CECS 551 Programming Assignment 2

Dr. Todd Ebert

Due Date: 3:30pm on March 1st

Note: Exercise 1 of this Assignment will be assessed as problem B of Quiz 2.

Preparatory Reading

Read Sections 4.2 and 5.1.

Presenting Your Work.

Submit via email an R script named `assign2-StudentFirstName.R`. For example, if your first name is George, then you would submit the file `assign2-George.R`. Line 1 of the file should include your full name as a comment. Also, for each block of code that pertains to some exercise, say Exercise 1, preface the code with the comment (all in caps)

```
#EXERCISE 1
```

Each function should be named *exactly* as it appears in the exercise, and have an input signature that follows the order specified in the exercise.

Important: make sure your file produces no errors when sourced by the R interpreter. Otherwise your work will not be graded. Therefore, the answers to Exercises 4 and 5 should be placed in comments. For example,

```
#EXERCISE 5
# w = (-4.5,10.8, -5.6, 12.5, 11.2), b = -0.3
```

Your email submission should be postmarked no later than the time shown on the due date. Submissions received after that time, but on the same day will lose one letter grade. Submissions emailed on a later day will either not be accepted, or will lose two letter grades.

Exercises

1. Write an R function called **simple_learner** that takes as input a data frame `df` (you may assume that all columns of `df` are numeric, and the final column provides a class label of 1 or -1) and returns a list having two attributes w and b, where $w = \overline{c}_+ - \overline{c}_-$ and $b = \overline{w} \cdot \overline{c}$, where $\overline{c}_+$ and $\overline{c}_-$ are the respective centers of mass of the positive and negative datapoints, and $\overline{c} = (\overline{c}_+ + \overline{c}_-)/2$ is a point on the linear decision surface.

2. Write an R function called **perceptron_learner** that takes as input a data frame `df` (you may assume that all columns of `df` are numeric, and the final column provides a class label of 1 or -1) and returns a list having two attributes w and b, where w and b are obtained by performing the perceptron learning algorithm using the vectors from `df`.

3. Write an R function called **classify** that takes as input i) data frame `df`, and ii) linear decision surface parameters w and b, and returns a vector v that provides the classifications of each of the vectors of `df`. Thus each component of v is either 1 or -1, and the length of v equals the number of vectors of `df`. Moreover, we assume that the number of columns of `df` is equal to the dimension of $\overline{w}$.

4. Use the data in file `exercise-4.csv` to build a data frame `df`, and provide the $\overline{w}$ and b values that are returned by `simple_learner(df)`.

5. Repeat Exercise 4, but now provide the $\overline{w}$ and b values that are returned by `perceptron_learner(df)`.