Q1)

Using batch normalization makes the network more stable during training. Batch normalization offers some regularization effect, reducing generalization error.

Dropout is a way to prevent Neural Networks from overfitting because the outputs of a layer under dropout are randomly subsampled, it has the effect of reducing the capacity or thinning the network during training.

Adding an extra duplicate layer helps in fitting training data well.

If we used three blocks instead of two, then ideally the training loss should decrease as training data is fit well. But, in this assignment we are using dropout and batch normalisation together. Ideally dropout and batch normalisation should not be used together.

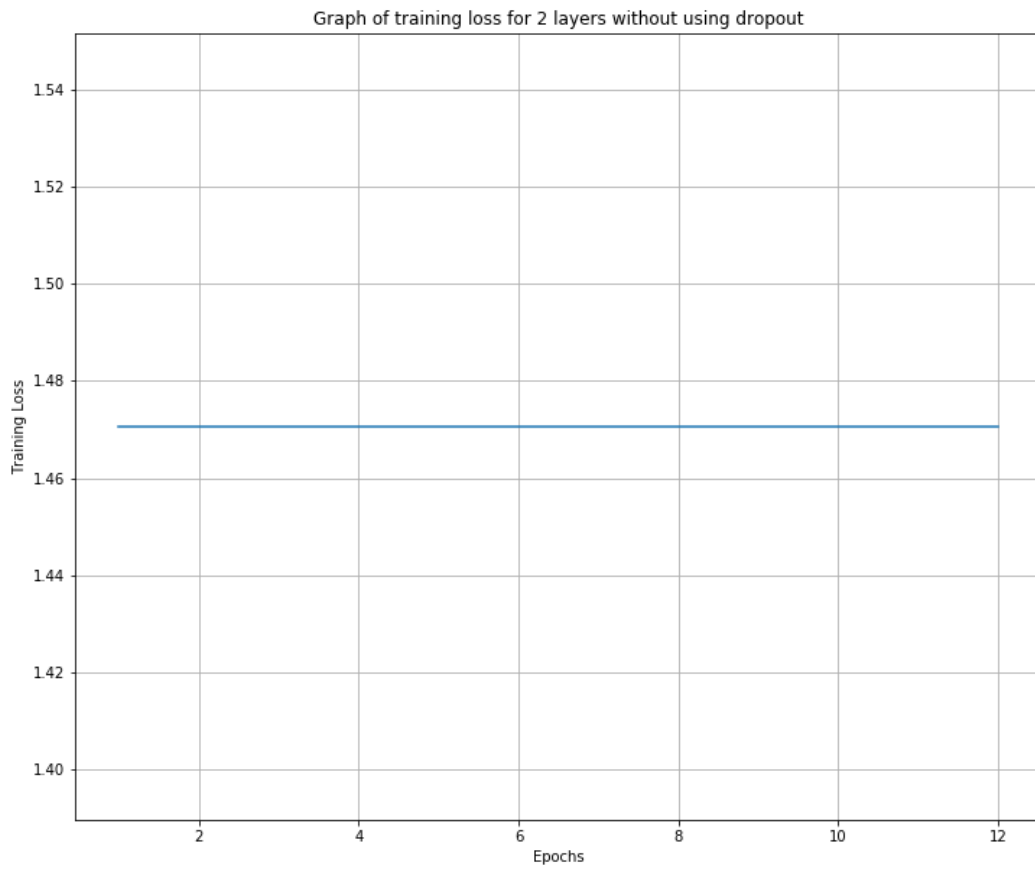## Q2)

Training Loss for the given Neural Network Architecture in the Assignment :
**1.4688057899475098**

Plot of Training Loss vs Number of epochs for 2 Layers :

