Unit - 2

Feed Forward Neural Network

Feed Forward Network:

It is an artificial neural network in which the connection between nodes does not form a cycle. Information is processed in one direction.

Components:

i) Layer of input
ii) Hidden layer
iii) Output layer
iv) Weights

Cost function:

$$C(w, b) = \frac{1}{2n} \sum_{x} \|y(x) - a\|^2$$ MSE - Mean Square Error

Loss function:

Cross entropy loss:

$$L(\theta) = \begin{cases} -\log(\hat{y}) & \text{if} \\ -\log(1-\hat{y}) \end{cases}$$

Back propagation:

It is a method of fine tuning the weights of the neural network based on the error rate obtained in the previous iteration (or) epoch.

Proper tuning of the weights allows you to reduce error rates and make the model reliable by increasing its generalizing.

Regularisation:

To overcome the problem of overfitting.

Overfitting - It gives good result during training process but fails in testing.

Underfitting - Fails in training process

Regularisation techniques:

i) L2 Regularation
ii) L1 Regularisation
iii) Dropout
iv) Data Augmentation
v) Early stopping

X. 16m

Kohonen Self Organising Map

It refers to a neural network, which is trained using competitive learning.

i) Competetive learning - Select a wining processing element based on some criteria (Euclidean distance), after the winning processing element is selected, its weight is adjusted according to learning law.

ii) Topology Ordered Map - It implies that if two inputs are similar characteristics, the most active processing element answering two inputs that are located close to each other on the map.

iii) Dimensionality reduction - High dimension to low dimension.

iv) Feature Map - It returns the features of input and simply grouping distance themselves as indicated by similarity between each other.

v) SOM - All the entire learning process occurs without supervision because the nodes are self organising.

Structure

- 1 D
- 2D
- 3D
- vector
- rectangular
- hexagonal

SOM training algorithm:

Step 1 : Initialisation

i) $w_{ij} \rightarrow$ weight matrix

$i \rightarrow$ input neuron

$j \rightarrow$ cluster neuron

ii) Check the topology radius R.

R = 0 : means only winning unit weight updated

R = 1 : nearby neuron weights also updated

iii) Alpha $\rightarrow$ Learning Rate

Step 1 : While the stopping criteria is not satisfied. repeat step 2 to 8.

Step 2 : For each, every input pattern do 3 - 5

Step 3 : Find distance between input and cluster

$$D[j] = \sum_{i=1}^{n} \sum_{j=1}^{m} [x_i - w_{ij}]^2$$

Step 4 : Select winning unit

Step : Weight updation

$$w_{ij}(new) = w_{ij}(old) + \alpha [x_i - w_{ij}(old)]$$

Step 6 : Update α

Step 7 : Reduce R

Step 8 : Check for stopping criteria

Problem :

Construct KSOM to cluster 4 given vectors [0 0 1 1], [1 0 0 0], [0 1 1 0] and [0 0 0 1].

No of cluster formed is 2. Assume learning rate of 0.5.

Soln :

No of i/p vectors $n = 4$

No. of cluster $m = 2$

$$W_{ij} = \begin{array}{c c} & \begin{array}{cc} y_1 & y_2 \end{array} \\ \begin{array}{c} w_{11} \\ w_{21} \\ w_{31} \\ w_{41} \end{array} & \begin{bmatrix} 0.2 & 0.9 \\ 0.4 & 0.7 \\ 0.6 & 0.5 \\ 0.8 & 0.3 \end{bmatrix} \begin{array}{c} w_{12} \\ w_{22} \\ w_{32} \\ w_{42} \end{array} \end{array}$$

1^{st} i/p vector : $\begin{array}{c} \begin{array}{cccc} x_1 & x_2 & x_3 & x_4 \end{array} \\ \begin{bmatrix} 0 & 0 & 1 & 1 \end{bmatrix} \end{array}$

$$D[j] = \sum_{i=1}^{n} \sum_{j=1}^{m} [x_i - w_{ij}]^2$$

$j=1$:

$$D[1] = \sum_{i=1}^{4} \sum_{j=1}^{1} [x_i - w_{ij}]^2$$

$$= (x_1 - w_{11})^2 + (x_2 - w_{21})^2 + (x_3 - w_{31})^2 + (x_4 - w_{41})^2$$

$$= (0-0.2)^2 + (0-0.4)^2 + (1-0.6)^2 + (1-0.8)^2$$

$$= 0.04 + 0.16 + 0.04 = 0.4$$
$$+ 0.16$$

$$D[1] = 0.4$$

$j=2$:

$$D[2] = \sum_{i=1}^{4} \sum_{j=2}^{2} [x_i - w_{ij}]^2$$

$$= (x_1 - w_{12})^2 + (x_2 - w_{22})^2 + (x_3 - w_{32})^2 + (x_4 - w_{42})^2$$

$$= (0-0.9)^2 + (0-0.7)^2 + (1-0.5)^2 + (1-0.3)^2$$

$$= 0.81 + 0.49 + 0.25 + 0.49$$

$$= 2.04$$

$$D[2] = 2.04$$

$$D[1] < D[2]$$

$$0.4 < 2.04$$

$D[1]$ is the winning element.

