

MBA⁺

**Artificial Intelligence &
Machine Learning**





Naive Bayes

Classificador Bayesiano



- O teorema de Bayes permite o cálculo da posteriori de h
- MAP (Maximum a Posteriori) Hipótese Máxima a Posteriori
 - $P(h|D) = (P(D|h) P(h)) / P(D)$
 - Por que $P(h|D)$ deve decrescer quando $P(D)$ cresce?
 - Resposta: quanto mais provável D , menos evidência ele é para h

Classificador Bayesiano



- Considere:
 - $P(h)$: probabilidade da hipótese h (ser verdadeira) – priori
 - O que fazer quando não se conhece $P(h)$?
 - $P(D)$: probabilidade do conjunto de dados D ser observado
 - $P(x|y)$: probabilidade de x dado y
 - $P(D|h)$: probabilidade do conjunto de dados D ser observado dado que h é verdadeira
 - $P(h|D)$: probabilidade de h ser verdadeira dado o conjunto de dados D
- $P(h|D)$ é o foco quando se procura a hipótese mais provável
 - É a probabilidade a posteriori de h

Classificador Bayesiano



- Se calcularmos a posteriori de todas as hipóteses $h \in H$, podemos encontrar a hipótese (h) mais provável

$$\begin{aligned} h_{MAP} &\equiv \operatorname{argmax}_{h \in H} P(h|D) \\ &= \operatorname{argmax}_{h \in H} \frac{P(D|h)P(h)}{P(D)} \\ &= \operatorname{argmax}_{h \in H} P(D|h)P(h) \end{aligned}$$

- Como D é constante para todas as hipóteses, $P(D)$ pode ser retirada do cálculo



Classificador Bayesiano

- Quando todas as hipóteses são equiprováveis a priori, então a verossimilhança (likelihood) é quem domina o cálculo

- Máxima a Posteriori

Máxima Verossimilhança

$$\begin{aligned}h_{MAP} &\equiv \operatorname{argmax}_{h \in H} P(h|D) \\&= \operatorname{argmax}_{h \in H} \frac{P(D|h) P(h)}{P(D)} \\&= \operatorname{argmax}_{h \in H} P(D|h) P(h)\end{aligned}$$

$$h_{ML} \equiv \operatorname{argmax}_{h \in H} P(D|h)$$

Classificador Bayesiano



- Se o conhecimento a priori é uniforme (equiprovável), então, $h_{ML} = h_{MAP}$
- Nem sempre a hipótese de máxima probabilidade a posteriori (h_{MAP}) traz a melhor classificação.

Classificador Bayesiano



- Exemplo: considere 4 hipóteses h_1 , h_2 , h_3 e h_4 com probabilidades 0.3, 0.25, 0.2 e 0.25 respectivamente. Considere ainda uma nova instância (x) é classificada como negativa por h_1 (h_{MAP}) e como positiva por h_2 , h_3 e h_4 . Neste caso, a classificação mais provável é diferente da hipótese mais provável.

Classificador Bayesiano



- O CBO é computacionalmente caro
 - Probabilidade posterior de todas as hipóteses em H
 - Combinação da predição de cada h para classificar cada nova instância.

Classificador Bayesiano



- Pode-se minimizar este custo através de algumas técnicas que tornam o classificador não-ótimo, mas gerando bons resultados a um custo computacional aceitável.

- Naive Bayes Classifier

Classificador Naive Bayes



- Um classificador muito simples (ingênuo?)
- Performance surpreendente em vários domínios
- Muito utilizado
- Cada instância de treinamento é considerada uma conjunção dos valores dos atributos
- Pode-se dizer que não é necessária a construção de um modelo (mesmo sabendo-se que o modelo é gerado)

Classificador Naive Bayes



- O NB, simplifica o cálculo de $P(a_1, a_2, \dots, a_n | v_j)$ assumindo que todos os atributos são condicionalmente independentes entre si dado o valor da classe.
- Independência dos atributos condicionada à classe:

$$P(a_1, a_2 \dots a_n | v_j) = \prod_i P(a_i | v_j)$$

Classificador Naive Bayes



- Os termos $P(v_j)$ e $P(a_i|v_j)$ são calculados com base na frequência com que ocorrem no conjunto de instâncias de treinamento (contagem)
- O número de probabilidades a serem estimadas é linear em relação ao número de atributos (e não mais exponencial)

Classificador Naive Bayes



- Quando o conjunto de instâncias que descrevem o problema respeita a suposição de que todos os atributos são condicionalmente independentes dada a classe, o NB é um CBO.



Classificador Naive Bayes

Exemplo: pelo teorema de Bayes, podemos calcular a probabilidade de uma nova instância x ser classificada como 0, e a probabilidade de x ser classificada como 1.

