{\text{number_classifications}}$). This result is given in a 95% confidence interval.
- **Confusion matrix**
- **Sensitivity:** represents the rate of the correct classifications in the buyer's class.

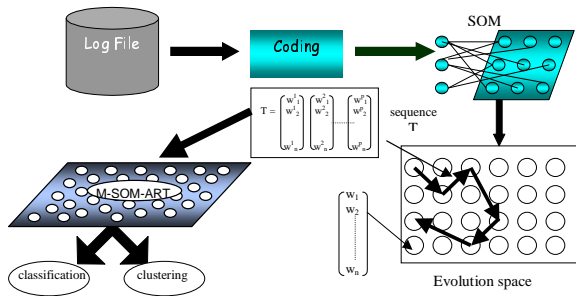


Figure 2. Data processing

- *Specificity*: represents the rate of the correct classifications in the non-buyer's class.
- *Receiver Operating Characteristics (ROC) space*²: is a graphical representation of the trade off between the correct buyer's classifications and the non-correct non-buyer's classifications with (1-specificity) on the X axis and (sensitivity) on the Y-axis. Each point represents one classifier. The best classifier is the closest to the western north.
- *Number of neurons*: This number has a direct influence on the data processing time. This processing time decreases when the number of neurons decreases.

Results

We have first compared M-SOM-ART classification results to those of other SOM models. These models are SOTPAR (SOTPAR) that uses an internal representation of the temporal information, the Kohonen map (SOM), which does not take into account the temporal aspect of data and SOM with exponentially weighted decay (SOMTemp) that uses an external representation of the temporal information.

In our approach each sequence is represented by a matrix, the classification is obtained directly by the map. In the other approaches, the classification is done for each vector in the sequence. The classification of the whole sequence is obtained by the majority vote on the results obtained for the classification of each vector in the sequence³. We have compared the results obtained by the different maps for the classification of navigations. We have used two data sets: a learning base, which contains 7000 navigations and a test base, which contains 3000 navigations. In these data sets, there are less than 8% of buyers.

The results of the first experimentation (see the table in figure 3) show that our approach gives better total result and sensitivity than SOM, SOTPAR and SOMTemp.

Apart SOM, which is not adapted to process temporal input, the specificity in all approaches are close. All temporal approaches succeed in the classification of the most frequent class (the non-buyers sequences represent more than 92% of the sequences in the data sets), but M-SOM-ART gives better results in the classification of the rare sequences (the buyer's sequences).

The ROC space (see Figure 4) shows that M-SOM-ART gives better classification results than the other approaches (its associated classifier is the closest to the western north: the Euclidian distances to this

² The figure scales are chosen for better seeing the classifiers.

³ If we find the same number of buyer and non-buyer classes in a sequence, we choose the rare class (the buyer sequences represent less than 8% of the sequences in the data sets).

	Total result	Confusion matrix	Sensitivity	Specificity
SOM	[94.54%, 96.05%]	$\begin{pmatrix} 109 & 28 \\ 111 & 2752 \end{pmatrix}$	0.495	0.989
SOTPAR	[95.48%, 96.85%]	$\begin{pmatrix} 114 & 07 \\ 106 & 2773 \end{pmatrix}$	0.518	0.997
SOMTemp	[97.27%, 98.31%]	$\begin{pmatrix} 160 & 04 \\ 60 & 2776 \end{pmatrix}$	0.727	0.998
M-SOM-ART	[99.26%, 99.75%]	$\begin{pmatrix} 215 & 08 \\ 05 & 2772 \end{pmatrix}$	0.977	0.997

Figure 3. Sequences classification results

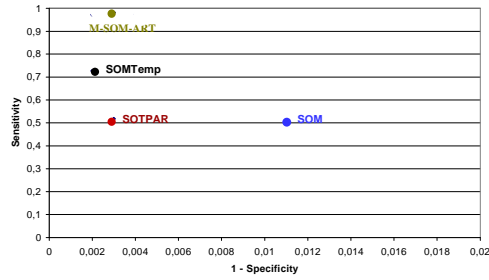


Figure 4. ROC space: Comparison of sequences classification results

point can be easily computed).

