



AKADEMIA GÓRNICZO-HUTNICZA
IM. STANISŁAWA STASZICA W KRAKOWIE

ELEKTRONICZNE SYSTEMY DIAGNOSTYKI MEDYCZNEJ I
TERAPII

Klasyfikacja pulsu - Naiwny Bayes

Autorzy:

Piotr JANUS

Kamil PISZCZEK

1 Wstęp

Naiwny klasyfikator Bayesa jest prostym klasyfikatorem probabilistycznym opartym na twierdzeniu Bayesa. Nazywany jest naiwnym ze względu na przyjęte założenie, które mówi, że poszczególne cechy są wzajemnie niezależne. Pomimo tak dużego uproszczenia, klasyfikator wypadła niespodziewanie dobrze w wielu rzeczywistych problemach. Dużą zaletą tego klasyfikatora jest dobra skalowalność, metoda operuje jedynie na jawnych wzorach w przeciwieństwie do innych metod wykorzystujących podejście iteracyjne.

2 Algorytm

2.1 Twierdzenie Bayesa

Twierdzenie w teorii prawdopodobieństwa określające zależność między prawdopodobieństwem warunkowym wystąpienia zdarzeń $A|B$ i $B|A$. Przyjmijmy zbiór zdarzeń X , w którym zdarzenia $B_i \in X$, $P(B_i) > 0$ ($i = 1, 2, \dots, n$) tworzą układ zupełny (iloczyn każdych dwóch zdarzeń jest zdarzeniem niemożliwym, natomiast suma wszystkich zdarzeń jest zdarzeniem pewnym). Wówczas dla dowolnego $A \in X$ zachodzi następująca zależność:

$$P(B_i|A) = \frac{P(A|B_i)P(B_i)}{P(A)} \quad (1)$$

Wykorzystując dodatkowo wzór na prawdopodobieństwo całkowite, powyższa zależność może zostać przekształcona do następującej postaci:

$$P(B_i|A) = \frac{P(A|B_i)P(B_i)}{\sum_{k=1}^n P(A|B_k)P(B_k)} \quad (2)$$

2.2 Model probabilistyczny

Zdefiniujmy k -elementowy zbiór klas $C = \{C_1, C_2, \dots, C_k\}$ oraz dane do klasyfikacji opisane jako wektor $x = \{x_1, x_2, \dots, x_n\}$ zawierający n niezależnych cech. Prawdopodobieństwo przynależności do danej klasy może zostać zapisane z wykorzystaniem twierdzenia Bayesa (rozdział 2.1):

$$P(C_i|x) = \frac{P(x|C_i)P(C_i)}{P(x)} \quad (3)$$

Prawdopodobieństwo $P(x)$ występujące w mianowniku wzoru (3) nie zależy od C i jest stałe, licznik może natomiast zostać przekształcony poprzez wykorzystanie definicji prawdopodobieństwa warunkowego:

$$P(x_1, \dots, x_n|C_i)P(C_i) = P(x_1, \dots, x_n, C_i) \quad (4)$$

$$P(x_1, \dots, x_n, C_i) = P(x_1|x_2, \dots, x_n, C_i)P(x_2|x_3, \dots, x_n, C_i) \dots P(x_{n-1}|x_n, C_i)P(x_n|C_i)P(C_i) \quad (5)$$

Wykorzystując przyjęte na początku założenie, że cechy x_1, \dots, x_n są niezależne można wyprowadzić następującą zależność:

$$P(x_j|x_{j+1}, \dots, x_n, C_i) = P(x_j|C_i) \quad \text{dla } j = \{1, 2, \dots, n-1\} \quad (6)$$

Podstawiając (6) do równania (5) otrzymujemy:

$$P(x_1, \dots, x_n, C_i) = P(C_i) \prod_{j=1}^n P(x_j|C_i) \quad (7)$$

Ostatecznie wzór (3) można zapisać w postaci:

$$P(C_i|x) = \frac{P(C_i) \prod_{j=1}^n P(x_j|C_i)}{P(x)} \quad (8)$$