Project 1: 2021 (Adaline, Back Propagation and Transfer Learning)

Deadlines: Part A and B are due by April 27. Part C and Part D are due by May 11.
Late penalties as below.

You should submit in pairs.

Collaboration: You may *NOT* share code between groups, but you may consult between groups freely on difficulties. I suggest you use the Forum to share questions and you may even schedule zoom meetings for discussions.

Part A and Part B should be submitted together before April 27. Late submissions will be accepted with a penalty of 5 points for one day to one-week lateness; -10 points for between one week and two weeks late (May 11), -20 points until May 20.

Part C and Part D should be submitted by May 11 with -5 for one week late (May 18), -10 for up to two weeks(May 25) and -20 for three weeks (i.e. by June 5).

Note: There will be an additional project two assigned later in the semester.

All data Is two dimensional, <x, y> where -1 <= x, y <= 1. The data is all data points <x, y> where x is of the form m/100 where m is an integer between -100 and +100 and y is of the form n/100 with n an integer between -100 and +100. Suppose that all data points with x > 1/2 and y >1/2 have the value 1; all other points have the value -1.

Now suppose you do not know this; but you are given a random sample of data of size 1000 together with its value (e.g., the point <60/100, 80/100> has value 1; while the point <20/100, 70/100> has the value -1.

(Part A and Part B are due by **April 27)**

Part A: Implement the Adaline learning algorithm and show how it generalizes to develop a decision that works on all the set. Does the accuracy of the result depend on the training set? Present tables and possibly a picture indicating your results. Suppose the data is of the size n/10,000 with n an integer between -10,000 and +10,000. How does this affect your choice of training data and testing data?

Part B: Now change the problem so that points such that <x, y> has value 1 only if
1/2 <= x**2 + y**2 <= 3/4. What are the best results you obtain using an Adaline? Does the quality of the results change if you use more data? Present tables and perhaps a figure. Part C and Part D are due by **May 11**

Part C: Now try the same with a back-propagation algorithm instead of the Adaline. You will have to define the architecture (i.e number of neurons and number of levels) you may either implement the algorithm or use a package. HOWEVER, YOU WILL NEED TO LOOK INSIDE the results of each neuron separately for Part D.
Show a geometric diagram in terms of the **inputs** of the training set for the *output* of *each neuron separately* in the neural network as well as for the output neuron. Present tables of results both for training and for testing.

Part D: Now use the trained neurons from the next to last level of Part 3 as input and only an Adaline for the output. (That is, you will give the Adaline the output of the neurons from Part 3 in the level below the output, and train only the Adaline). Describe how accurate the Adaline can be. Give diagrams. Draw whatever conclusions you think are appropriate from your results.