Project 1: 2023 Prof. Manevitz

(Adaline, Back Propagation and Some Transfer Learning)

Note: There are 3 deadlines.

1. Part 1 has to be in by this Monday. No extensions
2. Part 2 deadline is May 9. Late penalty: 10 points until May 16, -20 points afterwards.
3. Part3 (Backpropagation) deadline is May 23, -10 points until May 30, -20 points afterwards.

Please submit on time!!

Remember there will be another project so much better to get this done quickly..

Collaborations

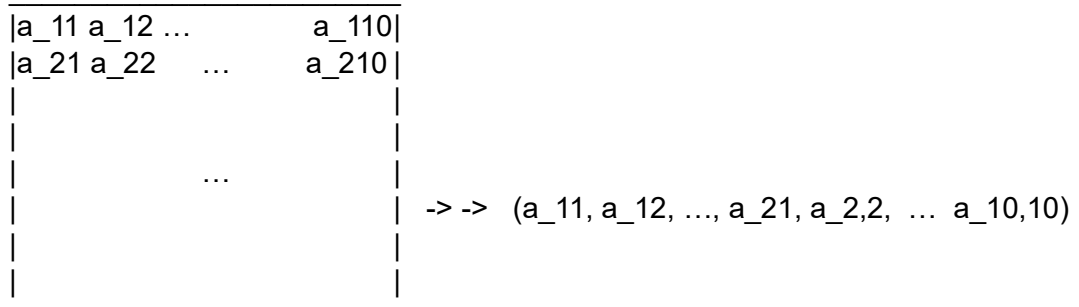
1. I prefer you to work in teams of two, but on this project I will allow three and even four. However all collaborators are obligated to know ALL details of all parts of submission.
2. You may ***NOT*** share code between groups,
   nor should you use code from other sources (except for part 3) but you may consult between groups freely on difficulties. I suggest you use the Forum to share questions and answers.

   Good luck and enjoy!

Part I: (DUE Monday April 24 by 23:59 midnight.) (10 points)

(Data Generation). Each student should submit 10 data points for “aleph” handwriting and 10 data points for “bet” handwriting and 10 data points for “mem”.

a) To do this you can simply write over a 10 x 10 matrix by hand and then generate the data as 10 x 10 matrices. Just prepare a table and write the letters (handwritten) and record them as +1 for an entry which has a part of the letter , and -1 for the others.



|a_101 a_102 … a_1010 |
|_____ |

If this is a matrix of your handwriting showing a bet insert an extension of the vector with a “1” in position 0.
(1,a_11, a_12, …, a_21, a_2,2, … a_10,10).

Similarly if it a lamed insert a “2”.
(2,a_11, a_12, …, a_21, a_2,2, … a_10,10).

If it was a “mem” insert a 2..

b) Now add another 10 letters by rotating the ones you traced out by 15degrees and by -15 degrees.

c) Each *individual* (not team) should submit their data to the package on the moodle by Monday2359, . (It must be done by Monday midnight as all the students will be using this database.)

There are about 100 students in the two classes, so each student should thereby submit 27 letters. This should give you about 3000 letters as data.

Part 2 Adaline classifications (Submissions by May 9 midnight. Late submission – 10 points until May 16 midnight). After May 23 -20 points.)

2. Classify “mem” versus “bet”. (This is done by a team of two or three.)

Build a classifier using the Adaline algorithm which classifies “mem” from “bet”. You should write the Adaline algorithm yourself.

You have to encode (yourself) a classifier that is trained on 80% of the data and tested on the remaining 20%. You should do this 5 times, so all the data is used for training and testing.

Report all the results. Include the number of interations the code had to run before converging.

Include in your full reports, your names and T.Z.s , the code you have written and the results for each choice of training and testing data, and your average of all the results.

3. Classify “lamed” versus “bet”. And report as above.

4. Classify “lamed” versus “mem”. And report as above

Part 3: (Due May 23 midnight. -10 points until May 30; -20 points after that.)

5. Now replace the Adaline with a feed forward neural network which you can either write yourself or use a package. Repeat the classifications from earlier parts and report the results. You should explain how you decided on number of neurons and levels.