Homework 1: Deep Learning

Hugo Mantinhas 95592, João Silveira 95597

November 30, 2023

Question 1

- **1. a)**

	Train	Validation	Test
Perceptron	0.4654	0.4610	0.3422

Figure 1: Train and validation accuracies for the perceptron model

- **1. b)** Analysing the plot with the learning rate of 0.01, we find that the weight update many times results in accuracy changes in the order of 5%, in both directions. This suggests that the weight adjustments might be overshooting the optimal value. On the other hand, the learning rate of 0.001 results in a more stable learning process, with the accuracy changes being in the order of 1% or less. Unlike the model trained with 0.01 learning rate, this one seems to be tending to grow over time. This suggests that the weight adjustments are closer to the optimal value. Although the accuracies don't look much different: 0.5784 for learning rate 0.01 and 0.5936 for learning rate 0.001; for the latter it shows a lot more stability in its accuracy across epochs.
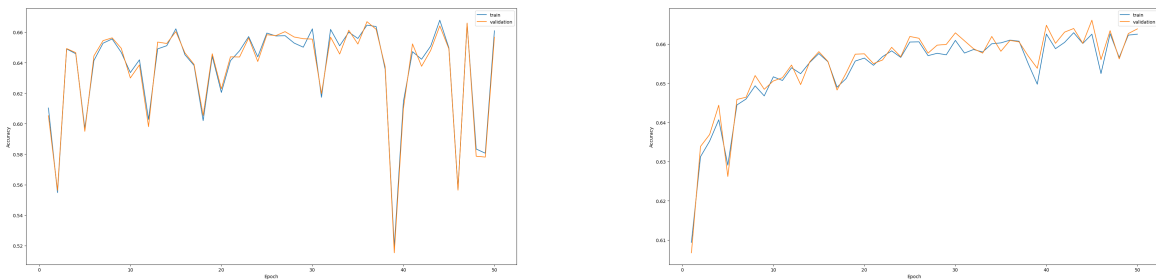

Figure 2: Train and validation accuracies for learning rates of 0.01 and 0.001

	Train	Validation	Test
Learning rate 0.01	0.6609	0.6568	0.5784
Learning rate 0.001	0.6625	0.6639	0.5936

Figure 3: Train and validation accuracies for learning rates of 0.01 and 0.001

- **2.** **a)** This statement is true. A logistic regression model is a linear model, which means that it can only learn linearly separable data. On the other hand, a multi-layer perceptron using relu activations can learn to separate non-linearly separable data because of the non-linearity introduced by the relu activation function in between layers. However, it can be shown, by computing the Hessian matrix of the loss function, that the loss function of a logistic regression model is always convex, while the loss function of a multi-layer perceptron, in general, is not. This means that the logistic regression model can be trained to a global optimum, while, with a multi-layer perceptron, we can never be sure that we have reached a global optimum.