Nome: José Elias Cangombe

Matrícula: 799

Intelência Artificial e Machine Learning

Lista 4

Exercício 1 - Neste exercício você utilizará o teorema de Bayes. Considere dois exames médicos, A e B, para um vírus. O teste A é 95% eficaz no reconhecimento do vírus quando ele está presente, mas tem uma taxa de falso positivo de 10% (indicando que o vírus está presente, quando ele não está). O teste B é 90% eficaz no reconhecimento do vírus, mas possui uma taxa de falso positivo de 5%. Os dois testes usam métodos independentes para identificar o vírus. 1% de todas as pessoas possuem o vírus. Digamos que uma pessoa é testada para o vírus usando apenas um dos testes e que o teste é positivo para o vírus. Qual teste, retornando positivo, é mais indicativo de alguém realmente estar com o vírus?

Teste A
Dados

$P_A(+|com\_virus) = 0.95$

$P_A(virus) = 0.01$

$P_A(+|sem\_virus) = 0.1$

$P_A(sem\_virus) = 1 - 0.01 = 0.99$

Fórmula

$P_A(+) = P_A(+|com\_virus) * P_A(virus) + P_A(+|sem_{virus}) * P_A(sem\_virus)$

$P_A(+) = 0.95 * 0.01 + 0.1 * 0.99 = 0.1085$

Fórmula

$$P_A(+|com_{virus}) = \frac{P_A(+|com_{virus}) * P_A(virus)}{P_A(+)}$$

$$P_A(+|com_{virus}) = \frac{0.95 * 0.01}{0.1085} = 0.0875$$

Fórmula

$$P_A(+|sem_{virus}) = \frac{P_A(+|sem_{virus}) * P_A(sem_{virus})}{P_A(+)}$$

$$P_A(+|sem_{virus}) = \frac{0.1 * 0.99}{0.1085} = 0.0912$$

Teste B

Dados

$P_B(+|com\_virus) = 0.9$

$P_B(virus) = 0.01$

$P_B(+|sem\_virus) = 0.05$

$P_B(sem\_virus) = 1 - 0.01 = 0.99$

Fórmula

$$P_B(+) = P_B(+|com\_virus) * P_B(virus) + P_B(+|sem_{virus}) * P_B(sem\_virus)$$

$$P_B(+) = 0.9 * 0.01 + 0.05 * 0.99 = 0.0585$$

Fórmula

$$P_B(+|com_{virus}) = \frac{P_B(+|com_{virus}) * P_B(virus)}{P_B(+)}$$

$$P_B(+|com_{virus}) = \frac{0.9 * 0.01}{0.0585} = 0{,}15386$$

Fórmula

$$P_B(+|sem_{virus}) = \frac{P_B(+|sem_{virus}) * P_B(sem_{virus})}{P_B(+)}$$

$$= \frac{0.05 * 0.99}{0.0585} = 0.84615$$

R: Segundo a probabilidade dos dois resultados apresentados nos dois testes. O teste B é mais indicativo de alguém realmente estar com o vírus com relação ao teste A.

Exercicio 3- Neste exercício você vai prever, baseado em alguns atributos físicos de uma pessoa, se ela é do sexo masculino ou feminino.

Dados os seguintes atributos físicos de uma pessoa:
Altura = 1.83 metros,
Peso = 58.97
Quilos e tamanho do calçado = 20.32 centímetros.

Baseado nas informações anteriores, qual classe tem maior probabilidade, ou seja, qual dos 2 sexos teria a maior probabilidade?

Para calcular as probabilidades, utilize os dados da tabela abaixo. **OBS**.: Apresente todos os cálculos feitos para se encontrar as probabilidades de cada classe, ou seja, neste exercício você não deve utilizar a biblioteca SciKit-learn.

