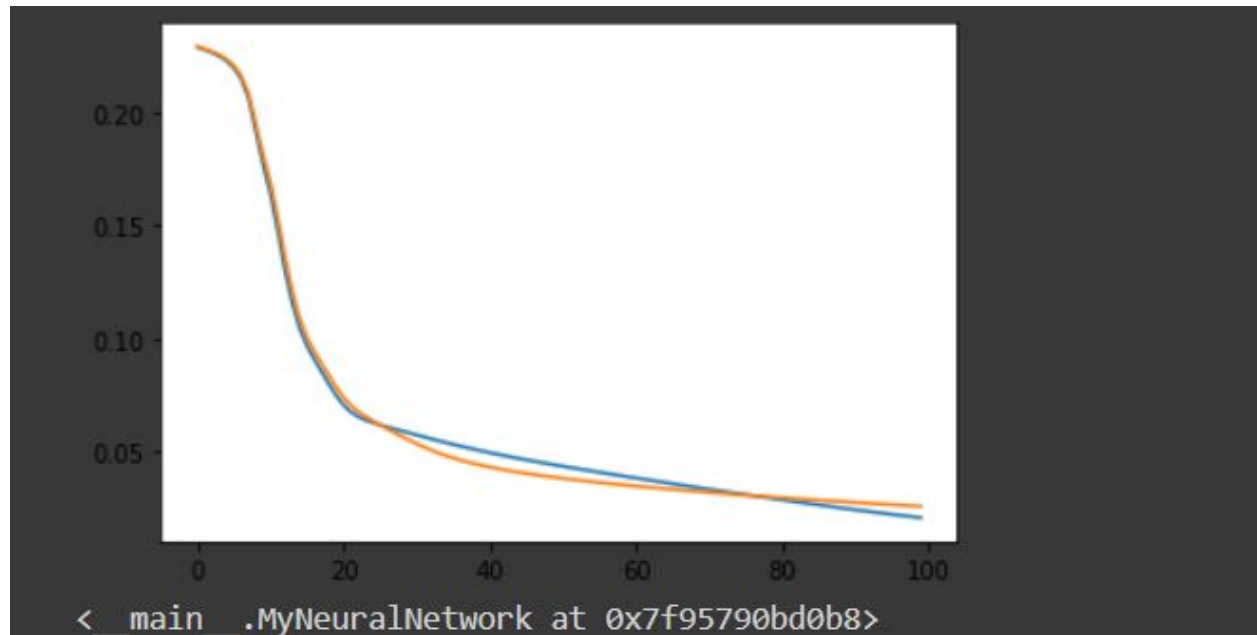


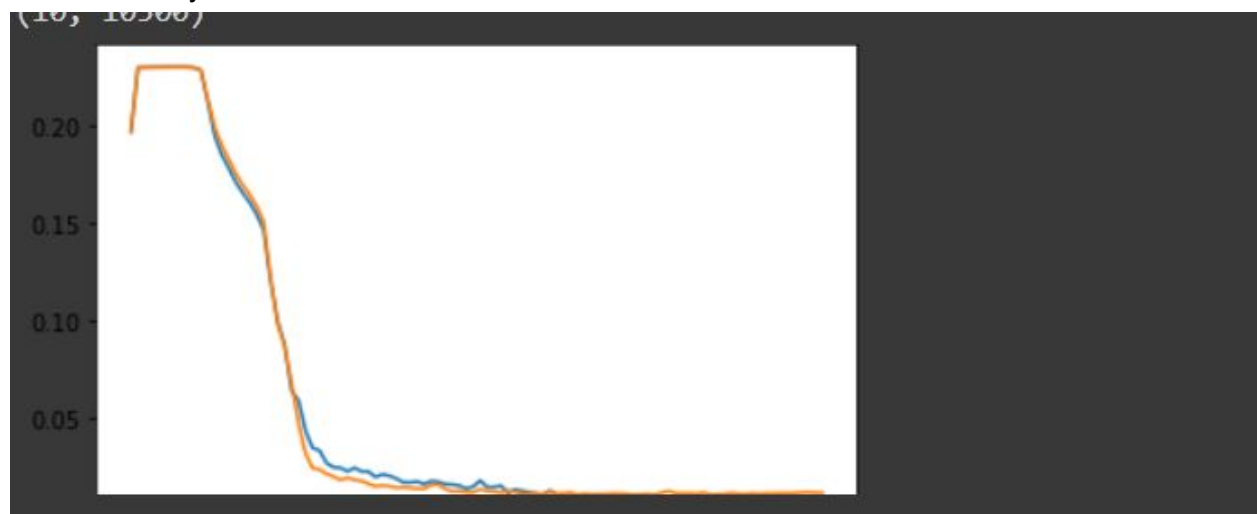
Q2.

