Dis Fri

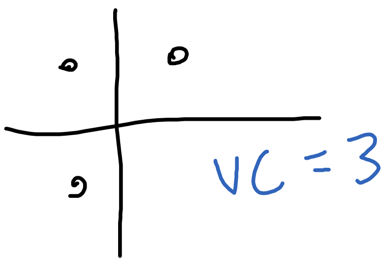
Friday, November 16, 2018   12:12 PM

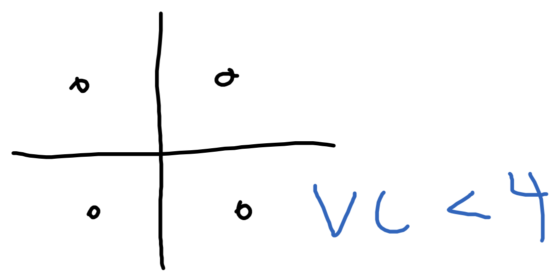
The **VC Dimension** of h is d if for a set of examples $x_1, \ldots, x_d$, there exists some labeling $y_1, \ldots, y_d$ s.t. some h can fit the data with zero training error

- a measure of the complexity of hypothesis space H

Our model **shatters** a given data set with specific labels if it has zero training error

Ex:
$h \in \operatorname{sgn}\left(b_1 x_1 + b_2 x_2 + b\right)$





SVM
We want to find a separating hyperplane with the maximum margin
For hyperplane $B^T x + B_0$
The margin is $\frac{1}{\|B\|}$
so we want to

$$\min \frac{1}{2}|B|^2 \;\; s.t.\; y\left(B^T x + B_0\right) \geq 1$$

The Lagrangian of the optimization problem:

$$L\left(B, B_0, \alpha\right) = \frac{1}{2}|B|^2 - \sum \alpha_i \left(y_i\left(B^T x_i + B_0\right) - 1\right)$$

$$\frac{\partial L}{\partial B} = B - \sum_i^N \alpha_i y_i x_i = 0 \;\; \Rightarrow B^* = \sum_1^N \alpha_i y_i x_i$$

$$\frac{\partial L}{\partial B_0} = \sum_1^N \alpha_i y_i = 0$$

1. Stationary point
2. Primal feasibility
3. Dual feasibility
4. Complementary slackness

Homework
1. VC=3
2.
3.
    a. -1, 2
    b. want the smallest possible w's satisfying the equations so that we minimize the norm