Pattern Recognition and Machine Learning

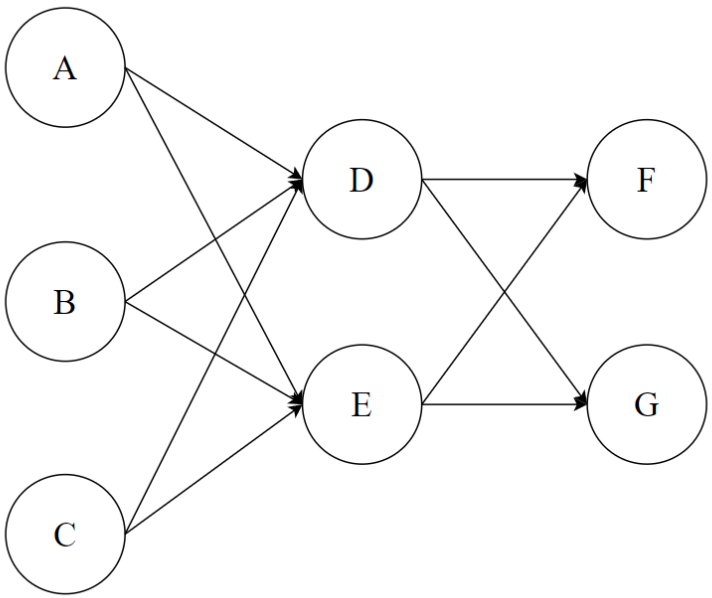
Tutorial on Artificial Neural Networks (ANN)

29^{th} October 2021

Question 1

Consider the following Artificial Neural Network (ANN). Answer the following:

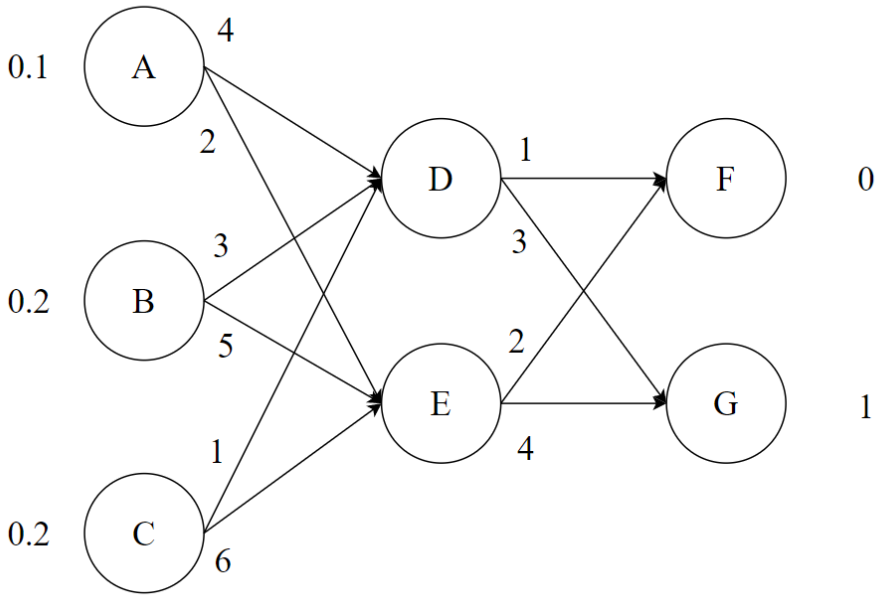
1. Identify the input nodes, hidden nodes, and output nodes.
2. How many parameters are to be learnt in this ANN? (exclude the bias terms)



Question 2.1

In the following figure, the ANN is populated with the input, weights corresponding to all the connections, and the expected output for a **classification** task. Answer the following (exclude the bias terms):

1. What values are input to the nodes D and E?

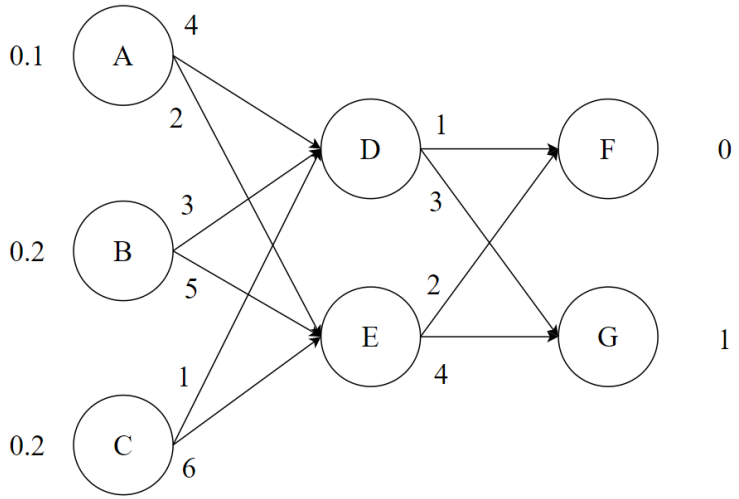


Question 2.2

In the following figure, the ANN is populated with the input, weights corresponding to all the connections, and the expected output for a **classification** task. Answer the following (exclude the bias terms):

