

# Deep Learning

## Assignment 1 Report

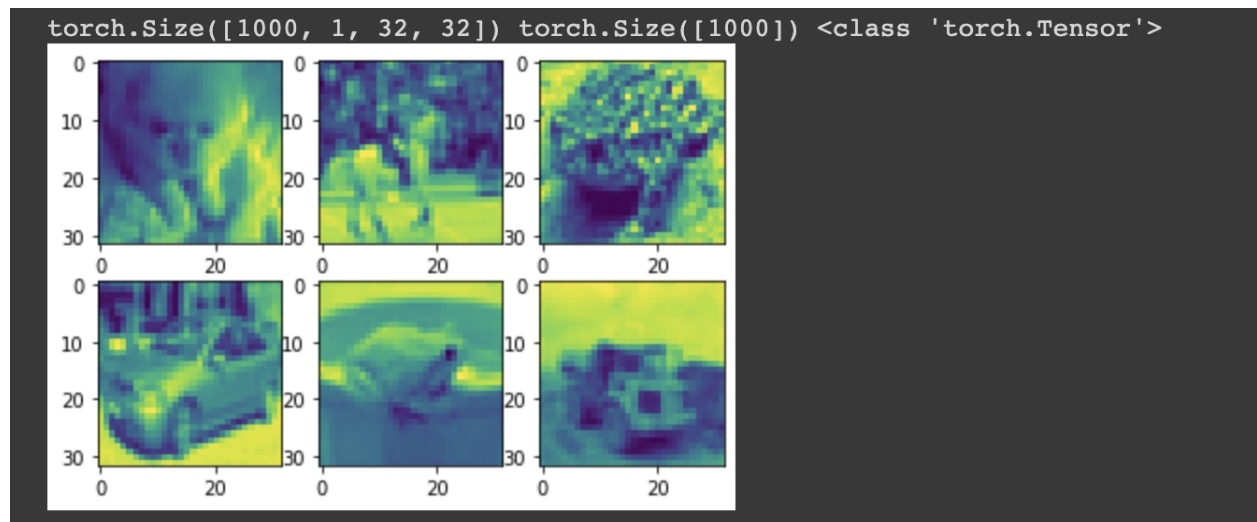
### M22MA003

Question 1 : Implement a neural network and utilize the CIFAR-10 dataset for the analysis.

(A) Utilize various activation functions like sigmoid, tanh and critique the performance in each case.

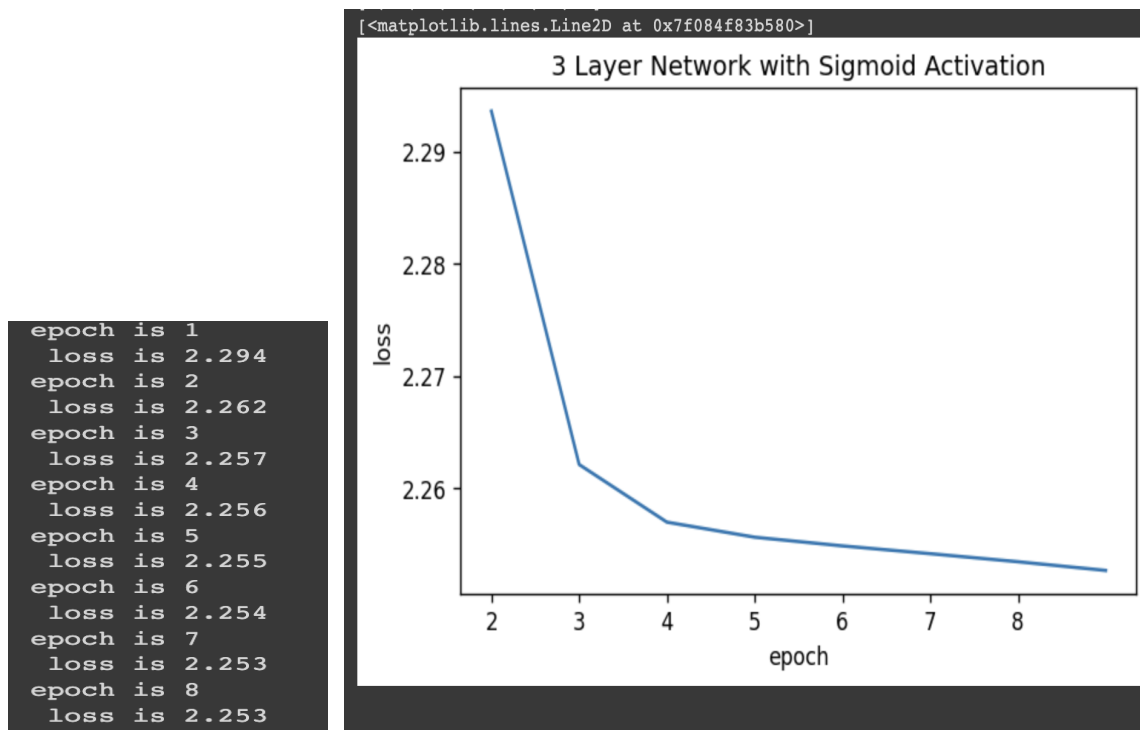
Imported the dataset CIFAR-10 using library dataloaders and transformed the 3 channel image into a single channel image right after importing the data to avoid redundant processing of 3 channel data.

