Proceedings of the 6th WSEAS Int. Conf. on NEURAL NETWORKS, Lisbon, Portugal, June 16-18, 2005 (pp84-89)

The LNFC: Labeled Neuro-Fuzzy Classifier

M. NEMISSI[1], H SERIDI[1,2], H. AKDAG[2,3]
[1]LAIG, Université de Guelma, B. P. 401, Guelma 24000, Algérie
[2] CReSTIC-LERI, Université de Reims Champagne Ardenne, France
[3]LIP6 Université P. & M. Curie, Paris, France

Abstract: This paper presents a model of Neuro-Fuzzy classification, which its conception is inspired from the labeled classification using Neural Networks. This last aims to improve the classification performances and to accelerate the training of the used classifier. It is based on the addition of a set of labels to all training examples. Tests will be then carried out with each of these labels to classify a new example. The advantage of this approach is the simplicity of its implementation, which does not require modification of the training algorithm. The proposed model is based on the use of this method with the NFC (Neuro Fuzzy Classifier). To appreciate its performances, tests are carried out on the Iris and human thigh data basis by the NFC with and without labels.

Words keys: Neural Networks, Neuro-fuzzy, classification, pattern recognition, supervised training

1 Introduction

Neural Networks are successfully used in several applications of the Pattern Recognition as regard to their big capacities of learning, generalizing and parallel computing. But among their disadvantage is the slowness of their learning. Thus, the labeled classification aims essentially to accelerate this phase. In addition to the simplicity of its implementation, this method is adaptable with different approach of acceleration and stabilization such as the adjustment of the training rate [1] [2] and the momentum [3] in the case of the MLP (i.e. Multi Layered Perceptron). The application of the labeled classification with the MLP permits to exploit its properties and to compensate its inconveniences [4] as it is showed on the table 1:

Exploited proprieties	**Compensates inconvenient**
Its output can be considered as posterior probability	Local minima
Fastness classification	Slowness of training

Table 1: Proprieties of labelled classification using the MLP

Based on works [4][5] that describe the Labeled Classification using ANN: The MLP and the RVFLNN (Random Vector Functional Link Neural Network) [6][7], a new model of the labeled classification based on the use of the NFC [8] [9] is proposed. The goal is to exploit and improve proprieties of the Neuro-Fuzzy systems. In fact, these systems allow, not only, to combine advantages of the ANN and of FIS (i.e. Fuzzy Inference system), but also, to compensate their inconveniences as indicate in table 2:

	Advantages	**Disadvantage**
ANN	Self-adaptation Capacity of generalisation Parallel computing	Black box Lack of initialisation techniques
FIS	Possibility to use a prior knowledge	Lack of training techniques

Table 2: Advantages and disadvantages of ANN and FIS

2 Neuro-Fuzzy Classifier

2.1 Architecture

The NFC (Neuro-Fuzzy Classifier) [8] [9] has a Neuro-Fuzzy architecture predisposed to be used in the applications of the pattern recognition. The NFC presents the advantage to use a priori knowledge to initialize its parameters allowing it to begin the training of an initialization point not far away from the optimal one [9]. Moreover, parameters gotten after the training can be transformed in a structure based on the fuzzy if-then rules. Fig. 1 illustrates the Neuro - Fuzzy Classifier. This last is composed of three layers permitting to establish a classification system based on the fuzzy rules.

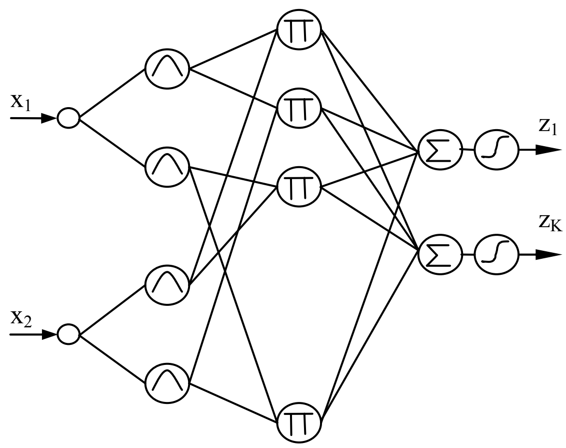

Fig.1 : Neuro-Fuzzy Classifier

In the first layer, every neuron corresponds to a linguistic term. Its outputs have the form:

$$s_i = \mu_{Aj}\left(x_n\right)$$

Where μ_{Aj} is the membership function of Aj and x_n is the n^{th} input. We use Gaussian function, so μ_{Aj} has the following form:

$$\mu_{Aj}\left(x_n\right) = \exp\left(-\left(\frac{x_n - c_{Aj}}{a_{Aj}}\right)^2\right)$$

Where c_{Aj} and a_{Aj} are the parameters corresponding to A_j

Neurons of the second layer send the product of the incoming signals. For example, the output of the first neuron has the form:

$$y_1 = \mu_{A1}(x_1) * \mu_{B1}(x_2)$$

Every neuron of the third layer corresponds to a class. The output of the k^{th} neuron is:

$$z_k = h\left(\sum_{m=1}^{M} w_{mk} \cdot y_m\right)$$

Where *h(.)* is the sigmoid function.