To evaluate the plasticity and the stability of our model, we have

Vigilance parameter	M-SOM-ART				M-SOM
	0.85	0.9	0.95	0.97	--
Number of neurons	42	114	154	289	416
Total result	[91.68%, 93.55%]	[94.51%, 96.03%]	[98.03%, 98.90%]	[99.26%, 99.75%]	[99.30%, 99.77%]
Confusion matrix	$\begin{pmatrix} 00 & 00 \\ 220 & 2780 \end{pmatrix}$	$\begin{pmatrix} 206 & 126 \\ 14 & 2654 \end{pmatrix}$	$\begin{pmatrix} 201 & 25 \\ 19 & 2755 \end{pmatrix}$	$\begin{pmatrix} 215 & 08 \\ 05 & 2772 \end{pmatrix}$	$\begin{pmatrix} 214 & 06 \\ 06 & 2774 \end{pmatrix}$
Sensitivity	0	0.936	0.913	0.977	0.972
Specificity	1	0.954	0.991	0.997	0.998

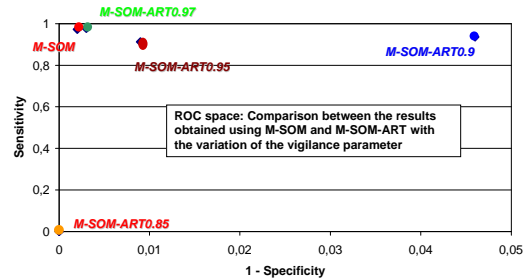


Figure 5. M-SOM-ART and M-SOM results comparison

compared it to the M-SOM model, which uses the same sequences modelling that M-SOM-ART without using the ART paradigm.

In the second experiments, we have computed the M-SOM-ART results for different values of the vigilance parameter. The results using some of these values are shown in Figure 5.

Figure 5 shows that the increasing of the vigilance parameter pro-

duces the increasing of the number of neurons and the total result. We obtain close results to those of M-SOM for the value 0.97 of the vigilance parameter with fewer neurons than SOM. This means that we can obtain close results to those of M-SOM in less time (this is needed for the time constraints of our application). The ROC space (Figure 5) shows that the results obtained using M-SOM and M-SOM-ART for the value 0.97 of the vigilance parameter⁴ are the best ones because the two classifiers are the closest to the north western. In the third experiments, we have partitioned the initial learning base into four small bases to test the stability and the plasticity of M-SOM-ART. These bases are used successively in the learning process of the two models. Figure 6 shows that the results obtained using M-SOM-ART after the partition of the learning base are better than those obtained using the hole learning base but the M-SOM results have decreased. This shows that M-SOM-ART can learn new knowledge without forgetting the old ones.

	M-SOM-ART	M-SOM
Total result	[99,38%, 99,81%]	[92,09%, 93,91%]
Confusion matrix	$\begin{pmatrix} 216 & 06 \\ 04 & 2774 \end{pmatrix}$	$\begin{pmatrix} 21 & 09 \\ 199 & 2771 \end{pmatrix}$
Sensitivity	0,998	0,09
Specificity	0,982	0,996

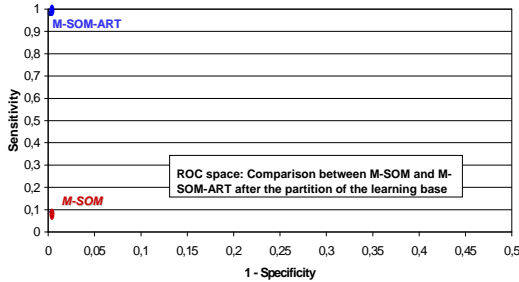


Figure 6. M-SOM-ART and M-SOM results comparison after the partition of the learning base

5 Conclusion and future work

We have proposed a growing temporal SOM, which models the sequences using covariance matrices. This modelling is used in the field of speaker identification but was never associated to the SOM maps. We have also proposed a new way to introduce the dynamics in the covariance matrix and we have incorporated the stability and the plasticity properties in SOM. We have obtained good results for the classification of a Web site's users. More experiments will be done using other data sets with a comparison to other SOM based algorithms. The results concerning the topology preservation and the data visualisation will be presented. We are also interested in extending M-SOM-ART to predict future events.

In M-SOM-ART, the inputs are modelled by covariance matrices. We have proposed a generalization of the SOM where the inputs and the neuron weights are represented by matrices instead of vectors.

⁴ When the value of the vigilance parameter is greater than 0.97, the number of neurons increases without a significant improvement of the results. For this, we have chosen the value 0.97.

