Elie Weintraub Artificial Intelligence - Project 2: Neural Network 12/6/13

**1. Overview**

For this project two programs, train.cpp and test.cpp were written in order to implement the training and testing, respectively, of a neural network. The training of the network employed back propagation. For simplicity a single hidden layer was used. As an additional component of this project a dataset was found, preprocessed, and used to train and test the neural network.

**2. Data Set**

**2.A. Origin**

The data set I used was derived from a data set available in MATLAB called wine\_dataset. It originates from the UCI Machine Learning Repository and is available at <http://archive.ics.uci.edu/ml/datasets/Wine>.

**2.B. Description of Data Set**

The description of the data set given below is largely derived from the documentation in MATLAB available by entering the command doc wine\_dataset in the MATLAB command window together with the documentation provided on the site listed above.

The data set is the results of a chemical analysis of wines grown in the same region in Italy but derived from three different wineries. The analysis determined the quantities of 13 different attributes found in each of the three types of wines. These 13 different attributes thus constitute the input attributes.

The input attributes are listed below:

1) Alcohol

2) Malic acid

3) Ash

4) Alcalinity of ash

5) Magnesium

6) Total phenols

7) Flavanoids

8) Nonflavanoid phenols

9) Proanthocyanins

10)Color intensity

11)Hue

12)OD280/OD315 of diluted wines

13)Proline

The outputs then correspond to which winery the wine was derived from and is represented with a one-of -N scheme (i.e. one for the winery from which the wine was derived, and 0 for the others).

**2. B. Modification of the Original Data Set**

The original data set consisted of 178 examples. This was then randomly divided into a training set and testing set, each consisting of 89 examples. Additionally, simple normalization was done where the values of each input attribute were divided by that attribute’s maximum value. Finally, some reformatting was done to get the data into the form specified for the assignment.

**2. C. Description of Neural Network, Initial Weight Generation, Parameters, and Results**

It was decided that the single hidden layer would consist of 10 nodes. Thus the neural network consisted of 13 input nodes, 10 hidden nodes, and 3 output nodes.

The initial weights were pseudo-randomly generated from a uniform distribution with range [0,1] using the MATLAB command rand.

The learning rate used was 0.1. It was empirically determined that this learning rate leads to good results.

Various number of epochs were used. Specifically, the values 100,200,500, and 1000 were explored. It was found that using 500 epochs yielded slight improvement over 200 epochs, which yielded slight improvement over 100 epochs. However, when increasing the number of epochs to 1000 results did not improve and were identical to those that resulted from using 500 epochs. Both the micro- and macro-averaged results for each of the metrics when using 500 epochs were in the range 93.3%-96.3%. Further improvement might be possible if a different number of hidden nodes is used, but these results seemed sufficient and so varying the number of hidden nodes was not explored.