Labels are: airplanes , cars, birds, cats, deer, frogs, horses, ships, and trucks.

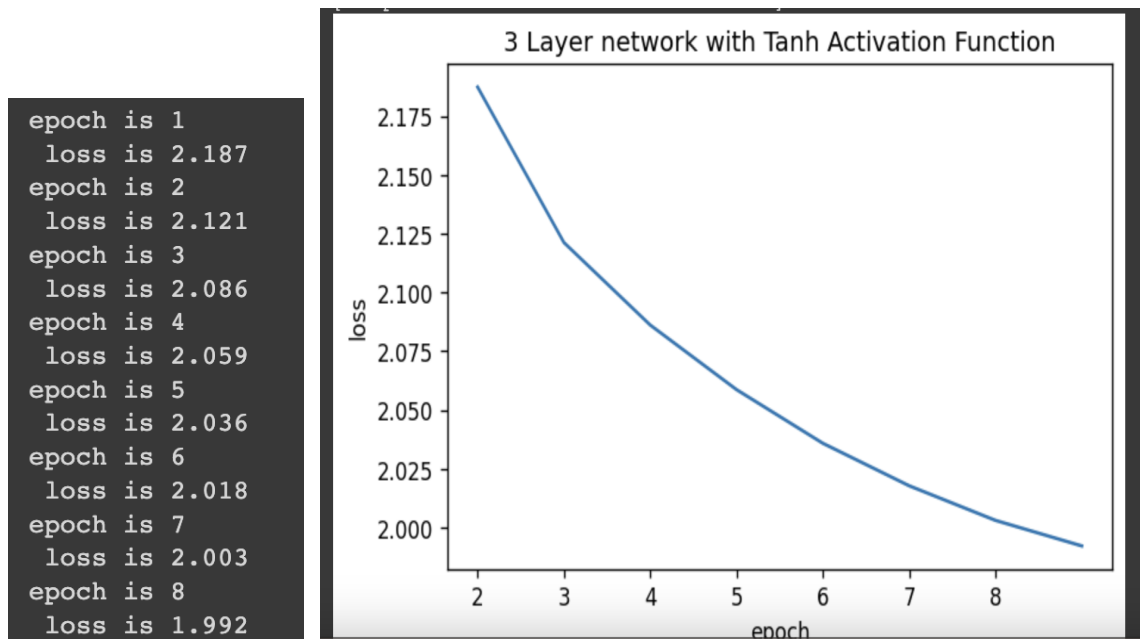


Created neural network with batch size 1000. Hidden Layer Nodes - 10. Learning rate-0.1/0.01. Different models are created for different activation functions and also when introducing more depth into the network. General training/testing method is defined to be called by different models with model objects to avoid any overlapping between the models parameters.

