作业2

2017年4月10日 8:19

1，分两种情况

① 样本正确，预测错误 ② 样本出错，预测正确

故 $p = \lambda\mu + (1-\lambda)(1-\mu) = 1-\lambda-(1-2\lambda)\mu$

2，h的表现与μ无关，则 $1-2\lambda = 0 \Rightarrow \lambda = \frac{1}{2}$

3，根据公式 $4(2N)^{d_{vc}}\exp(-1/8\,\varepsilon^2 N) = \delta$

$\varepsilon = 0.05$ $\delta = 1-95\% = 0.05$ $d_{vc} = 10$

求得 $N \approx 453000$ （可用 sympy 求解，参见代码）

4. d. e 可以化简

(d) $\varepsilon \le \sqrt{\frac{1}{N}\ln\frac{6m_H(2N)}{\delta}} + \frac{1}{N}$

(e) $\varepsilon \le \sqrt{\frac{1}{2N-4}\ln\frac{4m_H(N^2)}{\delta} + \left(\frac{1}{N-2}\right)^2} + \frac{1}{N-2}$

$m_H(N) = N^{d_{vc}}$ $d_{vc} = 50$ $\delta = 0.05$ $N = 10000$

计算易得 $\varepsilon_a = 0.63$ $\varepsilon_b = 0.86$ $\varepsilon_c = 0.33$ $\varepsilon_d = 0.22$. $\varepsilon_e = 0.21$ ε_e最小

（绘图参见代码）

5. $N=5$时 $\varepsilon_a = 13.83$ $\varepsilon_b = 16.26$ $\varepsilon_c = 7.05$ $\varepsilon_d = 5.09$ $\varepsilon_e = 5.55$ ε_d最小

6，N个数据点不同的符号，最多能分成三段，否则假设空间无法实现

分段型如：

$$\begin{cases}+\\-\end{cases}\quad\begin{cases}+\,-\\-\,+\end{cases}\quad\begin{cases}+\,-\,+\\-\,+\,-\end{cases}$$

故 $m_H(N) = 2\times\left(1+\binom{N-1}{1}+\binom{N-1}{2}\right) = N^2-N+2$

7，$N\le 3$时 $m_H(N) = 2^N$

$N=4$时 $m_H(N) < 2^N$

故 break point 为 4， $d_{vc} = 3$

8、令 $r = \sqrt{x_1^2+x_2^2}$，则原假设等价于

$$h(x) = \begin{cases}1 & r\in[a,b]\\ -1 & \text{otherwise}\end{cases}$$

故可根据N点的半径r形成的N+1个区间选择a，b值，

还有 [a b] 完全位于N点的区间之外

$$m_H(N) = \binom{N+1}{2}+1$$

9、证明 $d_{vc} = D+1$

① $d_{vc} \le D+1$

① $d_{vc} \le D+1$

实际上，可以证明对任意 D+1 个点均可以 shatter

D+1 个点按 ± 最多划分为 D+1 个区间，故最多需要多项式曲线与 X 轴的 D 个交点划分一般的，可取符号转换的两点的中点为交点，设为 $x_0, x_1, \dots x_j$ ($j<D$)

不失一般性，设最后一个区间的点取 +，

则 $h_c(x) = \mathrm{sign}\left(\prod_{k=0}(x-x_k)\right)$ 可以满足要求

② $d_{vc} \ge D+2$

对任意 D+2 个点，对其交错设置 ± 号，故需要多项式曲线与 X 轴有 D+1 个交点，

D 次多项式显然无法达成，故任意 D+2 个点均不能 shatter

10. 证明：$d_{vc} = 2^d$

① $d_{vc} \le 2^d$

考虑 d 维空间中的立方体的 2^d 个顶点 ($x_i = \pm 1$, $t_i = 0$)