A and B

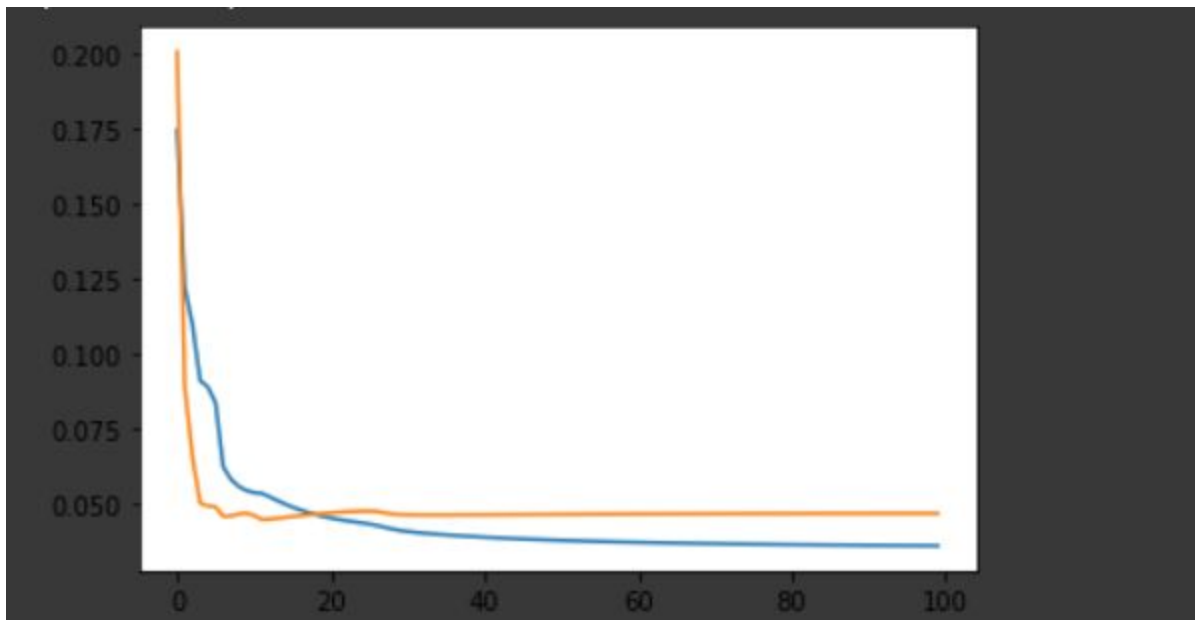
Sigmoid Accuracy= 0.945



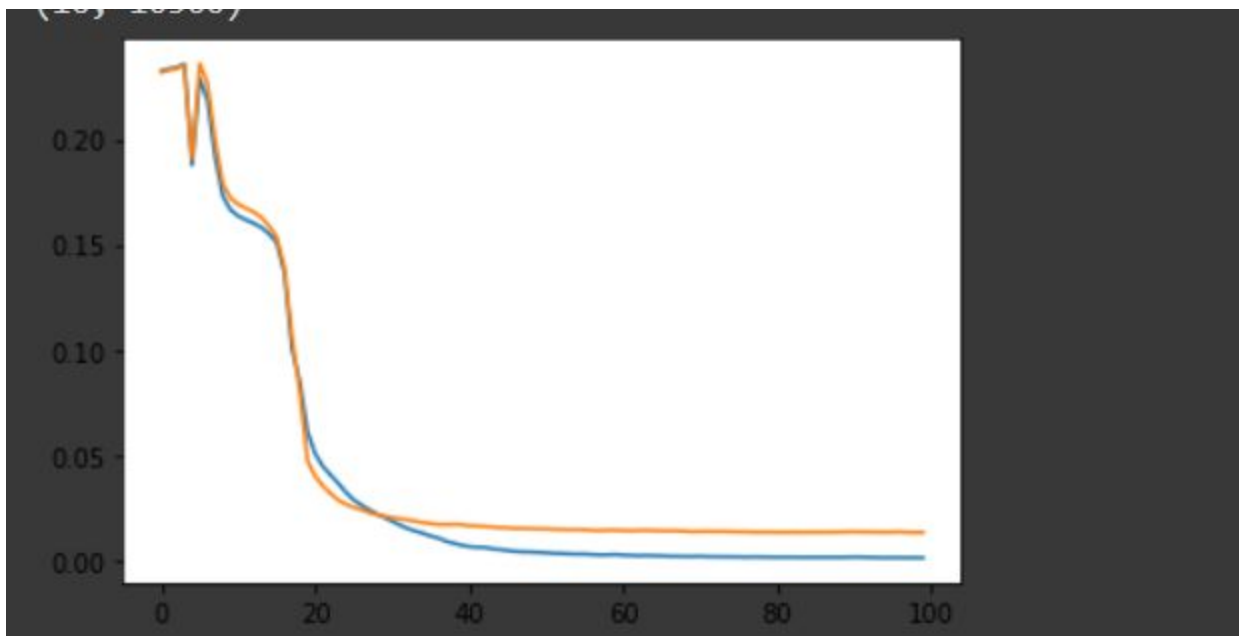
Relu Accuracy= 0.963



Linear Accuracy=0.900



Tanh Accuracy=0.971



C.

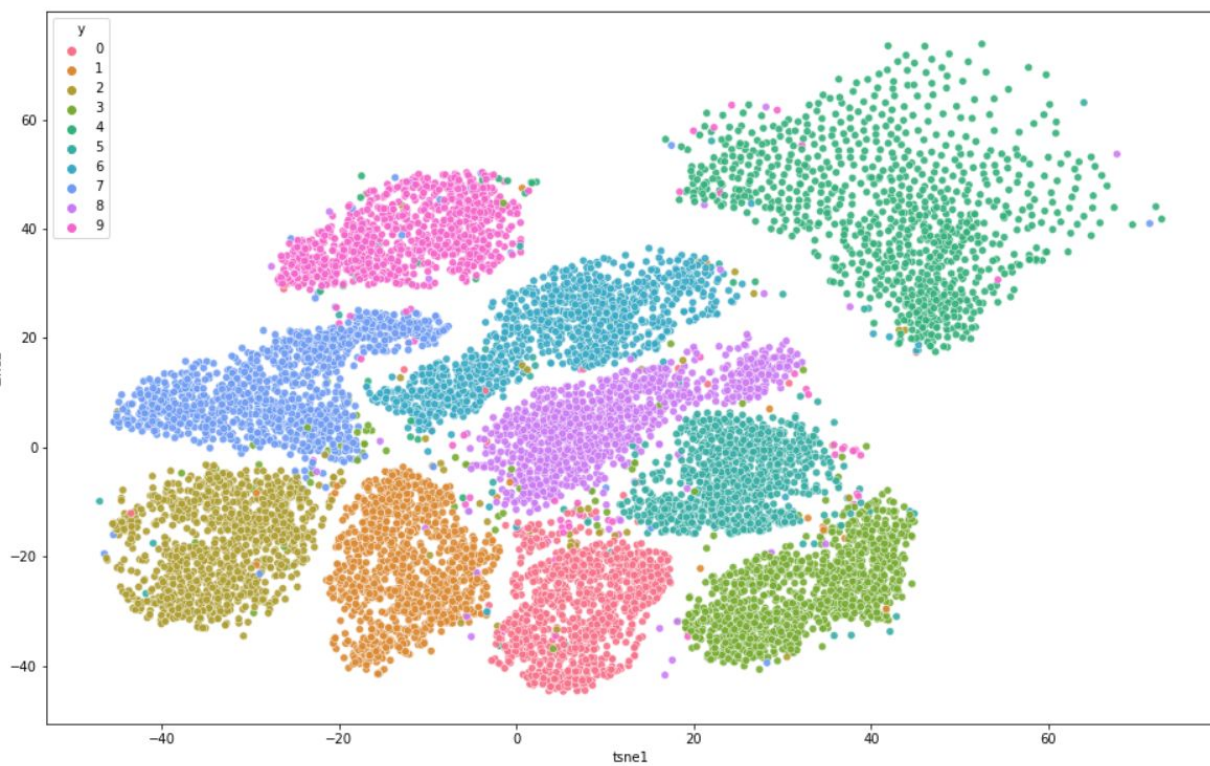
But in multi-classification, we need more than one output to know the probabilities of all the classes unlike in binary classification, so softmax is better option and gives the probabilities and distributes the probability throughout each output node.

