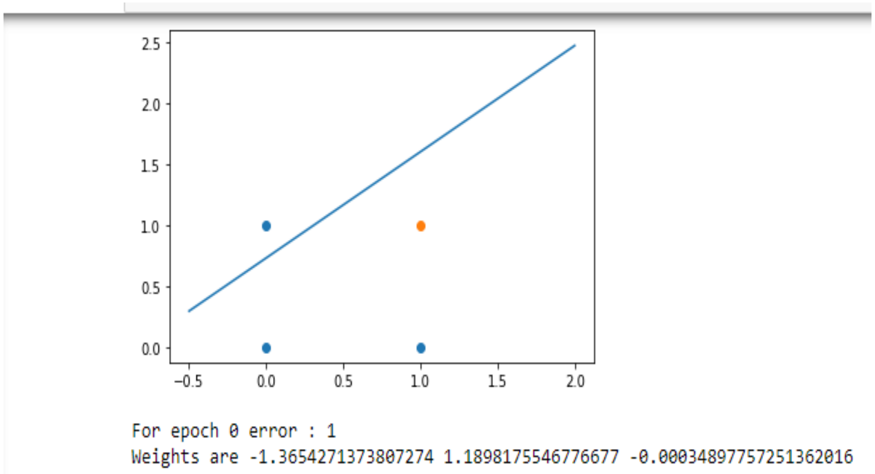
PART I: PERCEPTRON TRAINING ALGORITHM (PTA) (Shivam Dalmia MT20304)

1.

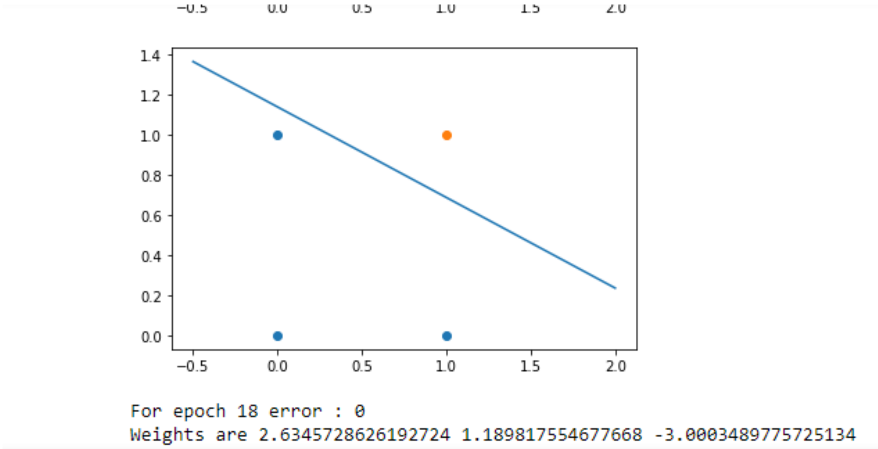
The decision boundary for AND and OR gates is :

AND:



```
For epoch 0 error : 1
Weights are -1.3654271373807274 1.1898175546776677 -0.00034897757251362016
```

This is for the first epoch the error and weights are also given.



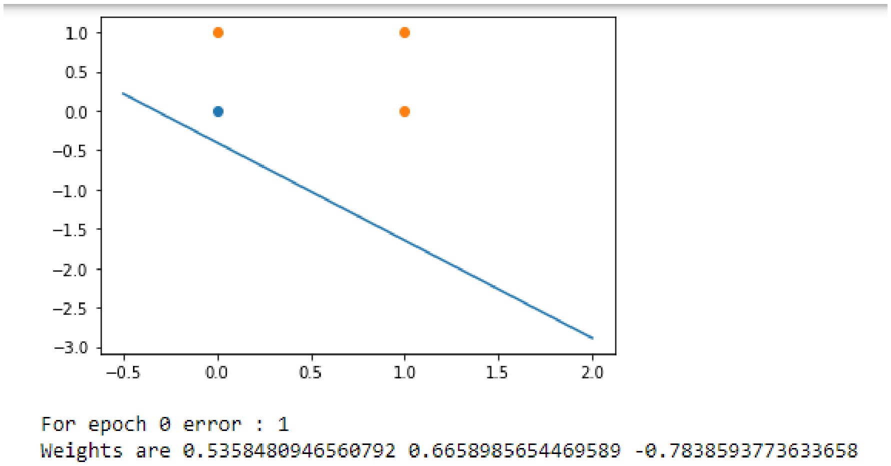
```
For epoch 18 error : 0
Weights are 2.6345728626192724 1.189817554677668 -3.0003489775725134
```

This is of epoch 18 when the model converges giving the error as 0 and the respective weights and the decision boundary also isolating all the classes point.

