Deep Learning in Computer Vision Lab 2: Neural Networks - Implementing a Neural Network from scratch

02/10/2019

Tara Vanhatalo
M2 Modélisation Mathématique pour le Signal et l'Image
Université de Bordeaux

Overview

The goal of the first part of this lab is to create a neural network containing at first only a single neuron. We will then introduce a single hidden layer composed of 64 neurons, keeping our binary output and finally we'll implement a neural network composed of one hidden layer (of 64 neurons also) and an output of 10 neurons. To code the non-binary output of the last part of this lab, we'll be using one hot encoding.

To do this, we will be using the MNIST dataset that consists of 70000 grayscale images of handwritten numbers of size 28 by 28.

The first 60000 images will be used to train our model and the other 10000 will be used to test it. To do this, we'll be implementing a gradient descent method to minimise our loss function using the cross entropy.

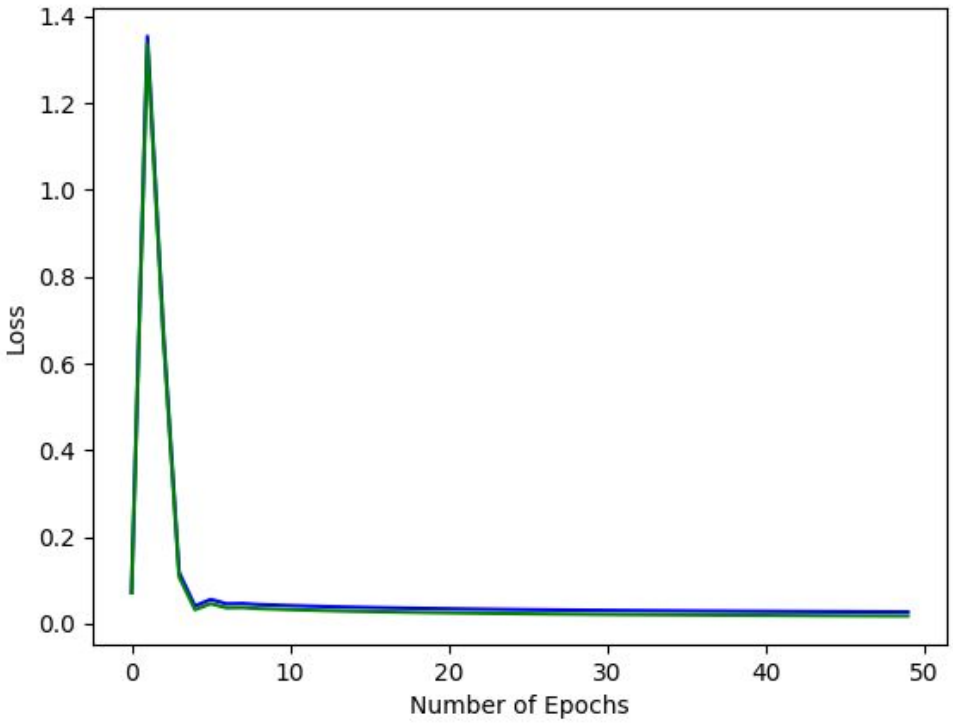
We will provide the graphs plotting the loss function against the number of epochs (i.e. number of iterations) for both our train set and test set. We will do this for various numbers of epochs for each model.

In all figures presented here the loss for the train set will be plotted in blue and for the test set it will plotted in green.