- Results using 3 layer network with Sigmoid Activation:-



- Results using 3 layer network with Tanh Activation :-



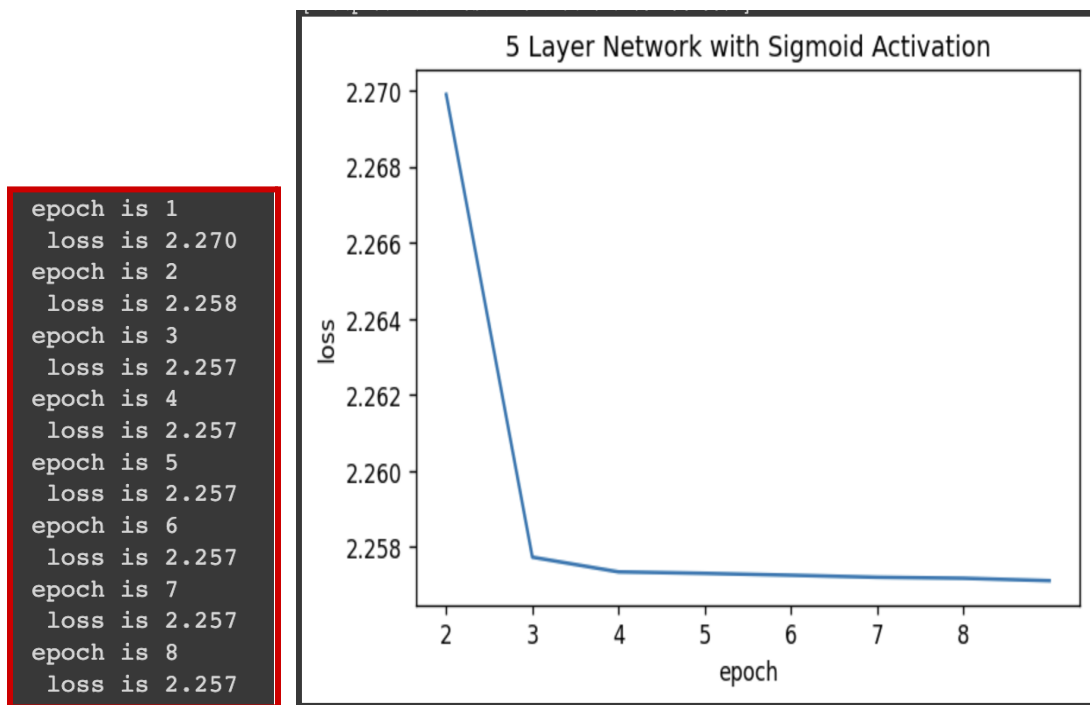
Tanh shows almost the same relative growth between the epochs whereas the Sigmoid function showed a good performance in early epochs and as the number of epochs increased, the loss was decreasing at a comparatively lesser rate.

---

(B) Introducing Vanishing Gradient Problem by increasing the depth of the network.

In the below image, one can see that the loss is stagnant after a few epochs as no new learning is happening in the model.

5 layer network with Sigmoid Activation Function:-



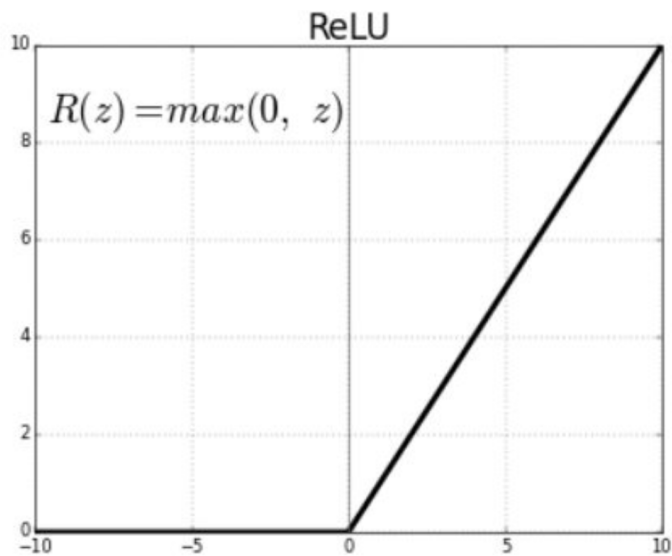
After Epoch 4, the loss became stagnant.

