Zestaw 2 - Zadanie 1

Metody probabilistyczne w uczeniu maszynowym

Łukasz Trzos

Treść zadania

Mamy dwie urny. Pierwsza z nich zawiera 10 kul czerwonych, 20 białych i 70 zielonych, druga natomiast 40 kul czerwonych, 40 białych i 20 zielonych. Złodziej ukradł nam jedną z urn: z prawdopodobieństwem λ pierwszą z nich, z $1 - \lambda$ drugą. Następnie losowo wyciągnął z ukradzionej urny jedną kulę. Na podstawie jej koloru chce określić, którą z urn ukradł. Jeśli się pomyli, boss szajki każe go zlinczować. Funkcja straty (z punktu widzenia złodzieja) zadana jest wzorem

$$\ell(a, y) = \begin{cases} 0 \text{ batów}, & \text{gdy } a = y, \\ 10 \text{ batów}, & \text{gdy } a = 1, y = 2, \\ 6 \text{ batów}, & \text{gdy } a = 2, y = 1. \end{cases}$$

Choć złodziej swoim zachowaniem zasługuje na naganę, to - brzydząc się przemocą - chcemy mu pomóc. Jaką decyzję powinien podjąć w zależności od koloru kuli i parametru λ?

Rozwiązanie

Chcemy znaleźć funkcję $f : \{C, B, Z\} \to \{1, 2\}$, która zminimalizuje ryzyko określone jako:

$$R(f) = E[\ell(f(x), y)]$$

Możemy podzielić ten problem na 3 przypadki w zależności od wylosowanego koloru kuli, i wyznaczyć optymalny wybór uwzględniając parametr λ. Przykładowo, jeśli wylosowaliśmy kulę czerwoną, to oczekiwana strata wynosi:

$$E[\ell(f(x), y)|C] = \ell(f(x), 1)P(U_1, C) \ + \ \ell(f(x), 2)P(U_2, C)$$

gdzie $P(U_i, C)$ oznacza prawdopodobieństwo wylosowania i-tej urny, w przypadku otrzymania kuli czerwonej. Oczekiwana wyniesie więc:

- $\ell(1, 2)(1 - P(U_1, C))$, jeśli przyjmiemy $f(C) = 1$
- $\ell(2, 1)P(U_2|C)$, jeśli przyjmiemy $f(C) = 2$

Obliczamy $P(U_1, C)$ z tw. Bayesa:

$$P(U_1, C) = \frac{P(C|U_1)P(U_1)}{P(C)} = \frac{P(C|U_1)P(U_1)}{P(C|U_1)P(U_1) + P(C|U_2)P(U_2)} = \frac{\frac{1}{10}\lambda}{\frac{1}{10}\lambda + \frac{4}{10}(1 - \lambda)} = \frac{\lambda}{4 - 3\lambda}$$

Podstawiając, otrzymujemy, że oczekiwana strata wynosi:

- $10\frac{(4-4\lambda)}{4-3\lambda} = \frac{40-40\lambda}{4-3\lambda}$, jeśli przyjmiemy $f(C) = 1$
- $\frac{6\lambda}{4-3\lambda}$, jeśli przyjmiemy $f(C) = 2$

Wybieramy tę mniejszą wartość. $\frac{40-40\lambda}{4-3\lambda} > \frac{6\lambda}{4-3\lambda}$ dla $\lambda < \frac{40}{46}$. Zatem optymalnym wyborem jest:

$$f(C) = \begin{cases} 1, & \lambda > \frac{20}{23}, \\ 2, & wpp. \end{cases}$$

Prowadzimy identyczne obliczenia dla pozostałych kolorów kul, i uzyskujemy następujące wyniki:

$$f(B) = \begin{cases} 1, & \lambda > \frac{10}{13}, \\ 2, & wpp. \end{cases}$$

$$f(Z) = \begin{cases} 1, & \lambda > \frac{10}{81}, \\ 2, & wpp. \end{cases}$$