CSCI 635 Homework 3

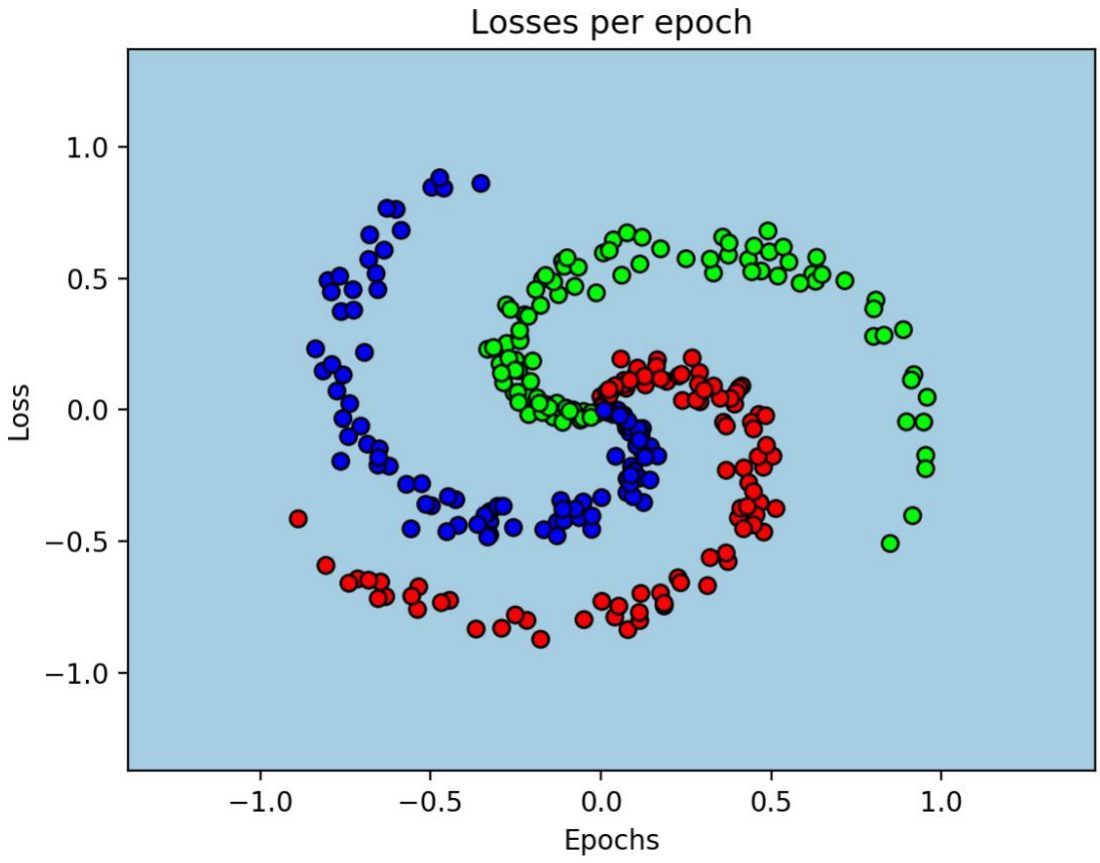
Name: Hitesh Vaidya (hv8322)

1.a)

With the selected combination of weights and biases, the secant approximation method indicated that the difference between theta parameters was 'incorrect'. Hence the graphs were not obtained further.
The maximum entropy model in last homework has no hidden layer while this has one. Due to this, the number of parameters change and even the backpropagation process changes.

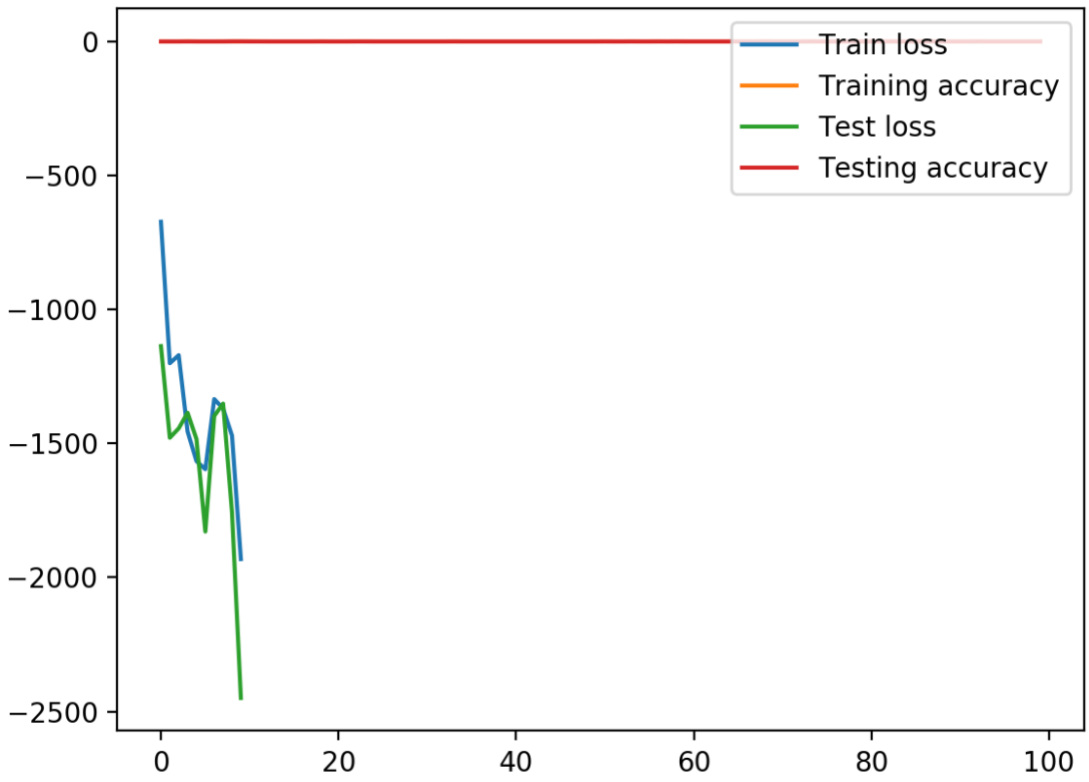
1.b)

