

Course Code: CS4045	Course Name: Deep Learning with Perception
Instructor Name: Dr. Jawwad A Shamsi	
Student Roll No:	Section No:

Time: 60 minutes.

Max Marks: 25 points

Question 1: 5 points

Design a Neural network based classifier which can be used to design an electrical circuit. The classifier takes three inputs (three phase wires) and outputs positive if any of the wire has live current. Your design should include a diagram with relevant weights and biases, no. of neurons at each layer, activation function, and no. of hidden layers. Explain the complete forward propagation path.

Solution:

This is an example of OR gate. The neural network will consist of a single perceptron, which will take three inputs and estimate y .

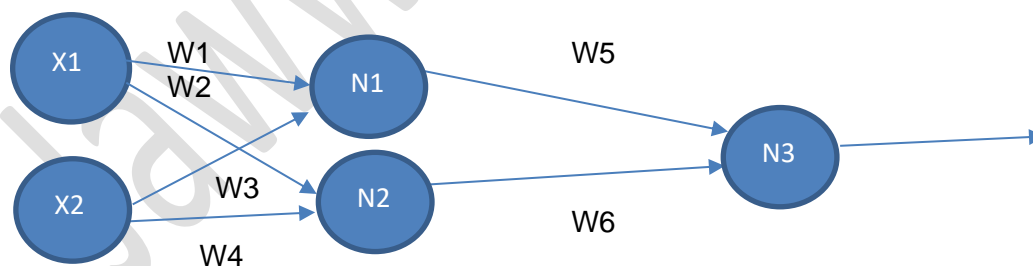
You have to find values of three weights, one for each input and the of bias.

You can derive these values using different cases such as (0,0,0), (0,0,1) and so on. Please note that there will be similarities between some cases e.g. (0,0,1) is same as (0,1,0) and (1,0,0). Similarly, (1,1,0) is same as (1,0,1) and (0,1,1).

One possible solution is (2,2,2,-1). However, other combinations are also possible. You may use lecture as a reference

Question 2: 5 points

Consider the below figure as a NN-based classifier. Explain the backward propagation process. Write proper equations. Describe the optimization algorithm and the choice of activation function.



Solution:

In this question, you will have to apply chain rule to derive different equations. This has been well explained in the lecture notes.

In addition, you will specify optimization algorithm. This could be Gradient Descent or any of its variation such as either Stochastic Gradient Descent or Batch Gradient Descent

Question 3: 5 points

Design a CNN which can detect occurrence of a rectangle in an image. The input images are in grey scale. You may assume an input image of size of 10x10.

- a) Describe no. of layers, actions performed at each layer, no. filters, no. of parameters to be trained at each layer, pooling, activation function.

This answer will depend upon different hyper parameters such as the no. of filters, size of filters, layers etc. However, you have to explain the process through a diagram. Explain the pooling process as well. No. of parameters to be learned will depend upon filter size and no. of filters.

- b) You should also describe appropriate values of filter values in the first hidden layer which are to be used to detect a rectangle and explain the process of forward propagation

Values in filters should be able to detect vertical as well as horizontal edges. As a rectangle contains both. A possible value is shown below:

-1	0	1
-1	0	1
-1	0	1

-1	-1	-1
0	0	0
1	1	1

If sum of all weights are not equal to 1 then the resultant feature map may be lighter or darker.

Question 4: Short Answers 10 points (2 points each)

- a) What Challenges exist when you solve Q3 with an MLP

Solution: MLP will have large no. of features in the input. This will make it challenging to extract features. CNN are good for feature extraction. Further, Spatial information will not be able to be captured in MLP.

- b) Differentiate choice of activation functions for multi-label and multi-class problems

For multi-label we will use sigmoid as sum of all probabilities is greater than 1. For multi-class, we will use softmax as sum of all probabilities is equal to 1.

- c) Explain the vanishing gradient problem

During the training process, weights are adjusted according to the slope. When the error is low, the derivative of the error is further low (approaching zero). This scenario leads to vanishing gradient as it becomes challenging to adjust weights when the gradient is low.

- d) Explain how would you decide number of perceptrons in the input as well as the output layers of a DNN

Number of perceptrons at the input are related to no. of input features, whereas no. of perceptron at the output are related to problem. For instance, for a multi-class /multi-label problem no. of output perceptrons are equal to no. of labels and no. of classes.

- e) How would you avoid overfitting in a DNN

Overfitting occurs when the model tries to remember the values. This could happen due to multiple reasons. One possible reason is lack of data. There could be other factors as well. Some possible solutions are mentioned below:

- Data Augmentation
- Creating more data through Synthesis
- Stop the training process when training error is decreasing and validation error is not decreasing. In such a case hyper parameters can be changed further.