AIML ASSIGNMENT II

SANTHOSHINI REDDY - IS201401040
CHRIS ANDREW - IS201401015

Classifying the IRIS Data-Set;
The data set has three classes: **Iris-setosa, Iris-versicolor, Iris-virginica**
Classification is done on the basis of four features given as input:
[sepal length(cm)] , [sepal width(cm)] , [petal length(cm)] , [petal width(cm)]
The target array are defined for each of the classes as follows
Iris-setosa : [1 0 0]
Iris-versicolor : [0 1 0]
Iris-virginica : [0 0 1]

The data-set has **150** entries and is divided into two parts randomly;
One file of **100** entries for training the perceptron.
One file of **50** entries for testing the perceptron.
The data is in a file called "**data**".
The program **divider.c** divides the data into two file: "**train**" and "**test**"
Running **train.c** trains the perceptron and writes the weights to the file "**weight**".
Running **test.c** tests the data by taking weights from "**weight**" and prints the results.
We have run the train files for multilayer perceptrons with **3**, **4**, **5** nodes each.
The set has been run for **4, 10** and **50 EPOCHS** for each of them.
We have not run the perceptron for more than **50 EPOCHS** because of the large time needed to train.
Gradient descent is done on each entry till sum-squared error is less than **0.000001**. The step size is taken as **0.1**. The test file classifies entries correctly if sum-squared error is less than **0.1**.
The results obtained are summarised as follows:
For **4 EPOCHS**:
3 hidden-nodes: Correctly classifies - **32/50** test entries: **36% Error Rate**
4 hidden-nodes: Correctly classifies - **47/50** test entries: **6% Error Rate**
5 hidden-nodes: Correctly classifies - **47/50** test entries: **6% Error Rate**

For **10 EPOCHS**:
3 hidden-nodes: Correctly classifies - **47/50** test entries: **6% Error Rate**
4 hidden-nodes: Correctly classifies - **46/50** test entries: **8% Error Rate**
5 hidden-nodes: Correctly classifies - **46/50** test entries: **8% Error Rate**

For **50 EPOCHS**:
3 hidden-nodes: Correctly classifies - **47/50** test entries: **6% Error Rate**
4 hidden-nodes: Correctly classifies - **47/50** test entries: **6% Error Rate**
5 hidden-nodes: Correctly classifies - **47/50** test entries: **6% Error Rate**

We found that **50 EPOCHS** gives us the best possible result for this training-set, the details are:

For 3 hidden-nodes, misclassified results are:
Target: 0.000000 1.000000 0.000000 | **Output**:0.000002 0.012757 0.987237
Entry: 1.000000 5.900000 3.200000 4.800000 Iris-versicolor

_____
Target: 0.000000 1.000000 0.000000 | **Output**:0.000013 0.642843 0.359037
Entry: 1.000000 6.200000 2.200000 4.500000 Iris-versicolor

_____
Target: 0.000000 1.000000 0.000000 | **Output**:0.000001 0.004027 0.996003
Entry: 1.000000 6.000000 2.700000 5.100000 Iris-versicolor

_____
Entries Classified correctly:**47**

For 4 hidden-nodes, misclassified results are:
Target: 0.000000 1.000000 0.000000 | **Output**:0.000002 0.012757 0.987237
Entry: 1.000000 5.900000 3.200000 4.800000 Iris-versicolor

_____
Target: 0.000000 1.000000 0.000000 | **Output**:0.000013 0.642843 0.359037
Entry: 1.000000 6.200000 2.200000 4.500000 Iris-versicolor

_____
Target: 0.000000 1.000000 0.000000 | **Output**:0.000001 0.004027 0.996003
Entry: 1.000000 6.000000 2.700000 5.100000 Iris-versicolor

_____
Entries Classified correctly:**47**

For 5 hidden-nodes, misclassified results are:
Target: 0.000000 1.000000 0.000000 | **Output**:0.000002 0.012757 0.987237
Entry: 1.000000 5.900000 3.200000 4.800000 Iris-versicolor

_____
Target: 0.000000 1.000000 0.000000 | **Output**:0.000013 0.642843 0.359037
Entry: 1.000000 6.200000 2.200000 4.500000 Iris-versicolor

_____
Target: 0.000000 1.000000 0.000000 | **Output**:0.000001 0.004027 0.996003
Entry: 1.000000 6.000000 2.700000 5.100000 Iris-versicolor

_____
Entries Classified correctly:**47**

*We observe that the same entries are misclassified in each case.

*After generating new test and train files for the same data set, we observe that there is not much difference in the above results obtained.