With an additional number of layers to the network, the partial derivative of the loss function reaches almost zero and furthermore, the value of weights does not change. And so, in the epoch the loss value remains the same. The last epoch value of loss is "2.257"

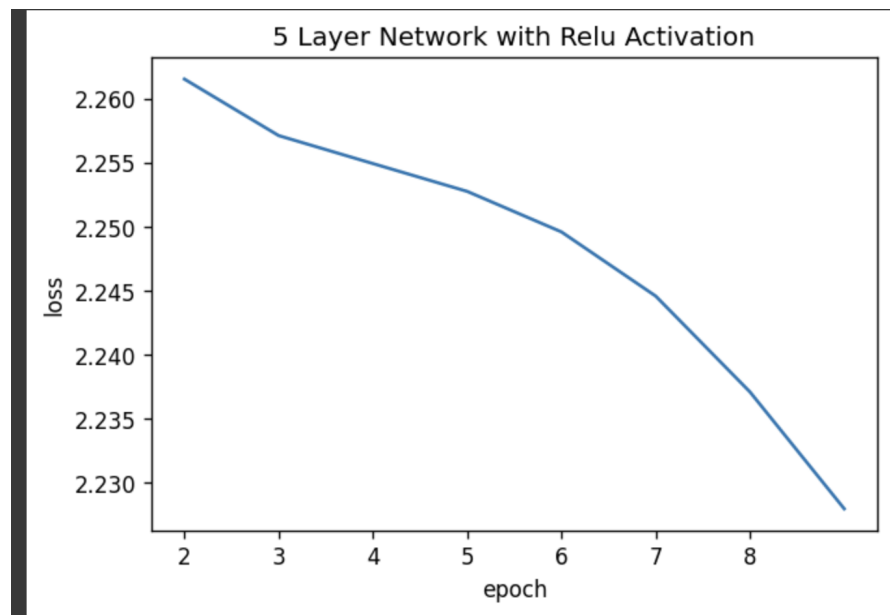
---

(C) Method to overcome the problem of Vanishing gradient

Using ReLU activation function in a 5 Layer network. ReLU activation has range from  $[0, z]$ .



```
epoch is 1  
  loss is 2.262  
epoch is 2  
  loss is 2.257  
epoch is 3  
  loss is 2.255  
epoch is 4  
  loss is 2.253  
epoch is 5  
  loss is 2.250  
epoch is 6  
  loss is 2.245  
epoch is 7  
  loss is 2.237  
epoch is 8  
  loss is 2.228
```



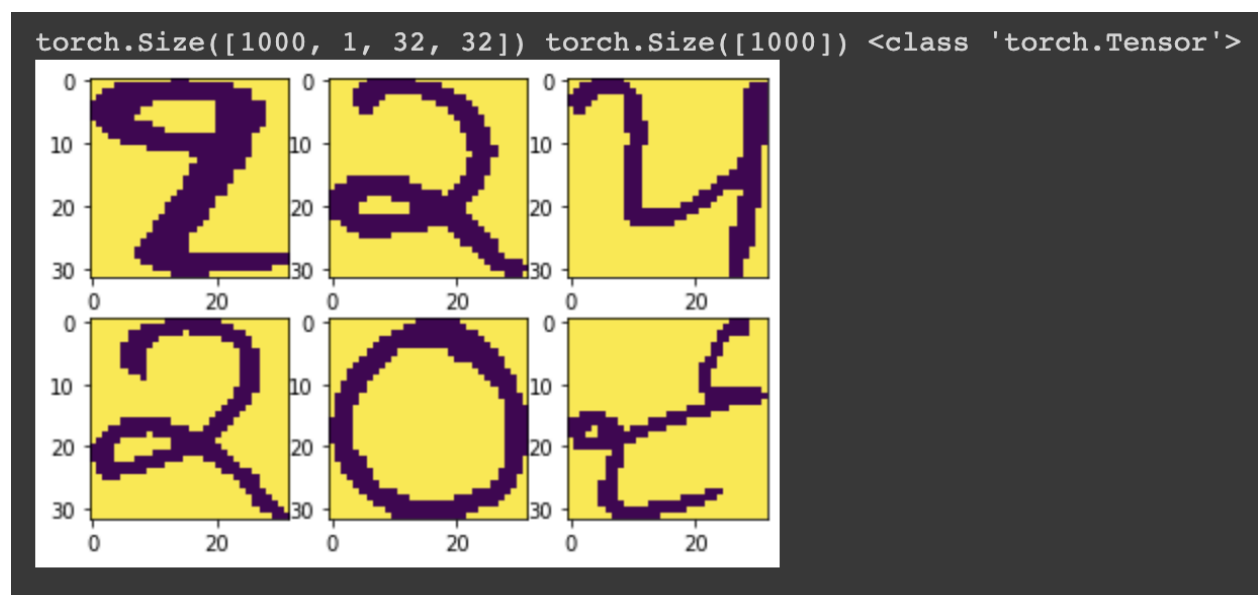
Here, after the 8th epoch, the loss value is "2.228" ( percentage decrease of 2.9% than sigmoid).

