

6. Дана следующая формулировка теор. NFL - Th.:

Пусть $\mathcal{H}$ - класс функций f заданн. фин. класс $\frac{m}{3} \text{ VCdim}(\mathcal{H}) \geq d$. Тогда $\forall$ алгоритма A , что работает на $\mathcal{H}$ (т.е. $E_{RM, \mathcal{H}} - \text{ан.}$) $\forall m \leq \frac{d}{k}$, $\exists$ некий $\mathcal{D}$ на $\mathcal{H} \times \{0, 1\}$ таков, что

- $L_{\mathcal{D}}(h) = 0$ для всех $h \in \mathcal{H}$;
- $E_{S \sim \mathcal{D}^m} (L_{\mathcal{D}}(A(S))) \geq \frac{d-m}{2d}$, а также $P_{S \sim \mathcal{D}^m} (L_{\mathcal{D}}(A(S)) \geq \frac{d-m}{4d}) \geq \frac{d-m}{3d+m}$.

З.ч. это все раз формулирует, что если $\text{VCdim}(\mathcal{H}) = \infty$ (т.е. d , а также m , могут быть любыми), $\mathcal{H}$ не PAC обуч.

У задачи 7. попер. заданная постановка задачи ан. или фикт для $\mathcal{H}$ - класс функций i $m < \frac{|\mathcal{H}|}{k}$. Для этого воспользуемся след. лем. C : $|C| = mk$ и тот факт, что $\mathcal{H}$ и

розбиває. Нехай менше C - з умови $VCdim(\mathcal{H}) \geq d$:
 $|C| = d \geq mk$. Для простоти візьмемо $d = mk$ (менше
 м'якшувати це зрештою можна). Злигнот з поперед. рез.,

$\forall A \exists \oplus : L_{\oplus}(h) = 0$ для $h \in \mathcal{H}$, і

$$\sum_{S \sim \mathcal{D}^m} (L_{\oplus}(A(S))) \geq \frac{k-1}{2k} = \frac{mk-m}{2mk} = \frac{d-m}{2d} \text{ частота ознак на } P.$$

Лема 4.1. Нехай $\mathcal{H}_1, \dots, \mathcal{H}_k$ - класи впрямоз. фин. клас. на X , і
 $d := \max_{i=1, \dots, k} VCdim(\mathcal{H}_i) \geq 3$. Доведемо наступну оцінку на
 VC-вим. об'єднання цих класів:

$$VCdim\left(\bigcup_{i=1}^k \mathcal{H}_i\right) \leq 4d \log_2 d + 2 \log_2 k.$$

Отже, нехай $\bigcup_{i=1}^k \mathcal{H}_i$ розбиває деяку C з $|C| = m$,
 тоді $\left|\left(\bigcup_{i=1}^k \mathcal{H}_i\right)_C\right| = 2^m$. З іншого боку,
 $2^m = \left|\left(\bigcup_{i=1}^k \mathcal{H}_i\right)_C\right| = \left|\bigcup_{i=1}^k (\mathcal{H}_i)_C\right| \leq \sum_{i=1}^k \left|(\mathcal{H}_i)_C\right| \leq \left[\frac{\text{deb. } T_{\mathcal{H}}}{\text{сп-гін}}
 нескінченно розбігається$

$$\leq \left[\begin{array}{l} T_K \\ \text{зрощ. сав. } K, \text{ малы } \text{визначено} \\ K = \arg\max_{K_i} V(\dim(K_i)) \text{ з } V(\dim(K)) = d \end{array} \right] \leq \psi \cdot T_K(m) < \left[\begin{array}{l} \text{дека } 3 \text{ аггера} \\ \text{для } m > d+1 \end{array} \right] <$$

$$< \psi \cdot \left(\frac{em}{d} \right)^d < [d \geq 3 > e] < \psi \cdot m^d. \text{ Визначено } \log_2$$

$$m = \log_2 2^m < \log_2 (\psi \cdot m^d) = d \log_2 m + \log_2 \psi.$$

$$\exists \text{ число } \delta \text{ таке } V(\dim(\bigcup_{i=1}^u K_i)) \geq \psi d \log_2 2d + 2 \log_2 \psi$$

$$\text{означено } \delta, \text{ що } \exists C, \text{ що розд. } \bigcup_{i=1}^u K_i \text{ з } |C| = m \geq$$

$$\geq \psi d \log_2 2d + 2 \log_2 \psi > d+1. \text{ Погоди за } \underline{\text{лем. 3}} \text{ рекурсії (що випливає}$$

