Pattern Classifier

Atilla E. Günhan,

Department of Information Science, University of Bergen.
Hightechnology center, N-5020 Bergen, Norway

E-mail: gunhan@cc.uib.no Telefax: (47-5) 544107

Abstract: In the present work, an alternative multi-layer unsupervised neural network model that may approximate certain neurophysiological features of Natural Neural Systems has been studied. The Network is formed by two parts. The first part of the network plays a role as a Short Term Memory that is a temporary storage for each pattern. The task for this part of the network is to preprocess incoming patterns without memorizing, in other words, to reduce the dimensions and the linear dependency among patterns by determining their relevant representations. This preprocessing ability is obtained by a dynamic lateral inhibition mechanism on the hidden layer. These representations are the input patterns for the next part of the network. The second part of the network may be accepted as a Long Term Memory which classifies and memorizes incoming pattern informations that come from hidden layer.

1 Architecture of the Network and Training process

The network is composed of two feed forward layers (Fig. 1).

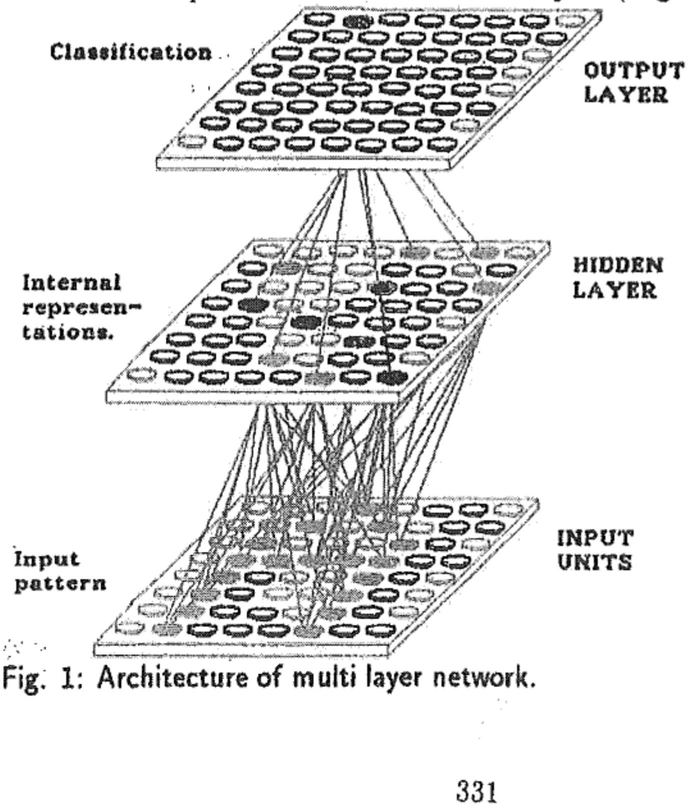

Fig. 1: Architecture of multi layer network.

64 input units, a hidden layer with 64 hidden units and an output layer with 64 units. The layers are fully interconnected to the next layer. The output and hidden layers have internal lateral inhibitions. The first part that is between input units and hidden layer accepts input patterns one at a time (Fig 2.) In other words. when we present the first pattern to this part, only this part of the network will be trained until we obtain internal representations in the hidden unit for this pattern.

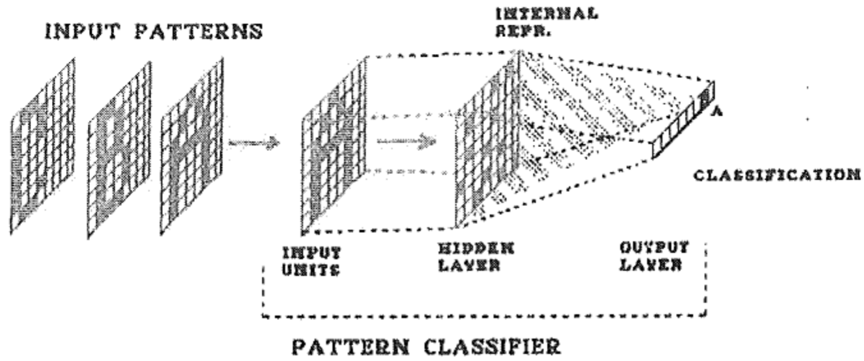

Fig. 2: Training process.

Before we can present the second pattern this part of the network will be initialized. After we have obtained an internal representation for the first pattern in the hidden layer, this representation presented to the second part of the network as an input for final classification of the pattern. There are two possible preprocessing in the first layer. We can train the network until we obtain internal representations of the input patterns. In this type of training, it is possible that patterns can have common elements. In the second type of trining, we can train the network by eliminating common activated units until we obtain unique representations for each input patterns. In this second method patterns will be reduced into their orthogonal representations.

