Assignment 2 Bonus

Hanlin Sun

April 18, 2018

1 Part 1 :Optimize the performance of the network

We mainly tried method (a), (b), (c) and (d).

1.1 Combination of (a) and (b)

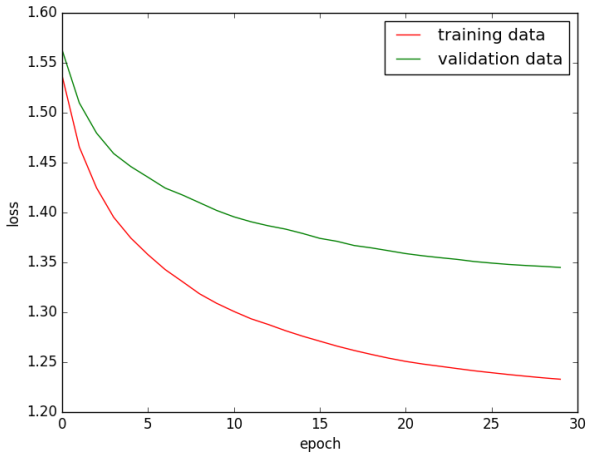

Figure 1: Loss function (without regularization)

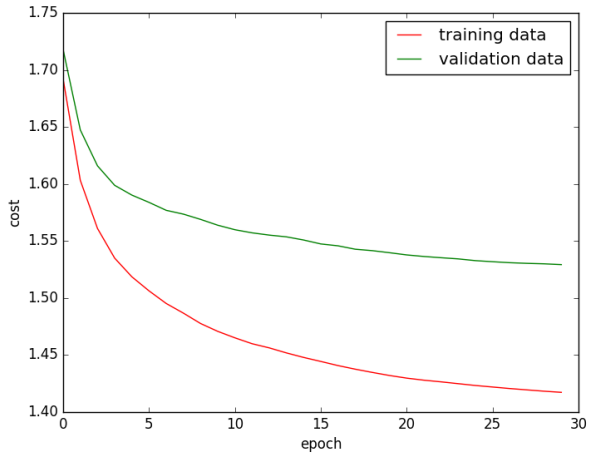

Figure 2: Cost function (with regularization)

We use all available data and He initialization. Using the hyperparameter given by former training, lambda = 0.004 and learning rate = 0.028, we achieve a higher accuracy 0.5204 after training 30 steps.

1.2 Combination of (a), (b) and (c)

We did a more exhausting searching of the optimal hyperparameters for this network. The first search used the range we fixed in Assignment 1 and we found the best hyperparameters as lambda = 0.0014, eta = 0.017; lambda = 0.0013, eta = 0.015; lambda = 0.0015, eta = 0.021. For these 3 sets of hyperparameters, the accuracies on test data are all over 0.51. We searched again, using lambda from 0.001 to 0.0015, eta from 0.015 to 0.025. Finally we picked the hyperparameters as: lambda = 0.0013 and eta = 0.025 and we get a higher accuracy on test data as 0.5254.

1.3 Combination of (a), (b), (c) and (d)

Finally we combined everything we had done to optimize the network. We used the network which has 300 hidden units and we turned lambda a little larger as 0.0015. Finally we got the best accuracy we had ever achieved as 55.23.

2 Part 2 :Train network using a different activation from ReLu

We use Sigmoid function to replace ReLU function and we use the hyperparameters which have been optimized. The accuracy on validation data after training 20 steps is 0.3821.

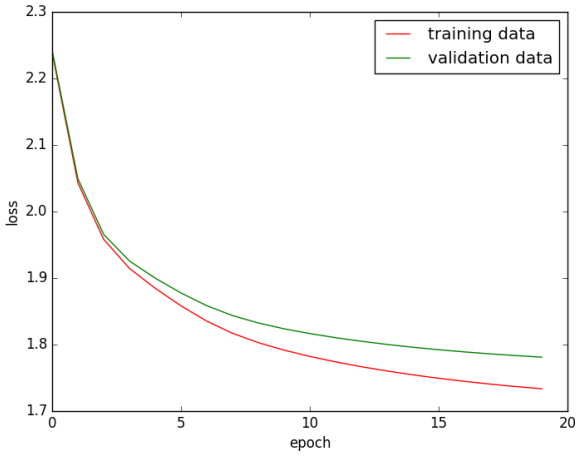

Figure 3: Loss function (without regularization)

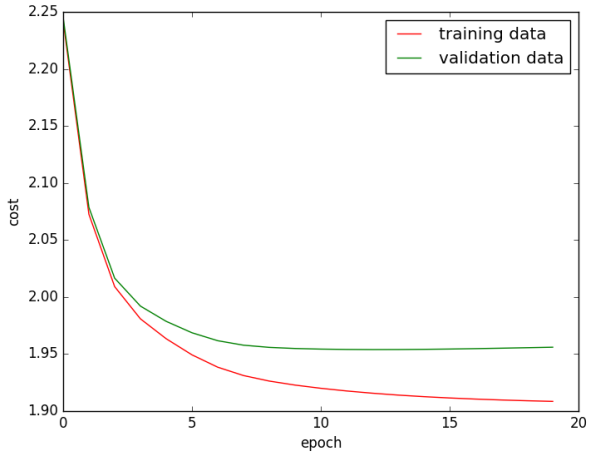

Figure 4: Cost function (with regularization)

On this network, we also tried to search best hyperparameters for this network. For coarse search, we set lambda from 0 to 0.005 and eta from 0.01 to 0.03. The best hyperparameters here is lambda =

0.0018, eta = 0.029; lambda = 0.0012, eta = 0.026; lambda = 0.0028, eta = 0.029. Therefore in next searching, we use lambda ranges from 0 to 0.002 and eta ranges from 0.025 to 0.05. Finally we got the best hyperparameters as: lambda = 0.0005 and eta = 0.045 and the accuracy on validation set is 0.4195.