Assignment 3

Bilal Baig (215734320)
bilalb@yorku.ca

March 20, 2021

Note: This assignment is mainly for you to review several important discriminative models. You have to work individually. **You must use the same mathematical notations in textbook or lecture slides to answer these questions.** You must use this latex template to write up your solutions. Remember to fill in your information (name, student number, email) at above. No handwriting is accepted. Direct your queries to Hui Jiang (hj@eecs.yorku.ca)

Exercise 1

(30 marks) **Fully-Connected Neural Networks**

(a) (5 marks) Q8.2 on page 199 (see the margin note on page 175 for some examples.)

(b) (25 marks) You will use the MNIST data set for this question. Implement the forward and backward passes for fully-connected deep neural networks as in Figure 8.19. Use all MNIST training data to learn a 10-class classifier using your own back-propagation implementation, investigate various network structures (such as different number of layers and nodes per layer), and report the best possible classification performance in the held-out MNIST test images. Note that you are only allowed to use libraries for linear algebra operations, such as matrix multiplication, matrix inversion, and etc. You are not allowed to use any existing machine learning or statistics toolkits or libraries or any open-source code for this question.

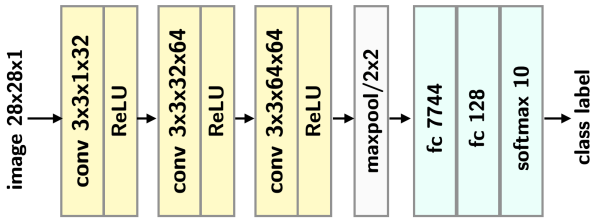
Your answers:

1. This would use the sigmoid function instead
   $L = \sum_x (\sum_{i=1}^{N} r_i y_i - \sum_{i=1}^{N} (1 - r_i)(1 - y_i))_x$
   Where r = a vector correctly labelled with 0 or more 1s for each samplex x = class and y=output
2. I found that 2 hidden layers and 1 output layer with 10 nodes generally worked the best as it was relatively accurate while reducing the time required for each epoch by a great deal and we need 10 nodes to be able to predict 1 of the 10 classes. After 2-3 epochs the neural network had a great result and so it would be best to leave it at that much.

Exercise 2

(30 marks) **Convolutional Neural Networks**
Note that 4404 students are required to do part (b) only (30 marks) while 5327 students need to do both parts (a) and (b) (10 marks + 20 marks).

(a) Q8.6 on page 200



(b) Implement the forward and backward passes for the following convolutional neural network:
Use all MNIST training data to learn a 10-class classifier and report the best possible classification performance in the held-out MNIST test images. Note that you are allowed to use any machine learning or statistics toolkits or libraries for this question. Do some investigations to ensure you use a suitable toolkit for this question.

Your answers:

1. From my testing, a batch size of 128 worked very well while only a few epochs were needed. 1 epoch gave an accuracy at 98.4%, a 2nd epoch pushed it to an accuracy of 98.8% and with the runtime of each epoch there is diminishing returns with the slight change in accuracy.

Exercise 3

(20 marks) **Ensemble Learning**

(a) (10 marks) Adaboost: Q9.1 on page 215

(b) (10 marks) Gradient Tree Boosting: Q9.5 on page 215

Your answers:

What to submit?

You must submit:

1. one PDF document (using this latex template) for your solutions to all written questions and all results and discussions for your programming assignments;
2. one zip file that includes all of your Python codes (e.g., *.ipynb if you use Jupyter notebooks) and a readme file for TA to run your codes;

from eClass before the deadline. No late submission will be accepted.