NPTEL Week-11 Live Session

on Machine Learning and Deep Learning - Fundamentals and Applications (noc24_ee146)

A course offered by: Prof. Manas Kamal Bhuyan, IIT Guwahati

Powered by:

NPTEL Quiz Solution: Week-10: Artificial neural networks

By

Arka Roy

NPTEL PMRF TA

Prime Minister's Research Fellow

Department of Electrical Engineering, IIT Patna

Web: https://sites.google.com/view/arka-roy/home

PMRF

Prime Minister's Research Fellows

Ministry of Education

Government of India

2) Which of the following is not a type of neural network architecture?

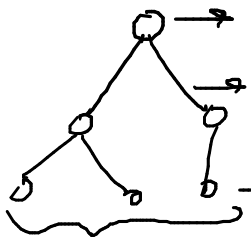
- Convolutional Neural Network
- Long Short-Term Memory
- Self-Organization Map
- Decision Tree Network

NN family

Long Short-Term Memory → Sequential data (Time series type of data)
→ Recurrent Neural Network architecture.
→ Causal data propagation

Prediction α f (current data, Past information)

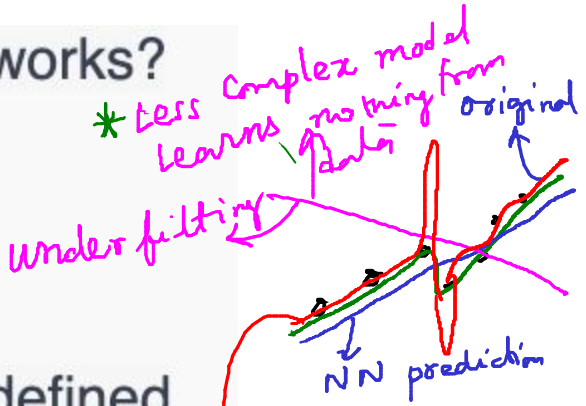
Decision Tree Network → Not a neural network.



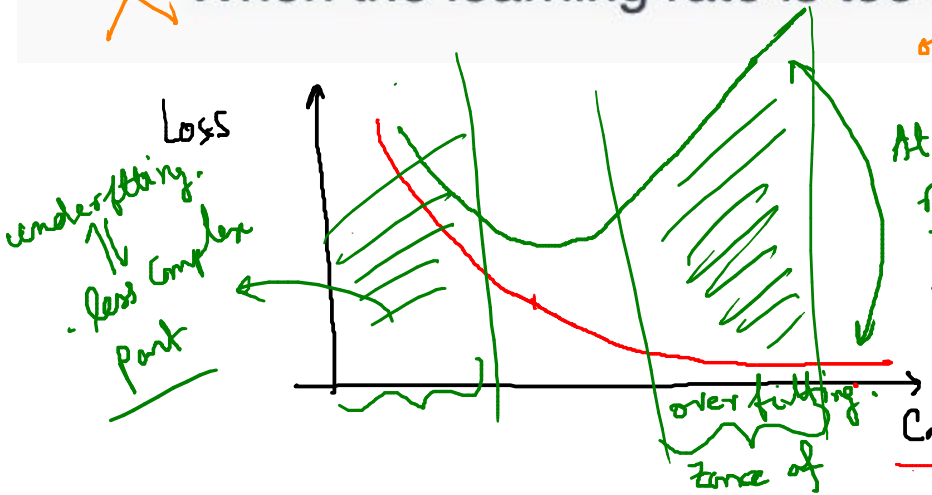
→ Root node.
→ branches
→ Do the decision making
leaf nodes. (classification / Regression)

3) What is overfitting in the context of neural networks?

- When the variance of the model is high
- When the model has too few layers
- When the activation function is not properly defined
- When the learning rate is too high



- highly comput. Complex model.
- In training it fits data very well but on unseen data based Testing model may show degraded performance



At last time result is degraded

And have high varriance with respect to training loss data.

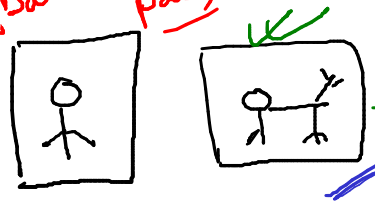
4) What is the purpose of backpropagation in training a neural network?

