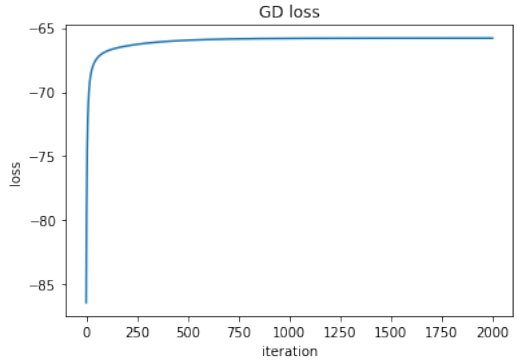
ROB 313: Assignment 3

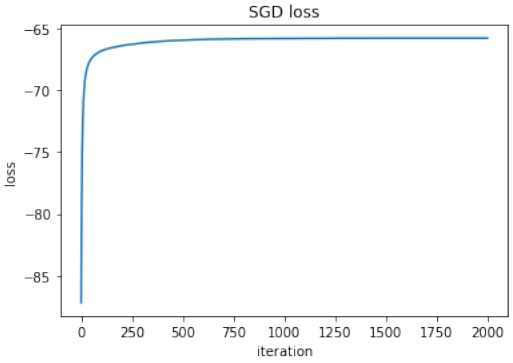
Ali Seifeldin 1003894431

In this assignment, we implement a GD and SVD model to predict on the iris dataset, before implementing a hidden layer with 100 neurons to predict hand written numbers.

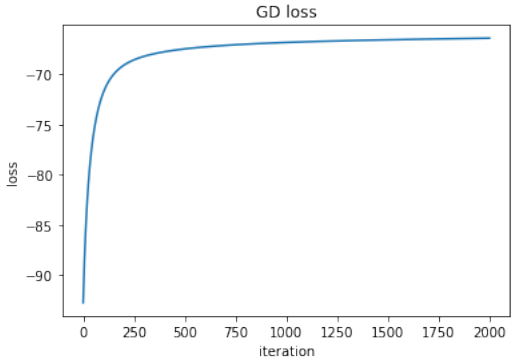
1) The value of the log likelihood would be -∞. This is reasonable behavior since the predicted value is completely off and the error would be at its maximum

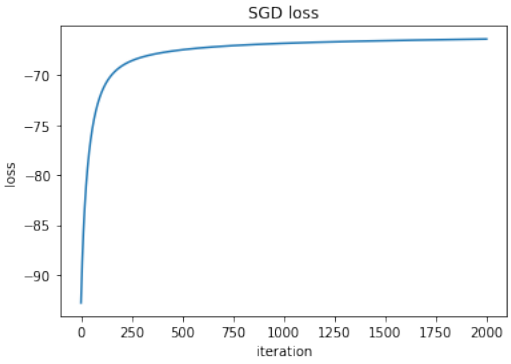
Plots of losses for a 0.005 learning rate.



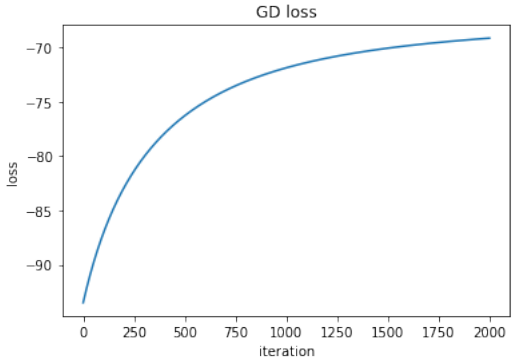


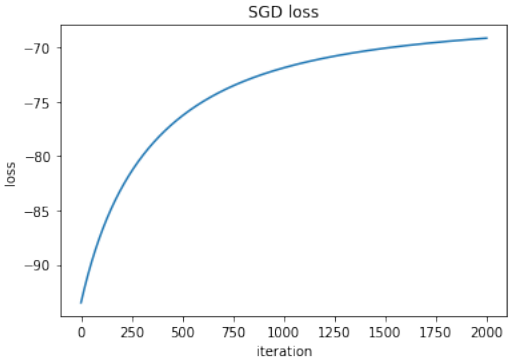
Plots of losses for a 0.0005 learning rate.





Plots of losses for a 0.00005 learning rate.





Using a learning rate of 0.005

	Test accuracy	Log-likelihood
GD	0.73	-6.9171321
SGD	0.73	-6.9324264

Using a Log-likelihood estimate would better, since we would get a better relative estimate of the error, as opposed to whether the result is above or below a threshold.

2)

a)

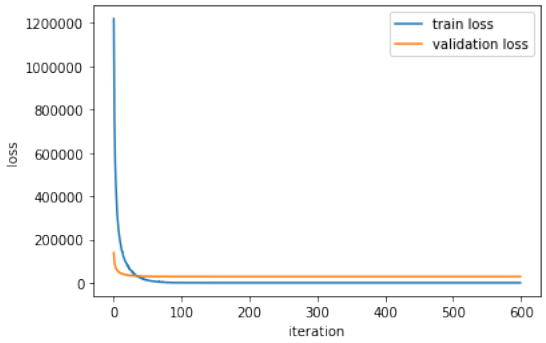
```
maxVals = np.max(Fhat,axis=1).reshape(-1,1)
Fhat = Fhat - maxVals
p = np.exp(Fhat)
p = np.sum(p,axis = 1,keepdims = True)
Fhat = Fhat - np.log(p)
```

Implementation is as follows. We find the max value per row, and minus the Fhat values by that. Thus the maximum value in Fhat will be 0 and all other numbers will be negative, thus never exploding

b)

```
nll = -np.sum(np.multiply(Fhat,y))
```

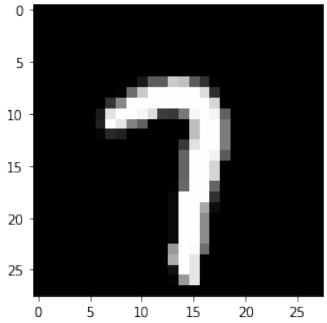
c)



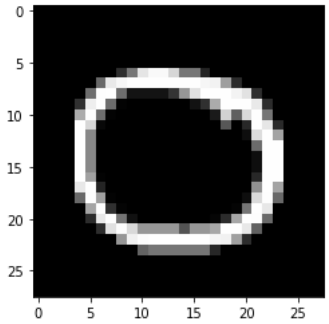
We can see that the training loss approaches 0, as the validation loss get more stable at about 28400. The network preforms with an accuracy of 1 for the training, .834 for the validation and .847 for the test set.

d)

predicts a 7 which is correct, but I see it as a one.



This 0 is predicted to be a 5, maybe due to certain aspects of its curve



This is predicted to be a 7 when it's a 2