ESANN'1999 proceedings - European Symposium on Artificial Neural Networks
Bruges (Belgium), 21-23 April 1999, D-Facto public., ISBN 2-600049-9-X, pp. 289-294

SUPERVISED ART-II: A New Neural Network Architecture, With Quicker Learning Algorithm, For Learning and Classifying Multivalued Input Patterns.

Kamal R. Al-Rawi, Consuelo Gonzalo, and Aqueda Arquero;

Departamento de Arquitectura y Tecnología de Sistemas Informáticos;
Facultad de Informática;
Universidad Politécnica de Madrid;
28660 Boadilla del Monte,Madrid, SPAIN
e-mail:kamal@palmera.datsi.fi.upm.es

ABSTRACT

A new artificial neural network (ANN) architecture for learning and classifying multivalued input patterns has been introduced, called Supervised ART-II. It represents a new supervision approach for ART modules. It is quicker in learning than Supervised ART-I when the number of category nodes is large, and it requires less memory.

The architecture, learning, and testing of the newly developed ANN have been discussed.

I- INTRODUCTION

Since 1976, when the first unsupervised ART ANN has been developed by Groosberg , many unsupervised ART modules have been developed, ART1, ART2, ART3, SART, and Fuzzy ART. However, in the early nineties, two supervised ART architectures have been developed. The first one is ARTMAP, which has the ability of learning and classifying binary multivalued input patterns [1]. The second one is Fuzzy ARTMAP, which has the ability of learning and classifying analog input patterns, in addition to the binary one [2]. Architecture of ARTMAP has been built from a pair of ART1´s modules, while architecture of Fuzzy ARTMAP has been built from a pair of Fuzzy ART modules.

Recently, Supervised ART-I, has been developed [3]. This ANN, as the Fuzzy ARTMAP, has the ability of learning and classifying of both binary and analog multivalued input patterns. However, it has a simple architecture, fewer parameters, and requires less memory than Fuzzy ARTMAP. This lead to quicker learning and classifying algorithms, with the same accuracy as Fuzzy ARTMAP. The great achievement of Supervised ART-I is that, its architecture has been built from a

ESANN'1999 proceedings - European Symposium on Artificial Neural Networks
Bruges (Belgium), 21-23 April 1999, D-Facto public., ISBN 2-600049-9-X, pp. 289-294

single Fuzzy ART, instead of a pair of them as in Fuzzy ARTMAP. This led to the elimination of the map field.

This article has been conducted to build a new generation of Supervised ART-I, called Supervised ART-II. It is quicker in learning and requires less memory. The classification accuracy and parameters are like those of Supervised ART-I.

II- THE SUPERVISED ART-II

A) THE ARCHITECTURE

Supervised ART-II, as Supervised ART-I, has been built from a single Fuzzy ART module. The one-dimensional memory N of the category nodes of Supervised ART-I is divided into L-one-dimensional memories ($N_k ; k = 1 ... L$) in Supervised ART-II, (see figure 1). Each of these one-dimensional memory N_k is called “stack”. The stack number k represents the class code for all its committed nodes.

The size of N_k (number of nodes which are available to be committed) in each stack are not necessarily equal. It depends on the nature and size of the data of each class. However, if no previous knowledge about the data is available, an equal memory size is recommended.

B) THE TRAINING

During training phase, a stream of multivalued input patterns $A^{(t)}$ and their class codes $b^{(t)}$ are introduced simultaneously to the network. The choice value is computed for each committed node in all the stacks;

$$T_{j_k k} = \frac{\sum_{i=1}^{2M} (A_i \wedge w_{ij_k k})}{\alpha + \sum_{i=1}^{2M} w_{ij_k k}}; \; j_k = 1 ... C(k); \; k = 1 ... L$$

where C(k) is the number of committed nodes in the stack number k, and $w_{ij_k k}$ are the weights, which connect each category node j_k in each stack k with all input nodes i (i=1...2M) where M is the dimension of the input vector A. α is the choice parameter (α >0).

The node, which has the maximum choice value for each stack is determined;

$$T_{J_k} = max \{T_{j_k k} ; j_k = 1 ... C(k)\}$$

These maximum choice value nodes are the candidates of their stacks to represent the current input. The node, which has the highest choice value T_{JK} among all the candidate nodes, is chosen to represent the current input;

$$T_{JK} = max \{T_{J_k} ; J_k = J_1 ... J_L \}$$

The match value is computed for the winning node;

ESANN'1999 proceedings - European Symposium on Artificial Neural Networks
Bruges (Belgium), 21-23 April 1999, D-Facto public., ISBN 2-600049-9-X, pp. 289-294

