CS 3891/5891

Spring 2018

Assignment 3

Two-Layer Neural Network

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*.

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.
- The training data set contains 60,000 examples. Create smaller training datasets for development and testing. Investigate the impact of the training data size.

Learning

- Use vectorized implementation.
- You can use 5-10 units in the hidden layer with tanh.
- Use sigmoid in the output layer.
- Plot the learning curve (cost function vs. number of iterations).
- Investigate the impact of the learning rate, number of hidden units, and activation function of the hidden layer.

Evaluation

- Compute the training and the test error.
- Try test examples.

Submission

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