(**Dica**: Assuma que as probabilidades condicionais dos atributos seguem uma distribuição Gaussiana).
(**Dica**: Assuma que a probabilidade da pessoa ser do sexo masculino ou do feminino é de 0.5, respectivamente).
(**Dica**: utilize a teoria do classificador naive Bayes e lembre-se que o numerador da equação do classificador não influencia na maximização das probabilidades).

$$P_{masculino} = \frac{4}{8} = 0.5$$

$$P_{feminino} = \frac{4}{8} = 0.5$$

$$P_{masculino} = P_{feminino}$$

1° CÁLCULO DE ALTURA

a)

Fórmula

$$\mu = \frac{1}{N}\sum_{i=1}^{N} x_i$$

$$\mu_{altura\_masculino} = \frac{\sum \mu\_(altura\_masculino)}{4}$$

$$\mu_{altura\ masculino} = \frac{1.83 + 1.80 + 1.70 + 2 * 1.80}{4} = 1.7825$$

Fórmula

$$\sigma = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(x_i - \mu)^2}$$

$$\sigma_{altura\ masculino} = \sqrt{\frac{(1.83 - 1.7825)^2 + (1.80 - 1.7825)^2 + (1.70 - 1.7825)^2 + (1.80 - 1.7825)^2}{4}}$$

$$\sigma_{altura\ masculino} = 0.043946$$

b) Cálculo Probabilístico da Altura Masculina

Fórmula

$$P(x_k|C_q) = \frac{1}{\sigma_{x_k C_q}^2 \sqrt{2\pi}} e^{-(x_k - \mu_{x_k C_q})^2 / 2\sigma_{\mu_{x_k C_q}}^2}$$

$$P(1.83|masculino) = \frac{1}{0.043946^2\sqrt{2\pi}} e^{-(1.83-1.7825)^2 / 2*(0.043946)^2} = 115{,}18122$$

c)

Para Altura Femenino

$$\mu_{altura\ feminino} = \frac{\sum \mu\_(altura\ feminino)}{4}$$

$$\mu_{altura\ feminino} = \frac{1.52 + 1.68 + 1.65 + 1.75}{4} = 1.65$$

$$\sigma_{altura\ feminino} = \sqrt{\frac{(1.52-1.65)^2 + (1.68-1.65)^2 + (1.65-1.65)^2 + (1.75-1.65)^2}{4}}$$

$$\sigma_{altura\ feminino} = 0.083367$$

d)

Cálculo Probabilístico da Altura Femenino

Fórmula

$$P(x_k|C_q) = \frac{1}{\sigma_{x_k C_q}^2 \sqrt{2\pi}} e^{-(x_k - \mu_{x_k C_q})^2 / 2\sigma_{\mu_{x_k C_q}}^2}$$

$$P(1.83|feminino) = \frac{1}{0.083367^2\sqrt{2\pi}} e^{-(1.83-1.65)^2 / 2*(0.083367)^2} = 5{,}57978748$$

2° CÁLCULO DE PESO

a)

Para o Peso Masculino Temos:

<u>Fórmula</u>

Fórmula

$$\mu = \frac{1}{N}\sum_{i=1}^{N} x_i$$

$$\mu_{peso\ masculino} = \frac{\sum \mu\_(peso\ masculino)}{4}$$

$$\mu_{peso\ masculino} = \frac{81.65 + 86.18 + 77.11 + 74.84}{4} = 79.945$$

$$\sigma = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(x_i - \mu)^2}$$

$$\sigma_{peso\ masculino} = \sqrt{\frac{(81.65 - 79.945)^2 + (86.18 - 79.945)^2 + (77.11 - 79.945)^2 + (74.84 - 79.945)^2}{4}}$$

$$\sigma_{peso\ masculino} = 4.35547$$

b) Cálculo Probabilístico do Peso Masculina

Fórmula

$$P(x_k|C_q) = \frac{1}{\sigma^2_{x_k C_q}\sqrt{2\pi}} e^{-\left(x_k - \mu_{x_k C_q}\right)^2 / 2\sigma^2_{\mu_{x_k C_q}}}$$

$$P(58.97|masculino) = \frac{1}{4.35547^2\sqrt{2\pi}} e^{-(58.97-79.945)^2 / 2*(4.35547)^2} = 1{,}935580585 * 10^{-7}$$

c)

