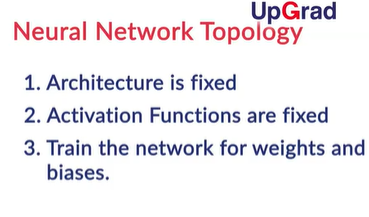
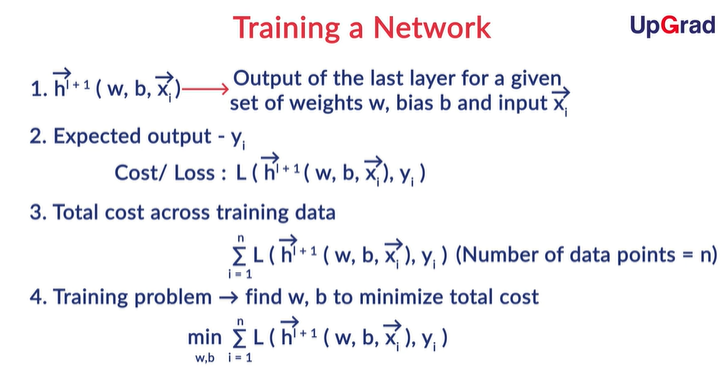
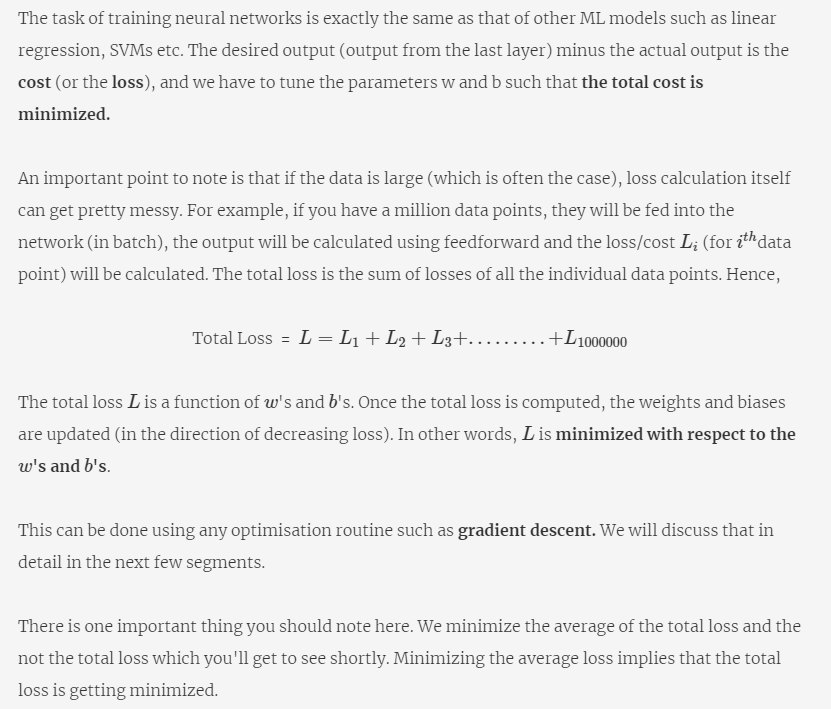
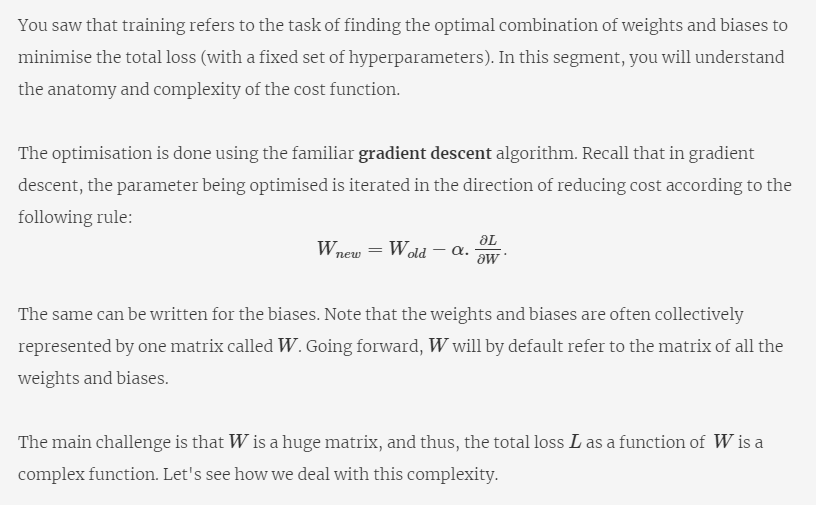
**Back Propogation**

