Machine Learning Homework 6

Michael Panicci and Kyle Richards

April 25 2019

Question 1: 3-Layer Neural Network

Part A

The attached code implements stochastic gradient descent for a neural network with three layers.

Part B

The `check_grad` function was used to check if our gradient was calculated correctly. It produced a small number below 0.1, specifically, `1.526e-06`.

Part C

Table 1: Hyperparameter Tuning

Iteration	Hidden Neurons	Learning Rate	Batch Size	Cross Entropy	Percent Correct
9	50.0	0.25	100.0	0.0944	0.9732
6	50.0	0.10	50.0	0.0955	0.9716
3	40.0	0.25	100.0	0.0960	0.9740
0	40.0	0.10	50.0	0.0998	0.9702
1	40.0	0.10	100.0	0.1003	0.9716
11	50.0	0.50	100.0	0.1005	0.9756
7	50.0	0.10	100.0	0.1023	0.9696
8	50.0	0.25	50.0	0.1091	0.9738
2	40.0	0.25	50.0	0.1198	0.9704
5	40.0	0.50	100.0	0.1326	0.9690
10	50.0	0.50	50.0	0.1403	0.9704
4	40.0	0.50	50.0	0.1460	0.9714

This table shows the hyperparameter tuning accuracies. The cross entropy and percent correct values were calculated on the validation set. The table is sorted by cross entropy, where the lowest value is at the top. Therefore, our selected parameters are `Hidden Neruons = 50`, `Learning Rate = 0.25`, `Batch Size = 100`.

Part D

Table 2 shows the last 20 epochs when training the neural network with the optimal parameters. Running on the test set gives a cross entropy of 0.1103 and a percent correct of 0.9731.

Table 2: Last 20 Epochs

Epoch	Cross Entropy	Percent Correct
20	0.0359	0.988309
21	0.0339	0.988909
22	0.0317	0.989782
23	0.0300	0.990127
24	0.0283	0.990636
25	0.0273	0.990982
26	0.0259	0.991455
27	0.0242	0.992236
28	0.0227	0.993055
29	0.0212	0.993418
30	0.0202	0.993727
31	0.0187	0.994527
32	0.0179	0.994745
33	0.0168	0.995036
34	0.0160	0.995400
35	0.0148	0.995964
36	0.0142	0.996000
37	0.0135	0.996291
38	0.0127	0.996636
39	0.0123	0.996727

