Tutorial 4

1. what are the parameters of a neural network, i.e. which parts of a neural network are adjusted to learn?
2. What are the hyper-parameters of a simple feed-forward neural network?
3. What is the difference between a loss function and a cost function?
4. Given a set of inputs **x1...xn**, a unit has a set of corresponding weights **w1...wn** and a bias **b**. So the weighted sum **z** can be represented as?
5. What is the role of bias in neural network
6. Consider following statements
    i. In the input, we have a features vector of dimension 1*768.
    ii. The number of hidden units (neurons) is 15.
    iii. The number of neurons at the output layer is 10.

    Answer the following question based on the above statement

    A. What will be the size of the input layer?
    B. What will be the dimension of the weights vectors
        a. between the input layer and hidden layer
        b. between the hidden layer and output layer

7. Given a string of words $w_1w_2w_3$ ... w_n what do **language models** (LMs) tell us?

8. Why do we take the squared loss (i.e. $Y = (Y_t-Y_p)^2$)?

9. Is it necessary to include learning rate and dropout when we save our best-trained model?