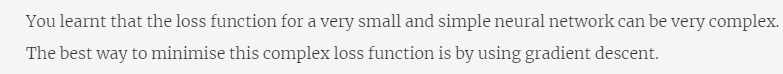




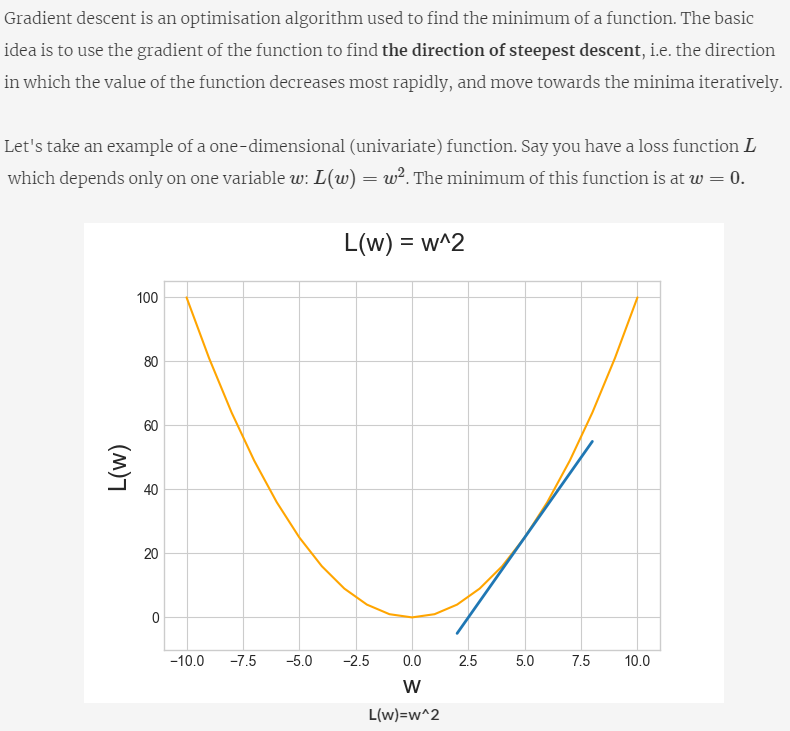


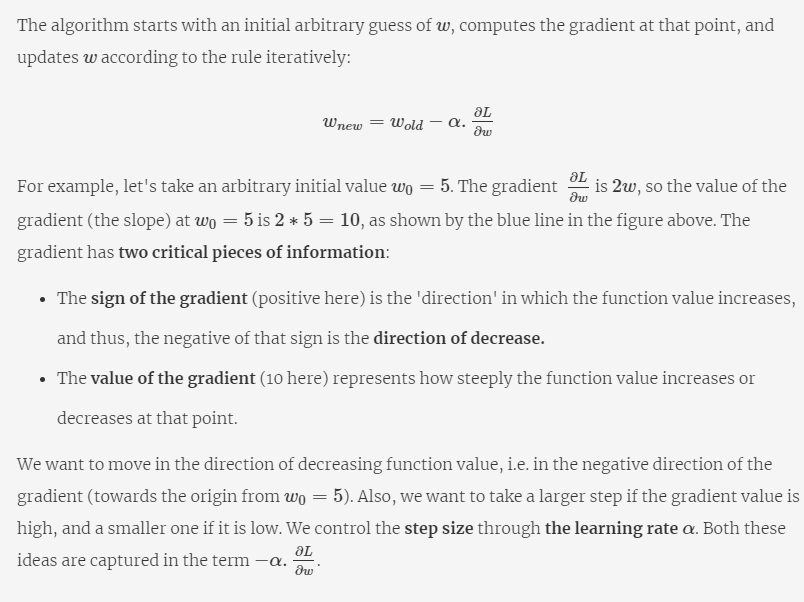
# Complexity of the Loss Function

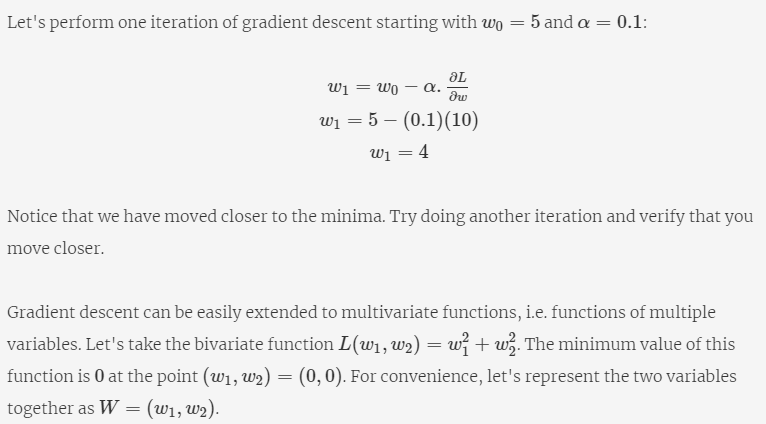


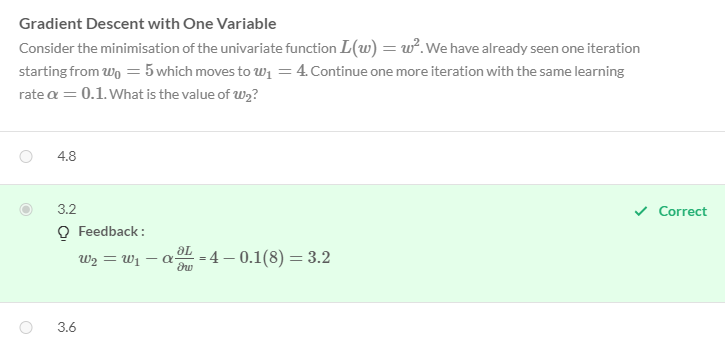


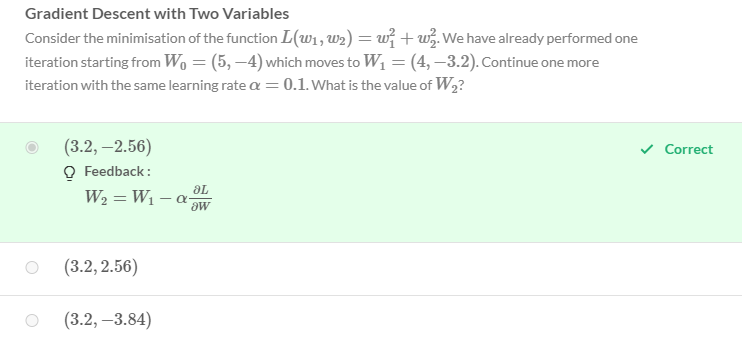
## **Comprehension - Gradient Descent**

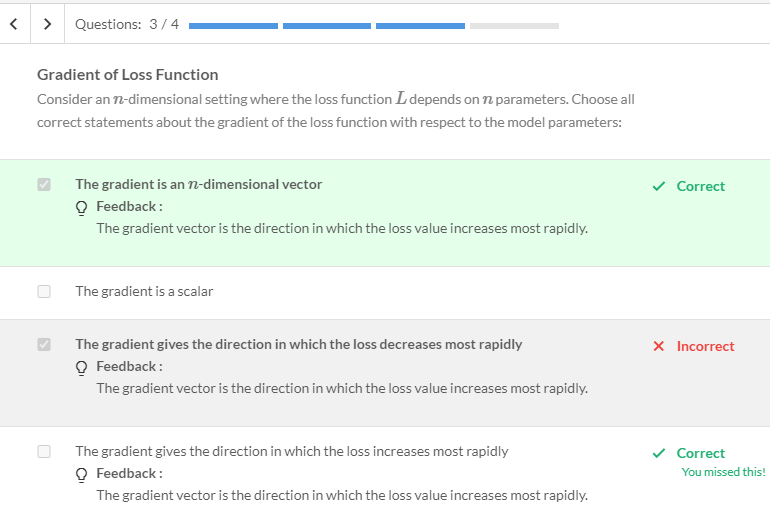


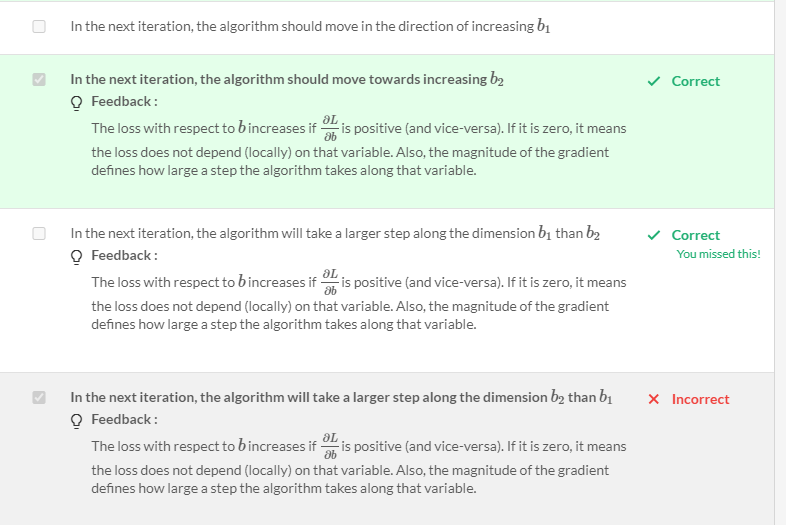
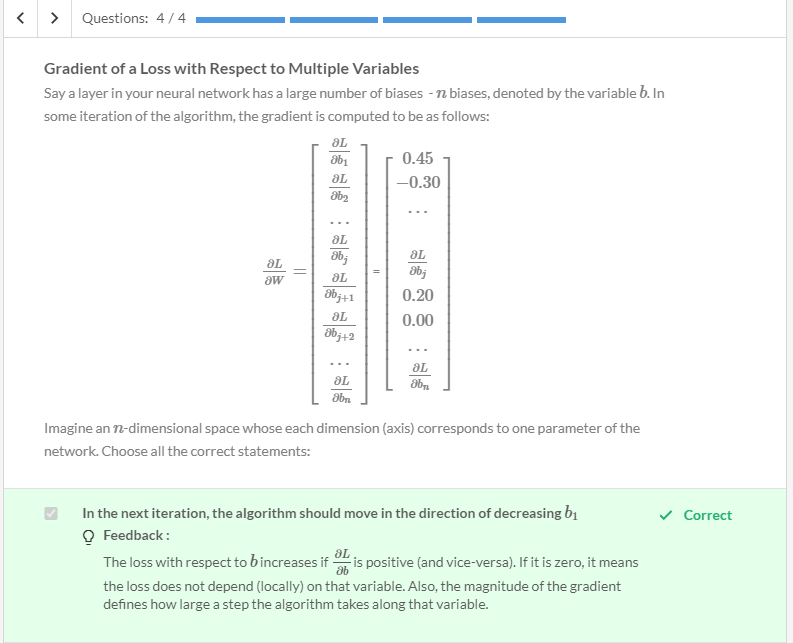