- [x] Forward pass calculation
- [x] Calculating the cost function (error)
- [ ] Updating weights and biases based on prediction errors
- [x] Regularization of the network

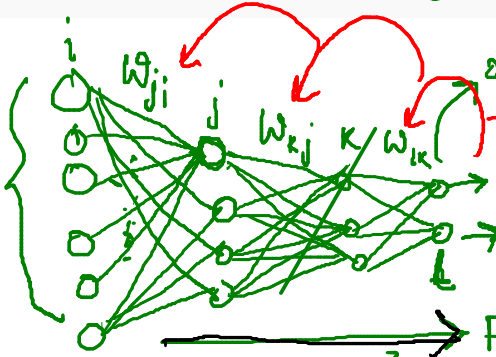
$D = \{x_i, y_i\}_{i=1}^{N}$ → Labels of images.

$\{W_{ji}, W_{kj}, W_{lk}\}$ → Randomly initialized

* Back propagation (Backward pass) through Gradient descent method.



Cat (True)



① Feed that data → forward pass

② calculate loss

③ Based on loss update weights through Back prop.

Prediction ⇒ Human.

Calculate Error

5) You are using the sigmoid activation function in a neural network. If the network's weighted input is 2.5, what is the output of the sigmoid activation function?

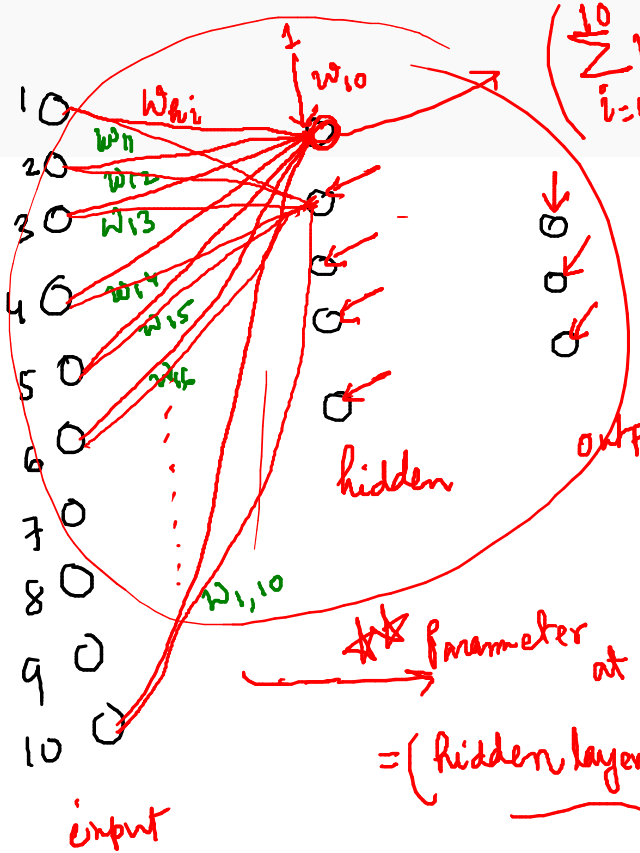
- 0.75
- 0.07
- 0.92
- None

$V = 2.5$

$\text{Sigmoid}(2.5) = \frac{1}{1+e^{-2.5}} = 0.92$

6) In a feedforward neural network with one hidden layer, the input layer has 10 neurons, the hidden layer has 5 neurons, and the output layer has 3 neurons. How many total weights (including biases) are there in the network?

- 70
- 73
- 50
- 65



hidden-output parameter

$= (3 \times 5) + 3$

$= 15 + 3$

$= 18$

$\times$

Total $= 55 + 18 = 73$

10×5

$\left(\sum_{i=1}^{10} W_{hi} \cdot x_i + W_{h0}\right)$ — $\sum_{i=1}^{10} W_{hi}$: 10 weight; $h=1$: 1 bias; for a single neuron in hidden layer

11 entries.

$11 \times 5 = 55$ parameters

$= (5 \times 10) + 5 = 55$

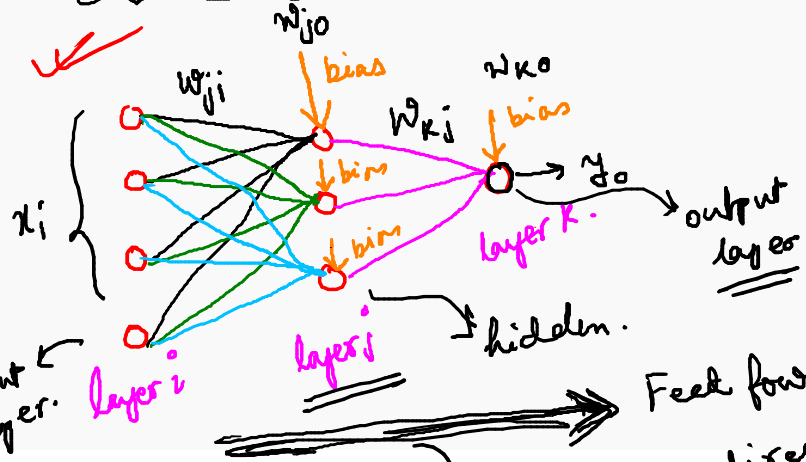
** parameter at W_{hi} (hidden − input)

