ComS 573 Machine Learning
Lab 2 (100 points)
Neural Network Classifier

Note:

- For help with laboratory assignments, please contact TA.
- You may perform the experiments using a machine learning package, such as Keras (with TensorFlow).

In this assignment, you will experiment with the Neural Network classifier.

1 Dataset

We will use the following data set from the UC Irvine Machine Learning Repository:

- Optical Recognition of Handwritten Digits Data Set (use optdigits.names, optdigits.tra as training data, and optdigits.tes as test data)
(`https://archive.ics.uci.edu/ml/datasets/Optical+Recognition+of+Handwritten+Digits`).

2 Tasks

You will experiment with the Neural Network classifier. You will use the softmax function for the output layer and use 1-of-c output encoding with target values such as (1, 0, 0, ...). Use the early stopping technique to decide when to stop training. For example, you may use 20% of training data in optdigits.tra as the validation set.

1. Experiment with fully-connected feed-forward neural networks.

   (a) **Sum-of-squares error (aka mean-squared-error) vs. cross-entropy error function**. Use the ReLU units for the hidden layers. For each of the two types of error functions, experiment with different values of hyper-parameters, including number of hidden layers, number of hidden units in each layer, learning rates, momentum rates, input scaling, and so on. Compare their classification accuracy and convergence speed (the number of iterations or actual time till training stops). Report your experimental results, and the best hyper-parameter values you find. Report for the best model you learned, the corresponding hyper-parameters and the performance including overall classification accuracy, class accuracy[1], and confusion matrix for both training and testing data. Discuss the results.

[1] class accuracy = (predicted as class i)/(number of examples in class i)

(b) **tanh vs. ReLU hidden units**. Use the cross-entropy error function. For each of the two types of hidden units, repeat the above experiments, that is, experiment with different values of hyper-parameters and report the results. Discuss the results.

2. Experiment with convolutional networks (CNNs). Use the cross-entropy error function, and ReLU hidden units. Repeat previous experiments, that is, experiment with different values of hyper-parameters (note CNNs may have different types of hyper-parameters, eg. filter size) and report the results. Discuss the results.

3 What to turn in

Turn in via Canvas (a compressed .zip file if necessary) the following:

- A lab report (in pdf file) with your experimental results and discussions of these results. It should include a brief description of all the design choices made. Try to use tables or plots to summarize your results. You should specify the parameters of every experiment in such a way that they can be replicated by the TA.
- Readme file with instructions on how to reproduce your experiments.
- Any source code that you may have written.