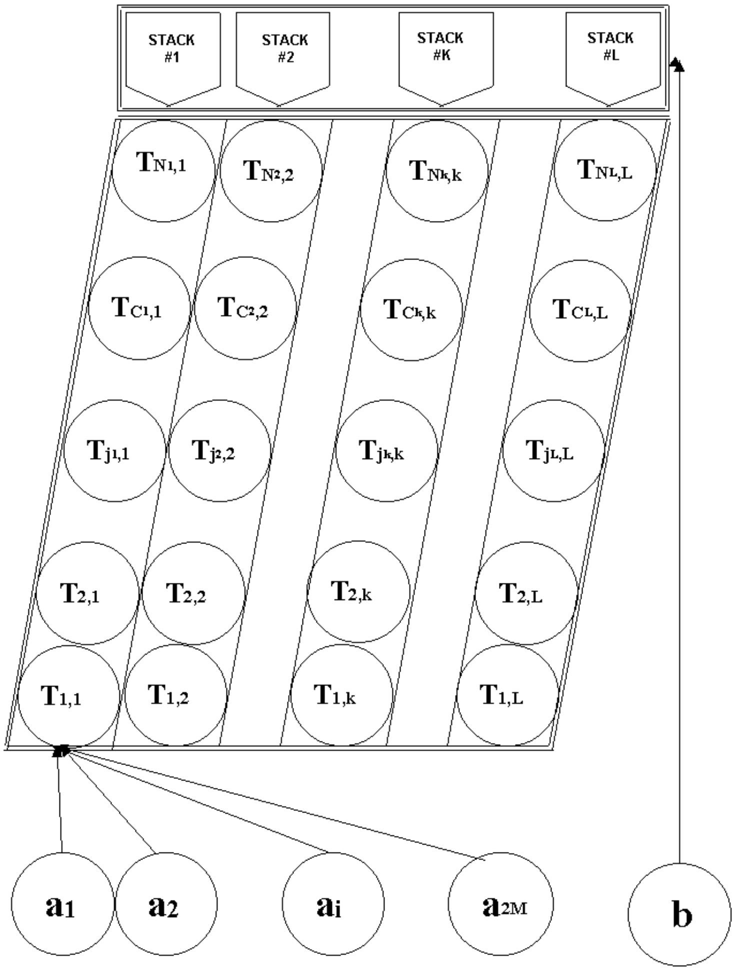

Figure 1. Architecture of Supervised ART-II. Weights should be connected to all the category nodes. The weights connections for the first node are shown here.

ESANN'1999 proceedings - European Symposium on Artificial Neural Networks
Bruges (Belgium), 21-23 April 1999, D-Facto public., ISBN 2-600049-9-X, pp. 289-294

$$\sum_{i=1}^{2M} (A_i \wedge w_{iJK}) / M$$

If the match value of the winning node has passed the vigilance parameter ρ, the class matching should be checked. The class matching is passed if the stack label K of the winning node matches the current class code $b^{(t)}$. Then, the weights of the winning node are trained;

$$w_{iJK}^{new} = \beta(A_i \wedge w_{iJK}^{old}) + (1-\beta)w_{iJK}^{old}; i = 1...2M$$

where, β is the dynamic learning parameter.

If either, the match value, or the class matching, for the current node, has failed, a value of –1 is assigned to the choice value T_{JK} of this node. This is to put it out of competition. Another node, with the maximum choice value should be selected among all the committed nodes of the stack K only. This node is the new candidate for its stack. The candidates of all other stacks are remaining the same. The candidate node, which has the highest choice value among all the candidate nodes of all the stacks is redetermined. We keep doing this until either one of the committed node can represents the current input, then the weights are trained, or a new node should be committed from the stack, which has the label of the class code of the current input $b^{(t)}$. This leads to stacking all the committed nodes according to their classes.

In the case of a new committed node, its weights are assigned the value of the current input $A^{(t)}$, which forces it to be committed;

$$C(b) = C(b) + 1;$$

$$w_{i,C(b),b} = A_i, i = 1...2M$$

where C(b) is the number of committed nodes in the stack b. Therefore, the weights initial values are not required.

As ARTMAP, Fuzzy ARTMAP, and Supervised ART-I, if class matching has failed, the vigilance parameter ρ, should be assigned the match value of the current category node, plus a small value ε, in order to treat the class correction;

$$\rho = \sum_{i=1}^{2M} (A_i \wedge w_{iJK}) / M + \varepsilon$$

However, the vigilance parameter should be relaxed to its base value, before introducing the next input; $\rho = \bar{\rho}$,where $\bar{\rho}$ is the predetermined minimum accepted matching value.

C) THE TESTING

During the testing phase, the choice value for each committed node, in all the stacks are computed;

$$T_{j_k k} = \frac{\sum_{i=1}^{2M} (A_i \wedge w_{ij_k k})}{\alpha + \sum_{i=1}^{2M} w_{ij_k k}}; j_k = 1...C(k); k = 1...L$$

The node, which has the highest choice value T_{JK}, among all the committed nodes, is selected, to represent the current input $A^{(t)}$;

$$T_{JK} = max\{T_{j_k k}; j_k = 1..C(k); k = 1..L\}$$

The match value of the winning node is computed;

$$\sum_{i=1}^{2M} (A_i \wedge w_{iJK}) / M$$

If the match value of the winning node, is greater than, or equal to, the vigilance parameter, the current input belongs to class K. If not the network fails to classify the current input.