$= ($hidden layer neuron $\times$ input layer neuron$)$ weight $+$ hidden layer Bias

7) Consider a standard three-layer backpropagation net with d input units, h hidden units, c output units, and bias, find the expression for output of a node k is

- $\sum_{j=1}^{h} f(w_{ki} f(\sum_{i=1}^{d} w_{ji} x_i + w_{j0}) + w_{ko})$
- $\sum_{j=1}^{h} f(w_{kj} f(\sum_{i=1}^{d} w_{ji} x_i + w_{j0}) + w_{ko})$
- $\sum_{j=1}^{h} f(w_{kj} f(\sum_{i=1}^{d} w_{ij} x_i + w_{j0}) + w_{ko})$
- $\sum_{j=1}^{h} f(w_{ki} f(\sum_{i=1}^{d} w_{ji} x_i + w_{j0}))$

activation function



generalised

output at layer j :- / hidden layer $\rightarrow$ $Y_j = f\left(\sum_{i=1}^{d} w_{ji} \cdot x_i + w_{j0}\right)$

output of final layer / layer k :- $Y_0 = f\left(\sum_{j=1}^{h} w_{kj} \cdot Y_j + w_{k0}\right)$

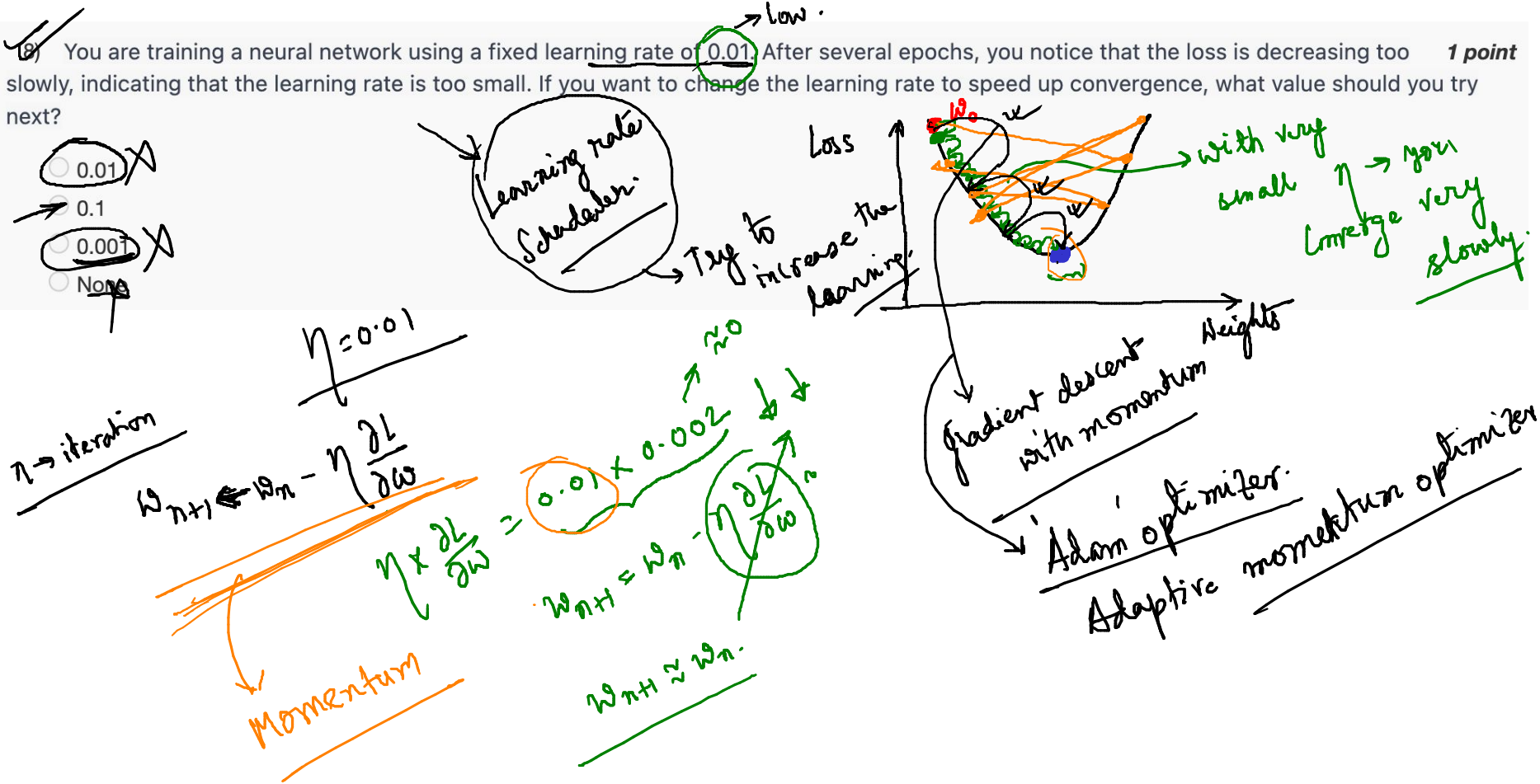
output layer response

$$Y_0 = f\left(\sum_{j=1}^{h} w_{kj} f\left(\sum_{i=1}^{d} w_{ji} x_i + w_{j0}\right) + w_{k0}\right)$$

hidden layer

(8) You are training a neural network using a fixed learning rate of 0.01. After several epochs, you notice that the loss is decreasing too slowly, indicating that the learning rate is too small. If you want to change the learning rate to speed up convergence, what value should you try next? **1 point**

- 0.01
- 0.1
- 0.001
- None



9) For a Kohonen Self-organizing map (SOM) to cluster four vectors. The online vectors to be clustered are $(1,1,0,0)$; $(0,0,0,1)$; $(1,0,0,0)$; $(0,0,1,1)$ sequentially. The maximum number of clusters to be formed is m=2. Suppose learning rate (Geometric decrease) is $\alpha(0) = 0.6$ and $\alpha(t+1) = 0.6\alpha(t)$. Initial weight matrix is $\begin{bmatrix} 0.3 & 0.7 \\ 0.4 & 0.3 \\ 0.6 & 0.6 \\ 0.8 & 0.3 \end{bmatrix}$ The first iteration for KSOM to cluster the vector $(1,0,0,0)$ will be

