DEEP LEARNING

DEEP NEURAL NETWORKS

February 10, 2019

NAME : HIMANSHU
ROLL NO: 16MA20020

Assignment Number 2

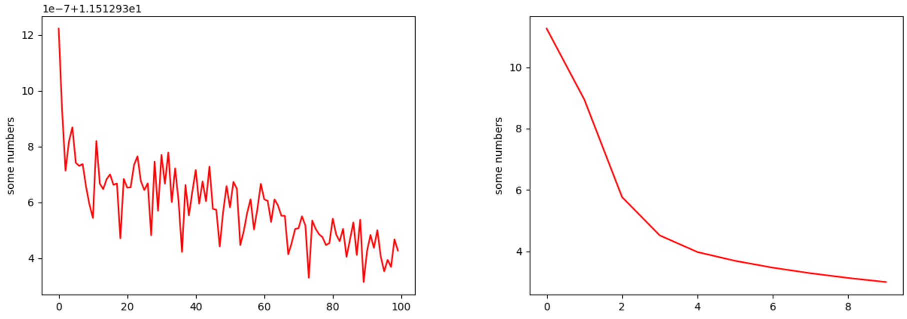
Contents

0.1 Task(a)

We have to build a neural network with the following details(use relu activation for hidden layers and softmax for output layer): Network 1 : âĂŃ 5 hidden layers and 1 output layer [deep and narrow] (512,128,64,32,16) Network 2 : âĂŃ 3 hidden layers and 1 output layer [shallow and wide] (1024,512,256)

0.1.1 Network 1

Deeper network which gradually decreases in width. It takes longer to train and accuracy is around 40%.



(a) Network 1 (loss)

(b) Network 2 (loss)

0.1.2 Network 2

Got 80% test accuracy with just 10 epochs using learning rate 0.0001 and stochastic gradient descent.

0.2 TASK(B)

Perform the following experiments on the network 2.

0.2.1 Experiment 1

Use Normal, Xavier and Orthogonal initialization.

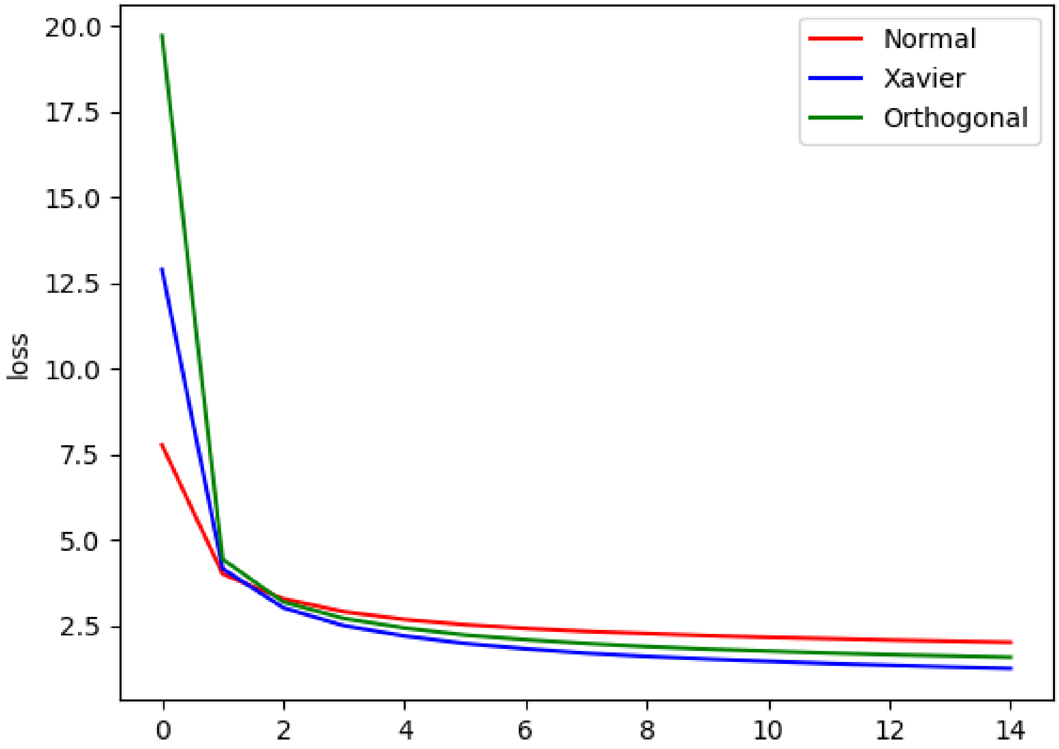


Figure 2: Comparison using various initialization techniques

From figure above observe that Xavier initialization provides fast and least loss for the same values of other hyperparameters in this case.

