DeepLearning Lab Report - Exercise 1

Sara Al-Rawi

30. October 2018

1 Introduction

In the first exercise, a neural network and different activation functions have been implemented. The neural network performs a multi-class classification problem.
In order to optimize the weight and the bias, a stochastic gradient descent and gradient descent algorithms have been implemented.
One of the goals of this lab is investigating the neural network architecture, the hyperparameters (e.g. epoch numbers, learning_rate) and their effect on the training. For that purpose, different learning rates have been applied and the results have been collected and compared.

2 The Implementation

The TODo parts in the code have been implemented and comments have been added. The theory behind the approach has been also added to the code to make it as clear as possible. I will avoid the repetition.

3 Observations

The hyperparameters such as learning_rate and number epochs have been at first set to 0.05 and 30 respectively. However, it has been observed that the algorithm started to converge in the epoch 25 achieving a validation error 0.0278. By reaching epoch 30 it achieved validation error 0.0283 and the duration was 890.5s. Therefore, the number of epochs has been reduced to 25 and keeping the learning_rate 0.05. In this case the validation error converged to 0.0308
The network of four layers has been considered and the activation relu. At the end the network achieved validation error 0.0241 with learning_rate 0.06 and 30 number of epochs and duration 896.02S. After achieving this result on the validation set, the network has been applied to the test set achieving error 0.0304