Годишник на секция “Информатика”
Съюз на учените в България
Том 6, 2013, 32–38

Annual of “Informatics” Section
Union of Scientists in Bulgaria
Volume 6, 2013, 32–38

GENERALIZED NET MODEL OF THE ART1 NEURAL NETWORKS

Todor Petkov, Sotir Sotirov

University of Prof. d-r Assen Zlatarov – Bourgas, Bulgaria
e-mail: t-petkov@mail.bg; ssotirov@btu.bg

Abstract: ART stands for Adaptive Resonance Theory invented by Stephen Grossberg in 1976. ART represents a family of Neural Networks. In this paper we introduce the work of GN model with ART1 Neural Network static structure. The model will explain how individual modules can calculate input information and how it passes it to others. It can also be used for understanding and optimization of ART1.
Keywords: Neural Networks, Generalized Nets, ART1

1. Introduction

Each learning system has to be able to adapt in changing environment (i.e. be flexible), but constant changes can make the system unstable, because the system can learn new information only by forgetting older information [3, 4].

ART is a type of neural network which solves this problem [5, 6]. ART supports self-regulating control that allows automatic recognition and learning. In general, it consists of two layers. The first layer is known as Comparison layer, the second one is known as Recognition layer that is fully connected with “bottom-up” and “top-down” weights and a Reset module, which controls the degree of similarity between both layers.

There are two methods of learning: supervised and unsupervised. Unsupervised learning is used for neural networks ART1, ART2, ART3 [7, 8]. Supervised is known as ARTMAP [7, 8].

Here we will describe the work of ART1 which operates with binary input vectors [3]. ART1 consist of:

- Comparison layer – It takes the input vector and transfer through “bottom-up” weights to Recognition layer where it takes place.
- Recognition layer – It connects the Comparison layer with “top-down” weights. Recognition layer clusters and receives data from the Comparison layer.
- Reset module – Controls the degree of similarity between input vectors and makes a decision whether the winning neuron will obtain input vector or not.
- Gain 1 – Controls the activity of the Comparison layer.
- Gain 2 – Controls the activity of the Recognition layer.

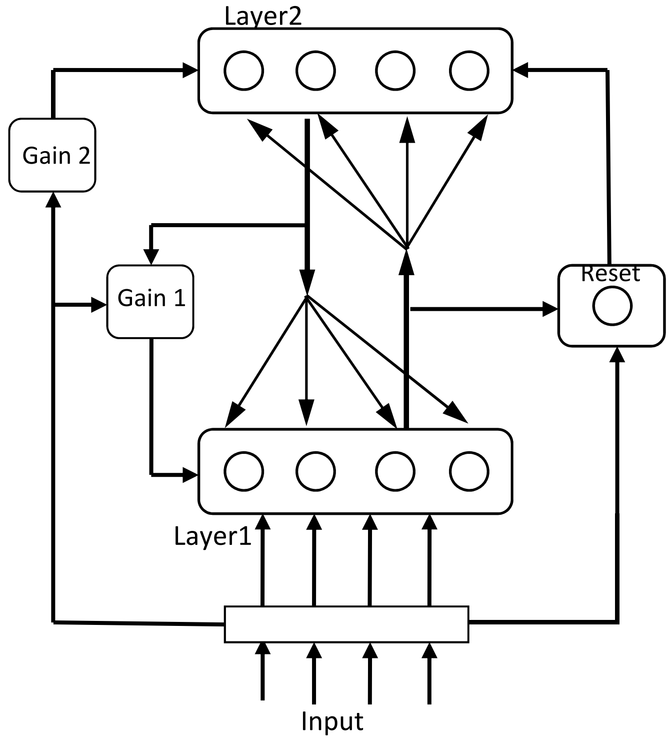

Figure 1. Basic ART1 Architecture

According to [2, 9] the learning algorithm is as follows:

1. Initial weights „W_{ij}” , „W_{ji}” are initialized:

$$w_{ij}(0) = \frac{1}{1+N}, w_{ji}(0) = 1$$

   where, “N” is the number of neurons in the comparison layer.
   Enter vigilance threshold “ρ” between $0 \le \rho \le 1$
2. Application of new input vector “I”.
3. The activating values on each neuron „y_j” are calculated in the recognition layer.

$$y_j = \sum_{j=1}^{N} w_{ij}(t) x_i;$$

   where, „x_i” is the value on each neuron from the “Comparison layer”.
4. Appropriate winning neuron “k”

$$y_k = \max(y)$$

5. Vigilance threshold is checked by the formula:

$$\rho \leq \frac{X}{I}$$

where, "X" is the calculated value from Recognition layer.

6. If vigilance threshold is not reached, then return to step 3 and the winning neuron is inhibited during next training.
7. If vigilance threshold is reached then update weights:

$$w_{kl}(t+1) = w_{kl}(t)x_l$$

$$w_{ik}(t+1) = \frac{w_{kl}(t)x_l}{\frac{1}{2} + \sum_{i=1}^{N} w_{kl}(t)x_l}$$

8. After all weights are updated, go back to step 2.

2. GN–Model

Initially the following tokens enter the generalized net:

- In place L_1 – one token with initial characteristic "Input vectors";
- In place L_9 – one token with initial characteristic "Structure of Neural Network in the first layer and weight matrix";
- In place L_{12} – one token with initial characteristic "Enter vigilance threshold";
- In place L_{21} – one token with initial characteristic "Structure of Neural Network in the second layer and weight matrix".

The GN consists of six transitions:

$$A = \{Z_1, Z_2, Z_3, Z_4, Z_5, Z_6\}$$

where the transitions describe the following processes:

- Z_1 – Dispensation input signal;
- Z_2 – Operation of the Comparison Layer;
- Z_3 – Operation of the Gain 1;
- Z_4 – Operation of the Orienting Sub-system;
- Z_5 – Operation of the Gain 2;
- Z_6 – Operation of the Recognition Layer;

All transitions will be formally described below.

$$Z_1 = \langle \{L_1, L_6\}, \{L_2, L_3, L_4, L_5, L_6\}, R_1, \vee(L_1, L_6) \rangle$$

$R_1 =$

	L_2	L_3	L_4	L_5	L_6
L_1	*False*	*False*	*False*	*False*	*True*
L_6	$W_{6,2}$	$W_{6,3}$	$W_{6,4}$	$W_{6,5}$	*True*

,

where $W_{6,2} = W_{6,3} = W_{6,4} = W_{6,5}$ = "There are data for sending".

The token moves from place L_1 to place L_6 and does not change its characteristic.

The tokens from place L_6 enter places L_2, L_3, L_4, L_5, changing their characteristics and obtaining one input vector.

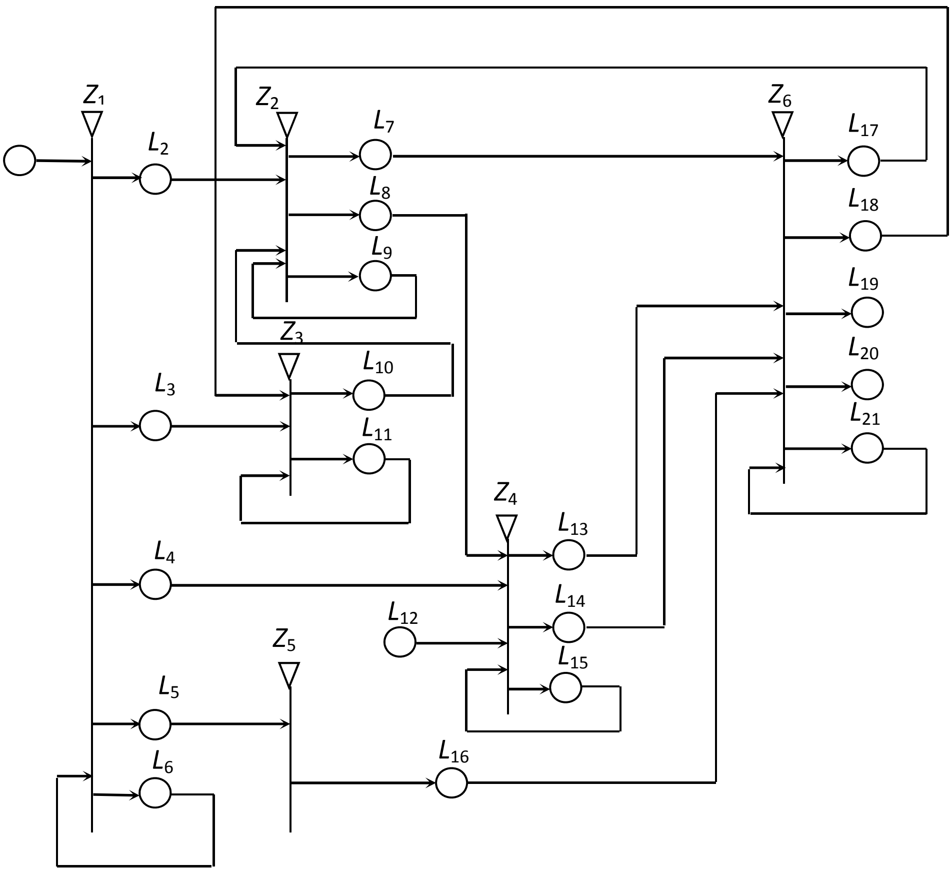

Figure 2. The GN model

The second transition has the form:

$$Z_2 = \langle \{L_2, L_9, L_{10}, L_{17}\},\ \{L_7, L_8, L_9\},\ R_2,\ \vee(\wedge(L_2, L_{10}), L_9, L_{17})\rangle$$

$R_2 =$

	L_7	L_8	L_9
L_2	*False*	*False*	*True*
L_9	$W_{9,7}$	$W_{9,8}$	*True*
L_{10}	*False*	*False*	*True*
L_{17}	*False*	*False*	*True*

,

where $W_{9,7} = W_{9,8} =$ "Computed outputs of the first layer".

The tokens transfer from places L_2, L_9, L_{10} and L_{17} to place L_9, where they get united in a new token that receives characteristic "Structure of Neural Network and weight matrix".

The tokens move from place L_9 to places L_7 and L_8 and do not change their characteristics.

The third transition has the form:

$$Z_3 = \langle \{L_3, L_{11}, L_{18}\},\ \{L_{10}, L_{11}\}, R_3 \vee(L_3, L_{11}, L_{18})\rangle$$

$$R_3 = \begin{array}{c|cc} & L_{10} & L_{11} \\ \hline L_3 & False & W_{3,11} \\ L_{18} & False & True \\ L_{11} & W_{11,10} & True \end{array},$$

where:

- $W_{3,11}$ = "There is no token in L_{18}";
- $W_{3,11}$ = "There is processing data".

The token from place L_3 enters place L_{11} and receives the characteristic "Activate Comparison Layer".

The token from place L_{18} enters place L_{11} and receives the characteristic "Comparison Layer is inhibit".

The token from place L_{11} enters in place L_{10} and receives the characteristic "Appropriate active signal".

The fourth transition has the form:

$$Z_4 = \langle \{L_4, L_8, L_{12}, L_{15}\},\{L_{13}, L_{14}, L_{15}\}, R_4, \vee(L_4, L_8, L_{12}, L_{15})\rangle$$

$$R_4 = \begin{array}{c|ccc} & L_{13} & L_{14} & L_{15} \\ \hline L_4 & False & False & True \\ L_8 & False & False & True \\ L_{12} & False & False & True \\ L_{15} & W_{15,13} & W_{15,14} & True \end{array},$$

where:

- $W_{15,13}$ = "Vigilance threshold was not reached";
- $W_{15,14} = \neg W_{15,13}$.

