Parker Hague

Professor Aakur

Assignment 2

09/25/21

Question 1:

*See included PDF for handwritten derivations. Hague\_derivations.pdf*

Question 2:

1. When training the model, I was able to get the cost down to **< 0.0001**. After testing the model with the Y\_test data I was able to get the cost down to **79.32899**.
2. The line in this graph converges towards 0 and I have it stopping when cost **< 0.0001**.

Graphical user interface

Description automatically generated

1. I tested various learning rates and here are my findings. The learning rate affects the rate in which it converges and also affects the cost. Usually decreasing the learning rate made my cost go down a little, but that’s not always going to be the case.
   1. LR: .0001 🡪 Cost: 79
   2. LR: .001 🡪 Cost: 184
   3. LR: .01 🡪 Cost: nan
   4. LR: .1 🡪 Cost: nan
   5. LR: 1 🡪 Cost nan

As you can see, I wasn’t able to drop the cost below .001 without getting nan results. Here is a little plot of the two values that I was actually able to compute.

Chart

Description automatically generated

1. The number of neurons in the hidden layer has a direct impact on the cost. For the most part, as I increase the number of neurons, the cost decreases. I was able to get my lowest cost of 79.32899 on the training data when I used about 50 neurons in the hidden layer. The model seems to take longer to train with the increase in number of neurons. Here’s a plot of the cost vs the number of neurons (only plotting 1 – 10).

Chart, line chart

Description automatically generated

1. The activation functions also play a huge role in the network because they can completely transform your data to be very different. These functions dictate how “active” a particular neuron is. For this part, I explored two different activation functions in the hidden layer (while still using identity function for the output layer). The two functions that I tested were the Tanh and Identity activation functions.

Model with Tanh in hidden layer cost: 197.8

Model with Identity in hidden layer cost: 50.6

*See included .ipynb for the implementations.*

I put each separate model in it’s own code block for organization.