则这些点形成了 2^d 个不同的 v 向量，通过 S 包含不同的向量，可以 shatter 这 2^d 个顶点

② $d_{vc} \ge 2^d+1$

由于 v 向量最多只有 2^d 个，故 2^d+1 个点中至少有两个点的 v 向量相同

这两个点符号不能相异，故这 2^d+1 个点不能 shatter

11. 直观上来说，该曲线为周期为 $\frac{1}{|\alpha|}$ 的三角波函数，周期可以任意小，

X 轴可以被该曲线划分为无限多区域，故 $d_{vc} = \infty$

严格证明，需要对任意 N，构造 N 个点，可以被三角波 shatter

基本思路：N 个点可取 ±1，分别用二进制 0, 1 表示，令 $k \in [0, 2^N-1]$

k 的第 i 位 ($i \in [0, N-1]$) 记为 $k(i)$，则构造一组 x_i, α_k

使得 $\begin{cases} k(i)=0 \text{ 时} & h_{\alpha_k}(x_i) = 1 \\ k(i)=1 \text{ 时} & h_{\alpha_k}(x_i) = -1 \end{cases}$ 暂未构造出来

12. $m_H(N) \le 2m_H(N-1)$ 因为增加 1 个点，该点可取两种情况

$\Rightarrow m_H(N) \le 2^2 m_H(N-2)$

$\Rightarrow \dots\dots$

$\Rightarrow m_H(N) \le 2^i m_H(N-i) \Rightarrow m_H(N) \le \min_{1\le i\le N-1} 2^i m_H(N-i)$

13. 假如 $m_H(N) = 2^{\lfloor\sqrt{N}\rfloor}$ 是成长函数，那么易知 $k=2$ 是 break point

而 $k=2$ 时 $m_H(N) \le N^{k-1}+1 = N+1$，N 足够大时，矛盾

故不是成长函数

14. 若 H_n 为空，则 $d_{vc} = 0$，故不等式左边成立是显然的

不等式右边等价于 $d_{vc} \le d_{vc}(H_k)$ $(k \in [1, K])$

反证法，若存在某个 k，$d_{vc} > d_{vc}(H_k)$

则 $\bigcap_{k=1}^{K} H_k$ 可以 shatter 超过 $d_{vc}(H_k)$ 个点

则由交集的含义，H_k 也可以 shatter 超过 $d_{vc}(H_k)$ 个点，矛盾

（通俗来看，H_k 可取参数范围更大）

15. H_k 能 shatter $d_{vc}(H_k)$ 个点，

则由并集的含义，$\bigcup_{k=1}^{K} H_k$ 也能 shatter $d_{vc}(H_k)$ 个点，不等式左边成立，

对并集而言，成长函数 $m_{H' \cup H}(N) \le m_H(N) + m_{H'}(N)$

设 $d_{vc}(H) = d$，$d_{vc}(H') = d'$，则

$$m_{H \cup H'}(N) \le \sum_{i=0}^{d} \binom{N}{i} + \sum_{i=0}^{d'} \binom{m}{i} = \sum_{i=0}^{d} \binom{N}{i} + \sum_{i=0}^{d'} \binom{N}{N-i} \le \sum_{i=0}^{d} \binom{m}{i} + \sum_{i=N-d'}^{N} \binom{m}{i}$$

如果 $N - d' > d+1$ 即 $N \ge d + d' + 2$ 则 $m_{H \cup H'} \le \sum_{i=0}^{N} \binom{N}{i} - \binom{N}{d+1} < 2^N$

故 $d_{vc}(H \cup H') \le d + d' + 1$

将该式推广易知不等式右边成立（逐个合并）

16. 考虑预测的错误率 μ

θ	s	μ
>0	$+$	θ
>0	$-$	$1-\frac{\theta}{2}$
<0	$+$	$-\theta$
<0	$-$	$1+\frac{\theta}{2}$

$$\Longrightarrow \mu = \begin{cases} |\theta|/2 & s = +1 \\ 1 - |\theta|/2 & s = -1 \end{cases} \Rightarrow \mu = \frac{1}{2} + s\frac{|\theta| - 1}{2}$$

则根据第1题结论

$$E_{out}(h_{s,\theta}) = \lambda\mu + (1-\lambda)(1-\mu) \qquad \lambda = 1 - 0.2$$

$$= 0.5 + 0.3\, s(|\theta| - 1)$$

17. 约为 0.17 参见代码

18. 约为 0.25 参见代码

19. 约为 0.25 参见代码

20. 约为 0.35 参见代码