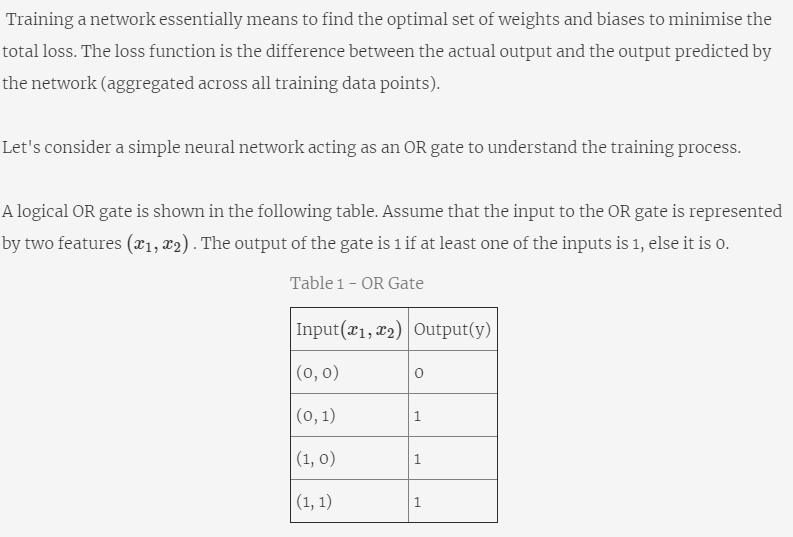


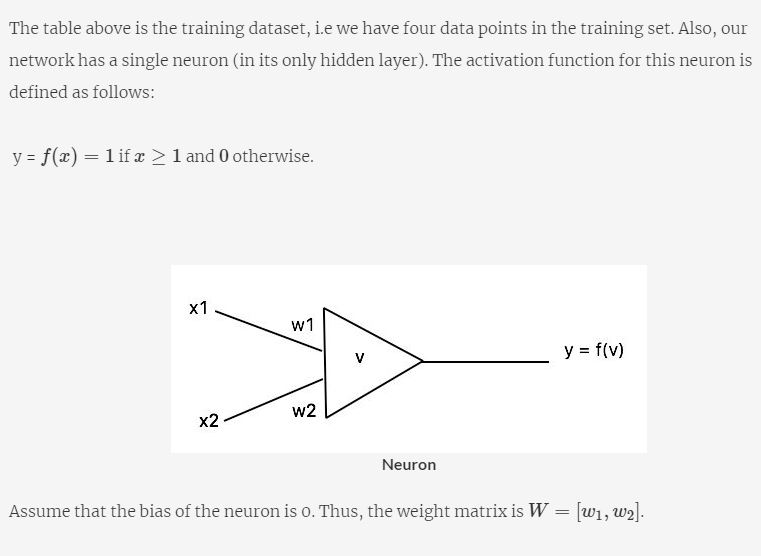


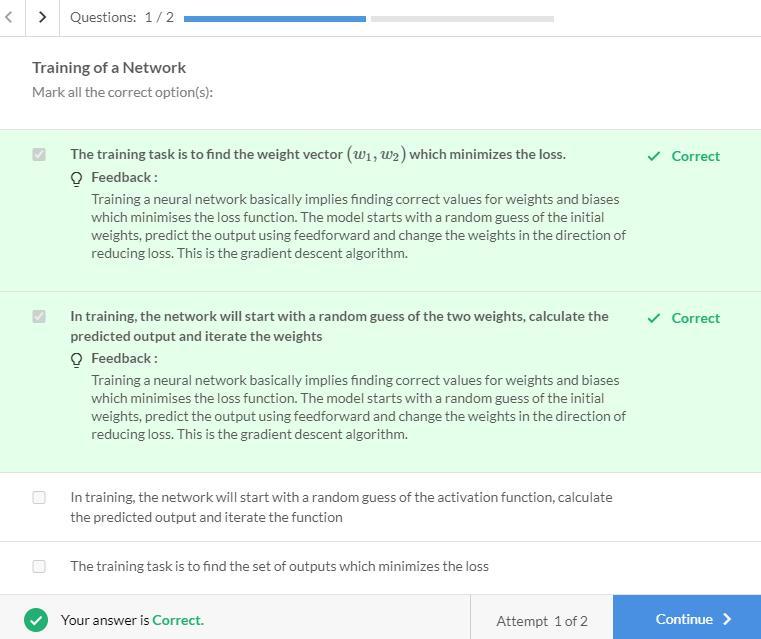


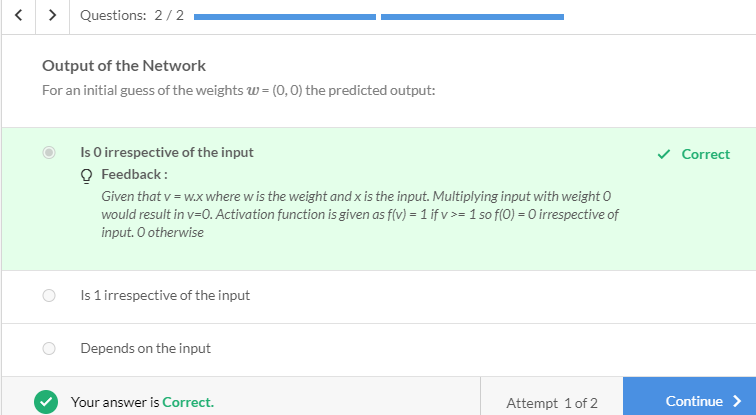


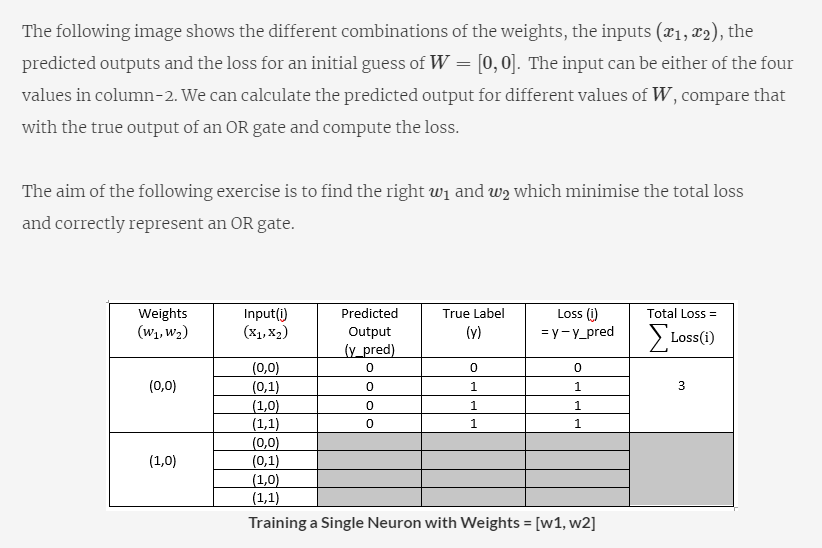
# Comprehension - Training a Neural Network

