CMSC5724: Exercise List 3

Problem 1. Consider d boolean variables $x_1, x_2, ..., x_d$. An *expression* is a conjunction of d literals where the i-th *literal* ($i \in [1, d]$) is either x_i or its negation $\overline{x_i}$. For example, when $d = 3$, $x_1 \wedge x_2 \wedge \overline{x_3}$ and $\overline{x_1} \wedge x_2 \wedge \overline{x_3}$ are both expressions. Let $\mathcal{H}$ be the set of all expressions. Define the instance space $\mathcal{X} = \{0, 1\}^d$ and label space $\mathcal{Y} = \{-1, 1\}$. Note that each $x \in X$ is a d-dimensional boolean vector $x = (x_1, ..., x_d)$. For each expression (i.e., classifier) $h \in \mathcal{H}$, $h(x)$ equals 1 if evaluating the expression h on x gives 1, or -1 if the evaluation gives 0. Consider $d = 10$. Denote by $\mathcal{D}$ a distribution over $\mathcal{X} \times \mathcal{Y}$. Let S be a training set of objects drawn independently from $\mathcal{D}$. Prove: when $|S| \geq 50000$, with probability at least 0.9, the error of h on $\mathcal{D}$ is higher than the error of h on S by at most 0.01, for every $h \in \mathcal{H}$.

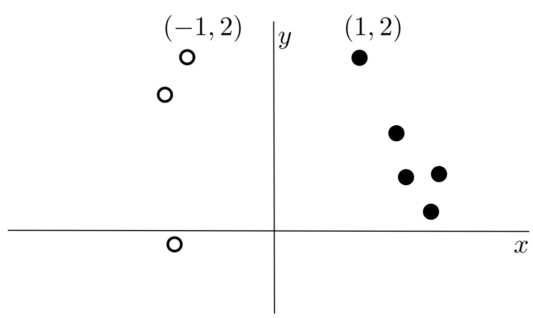
Problem 2. Let P be a set of 4 points: $A = (1, 2, 1)$, $B = (2, 1, 1)$, $C = (0, 1, 1)$ and $D = (1, 0, 1)$. A and B have label 1, while C and D have label -1. Execute Perceptron on P. Give the weight vector $\boldsymbol{w}$ maintained by the algorithm after each iteration.

Problem 3. Let P be a set of multidimensional points where each point has a label equal to 1 or -1. We want to design an algorithm to achieve the following purpose:

- Either return a separation plane (see the lecture notes for the definition of separation plane);
- Or declare that P has no separation planes with a margin at least γ.

Your algorithm must still work even if no separation planes exist.

Problem 4. Consider the set of points below where points of different colors carry different labels. Only two points have their coordinates shown. Apply Perceptron to find a separation plane on the set. Prove: Perceptron finishes after at most 5 iterations.



Problem 5. Some people prefer the following variant of the Perceptron algorithm:

1. $\boldsymbol{w} = \boldsymbol{0}$
2. **while** there is a violating point $\boldsymbol{p}$
3.   **if** $\boldsymbol{p}$ has label 1
4. $\boldsymbol{w} = \boldsymbol{w} + \lambda \cdot \boldsymbol{p}$
    else
5. $\boldsymbol{w} = \boldsymbol{w} - \lambda \cdot \boldsymbol{p}$

where λ is a positive real-value constant. In the version we discussed in the lecture, $\lambda = 1$. Prove: regardless of λ, Perceptron always terminates in R^2/γ^2 iterations, where R is the maximum distance of the points to the origin and γ the largest margin of all separation planes.

Problem 6. Let P be a set of points in $\mathbb{R}^d$, where each point is labeled 1 or -1. A d-dimensional plane π is a *separation plane* of P if

- π does not pass any point in P;
- the points of the two labels in P fall on different sides of π.

Note that we do not require π to pass the origin.

Construct a $(d+1)$-dimensional point set P' as follows: given each $p \in P$, add to P' the point $(p[1], p[2], ..., p[d], 1)$ (i.e., adding a new coordinate 1), carrying the same label as p. Prove: P has a separation plane if and only if P' has a separation plane passing the origin of $\mathbb{R}^{d+1}$.