CSE 251A Winter 2021 - Take Home Exam 3

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) In class we looked at convex and strongly convex functions. A function $f : \mathbb{R}^d \to \mathbb{R}$ is said to be positively convex if one of the following three conditions hold:
  (a) For all $x, y \in \mathbb{R}^d$ and $0 < t < 1$, $f(tx + (1-t)y) < tf(x) + (1-t)f(y)$.
  (b) If f is differentiable then for all $x, y \in \mathbb{R}^d$, $f(y) > f(x) + \nabla f(x)^\top (y - x)$.
  (c) If f is doubly differentiable, then for all $x \in \mathbb{R}^d$, $\nabla^2 f(x)$ is positive definite. In other words for all $z \in \mathbb{R}^d$ such that $z \neq 0$, we have $z^\top \nabla^2 f(x) z > 0$.

Notice that the difference between convexity and positive convexity is that the inequalities are strict. In what follows, w and x are vectors in $\mathbb{R}^d$ and y is a scalar. Answer the following questions.
  (a) (5 points) Is the function $f(w) = \log(1+e^{-yw^\top x})$ a positively convex function of w? Justify your answer.

  (b) (5 points) Let $g(w) = (w^\top x - y)^2 + \frac{1}{2}\|w\|^2$. Is g a positively convex function of w? Justify your answer.

(2) We have a training set $S = \{(x^{(i)}, y^{(i)}), i = 1, \ldots, n\}$. Suppose we transform the feature vectors by scaling each vector by a constant $c > 0$ to get a new training set T. In other words, $T = \{(cx^{(i)}, y^{(i)}, i = 1, \ldots, n\}$. Assume that we are looking for linear classifiers whose decision boundary pass through the origin – namely, the term $b = 0$.

(a) (5 points) Write down the Perceptron algorithm for training set T, with the weight vector initialized as $w = 0$.

(b) (5 points) Suppose w_S is the output of Perceptron for S and w_T is the output of Perception for T. Is $w_S = w_T$? Justify your answer through a brief proof or counterexample.