DEEPLEARNING(worksheet-2)

1. (C)
2. (A)
3. (C)
4. (B)
5. (C)
6. (B)
7. (D)
8. (B)
9. (A,B)
10. (D)
11. Y=w1x1+w2x2+w3x3
12. .If the learning rate is too small then the training process becomes very slow as it does very small updates to the weights in the network and for the best value, we need too many iterations.

If the learning rate is too high, we may skip the optimal solution and can cause the undesirable divergent behaviour in loss function.

1. .y=w1x1+w2x2+w3x3, y=w0 + w1x1 + w2x2 + w3x3 + …. wnxn
2. .In a network of n hidden layers, n derivatives will be multiplied together. If the derivatives are large then the gradient will increase exponentially as we propagate down the model until they eventually explode, and this is what we call the problem of exploding gradient. Alternatively, if the derivatives are small then the gradient will decrease exponentially as we propagate through the model until it eventually vanishes, and this is the vanishing gradient problem.
3. .Epochs: one epoch is basically when entire dataset is passed once through forward and backward propagation of the neural network. As we increase the number of epochs the greater number of times the weight is changed in the neural network and the curve goes from underfitting to optimal to overfitting curve. So, the right number of epochs is that who gives the highest accuracy of the model.

Batch: Batch is basically, we grouped a set of training examples. Number of groups we set is equal to our batch size. It is more like for loop which iterates over and over to give the desired output that we need.

Iteration: Iteration is defined as the number of batches needed to complete one epoch. For one epoch, number of batches is equal to the number of iterations. To obtain the number of iterations, we have to specify the batch size and the number of epochs for a learning algorithm.