The tokens from places L_4, L_8 and L_{12} enter place L_{15} and do not change their characteristics.

The tokens from place L_{15} enter places L_{13}, L_{14} and receive characteristics accordingly "Vigilance threshold has not been reached" and "Vigilance threshold was reached".

The fifth transition has the form:

$$Z_5 = \langle \{L_5\},\ \{L_{16}\}, \mathrm{R}_5, \vee(L_5)\rangle$$

$$R_5 = \begin{array}{c|c} & L_{16} \\ \hline L_5 & True \end{array},$$

The token from place L_5 enters place L_{16} and receives the characteristic "Activate Recognition Layer".

The sixth transition has the form:

$$Z_6 = \langle \{L_7, L_{13}, L_{14}, L_{16}, L_{21}\},\ \{L_{15}, L_{16}, L_{17}, L_{18}, L_{19}\}, R_6, \wedge(\vee(L_7, L_{13}, L_{19}), L_{12})\rangle$$

$$R_6 = \begin{array}{c|ccccc} & L_{17} & L_{18} & L_{19} & L_{20} & L_{21} \\ \hline L_7 & False & False & False & False & True \\ L_{13} & False & False & False & False & True \\ L_{14} & False & False & False & False & True \\ L_{16} & False & False & False & False & True \\ L_{21} & W_{21,17} & W_{21,18} & W_{21,19} & W_{21,20} & True \end{array},$$

where:

- $W_{21,17} = W_{21,18}$ = "Output signal was calculated";
- $W_{21,19}$ = "Input signal has hit the cluster";
- $W_{21,20}$ = "Create new cluster".

The tokens from places L_7, L_{13}, L_{14} and L_{16} enters place L_{21} and merge in a new token which receive the characteristic "Structure of Neural Network in the second layer; weight matrix".

The tokens from place L_{21} enter places L_{17} and L_{18} and receive there the characteristic "Output signal was calculated".

The token from place L_{21} enters place L_{19} and receives there the characteristic "Clustered data".

The token from place L_{21} enters place L_{20} and receives there the characteristic "New cluster was created".

3. Conclusion

ART is the one of the neural networks that takes its inspiration not from the human brain. Its model is taken from human eyes. That determines its behavior. The ART have two different algorithms: slowest and fastest. In this paper we use fastest algorithm. This is the first of series of papers devoted to neural network of type ART, where is described operation of ART1. In future, we will pay attention to ART1's algorithm of functioning.

References

[1] Atanassov, K., *Generalized Nets*, World Scientific, Singapore, 1991.

[2] Carpenter, G. A. *The ART of Adaptive Pattern Recognition by Self Organizing Neural Network*, 1988.

[3] Carpenter, G.A., S. Grossberg, A massively parallel architecture for a self-organizing neural pattern recognition machine, *Computer Vision, Graphics, and Image Processing*, Vol. 37, 1987, 54–115.

[4] Grossberg, S. *Neural Network and Natural Intelligence*, Cambridge, MA: MIT Press, 1988.

[5] Grossberg, S. Adaptive pattern classification and universal recoding. II. Feedback, expectation, olfacation, and illusions, *Biol. Cybernet*. Vol. 23, 1976, 187–202.

[6] Grossberg, S. Adaptive pattern classification and universal recoding. I. Parallel development and coding of neural feature detectors, *Biol. Cybernet*, Vol. 23, 1976, 121–134.

[7] Grossberg, S. The adaptive self-organization of serial order in behavior: Speech, language, and motor control, in *Pattern Recognition by Humans and Machines* (E. C. Schwab and H. C. Nusbaum, Eds.), Vol. 1, Academic Press, New York, 1986.

[8] Grossberg, S. Do all neural networks really look alike? A comment on Anderson, Silverstein, Ritz, and Jones, *Psychol. Rev*. Vol. 85, 1978, 592–596.

[9] Kumar, N., R. S. Joshi, *Data Clustering Using Artificial Neural Networks*, Mechanical Dept., DBPC Dhuri, Punjab, 2007.