X

Week 3 : Assignment 3

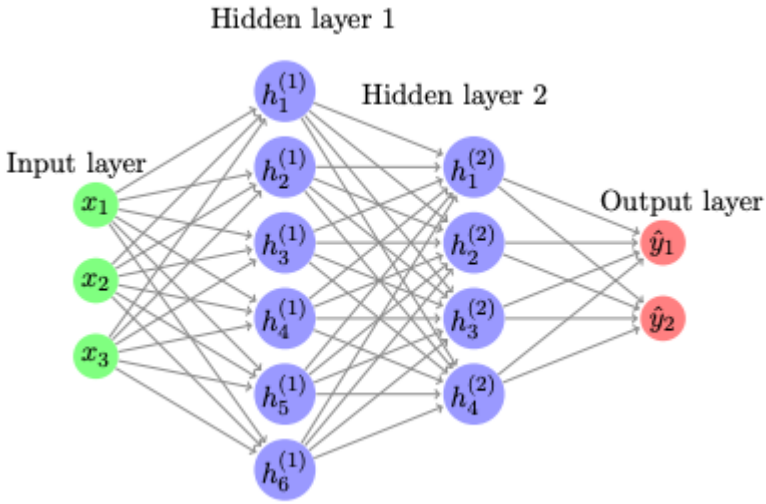
The due date for submitting this assignment has passed.

Due on 2024-08-14, 23:59 IST.

Assignment submitted on 2024-08-08, 16:02 IST

Use the following data to answer the questions 1 to 3

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 6 neurons, the second hidden layer contains 4 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, $\mathbf{W_1}$ is a matrix and $\mathbf{x}, \mathbf{a_1}, \mathbf{h_1}$, and $\mathbf{O}$ are all column vectors. The notation $\mathbf{W_i[j,:]}$ denotes the j^{th} row of the matrix $\mathbf{W_i}$, $\mathbf{W_i[:,j]}$ denotes the j^{th} column of the matrix $\mathbf{W_i}$ and $\mathbf{W_{k_{ij}}}$ denotes an element at i^{th} row and j^{th} column of the matrix $\mathbf{W_k}$.

1) Choose the correct dimensions of $\mathbf{W_1}$ and $\mathbf{a_1}$ ***1 point***

- [ ] $\mathbf{W_1} \in \mathbb{R}^{6\times 3}$
- [ ] $\mathbf{W_1} \in \mathbb{R}^{3\times 6}$
- [ ] $\mathbf{W_1} \in \mathbb{R}^{1\times 6}$
- [x] $\mathbf{a_1} \in \mathbb{R}^{6\times 1}$
- [ ] $\mathbf{a_1} \in \mathbb{R}^{1\times 6}$
- [ ] $\mathbf{a_1} \in \mathbb{R}^{1\times 1}$

Yes, the answer is correct.
Score: 1

Accepted Answers:
$\mathbf{W_1} \in \mathbb{R}^{6\times 3}$
$\mathbf{a_1} \in \mathbb{R}^{6\times 1}$

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

62

Yes, the answer is correct.
Score: 1

Accepted Answers:
(Type: Numeric) 62

1 point

3) Which of the following loss functions can be used for the classification problem? ***1 point***

- ○ Means Squared error
- ◉ Cross entropy
- ○ Both Mean Squared Error and Cross Entropy loss can be used

No, the answer is incorrect.
Score: 0

Accepted Answers:
Both Mean Squared Error and Cross Entropy loss can be used

4) 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

5) We have data x with the following labels y = [‘car’, ‘bus’, ‘bike’, ‘car’, ‘bus’, ‘car’, ‘bus’, ‘bike’, ‘car’]. Which of the following distribution will give the lowest cross-entropy loss with y? (Distribution is given in the following order [‘car’, ‘bus’,‘bike’])? ***1 point***

- $[0.49, 0.30, 0.21]$
- $[0.44, 0.33, 0.22]$
- $[0.50, 0.30, 0.20]$
- $[0.44, 0.39, 0.19]$

Yes, the answer is correct.
Score: 1

Accepted Answers:
$[0.44, 0.33, 0.22]$

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 initialize the weights in a neural network.
- It is used to optimize 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 lower than corresponding values in q
- $p = 0$(0 is a vector)
- $p = q$

Yes, the answer is correct.
Score: 1

Accepted Answers:
$p = q$

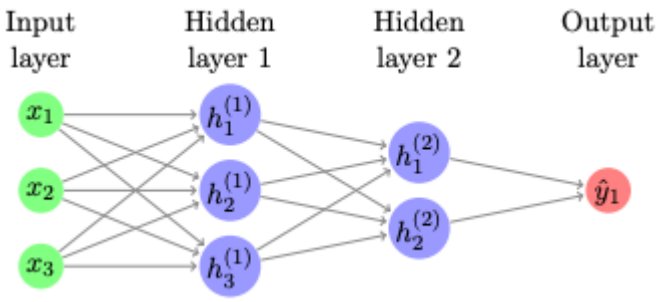
8) Given that the probability of Event A occurring is 0.80 and the probability of Event B occurring is 0.20, which of the following statements is correct? ***1 point***

- [x] Event A has a low information content
- [ ] Event A has a high information content
- [ ] Event B has a low information content
- [x] Event B has a high information content

Yes, the answer is correct.
Score: 1

Accepted Answers:
Event A has a low information content
Event B has a high 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 $x \in R^3$. 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:

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

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

The target value y is: $\mathbf{y} = 8$

9) What is the predicted output for the given input x after doing the forward pass? (Choose the option closest to your answer)

4.8949

Yes, the answer is correct.
Score: 1

Accepted Answers:
(Type: Range) 4.7,5.2

1 point

10) Compute and enter the loss between the output generated by input $\mathbf{x}$ and the true output $\mathbf{y}$

9.6478

Yes, the answer is correct.
Score: 1

Accepted Answers:
(Type: Range) 8.82,9.74

1 point