The second part of the network can be considered as another single layer network between the hidden layer and output layer. The function of this second network differs from the first part of the network in terms of the way of training. This part accepts the internal representation as input and the layer is trained untill the pattern is classified. The next internal representation, which is the result of the second pattern that is presented in the first part. will also be accepted as second input pattern to the second part of the network without any initialization. In this way, the second part of the network is capable of training and memorizing a sequence of patterns as in traditional training.

A "neighbour inhibition" method with an Eight-directional Inhibition Strategy (Fig. 3) is used in the layer of the first part of the network that is in fact the hidden layer for the whole network. In this strategy most active unit inhibits the activity of its closest neighbour units. Units in the layer organized in a two dimensional array. In the output layer of the network, the "Winner-take-all" method is used.

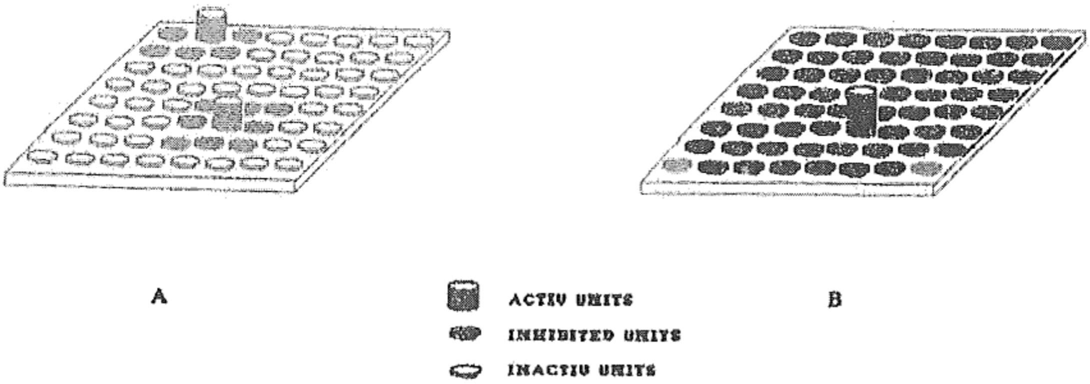

Fig. 3: a) Neighbour neuron inhibition with Eight-directional Inhibition Strategy. Most active unit inhibits the activity of its closest neighbour units. b) Winner-take-all Inhibition Strategy. Winner unit pushes other neurons into a constant minimum value.

2 Learning algorithm of the Network

There is a large number of fibers that provide synaptic connection to a given Purkinje cell (Fig. 4) in a Natural Neural Network[9]. We denote the number of all parallel fibers by N. The input carried by the k^{th} parallel fiber at the time t is denoted by $s_k(t)$: it can have the value of 1 or 0 according to whether the fiber carries an impulse or not. The effect of the input from the k^{th} fiber on the i^{th} Purkinje cell is determined by the synaptic coupling strength $W_k^{(i)}$.

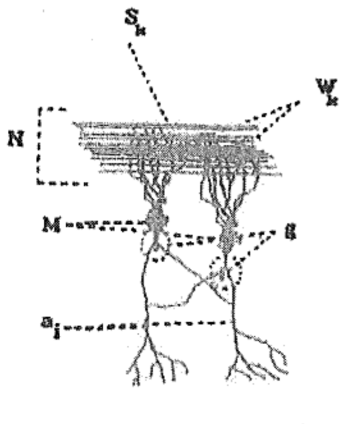

Fig. 4: Symbolic notations for two Purkinje cells.

The output of the network is carried by the axons of the M Purkinje cells and acts to inhibit muscle activity. The strength of the effective inhibition of a cell denoted by g. This activity of the i^{th} Purkinje cell is characterized by the quantity $a^{(i)}(t)$, which is 1 or 0 depending on whether the cell it originates at is active (*i.e.* has fired) or not. The activity state of the Purkinje cell in turn depends on its axon hill potential, $v^{(i)}(t)$.

2.1 Neuronic equation (Netinput)

To describe the behaviour of a neuron, I will employ a formulation based on ref. 3. The dynamic behaviour of a neuron is governed by the *neuronic equation*:

$$v^{(i)}(t+\tau) = \Big[(1-\lambda)v^{(i)}(t) + \sum_{k=1}^{N} W_k^{(i)}(t)s_k(t) - \sum_{j=1}^{M} g_j^{(i)}a^{(j)}(t)\Big](1-a^{(i)}(t)) \,, \qquad (1)$$

where

- $v^{(i)}(t+\tau)$:Hill potential at time $(t+\tau)$ of cell i. (In natural systems this time is $\tau \approx 1$ msec.)
- $(1-\lambda)v^{(i)}(t)$:The remaining potential of neuron i from time t at time $(t+\tau)$. If no other input arrives from fibers then the potential of the neuron decreases exponentially with λ decay constant ($\lambda = 0.1$ in my appl.)
- $W_k^{(i)}(t)s_k(t)$:The excitation arriving from the k^{th} parallel fiber to neuron i at time t.
- $g_j^{(i)}a^{(j)}(t)$:The inhibition arriving from other neurons to neuron i at time t.
- $(1-a^{(i)}(t))$:A multiplicative factor. This term resets the potential to its resting value after the firing and keeps it there the refractory period (the minimum time between two firing), during which the the neuron is not excitable.