D.

Total number of layers=5

Hidden layers=3

E



F

relu

acc=0.9745714285714285

Tanh

0.9657142857142857

Sigmoid

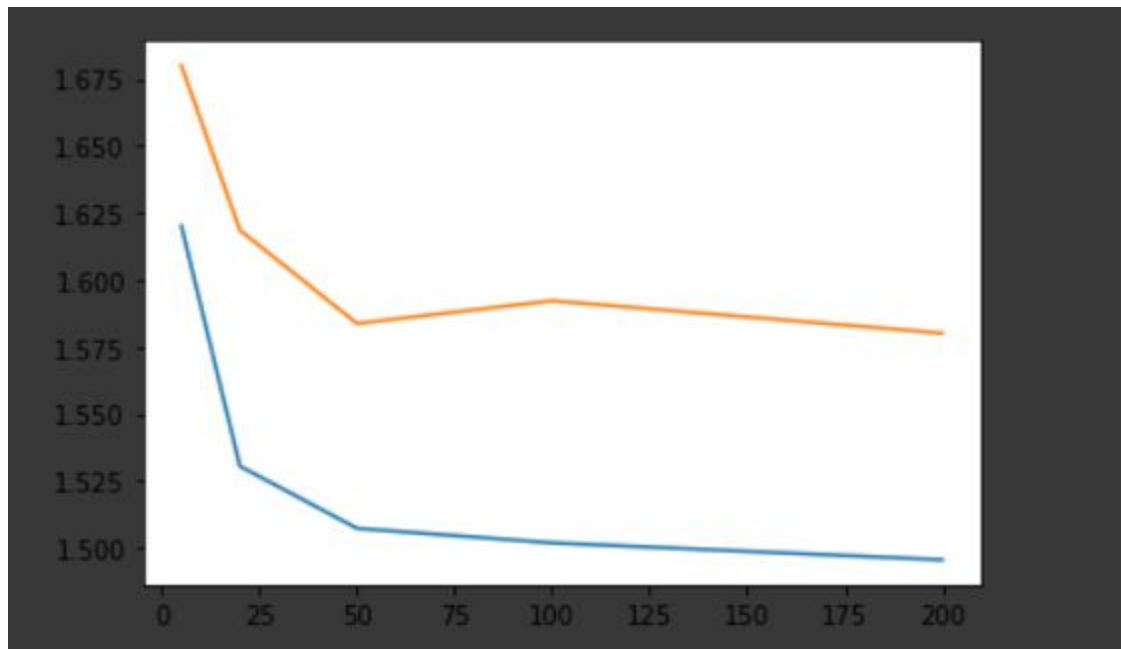
0.9064761904761904

Linear

0.8002857142857143

Q3

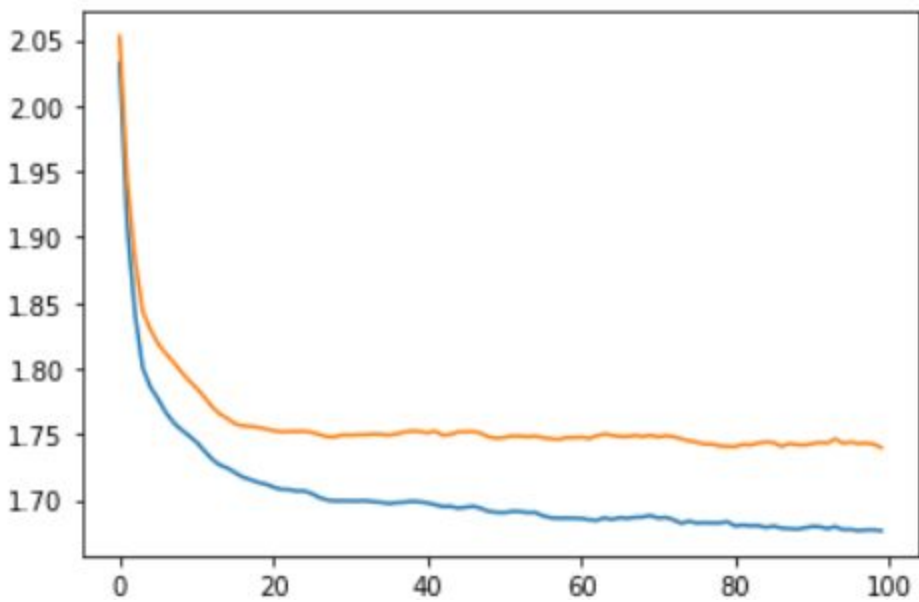
a



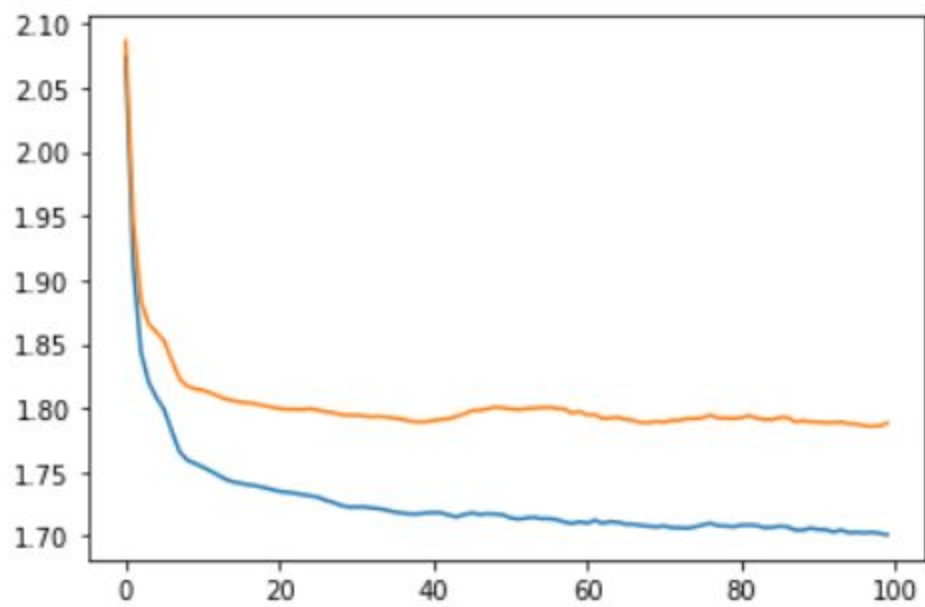
Training Loss goes down as we increase the hidden units to converge