Wen Sheng
Taylor Arnold
March 7, 2016

Problem Set 05 Data Analysis

Data Mining and Machine Learning

The highest validation accuracy rates I get for MNIST is 0.975 and for CIFAR is 0.4829. The parameters that affects the model accuracy are: hidden layer node numbers, mini_bach_size, epochs number, lambda number, cost function method and eta number.

The best combination of parameters for MNIST is:

[80, 10]

cost=network.QuadraticCost

epochs=40

mini_batch_size=5

eta=0.2

lmbda=0

The best combination of parameters for CIFAR is:

hidden layer: [15, 10]

cost: network.CrossEntropyCost

epochs=20

mini_batch_size=5

eta=0.01

lmbda=0

The best cost function for the model depends on the dataset. The best cost function for MNIST is QuadraticCost function and the best cost function for CIFAR is CrossEntropyCost.

When the first layer node number increases, the model accuracy increases; while the second layer node number dons't have such noticeable influence on the model accuracy as the first layer.

Halving the learning rate will make learning speed for each epic is lowered, while doubling the learning rate will make learning quicker with the risk of skipping the local maximum point. In these 2 model training and enough epoch numbers, halving the learning

rate will make the model advancing to the local maximum point and improves the model accuracy.

The regularization value means how much the cost function value (or weight value) are considered in the false scoring function. If regularization value is 0, the false scoring only depends on the cost function result. Otherwise, the scoring function will take the weight into consideration.

Compared with MNIST datasets, CIFAR data(with an accuracy of 0.48) is more difficult to predict as **common** features existed in every two output categories and with **complex contexts**.