In the above graph, we can observe that there is no break in the curve, as one observed in the case of Sigmoid activation function. There is continuous decrement in the loss value even after the 4th epoch.

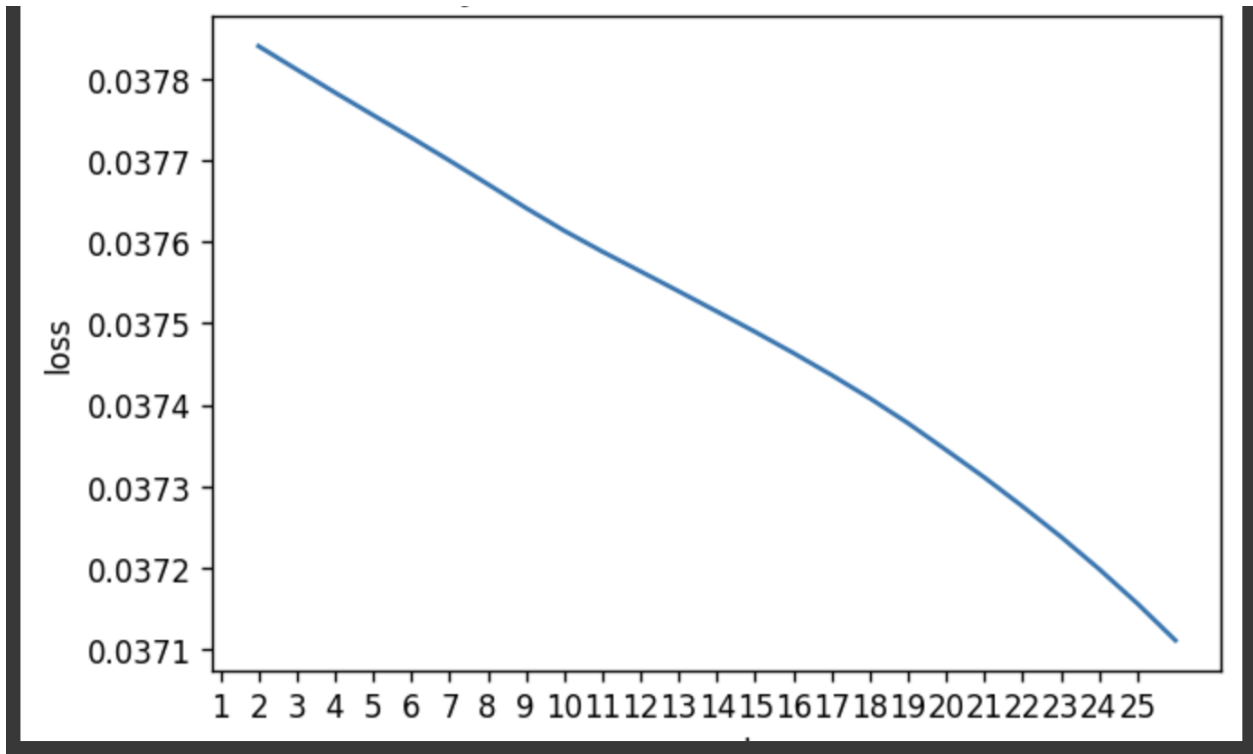
Question 2.