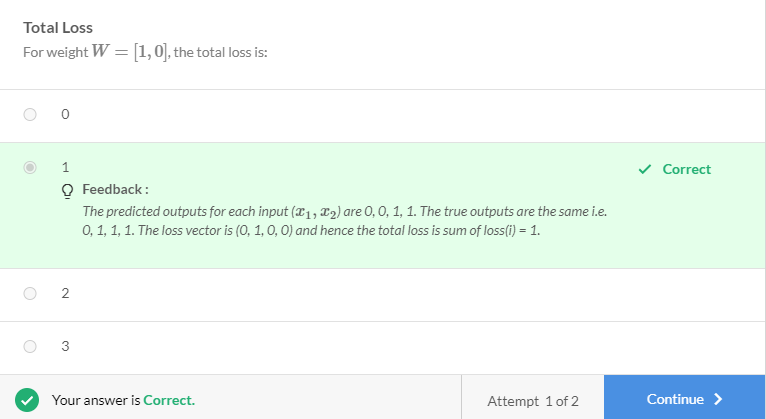


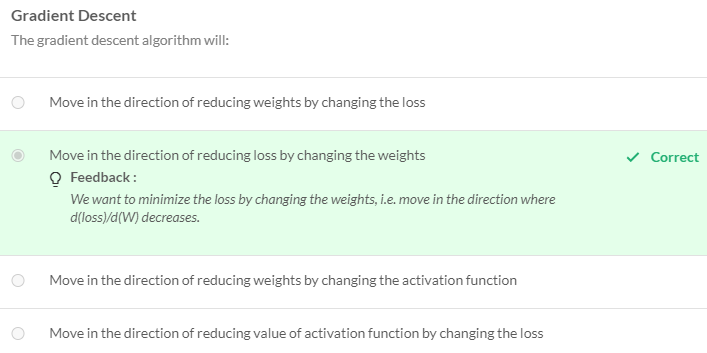


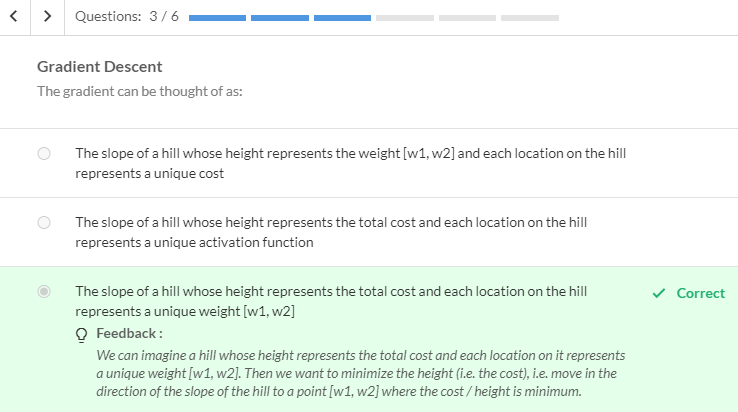


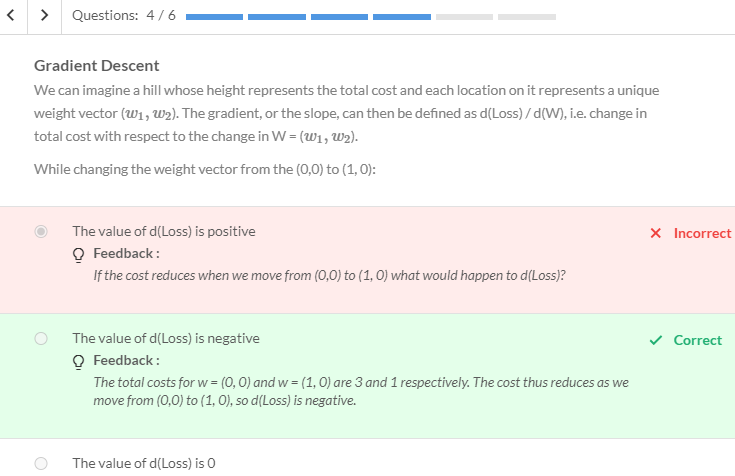


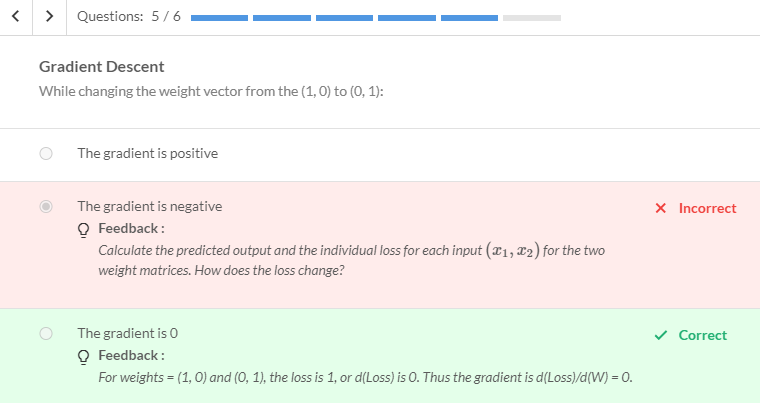


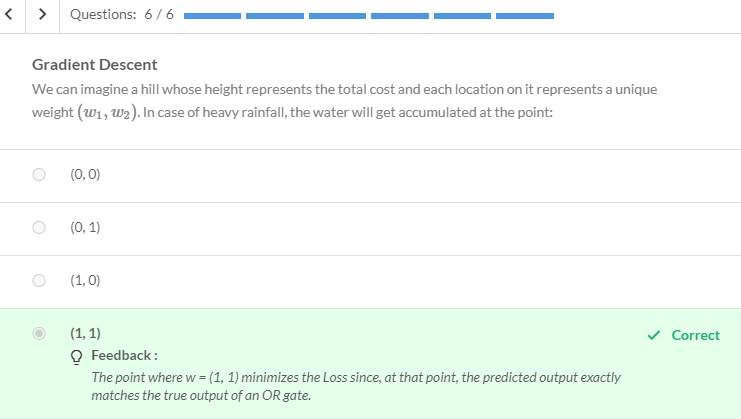




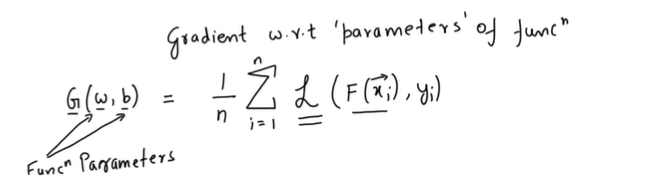


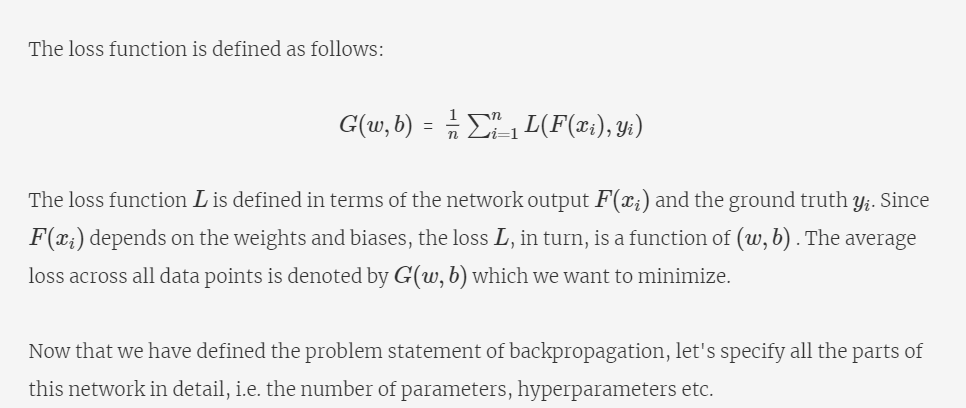




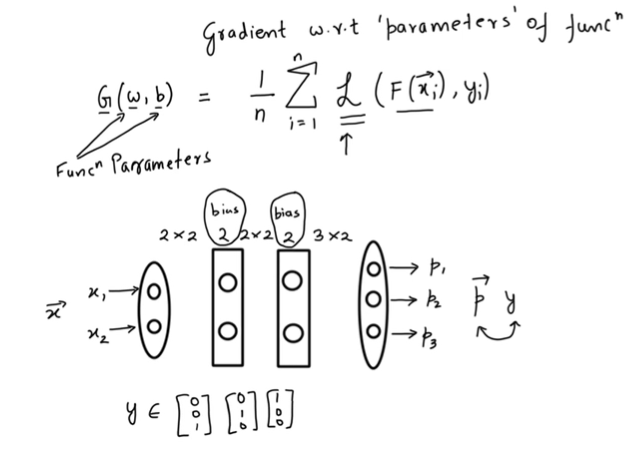


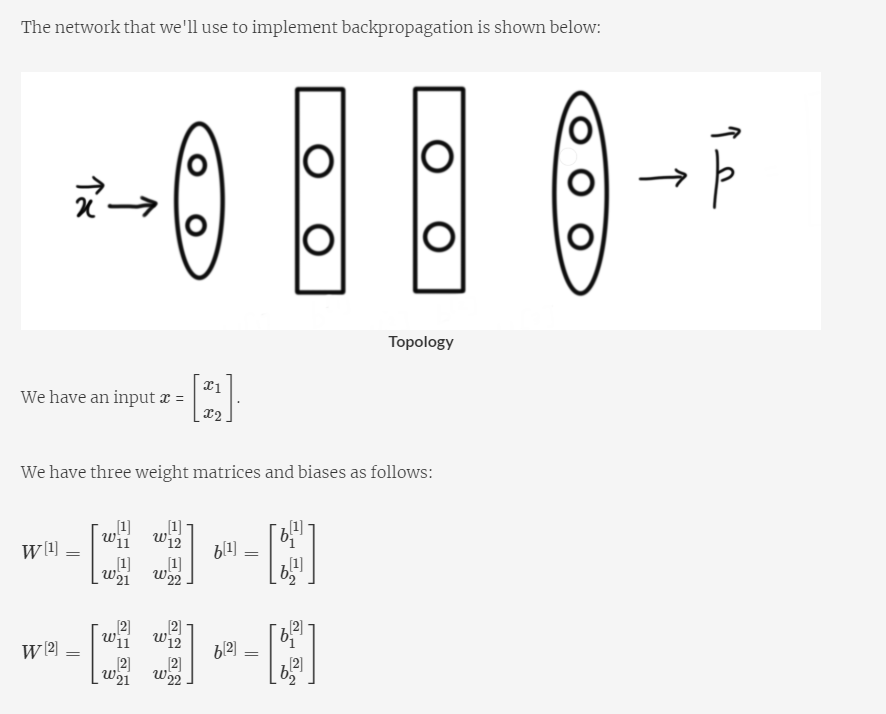
# Updating the Weights and Biases - I

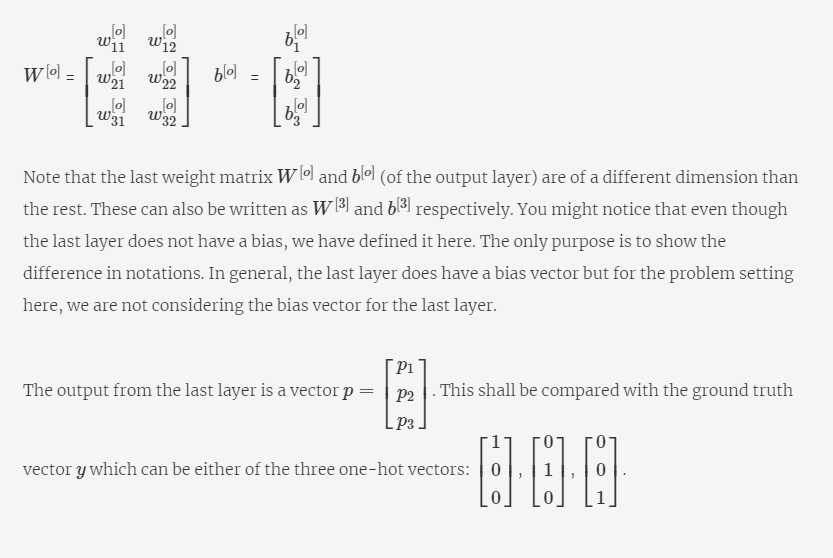


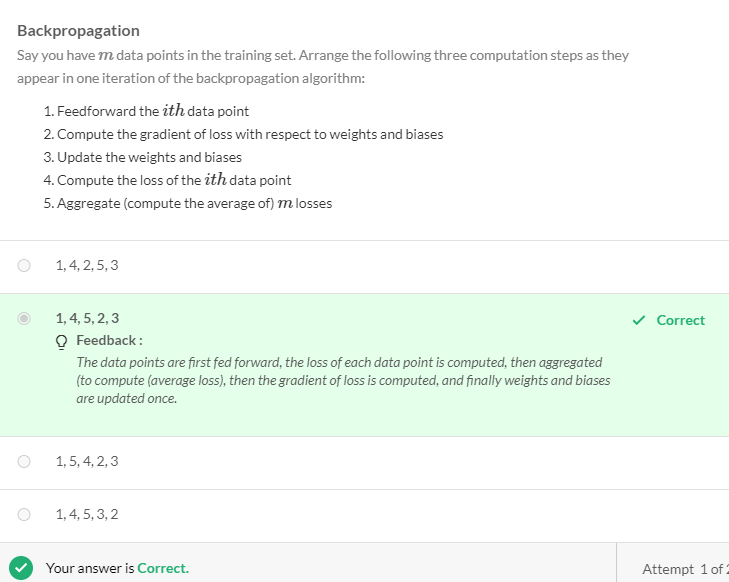


G here is the gradient descent function. W is the weight and b is the biases. So, the Gradient Descent function is affected here only by the weights and biases (that we control) and the aim is to optimize the loss function.

