CSE 417T: Introduction to Machine Learning

Lecture 5: VC-Dimension

Henry Chai

09/11/18

Recall: Hoeffding's Inequality

- $P\{|E_{in}(g) - E_{out}(g)| > \epsilon\} \leq 2(m)e^{-2\epsilon^2 n}$ where $m = |\mathcal{H}|$
- Growth function: $m_{\mathcal{H}}(n) = \max_{(\overrightarrow{x_1},\ldots,\overrightarrow{x_n}) \in \mathcal{X}} |\mathcal{H}(\overrightarrow{x_1}, \ldots, \overrightarrow{x_n})|$
- Can we simply replace m in Hoeffding's inequality with the growth function?

Vapornik-Chervonenkis (VC)-Bound

$$P\{|E_{in}(g) - E_{out}(g)| > \epsilon\} \leq 4m_{\mathcal{H}}(2n)e^{-\frac{1}{8}\epsilon^2 n}$$

$$\Updownarrow$$

$$E_{out}(g) \leq E_{in}(g) + \sqrt{\frac{8}{n}\log\left(\frac{4m_{\mathcal{H}}(2n)}{\delta}\right)}$$

with probability at least $1 - \delta$

Recall

- Claim: if $\exists\ k$ s.t. k is a breakpoint for $\mathcal{H}$, then $m_{\mathcal{H}}(n)$ is bounded by a polynomial in n

- If $m_{\mathcal{H}}(n)$ is bounded by a polynomial in n, then

  $\sqrt{\frac{8}{n}\log\left(\frac{4m_{\mathcal{H}}(2n)}{\delta}\right)} \to 0$ as $n \to \infty$

Bounding $m_{\mathcal{H}}(n)$

- Let $B(n,k) =$ the maximum number of dichotomies on n points s.t. no subset of k points is shattered
- If k is a breakpoint for $\mathcal{H}$, then $m_{\mathcal{H}}(n) \leq B(n,k)$
- If $B(n,k)$ is bounded by a polynomial in n and $m_{\mathcal{H}}(n)$ is bounded by $B(n,k)$, then $m_{\mathcal{H}}(n)$ is bounded by a polynomial in n

$B(n,k)$: Example

- What is $B(4,2)$?

$\vec{x_1}$	$\vec{x_2}$	$\vec{x_3}$	$\vec{x_4}$
+1	+1	+1	+1
+1	+1	+1	-1
+1	+1	-1	+1
+1	-1	+1	+1
-1	+1	+1	+1

$B(n,k)$: Example

- What is $B(4,2)$?

$\overrightarrow{x_1}$	$\overrightarrow{x_2}$	$\overrightarrow{x_3}$	$\overrightarrow{x_4}$
+1	+1	+1	+1
+1	+1	+1	-1
+1	+1	-1	+1
+1	-1	+1	+1
-1	+1	+1	+1
-1	-1	+1	+1

$B(n,k)$: Example

- What is $B(4,2)$?

$\vec{x_1}$	$\vec{x_2}$	$\vec{x_3}$	$\vec{x_4}$
+1	+1	+1	+1
+1	+1	+1	-1
+1	+1	-1	+1
+1	-1	+1	+1
-1	+1	+1	+1
+1	-1	-1	-1

$B(n, k)$: Example

- $B(4, 2) = 5$

$\vec{x_1}$	$\vec{x_2}$	$\vec{x_3}$	$\vec{x_4}$
+1	+1	+1	+1
+1	+1	+1	-1
+1	+1	-1	+1
+1	-1	+1	+1
-1	+1	+1	+1

$B(n,k)$: Example

- $B(4,2) = 5$

$\overrightarrow{x_1}$	$\overrightarrow{x_2}$	$\overrightarrow{x_3}$	$\overrightarrow{x_4}$
+1	+1	+1	+1
+1	-1	-1	-1
+1	+1	-1	+1
+1	-1	-1	+1
-1	+1	+1	+1