III- DISCUSSION

It is clear from the algorithms of Fuzzy ARTMAP [2], Supervised ART-I [3], and Supervised ART-II, that, they have the same classification accuracy (for classification performance see [1, 2, and 4]). However, the last two are quicker in learning due to their simple architectures, because they have been built from a single ART's module.

The Supervised ART-II is quicker in learning binary and analog input patterns, then Supervised ART-I. Stacking the category nodes according to their classes, makes the redetermination of the maximum choice value node quicker than the tagging approach of the Supervised ART-I, when the previous one has failed to pass the vigilance parameter or the class matching. The number of comparisons Nc which are required to redetermine the winning node, among all the committed category nodes C, are Nc=C-1 comparisons for Supervised ART-I, while they are Nc= (C/L-1)+(L-1)=(1/L)C+(L-2) comparisons, as an average, for Supervised ART-II. Therefore, learning time for Supervised ART-II is quicker, compared to Supervised ART-I, as C increases [5]. The training time requirement, for (Supervised ART-I and Supervised ART-II) in a classification task of Landsat TM images, using ALPHAstation 500 (400MHz), are (18.06s, 21.15s), (33.73s, 33.45s), (1m5.23s, 52.41s), and (8m10.66s, 4m8.31s), for number of category nodes of 473, 720, 1033, and 2319 nodes, respectively. The value of vigilance parameter and achieved classification accuracy are (0.95, 79.35%), (0.97, 82.79%), (0.98, 83.39%), and (0.99, 83.82%). A constant dynamic learning rate β =0.1 is used. A total 0f 9 000 pixels are used in training to classify 52 440 pixels into 13 classes. They are also employed for automatic monitoring of forest fires in real time, using the satellite NOAA-AVHRR images [6, 7]. The minimum accuracy is more than 98%.

ESANN'1999 proceedings - European Symposium on Artificial Neural Networks
Bruges (Belgium), 21-23 April 1999, D-Facto public., ISBN 2-600049-9-X, pp. 289-294

The maximum number of category nodes of each stack are predetermined before the learning process. When all the nodes of a particular stack have been committed, borrowing an uncommitted node from another stack is not possible, while in the tagging approach of Supervised ART-I, uncommitted nodes are free to represent any class during learning process. This is the main constrain of the stacking supervision approach of Supervised ART-II.

This limitation of the stacking supervision approach, of Supervised ART-II, can be overcome by increasing the memory size of each stack. This additional memory is compensated by employing only (1/L)th of the released tagging memory N of the tagging supervision of Supervised ART-I. The released memory can be used to increase the memory size of each stack by one fold.

IV- CONCLUSIONS

Supervised ART-II should be employed when a large number of committed nodes (>1000) are expected. Otherwise Supervised ART-I should be used.

REFERENCES

[1] G. A. Carpenter, S. Groosbergh, and J. H. Renold, “ ARTMAP: Supervised real-time learning and classification of non-stationary data by a self organizing neural network,” Neural Network, Vol. 4, pp. 565-588, 1991.

[2] G. A. Carpenter, S. Groosberg, N. Markuzon, J. H. Renold, and D. B. Rosen, “ Fuzzy ARTMAP: A neural network architecture for incremental supervised learning of analog multidimensional maps,” IEEE Trans. On Neural network, Vol. 3, pp. 698-713, 1992.

[3] K. R. AL-Rawi , “ Supervised ART-I: A new neural network architecture for learning and classifying multivalued input patterns,”,to be presented in the International Work-Conference on Artificial Neural Network IWANN, Alicante, Spain, 1999.

[4] G. A. Carpenter, M. N. Gaja, S. Gapa, and C. E. Wooddock, “ ART neural for remote sensing: Vegetation classification from Landsat TM and terrain data”, IEEE Trans. Geosc. Remote Sensing, Vol. 35, pp 308-325, 1997.

[5] K. R. Al-Rawi, C. Gonzalo, and E. Martinez, “Supervised ART-II neural network for Landsat TM image classification,” to be presented in the 19th Symposium of European Association of Remote Sensing Laboratories EARSeL, LATUV, University of Valladolid, Valladolid, Spain, 1999.

[6] K. R. Al-Rawi and J. L. Casanova, “A neural network approach for automatic monitoring of forest fires with NOAA-AVHRR imagery. Part-I: mapping burned areas,” sent for publication to the Int. J. of Remote Sensing, 1999.

[7] K. R. Al-Rawi, J. L. Casanova, and A. Calle, “A neural network approach for automatic monitoring of forest fires with NOAA-AVHRR imagery. Part-II: fires detection,” sent for publication to the Int. J. of Remote Sensing, 1999.