2.2 Training

We use the gradient descent method to adjust the NFC parameters [10]. The adaptation task is to minimize the total sum-squared error E between the classifier outputs and the target outputs. E is defined over all *(Q)* training examples and all *(K)* outputs values as:

$$E = \sum_{q=1}^{Q}\sum_{k=1}^{K}(t_k^{(q)} - z_k^{(q)})^2$$

The adaptation expressions of the weight w_{mk} at the iteration $(r+1)$ is:

$$w_{mk}^{(r+1)} = w_{mk}^{(r)} - \eta \frac{\partial E^{(r)}}{\partial w_{mk}}$$

3. The labeled classification

The labeled classification is destined to improve performances of the classifiers. It essentially based on the fact that the training is faster when classes are linearly separable. Its application is based on the addition of an additional feature (labels) for all training examples without modification of the training algorithm [4] [5]. Each class C_i corresponds to a label L_i. Then, every new example will be classified according to the following decision rule:

$$X \in C_i \ \text{ if } \ Er_i(X) = \min \{Er_1(X), Er_2(X), \ldots Er_K(X)\},$$

Where Er_i is the sum-squared error between the target output $T^{(i)}$ $(t_1\ t_2\ \ldots t_K)$ corresponding to the class C_i and the classifier output $Z^{(i)}$ $(z_1\ z_2\ \ldots z_K)$ using the label L_i. E_i is defined as:

$$Er_i = \sum_{k=1}^{K}(t_k - z_k)^2$$

The procedure of the labelled classification is shown on fig. 2 :

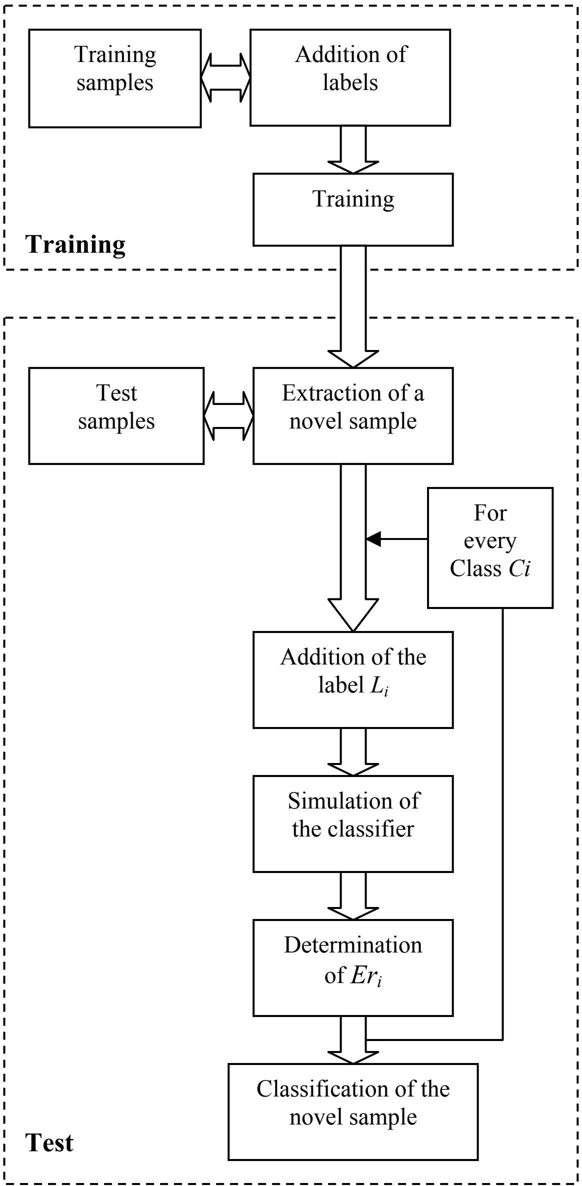

Fig. 2: Scheme of Labeled Classification

4 The LNFC (Labeled Neuro-fuzzy Classifier)

4.1 Architecture

Based on the use of the labelled classification with the NFC, we proposed a new model of Neuro-Fuzzy classification. Its conception aims to exploit and improve proprieties of the NFC. The use of LNFC leads to replace rules of the form:

