

Answers for question 5

by Amogha Sekhar, A53301791

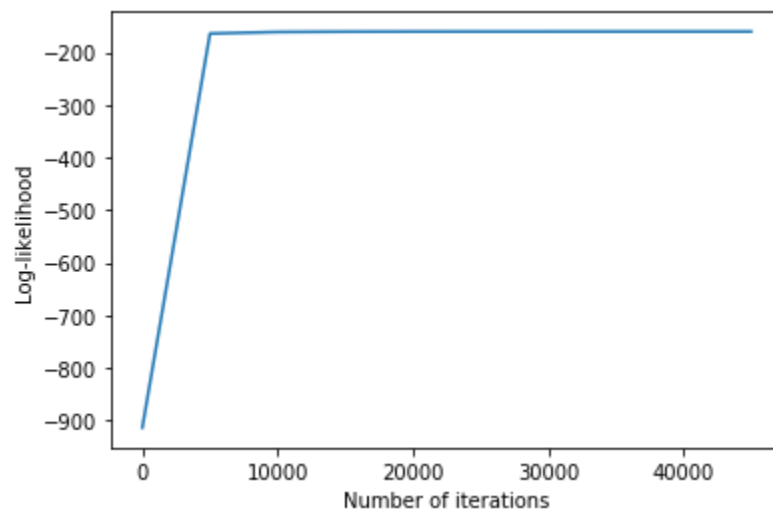
Part a)

Algorithm used: Gradient Descent

3: label 0, 5: label 1

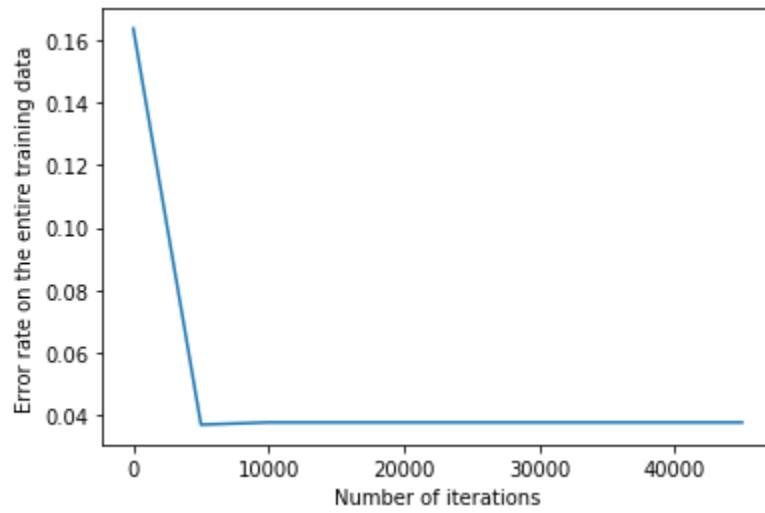
Log likelihood values:

Iteration: 0
-913.8316798294908
Iteration: 5000
-164.4766619007139
Iteration: 10000
-161.49202924399023
Iteration: 15000
-160.906220873847
Iteration: 20000
-160.75665309705118
Iteration: 25000
-160.71391112584024
Iteration: 30000
-160.70097193251004
Iteration: 35000
-160.69693340451414
Iteration: 40000
-160.69565612672773
Iteration: 45000
-160.69525337024479

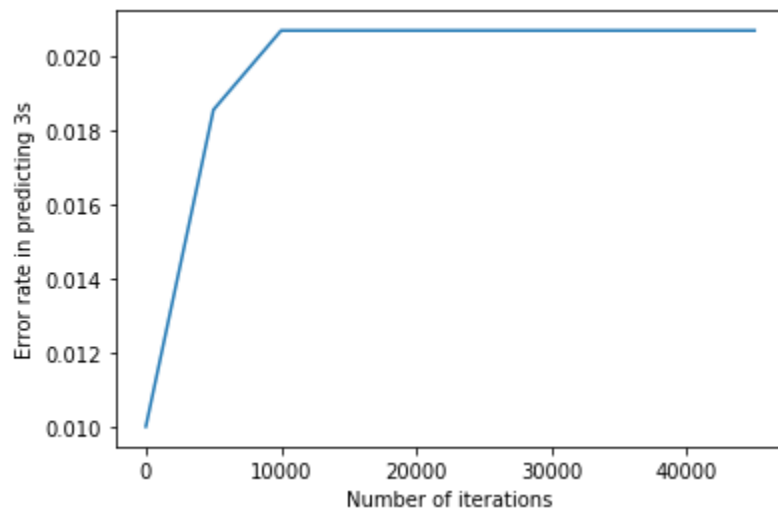


Error plots:

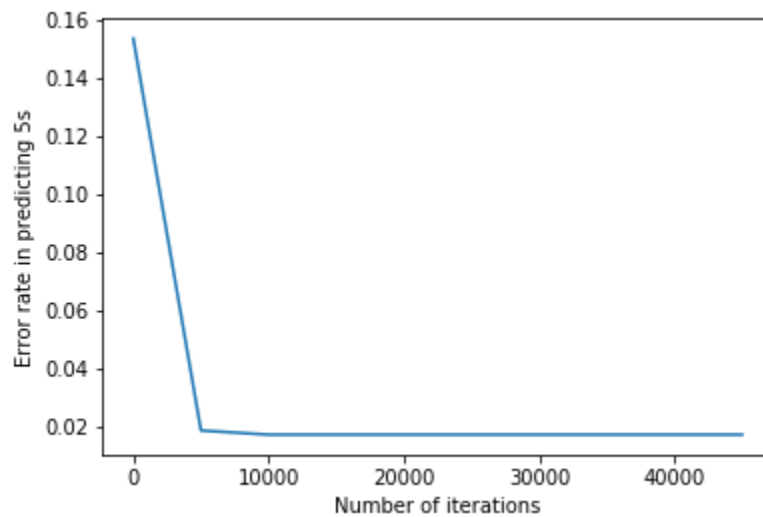
Error on the training set for the entire data



Error on the training set for classifying 3s



Error on the training set for classifying 5s



The weight vector is as follows:

[-0.70142002 -1.78974526 -1.09517625 -1.56056707 -0.61186491 -1.19527036 0.80513707 1.97996181 -
0.30630571 -0.27336541 0.33760441 -0.03573582 -0.7001968 1.00595075 -1.49952549 -1.51410911
4.53457038 1.39801323 1.62944226 0.09589359 1.03627812 -2.47825194 -2.46698309 -2.94432925
0.7533847 0.36389173 0.79236113 -0.36526679 -0.53075228 -2.81149831 0.53303721 -0.06521238
0.66903417 1.33370927 0.11273271 -0.48363481 -0.63321262 -0.02991974 -0.67736535 -0.06153866
1.34318369 -0.30175553 -0.45830022 -0.22661753 -0.05369975 -1.16850445 1.03719343 -1.89507721
1.75919754 -0.77950427 1.42494901 0.7411996 0.54165717 -0.47592851 0.12202656 -1.76627149
0.74513701 0.35908298 0.7887806 2.71257091 0.42891639 0.75510338 0.99147507 -0.63298529]

Error rate on training set: 0.03785714285714286

Error rate for predicting 3 on training set: 0.020714285714285713

Error rate for predicting 5 on training set: 0.017142857142857144

Part b)

Error rate on test set: 0.06625

Error rate for predicting 3 on test set: 0.0375

Error rate for predicting 5 on test set: 0.02875