DEEP LEARNING – WORKSHEET 3

1. Which of the following is true about model capacity (where model capacity means the ability of neural network to approximate complex functions)?

Answer: A) As dropout ratio increases, model capacity increases

2. Batch Normalization is helpful because?

Answer: C) It normalizes (changes) all the input before sending it to the next layer

3. What if we use a learning rate that’s too large?

Answer: C) either A or B

4. What are the factors to select the depth of neural network?

Answer: C) 1, 3, 4, 5

5. Suppose you have inputs as x, y, and z with values -2, 5, and -4 respectively. You have a neuron ‘q’ and neuron ‘f’ with functions: q = x + y f = q \* z Graphical representation of the functions is as follows: What is the gradient of F with respect to x, y, and z? (use chain rule of derivatives to find the solution)

Answer: C) (-4, -4, 3)

6. Which of the following statement is the best description of early stopping?

Answer: B) Simulate the network on a test dataset after every epoch of training. Stop training when the generalization error starts to increase

7. Which gradient descent technique is more advantageous when the data is too big to handle in RAM simultaneously?

Answer: A) Mini Batch Gradient Descent

8. Consider the scenario. The problem you are trying to solve has a small amount of data. Fortunately, you have a pre-trained neural network that was trained on a similar problem. Which of the following methodologies would you choose to make use of this pre-trained network?

Answer: B) Assess on every layer how the model performs and only select a few of them

9. Which of the following neural network training challenge can be solved using batch normalization?

Answer: A) Overfitting

10. For a binary classification problem, which of the following activations may be used in output layer?

Answer: A) ReLU B) sigmoid C) softmax

11. What will happen if we do not use activation function in artificial neural networks?

Answer: Activation functions are really important for a Artificial Neural Network to learn and make sense of something really complicated and Non-linear complex functional mappings between the inputs and response variable. They introduce non-linear properties. Their main purpose is to convert a input signal of a node in a A-NN to an output signal.

If we do not apply a Activation function then the output signal would simply be a simple linear function.A linear function is just a polynomial of one degree. a linear equation is easy to solve but they are limited in their complexity and have less power to learn complex functional mappings from data. A Neural Network without Activation function would simply be a linear regression Model, which has limited power and does not perform good most of the times.

12. How does forward propagation and backpropagation work in deep learning?

Answer: