Zestaw 7 - Zadanie 7

Metody probabilistyczne w uczeniu maszynowym

Łukasz Trzos

Treść zadania

Niech $\mathcal{A}$ i $\mathcal{B}$ będą klasami pojęć. Wykaż, że dla $\mathcal{C} = \{a \cap b : a \in \mathcal{A}, b \in \mathcal{B}\}$ zachodzi $\Pi_{\mathcal{C}}(m) \leqslant \Pi_{\mathcal{A}}(m)\Pi_{\mathcal{B}}(m)$. Korzystając z powyższego faktu wykaż, że dla klasy pojęć $\mathcal{C}$ o wymiarze $VC(\mathcal{C}) = d$ i klas $\mathcal{C}_k = \{\bigcap_{i=1}^{k} c_i : c_i \in \mathcal{C}\}$ zachodzi

$$VC(\mathcal{C}_k) \leqslant 2dklog_2(3k)$$

dla każdego $k \geqslant 1$.

Rozwiązanie

Ograniczenie dla przecięcia klas pojęć

Weźmy dowolny zbiór m elementowy. Przyjmijmy, że klasa $\mathcal{A}$ realizuje na tym zbiorze m_A różnych klasyfikacji, analogicznie klasa $\mathcal{B}$ realizuje ich m_B. W klasie $\mathcal{C}$ każdy klasyfikator jest postaci $a \cap b$. Wiemy, że możemy wygenerować maksymalnie $m_A m_B$ parami różnych klasyfikatorów poprzez skrzyżowanie każdego z $\mathcal{A}$ z każdym z $\mathcal{B}$. Zatem klasa C może na tym zbiorze realizować maksymalnie $m_A m_B$ różnych klasyfikacji.

Ponieważ zachodzi to dla dowolnego zbioru, możemy rozważyć zbiór, który świadczy $\Pi_{\mathcal{C}}(m)$. Wiemy, że na tym zbiorze $\mathcal{A}$ i $\mathcal{B}$ realizowały kolejno jakieś m_A i m_B klasyfikacji, oraz, że $\Pi_{\mathcal{C}}(m) \leqslant m_A m_B$, ale skoro znaleźliśmy zbiory świadczące granice: $\Pi_{\mathcal{A}}(m) \geqslant m_A$ oraz $\Pi_{\mathcal{B}}(m) \geqslant m_B$, to ostatecznie dochodzimy do wniosku, że:

$$\Pi_{\mathcal{C}}(m) \leqslant m_A m_B \leqslant \Pi_{\mathcal{A}}(m)\Pi_{\mathcal{B}}(m)$$

□

Ograniczenie dla klasy $\mathcal{C}_k$

W rozwiązaniu skorzystamy z Lematu Sauer'a, według którego jeśli dla pewnej klasy pojęć $\mathcal{C}$ zachodzi $VC(\mathcal{C}) = d$, to dla $m \geqslant d$:

$$\Pi_{\mathcal{C}}(m) \leqslant \sum_{i=0}^{d} \binom{m}{i} \leqslant \left(\frac{em}{d}\right)^d$$

W naszym przypaku bierzemy $m = 2dklog_2(3k)$ i otrzymujemy:

$$\Pi_{\mathcal{C}_\|}(m) \leqslant \prod_{i=1}^{k} \Pi_{\mathcal{C}}(m) \leqslant \left(\frac{em}{d}\right)^{dk} = (2eklog_2(3k))^{dk}$$

Chcemy pokazać, że $\Pi_{\mathcal{C}_\|}(m) < 2^m$, z czego będzie bezpośrednio wynikało $VC(\mathcal{C}_k) \leqslant m$. Przekształcamy równoważnie tezę:

$$\Pi_{\mathcal{C}_\|}(2dklog_2(3k)) < 2^{2dklog_2(3k)}$$

$$(2eklog_2(3k))^{dk} < 2^{2dklog_2(3k)}$$

$$dklog_2(2eklog_2(3k)) < 2dklog_2(3k)$$

$$log_2(2eklog_2(3k)) < log_2((3k)^2)$$

$$2eklog_2(3k) < 9k^2$$

$$2elog_2(3k) < 9k$$

Aby wykazać ostatnią nierówność obliczamy pochodną funkcji $f(k) = 9k - 2elog_2(3k)$:

$$\frac{df}{dk} = 9 - \frac{2e}{3kln(2)}3 = 9 - \left(\frac{2e}{ln(2)}\right)\frac{1}{k}$$

Widzimy, że pochodna rośnie wraz ze wzrostem k, najmniejszą wartość przyjmuje dla $k = 1$, jest ona równa $9 - \left(\frac{2e}{ln(2)}\right) \approx 1.15 > 0$, zatem funkcja jest rosnąca na interesującym nas przedziale. Pozostaje sprawdzić, czy dla $k = 1$ zachodzi $f(k) > 0$: $f(1) = 9 - 2elog_2(3) \approx 0.38 > 0$. Tym samym udowodniliśmy nierówność z której, jak wcześniej wspomniałem, wynika teza postawiona w zadaniu. $\square$