$$\text{для } \log_2 \text{ - невідомо) } m \geq d \log_2 m + \log_2 \psi, \text{ що, як}$$

$$\text{показує, означає } \delta \text{ мб. про розд. } \psi.$$

$$\underline{6.} \text{ Для } \psi = 2 \text{ означає з } a \text{ що } V(\dim(K_1 \cup K_2)) \leq$$

$$\leq \psi d \log_2 2d + 2 \log_2 2 = \psi d \log_2 d + \psi d + 2. \text{ Каснабери}$$

$$\text{можна показати, що } V(\dim(K_1 \cup K_2)) \leq 2d+1.$$

$$\text{Діагно, нехай це не так, тоді } \exists C: m = |C| \geq 2d+2, i$$

$\mathcal{H}_1 \cup \mathcal{H}_2$ может. С. Так как

$$Z^m \leq T_{\mathcal{H}_1}(m) + T_{\mathcal{H}_2}(m) \leq [\text{Лема Зигмунда}] \leq 2 \sum_{i=0}^d C_m^i = \sum_{i=0}^d C_m^i + \sum_{i=0}^d C_m^{m-i}$$

$$\left[\begin{matrix} i \leq d \\ m-i \leq d \end{matrix} \right] = \sum_{i=0}^d C_m^i + \sum_{i=m-d}^m C_m^i \leq \left[\begin{matrix} m \geq 2d+2 \Rightarrow \\ m-d \geq d+2 \end{matrix} \right] \leq \sum_{i=0}^d C_m^i + \sum_{i=d+2}^m C_m^i < \sum_{i=0}^m C_m^i = Z^m \quad \Downarrow$$

8. Показати, що $\forall$ скінч. класу $\text{insets } \mathcal{K}$ з заданим fin. клас. і новим опису $d: \mathcal{K} \rightarrow \{0, 1\}^*$

$$\text{VCdim}(\mathcal{K}) \leq \max \{|d(h)| \mid h \in \mathcal{K}\} + 1.$$

Якщо не d префіксна, то

$$\text{VCdim}(\mathcal{K}) \leq \max \{|d(h)| \mid h \in \mathcal{K}\}.$$

Задаємо, що $\text{VCdim}(\mathcal{K}) \leq \log_2 |\mathcal{K}|$ (Зілом менше, $\leq \lfloor \log_2 |\mathcal{K}| \rfloor$, бо це натуральне число). Кількість усіх рядків (власно з перонієм) з довжиною $\leq n := \max \{|d(h)|\}$ - це $\underset{\text{нєл.}}{1} + \underset{\text{довж. 1}}{2} + \underset{\text{довж. 2}}{4} + \dots + 2^n = 2^{n+1} - 1$.

Оскільки $\mathcal{K}$ можна описати такою мовою ($\exists$ інж. відобр. з рядки), то ч.

$$\text{VCdim}(\mathcal{K}) \leq \log_2 |\mathcal{K}| \leq \log_2 (2^{n+1} - 1) < \log_2 2^{n+1} = n+1.$$

Якщо d префіксна, $\forall$ рядка $z \neq$ імовірно з початком z ,

матрицу, генерируемую S до n случайных чисел, попарно независимо,
 что S - матрица $n \times n$. Т.ч. каждая строка матрицы S имеет 2^n значений.

$$VCdim(K) \leq \log_2 |K| \leq \log_2 2^n = n$$

д Пусть K - звучащий класс n -омов γ задачи ML , d - предиктор, d - новый набор оценок,
 $B \in \mathbb{N}$, $K_B := \{h \in K \mid |h| \leq B\}$, где $|h| := |d(h)|$ - длина строки h . Пусть h_S - результат работы MDL -алгоритма
 на выборке S , а $h_B^* \in \argmin_{h \in K_B} L_\Phi(h)$ (т.е. оптимальная строка).

Оценим сверху $L_\Phi(h_S) - L_\Phi(h_B^*)$ (как известно, $\forall$ произв.
 Φ и $\text{им.} \geq 1 - \delta$) в терминах $|S| = m$, $B := \delta$ (oracle inequality).
 на $\mathbb{Z}$

За люб. MDL , $h_S \in \argmin_{h \in K} \left(L_S(h) + \sqrt{\frac{1}{2m} (|h| + \ln \frac{2}{\delta})} \right)$, и