$B(n, k)$: Example

- $B(3, 2) = 4$

$\vec{x_1}$	$\vec{x_2}$	$\vec{x_3}$
+1	+1	+1
+1	+1	-1
+1	-1	+1
-1	+1	+1

$B(n, k)$: Example

- $B(4, 3) = 11$

$\overrightarrow{x_1}$	$\overrightarrow{x_2}$	$\overrightarrow{x_3}$	$\overrightarrow{x_4}$
-1	-1	+1	+1
-1	+1	-1	+1
+1	-1	-1	-1
+1	+1	+1	+1
-1	+1	+1	+1
+1	-1	+1	+1
+1	+1	-1	+1
+1	+1	+1	-1
-1	+1	+1	-1
+1	-1	+1	-1
+1	+1	-1	-1

$B(n, k)$: Example

- $B(4,3) = \alpha + 2\beta$

$\overrightarrow{x_1}$	$\overrightarrow{x_2}$	$\overrightarrow{x_3}$	$\overrightarrow{x_4}$	
-1	-1	+1	+1	
-1	+1	-1	+1	$S_1, \lvert S_1 \rvert = \alpha$
+1	-1	-1	-1	
+1	+1	+1	+1	
-1	+1	+1	+1	$S_2^+, \lvert S_2^+ \rvert = \beta$
+1	-1	+1	+1	
+1	+1	-1	+1	
+1	+1	+1	-1	
-1	+1	+1	-1	$S_2^-, \lvert S_2^- \rvert = \beta$
+1	-1	+1	-1	
+1	+1	-1	-1	

$B(n,k)$: Example

- $B(4,3) = \alpha + 2\beta$
- $\alpha + \beta \le B(3,3)$
- $\beta \le B(3,2)$
- $B(4,3) \le B(3,3) + B(3,2)$

$\vec{x_1}$	$\vec{x_2}$	$\vec{x_3}$	$\vec{x_4}$
-1	-1	+1	+1
-1	+1	-1	+1
+1	-1	-1	-1
+1	+1	+1	+1
-1	+1	+1	+1
+1	-1	+1	+1
+1	+1	-1	+1
+1	+1	+1	-1
-1	+1	+1	-1
+1	-1	+1	-1
+1	+1	-1	-1

$S_1, |S_1| = \alpha$

$S_2^+, |S_2^+| = \beta$

$S_2^-, |S_2^-| = \beta$

$B(n,k)$: Recursion

- $B(n,k) \le B(n-1,k) + B(n-1,k-1)$
- $B(n,k) = \alpha + 2\beta$

$\overrightarrow{x_1}$	$\overrightarrow{x_2}$	...	$\overrightarrow{x_{n-1}}$	$\overrightarrow{x_n}$	
-1	-1	...	+1	-1	$S_1, \lvert S_1 \rvert = \alpha$
-1	+1	...	-1	+1	
⋮	⋮	⋮	⋮	⋮	
+1	-1	...	-1	-1	
+1	+1	...	+1	+1	$S_2^+, \lvert S_2^+ \rvert = \beta$
-1	+1	...	+1	+1	
⋮	⋮	⋮	⋮	⋮	
+1	-1	...	+1	+1	
+1	+1	...	+1	-1	$S_2^-, \lvert S_2^- \rvert = \beta$
-1	+1	...	+1	-1	
⋮	⋮	⋮	⋮	⋮	
+1	-1	...	+1	-1	

$B(n,k)$: Recursion

- $B(n,k) \leq B(n-1,k) + B(n-1,k-1)$
- $B(n,k) = \alpha + 2\beta$
- $\alpha + \beta \leq B(n-1,k)$

$\overrightarrow{x_1}$	$\overrightarrow{x_2}$	…	$\overrightarrow{x_{n-1}}$	$\overrightarrow{x_n}$
-1	-1	…	+1	-1
-1	+1	…	-1	+1
⋮	⋮	⋮	⋮	⋮
+1	-1	…	-1	-1
+1	+1	…	+1	+1
-1	+1	…	+1	+1
⋮	⋮	⋮	⋮	⋮
+1	-1	…	+1	+1
+1	+1	…	+1	-1
-1	+1	…	+1	-1
⋮	⋮	⋮	⋮	⋮
+1	-1	…	+1	-1