We have chosen the Frobenius distance, which is close to Euclidian distance, to compare matrices. More suitable measures can be used to compare two covariance matrices. These measures are used in the field of speaker recognition and are based on the sphericity measure [3]. Future work reports on the use of an Auto-Regressive (AR) model [16], also used in the field of speaker recognition. It takes into account the order of the vectors in the sequences by using lagged matrices. This approach is considered as an efficient way to extract dynamic speaker characteristics.

REFERENCES

- [1] A. Baraldi and E. Alpaydin, 'Simplified ART: A new class of ART algorithms', Technical Report TR-98-004, Berkeley, CA, (1998).
- [2] H.-U. Bauer and T. Villmann, 'Growing a hypercubical output space in a self-organizing feature map', *IEEE Transactions on Neural Networks*, **8**(2), 218–226, (March 1997).
- [3] F. Bimbot and L. Mathan, 'Text-free speaker recognition using an arithmetic harmonic sphericity measure', in *Eurospeech'93*, volume 1, pp. 169–172, Berlin, Germany, (1993).
- [4] G.A. Carpenter and S. Grossberg, 'The art of adaptive pattern recognition by a self-organizing neural network', *IEEE Computer*, **21**(3), 77–88, (march 1988).
- [5] J.C. Chappelier, M. Gori, and Alain Grumbach, 'Time in connexionist models', in *Sequence Learning: Paradigms, Algorithms and Applications*, eds., R. Sun and G. Lee Giles, 105–134, Springer-Verlag, (November 2000).
- [6] G.J. Chappell and J.G. Taylor, 'The temporal kohonen map', *Neural Networks*, **6**, 441–445, (1993).
- [7] S. Durand and F. Alexandre, 'Spatio-temporal mask learning: Application to speech recognition', in *Proceedings International Conference on Artificial Neural Networks and Genetic Algorithms*, Alès, (Avril 1995).
- [8] N.R. Euliano and J.C. Principe, 'Spatio-temporal self-organising feature maps', in *Proceedings of the ICNN '96*, volume 1, pp. 1900–1905, Washington DC, (June 1996).
- [9] B. Fritzke, 'Growing cell structures – a self-organizing network for unsupervised and supervised learning', TR-93-026, International Computer Science Institute, Berkeley, (1993).
- [10] B. Fritzke, 'Growing grid - a self-organizing network with constant neighborhood range and adaptation strength', *Neural Processing Letters*, **2**(5), 9–13, (1995).
- [11] B. Fritzke, 'A growing neural gas network learns topologies', in *Advances in Neural Information Processing Systems*, eds., In G. Tesauo, D. S. Touretzky, and T. K. Leen, volume 7, pp. 625–632, Cambridge MA, (1995).
- [12] D.L. James and R. Miikkulainen, 'Sardnet: A self-organising feature map for sequences', in *Advances in Neural Processing Systems*, eds., D.S. Touretzky G. Tesauo and T.K. Leen, volume 7, pp. 577–584, Cambridge, (1995). MA: MIT Press.
- [13] J. Kangas, 'Phoneme recognition using time-dependent versions of self organising maps', in *Speech and Processing (ICASSP'91)*, ed., Proc. IEEE Int. Conf. on Acoustics, volume 1, pp. 101–104, (1991).
- [14] T. Kohonen, 'The hypermap architecture', in *Artificial Neural Network*, eds., O. Simula T. Kohonen, K. Makiasara and J. Kangas, volume 1, pp. 1357–1360, Amstamd, Netherlands: North-Holland, (1992).
- [15] T. Kohonen, *Self-Organizing Maps*, volume 30 of *Springer Series in Information Sciences*, Springer, Berlin, Heidelberg, 1995. (Second Extended Edition 1997).
- [16] I. Magrin-Chagnolleau, J. Wilke, and F. Bimbot, 'A further investigation on AR-vector models for text-independent speaker identification', in *Proc. ICASSP '96*, pp. 101–104, Atlanta, GA, (1996).
- [17] A. Waibel et al., 'Phoneme recognition using time-delay neural networks', *IEEE, Trans. Acoustics, Speech and Signal Processing*, (March 1989).
- [18] A. Zeboulon, Y. Bennani, and K. Benabdeslem, 'Hybrid connectionist approach for knowledge discovery from web navigation patterns', in *Proceedings of ACS/IEEE International Conference on Computer Systems and Applications*, Tunisia, (July 2003).
- [19] F. Zehraoui and Y. Bennani, 'M-som: Matricial self organizing map for sequences clustering and classification.', in *International Joint Conference on Neural Networks (IJCNN 2004)*, Budapest, Hungary, (july 2004).