за лю. $\mathbb{Z}$ распределением, где $\forall$ произв. Φ на $\mathbb{Z}$: $S \sim \Phi^m$ и $\text{им.} \geq 1 - \delta$

$$L_\Phi(h_S) \leq L_S(h_S) + \sqrt{\frac{1}{2m} (|h_S| + \ln \frac{2}{\delta})} \leq \left[\underset{h \in K}{h_S \text{ минимизирует}} \right] \leq L_S(h_B^*) +$$

$$+ \sqrt{\frac{1}{2m} (|h_B^*| + \ln \frac{2}{\delta})} \leq [\text{люб. } K_B] \leq L_S(h_B^*) + \sqrt{\frac{1}{2m} (B + \ln \frac{2}{\delta})} \leq [\text{мы не } \mathbb{P}_\gamma] \leq$$

$$\leq L_{\Phi}(h_B^*) + \sqrt{\frac{1}{2m}(|h_B^*| + \ln \frac{2}{\delta})} + \sqrt{\frac{1}{2m}(B + \ln \frac{2}{\delta})} = L_{\Phi}(h_B^*) + 2\sqrt{\frac{1}{2m}(B + \ln \frac{2}{\delta})}.$$

Тому ми ніде не використовуємо, що h_B^* мінімізує L_{Φ} :
 просто для кожного $L_{\Phi}(h_S) - L_{\Phi}(h_B^*)$ надіможе. Т.ч.
 воно $\leq 2\sqrt{\frac{1}{2m}(B + \ln \frac{2}{\delta})}$.

10 Нехай клас гіпотез $\mathcal{H}$ має визн. зм. об'єкту
 $\bigcup_n \mathcal{H}_n$, де усі $\mathcal{H}_n$ - (не дійсн. лінійні) зм. Тоді
 $\mathcal{H}$ зліченний. Як можна побудувати вичову ф-цію
 $w: \mathcal{H} \rightarrow [0, 1]$ чи навпаки? Обрати її так, щоб $w(h)$
 визначалася перем $n(h) = \min \{n \mid h \in \mathcal{H}_n\}$ і $|\mathcal{H}_n(h)|$,
 якщо $\mathcal{H}_n(h)$ скінч.

Спочатку оберемо вичову $w: \mathcal{N} \rightarrow [0, 1]$ для об'єкт.

$\mathcal{H} = \bigcup_n \mathcal{H}_n$, скажімо, $w(n) = \frac{1}{2^n}$.

Якщо $K_n(h)$ скінч., на певну константу розн. ~~вони~~ ~~визначені~~
 однозначно: $w(h) := \frac{1}{|K_n(h)|} \cdot \frac{1}{2^{n(h)}}$

Якщо $K_n(h)$ зліч. (нескінч.), безперервно іноді єд-на і
 розн. на пев. q -гріо $K \mapsto \frac{1}{2^K}$. Якщо $K(h)$ - певн. h
 $\notin K_n(h)$, то $w(h) := \frac{1}{2^{K(h)}} \cdot \frac{1}{2^{n(h)}}$

Ми гіпотезно побудували базову q -гріо, бо

$$\sum_{h \in K} w(h) = \sum_{\substack{h \in K: \\ K_n(h) \text{ скінч.}}} \frac{1}{|K_n(h)|} \cdot \frac{1}{2^{n(h)}} + \sum_{\substack{h \in K: \\ K_n(h) \text{ зліч.}}} \frac{1}{2^{K(h)}} \cdot \frac{1}{2^{n(h)}} \leq \sum_{h \in K} \frac{1}{2^{K(h)}} \leq \sum_{h \in K} \frac{1}{2^{n(h)}}$$

$$\leq \left[\text{голомоза класами} \right] \leq \sum_n \frac{1}{2^n} = 1.$$

Наприклад, для поліномів $K = \bigcup_n K_n$, де K_n - поліноми
 ст. $n-1$, K_n зліч., якщо розглядати ^{лише} раціональні коеф.,
 і скінч. - якщо коеф. зі скінч. множин (напр. в силу дискретизації).

11. Нехай $\mathcal{H}$ - клас порогових функцій на $\mathbb{R}^d$ з заданою системою класифікації. Покажати, що $V(\dim(\mathcal{H})) = \infty$.

Ми вже показали це для $d=1$ на лекції: $\forall$ скін.
 $C \subset \mathbb{R}$ $\mathcal{H}$ розб. C . Тепер $\forall$ m розмірено m -елементному підмнож.
 $C \subset \mathbb{R}^d$ на осі x^1 : $C = \{(c_i, 0, \dots, 0)\}_{i=1}^m$.
Потім за зведеним принципом від першої коорд. побудуємо
що неможливо. Але бачи, зокрема, $\in$ принципом від d
коорд.: $\forall f: C \rightarrow \{0,1\} \quad \exists h \in \mathcal{H}: h|_C = f$. Це й потрібно.