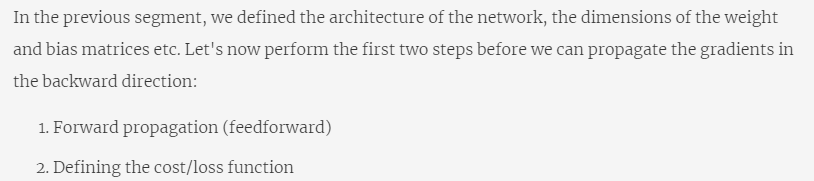


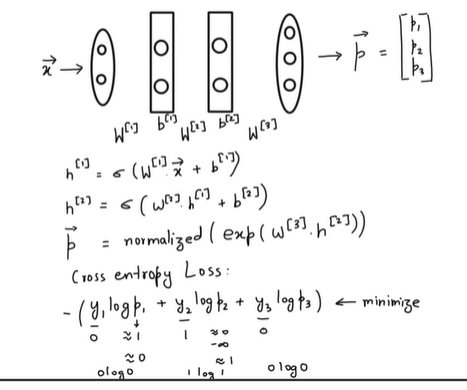




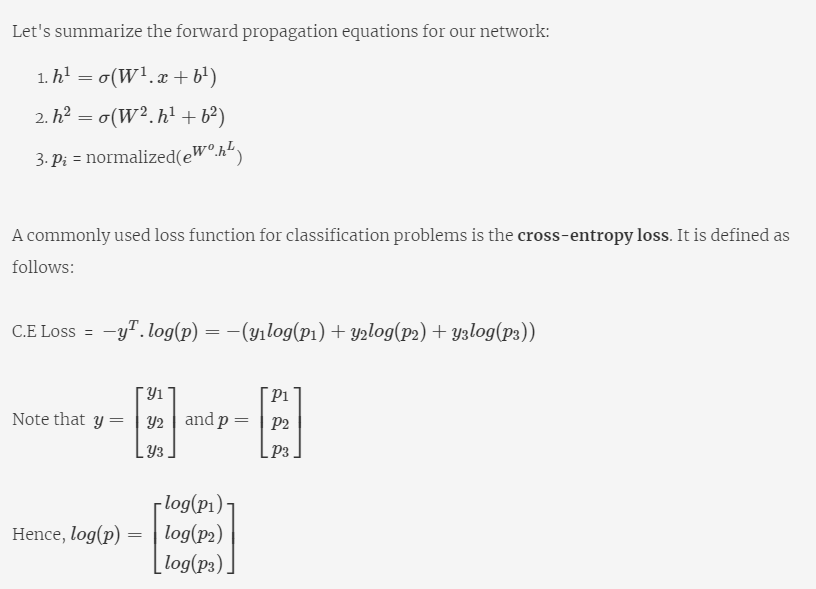


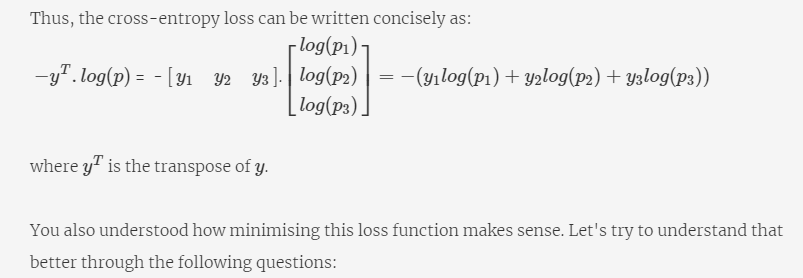
# Updating the Weights and Biases - II

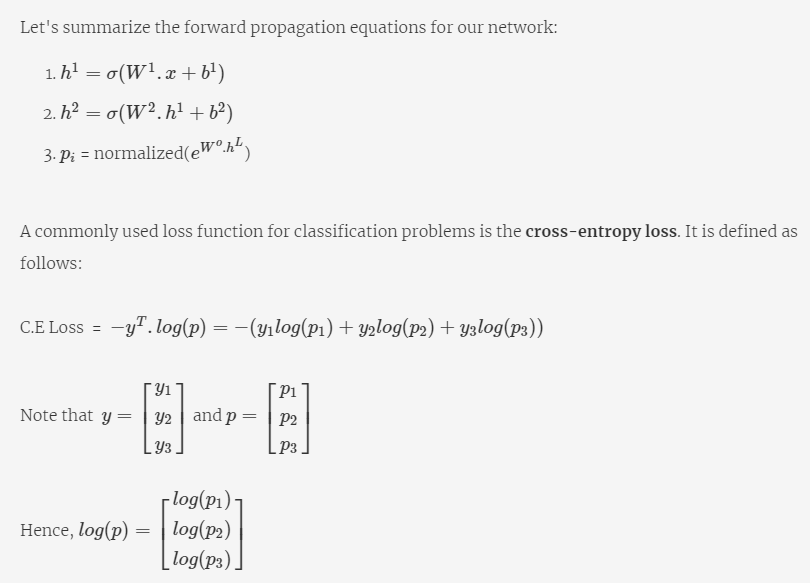


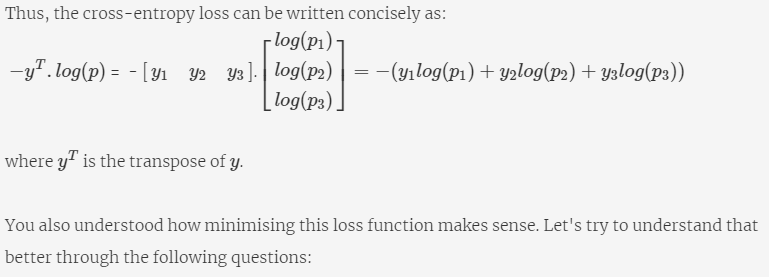


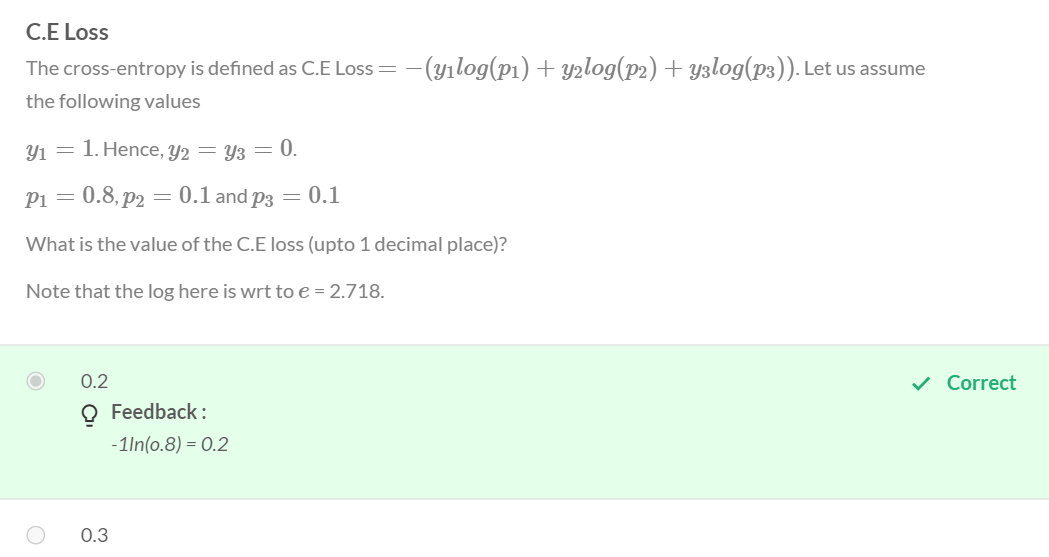
Cross entropy loss for classification

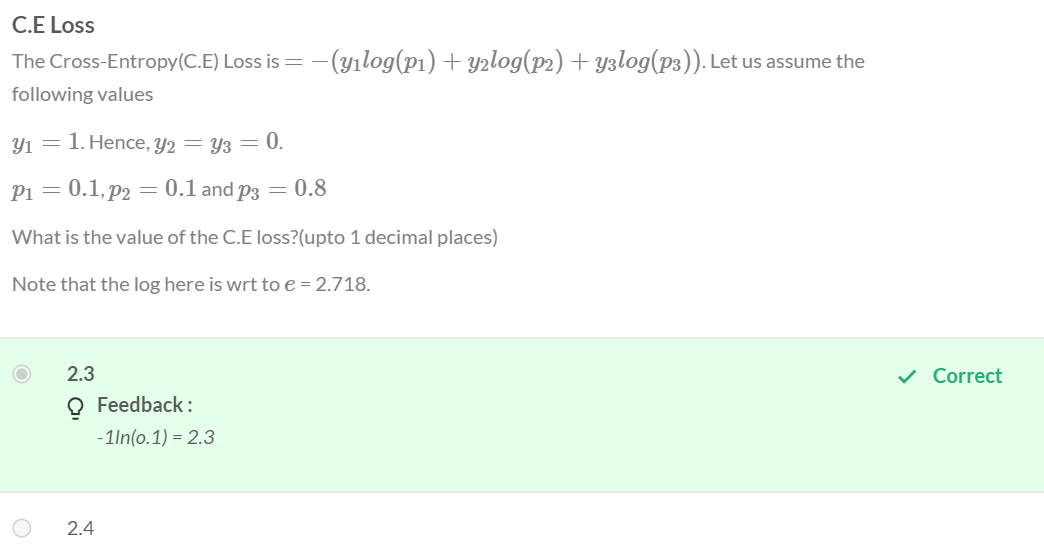


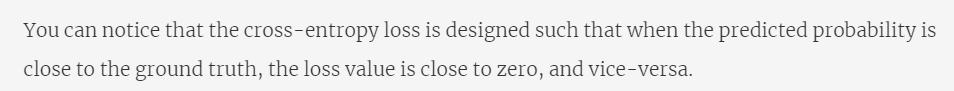


