Mythreyi R

Subject Code: MDS 6304

Subject Name: Deep Learning Principles and Applications

Segment 3: Shallow Neural Network

1. Which one of the following statements is not true about a softmax activation layer in a zero hidden layer neural network architecture?
2. It converts raw scores to probabilities
3. It is fully connected to the dense layer before
4. It is node-wise connected to the dense layer before
5. It has the same number of nodes as the dense layer before

Correct Answer: c. It is node-wise connected to the dense layer before

1. Which one of the following activation functions does not have the vanishing gradient problem?
2. Sigmoid
3. tanh
4. ReLU
5. Leaky ReLU

Correct Answer: c. ReLU

1. Which one of the following activation functions clips all negative raw scores to zeros?
2. Sigmoid
3. tanh
4. ReLU
5. Leaky ReLU

Correct Answer: c. ReLU

1. Which one of the following activation functions results in mean-centered activated values?
2. Sigmoid
3. tanh
4. ReLU
5. Leaky ReLU

Correct Answer: b. tanh

1. Suppose we are dealing with a classification problem in which a sample can belong to one of 5 possible output categories labeled from 0 through 4 (Python indexing style). Which one of the following is the correct one-hot encoded representation for a sample with output label 3?
2. [0, 0, 1, 0, 1]
3. [0, 0, 0, 0, 1]
4. [0, 0, 1, 0, 0]
5. [0, 0, 0, 1, 0]

Correct Answer: d. [0,0,0,1,0]

1. The CCE loss for a sample with correct probability vector [0, 1, 0] and predicted probability vector [0.75, 0.15, 0.1] is
2. -log(0.75)
3. -log(0.1)
4. -log(0.15)
5. -log(0.85)

Correct Answer: c. -log(0.15)

1. Suppose we apply a softmax classifier for a classification problem with 10 possible output categories. What is the shape of the gradient ?
2. 11 x 11
3. 10 x 10
4. 9 x 9
5. 10 x 11

Correct Answer: d. 10 x 11

1. Suppose we use a single hidden layer neural network for a classification problem with 10 features per sample. If there are 6 nodes in the dense layer of the hidden layer, what is the shape of the weights matrix  if we ignore the bias feature?
2. 6 x 10
3. 6 x 11
4. 10 x 6
5. 11 x 6

Correct Answer: a. 6 x10

1. What is the shape of the gradient  for a 5 x 10-matrix  and a 10-vector ?
2. 5 x 10 x10
3. 10 x 5 x 10
4. 5 x 10
5. 10 x 10 x 5

Correct Answer: c. 5x10

1. Which one of the following is not a hyperparameter for a neural network?
2. Weights
3. Learning rate
4. Regularization strength
5. Number of nodes in the hidden layer

Correct Answer: a. Weights