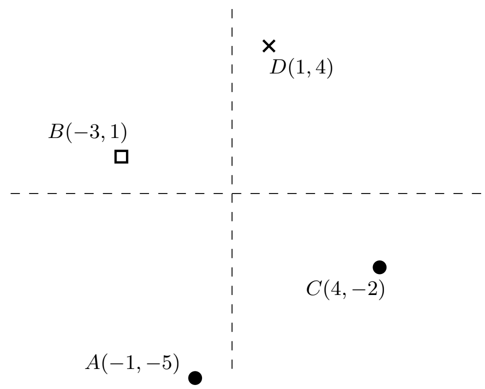
CMSC5724: Exercise List 7

Problem 1. Consider the training set P of points shown below:



where the two dots have label 1, the cross has label 2, and the box has label 3. Run multiclass Perceptron to find a generalized linear classifier to separate P.

Problem 2. Calculate the margin of the classifier you obtained in the previous problem.

Problem 3. Suppose we run multiclass Perceptron on $k = 2$. Let $\{\vec{w_1}, \vec{w_2}\}$ be the set of weight vectors returned. Prove: $\vec{w_1} = -\vec{w_2}$.

Problem 4. Continuing on Problem 3, prove: the "margin" of $W = \{\vec{w_1}, \vec{w_2}\}$ as defined in multiclass Perceptorn is precisely the "margin" as defined in (the traditional) Perceptorn (i.e., the smallest distance from a point in the training set P to the separation plane).

Problem 5 (Multi-Class Generalization Theorem). Let $\mathcal{X}$ be an instance space, $\mathcal{Y} = \{1, 2., ..., k\}$ be a label space, and $\mathcal{D}$ a distribution over $\mathcal{X} \times \mathcal{Y}$. Let S be a set of independent samples drawn from $\mathcal{D}$. A *classifier* h is a function $h : \mathcal{X} \to \mathcal{Y}$. For every such h, define

$$\begin{aligned} \mathrm{er}(h) &= \Pr_{(x,y)\sim\mathcal{D}}[h(x) \neq y] \\ \mathrm{er}_S(h) &= \frac{|\{(x,y) \in S \mid h(x) \neq y\}|}{|S|}. \end{aligned}$$

Let $\mathcal{H}$ be a finite set of classifiers. Fix a value δ such that $0 < \delta \leq 1$. Prove: with probability at least $1 - \delta$, we have the property that

$$\mathrm{er}(h) \leq \mathrm{er}_S(h) + \sqrt{\frac{\ln(1/\delta) + \ln|\mathcal{H}|}{2|S|}}$$

holds true for every $h \in \mathcal{H}$.