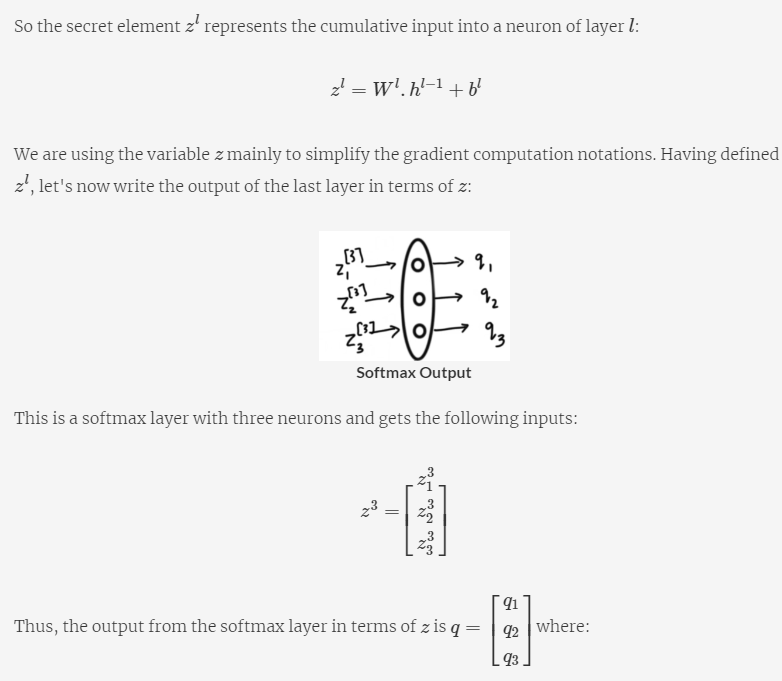


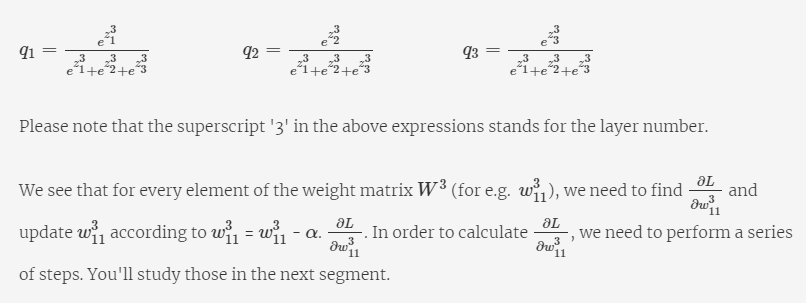


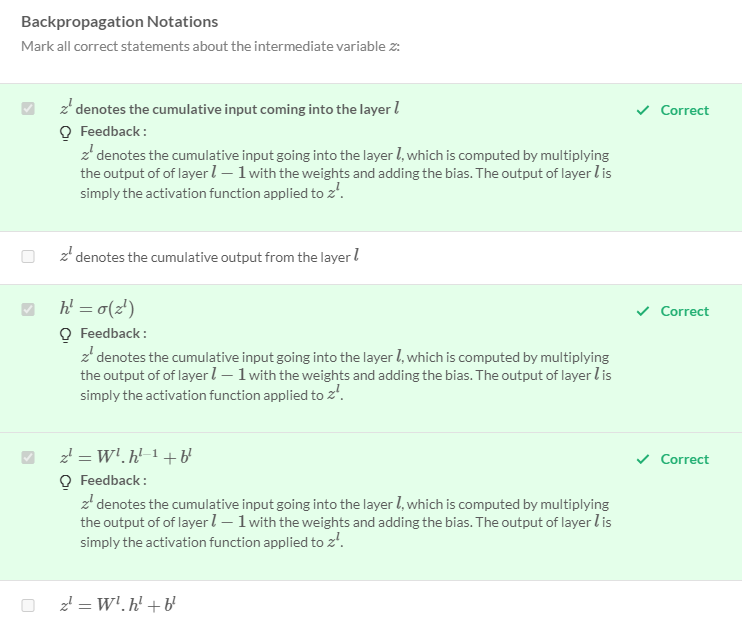




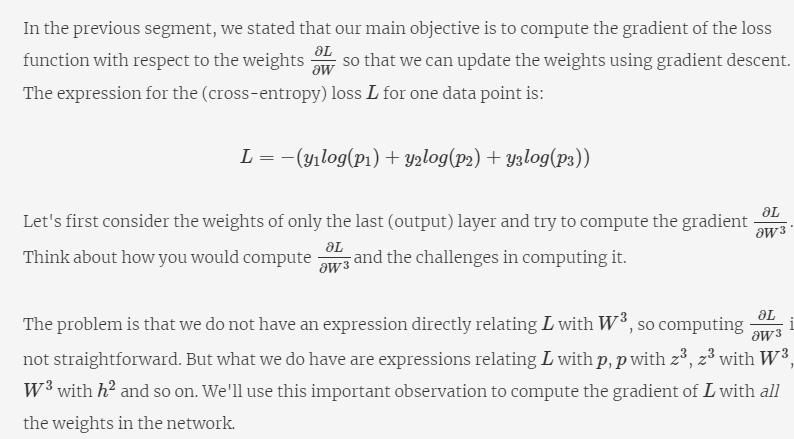


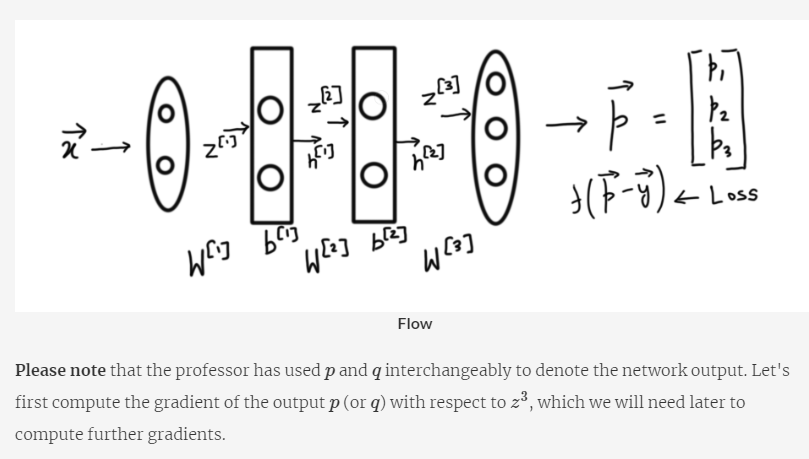


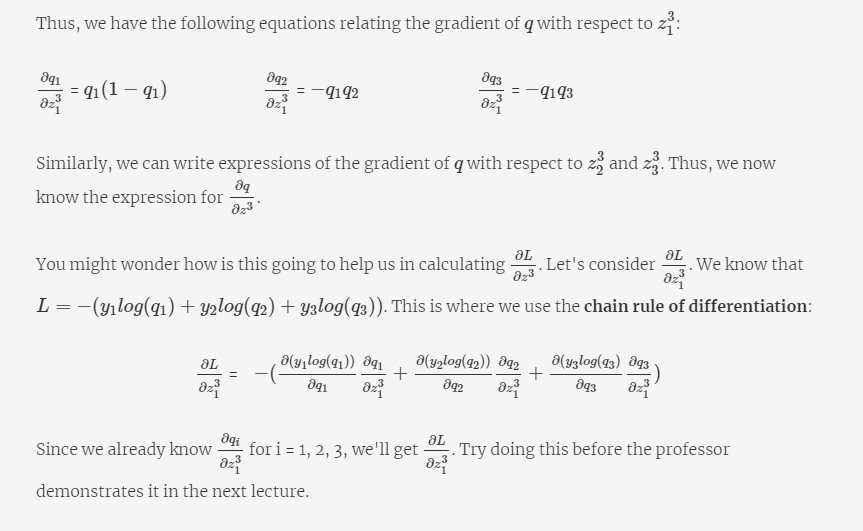


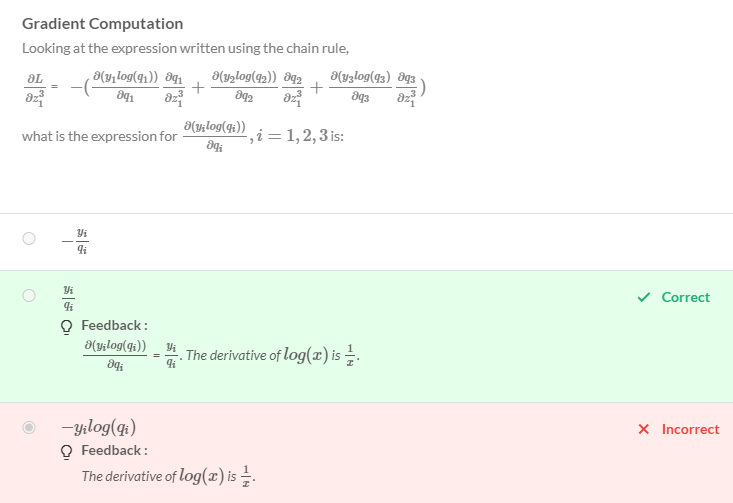


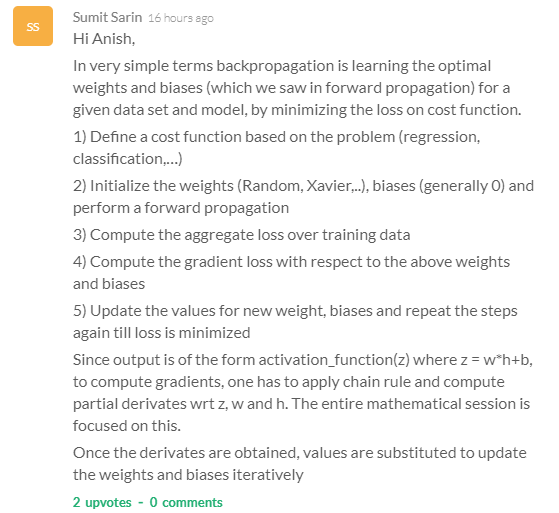
# Updating the Weights and Biases - III

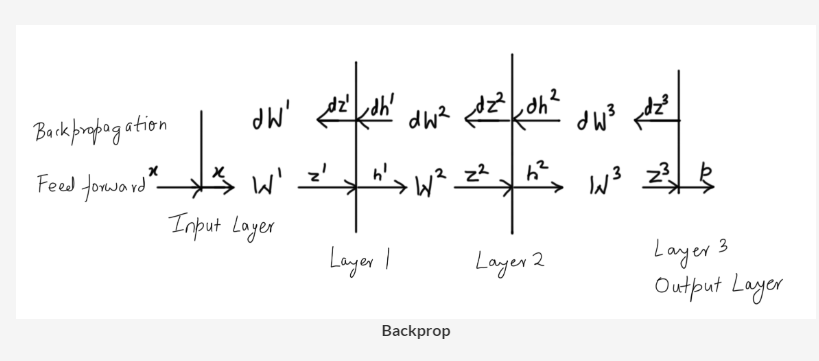


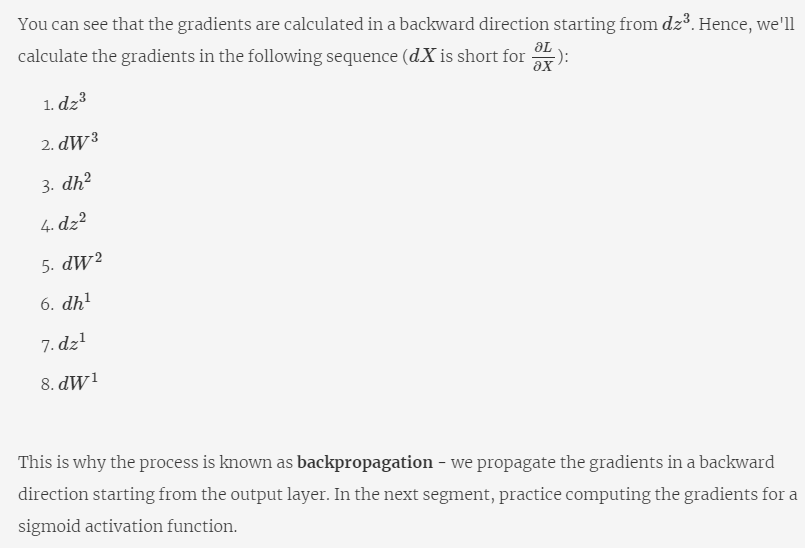