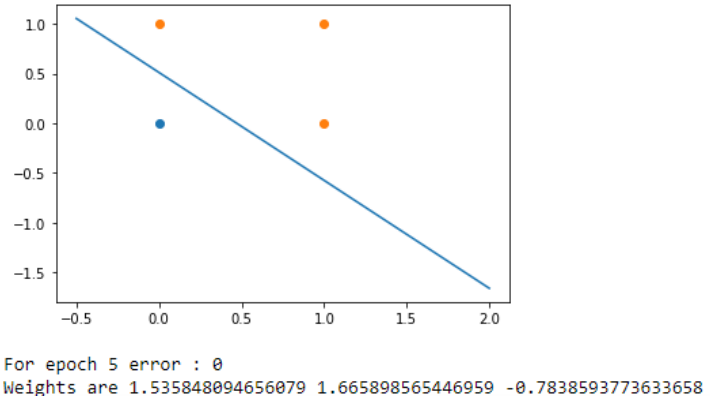
Pickle answer: W1: 2.6345728626192724 W2: 1.189817554677668 b: -3.0003489775725134

The pickle file is AND.pkl

OR:



```
For epoch 0 error : 1
Weights are 0.5358480946560792 0.6658985654469589 -0.7838593773633658
```



```
For epoch 5 error : 0
Weights are 1.535848094656079 1.665898565446959 -0.7838593773633658
```

This is of epoch 5 when the model converges giving the error as 0 and the respective weights and the decision boundary also isolating all the classes point

Pickle answer: W1: 1.535848094656079 W2: 1.665898565446959 b: -0.7838593773633658

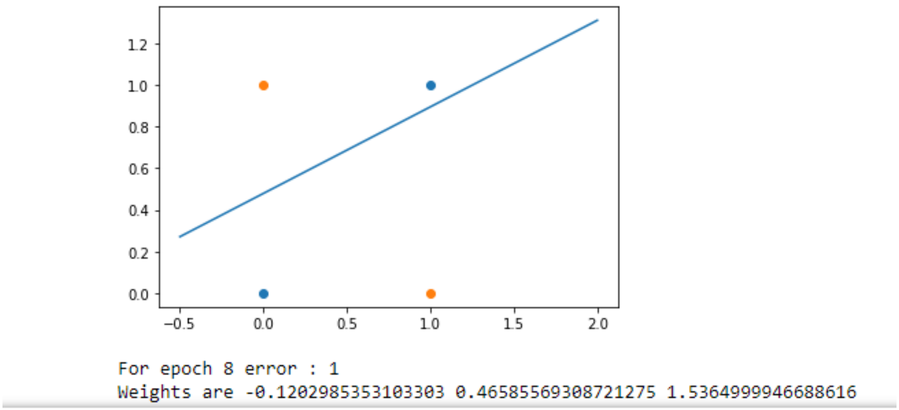
The pickle file is OR.pkl

NOT:

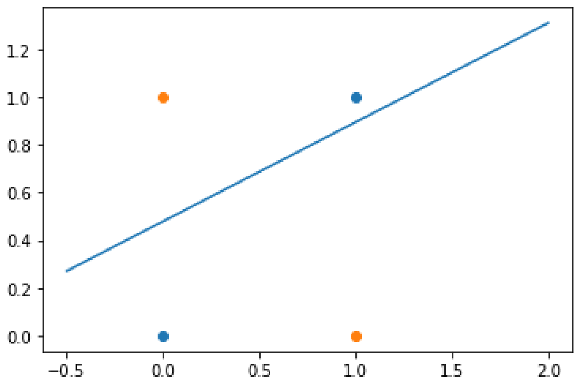
Pickle answer:
W1: -1.2063730732266231 b: 0.36549163213751756