A neuron will become active, if its potential $v^{(i)}$ at the axon hill exceeds a threshold value, $\theta^{(i)}$, ($\theta = 1$ in my appl.). In that case the neuron will emit a non-decreasing pulse of fixed duration τ, along its axon making inhibitory action at all of its synaptic connections.

This activity is described as,

$$a^{(i)}(t) = \begin{cases} 1, if\ v^{(i)}(t) \geq \theta^{(i)} \\ 0, \quad otherwise \end{cases}$$

Given the time-dependent input $s_k(t)$, the response of the network can be calculated with eq. (1). For a shorter period, the synaptic strengths $W_k^{(i)}$ and $g_j^{(i)}$ can be considered as constants. If the incoming pattern remains steady for several time steps; those Purkinje cells whose synaptic strength vectors have the best overlap with the input pattern vector will be most likely to fire; in turn, they will then inhibit their neighbouring neurons as a consequence of the lateral inhibition $g_i^{(j)}$. If the membrane potential, v, of a neuron exceeds a given threshold value, θ, then neuron becomes active (fired) and inhibits its neighbours with a certain inhibition value. Active neuron also resets itself into resting membrane potential value.

2.2 Memory equation (Learning)

On a longer time scale the coupling strengths may change, thus providing a learning ability for the net. When we assume that the inhibitory synapses are fixed, the excitatory synapses leading from the parallel fibers to the Purkinje cells may change according to Hebb's rule[6], which we formulate in the following *memory equation*:

$$W_k^{(i)}(t+\tau) = q^{(i)}\Big[(1-\varepsilon)W_k^{(i)}(t) + \delta a^{(i)}(t)s_k(t-\tau)\Big] \,. \tag{2}$$

where

- $W_k^{(i)}(t+\tau)$:Synaptic strength of coupling between fiber k and neuron i at the time $(t+\tau)$.
- $q^{(i)}$:Normalization constant for neuron i.
- $(1-\varepsilon)W_k^{(i)}(t)$:Exponential decay of synaptic strength ($\varepsilon = 0.001$ in my appl.)
- $\delta a^{(i)}(t)s_k(t-\tau)$:If neuron i active at time t (i.e., $a=1$), the connections between neuron i and fiber k will be strengthened by δ learning rate.

The normalization condition

$$q^{(i)} = \eta^{(i)}/\Big(\sum_{k=1}^{N}\Big[(1-\varepsilon)W_k^{(i)}(t) + \delta a^{(i)}(t)s_k(t-\tau)\Big]\Big) \tag{3}$$

where $\eta^{(i)}$ is a constant (i.e., $\eta^{(i)} = 1$). This condition ensures that the total synaptic strength remains constant,

$$\sum_{k=1}^{N} W_k^{(i)}(t) = \eta^{(i)}.$$

This learning mechanism simply describes the effect that if an excitatory impulse is arriving to a synaptic coupling which will lead in the following time step to a firing of the post-synaptic cell, then that connection will be strengthened by δ. Due to the normalization $\eta^{(i)}$, and the slow exponential decay of the couplings, all "non-successful" synapses will be somewhat weakened at the same time. Exponential decay $(1-\lambda)$ in *neuronic equation* and $(1-\varepsilon)$ in *memory equation* are ignored in the second layer of the network to obtain plausible classification of patterns.

Learning algorithm for the second layer of network is as follows *neuronic equation*,

$$v^{(i)}(t+\tau) = \left[v^{(i)}(t) + \sum_{k=1}^{N} W_k^{(i)}(t)s_k(t) - \sum_{j=1}^{M} g_j^{(i)}a^{(j)}(t)\right](1 - a^{(i)}(t)) \,. \tag{4}$$

and *memory equation*,

$$W_k^{(i)}(t+\tau) = q^{(i)}\left[W_k^{(i)}(t) + \delta a^{(i)}(t)s_k(t-\tau)\right] \,. \tag{5}$$

3 Features of the network

The learning algorithm and architecture of the network explained above are different from other network models. The advantages of the two layer model presented above are:

Ability of preprocessing the incoming pattern.
The network has ability to preprocess the incoming patterns and reduce the dimensions and the linear dependency among these patterns then memorize them. Experiments indicate that the network has ability to classify all input vectors as long as those vectors have less than 40% common elements with each other. When the patterns **A, B, C, 1** and **2** presented to the network (Fig. 5), patterns **C** and **2** have been classified into same class.

