



ARTIFICIAL NEURAL NETWORK

Unit-2: Perceptron

Ms. Swetha R.

Department of Electronics and
Communication Engineering
PES University

- **Sequential v/s Batch mode:**
 - For large data sequential processing is faster
 - For smaller data and network, batch processing is faster
- **Maximizing the information content**
- Every training example presented to the BPA should be chosen on the basis that its information content is the largest possible for the task at hand
- This can be achieved in 2 ways
 - a. use of an example that results in the largest training error
 - b. the use of example that is radially different from all those previously used

Activation Function:

- A M.L.P trained with the back-propagation algorithm may, in general, learn faster when the asymmetric sigmoidal activation function are used in neuron model, for example $\tanh(\cdot)$.

Target Value: it is important to choose the target value within the activation function range i.e +1 and -1

Normalization of the inputs:

The training samples must be preprocessed before presenting to the neural network. the preprocessing steps are

Artificial Neural Network-Perceptron

BPA: Heuristics to make performance better

- a. subtracting each input from mean
- b. decorrelate the training samples
- c. covariance: it ensures the different synaptic weight in the networks learn at same speed.

Learning Rate

Initialization

Artificial Neural Network-Perceptron

BPA-Output representation and decision rule



Consider a M-class classification problems, which means each input pattern belong to one of the M-distinct classes.

Therefore, we will have M number outputs in the network.



X_j denotes the j th input pattern. and it is m -dimensional input vector

The output is M -dimensional vector. y_{kj} is the k th output of network in correspondence to the j th input

Therefore, block diagram can be represented as

$$y_{kj} = F_k(x_j)$$

Let the output vector y_j

$$y_j = [y_{1j} \quad y_{2j} \quad \dots \quad y_{kj} \quad \dots \quad y_{Mj}]^T$$

$$y_j = [F_1(x_j) \quad F_2(x_j) \quad \dots \quad F_k(x_j) \quad \dots \quad F_M(x_j)]^T$$

$$y_j = F(x_j)$$

$F(.)$ is continuous function and minimizes the empirical risk function

$$R = \frac{1}{2N} \sum_{j=1}^N \|d_j - F(x_j)\|^2$$

Now train the network with binary values as follows:

- The output space has M dimension
- An input X_j belongs to class C_k
- let $d_k=1$ when X_j belongs to class C_k otherwise 0

Output Decision Rule stated as follows:

Classify the random vector X as belonging to C_k if

$$F_k(X) > F_j(X) \text{ for all } k \neq j$$



THANK YOU

Ms. Swetha R.

Department of Electronics and
Communication Engineering

swethar@pes.edu

+91 80 2672 1983 Extn 753