1 point

- $\begin{bmatrix} 0.83 & 0.88 \\ 0.42 & 0.72 \\ 0.21 & 0.24 \\ 0.12 & 0.12 \end{bmatrix}$
- $\begin{bmatrix} 0.12 & 0.952 \\ 0.16 & 0.29 \\ 0.24 & 0.096 \\ 0.92 & 0.048 \end{bmatrix}$
- $\begin{bmatrix} 0.12 & 0.88 \\ 0.16 & 0.72 \\ 0.24 & 0.24 \\ 0.92 & 0.12 \end{bmatrix}$
- $\begin{bmatrix} 0.63 & 0.35 \\ 0.18 & 0.21 \\ 0.25 & 0.15 \\ 0.61 & 0.6 \end{bmatrix}$

w_1, w_2; A; $\in R^{4\times 1}$

Step 1:→ Take the first sequential data = (1,1,0,0) ∈ $R^{4\times1}$

Step-2:→ Calculate the Eucledian distance from the given sequential data to the weight vector.

$d_1(w_1, A) = \sqrt{(1-0.3)^2 + (1-0.4)^2 + (0-0.6)^2 + (0-0.8)^2}$

$d_2(w_2, A) = \sqrt{(1-0.7)^2 + (1-0.3)^2 + (0-0.6)^2 + (0-0.3)^2}$

Step3:- Check the least distance between $d_1(w_1,A)$ & $d_2(w_2,A)$

Say, if $d_2(w_2,A) < d_1(w_1,A) \rightarrow w_2$ is best matching unit as it yields lowest distance value.

Step-4:- Now the best matching unit has to be updated now :- w_2^* (updated)

$w_2^* = w_2 + \alpha(0) \times (A - w_2)$

Repeat until reaches (1,0,0,0)

10) Suppose Mean Squared Error (MSE) is used as the loss function for training your neural network model. Your model predicts a value of 8 for a data point with an actual target value of 10. What is the MSE loss for this data point?

- 4
- 2
- 1
- 0

$$MSE(y_{true}, y_{pred}) = \frac{1}{N} \sum_{i=1}^{N} (y_{true} - y_{pred})^2$$

squared error.

Mean.

$N = 1$

Pred = 8.

True = 10.

$$MSE(10, 8) = \frac{1}{1} (10 - 8)^2$$

$$= 2^2 = 4.$$