The pickle file is NOT.pkl

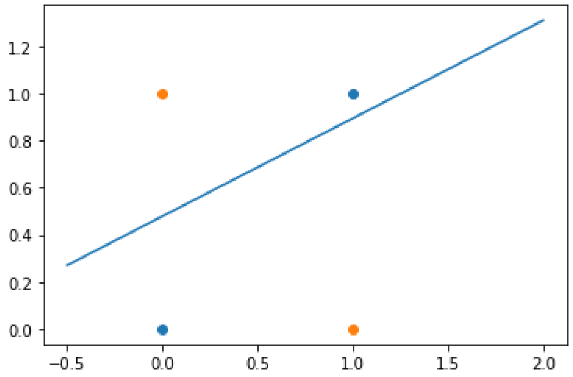
XOR:



```
For epoch 8 error : 1
Weights are -0.1202985353103303 0.46585569308721275 1.5364999946688616
```



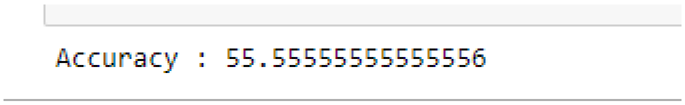
```
For epoch 12 error : 1
Weights are -0.1202985353103303 0.46585569308721275 1.5364999946688616
```



```
For epoch 16 error : 1
Weights are -0.1202985353103303 0.46585569308721275 1.5364999946688616
```

As you can see in epoch 8 the weights are -0.1202985353103303 0.46585569308721275 1.5364999946688616 and the same weights are also in epoch 12,16 and so on the weights of the 9th epoch is repeated in 13,17 and so on thus concluding the change in the weights are stuck in the infinite loop and the error will always remain 1.

2)



```
Accuracy : 55.55555555555556
```

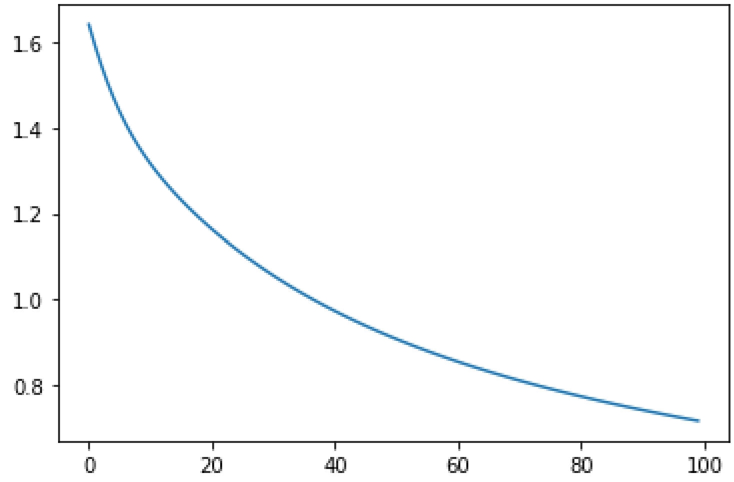
The accuracy of the Madeleine learning algorithm is 55.55% using 13 neurons.

The pickle file is Madeline.pkl

PART II: MULTILAYER PERCEPTRON (MLP) (Dipankar Sarkar MT20307)

1.
First, we trained with learning_rate 0.01 and we assume to have one hidden layer with 128 neurons. The last layer has 10 neurons. With each data having 784 features with hidden layer as tanh and last layer as softmax.

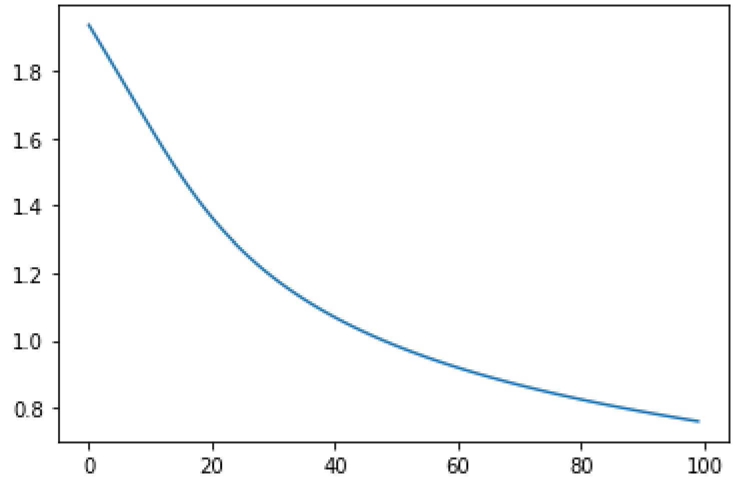
The accuracy of the Network is 42.91



The loss seems to decrease with each epoch and by 100 th epoch it reduces to 0.71

