

X

<https://swayam.gov.in>https://swayam.gov.in/nc_details/NPTEL

shweta_s@cs.iitr.ac.in ▾

NPTEL (<https://swayam.gov.in/explorer?ncCode=NPTEL>) » Deep Learning - IIT Ropar (course)

Click to register
for Certification
exam ([https://
examform.nptel.ac.in/2025_10/
exam_form/
dashboard](https://examform.nptel.ac.in/2025_10/exam_form/dashboard))

If already
registered, click
to check your
payment status

Course outline

About
NPTEL ()

How does an
NPTEL
online
course
work? ()

Week 0 ()

Week 1 ()

Week 2 ()

Week 3 ()

● Feedforward
Neural

Week 3 : Assignment 3

The due date for submitting this assignment has passed.

Due on 2025-02-12, 23:59 IST.

Assignment submitted on 2025-02-11, 11:38 IST

Use the following data to answer the questions 1 to 2

A neural network contains an input layer $h_0 = x$, three hidden layers (h_1, h_2, h_3) , and an output layer O. All the hidden layers use the Sigmoid activation function, and the output layer uses the Softmax activation function.

Suppose the input $x \in \mathbb{R}^{200}$, and all the hidden layers contain 10 neurons each. The output layer contains 4 neurons.

1) How many parameters (including biases) are there in the entire network?

No, the answer is incorrect.
Score: 0

Accepted Answers:
(Type: Numeric) 2274

1 point

2) Suppose all elements in the input vector are zero, and the corresponding true label is also 0. Further, suppose that all the parameters (weights and biases) are initialized to zero. What is the loss value if the cross-entropy loss function is used? Use the natural logarithm (ln).

No, the answer is incorrect.
Score: 0

Accepted Answers:
(Type: Range) 1.317,1.455

1 point

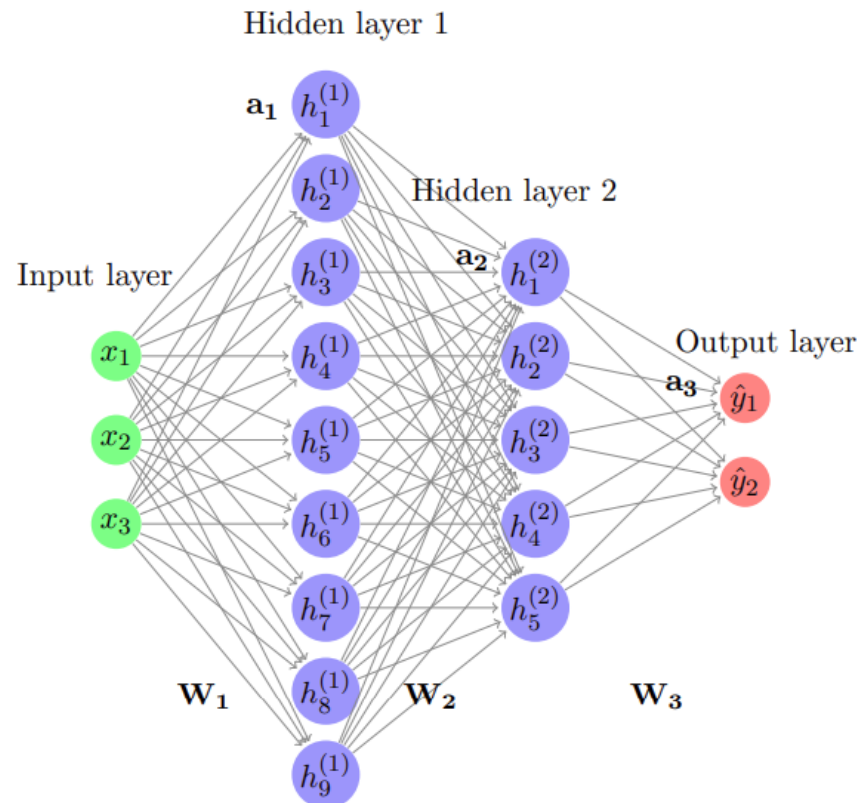
Networks

(a.k.a
multilayered
network of
neurons) (unit?
unit=46&lesso
n=47)

- Learning
Parameters of
Feedforward
Neural
Networks
(Intuition)
(unit?
unit=46&lesso
n=48)
- Output
functions and
Loss functions
(unit?
unit=46&lesso
n=49)
- Backpropagati
on (Intuition)
(unit?
unit=46&lesso
n=50)
- Backpropagati
on: Computing
Gradients w.r.t.
the Output
Units (unit?
unit=46&lesso
n=51)
- Backpropagati
on: Computing
Gradients w.r.t.
Hidden Units
(unit?
unit=46&lesso
n=52)
- Backpropagati
on: Computing
Gradients w.r.t.
Parameters
(unit?
unit=46&lesso
n=53)
- Backpropagati
on: Pseudo
code (unit?
unit=46&lesso

Use the following data to answer the questions 3 to 4

The diagram below shows a neural network. The network contains two hidden layers and one output layer. The input to the network is a column vector $x \in \mathbb{R}^3$. The first hidden layer contains 9 neurons, the second hidden layer contains 5 neurons and the output layer contains 2 neurons. Each neuron in the l^{th} layer is connected to all the neurons in the $(l+1)^{th}$ layer. Each neuron has a bias connected to it (not explicitly shown in the figure)



In the diagram, W_1 is a matrix and x , a_1 , h_1 , and O are all column vectors. The notation $W_i[j, :]$ denotes the j^{th} row of the matrix W_i , $W_i[:, j]$ denotes the j^{th} column of the matrix W_i and W_{kij} denotes an element at i^{th} row and j^{th} column of the matrix W_k .