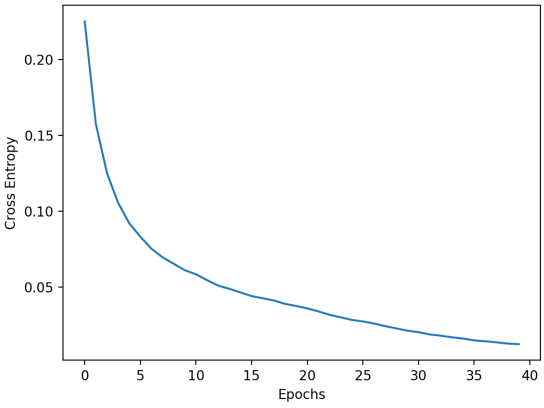

Figure 1: Gradient Descent

Question 2: Mountains and Valleys

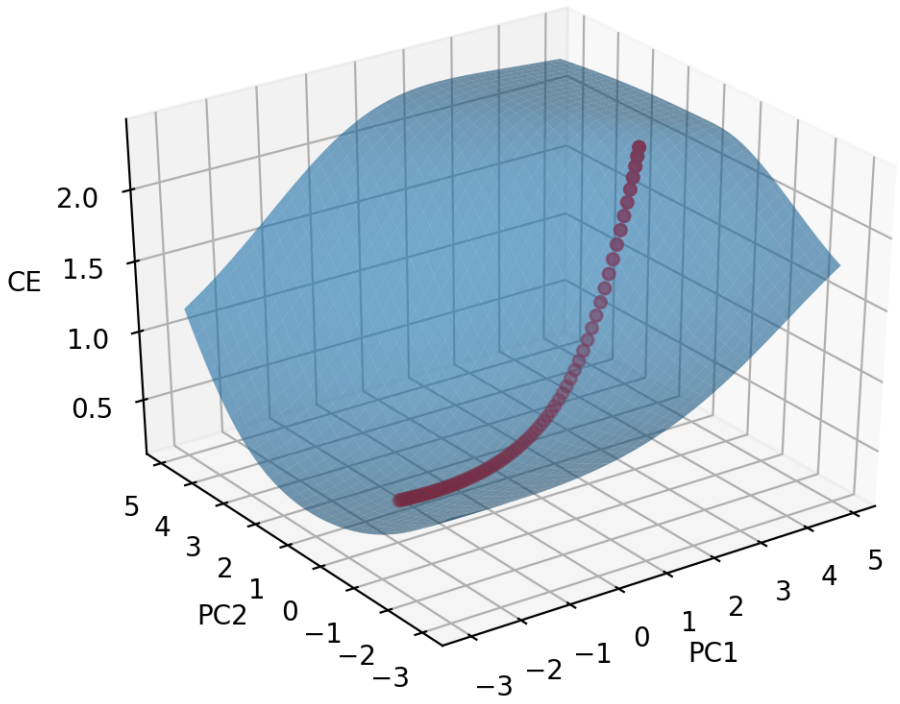

Figure 2: SGD Visualization 1

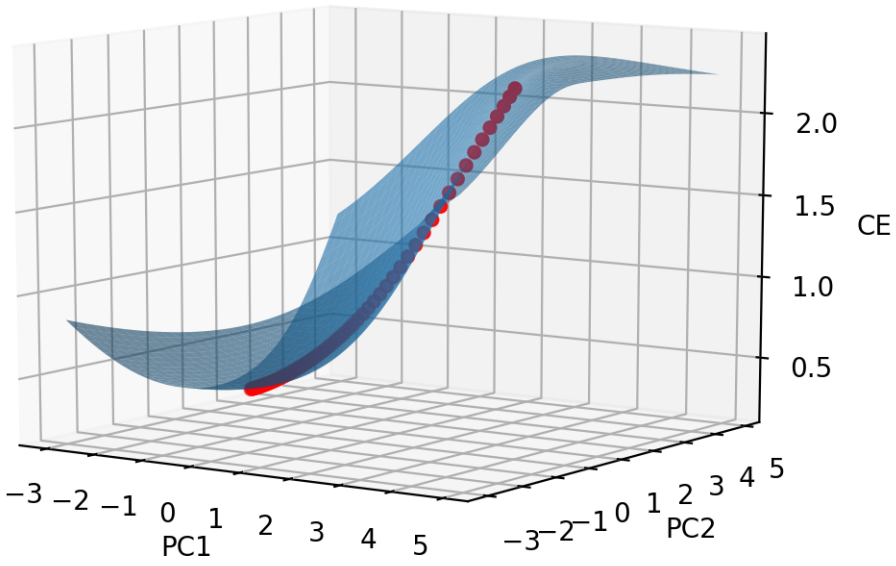

Figure 3: SGD Visualization 2

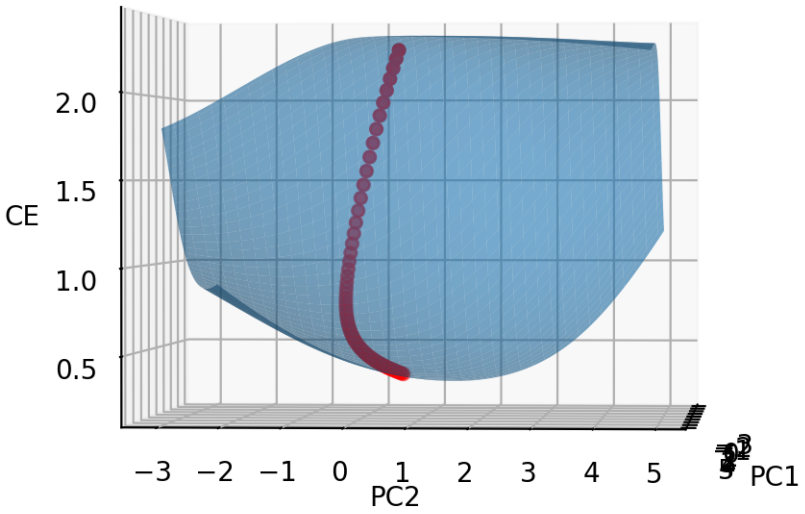

Figure 4: SGD Visualization 3