Assignment 2

How deep is your net

Out: 09/21/2023. **Due**: before midterm break.

Consider the following training set (gzipped) on 1000 dimensions: training data set, training label set.

Design a Boolean classifier (outputs +1/-1) which is (approximately) accurate on the training set. This means that the total number of misclassifications that the classifier makes on the training set should be (near) zero. Use a classifier that is a neural network with one output and one layer of hidden units.

Then run your classifier on this testing set (gzipped): testing data set, testing label set. Report the number of misclassifications of your classifier on the testing set.

What to submit

Collect the Boolean predictions of your classifier on the testing set in a single text file (as a sequence of +1 and -1 separated by spaces in a single line). Also store your net as a text file using the following format. The first line should be an integer that represents the number of hidden units. The second line should be the sequence of weights used by the output units. The subsequent lines should be the sequence of weights used by each hidde unit on the inputs (ordered in the same manner as the data). So, if you use M hidden units, this file should ha M+2 lines. The last M lines should have 1000 entries in them. Submit via email the two text files along with your source code for the homework.

You may work in groups of two.

Reference: Cybenko's theorem.