I tried testing model with different hyperparameters. The performance of model depends totally on how the parameters are initialized. The decision boundary in this problem is as follows:



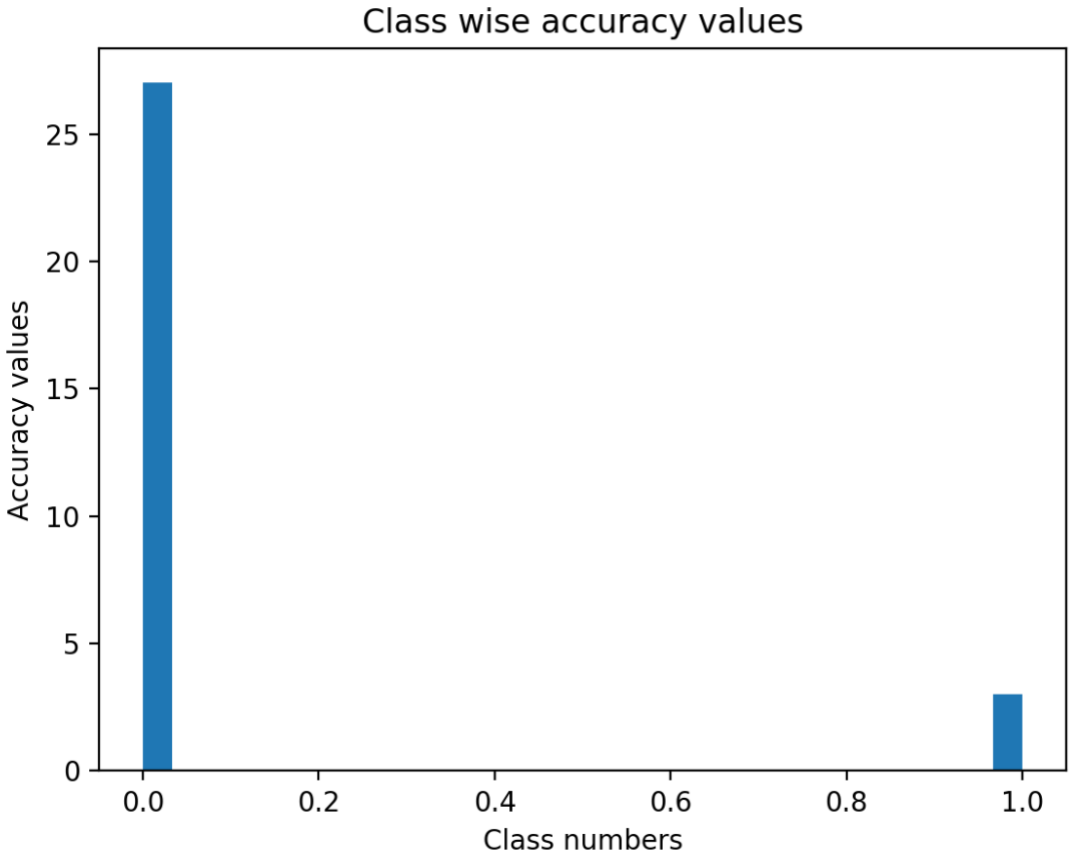
1.c)

The error graph Is as follows:



If the model is trained for more number of epochs, it goes on overfitting. The validation loss increases again (parabola) while the training loss goes on decreasing. It could be controlled by early stopping. Training could be stopped at the point from where the validation loss goes on increasing.

2)



Confusion matrix:

```
[[6903    0    0    0    0    0    0    0    0    0]
 [   0    0    0    0    0    0    0    0    0    0]
 [   0    0    0    0    0    0    0    0    0    0]
 [   0    0    0    0    0    0    0    0    0    0]
 [   0    0    0    0    0    0    0    0    0    0]
 [   0    0    0    0    0    0    0    0    0    0]
 [   0    0    0    0    0    0    0    0    0    0]
 [   0    0    0    0    0    0    0    0    0    0]
 [   0    0    0    0    0    0    0    0    0    0]
 [   0    0    0    0    0    0    0    0    0    0]]
```

As visible from results, this model does not perform well. It needs proper training and hyperparameter initialization.