ComS 573: Machine Learning

Laboratory Assignment 2 (100 points)

For help with laboratory assignments, please contact TAs.

In this assignment, you will experiment with the Neural Network classifier. You may use a machine learning tool (we recommend R because it is more flexible) to perform the experiments. You may use machine learning packages such as h2o, etc..

Data

- Optical Recognition of Handwritten Digits (use optdigits.names, optdigits.tra as training data, and optdigits.tes as test data) from the UC Irvine Machine Learning Repository.

You will use fully-connected feed-forward neural networks. 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.

Tasks

1. **Sum-of-squares error vs. cross-entropy error function**. Use the logistical sigmoid units for the hidden layers. For each of the two types of error functions, experiment with different values of hyper-parameters, such as number of hidden layers, number of hidden units in each layer, learning rates, momentum rates, and input scaling etc.. Compare their convergence speed and classification accuracy performance. 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, and confusion matrix for both training and testing data. Discuss the results.
2. **Logistical sigmoid vs. tanh vs. ReLU hidden units**. Use the cross-entropy error function. For each type of hidden units, repeat the above experiments, that is, experiment with different values of hyper-parameters and report the results. Discuss the results.

What to Turn In

Turn in via Blackboard (a compressed file if necessary) the following:

- A report of the 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.
- Any source code that you may have written.