

DEEP LEARNING WEEK 8

1. We have observed that the sigmoid neuron has become saturated. What might be the possible values of output at this neuron? (MSQ)

- a)0.96
- b)0.05
- c)0.65
- d)0.35

Answer: a), b)

Solution: Since the neuron has saturated its output values are close to 0 or 1.

2. We train a feed-forward neural network and observe that the weights learned are all equal for a single neuron. What are the possible causes of this problem? (MSQ)

- a)Weights were initialized randomly.
- b)Weights were initialized to high values.
- c)Weights were initialized to equal values.
- d)Weights were initialized to zero.

Answer: c),d)

Solution: All weights remain equal for a single neuron due to the symmetry breaking problem which is caused by equal initialization of weight

3. Which of the following functions produces zero-centered outputs?

- a)Relu
- b)Sigmoid
- c)Tanh
- d) None of These

Answer: c)

Solution: Sigmoid is centered around 0.5 while is not centered around any real number. Tanh is centered around 0.

4. Which of the following is a common solution to the symmetry-breaking problem?

- a) Increasing the learning rate
- b) Reducing the learning rate
- c) Adding more layers to the network
- d) Randomly initializing the weights

Answer: d) Randomly initializing the weights

Solution: One common solution to the symmetry-breaking problem is to randomly initialize the weights in the network. This can break the symmetry and allow the network to learn complex features and patterns in the data.

5. How can you tell if your network is suffering from the Dead ReLU problem?

- a) The loss function is not decreasing during training
- b) The accuracy of the network is not improving
- c) A large number of neurons have zero output
- d) The network is overfitting to the training data

Answer: c) A large number of neurons have zero output

Solution: The Dead ReLU problem can be detected by checking the output of each neuron in the network. If a large number of neurons have zero output, then the network may be suffering from the Dead ReLU problem. This can indicate that the bias term is too high, causing a large number of dead neurons.

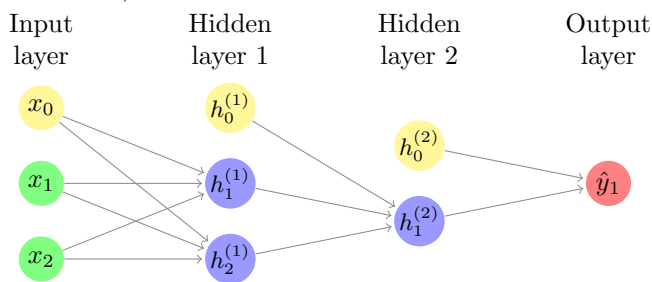
6. Which of the following is incorrect concerning the batch normalization process in neural networks?

- a) We normalize the output produced at each layer before feeding it into the next layer
- b) Variance and mean are not learnable parameters.
- c) Batch normalization leads to faster convergence
- d) Backpropagation can be used after batch normalization

Answer: b)

Solution: The network is allowed to adjust variance and mean by making them a learnable parameter.

7. We observe that in the following neural network with Relu as the activation function, weights in the first hidden layer and input layer are not getting updated. Choose the correct option for the value of bias b going to 2nd hidden layer. (Yellow nodes represent the bias and inputs are normalized)



- a) $b=20$
- b) $b=0$
- c) $b=-0.4$
- d) $b=-15$

Answer: d)

Solution: Since the weights are not getting updated our neuron is dead which is caused due to high negative bias.

8. Which of the following methods can help to avoid saturation in deep learning?

- a) Using a different activation function.
- b) Increasing the learning rate.
- c) Increasing the model complexity
- d) All of the above.

Answer: a)

Solution: Using a different activation function such as ReLU can avoid saturation.

9. Which of the following is true about the role of unsupervised pre-training in deep learning?

- a. It is used to replace the need for labeled data
- b. It is used to initialize the weights of a deep neural network
- c. It is used to fine-tune a pre-trained model
- d. It is only useful for small datasets

Answer: b

Solution: Unsupervised pre-training is used to initialize the weights of a deep neural network before being fine-tuned on labeled data

10. Which of the following statements is true with respect to exponential Relu given by $ae^x - 1$ if $x \leq 0$ and x if $x \geq 0$? ($a \neq 1$)
- a) It is discontinuous at 0.
 - b) It is non-differentiable at 0.
 - c) is less computationally expensive than Relu
 - d) Exponential Relu can have negative values.

Answer:d)

Solution: Exponential Relu can have negative values and since $a \neq 1$, it is non-differentiable at 0.