If x_1 is A_1 and x_2 is B_1 then $X \in C_k$

By rules of the form:

If x_1 is A_1 and x_2 is B_1 and x_3 is L_1 then $X \in C_k$

Or by rules :

If x_1 is 'small' and x_2 is 'big' and its label is L_1 then this example belongs to C_k

The first step of the proposed method consists in adding labels to all training examples. Therefore, to add a neuron to the first layer and K neurons at the second (K is the number of classes). Every neuron added to the second layer corresponds to a membership function.

Fig. 3 shows an example of LNFC with two input variables (x_1 x_2) and two output variables (z_1 z_2). Every input is represented by two linguistic variables: A_1 and A_2 are the linguistic variables characterized by the membership functions μ_{A1} and μ_{A2}; B_1 and B_2 are characterized by μ_{B1} and μ_{B2}; L_1 and L_2 are respectively the corresponding labels to the C_1 and C_2 and that are characterized by μ_{L1} and μ_{L2}.

We choose the membership functions of labels to be Gaussian with labels as centres. Then, μ_{L1} has the following form:

$$\mu_{Li}(x_n) = \exp\left(-\left(\frac{x_n - L_i}{a_{li}}\right)^2\right)$$

Where L_i is the label and a_{L1} is the parameter corresponding to L_i

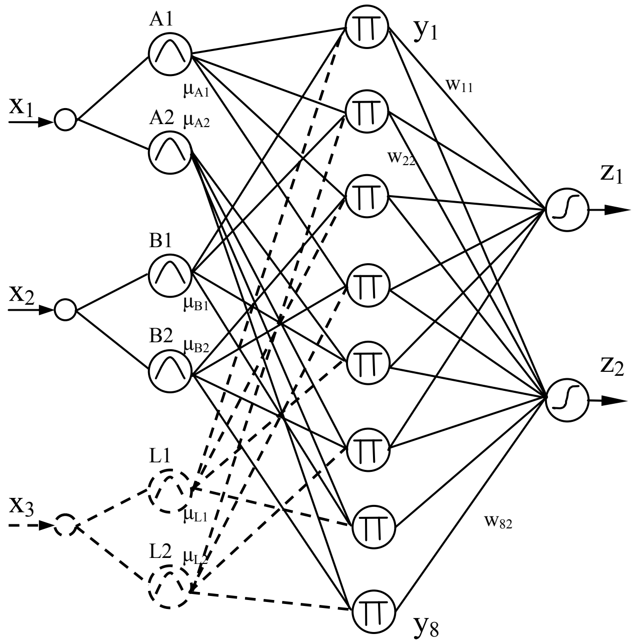

Fig. 3 : LNFC, Labeled Neuro-Fuzzy Classiffier

Premises of rules sent by the third layer are:

$y_1 = \mu_{A1}(x_1) \,.\, \mu_{B1}(x_2) \,.\, \mu_{L1}(x_3)$

$y_2 = \mu_{A1}(x_1) \,.\, \mu_{B1}(x_2) \,.\, \mu_{L2}(x_3)$

…

$y_8 = \mu_{A2}(x_1) \,.\, \mu_{B2}(x_2) \,.\, \mu_{L3}(x_3)$

4.2 effect of labels

The premises of Fuzzy rule established by the third layer are affected by the membership functions of labels, rather than by the labels themselves (fig. 4). That is to say, contrary to the case of the ANN when the choice of labels values directly influences the classification performances.

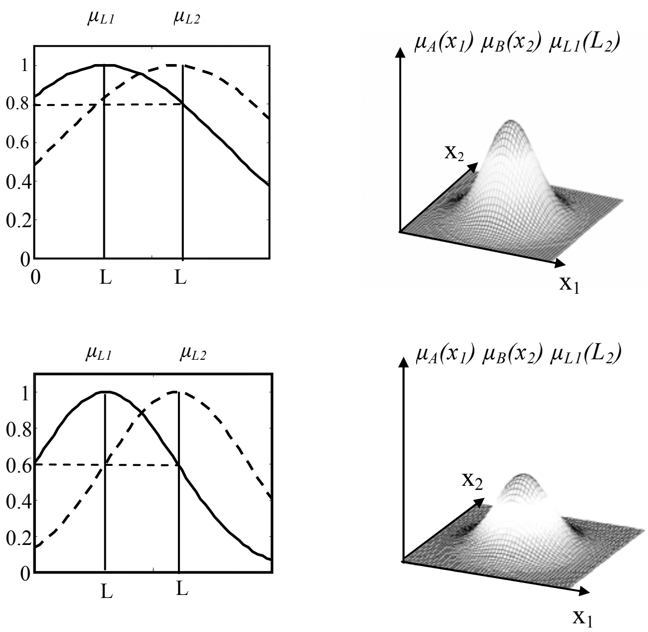

Fig. 4: Effect of labels membership functions.

