Deep Learning Assignment - 2

- Krishna Sai Jaswanth Rongali 16CE10024

In this Assignment, We try to play with various optimizers, hyperparameters and also techniques like Batch normalisation, Dropout normalisation and study the behaviour of a fully connected network with variations of these parameters.

Task - a

In the first task, We design two fully connected networks where one is deeper and the other is a shallow network.The deeper network is **Network_1** and the shallow network is **Network_2**. The final test accuracies of the two networks are .If we watch the runtime of the program for both the networks, at an average the **Network_1** is taking around 5.2 sec per epoch and the **Network_2** is taking around 14.5 sec per epoch where both are trained for the same number of epochs i.e. 20 when I run the models on my gpu. The final training, validation accuracies of both networks and the plots of validation loss vs epoch and training loss vs epoch are as follows:

For **Network_1** (for 20 epochs and learning rate = 0.01)-

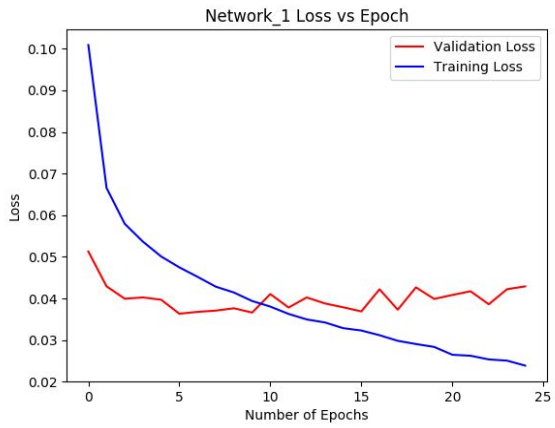
Validation Accuracy = 89.5%

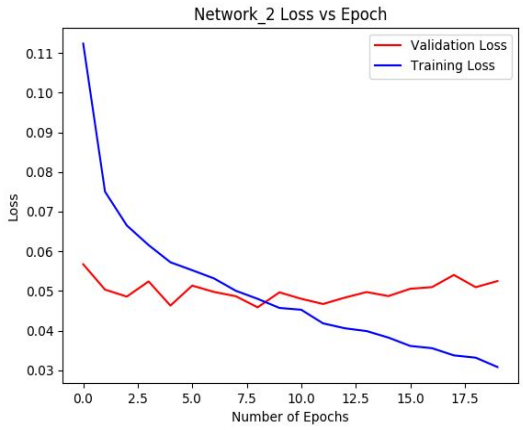
Training Accuracy = 94.9%

For **Network_2** (for 20 epochs and learning rate = 0.01)-

Validation Accuracy = 89.4%

Training Accuracy = 94.6%



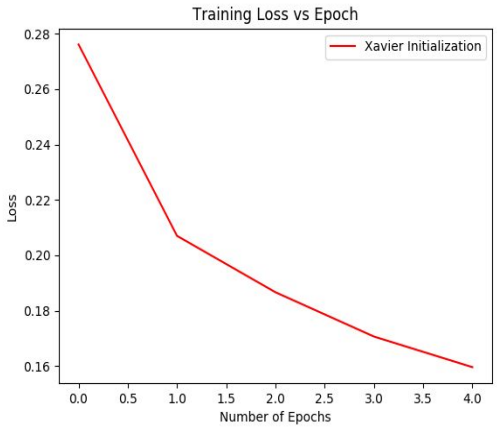


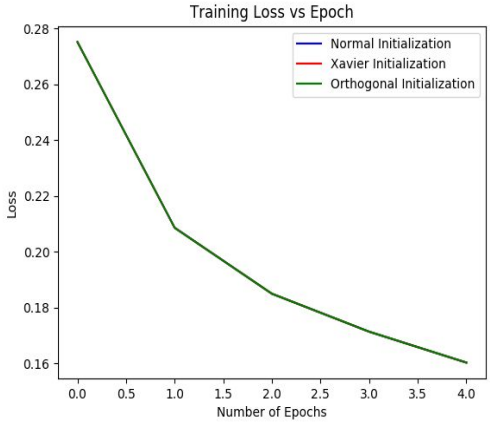
As, we can see that when try to increase the number of epochs the model overfits the data and so the loss increases on validation set. In this both neural networks have the following parameters:

- Normal initialization of weights
- Optimizer - Adam

Task - b

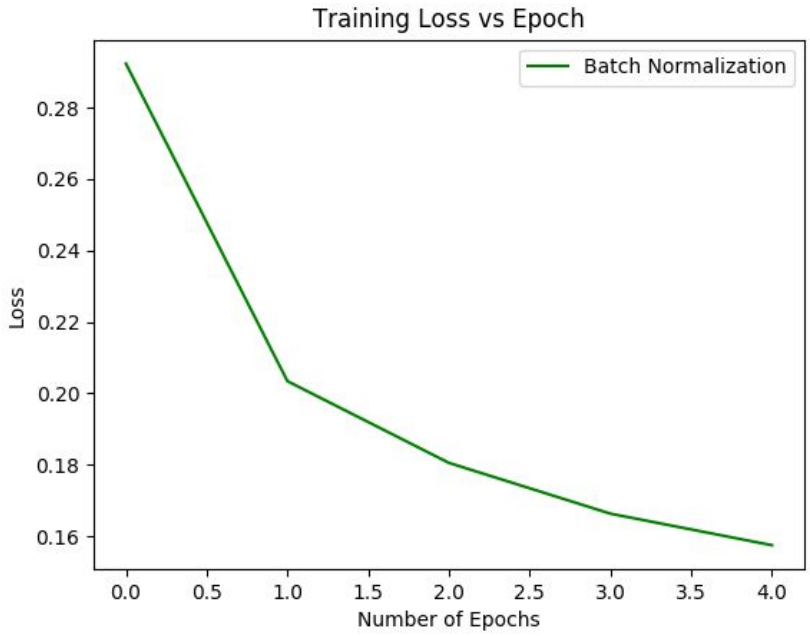
1. Using Normal,Xavier and Orthogonal Initialization



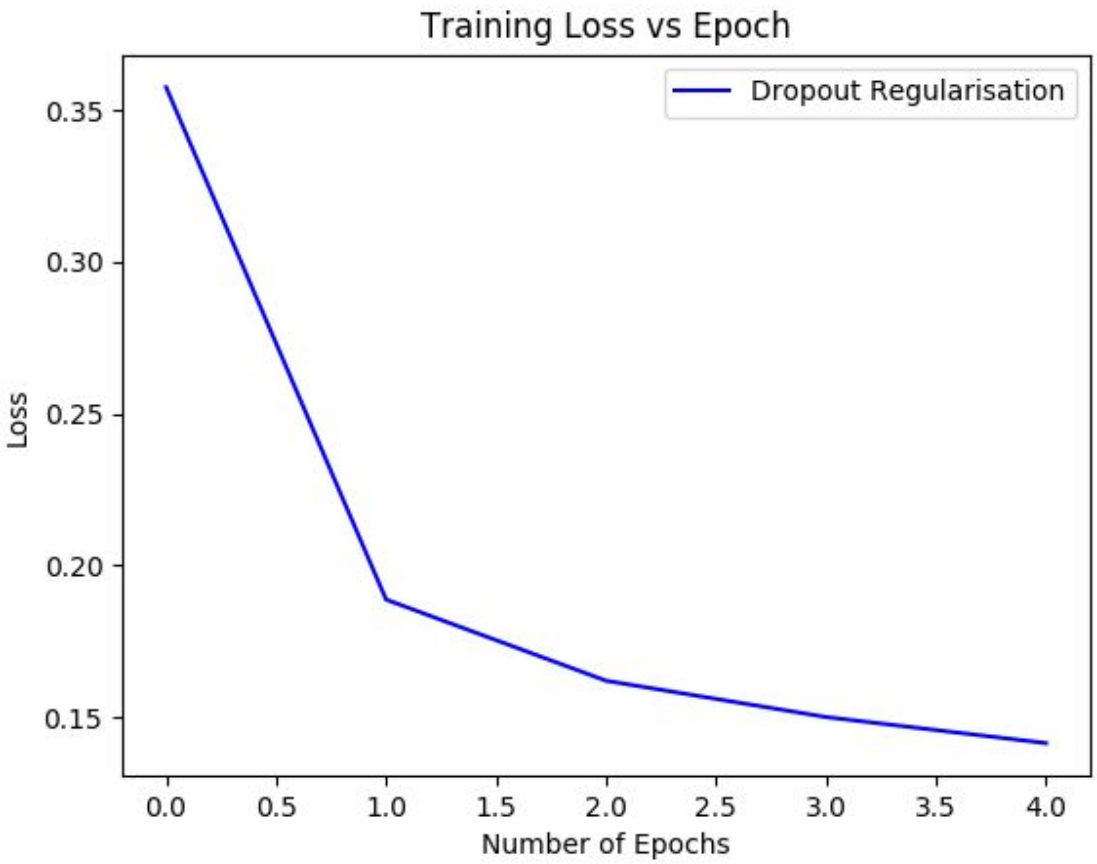


Normal Initialisation is shown in task - a

2. Using Batch Normalisation



3. Using Dropout Regularization



4. Using Various Optimizers