Machine Learning Diary

25/8

changed gradient descent algorithm. Used to do batch gd with the batch specified on initialisation of network class but now have gone full sgd.

Initial algorithm had a quadruple nested for loop but changing to sgd along with some other changes has reduced it to a single nested one.

Tested the algorithm – 27% accuracy after taking about an hour to train. Not good at all

Changed cost function from quadratic to cross-entropy

Resulted in 52% accuracy

26/8

There was an issue with the way I implemented the cost function

Fixed it and plotted a graph of cost over time to make sure it was converging

A graph with blue lines

Description automatically generated

It should be working – y label is incorrect, it is the general cost function of the network, not of a single neuron. The mis labelling was due to my misunderstanding of how the function worked which is probably why I was getting rubbish results

Retraining with these parameters:

Batch size = 10

Learning rate = 0.01

Number of hidden neurons = 30

Number of hidden layers = 1

A graph with numbers and lines

Description automatically generated

After training w 10 000 imgs – accuracy of 8% - possible overfitting?

Overfitting isn’t too much of an issue with this dataset

Changed cost function again – another misunderstanding of how it works. I didn’t realise for multiple output neurons, you need to calculate the cost for each neuron and then sum them.

Learnt the theory behind how neural networks work, gradient descent algorithm, backpropagation, different loss functions (cross entropy, MSE),

**Pytorch**

w/o regularisation

training acc = 98.1%

testing acc = 95.9%

A graph with blue and orange lines

Description automatically generated

w/ l2 regularisation

training acc

testing acc = 96.22 lambda = 0.8