ODAY BANI AHMAD

A20344588

Q1: Explain briefly how this is different   
from the update rule for the network trained for binary classification using log loss.

Answer:

Q2 : 1: SoftMax : because the model predicts a multinomial probability.

Q2 : 2: 10 : based on the number of multinomial probability.

Q2: 3: Report of the average MSE loss and the accuracy. Calculated in the python code with the

plot.

Q2 :4: Plot the loss and accuracy as a function of the number of iterations. In the python code

Q2: 5: too small may result in a long training process that could get stuck, whereas a value too large may result in learning a sub-optimal set of weights too fast or an unstable training process.

Q2:6: large number of neurons in the hidden layers can increase the time it takes to train the network. The amount of training time can increase

A: it will not change but the amount of work and iteration will increase.

B:I notesed that the time increased and the loss function become smaller as will as the predction inhanced and the error minimized.

Q7: A: A neuron's activation status is determined by an activation function. By employing simpler mathematical procedures, it will determine whether or not the neuron's input to the network is significant during the prediction process.

B: No it will not affect the update rule

C: in python code all the observation.