Instâncias	Classe	A1	A2	A3	A4
1	1	0	1	1	1
2	1	0	1	1	1
3	0	1	0	0	0
4	1	0	1	1	1
5	1	0	1	1	1
6	0	0	0	0	1
7	0	1	1	0	0
8	1	0	1	1	1
9	0	1	0	0	0
10	0	1	1	0	0

$$P(A|B) = \frac{P(B|A) P(A)}{P(B)}$$



Classificador Naive Bayes

Exemplo: pelo teorema de Bayes, podemos calcular a probabilidade de uma nova instância x ser classificada como 0, e a probabilidade de x ser classificada como 1.

Na sequência, podemos identificar qual classe é a mais provável. Ou seja, o que temos que calcular é:

$P(\text{Classe}=0|A1,A2,A3,A4)$ e

$P(\text{Classe}=1|A1,A2,A3,A4)$,

usando o **teorema de Bayes** e assumindo a independência condicional entre as variáveis.

Instâncias	Classe	A1	A2	A3	A4
1	1	0	1	1	1
2	1	0	1	1	1
3	0	1	0	0	0
4	1	0	1	1	1
5	1	0	1	1	1
6	0	0	0	0	1
7	0	1	1	0	0
8	1	0	1	1	1
9	0	1	0	0	0
10	0	1	1	0	0

$$P(A|B) = \frac{P(B|A) P(A)}{P(B)}$$

Classificador Naive Bayes



Considere por exemplo, que a partir do conjunto de instâncias de treinamento dado, deseja-se classificar uma nova instância (11):

Instâncias	Classe	A1	A2	A3	A4
1	1	0	1	1	1
2	1	0	1	1	1
3	0	1	0	0	0
4	1	0	1	1	1
5	1	0	1	1	1
6	0	0	0	0	1
7	0	1	1	0	0
8	1	0	1	1	1
9	0	1	0	0	0
10	0	1	1	0	0

$$P(A|B) = \frac{P(B|A) P(A)}{P(B)}$$

Classificador Naive Bayes



Considere por exemplo, que a partir do conjunto de instâncias de treinamento dado, deseja-se classificar uma nova instância (11)

$P(\text{Classe}=0 | A1=0, A2=1, A3=0, A4=1) =$
 $P(A1=0, A2=1, A3=0, A4=1 | \text{Classe}=0) * P(\text{Classe}=0)$

Instâncias	Classe	A1	A2	A3	A4
1	1	0	1	1	1
2	1	0	1	1	1
3	0	1	0	0	0
4	1	0	1	1	1
5	1	0	1	1	1
6	0	0	0	0	1
7	0	1	1	0	0
8	1	0	1	1	1
9	0	1	0	0	0
10	0	1	1	0	0
11	?	0	1	0	1

$$P(A|B) = \frac{P(B|A) P(A)}{P(B)}$$

Classificador Naive Bayes



Considere por exemplo, que a partir do conjunto de instâncias de treinamento dado, deseja-se classificar uma nova instância (11)

$$P(\text{Classe}=0 | A1=0, A2=1, A3=0, A4=1) = \\ P(A1=0, A2=1, A3=0, A4=1 | \text{Classe}=0) * P(\text{Classe}=0)$$

Como assumimos a independência condicional, temos que:

$$P(A1=0, A2=1, A3=0, A4=1 | \text{Classe}=0) = \prod P(A_i=x_i | \text{Classe}=0)$$

Instâncias	Classe	A1	A2	A3	A4
1	1	0	1	1	1
2	1	0	1	1	1
3	0	1	0	0	0
4	1	0	1	1	1
5	1	0	1	1	1
6	0	0	0	0	1
7	0	1	1	0	0
8	1	0	1	1	1
9	0	1	0	0	0
10	0	1	1	0	0
11	?	0	1	0	1

$$P(A|B) = \frac{P(B|A) P(A)}{P(B)}$$

Classificador Naive Bayes



Considere por exemplo, que a partir do conjunto de instâncias de treinamento dado, deseja-se classificar uma nova instância (11

$$P(\text{Classe}=0 | A1=0, A2=1, A3=0, A4=1) = \\ P(A1=0, A2=1, A3=0, A4=1 | \text{Classe}=0) * P(\text{Classe}=0)$$

Como assumimos a independência condicional, temos que:

$$P(A1=0, A2=1, A3=0, A4=1 | \text{Classe}=0) = \prod P(A_i=x_i | \text{Classe}=0)$$

Assim, temos:

$$P(\text{Classe}=0 | A1=0, A2=1, A3=0, A4=1) = \\ (\prod P(A_i=x_i | \text{Classe}=0)) * P(\text{Classe}=0)$$