5 Iris data basis classification

To appreciate the performances of the proposed system, tests are carried out on the Iris data basis by the NFC with and without labels. In both cases, we used three linguistic variables: Small, Medium and Big (fig.5). The results obtained on the whole of all examples, without adjustment of the membership functions, show that the labeled training reduces the iteration number from 13 to 7; on the other hand, the classification rate which is equal to 97,33 % remains unchanged.

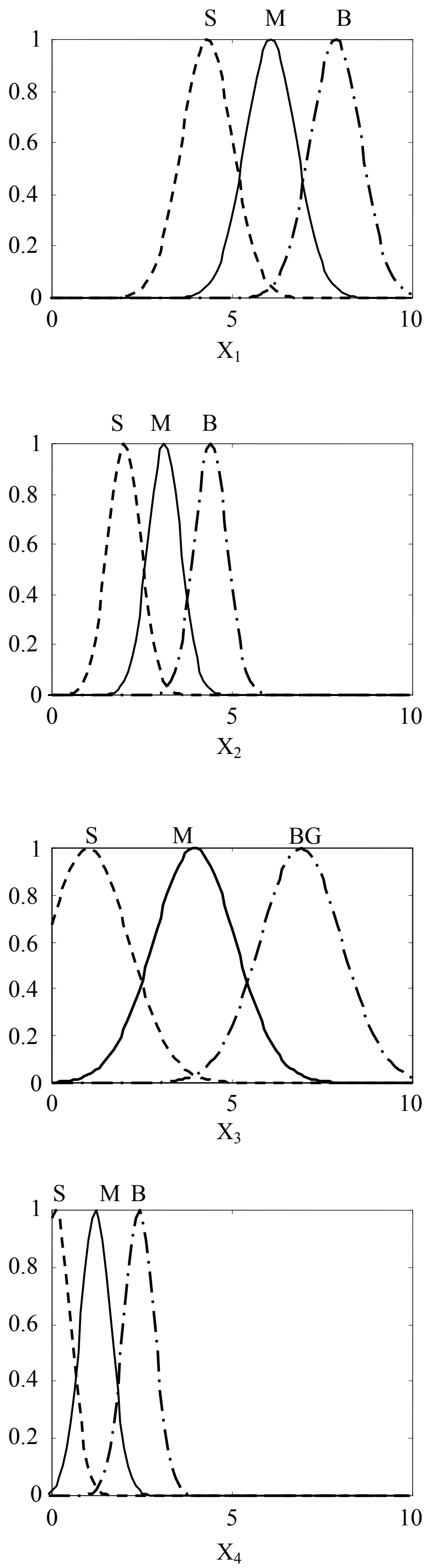

Fig. 5: Membership functions of Iris data basis features.

6 human thigh data basis classification

The image of fig. 6 is acquired by cryosection color photography. This image is put under format tiff. It has a size of 670*415 pixels. A manual classification was made by an expert and four component were identified (grease, bone, marrow and muscle). Each one of these component corresponds to a class and each class is represented by a file of 300 pixels. The sample obtained consists of 1200 pixels, 300 pixels of each class. The addition of components X and Y (to locate geometrical position of a pixel and to take account of its vicinity) improves the classification performances [11].

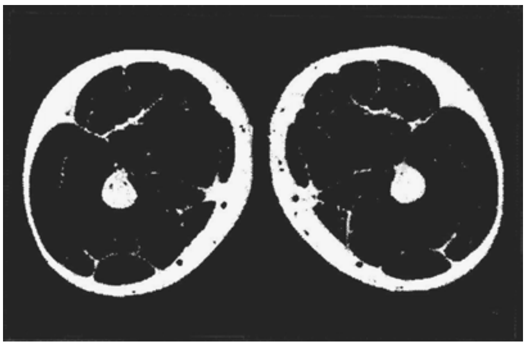

Fig. 6: Image of human thigh creyosection

Tests done on the data basis of the human thigh without and with labels are illustrated on fig. 7. In the both cases we used three linguistic variables: Small, Means, Big.