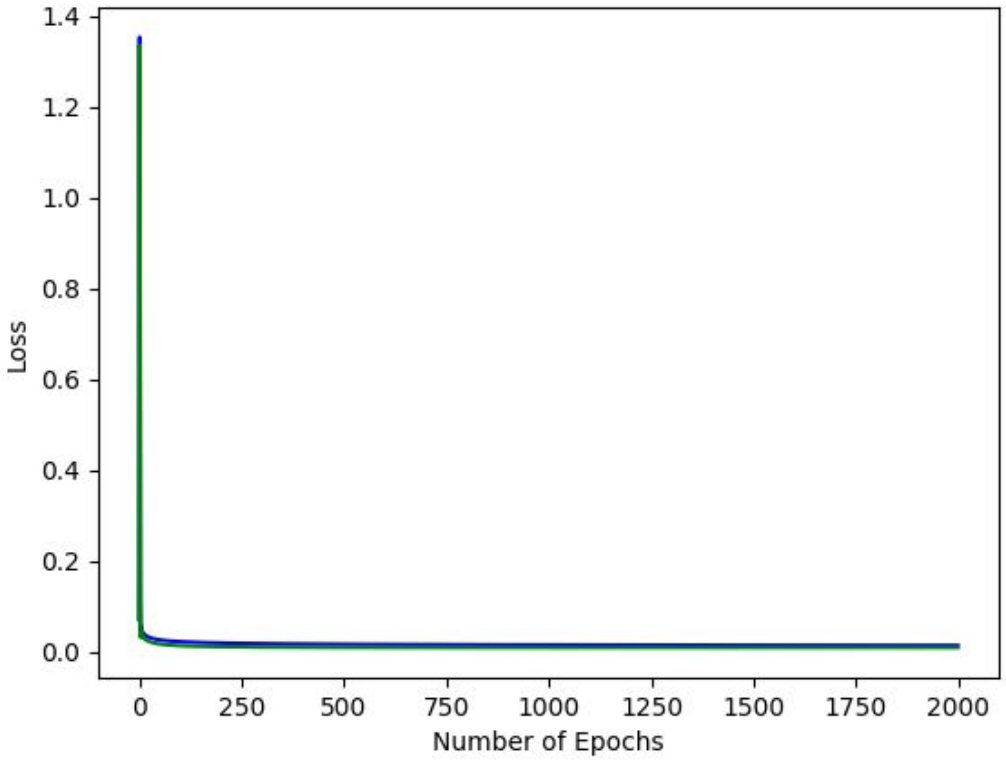
A Single Neuron

For this first model, the graphs provided are for 50 and 2000 epochs.

I. 50 Epochs



II. 2000 Epochs



III. Comments

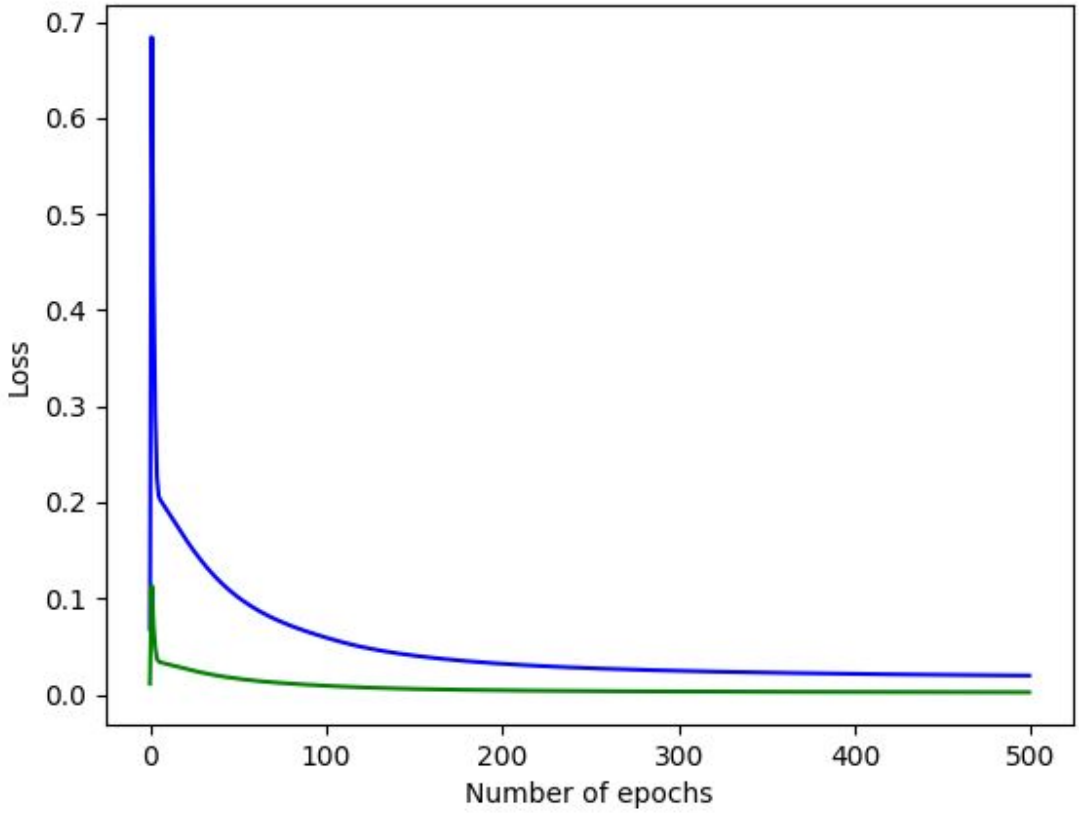
After running our code for 2000 epochs, we found out that the loss for our test set starts to increase after 997 epochs. Although the method done for this is inaccurate because the first instance of increase could be caused by fluctuation in our loss values not linked to over-fitting. Therefore we cannot conclude whether or not we have over-fitting for a number of epochs inferior to 2000 even though judging by the graph above, we do not have visible overfitting.