Instâncias	Classe	A1	A2	A3	A4
1	1	0	1	1	1
2	1	0	1	1	1
3	0	1	0	0	0
4	1	0	1	1	1
5	1	0	1	1	1
6	0	0	0	0	1
7	0	1	1	0	0
8	1	0	1	1	1
9	0	1	0	0	0
10	0	1	1	0	0
11	?	0	1	0	1

Classificador Naive Bayes



Considere por exemplo, que a partir do conjunto de instâncias de treinamento dado, deseja-se classificar uma nova instância (11):

Instâncias	Classe	A1	A2	A3	A4
1	1	0	1	1	1
2	1	0	1	1	1
3	0	1	0	0	0
4	1	0	1	1	1
5	1	0	1	1	1
6	0	0	0	0	1
7	0	1	1	0	0
8	1	0	1	1	1
9	0	1	0	0	0
10	0	1	1	0	0
11	?	0	1	0	1

$$P(\text{Classe}=0 | A1=0, A2=1, A3=0, A4=1) =$$

$$(\prod P(A_i=x_i | \text{Classe}=0)) * P(\text{Classe}=0) =$$

$$(P(A1=0 | C=0) * P(A2=1 | C=0) * P(A3=0 | C=0) * P(A4=1 | C=0)) * P(\text{Classe} = 0)$$

Classificador Naive Bayes



Considere por exemplo, que a partir do conjunto de instâncias de treinamento dado, deseja-se classificar uma nova instância (11):

$$P(\text{Classe}=0 | A1=0, A2=1, A3=0, A4=1) =$$

$$(\prod P(A_i=x_i | \text{Classe}=0)) * P(\text{Classe}=0) =$$

$$(P(A1=0 | C=0) * P(A2=1 | C=0) * P(A3=0 | C=0) * P(A4=1 | C=0)) * P(\text{Classe} = 0)$$

Instâncias	Classe	A1	A2	A3	A4
1	1	0	1	1	1
2	1	0	1	1	1
3	0	1	0	0	0
4	1	0	1	1	1
5	1	0	1	1	1
6	0	0	0	0	1
7	0	1	1	0	0
8	1	0	1	1	1
9	0	1	0	0	0
10	0	1	1	0	0
11	?	0	1	0	1

Classificador Naive Bayes



Considere por exemplo, que a partir do conjunto de instâncias de treinamento dado, deseja-se classificar uma nova instância (11):

Instâncias	Classe	A1	A2	A3	A4
1	1	0	1	1	1
2	1	0	1	1	1
3	0	1	0	0	0
4	1	0	1	1	1
5	1	0	1	1	1
6	0	0	0	0	1
7	0	1	1	0	0
8	1	0	1	1	1
9	0	1	0	0	0
10	0	1	1	0	0
11	?	0	1	0	1

$$\begin{aligned} P(\text{Classe}=0 | A1=0, A2=1, A3=0, A4=1) = \\ (\prod P(A_i=x_i | \text{Classe}=0)) * P(\text{Classe}=0) = \\ (P(A1=0 | C=0) * P(A2=1 | C=0) * P(A3=0 | C=0) * P(A4=1 | C=0)) * P(\text{Classe} = 0) \\ (1/5) * \end{aligned}$$

Classificador Naive Bayes



Considere por exemplo, que a partir do conjunto de instâncias de treinamento dado, deseja-se classificar uma nova instância (11):

Instâncias	Classe	A1	A2	A3	A4
1	1	0	1	1	1
2	1	0	1	1	1
3	0	1	0	0	0
4	1	0	1	1	1
5	1	0	1	1	1
6	0	0	0	0	1
7	0	1	1	0	0
8	1	0	1	1	1
9	0	1	0	0	0
10	0	1	1	0	0
11	?	0	1	0	1

$$P(\text{Classe}=0 | A1=0, A2=1, A3=0, A4=1) =$$

$$(\prod P(A_i=x_i | \text{Classe}=0)) * P(\text{Classe}=0)=$$

$$(P(A1=0|C=0) * P(A2=1|C=0) * P(A3=0|C=0) * P(A4=1|C=0)) * P(\text{Classe}=0) \\ (1/5) *$$

Classificador Naive Bayes



Considere por exemplo, que a partir do conjunto de instâncias de treinamento dado, deseja-se classificar uma nova instância (11):

Instâncias	Classe	A1	A2	A3	A4
1	1	0	1	1	1
2	1	0	1	1	1
3	0	1	0	0	0
4	1	0	1	1	1
5	1	0	1	1	1
6	0	0	0	0	1
7	0	1	1	0	0
8	1	0	1	1	1
9	0	1	0	0	0
10	0	1	1	0	0
11	?	0	1	0	1

$$P(\text{Classe}=0 | A1=0, A2=1, A3=0, A4=1) = (\prod P(A_i=x_i | \text{Classe}=0)) * P(\text{Classe}=0) =$$

$$(P(A1=0 | C=0) * P(A2=1 | C=0) * P(A3=0 | C=0) * P(A4=1 | C=0)) * P(\text{Classe} = 0) \\ (1/5) *$$

Classificador Naive Bayes



Considere por exemplo, que a partir do conjunto de instâncias de treinamento dado, deseja-se classificar uma nova instância (11):

Instâncias	Classe	A1	A2	A3	A4
1	1	0	1	1	1
2	1	0	1	1	1
3	0	1	0	0	0
4	1	0	1	1	1
5	1	0	1	1	1
6	0	0	0	0	1
7	0	1	1	0	0
8	1	0	1	1	1
9	0	1	0	0	0
10	0	1	1	0	0
11	?	0	1	0	1

$$P(\text{Classe}=0 | A1=0, A2=1, A3=0, A4=1) = (\prod P(A_i=x_i | \text{Classe}=0)) * P(\text{Classe}=0) =$$

$$(P(A1=0 | C=0) * P(A2=1 | C=0) * P(A3=0 | C=0) * P(A4=1 | C=0)) * P(\text{Classe} = 0)$$
$$(1/5) * (2/5)$$

Classificador Naive Bayes



Considere por exemplo, que a partir do conjunto de instâncias de treinamento dado, deseja-se classificar uma nova instância (11):

Instâncias	Classe	A1	A2	A3	A4
1	1	0	1	1	1
2	1	0	1	1	1
3	0	1	0	0	0
4	1	0	1	1	1
5	1	0	1	1	1
6	0	0	0	0	1
7	0	1	1	0	0
8	1	0	1	1	1
9	0	1	0	0	0
10	0	1	1	0	0
11	?	0	1	0	1

$$P(\text{Classe}=0 | A1=0, A2=1, A3=0, A4=1) = (\prod P(A_i=x_i | \text{Classe}=0)) * P(\text{Classe}=0) =$$

$$(P(A1=0 | C=0) * P(A2=1 | C=0) * P(A3=0 | C=0) * P(A4=1 | C=0)) * P(\text{Classe} = 0)$$
$$(1/5) * (2/5)$$

Classificador Naive Bayes



Considere por exemplo, que a partir do conjunto de instâncias de treinamento dado, deseja-se classificar uma nova instância (11):

Instâncias	Classe	A1	A2	A3	A4
1	1	0	1	1	1
2	1	0	1	1	1
3	0	1	0	0	0
4	1	0	1	1	1
5	1	0	1	1	1
6	0	0	0	0	1
7	0	1	1	0	0
8	1	0	1	1	1
9	0	1	0	0	0
10	0	1	1	0	0
11	?	0	1	0	1

$$P(\text{Classe}=0 | A1=0, A2=1, A3=0, A4=1) =$$

$$(\prod P(A_i=x_i | \text{Classe}=0)) * P(\text{Classe}=0) =$$

$$(P(A1=0 | C=0) * P(A2=1 | C=0) * P(A3=0 | C=0) * P(A4=1 | C=0)) * P(\text{Classe} = 0)$$

$$(1/5) * (2/5)$$

Classificador Naive Bayes



Considere por exemplo, que a partir do conjunto de instâncias de treinamento dado, deseja-se classificar uma nova instância (11):

Instâncias	Classe	A1	A2	A3	A4
1	1	0	1	1	1
2	1	0	1	1	1
3	0	1	0	0	0
4	1	0	1	1	1
5	1	0	1	1	1
6	0	0	0	0	1
7	0	1	1	0	0
8	1	0	1	1	1
9	0	1	0	0	0
10	0	1	1	0	0
11	?	0	1	0	1

$$P(\text{Classe}=0 | A1=0, A2=1, A3=0, A4=1) =$$

$$(\prod P(A_i=x_i | \text{Classe}=0)) * P(\text{Classe}=0) =$$

$$(P(A1=0 | C=0) * P(A2=1 | C=0) * P(A3=0 | C=0) * P(A4=1 | C=0)) * P(\text{Classe}=0)$$

$$(1/5) * (2/5) * (5/5) *$$

Classificador Naive Bayes



Considere por exemplo, que a partir do conjunto de instâncias de treinamento dado, deseja-se classificar uma nova instância (11):

