CSCI4230 Computatioal Learning Theory   Spring 2019
Lecturer: Siu On Chan   *Based on Rocco Servedio's notes*

Notes 5: VC dimension

1. VAPNIK–CHERVONENKIS DIMENSION

Related to mistake lower bounds in Online Learning

Usually an integer, telling us how expressive a concept class $\mathcal{C}$ is

Given concept class $\mathcal{C}$ over instance space X, subset $S \subseteq X$ is **shattered by** $\mathcal{C}$ if all "dichotomies" of S can be induced by $\mathcal{C}$, i.e.:

$$\forall T \subseteq S, \exists c \in C \text{ s.t. } c \cap S = T$$

$X =$ the plane $= \mathbb{R}^2 \qquad \mathcal{C} =$ axis-aligned rectangles
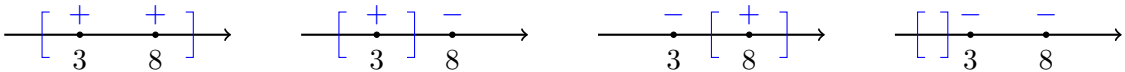
$S = \bullet \qquad$ has all dichotomies such as $\qquad$ and

---
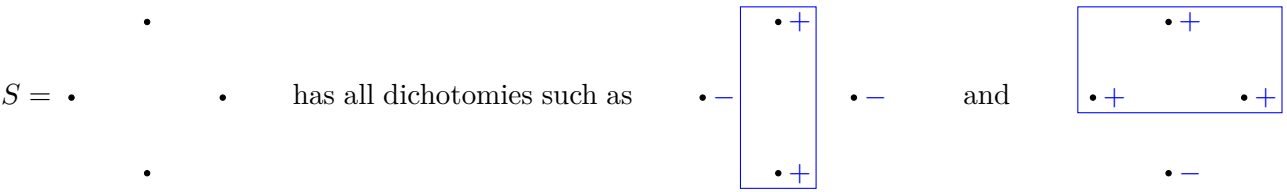
VCDim($\mathcal{C}$) is the size of the largest subset $S \subseteq X$ shattered by $\mathcal{C}$

VCDim($\mathcal{C}$) $= d$ if and only if

    (1) some subset $S \subseteq X$ with $|S| = d$ is shattered by $\mathcal{C}$; and

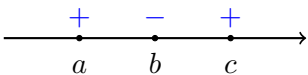   (2) all subsets of size $d + 1$ is not shattered by $\mathcal{C}$

VCDim($\mathcal{C}$) can be ∞
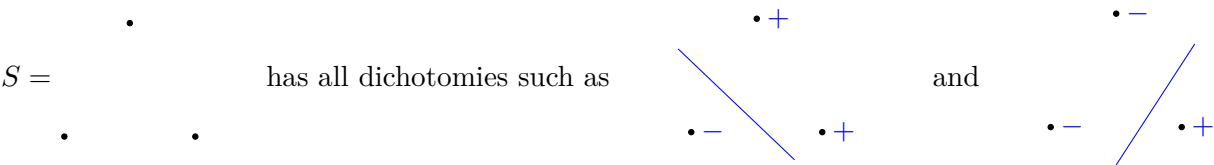

Example: Closed intervals of the real line

$X = \mathbb{R} \qquad \mathcal{C} =$ closed intervals $= \{[a, b] \mid a, b \in \mathbb{R}\}$ where $[a, b] = \{x \in \mathbb{R} \mid a \leqslant x \text{ and } x \leqslant b\}$

Every two points (e.g. 3 and 8) can be shattered $\qquad \Longrightarrow \qquad$ VCDim($\mathcal{C}$) $\geqslant 2$

No three points ($a < b < c$) can be shattered $\qquad \Longrightarrow \qquad$ VCDim($\mathcal{C}$) $\leqslant 2$

Example: Halfspaces in the plane $\qquad X = \mathbb{R}^2 \qquad \mathcal{C} =$ LTF

Any three non-collinear points can be shattered $\qquad \Longrightarrow \qquad$ VCDim($\mathcal{C}$) $\geqslant 3$

$S = \qquad\qquad$ has all dichotomies such as $\qquad\qquad$ and

No four points can be shattered $\qquad \Longrightarrow \qquad$ VCDim($\mathcal{C}$) $\leqslant 3$

Case 1: contains three collinear points
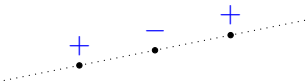
Case 2: No three points collinear

    Case 2a: Some point inside the triangle formed by three other points

    Case 2b: Four points form a convex quadrilateral $\qquad \Longleftrightarrow \qquad$ the two diagonals cross