We tried with hidden layer as sigmoid and last layer as softmax

The accuracy of the Network is 17.78

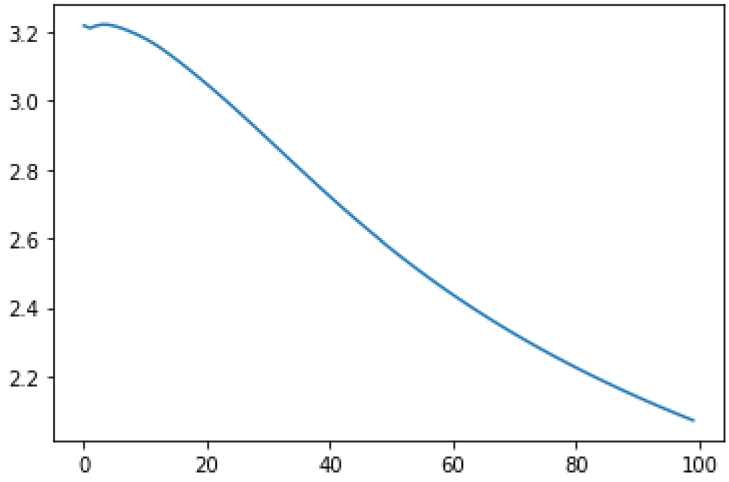


The loss seems to decrease with each epoch and by 100 th epoch it reduces to 0.76

We tried with hidden layer as reLu and last layer as softmax

The accuracy of the Network is 44.57

In terms of accuracy, it is slightly better than tanh but the tanh graph shows a better convergence. So we will proceed with tanh as activation in the hidden layer.

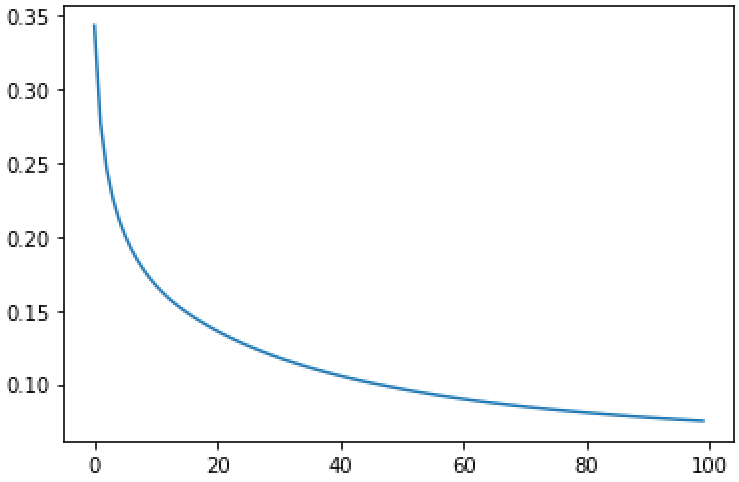


2.

