CAS CS 440/660 Artificial Intelligence - Spring 2019

This assignment is due on Wednesday, March 6, at 00:00 EST (i.e. the night from Wednesday to Thursday). You are encouraged to work in teams of two students. Each student must submit code along with the report.

Programming Assignment 1

In this assignment, you need to build a neural network by yourself and implement the backpropagation algorithm to train it.

Each team must write a webpage report on the assignment. For example, http://www.cs.bu.edu/faculty/betke/cs440/restricted/p1/p1-template.html. To ensure that each team member builds his or her own electronic portfolio, we ask that everybody submits his or her own report.

Learning Objectives

1. Understand how neural networks work
2. Implement a simple neural network
3. Apply backpropagation algorithm to train a neural network
4. Understand the role of different parameters of a neural network, such as epoch and learning rate.

You need to do

1. We provide code that you may use to implement a neural network with one hidden layer. (10pts)
2. Train your neural network with the provided linear and non-linear dataset (DATA/data_linear, DATA/data_nonLinear) respectively and evaluate your trained model by a **5-fold round robin cross-validation**, i.e. separate the whole dataset into 5 parts, pick one of them as your test set, and the rest as your training set. Repeat this procedure 5 times, but each time with a different test set. To evaluate your learning system, you will need to calculate a confusion matrix. The cross validation procedure enables you to combine the results of five experiments. Why is this useful? (15pts)
3. What effect does the learning rate have on how your neural network is trained? Illustrate your answer by training your model using different learning rates. Use a script to generate output statistics and visualize them. (5pts)
4. What is overfitting and why does it occur in practice? Name and briefly explain 3 ways to reduce overfitting. (5pts)
5. One common technique used to reduce overfitting is L2 regularization. How does L2 regularization prevent overfitting? Implement L2 regularization. How differently does your model perform before and after implementing L2 regularization?(5pts)

6. **Optional for CS440, Required for CS640:** Now, let's try to solve real world problem. You are given hand-written digits as below, all digits are stored in a csv file. You need to implement the neural network class with 1 hidden layer to recognize the hand-written digits, you should train your model on DATA/Digit_X_train.csv and DATA/Digit_y_train.csv, then test your model on DATA/Digit_X_test.csv and DATA/Digit_y_test.csv. Provide your results and a discussion of the performance of your AI system. (10pts)

7. Instructions:(1) For **CS440**, total points: **40 + 10 extra credits**. For **CS640**, total points: **50**
   (2) Your report should include the results and your analysis for part 2-6. In lab section (demo), we will ask you to run and explain your code.
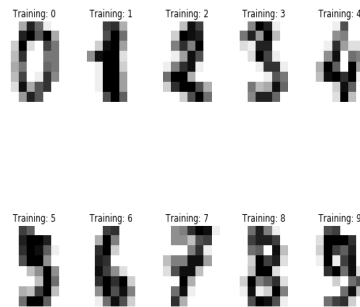
Submission The programming assignment, along with the webpage report, should be made into a **CS440_P1_FirstName_LastName.zip** file and submitted using gsubmit (we will post a link about how to use gsubmit soon).

Margrit Betke, Professor
Computer Science Department
Email: betke@cs.bu.edu