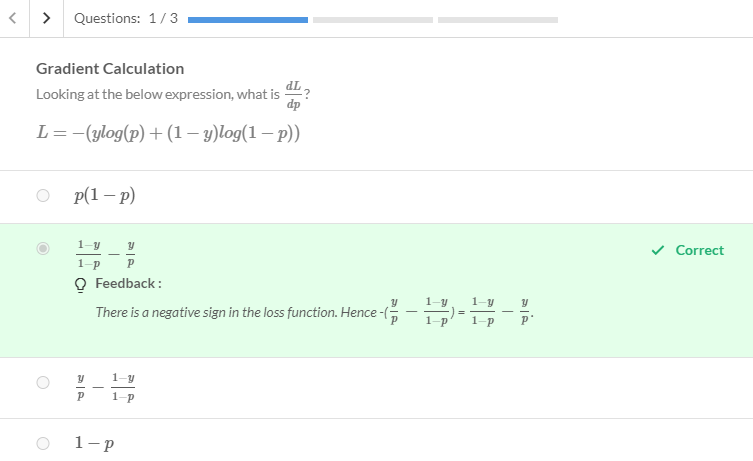


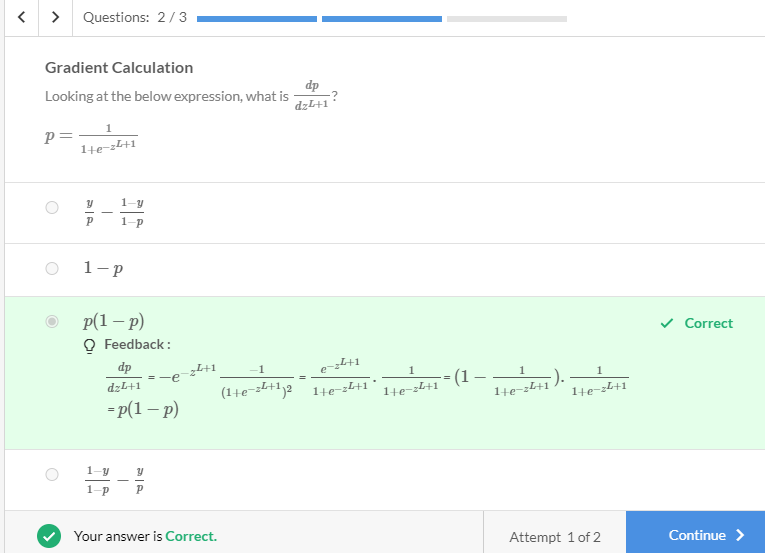


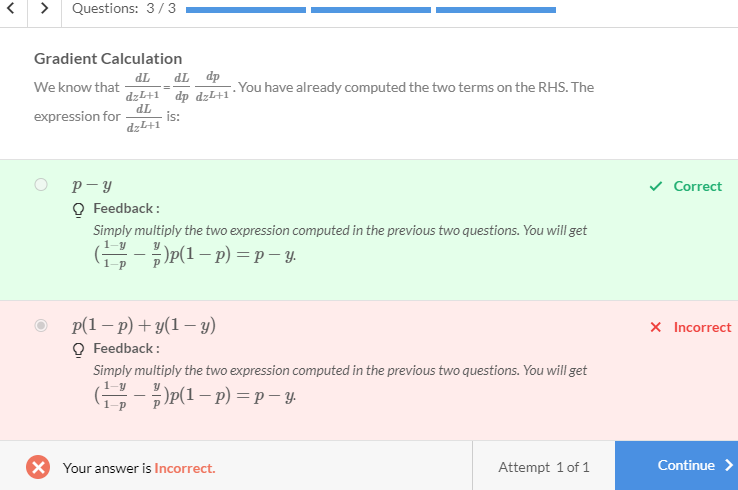




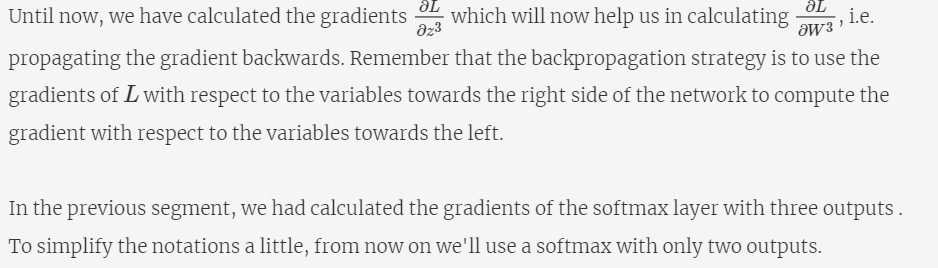


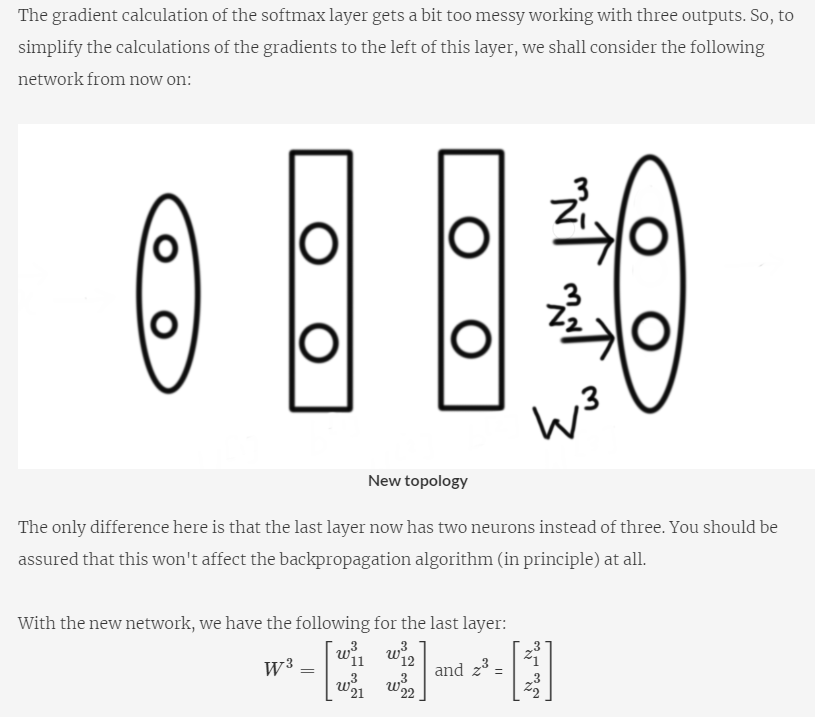


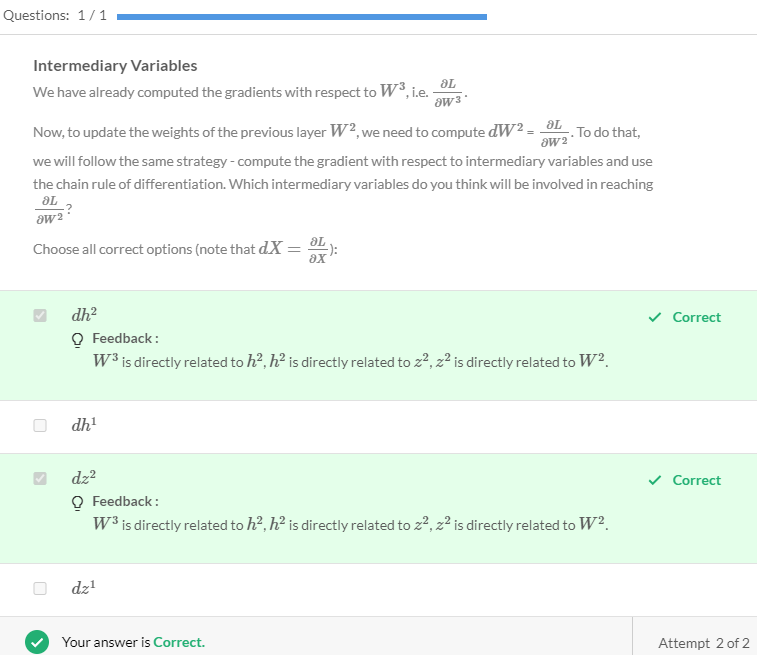


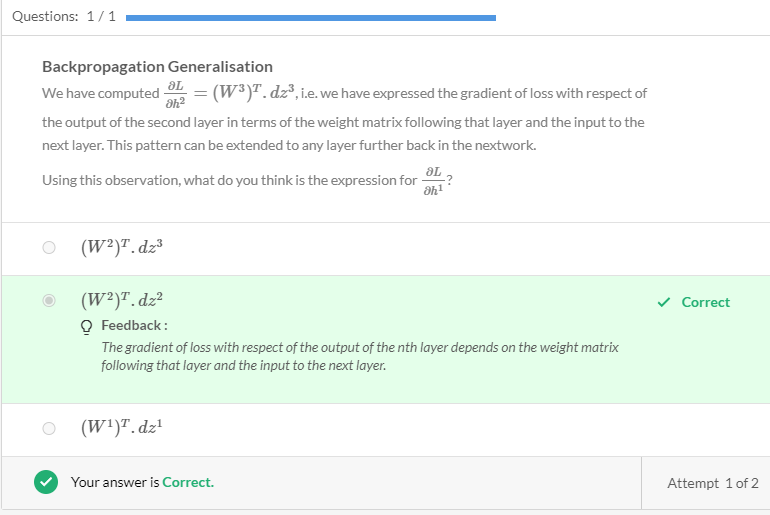


# Updating the Weights and Biases - IV

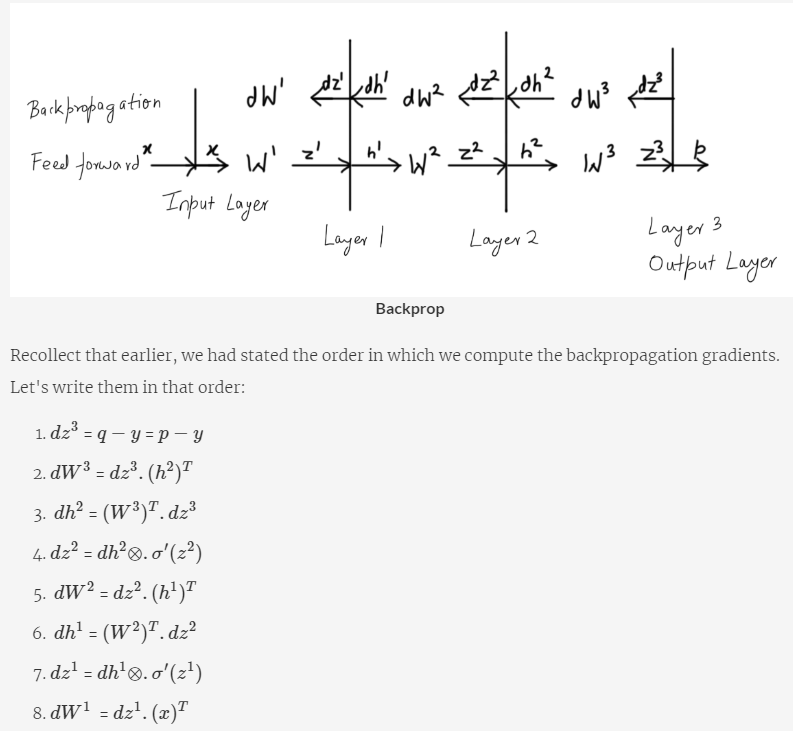


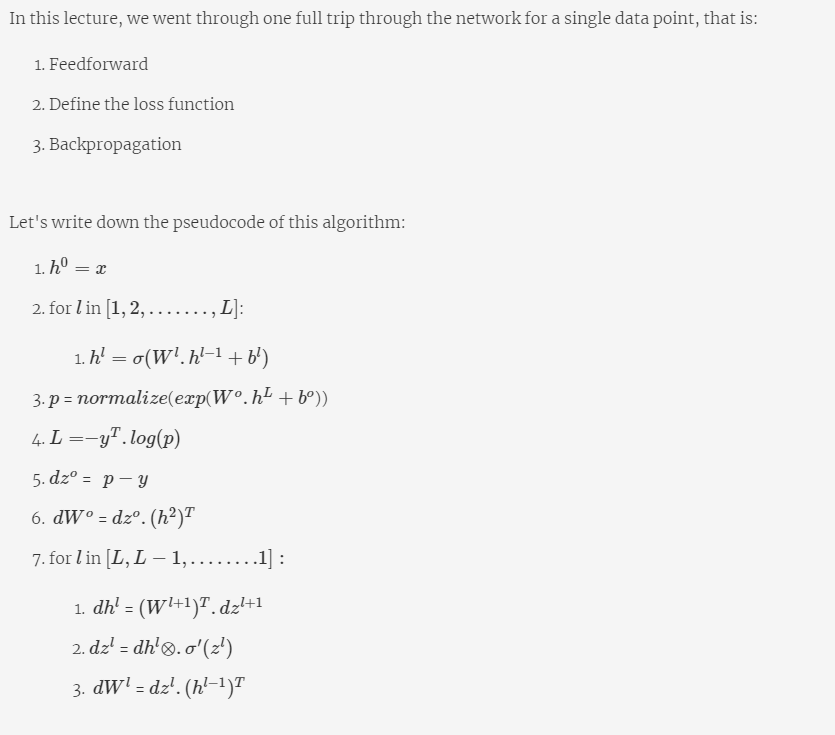
****

****

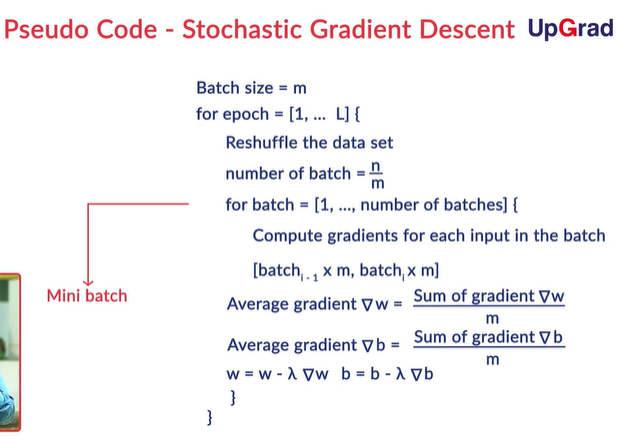
