CS 760 Homework 3: Neural Network

Qihong Lu

November 13, 2016

QUESTION1

To run the program, type the following command:

./nnet l h e <train-set-file> <test-set-file>

where l specifies the learning rate, h the number of hidden units and e the number of training epochs.

Code:

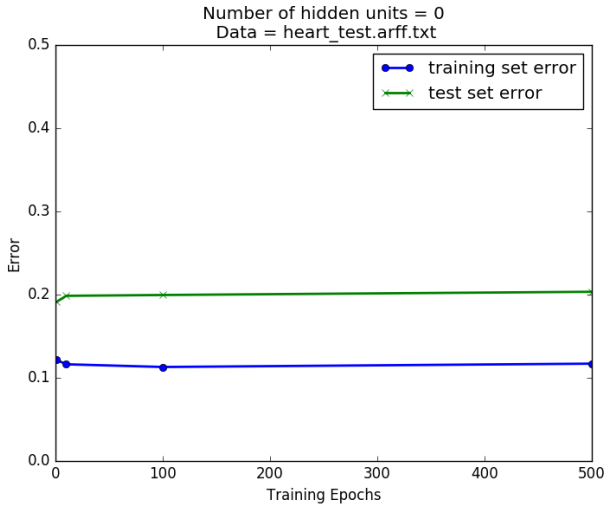
- nn_alg.py: implements binary classification network with zero or one hidden layer
- util.py: the definitions of some constants and helper functions
- nnet.py: run neural network, return test results

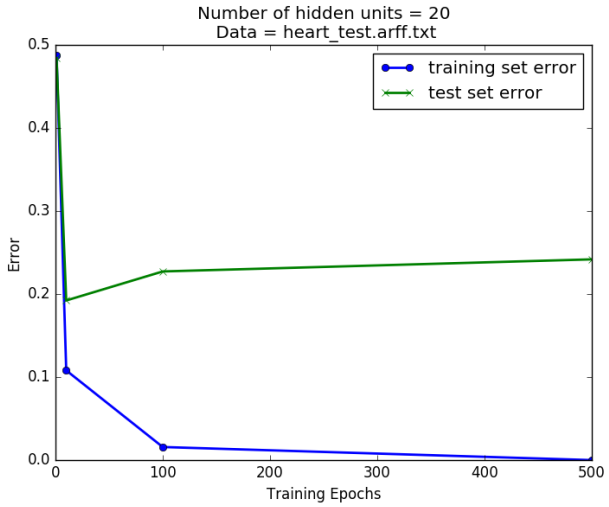
Dependencies:

- pyhton 2.7
- numpy
- scipy
- sys

QUESTION2

Using heart_train.arff and heart_test.arff, you should make two graphs showing error-rates versus the number of training epochs. For the first graph, plot training and testing error rates for a single-layer network trained for 1, 10, 100 and 500 epochs, using a learning rate of 0.1. For the second graph, plot similar curves for a network with 20 hidden units.