$S_1, |S_1| = \alpha$

$S_2^+, |S_2^+| = \beta$

$B(n,k)$: Recursion

- $B(n,k) \leq B(n-1,k) + B(n-1,k-1)$
- $B(n,k) = \alpha + 2\beta$
- $\alpha + \beta \leq B(n-1,k)$
- $\beta \leq B(n-1,k-1)$

$\overrightarrow{x_1}$	$\overrightarrow{x_2}$	…	$\overrightarrow{x_{n-1}}$	$\overrightarrow{x_n}$
-1	-1	…	+1	-1
-1	+1	…	-1	+1
⋮	⋮	⋮	⋮	⋮
+1	-1	…	-1	-1
+1	+1	…	+1	+1
-1	+1	…	+1	+1
⋮	⋮	⋮	⋮	⋮
+1	-1	…	+1	+1
+1	+1	…	+1	-1
-1	+1	…	+1	-1
⋮	⋮	⋮	⋮	⋮
+1	-1	…	+1	-1

$S_2^-, |S_2^-| = \beta$

Bounding $B(n,k)$

- Claim: $B(n,k) \leq \sum_{i=0}^{k-1} \binom{n}{i}$

- Intuition: **<u>if recursion, then induction</u>**

Bounding $B(n,k)$

- Claim: $B(n,k) \leq \sum_{i=0}^{k-1} \binom{n}{i}$
- Induction (base cases):
- $\forall\, n \geq 1,\ B(n,1) = 1 \leq \sum_{i=0}^{k-1} \binom{n}{i} = \binom{n}{0} = 1$
- $\forall\, k \geq 2,\ B(1,k) = 2 \leq \sum_{i=0}^{k-1} \binom{1}{i} = \binom{1}{0} + \binom{1}{1} = 2$

Bounding $B(n,k)$

- Claim: $B(n,k) \leq \sum_{i=0}^{k-1} \binom{n}{i}$
- Induction (inductive step):
- Suppose $B(n',k') \leq \sum_{i=0}^{k'-1} \binom{n'}{i}$ $\forall\, n' < n, k' < k$
- Prove $B(n,k) \leq \sum_{i=0}^{k-1} \binom{n}{i}$

Bounding $B(n,k)$

- Prove $B(n,k) \le \sum_{i=0}^{k-1} \binom{n}{i}$

- $B(n,k) \le B(n-1,k) + B(n-1,k-1)$

$$\le \sum_{i=0}^{k-1} \binom{n-1}{i} + \sum_{i=0}^{k-2} \binom{n-1}{i}$$

$$= 1 + \sum_{i=1}^{k-1} \binom{n-1}{i} + \sum_{i=1}^{k-1} \binom{n-1}{i-1}$$

$$= 1 + \sum_{i=1}^{k-1} \left(\binom{n-1}{i} + \binom{n-1}{i-1} \right) = 1 + \sum_{i=1}^{k-1} \binom{n}{i}$$

Bounding $B(n,k)$

- Prove $B(n,k) \leq \sum_{i=0}^{k-1} \binom{n}{i}$
- $B(n,k) \leq B(n-1,k) + B(n-1,k-1)$

$$\leq \sum_{i=0}^{k-1} \binom{n-1}{i} + \sum_{i=0}^{k-2} \binom{n-1}{i}$$

$$= 1 + \sum_{i=1}^{k-1} \binom{n-1}{i} + \sum_{i=1}^{k-1} \binom{n-1}{i-1}$$

$$= 1 + \sum_{i=1}^{k-1} \left(\binom{n-1}{i} + \binom{n-1}{i-1} \right) = \sum_{i=0}^{k-1} \binom{n}{i} \quad \blacksquare$$