

## Assignment4 :Toy Neural Network

Entry no: 2021MCS2132

Name: Gaurav Milind Chaudhari

Email: [mcs212132@iitd.ac.in](mailto:mcs212132@iitd.ac.in)

I implemented toy neural network in c . with following parameters

1st argument: number of layers including output layer eg: 3 in case 2 hidden 1 output layer

2nd argument : number of neurons in each layer in list format

Eg: 10, 10, 2

3rd argument :activation function choice

Eg:

1 for sigmoid

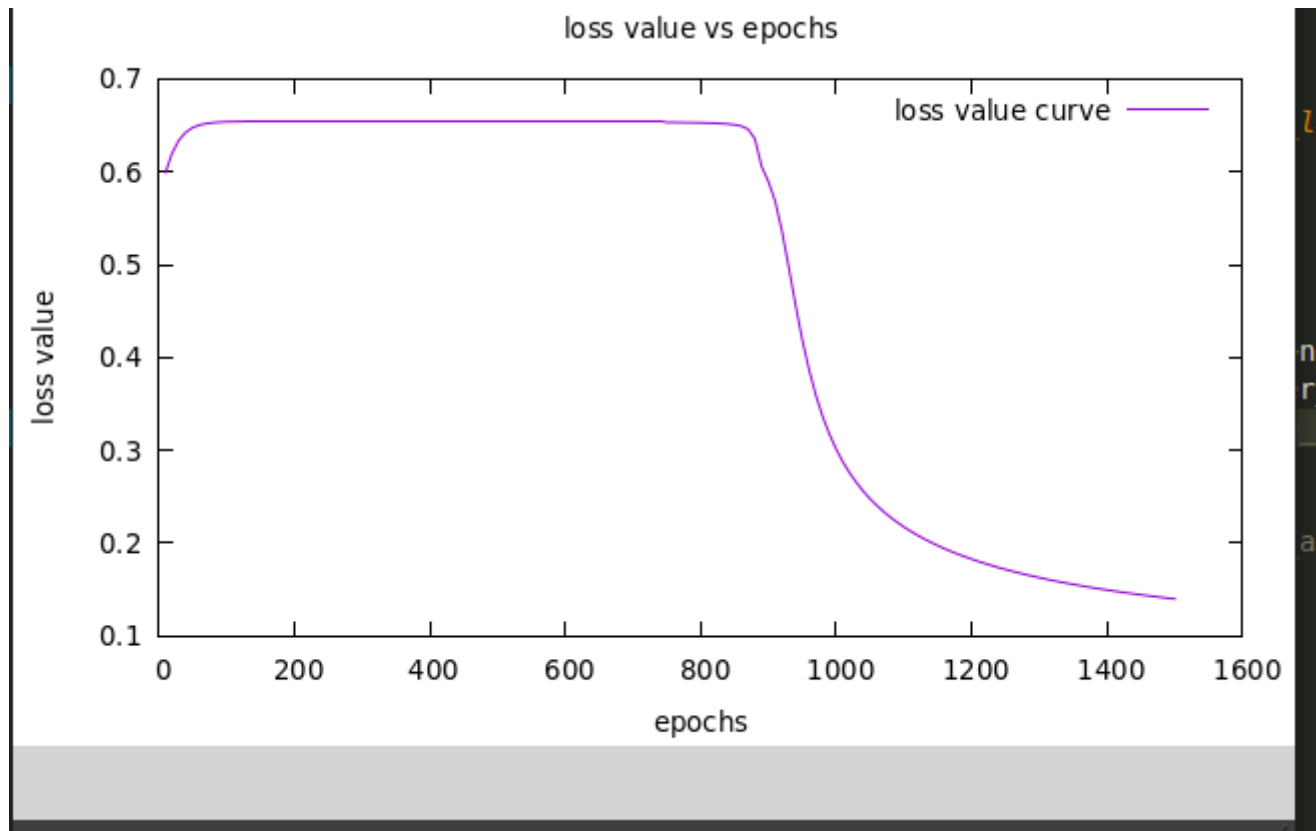
2 for tanh

3 for relu

4th argument is number of epochs to run eg: 1000

5th argument: learning rate e.g 0.1

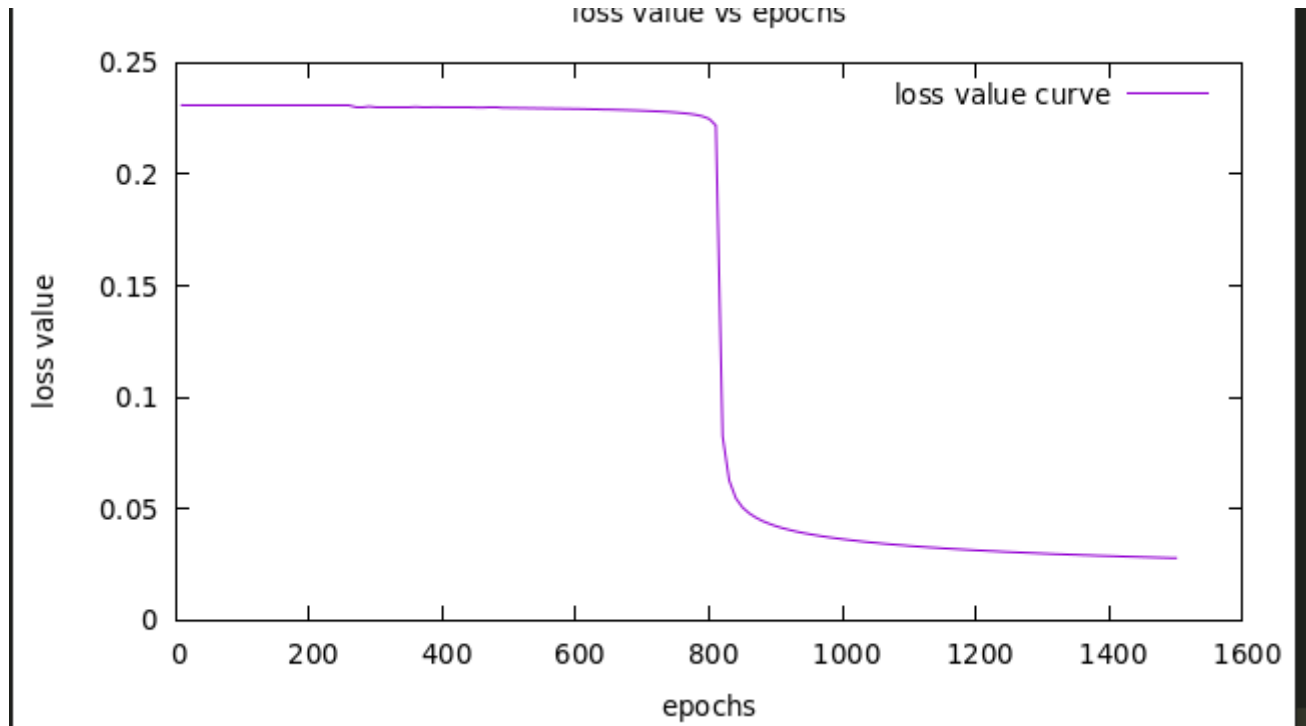
Argument output: fun sigmoid, layers 10,2, output sigmoid



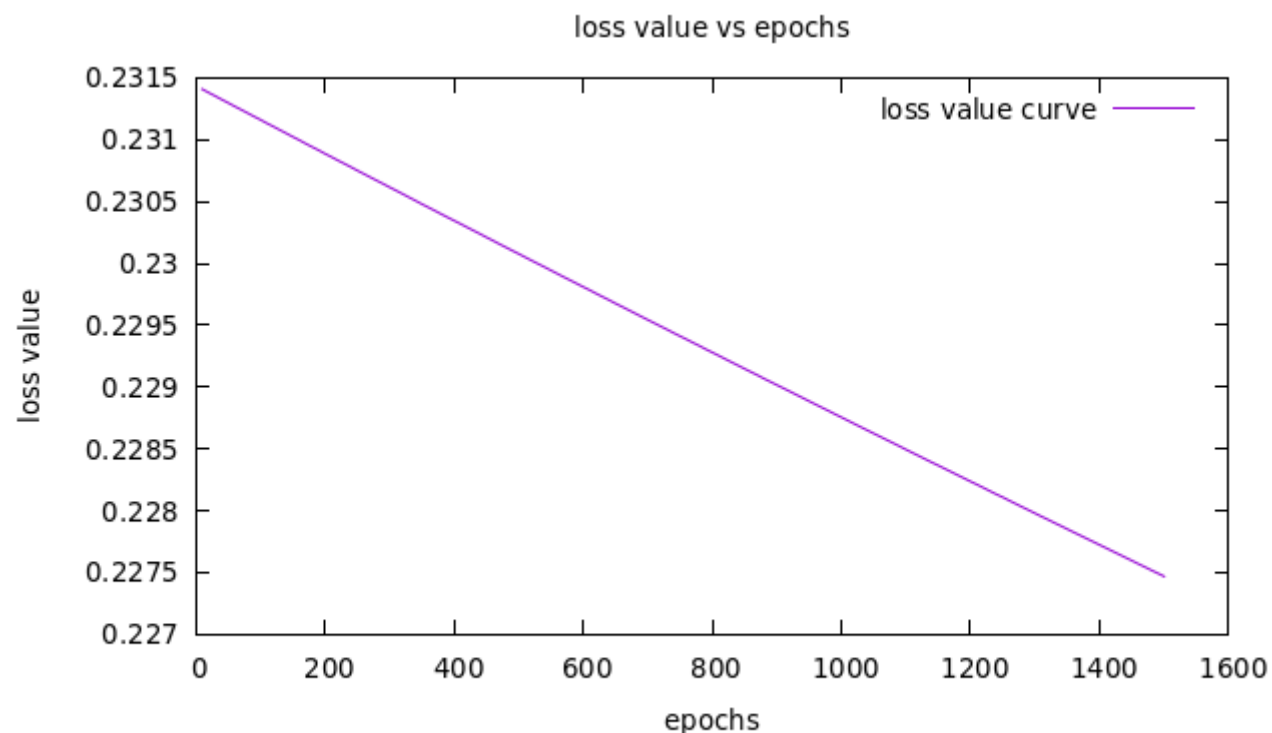
As we can see loss value remains constant at approximately 800 epochs and start decreasing from there