0.2.2 Experiment 2

Use Batch Normalization

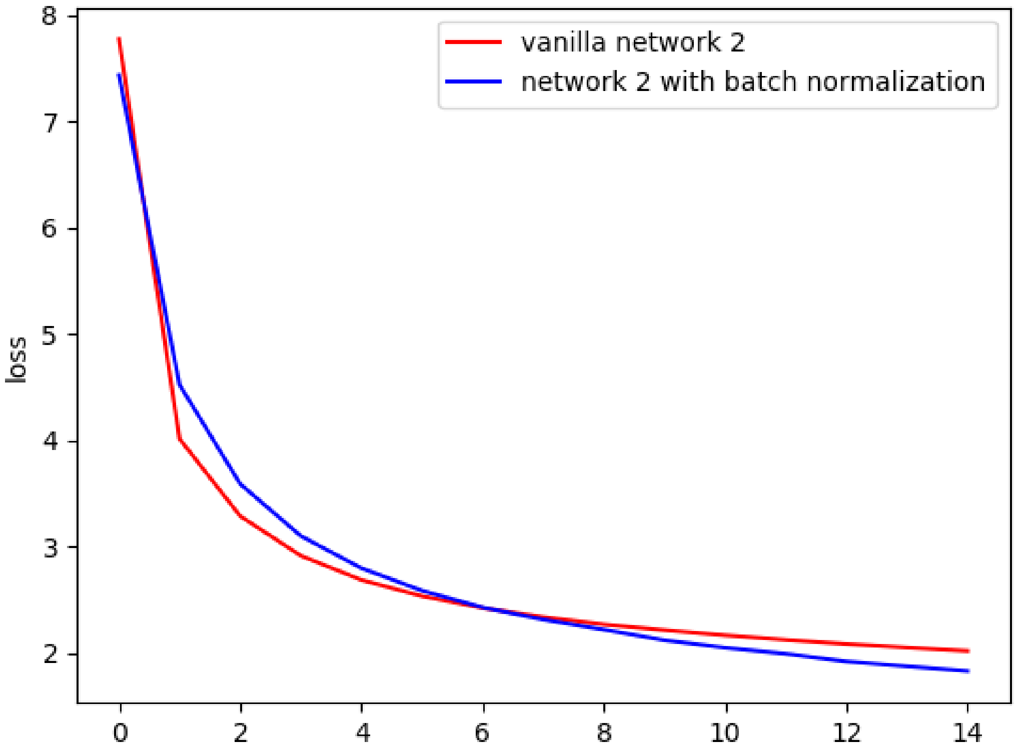


Figure 3: Using Batch Normalization

Batch Normalization gives better results. Loss decrease is better with the network using batch normalization.

0.2.3 Experiment 3

Use Different regularization techniques: Dropout(0.1, 0.4, 0.6)