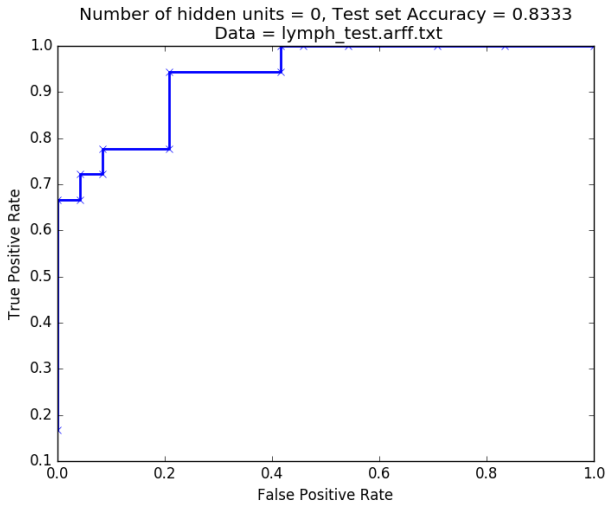


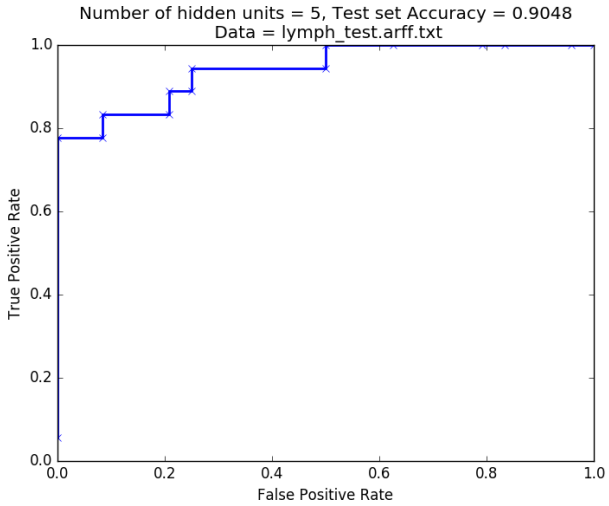
The learning curves above are the average results of 20 runs. Although the model with hidden units (the bottom figure) did better at reducing the training error, but its test error is higher, compared to the simpler perceptron (without hidden layer). This suggests that the model with 20 hidden units is overfitting the training set. This issue of overfitting (the bottom figure) can also be seen from the increasing trend of the test error over the training epochs.

QUESTION3

For this part, you should produce ROC curves for two data sets. Use the activation of the output unit as the measure of confidence that a given test instance is positive, and plot ROC curves for both the heart data set indicated above, and the lymphography data set lymph_train.arff, lymph_test.arff. Be sure to label the axes of your plots.

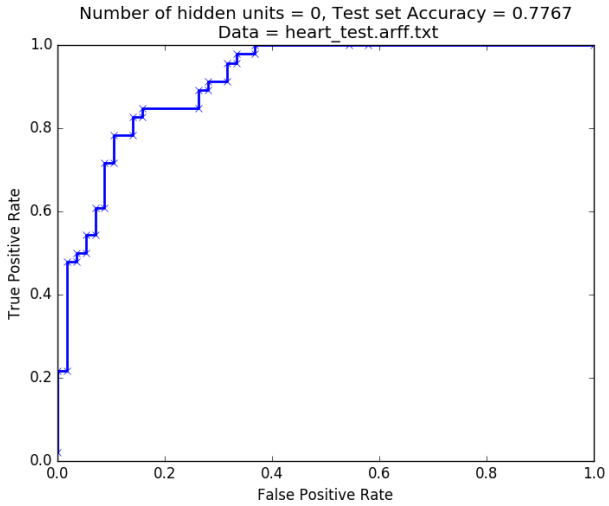
1. Here're the results for the **lymph** data set:

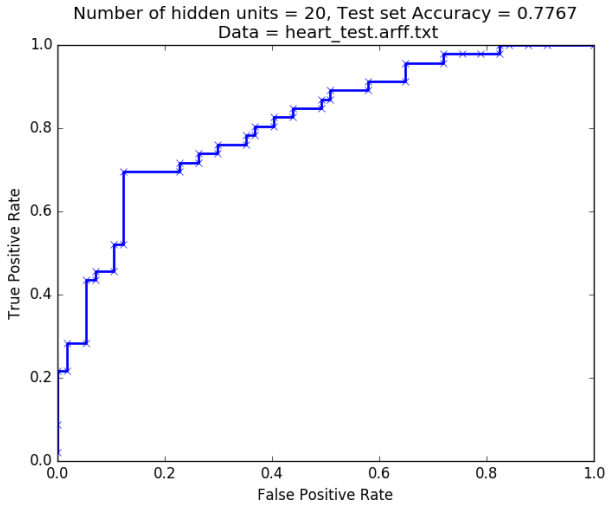




In this simulation, both models were trained for 500 epochs with learning rate of .1. The model with the hidden layer has five hidden units. It turned out that the model with five hidden units achieved a better accuracy and a slightly larger area under ROC curve, suggesting the simple perceptron is slightly underfitting the data.

2. Here're the results for the **heart** data set:





In this simulation, both models were trained for 500 epochs with learning rate of .1. The multilayered-model has 20 hidden units. The results show that the simpler model no hidden units had a significantly larger area under the ROC curve, indicating the model with 20 hidden units is overfitting the training set.