Instâncias	Classe	A1	A2	A3	A4
1	1	0	1	1	1
2	1	0	1	1	1
3	0	1	0	0	0
4	1	0	1	1	1
5	1	0	1	1	1
6	0	0	0	0	1
7	0	1	1	0	0
8	1	0	1	1	1
9	0	1	0	0	0
10	0	1	1	0	0
11	?	0	1	0	1

$$\begin{aligned}
 &P(\text{Classe}=0 | A1=0, A2=1, A3=0, A4=1) = \\
 &(\prod P(A_i=x_i | \text{Classe}=0)) * P(\text{Classe}=0) = \\
 &(P(A1=0 | C=0) * P(A2=1 | C=0) * P(A3=0 | C=0) * P(A4=1 | C=0)) * P(\text{Classe}=0) \\
 &(1/5) * (2/5) * (5/5) *
 \end{aligned}$$

Classificador Naive Bayes



Considere por exemplo, que a partir do conjunto de instâncias de treinamento dado, deseja-se classificar uma nova instância (11):

Instâncias	Classe	A1	A2	A3	A4
1	1	0	1	1	1
2	1	0	1	1	1
3	0	1	0	0	0
4	1	0	1	1	1
5	1	0	1	1	1
6	0	0	0	0	1
7	0	1	1	0	0
8	1	0	1	1	1
9	0	1	0	0	0
10	0	1	1	0	0
11	?	0	1	0	1

$$\begin{aligned} P(\text{Classe}=0 | A1=0, A2=1, A3=0, A4=1) &= \\ (\prod P(A_i=x_i | \text{Classe}=0)) * P(\text{Classe}=0) &= \\ (P(A1=0 | C=0) * P(A2=1 | C=0) * P(A3=0 | C=0) * P(A4=1 | C=0)) * P(\text{Classe}=0) \\ (1/5) * (2/5) * (5/5) * (1/5) \end{aligned}$$

Classificador Naive Bayes



Considere por exemplo, que a partir do conjunto de instâncias de treinamento dado, deseja-se classificar uma nova instância (11):

$P(\text{Classe}=0 | A1=0, A2=1, A3=0, A4=1) =$

$(\prod P(A_i=x_i | \text{Classe}=0)) * P(\text{Classe}=0) =$

$(P(A1=0|C=0) * P(A2=1|C=0) * P(A3=0|C=0) * P(A4=1|C=0)) * P(\text{Classe}=0)$

$(1/5) * (2/5) * (5/5) * (1/5) * (5/10)$

Instâncias	Classe	A1	A2	A3	A4
1	1	0	1	1	1
2	1	0	1	1	1
3	0	1	0	0	0
4	1	0	1	1	1
5	1	0	1	1	1
6	0	0	0	0	1
7	0	1	1	0	0
8	1	0	1	1	1
9	0	1	0	0	0
10	0	1	1	0	0
11	?	0	1	0	1

Classificador Naive Bayes



Considere por exemplo, que a partir do conjunto de instâncias de treinamento dado, deseja-se classificar uma nova instância (11):

Instâncias	Classe	A1	A2	A3	A4
1	1	0	1	1	1
2	1	0	1	1	1
3	0	1	0	0	0
4	1	0	1	1	1
5	1	0	1	1	1
6	0	0	0	0	1
7	0	1	1	0	0
8	1	0	1	1	1
9	0	1	0	0	0
10	0	1	1	0	0
11	?	0	1	0	1

$$\begin{aligned} P(\text{Classe}=0 | A1=0, A2=1, A3=0, A4=1) &= \\ (\prod P(A_i=x_i | \text{Classe}=0)) * P(\text{Classe}=0) &= \\ (P(A1=0 | C=0) * P(A2=1 | C=0) * P(A3=0 | C=0) * P(A4=1 | C=0)) * P(\text{Classe}=0) &= \\ (1/5) * (2/5) * (5/5) * (1/5) * (5/10) &= \end{aligned}$$

Classificador Naive Bayes



Considere por exemplo, que a partir do conjunto de instâncias de treinamento dado, deseja-se classificar uma nova instância (11):

Instâncias	Classe	A1	A2	A3	A4
1	1	0	1	1	1
2	1	0	1	1	1
3	0	1	0	0	0
4	1	0	1	1	1
5	1	0	1	1	1
6	0	0	0	0	1
7	0	1	1	0	0
8	1	0	1	1	1
9	0	1	0	0	0
10	0	1	1	0	0
11	?	0	1	0	1

$$\begin{aligned} P(\text{Classe}=0 | A1=0, A2=1, A3=0, A4=1) &= \\ (\prod P(A_i=x_i | \text{Classe}=0)) * P(\text{Classe}=0) &= \\ (P(A1=0 | C=0) * P(A2=1 | C=0) * P(A3=0 | C=0) * P(A4=1 | C=0)) * P(\text{Classe}=0) &= \\ (1/5) * (2/5) * (5/5) * (1/5) * (5/10) &= \\ 0.2 * 0.4 * 1 * 0.2 * 0.5 &= 0.008 \end{aligned}$$