A Neural Network with one hidden layer

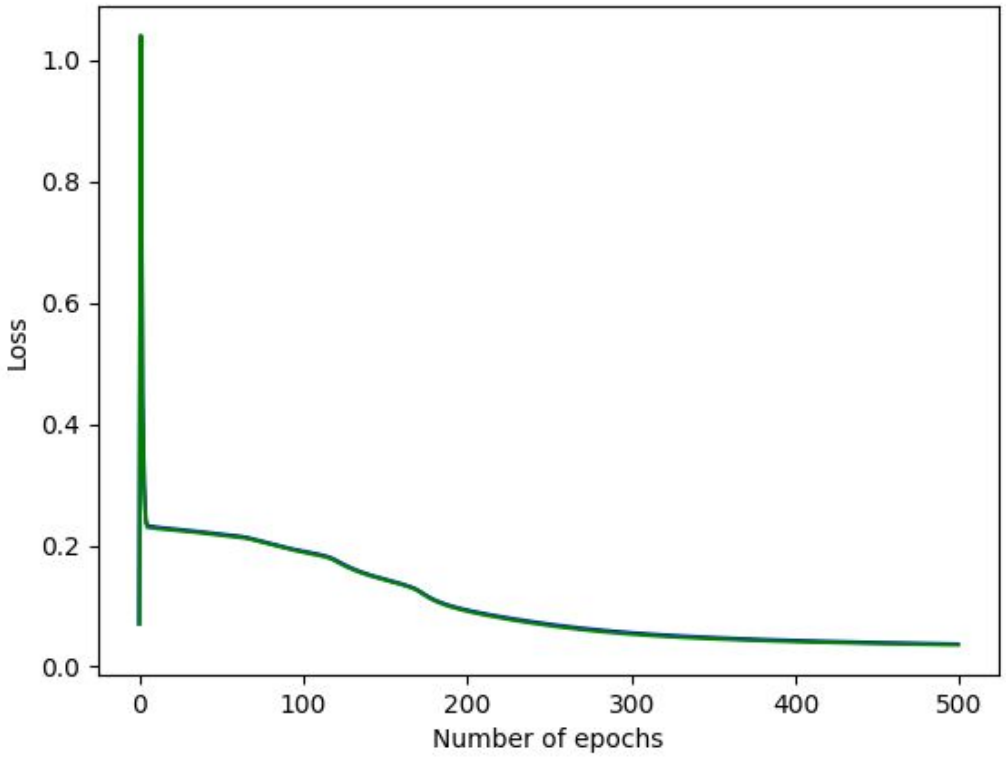
For these next two parts of the lab we'll simply provide the graphs for 500 epochs and the loss value at the last epoch.

Train loss value for 500 epochs is 0.019736216500029664).

Test loss value for 500 epochs is 0.002380718798755575.



Multiclass Neural Network



Train loss value for 500 epochs is 0.03708965237557871.

Test loss value for 500 epochs is 0.03587903684734151.