a)
Mini Batch 64 with gradient and momentum

We take momentum as 0.9

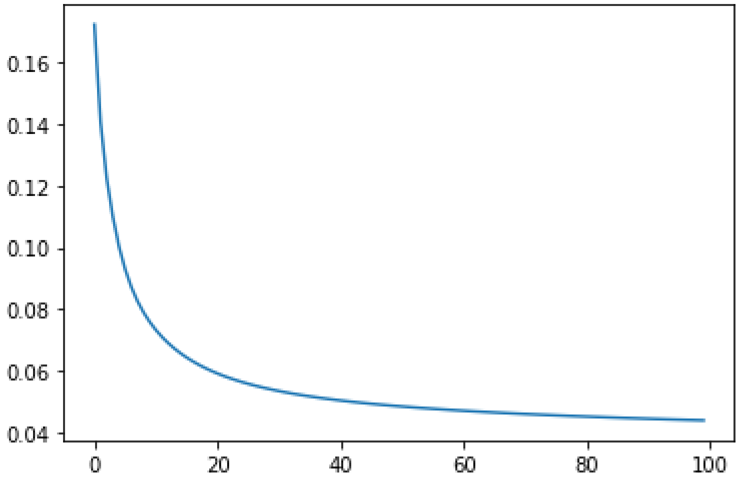
The accuracy of the Network is 82.28



The losses reduces significantly and at 100 th epoch it reaches 0.07

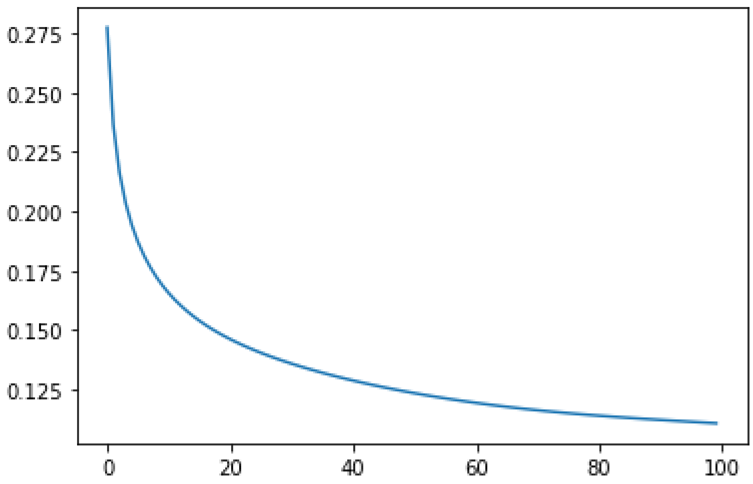
b)
Mini Batch 64 with Nesterov's Accelerated Gradient

Accuracy of the Network is 85.21



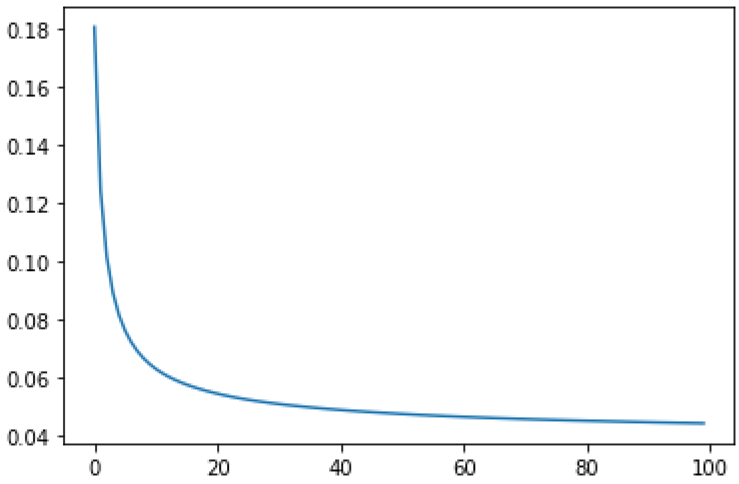
c)
Mini Batch 64 with AdaGrad

Accuracy of the Network is 78.17



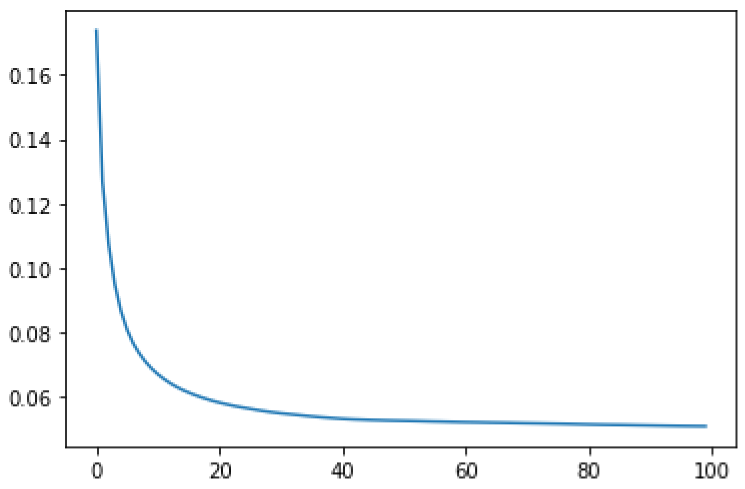
d)
Mini Batch 64 with Rmsprop

Accuracy of the Network is 81.33



e)
Mini Batch 64 with Adam

The accuracy of the Network is 83.28999999999999


0.16
0.14
0.12
0.10
0.08
0.06
0
20
40
60
80
100