Classificador Naive Bayes



Considere por exemplo, que a partir do conjunto de instâncias de treinamento dado, deseja-se classificar uma nova instância (11):

Instâncias	Classe	A1	A2	A3	A4
1	1	0	1	1	1
2	1	0	1	1	1
3	0	1	0	0	0
4	1	0	1	1	1
5	1	0	1	1	1
6	0	0	0	0	1
7	0	1	1	0	0
8	1	0	1	1	1
9	0	1	0	0	0
10	0	1	1	0	0
11	?	0	1	0	1

$$\begin{aligned} P(\text{Classe}=0 | A1=0, A2=1, A3=0, A4=1) &= \\ (\prod P(A_i=x_i | \text{Classe}=0)) * P(\text{Classe}=0) &= \\ (P(A1=0 | C=0) * P(A2=1 | C=0) * P(A3=0 | C=0) * P(A4=1 | C=0)) * P(\text{Classe}=0) &= \\ (1/5) * (2/5) * (5/5) * (1/5) * (5/10) &= \\ 0.2 * 0.4 * 1 * 0.2 * 0.5 &= 0.008 \end{aligned}$$

Agora fazemos o mesmo processo para a classe=1:



Classificador Naive Bayes

Considere por exemplo, que a partir do conjunto de instâncias de treinamento dado, deseja-se classificar uma nova instância (11):

Instâncias	Classe	A1	A2	A3	A4
1	1	0	1	1	1
2	1	0	1	1	1
3	0	1	0	0	0
4	1	0	1	1	1
5	1	0	1	1	1
6	0	0	0	0	1
7	0	1	1	0	0
8	1	0	1	1	1
9	0	1	0	0	0
10	0	1	1	0	0
11	?	0	1	0	1

$$\begin{aligned}
 &P(\text{Classe}=0|A1=0, A2=1, A3=0, A4=1) = \\
 &(\prod P(A_i=x_i | \text{Classe}=0)) * P(\text{Classe}=0)= \\
 &(P(A1=0|C=0) * P(A2=1|C=0) * P(A3=0|C=0) * P(A4=1|C=0)) * P(\text{Classe}=0) \\
 &(1/5) * (2/5) * (5/5) * (1/5) * (5/10) = \\
 &0.2 * 0.4 * 1 * 0.2 * 0.5 = 0.008
 \end{aligned}$$

Agora fazemos o mesmo processo para a classe=1:

$$\begin{aligned}
 &P(\text{Classe}=1|A1=0, A2=1, A3=0, A4=1) = \\
 &(\prod P(A_i=x_i | \text{Classe}=1)) * P(\text{Classe}=1)=
 \end{aligned}$$

Classificador Naive Bayes



Considere por exemplo, que a partir do conjunto de instâncias de treinamento dado, deseja-se classificar uma nova instância (11):

Instâncias	Classe	A1	A2	A3	A4
1	1	0	1	1	1
2	1	0	1	1	1
3	0	1	0	0	0
4	1	0	1	1	1
5	1	0	1	1	1
6	0	0	0	0	1
7	0	1	1	0	0
8	1	0	1	1	1
9	0	1	0	0	0
10	0	1	1	0	0
11	?	0	1	0	1

$$\begin{aligned}
 &P(\text{Classe}=0|A1=0, A2=1, A3=0, A4=1) = \\
 &(\prod P(A_i=x_i | \text{Classe}=0)) * P(\text{Classe}=0)= \\
 &(P(A1=0|C=0) * P(A2=1|C=0) * P(A3=0|C=0) * P(A4=1|C=0)) * P(\text{Classe} = 0) \\
 &(1/5) * (2/5) * (5/5) * (1/5) * (5/10) = \\
 &0.2 * 0.4 * 1 * 0.2 * 0.5 = 0.008
 \end{aligned}$$

Agora fazemos o mesmo processo para a classe=1:

$$\begin{aligned}
 &P(\text{Classe}=1|A1=0, A2=1, A3=0, A4=1) = \\
 &(\prod P(A_i=x_i | \text{Classe}=1)) * P(\text{Classe}=1)= \\
 &(P(A1=0|C=1) * P(A2=1|C=1) * P(A3=0|C=1) * P(A4=1|C=1)) * P(\text{Classe} = 1)
 \end{aligned}$$



Classificador Naive Bayes

Considere por exemplo, que a partir do conjunto de instâncias de treinamento dado, deseja-se classificar uma nova instância (11):