Para Peso Femenino

Fórmula

$$\mu_{peso\ feminino} = \frac{\sum \mu\_(peso\ feminino)}{4}$$

$$\mu_{peso\ feminino} = \frac{45.36 + 2 * 68.04 + 58.97}{4} = 60.1025$$

$$\sigma_{peso\ feminino} = \sqrt{\frac{2 * (68.04 - 60.1025)^2 + (45.36 - 60.1025)^2 + +(58.97 - 60.1025)^2}{4}}$$

$$\sigma_{peso\ feminino} = 9.28212899$$

d)

Cálculo Probabilístico da Altura Femenino

Fórmula

$$P(x_k|C_q) = \frac{1}{\sigma^2_{x_k C_q}\sqrt{2\pi}} e^{-(x_k - \mu_{x_k C_q})^2 / 2\sigma^2_{\mu_{x_k C_q}}}$$

$$P(58.97|feminino) = \frac{1}{9.28212899^2\sqrt{2\pi}} e^{-(58.97-60.1025)^2 / 2*(9.28212899)^2} = 0{,}0045960251$$

3° CÁLCULO DE CALÇADOS

a)

Para o Género MasculinoTemos:

Fórmula

$$\mu = \frac{1}{N}\sum_{i=1}^{N} x_i$$

$$\mu_{calçado\ masculino} = \frac{2 * 30.48 + 27.94 + 25.40}{4} = 28.575$$

Fórmula

$$\sigma = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(x_i - \mu)^2}$$

$$\sigma_{calçado\ masculino} = \sqrt{\frac{2 * (30.48 - 28.575)^2 + (27.94 - 28.575)^2 + (30.48 - 28.575)^2 + (25.40 - 28.575)^2}{4}}$$

$$\sigma_{calçado\ masculino} = 2.11$$

b) Cálculo Probabilístico Para o Tamanho dos Calçados

Para o Género Masculino Temos:

Fórmula

$$P(x_k|C_q) = \frac{1}{\sigma^2_{x_k C_q}\sqrt{2\pi}} e^{-\left(x_k - \mu_{x_k C_q}\right)^2 / 2\sigma^2_{\mu_{x_k C_q}}}$$

$$P(20.32|masculino) = \frac{1}{2{,}11^2\sqrt{2\pi}} e^{-(20.32-28.575)^2 / 2*(2.11)^2} = 42{,}52377 * 10^{-6}$$

c) Para o Género FemeninoTemos:

Fórmula

$$\mu_{calçado\ feminino} = \frac{\sum \mu\_(calçado\ femenino)}{4}$$

$$\mu_{calçado\ feminino} = \frac{15.24 + 20.32 + 17.78 + 22.86}{4} = 19.05$$

$$\sigma_{calçado\ feminino} = \sqrt{\frac{(15.24-19.05)^2+(20.32-19.05)^2+(17.78-19.05)^2+(22.86-19.05)^2}{4}} = 2.84$$

d) Cálculo Probabilístico Para o Tamanho dos Calçados

Fórmula

$$P(x_k|C_q) = \frac{1}{\sigma^2_{x_k C_q}\sqrt{2\pi}} e^{-(x_k-\mu_{x_k C_q})^2 / 2\sigma^2_{\mu_{x_k C_q}}}$$

$$P(20.32|feminino) = \frac{1}{2{,}84^2\sqrt{2\pi}} e^{-(20.32-19.05)^2 / 2*(2{,}84)^2} = 4{,}47558 * 10^{-6}$$

3° Calculando as Probabilidades dos dois (2) Géneros e Comparando os resultados

Para o Género Masculino

$P_{MASCULINO}$=(Altura, Peso, Tamanho do Calçados)

$P_{MASCULINO}$=$115{,}18122 * 1{,}935580585 * 10^{-7} * 42{,}52377 * 10^{-6} * 0{,}5 = 474{,}018 * 10^{-12}$

Para o Género Femenino

$P_{FEMENINO}$=(Altura, Peso, Tamanho do Calçados)

$P_{FEMENINO}$=$5{,}579787484 * 0{,}0045960251 * 4{,}47558 * 10^{-6} * 0{,}5 = 57{,}38777 * 10^{-9}$

R: Com base nos dois resultados posso dizer que o género femenino apresenta maior probabilidade com relação ao masculino.