Graph of training loss for 2 layers
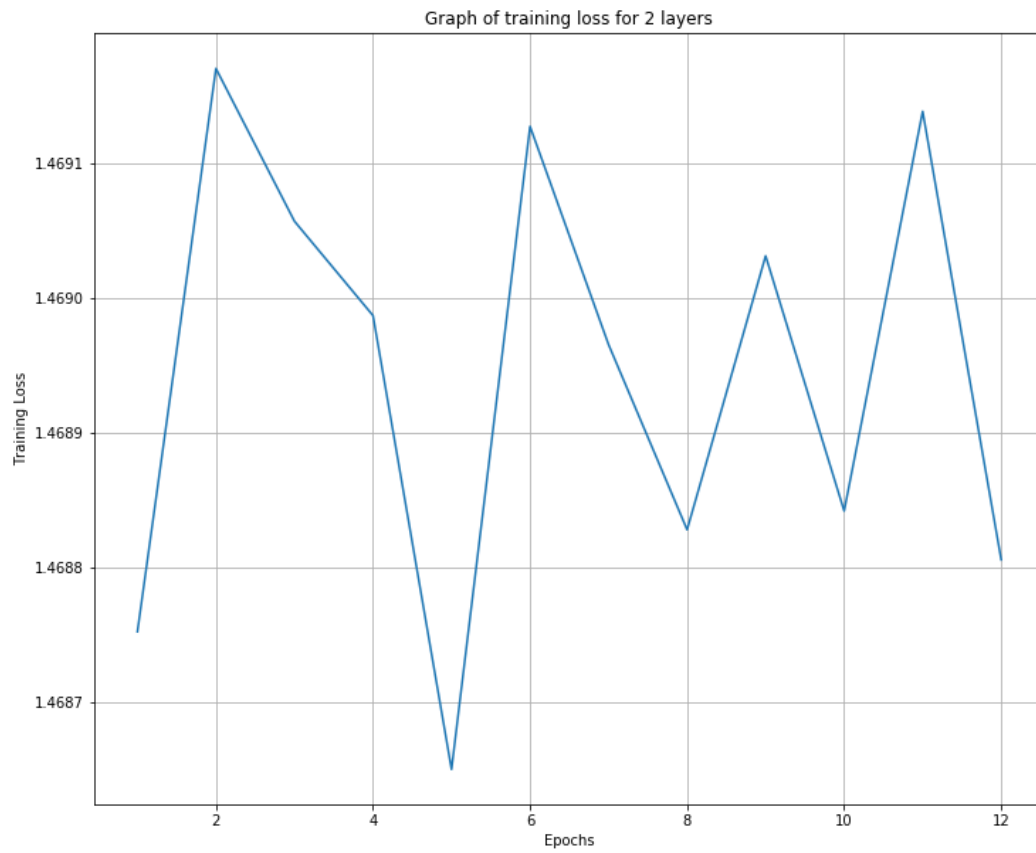
Plot of Training Loss vs Number of epochs for 2 Layers without using Dropout Layer in Residual Block :

Graph of training loss for 2 layers without using dropout

Plot of Training Loss vs Number of epochs for 2 Layers without using Batch Normalisation in Residual Block :

Graph of training loss for 2 layers without using batch normalisation
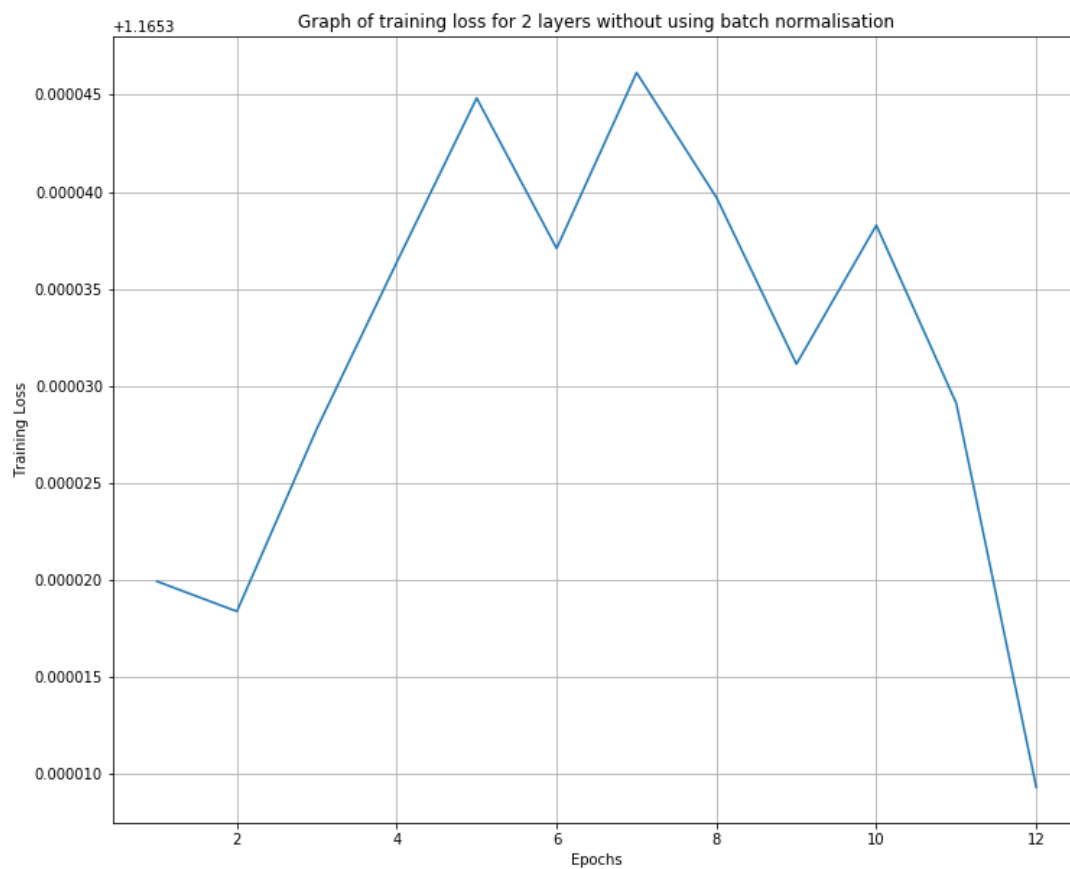
Plot of Training Loss vs Number of epochs for 3 Layers :

Graph of training loss for 3 layers