model however validation loss increases slightly on hidden unit=100, but goes down again at 200.

B

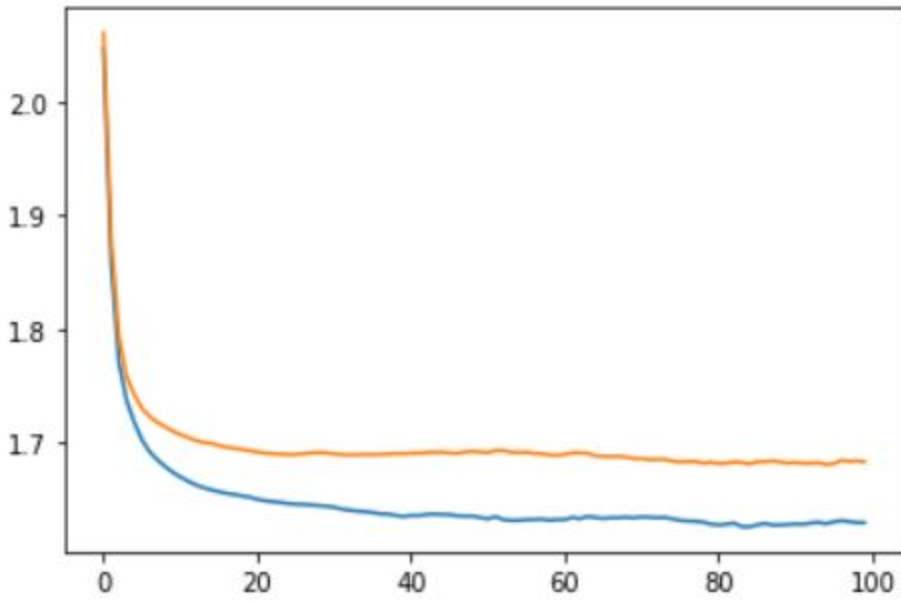
lr= 0.1



$lr=0.01$



$lr=0.001$



Graph converges after a few epochs at all the learning rates, but smoothly converges on alpha 0.001.

There is a noise at higher learning rate probably because of overshooting.

Q4

