

COSC 190: HW #4

Machine Learning I

Instructor: Dr. Fred Park

The goal for machine learning, in this context is, for the computer to be able to learn from data without being programmed. An example would be for the computer to learn a model for pass/no-pass based on exam scores and labeled data. By labeling, we mean that a value of 1 means pass and 0 no-pass. Thus, if two midterm scores were input that were not part of the training set, the computer would be able to tell you if they passed or not without ever knowing the specific grading criteria. (In fact, it can also learn the grading criteria if needed.) An example: if scores were 40 and 80 for midterms 1 and 2 respectively, the computer would tell you that the student did not pass. However, if the scores were 40 and 90, it would tell you that they did in fact pass. Thus, the computer learned a decision boundary as to what scores would allow a pass or not. Based on this, we can make predictions on pass/no-pass of future students.

1. Create a data set with at least 10 examples. Each data point must have 2 features associated to it to make a classification decision. For example, if we wanted to tell the difference between a square and a triangle, we would create 5 examples of each where the features would be
 - feature #1: number of corners
 - feature #2: average of the angles between adjacent sides.

we would then assign label value 1 if the shape is a square and label 0 if it is not. You can use college admission as an example, final exam+class attendance as a criteria for pass/no-pass, or any other example you can think of.

- (a) By hand, plot feature 1 vs feature 2 on a sheet of paper. Use an “x” if it is a square as a marker and an “o” if it is not.
- (b) Can you determine a decision boundary visually? Draw it with a dotted line. Now take two points on the line and find out what their x and y values are.
- (c) From the two points on your nicely drawn dotted line, find the equation of the line going through those two points.
- (d) Plot the line on your nicely hand-drawn plot. You have just trained a classifier!
- (e) Create 2 testing examples and find a way to use your decision boundary to determine the class that each testing example belongs to.