$$w_{11} = 0.2 + (0.5)(0-0.2) = 0.1$$

$$w_{21} = 0.4 + (0.5)(0.0.4) = 0.2$$

$$w_{31} = 0.6 + (0.5)(1-0.6) = 0.8$$

$$w_{41} = 0.8 + (0.5)(1-0.8) = 0.9$$

$$w_{ij} = \begin{array}{cc} y_1 & y_2 \\ \end{array}$$
$$\begin{bmatrix} 0.1 & 0.9 \\ 0.2 & 0.7 \\ 0.8 & 0.5 \\ 0.9 & 0.3 \end{bmatrix}$$

2^{nd} i/p vector = [1 0 0 0]

$j = 1$

$$D[j] = \sum_{i=1}^{n} \sum_{j=1}^{m} [x_i - w_{ij}]^2$$

$$D[1] = \sum_{i=1}^{4} \sum_{j=1}^{1} [x_i - w_{ij}]^2$$

$$= (1-0.1)^2 + (0.0.2)^2 + (0-0.8)^2 + (0-0.9)^2$$

$$= 0.81 + 0.04 + 0.64 + 0.81 = 2.3$$

$j = 2$

$$D[2] = \sum_{i=1}^{4} \sum_{j=2}^{2} [x_i - w_{ij}]^2$$

$$= [1-0.9)^2 + (0-0.7)^2 + (0-0.5)^2 + (0-0.3)^2$$

$$= 0.01 + 0.49 + 0.25 + 0.09 = 0.84$$

$$2.3 > 0.84$$

$$D[1] > D(2]$$

$$w_{12} = 0.9 + (0.5)(1-0.9) = 0.95$$

$$w_{22} = 0.7 + (0.5)(0-0.7) = 0.35$$

$$w_{32} = 0.5 + (0.5)(0-0.5) = 0.25$$

$$w_{42} = 0.3 + (0.5)(0-0.5) = 0.15$$

$$w_{ij} = \begin{array}{c} \begin{array}{cc} y_1 & y_2 \end{array} \\ \begin{bmatrix} 0.1 & 0.95 \\ 0.2 & 0.35 \\ 0.8 & 0.25 \\ 0.9 & 0.15 \end{bmatrix} \end{array}$$

3^{rd} i/p vector : $[0 \quad 1 \quad 1 \quad 0]$

$$D[j] = \sum_{i=1}^{n} \sum_{j=1}^{m} [x_i - w_{ij}]^2$$

$$D[1] = \sum_{i=1}^{4} \sum_{j=1}^{1} [0-0.1]^2 + [1-0.2]^2 + [1-0.8]^2 + [0-0.9]^2$$

$$= 0.01 + 0.64 + 0.04 + 0.81$$

$$= 1.5$$

$$D[2] = \sum_{i=1}^{4} \sum_{j=2}^{2} [0-0.95]^2 + [1-0.35]^2 + [1-0.25]^2 + [0-0.15]^2$$

$$= 0.90 + 0.42 + 0.56 + 0.02$$

$$= 1.9$$

$$D[1] < D[2]$$

$$1.5 < 1.9$$

$D[1]$ is the winning element

$$w_{11} = 0.1 + (0.5)(0-0.01) = 0.05$$

$$w_{21} = 0.2 + (0.5)(1-0.2) = 0.6$$

$$w_{31} = 0.8 + (0.5)\cdot(1-0.8) = 0.9$$

$$w_{41} = 0.9 + (0.5)(0-0.9) = 0.45$$

$$w_{ij} = \begin{array}{c} \\ w_{11} \\ w_{21} \\ w_{31} \\ w_{41} \end{array} \begin{array}{c} \begin{array}{cc} y_1 & y_2 \end{array} \\ \begin{bmatrix} 0.05 & 0.95 \\ 0.6 & 0.35 \\ 0.9 & 0.25 \\ 0.45 & 0.15 \end{bmatrix} \end{array} \begin{array}{c} \\ w_{12} \\ w_{22} \\ w_{32} \\ w_{42} \end{array}$$

4^{rd} i/p vector : [0, 0, 0, 1]

$$D[j] = \sum_{i=1}^{n} \sum_{j=1}^{m} [x_i - w_{ij}]^2$$

$$D[1] = \sum_{i=1}^{4} \sum_{j=1}^{2} [0-0.05]^2 + [0-0.6]^2 + [0-0.9]^2 + [1-0.45]^2$$

$$= 0.002 + 0.36 + 0.81 + 0.30$$

$$= 1.47$$

$$D[2] = \sum_{i=1}^{4} \sum_{j=1}^{2} [0-0.95]^2 + [0-0.35]^2 + [0-0.25]^2 + [1-0.15]^2$$

$$= 0.90 + 0.12 + 0.06 + 0.72$$

$$= 1.8$$

$D[1] < D[2]$

$1.47 < 1.8$

$D[1]$ is the winning element

$$w_{11} = 0.05 + (0.5)(0-0.05) = 0.02$$

$$w_{21} = 0.6 + (0.5)(0-0.6) = 0.3$$

$$w_{31} = 0.9 + (0.5)(0-0.9) = 0.45$$

$$w_{41} = 0.45 + (0.5)(1-0.45) = 0.72$$

$$w_{ij} = \begin{array}{c} \begin{array}{cc} y_1 & y_2 \end{array} \\ \begin{bmatrix} 0.2 & 0.95 \\ 0.3 & 0.35 \\ 0.4 & 0.25 \\ 0.72 & 0.15 \end{bmatrix} \end{array}$$