

Zaawansowane uczenie maszynowe: ćwiczenia do wykładu 2

Paweł Cichosz

x	a_1	a_2	a_3	c
1	1	1	2	0
2	1	2	3	0
3	1	3	4	0
4	2	1	1	0
5	2	2	2	0
6	1	1	1	1
7	1	2	1	1
8	1	3	4	1
9	2	2	3	1
10	2	3	2	1

1. Traktując atrybuty a_1, a_2, a_3 jako dyskretne, zastosować wygładzanie Laplace'a do prawdopodobieństwa wartości 3 dla każdego z nich (przyjmując np. $l = 1$).
2. Traktując atrybuty a_1, a_2, a_3 jako dyskretne, wskazać atrybut o minimalnej entropii.
3. Traktując atrybuty a_1, a_2, a_3 jako dyskretne, wskazać ten z nich, który jest najsilniej związany z pojęciem docelowym c według informacji wzajemnej.
4. Wykazać równość:

$$\sum_{v_1 \in A_1} \sum_{v_2 \in A_2} P_S(a_1 = v_1, a_2 = v_2) \log \frac{P_S(a_1 = v_1, a_2 = v_2)}{P_S(a_1 = v_1) \cdot P_S(a_2 = v_2)} = E_S(a_1) - E_S(a_1|a_2)$$

Wskazówka:

$$\begin{aligned} P_S(a_1 = v_1, a_2 = v_2) &= P_S(a_1 = v_1|a_2 = v_2)P_S(a_2 = v_2) \\ P_S(a_1 = v_1|a_2 = v_2) &= P_{S_{a_2=v_2}}(a_1 = v_1) \end{aligned}$$

5. Dla dziedziny X , opisywanej przez atrybuty nominalne $a_1 : X \mapsto \{1, 2\}$, $a_2 : X \mapsto \{1, 2\}$ i $a_3 : X \mapsto \{1, 2, 3, 4\}$, zbioru klas $C = \{0, 1\}$ oraz przestrzeni modeli reprezentowanych przez dowolne funkcje $\{1, 2\} \times \{1, 2\} \times \{1, 2, 3, 4\} \rightarrow \{0, 1\}$, oszacować liczbę przykładów trenujących wystarczającą, aby (dla ustalonych δ i ϵ) z prawdopodobieństwem co najmniej $1 - \delta$ uzyskać model o błędzie rzeczywistym nieprzekraczającym ϵ przy użyciu spójnego algorytmu uczenia się, posługując się rozmiarem przestrzeni modeli oraz jej wymiarem VC.
6. Dla dziedziny X , opisywanej przez atrybuty nominalne $a_1 : X \mapsto \{1, 2\}$, $a_2 : X \mapsto \{1, 2\}$ i $a_3 : X \mapsto \{1, 2, 3, 4\}$, zbioru klas $C = \{0, 1\}$ oraz przestrzeni modeli reprezentowanych przez drzewa decyzyjne zawierające dokładnie dwa poziomy węzłów z nierównościami warunkami podziału ($a \leq v$), oszacować liczbę przykładów trenujących wystarczającą, aby (dla ustalonych δ i ϵ) z prawdopodobieństwem co najmniej $1 - \delta$ uzyskać model o błędzie rzeczywistym nieprzekraczającym ϵ przy użyciu spójnego algorytmu uczenia się, posługując się rozmiarem przestrzeni modeli oraz jej wymiarem VC.
7. Używając wyłącznie atrybutów a_1 i a_2 traktowanych jako dyskretne, zbudować drzewo decyzyjne, w którym:
 - stosowane są binarne podziały równościowe wybierane według entropii,
 - liście tworzone są po osiągnięciu jednolitej klasy, mniej niż 3 przykładów lub wyczerpaniu możliwości podziału,
 oraz wyznaczyć jego błąd na zbiorze trenującym.
8. Używając wyłącznie atrybutu a_3 traktowanego jako ciągły, zbudować drzewo decyzyjne, w którym:
 - stosowane są binarne podziały nierównościami wybierane według entropii,
 - liście tworzone są po osiągnięciu jednolitej klasy, mniej niż 3 przykładów lub wyczerpaniu możliwości podziału,
 oraz wyznaczyć jego błąd na zbiorze trenującym.