****

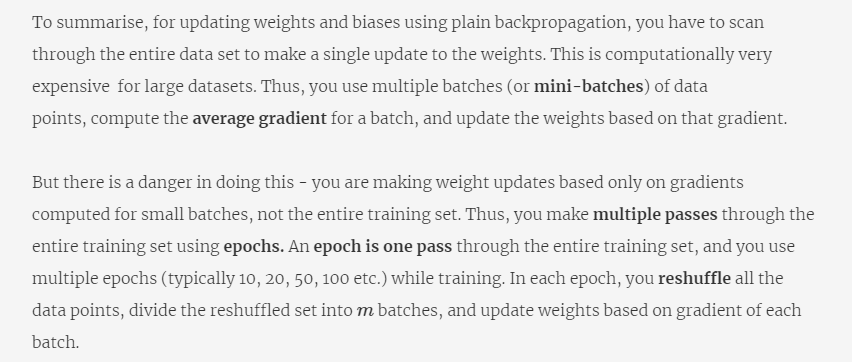
# Updating the Weights and Biases - VI

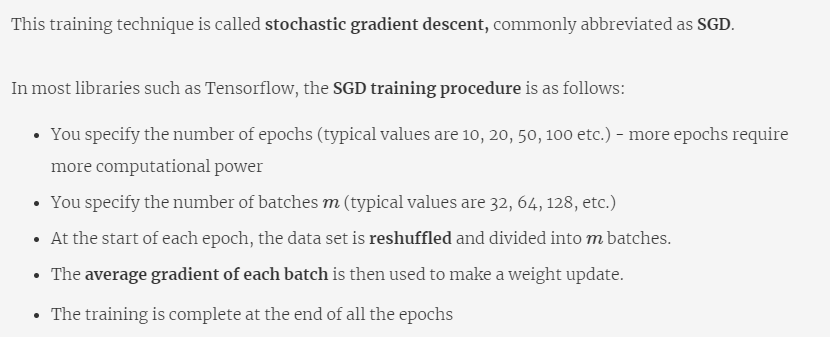
****

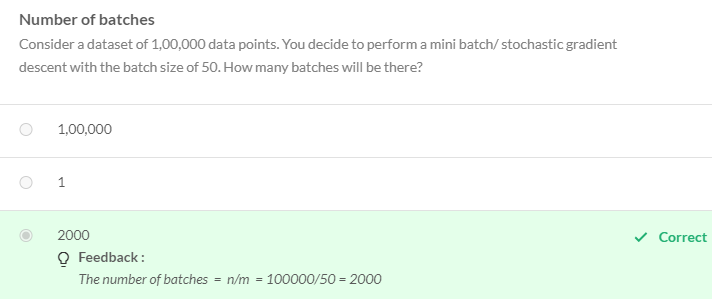
****

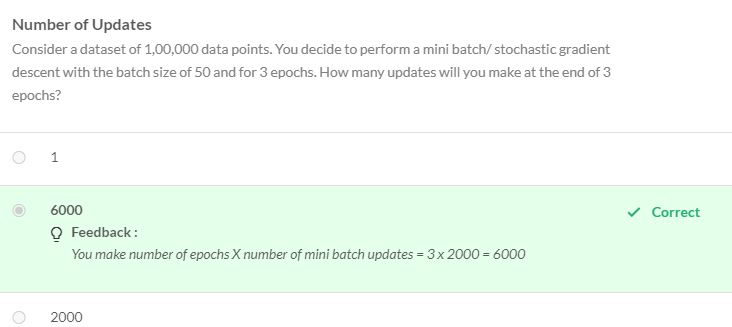
# Batch in Backpropagation

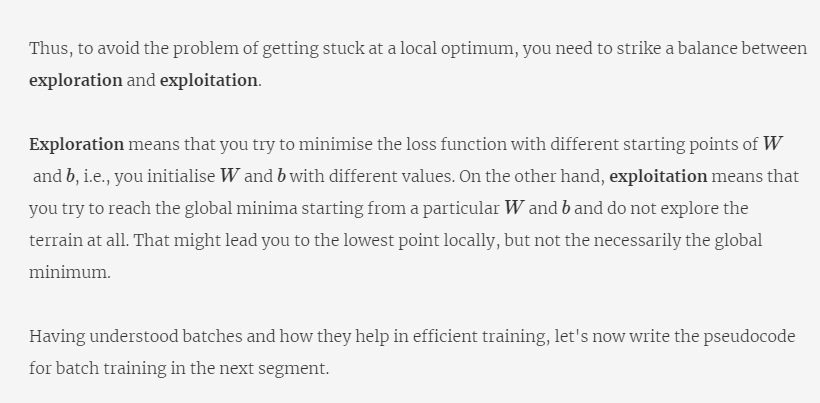
****

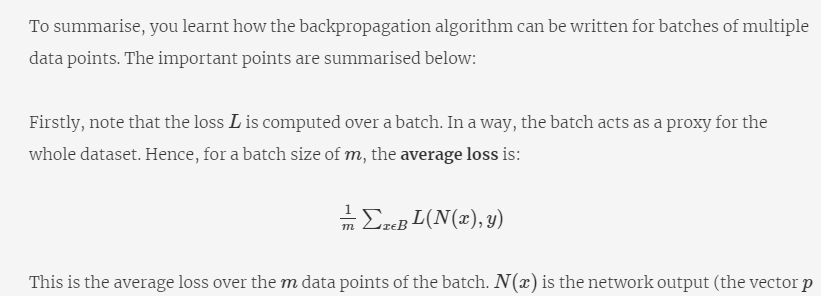
****

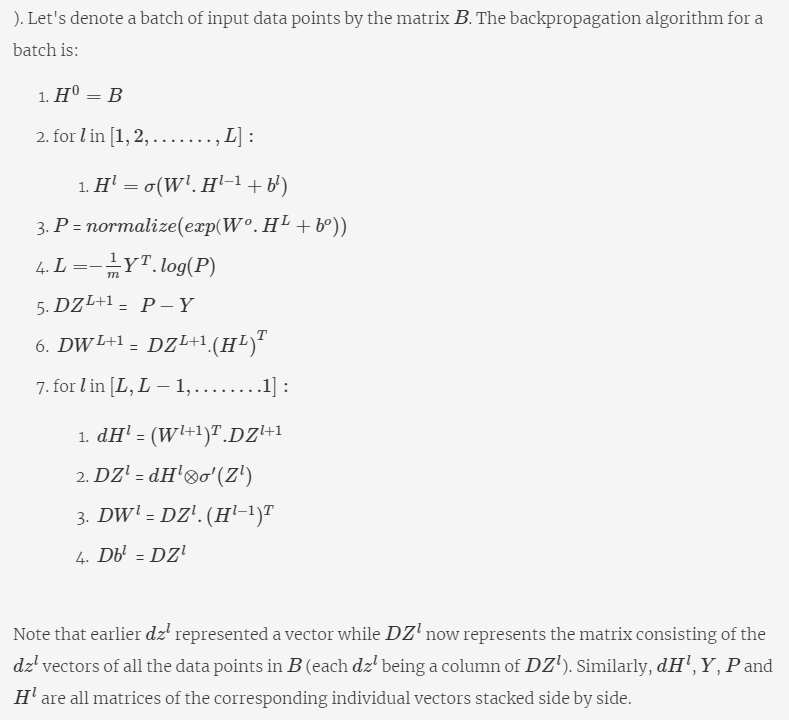
****

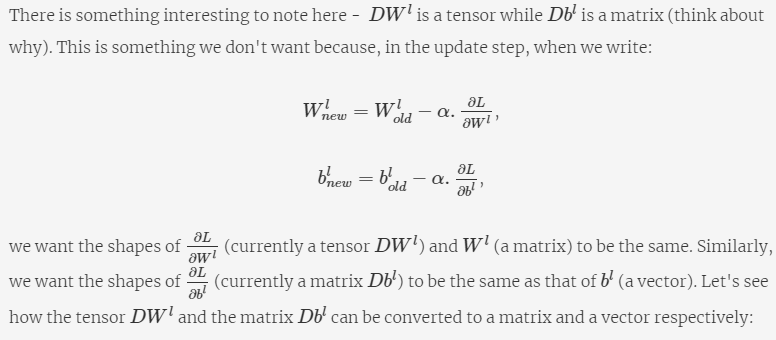
****

****

****

****

****

****