Epochs: 1500 final test accuracy 94.011976 percent

For argument: tanh epochs 1500 layer 10, 2 learning rate 0.1



The loss function for stipper for tan h function accuracy also high  
final test accuracy 95.209581 percent



Loss value for rely function not giving output as expected there might be error in  
implementation but, learning is also slow

Final test accuracy

final test accuracy 60.479042 percent

Sklearn implementation

Relu 1500 epochs sgd Gives 95.1048951048951% accuracy

Sigmoid 1500 epochs sgd gives 90.83916083916084 %accuracy

Tanh 1500 epochs sgd gives 95.804195804195 % accuracy

Gradient vanishing problem occurs for sigmoid and tanh function:

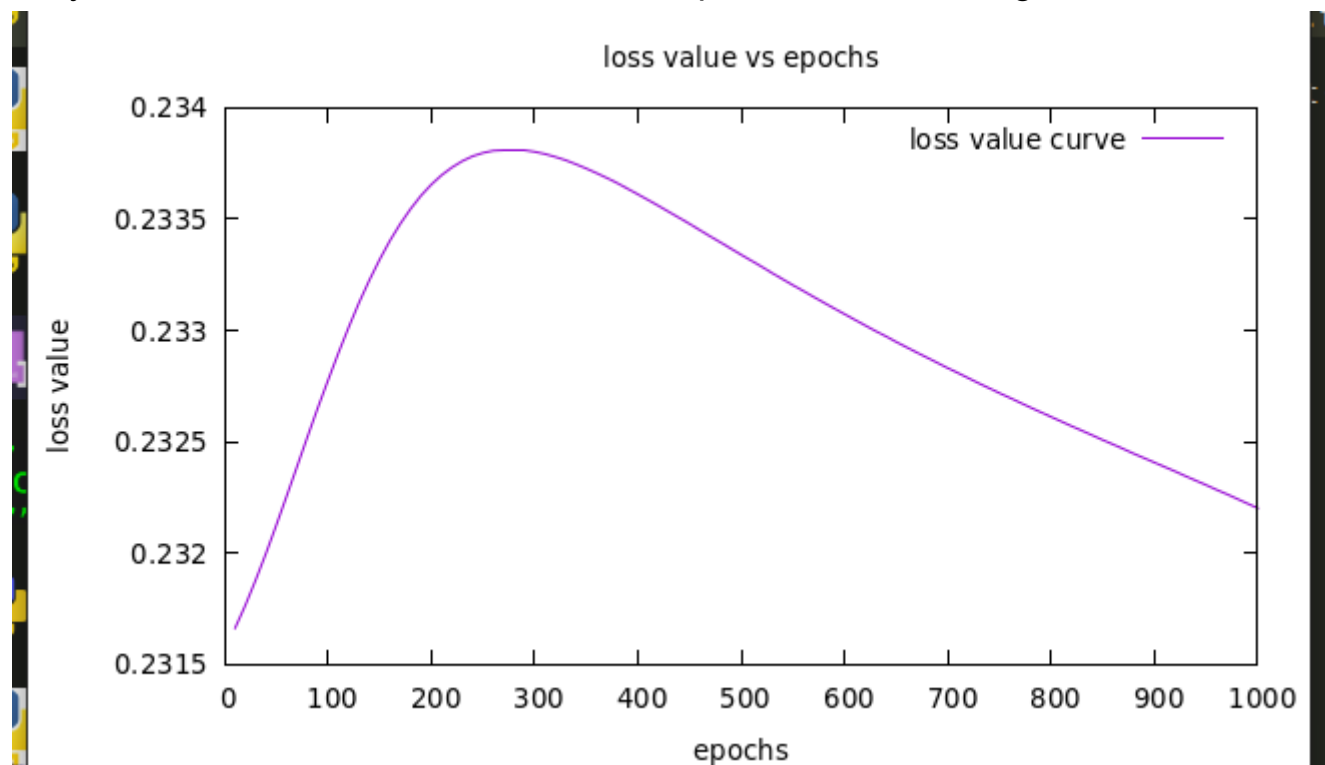
It is because the sigmoid function range is 0 to 1 but its derivative range is 0.25 hence we reduce hence we keep losing the range that we operate hence we face improper learning rate

Slow convergence or increase in the loss function

Here I tried to implement the lost gradient descent problem

These are specifications for the arguments

3 layers 10,20,2 activation tanh function epochs 1000 learning rate 0.1



Here as we can see for 3 layers

Conclusion:

Implementation works for sigmoid function and tanh function but still fails to get good learning rate of relu function. Problem faced in assignments is mostly incorrect implementation of mathematics I stripped down some requirements to get a working Network. I tested the same network on another multiclass dataset it works correctly.