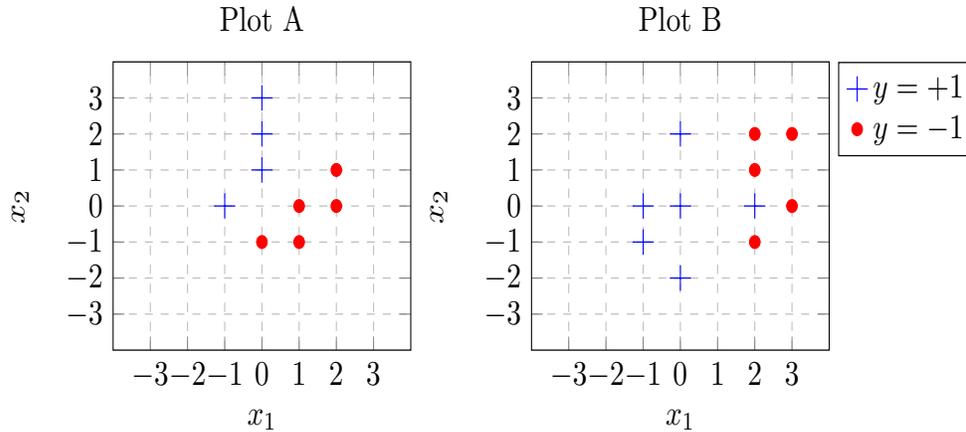


Linear Classifiers

1. (14 points) In the plots below, we give you 2D points with +1 and -1 labels.



Answer the following questions for both plot A and plot B:

- (a) Using a linear separator $h(p; \theta, \theta_0) = \text{sign}(\theta^\top p + \theta_0)$, what is the minimum possible number of misclassified points?

Plot A:

Plot B:

- (b) What are the values of $\theta \in \mathbb{R}^2$ and $\theta_0 \in \mathbb{R}$ that define your separator?

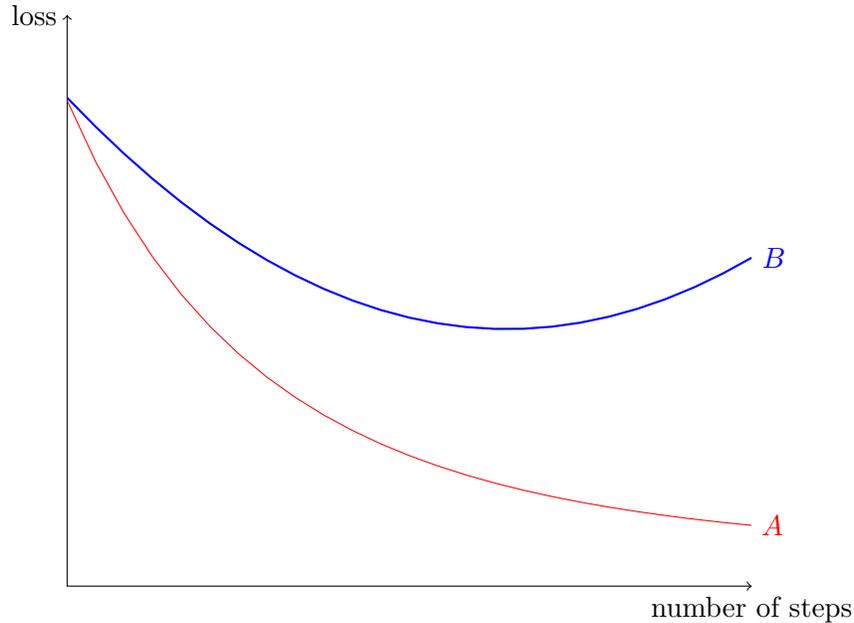
Plot A:

Plot B:

- (c) For a given point p , what does $\frac{\theta^\top p + \theta_0}{\|\theta\|}$ intuitively represent?

Name: _____

Consider the following plot from the previous classification task. The two curves show the train and test error vs. the number of steps in the optimization algorithm.



(d) Assign the appropriate labels:

Test error (select one): A B

Train error (select one): A B

(e) Which of the following options can improve the final performance of the trained classifier on the test data set? Note: augmentation of a data set refers to taking the existing data set and adding many points which are slightly perturbed versions of the original points. Select all that apply.

A. Augment the training data set and retrain the classifier.

B. Augment the test data set and retrain the classifier.

C. Terminate the training process earlier.

D. Add a penalty on the magnitude of the parameter values and retrain the classifier.