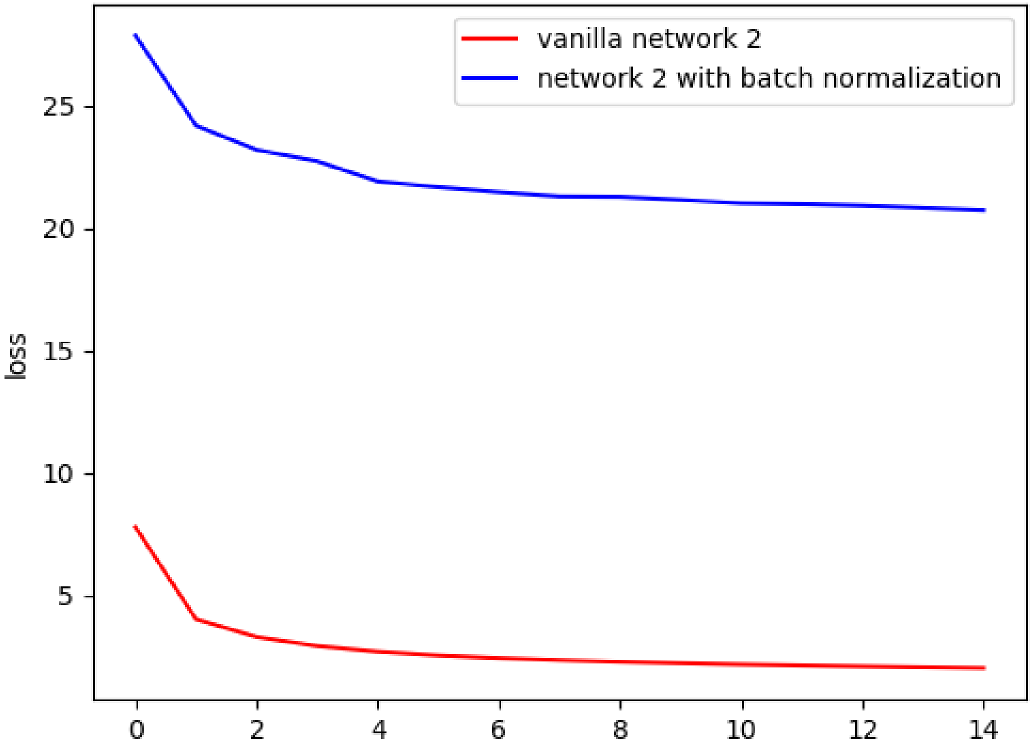


Figure 4: Comparison using various initialization techniques

In 15 epochs clearly there is no overfitting in both cases. Accuracy is similar for both cases i.e. with or without dropout.

0.2.4 Experiment 4

Use different optimizer SGD, Nesterov's accelerated momentum, AdaDelta, Adagrad, RmsProp, Adam.

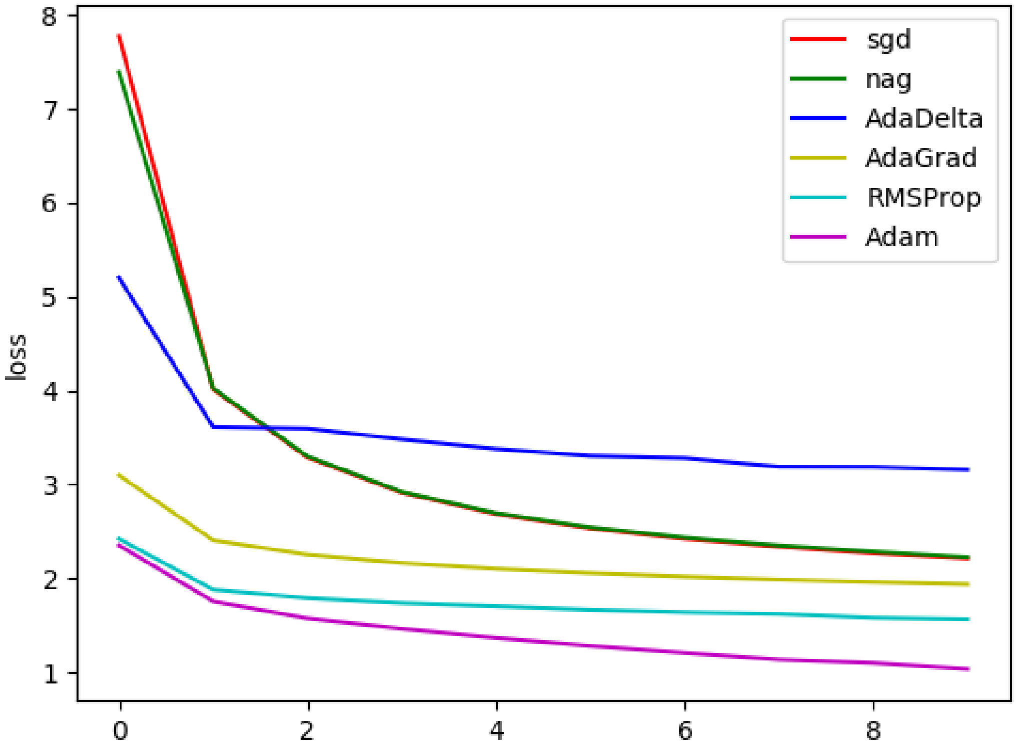


Figure 5: Comparison of different optimizers