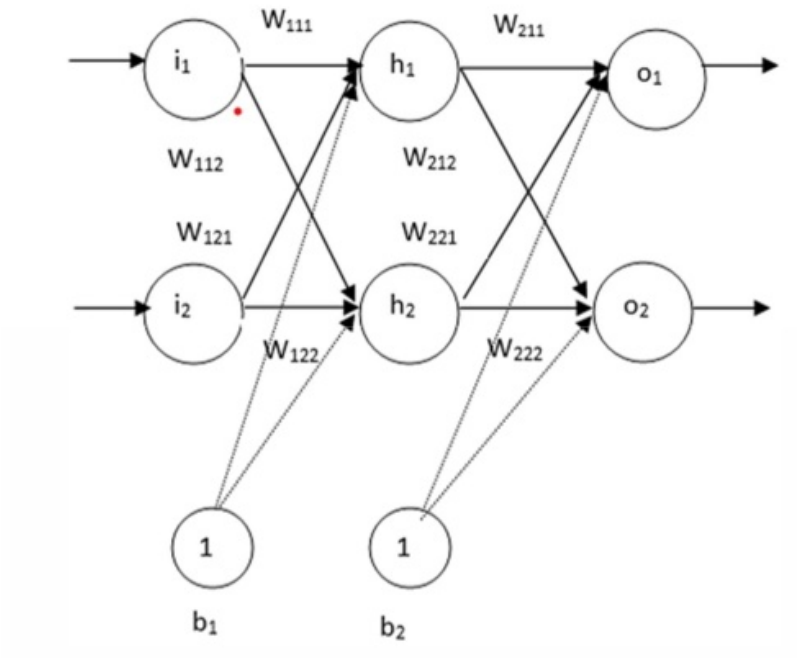
1. CDT402-Deep Learning for Data Science

2. **Module 1**
3.

1. Consider a 3 -dimensional input x = (x 1 ,x 2 ,x 3) = (2,2,1) fully connected with weights (0.5,0.3,0.1) to one neuron which is in the hidden layer with sigmoid activation function. Calculate the output of the hidden layer neuron.
2. Illustrate the limitation of a single layer perceptron with an example.
3. Consider a fully connected neural network with one hidden layer and two output nodes. The network takes a two-dimensional input vector x = [x1, x2] as input, and activation function used in both the hidden and output layers is the sigmoid activation function. The weight and biases of the network are initialized randomly as follows:

7.



8.
9.
10.
11.
12.
13.
14.
15.
16.
17.
18.
19.
20.
21.
22.
23.
24.

W1 = [[w111 = 0.2, w112 = -0.4], [w121 = 0.1, w122 = -0.1]], b1 = [-0.3, 0.4]
W2 = [w211 = 0.3, w212 = 0.1], [w221 = - 0.2, w222 = 0.2]], b2 = [-0.2, 0.1]
Consider a training example with input x = [1,2], and target output y =[0.9,0.1]. Compute the gradient with respect to one of the weights w 222 of the network using back propagation, assuming a mean squared error loss function. (Take η=0.5)

4. With an example of classification problem, explain the following terms:
   i) Hyper parameters ii) Training set iii) Validation sets iv) Bias.
5. Compare overfitting and underfitting. How can it affect model generalization?
6. You are training a fully connected neural network with 4 hidden layers, each containing 9 hidden units. The input dimension is 15, and the output is a scalar. The network includes bias terms for each neuron in both the hidden and output layers. Calculate the total number of trainable parameters in the network.

7. Design a neural network with two hidden layers having single neuron (using ReLU activation) and an output neuron to approximate a function $f(x)=x2-4x+4$ as accurately as possible in the range $x \in [0,5]$.
8. Specify the advantages of ReLU over sigmoid activation function.
9. Suppose you have a 3-dimensional input x = (x1, x2, x3) = (2, 2, 1) fully connected with weights (0.5, 0.3, 0.2) to one neuron which is in the hidden layer with sigmoid activation function. Calculate the output of the hidden layer neuron.
10. Consider the case of the XOR function in which the two points {(0, 0), (1, 1)} belong to one class, and the other two points {(1, 0), (0, 1)} belong to the other class. Design a multilayer perceptron for this binary classification problem.
11. Sketch the typical learning curves for the training and validation sets, for a setting where overfitting occurs at some point. Assume that the training set and the validation set are of the same size.
12. Can a single-layer perceptron successfully represent the Boolean AND, OR, and XOR functions? Justify your answer, and if there are any limitations, suggest possible ways to overcome them.
13. What advantages does a deep feedforward network offer over shallow networks?
14. What is the purpose of a validation set in neural network training?
15. Explain the terms overfitting and underfitting in the context of neural networks.
16. Explain back propagation algorithm for neural network training.
17. How does bias and variance trade-off affect machine learning algorithms?
18. Explain the significance of loss function in a deep learning algorithm.