Instâncias	Classe	A1	A2	A3	A4
1	1	0	1	1	1
2	1	0	1	1	1
3	0	1	0	0	0
4	1	0	1	1	1
5	1	0	1	1	1
6	0	0	0	0	1
7	0	1	1	0	0
8	1	0	1	1	1
9	0	1	0	0	0
10	0	1	1	0	0
11	?	0	1	0	1

$$\begin{aligned}
 &P(\text{Classe}=0|A1=0, A2=1, A3=0, A4=1) = \\
 &(\prod P(A_i=x_i | \text{Classe}=0)) * P(\text{Classe}=0)= \\
 &(P(A1=0|C=0) * P(A2=1|C=0) * P(A3=0|C=0) * P(A4=1|C=0)) * P(\text{Classe} = 0) \\
 &(1/5) * (2/5) * (5/5) * (1/5) * (5/10) = \\
 &0.2 * 0.4 * 1 * 0.2 * 0.5 = 0.008
 \end{aligned}$$

Agora fazemos o mesmo processo para a classe=1:

$$\begin{aligned}
 &P(\text{Classe}=1|A1=0, A2=1, A3=0, A4=1) = \\
 &(\prod P(A_i=x_i | \text{Classe}=1)) * P(\text{Classe}=1)= \\
 &(P(A1=0|C=1) * P(A2=1|C=1) * P(A3=0|C=1) * P(A4=1|C=1)) * P(\text{Classe} = 1) \\
 &(5/5) * (5/5) * (0/5) * (5/5) * (5/10) =
 \end{aligned}$$



Classificador Naive Bayes

Considere por exemplo, que a partir do conjunto de instâncias de treinamento dado, deseja-se classificar uma nova instância (11):

Instâncias	Classe	A1	A2	A3	A4
1	1	0	1	1	1
2	1	0	1	1	1
3	0	1	0	0	0
4	1	0	1	1	1
5	1	0	1	1	1
6	0	0	0	0	1
7	0	1	1	0	0
8	1	0	1	1	1
9	0	1	0	0	0
10	0	1	1	0	0
11	?	0	1	0	1

$$\begin{aligned}
 &P(\text{Classe}=0|A1=0, A2=1, A3=0, A4=1) = \\
 &(\prod P(A_i=x_i | \text{Classe}=0)) * P(\text{Classe}=0) = \\
 &(P(A1=0|C=0) * P(A2=1|C=0) * P(A3=0|C=0) * P(A4=1|C=0)) * P(\text{Classe} = 0) \\
 &(1/5) * (2/5) * (5/5) * (1/5) * (5/10) = \\
 &0.2 * 0.4 * 1 * 0.2 * 0.5 = 0.008
 \end{aligned}$$

Agora fazemos o mesmo processo para a classe=1:

$$\begin{aligned}
 &P(\text{Classe}=1|A1=0, A2=1, A3=0, A4=1) = \\
 &(\prod P(A_i=x_i | \text{Classe}=1)) * P(\text{Classe}=1) = \\
 &(P(A1=0|C=1) * P(A2=1|C=1) * P(A3=0|C=1) * P(A4=1|C=1)) * P(\text{Classe} = 1) \\
 &(5/5) * (5/5) * (0/5) * (5/5) * (5/10) = \\
 &1 * 1 * 0 * 1 * 0.5 = 0
 \end{aligned}$$

Classificador Naive Bayes



Considere por exemplo, que a partir do conjunto de instâncias de treinamento dado, deseja-se classificar uma nova instância (11):

Instâncias	Classe	A1	A2	A3	A4
1	1	0	1	1	1
2	1	0	1	1	1
3	0	1	0	0	0
4	1	0	1	1	1
5	1	0	1	1	1
6	0	0	0	0	1
7	0	1	1	0	0
8	1	0	1	1	1
9	0	1	0	0	0
10	0	1	1	0	0
11	?	0	1	0	1

$$P(\text{Classe}=0|A1=0, A2=1, A3=0, A4=1)=0.008$$

$$P(\text{Classe}=1|A1=0, A2=1, A3=0, A4=1)=\mathbf{0}$$

Normalizando, temos:

$$P(\text{Classe}=0|A1=0, A2=1, A3=0, A4=1) = 0.008/(0.008+0) = 1$$

$$P(\text{Classe}=1|A1=0, A2=1, A3=0, A4=1) = \mathbf{0}/(0.008+0) = 0$$



Classificador Naive Bayes

Considere por exemplo, que a partir do conjunto de instâncias de treinamento dado, deseja-se classificar uma nova instância (11):

