

Zad6 - Klasyfikator bayesowski

1. Zbiór uczący.

| Dreszcze | Katar | Ból głowy | Gorączka | Grypa? |
|----------|-------|-----------|----------|--------|
| Tak | Nie | Średni | Tak | Nie |
| Tak | Tak | Nie | Nie | Tak |
| Tak | Nie | Duży | Tak | Tak |
| Nie | Tak | Średni | Tak | Tak |
| Nie | Nie | Nie | Nie | Nie |
| Nie | Tak | Duży | Tak | Tak |
| Nie | Tak | Duży | Nie | Nie |
| Tak | Tak | Średni | Tak | Tak |

2. Klasyfikator bayesowski – opis.

- Klasyfikator bayesowski jest prostym probabilistycznym klasyfikatorem.
- Zakłada on, że zmienne są niezależne od siebie.
- Opiera się on na twierdzeniu Bayesa.

Twierdzenie Bayesa:

$$P(C|X) = \frac{P(X|C) * P(C)}{P(X)}$$

Powyższy wzór pozwala wyznaczyć prawdopodobieństwo zdarzenia w którym klasa C opisuje zdarzenie X.

$P(C)$ – oznacza prawdopodobieństwo a-priori wystąpienia klasy C.

$P(X|C)$ – oznacza prawdopodobieństwo a-posteriori, że zdarzenie X należy do klasy C.

$P(X)$ – oznacza prawdopodobieństwo a-priori wystąpienia przykładu X.

W przypadku n-atrybutów, które opisują zdarzenie $X=\{x_1, x_2, \dots, x_n\}$, wartość $P(X|C)$ jest wyrażona wzorem:

$$P(X|C) = \prod_{i=1}^n p(x_i | C)$$

3. Obliczenia.

3.1. Prawdopodobieństwo wystąpienia konkretnej klasy C.

| C | P(C) |
|-------------|------------|
| Grypa = Tak | =5/8=0,625 |
| Grypa = Nie | =3/8=0,375 |

3.2. Prawdopodobieństwo przynależności poszczególnych atrybutów zdarzenia X do klasy C.

| X C | P(X C) |
|----------------------------|-------------|
| Dreszcze=Tak Grypa=Tak | 0,6 |
| Dreszcze=Tak Grypa=Nie | 0,333333333 |
| Dreszcze=Nie Grypa=Tak | 0,4 |
| Dreszcze=Nie Grypa=Nie | 0,666666667 |
| Katar=Tak Grypa=Tak | 0,8 |
| Katar=Tak Grypa=Nie | 0,333333333 |
| Katar=Nie Grypa=Tak | 0,2 |
| Katar=Nie Grypa=Nie | 0,666666667 |
| Ból głowy=Nie Grypa=Tak | 0,2 |
| Ból głowy=średni Grypa=Tak | 0,4 |
| Ból głowy=duży Grypa=Tak | 0,4 |
| Ból głowy=Nie Grypa=Nie | 0,333333333 |
| Ból głowy=średni Grypa=Nie | 0,333333333 |
| Ból głowy=duży Grypa=Nie | 0,333333333 |
| Gorączka=Tak Grypa=Tak | 0,8 |
| Gorączka=Tak Grypa=Nie | 0,333333333 |
| Gorączka=Nie Grypa=Tak | 0,2 |
| Gorączka=Nie Grypa=Nie | 0,666666667 |

4. Klasyfikacja.

Zdarzenie do zaklasyfikowania:

| Dreszcze | Katar | Ból głowy | Gorączka | Grypa? |
|----------|-------|-----------|----------|--------|
| Tak | Nie | średni | Tak | ? |

Przynależność zdarzenia X do klasy Grypa=Tak wyraża się wzorem:

$$P(X|Grypa = Tak) = p(Dreszcze = Tak | Grypa = Tak) * \\ p(Katar = Nie | Grypa = Tak) * \\ p(Ból\ głowy = średni|Grypa = Tak) * \\ p(Gorączka = Tak | Grypa = Tak)$$

Zgodnie z powyższym wzorem otrzymujemy:

$$P(X|Grypa = Tak) = 0,6 * 0,2 * 0,4 * 0,8 = 0,0384$$

Przynależność zdarzenia X do klasy Grypa=Nie wyraża się wzorem:

$$P(X|Grypa = Nie) = p(Dreszcze = Tak | Grypa = Nie) * \\ p(Katar = Nie | Grypa = Nie) * \\ p(Ból głowy = średni | Grypa = Nie) * \\ p(Gorączka = Tak | Grypa = Nie)$$

Zgodnie z powyższym wzorem otrzymujemy:

$$P(X|Grypa = Nie) = 0,33 * 0,66 * 0,33 * 0,33 = 0,0237$$

Posiadając powyższe obliczenia policzmy $P(Grypa=Tak|X)$ oraz $P(Grypa=Nie|X)$:

$$P(Grypa = Tak|X) = P(X|Grypa = Tak) * P(Grypa = Tak)/P(X) \\ P(Grypa = Nie|X) = P(X|Grypa = Nie) * P(Grypa = Nie)/P(X)$$

Dodatkowo $P(X)$ wyraża się wzorem:

$$P(X) = P(X|Grypa = Tak) * P(Grypa = Tak) + \\ P(X|Grypa = Nie) * P(Grypa = Nie)$$

Podstawiając do wzorów otrzymujemy:

$$P(X|Grypa = Tak) * P(Grypa = Tak) = 0,0384 * 0,625 = 0,024$$

$$P(X|Grypa = Nie) * P(Grypa = Nie) = 0,0237 * 0,375 = 0,00889$$

$$P(X) = 0,024 + 0,00889 = 0,03289$$

$$P(Grypa = Tak|X) = \frac{0,024}{0,03289} = 0,7297$$

$$P(Grypa = Nie|X) = \frac{0,00889}{0,03289} = 0,2703$$

$P(Grypa = Tak|X) > P(Grypa = Nie|X)$ zatem zdarzenie klasyfikujemy do klasy Grypa=Tak.