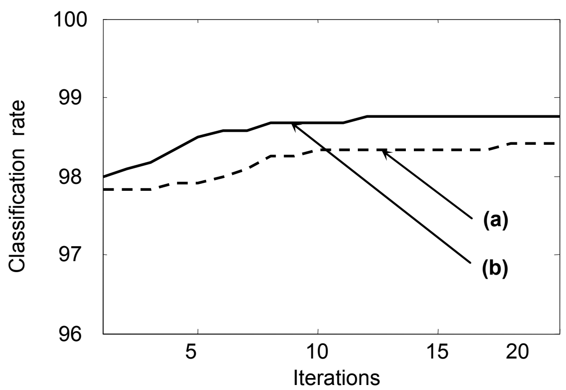

Fig. 7: Evolution of the classification rate during the training:
(a) : Without labels
(b) : Labelled classification

Using a cross validation of order 4, the gotten results are given on table 3:

	Classification without labels		Labeled classification					
			$\mu_{Li}(L_i)=1$ $\mu_{Li}(L_j)=0.8$		$\mu_{Li}(L_i)=1$ $\mu_{Li}(L_j)=0.85$		$\mu_{Li}(L_i)=1$ $\mu_{Li}(L_j)=0.9$	
	Rate	Iter.	Rate	Iter.	Rate	Iter.	Rate	Iter.
Data-set 1	99.67	7	100	8	100	5	100	6
Data-set 2	100	4	100	3	100	3	100	2
Data-set 3	99.67	2	99.67	2	99.67	1	99.67	1
Data-set 4	92.33	48	93	44	93	36	93	37
Mean	97.92	15.25	98.17	14.25	**98.17**	**11.25**	98.17	11.50

Table 3: Results of Iris data set classification using 3 labels membership functions

7 Conclusion

Ours investigation works shows that the classification performances depend intensely on the used training strategy. Thus, the conception of the proposed model allows to exploit the properties of the Neuro-Fuzzy Systems and to improve their performances by using the labelled classification. This latter makes it possible that the LNFC give satisfying results without adjustments of the membership functions parameters.

The experimental results carried out on the Iris and human thigh data basis with the Neuro-Fuzyy Classifier are satisfactory. Moreover, the choice of the labels is made less critical compared to the case of the ANN. In prospects, we plan to generalize this approach with other classification systems and others types of data.

References

[1] Rumelhart D.E, Hinton G.E, et Williams R.J., learning internal representation by error propagation, parallel distributed processing: exploration in the micro structure of cognition, *D.E. Rumelhart et J.L. McClelland (edition), MIT press Cambridge*, 1986, pp 318-362.

[2] Looney C., *Pattern Recognition Using Neural Networks*, Oxford University Press, New York, 1997.

[3] Jacobs R.A., Increased rates of convergence through learning rate adaptation, *Neural Networks*, vol.1, No.4, 1988, pp 295-307.

[4] Nemissi M. et Serdi H., Performances de l'application de l'apprentissage étiqueté avec le Perceptron Muliticouches, *dans Proc. Conférence Interbationale : Science Electroniques, Technologie de l'information et des Telecomunications (IEEE-SETIT'2004), Souse-Tunisie*, mars 2004, p159.

[5] Nemissi M., Seridi H., Akdag H., Performances de l'apprentissage étiqueté avec Random vector functional LinK Neural Network. *Dans Proc. 4ème Séminaire National en Informatiuque (SNIB'2004),* Biskra-Algérie, mai 2004, p320.

[6] Pao Y. H., Adaptive Pattern Recognition and Neural Network, Addison-Wesley, Reading, MA, 1989.

[7] Pao Y. H., Park G. H., and Sobjic D. J. , Learning and generalization characteristics of the Random vector functional link net, Neurocomputing, vol. 6, 163-180, 1994.

[8] Sun C. T., and Roger Jang J. S., A neuro-fuzzy classifier and its applications, *In proc. of IEEE international conference on fuzzy systems*, San Francisco, 1993, pp. 94-98.

[9] Roger Jang J. S., *Neuro-Fuzzy Modeling: Architecture, Analyses and Application*, PHD thesis, University of California, Berkeley, 1992.

[10] Shi Y., and Mizumoto, M., Some consideration on conventional neuro-fuzzy learning algorithms by gradient descent method, *Fuzzy Sets and Systems*, Vol. 112, 2000, pp. 51-63.

[11] Borgi, A., Apprentissage supervisé par génération de règles: le système SUCRAGE, Thèse de l'Université de Paris 6, 1999.