Neural Networks and Deep Learning

Cracow University of Technology

Lab Assignment 7:

Task: showing the effect of different types of weight initializations and implementing dropout technique to regularize the network.

Subtask 1:

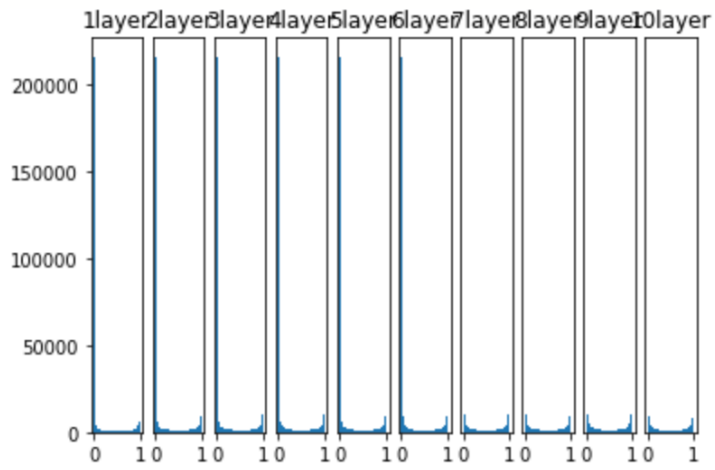
In the first part of this lab you are asked to show the effect of having different weight initializations in the forward pass as discussed in lecture 8 and according to slides 5 through 21. Please write a code to plot the same histograms shown in lecture slides for the outputs of all layers using the following weight initialization settings:

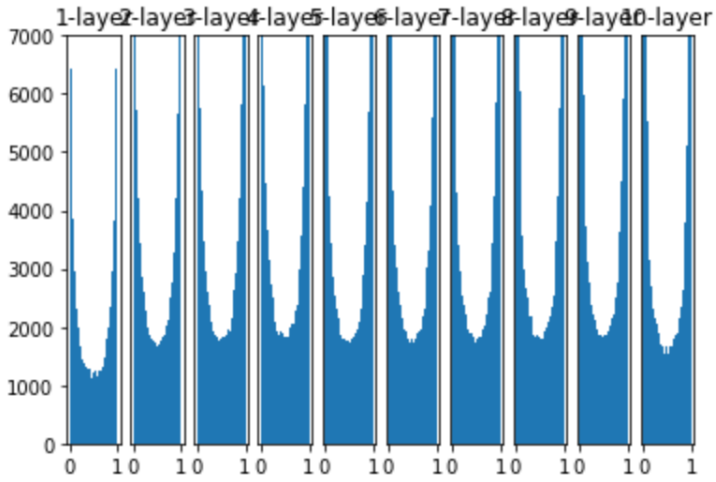
1- Initialize the weights to zero
2- Initialize the weights to small random values
3- Initialize the weights to big random values
4- Initialize the weights with Xavier initialization method
5- Initialize the weights with He initialization method

For your reference here is the configuration we talked about:

- Input is unit Gaussian distribution
- Weights are Gaussian distribution with zero mean.
- Try different activation functions: sigmoid, tanh, and ReLU for all layers.

Network has 10 layers and each layer has 500 neurons.





Subtask 2:

In the second part you are asked to add dropout to the neural network you implemented in lab 5 with retention probability of “0.5” for hidden layer and “1” for input layer. Then compare the result with the network without dropout in terms of train and test accuracies.