CS 3891/5891

Spring 2018

Assignment 6

Basic SGD Algorithms

The MNIST database contains the dataset for this problem. Let *ld* be the last digit of your student ID. Develop an image recognition algorithm using logistic regression based on gradient descent that can correctly classify handwritten images as *ld* or non-*ld*.

1. *Preprocessing*
    *   Typically, pictures are reshaped to column vectors. The images in the dataset are of size (28,28) and should be reshaped to column vectors of size (28x28,1).
    *   For scaling of picture datasets, it is simpler and more convenient and works almost as well to just divide every row of the dataset by 255 (the maximum value of a pixel channel).
    *   For this exercise, you need to change the labels of the training and testing data sets. A label of an image should be 1 if the image shows *ld* and 0 otherwise.

2. *Minibatch SGD*
    *   Train a 3-layer neural network to classify the data using minibatch SGD.
    *   You can use 2 hidden layers with 20 and 10 units respectively
    *   Use ReLU in the hidden layers and sigmoid in the output layer.
    *   Plot the learning curve (cost function vs. number of iterations).
    *   Compute the training and test error.
    *   Investigate the impact of the minibatch size.

3. *Minibatch SGD with Momentum*
    *   Train a 3-layer neural network to classify the data using minibatch SGD with momentum.
    *   You can use 2 hidden layers with 20 and 10 units respectively
    *   Use ReLU in the hidden layers and sigmoid in the output layer.
    *   Plot the learning curve (cost function vs. number of iterations).
    *   Compute the training and test error.
    *   Investigate the impact of the momentum parameter.

Submission
    *   A report describing your solution
    *   Your software
    *   Jupyter notebood (optional)