Implement a neural network on the Gurmukhi dataset and implement the following regularization techniques.

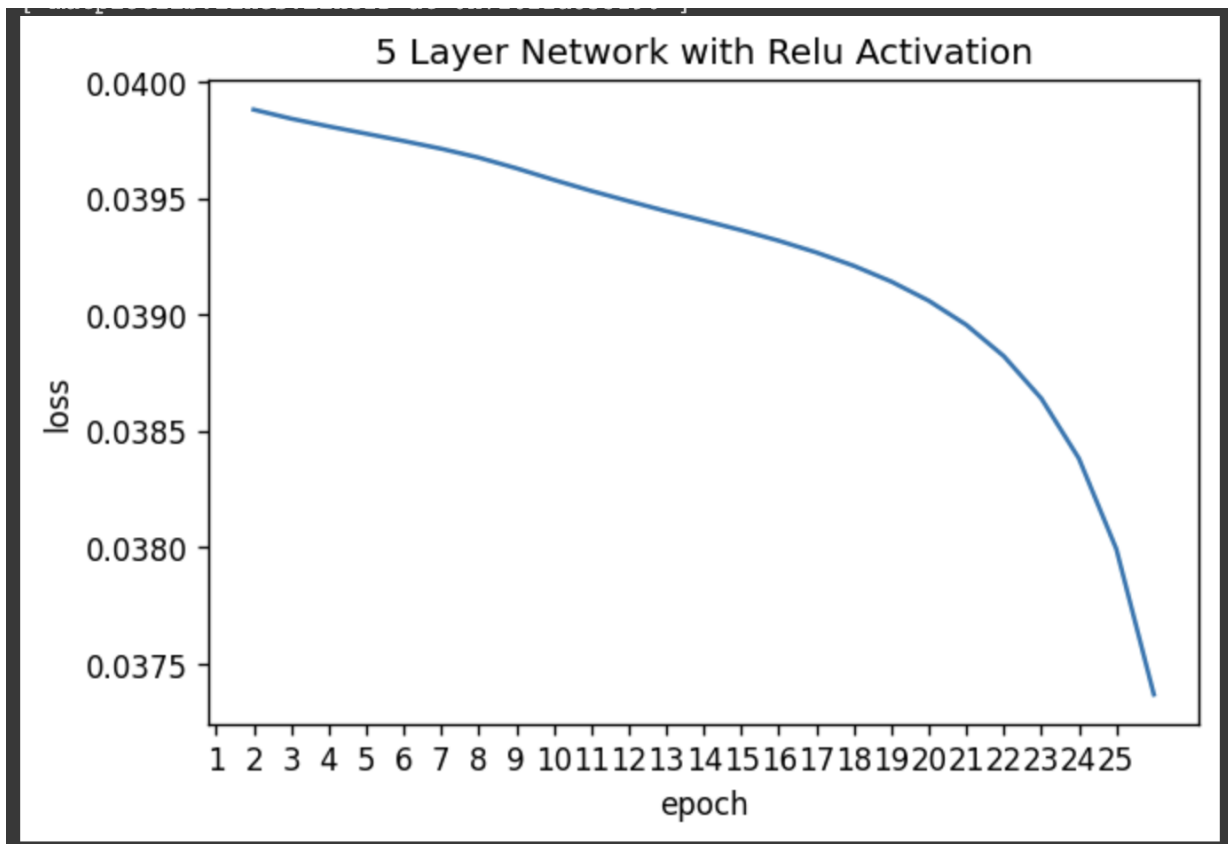
Imported the dataset using library ImageFolders and transformed the 3 channel image into a single channel image right after importing the data to avoid redundant processing of 3 channel data.



Results from training neural network with L1 Regularisation using the first norm of the weights parameter of model :-



Results from training neural network with L1 Regularisation using the first norm of the weights parameter of model :-



Without Regularization:-

```
test_accuracy = test_method(model_relu_5)

test accuracy = 8.98876404494382
```

With Regularization:-

```
test_accuracy = test_method(model_relu_5_12)

test accuracy = 17.97752808988764
```