CMSC 409: Artificial Intelligence

Fall 2022, Instructor: Dr. Milos Manic, http://www.people.vcu.edu/~mmanic

Project 1

CMSC 409: Artificial Intelligence
Project No. 1
Due Tuesday, September 15, 2022, noon

Student certification:
Team member 1:
Print Name: _____ *Date:* _____
I have contributed by doing the following: _____
Signed: _____ *(you can sign/scan or use e-signature)*

Team member 2:
Print Name: _____ *Date:* _____
I have contributed by doing the following: _____
Signed: _____ *(you can sign/scan or use e-signature)*

Team member 3:
Print Name: _____ *Date:* _____
I have contributed by doing the following: _____
Signed: _____ *(you can sign/scan or use e-signature)*

Pr.1.1

A) Understand and explore a data set (10 pts)

Three data sets (set A, B, and C) have been created following normally distributed classes. These data sets provide examples of car models where:

- The first column represents the cost in USD.
- The second column represents the weight in pounds.
- The third (last) column corresponds to the type (0 for small, 1 for big car).

Each data set contains 4,000 samples.

For each data set, do the following:

1. Normalize the data, then plot the data using software of your choice (for two vehicle types). (2 points)
2. Estimate a separation line and draw it manually (by hand) on that plot. This line will be a linear separator, which separates "small" cars from "big" cars. At this time, no running algorithm is needed (we will do that in the next assignment). (1 point)
3. Determine the mathematical definition of this linear separator. Based on this linear separator (equation), determine the inequality that selects "small" cars. What are the weights and threshold? Comment. (2 points)
    a. Important: inequality defines neuron's functionality (think of the inequality we covered in Session 04&05). Inequality "makes" the decision (chooses a portion of XoY space here). Deciding on inequality means determining weights and threshold of a neuron.
4. Provide a confusion matrix (false positives, false negatives, etc.) (1 point)
5. Calculate accuracy, error, true positive rate and true negative rate, false-positive rate, and false-negative rate. (Note: the true positive rate is different from the true positive). (2 points)
6. Compare results for each data set and explain the differences. How are these datasets different? Why was data normalization helpful? (2 points)

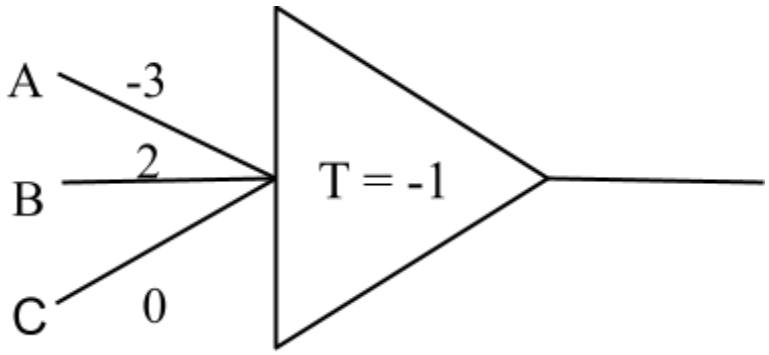
Note: An example of true positive: the class is "it is a small car" and prediction is "small car"

B) **McCulloch-Pitts neurons (5 pts)**

1. Create a truth table for the artificial neuron below. What is the **functionality of** this neuron? (3 points)
2. Given the same set of weights and the determined functionality of a neuron, what would be the **range of possible values for threshold**? (2 points)

Note: Consider unipolar hard threshold activation function (possible inputs/outputs are obviously 0 & 1). Always start with the unit definition (net, output). The functionality of the neuron is a Boolean function.

Hint: The truth table (similar to the one in class) should present inequalities that will evidence the functionality of a neuron (prove that it works as promised).



Note:

1. Compile all your deliverables into a single **pdf** file - this is YOUR REPORT.
2. Archive/zip report with other pertinent files (code, data, other).
3. Submit your file (Canvas). Please name the zip file as GroupName_Project1.zip.