3) Choose the correct dimensions of W_1 and a_1

1 point

- ☒ $W_1 \in \mathbb{R}^{3 \times 9}$
- ☐ $a_1 \in \mathbb{R}^{9 \times 5}$
- ☐ $W_1 \in \mathbb{R}^{9 \times 3}$
- ☐ $a_1 \in \mathbb{R}^{1 \times 9}$
- ☐ $W_1 \in \mathbb{R}^{1 \times 9}$
- ☒ $a_1 \in \mathbb{R}^{9 \times 1}$

No, the answer is incorrect.
Score: 0

Accepted Answers:

$$W_1 \in \mathbb{R}^{9 \times 3}$$

n=54)

- Derivative of the activation function (unit? unit=46&lesso n=55)
- Information content, Entropy & cross entropy (unit? unit=46&lesso n=56)
- Lecture Material for Week 3 (unit? unit=46&lesso n=57)
- Week 3: Solution (unit? unit=46&lesso n=247)
- Week 3 Feedback Form: Deep Learning - IIT Ropar!! (unit? unit=46&lesso n=186)
- **Quiz: Week 3 : Assignment 3 (assessment? name=307)**

week 4 ()**Week 5 ()****Week 6 ()****Week 7 ()****Week 8 ()****Week 9 ()****week 10 ()****Week 11 ()**

$$\mathbf{a}_1 \in \mathbb{R}^{9 \times 1}$$

4) How many learnable parameters(including bias) are there in the network?

85

No, the answer is incorrect.
Score: 0

Accepted Answers:
(Type: Numeric) 98

1 point

5) We have a multi-classification problem that we decide to solve by training a feedforward neural network. What activation function should we use in the output layer to get the best results? **1 point**

- ☐ Logistic
☐ Step function
☒ Softmax
☐ linear

Yes, the answer is correct.
Score: 1

Accepted Answers:
Softmax

6) Which of the following statements about backpropagation is true? **1 point**

- ☐ It is used to compute the output of a neural network.
☒ It is used to optimize the weights in a neural network.
☐ It is used to initialize the weights in a neural network.
☐ It is used to regularize the weights in a neural network.

Yes, the answer is correct.
Score: 1

Accepted Answers:
It is used to optimize the weights in a neural network.

7) Given two probability distributions p and q , under what conditions is the cross entropy between them minimized? **1 point**

- ☐ All the values in p are lower than corresponding values in q
☐ All the values in p are higher than corresponding values in q
☒ $p = 0$ (0 is a vector)
☐ $p = q$

No, the answer is incorrect.
Score: 0

Accepted Answers:
 $p = q$

8) Given that the probability of Event A occurring is 0.18 and the probability of Event B occurring is 0.92, which of the following statements is correct? **1 point**

Week 12 ()

Download
Videos ()

Books ()

Text
Transcripts ()Problem
Solving
Session -
Jan 2025 ()

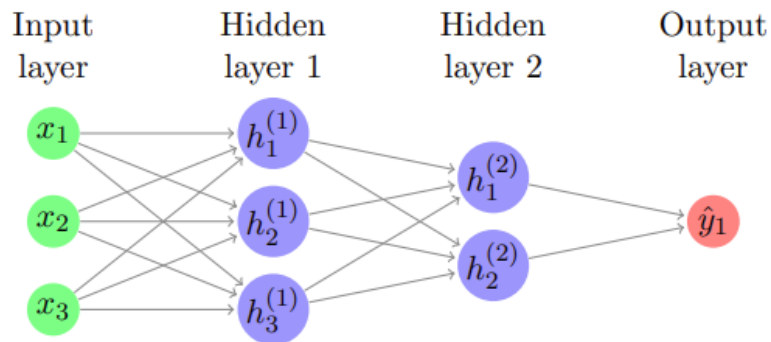
- ☒ Event A has a low information content
- ☐ Event A has a high information content
- ☒ Event B has a low information content
- ☐ Event B has a high information content

No, the answer is incorrect.
Score: 0

Accepted Answers:

*Event A has a high information content**Event B has a low information content***Use the following data to answer the questions 9 and 10**

The following diagram represents a neural network containing two hidden layers and one output layer. The input to the network is a column vector $x \in \mathbb{R}^3$. The activation function used in hidden layers is sigmoid. The output layer doesn't contain any activation function and the loss used is squared error loss $(pred_y - true_y)^2$.



The following network doesn't contain any biases and the weights of the network are given below:

$$W_1 = \begin{bmatrix} 1 & 1 & 3 \\ 2 & -1 & 1 \\ 1 & 2 & -2 \end{bmatrix} \quad W_2 = \begin{bmatrix} 1 & 1 & 2 \\ 3 & 1 & 1 \end{bmatrix} \quad W_3 = [1 \quad 2]$$

The input to the network is: $x = \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix}$

The target value y is: $y = 5$

9) What is the predicted output for the given input x after doing the forward pass?

Yes, the answer is correct.
Score: 1

Accepted Answers:

*(Type: Range) 2.9,3.0***1 point**

10) Compute and enter the loss between the output generated by input x and the true output y .

Yes, the answer is correct.
Score: 1

Accepted Answers:

(Type: Range) 3.97,4.39

1 point