2. Assuming the sigmoid activation function at nodes D and E, what are the outputs of these nodes?

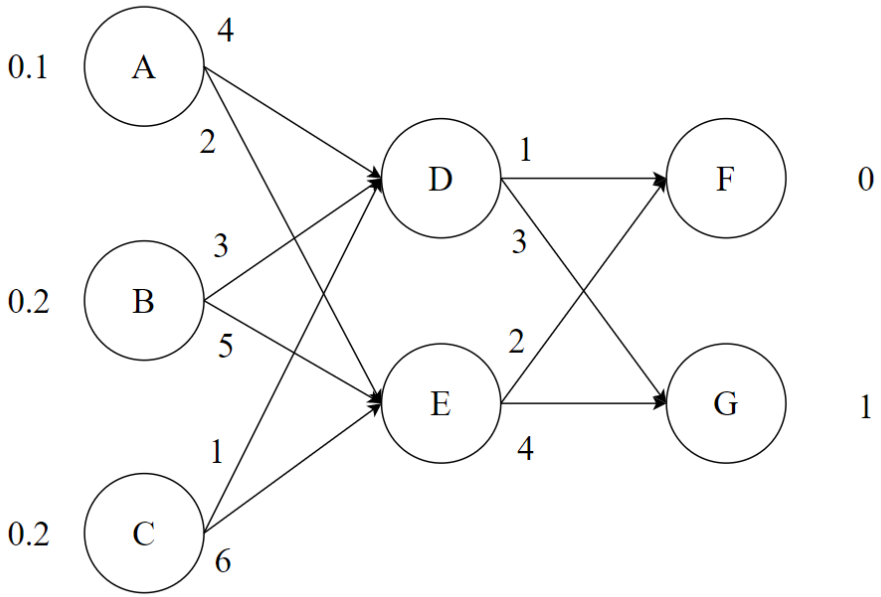
$$\sigma(x) = \frac{1}{1 + e^{-x}}$$



Question 2.3

In the following figure, the ANN is populated with the input, weights corresponding to all the connections, and the expected output for a **classification** task. Answer the following (exclude the bias terms):

3. What values are input to the nodes F and G?

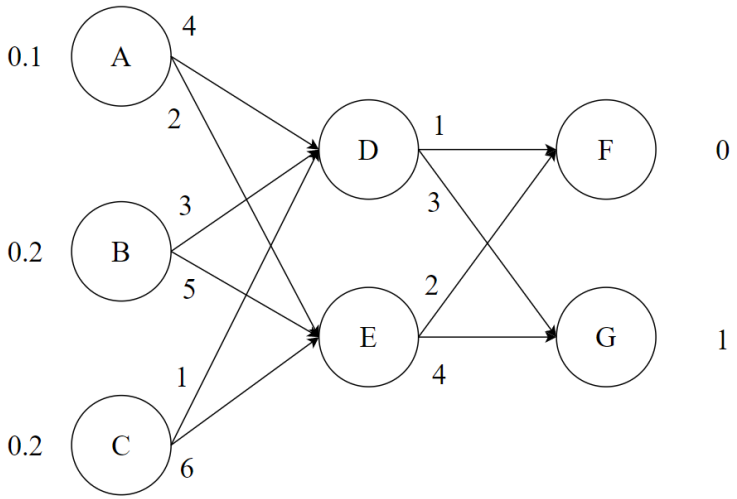


Question 2.4

In the following figure, the ANN is populated with the input, weights corresponding to all the connections, and the expected output for a **classification** task. Answer the following (exclude the bias terms):

4. Assuming the softmax activation function at nodes F and G, what are the outputs of these nodes?

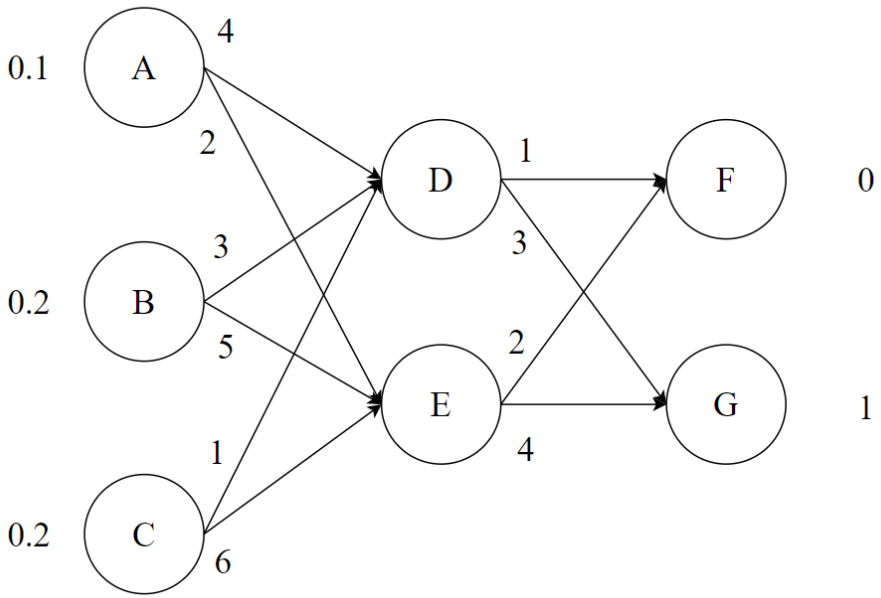
$$softmax(\vec{x})_i = \frac{e^{x_i}}{\sum_k e^{x_k}}$$



Question 2.5

In the following figure, the ANN is populated with the input, weights corresponding to all the connections, and the expected output for a **classification** task. Answer the following (exclude the bias terms):

5. What loss function will be suitable in this case? What is the loss value?



Question 3

Can the computation in the previous question be done using matrix multiplication?

Question 4

Using the insights drawn from the previous question, can you argue that if no non-linear function (such as σ and $softmax$) is used, the neural network will not learn a non-linear function?

Question 5

Express the derivative of $\sigma(x)$ in terms of itself.

$$\sigma(x) = \frac{1}{1 + e^{-x}}$$

Question 6

Express the derivative of $tanh(s)$ in terms of itself.

$$tanh(x) = \frac{e^x - e^{-x}}{e^x + e^{-x}}$$

Question 7

What is the relation between $\sigma(x)$ and $tanh(x)$?