CSE 151A Winter 2021 - Take Home Exam 4

1 hour **Name:** **PID:**

---

(1) This is an open book, open notes exam. You are free to consult any text book or notes. **You are not allowed to consult with any other person.**
(2) If you need any clarification, please post a private message to the instructors on Piazza.
(3) Remember that your work is graded on the *clarity* of your writing and explanation as well as the validity of what you write.
(4) This is a one-hour exam.

(1) We are given a training set $S = \{(x_1, y_1), \ldots, (x_n, y_n)\}$ where $y_i \in \{-1, 1\}$ and each x_i is a $d \times 1$ vector. Suppose we know that the feature vectors $x_1, \ldots, x_n$ lie on a k-dimensional subspace T of $\mathbb{R}^d$. State whether the following statements are true or false. Justify your answer.
   (a) (5 points) Suppose we run Perceptron for a single pass on S starting with an initial point $w_0 = 0$ (the all zeros vector). Does the output w_P lie in T? Justify your answer.

   (b) (5 points) Now suppose we run gradient descent for logistic regression on S for a 100 iterations starting with an initial point $w_0 = 0$. Does the output w_L lie in T? Justify your answer.

(2) (5 points) Write down an example of a dataset that is (a) linearly separable, but (b) where running a single pass of Perceptron does not lead to a classifier with zero training error.

(3) (5 points) Suppose Alice and Bob are given the same training dataset S. Alice finds a classifier w_A that exactly minimizes the logistic regression loss function. Bob finds a classifier w_B that minimizes the loss function:

$$w_B = \mathbf{argmin}_w \exp(\sum_{i=1}^{n} \log(1 + e^{-y_i w^\top x_i}))$$

Is $w_A = w_B$ for all training sets S? Justify your answer.