Deep Learning, DD2424

Assignment 1

William Obeng

April 20, 2021

1 objectives and scope of the assignment

- Train and test a one layer network with multiple outputs
- Train the network with mini batch gradient descent applied to a cost function
- The cost function computes the cross entropy loss of the classifier applied to labelled training data and L_2 regularisation

2 Remarks

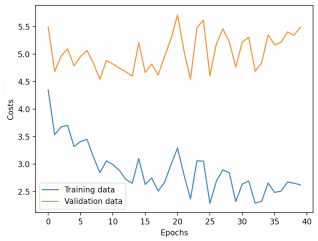
- All implementations where done in python.
- To make sure my gradient were correct I created a simple function to check the relative errors between the analytical gradient and numerical gradient.
- I also used testing library to compare my both gradients.
- I was able correctly create the gradient analytically. A good chunk of them where less than 1e-4 and about 80% where correct up to 4 decimal places.
- The accuracy of the data generally improves as the learning rate is reduced. With a smaller learning rate we are less likely to miss the local minimum.
- The training of the model is unstable if the learning rate is too high.
- When the regularization was added, we were able to have a really good fit.

2.1 Plots & Figures

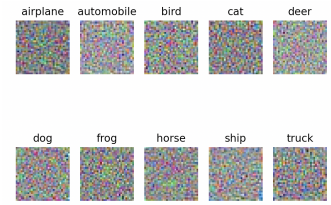
- Data was shuffled at the beginning of each epoch

2.1.1 lambda=0, n epochs=40, n batch=100, eta=.1

- Accuracy for training data was 46.68%
- Accuracy for validation data was 10.41%
- Accuracy for test data was 9.61%



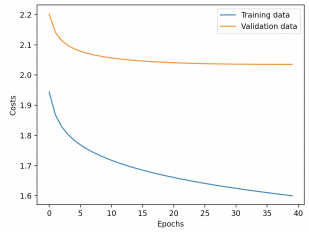
Figur 1: Graph of training and validation loss for each epoch



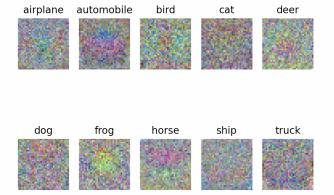
Figur 2: Learnt Weight Matrix

2.1.2 lambda=0, n epochs=40, n batch=100, eta=.001

- Accuracy for training data was 46.32%
- Accuracy for validation data was 10.01%
- Accuracy for test data was 10.08%



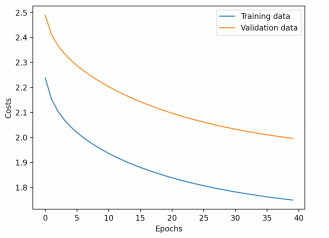
Figur 3: Graph of training and validation loss for each epoch



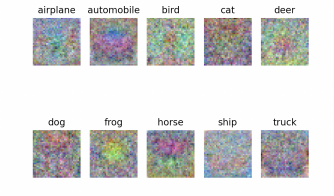
Figur 4: Learnt Weight Matrix

2.1.3 lambda=0.1, n epochs=40, n batch=100, eta=.001

- Accuracy for training data was 45.46%
- Accuracy for validation data was 9.81%
- Accuracy for validation data was 10.03%



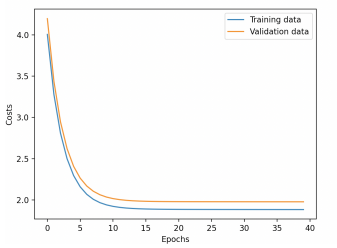
Figur 5: Graph of training and validation loss for each epoch



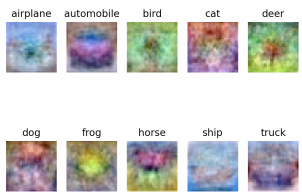
Figur 6: Learnt Weight Matrix

2.1.4 lambda=1, n epochs=40, n batch=100, eta=.001

- Accuracy for training data was 40.22%
- Accuracy for validation data was 10.01%
- Accuracy for test data was 9.99%



Figur 7: Graph of training and validation loss for each epoch



Figur 8: Learnt Weight Matrix