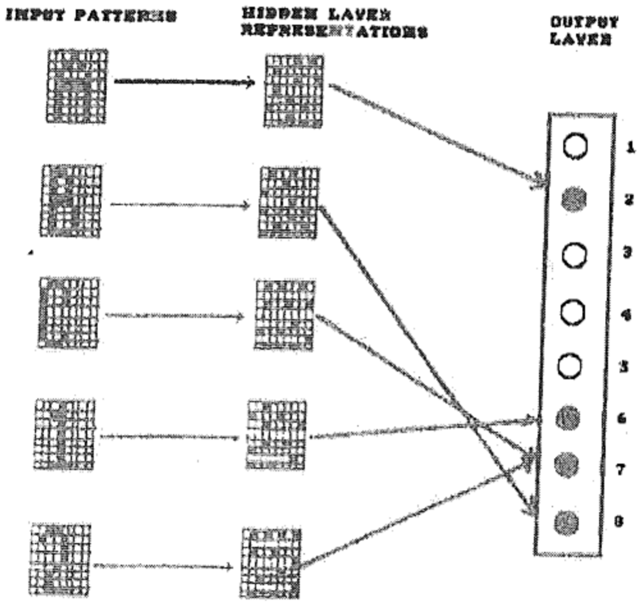

Fig. 5: Input patterns **A, B, C, 1, 2** , their Internal representations and response of the network in output layer.

The reason is that common elements of these two patterns are more than 40% of the number of representation elements. Patterns **A** and **C** had 4 common elements but they have been classified into their own classes since their uncommon elements were in majority. The second layer of the network is limited with properties of `''winner-takes-all''` method[10].

Network architecture and learning algorithm are similar to biological mechanisms.

The learning algorithm is based on neurophysiological activity of real neurons. If the membrane potential, v, of a neuron at a given time exceeds a certain threshold value, θ, then the neuron becomes active (fired) and inhibits its neighbours with a certain inhibition value. Only the weights of the active neuron are adjusted. Active neuron also resets itself into resting membrane potential value.

Network accepts repeated patterns.

Repeated patterns during training do not cause any side effect. They activate always the same output units.

Network has ability to learn a new pattern without training the whole set again.

Network trains with one at a time process. It makes possible to add a new pattern to the network in any given time.

Values of parameters are fixed.

Choice of parameter values for learning and the number of training iterations are fixed. The network may be trained with a value of learning rate 0.1, and 150 cycle for each pattern for each layer. These values are valid for any type of pattern. Inhibition values between the units are set to 0.9 when the method is neighbour inhibition.

Conclusion

In the present work, an alternative unsupervised neural network model that may approximate certain neurophysiological features of Natural Neural Systems has been studied. This work is a result of further investigations on ref. 3 and 4. However the present model shows us the ability of character recognition, It can also be investigated within the subject of signal processing. This model can work as a feature detector or signal classifier in many application areas like speech recognition, active sonar classification, *etc.*

References

[1] Caianiello, E. R. , *Outline of Theory of Thought-Processes and Thinking Machines*, J. Theoret. Biol. (1961) **2**, 204-235

[2] Caianello, E.R. , *Cybernetics of Neural Process* (Ricerca Scientifica, Roma, 1965).

[3] Günhan E. Atilla, Csernai L. P. & Randrup J., *"Unsupervised Competitive Learning in Neural Networks"*, International Journal of Neural Systems, World Scientific, London, Vol. **1**, No. **2** 177-186

[4] Günhan E. Atilla, *"Pattern recognition and Self-Organization of Neural Networks"*, Dept. of Information Science, University of Bergen, Norway. Master thesis, (March 1991).

[5] Günhan E. Atilla, *"Pattern Classifier, An Alternative Method of Unsupervised Learning"*, Neural Network World, International Journal on Neural and Mass-Parallel Computing, Czechoslovakia, Vol. **1**, No. **6** 349-354, (December 1991)

[6] Hebb, D. O. , *The Organization of Behaviour* , (Wiley, New York 1949)

[7] Hinton, G. E. , *Connectionist Learning Procedures*, Technical Report, CMU-CS-87-115, (University of Toronto, Canada 1987)

[8] Kohonen, T., *Self-Organization and Associative Memory*, (Springer-Verlag, New York 1984)

[9] Pellionisz, A. and Llinás, R. , *A computer model of Cerebellar Purkinje Cells*, Neuroscience, (1977) **Vol. 2**, 37-48

[10] Rumelhart, D. E. , McCelland, J. L. and The PDP research group, *Parallel Distributed Processing* (The MIT press, Cambridge 1986), **Vol. 1-2**