Instâncias	Classe	A1	A2	A3	A4
1	1	0	1	1	1
2	1	0	1	1	1
3	0	1	0	0	0
4	1	0	1	1	1
5	1	0	1	1	1
6	0	0	0	0	1
7	0	1	1	0	0
8	1	0	1	1	1
9	0	1	0	0	0
10	0	1	1	0	0
11	?	0	1	0	1

$$P(\text{Classe}=0|A1=0, A2=1, A3=0, A4=1)=0.008$$

$$P(\text{Classe}=1|A1=0, A2=1, A3=0, A4=1)=\mathbf{0}$$

Normalizando, temos:

$$P(\text{Classe}=0|A1=0, A2=1, A3=0, A4=1) = 0.008/(0.008+0) = 1$$

$$P(\text{Classe}=1|A1=0, A2=1, A3=0, A4=1) = \mathbf{0}/(0.008+0) = 0$$



Classificador Naive Bayes

Solução de Engenharia

Instâncias	Classe	A1	A2	A3	A4
1	1	0	1	1	1
2	1	0	1	1	1
3	0	1	0	0	0
4	1	0	1	1	1
5	1	0	1	1	1
6	0	0	0	0	1
7	0	1	1	0	0
8	1	0	1	1	1
9	0	1	0	0	0
10	0	1	1	0	0
11	?	0	1	0	1

$$\begin{aligned}
 &P(\text{Classe}=0|A1=0, A2=1, A3=0, A4=1) = \\
 &(\prod P(A_i=x_i | \text{Classe}=0)) * P(\text{Classe}=0)= \\
 &(P(A1=0|C=0) * P(A2=1|C=0) * P(A3=0|C=0) * P(A4=1|C=0)) * P(\text{Classe} = 0) \\
 &(1/5) * (2/5) * (5/5) * (1/5) * (5/10) = \\
 &0.2 * 0.4 * 1 * 0.2 * 0.5 = 0.008
 \end{aligned}$$

Agora fazemos o mesmo processo para a classe=1:

$$\begin{aligned}
 &P(\text{Classe}=1|A1=0, A2=1, A3=0, A4=1) = \\
 &(\prod P(A_i=x_i | \text{Classe}=1)) * P(\text{Classe}=1)= \\
 &(P(A1=0|C=1) * P(A2=1|C=1) * P(A3=0|C=1) * P(A4=1|C=1)) * P(\text{Classe} = 1) \\
 &(5/5) * (5/5) * (0/5) * (5/5) * (5/10) = \\
 &1 * 1 * 0 * 1 * 0.5 = 0
 \end{aligned}$$



Classificador Naive Bayes

Solução de Engenharia

Instâncias	Classe	A1	A2	A3	A4
1	1	0	1	1	1
2	1	0	1	1	1
3	0	1	0	0	0
4	1	0	1	1	1
5	1	0	1	1	1
6	0	0	0	0	1
7	0	1	1	0	0
8	1	0	1	1	1
9	0	1	0	0	0
10	0	1	1	0	0
11	?	0	1	0	1

$$\begin{aligned}
 &P(\text{Classe}=0|A1=0, A2=1, A3=0, A4=1) = \\
 &(\prod P(A_i=x_i | \text{Classe}=0)) * P(\text{Classe}=0)= \\
 &(P(A1=0|C=0) * P(A2=1|C=0) * P(A3=0|C=0) * P(A4=1|C=0)) * P(\text{Classe} = 0) \\
 &(1/5) * (2/5) * (5/5) * (1/5) * (5/10) = \\
 &0.2 * 0.4 * 1 * 0.2 * 0.5 = 0.008
 \end{aligned}$$

Agora fazemos o mesmo processo para a classe=1:

$$\begin{aligned}
 &P(\text{Classe}=1|A1=0, A2=1, A3=0, A4=1) = \\
 &(\prod P(A_i=x_i | \text{Classe}=1)) * P(\text{Classe}=1)= \\
 &(P(A1=0|C=1) * P(A2=1|C=1) * P(A3=0|C=1) * P(A4=1|C=1)) * P(\text{Classe} = 1) \\
 &(5/5) * (5/5) * (0/5) * (5/5) * (5/10) = \\
 &1 * 1 * 0.00001 * 1 * 0.5 = 0.000005
 \end{aligned}$$

Classificador Naive Bayes



Solução de Engenharia

$$P(\text{Classe}=0|A1=0, A2=1, A3=0, A4=1)=0.008$$

$$P(\text{Classe}=1|A1=0, A2=1, A3=0, A4=1)=\mathbf{0.000005}$$

Normalizando, temos:

$$P(\text{Classe}=0|A1=0, A2=1, A3=0, A4=1) = 0.008 / (0.008 + \mathbf{0.000005}) = 0.99$$

$$P(\text{Classe}=1|A1=0, A2=1, A3=0, A4=1) = \mathbf{0.000005} / (0.008 + \mathbf{0.000005}) = 0.01$$

A classe mais provável é a classe 0.

Instâncias	Classe	A1	A2	A3	A4
1	1	0	1	1	1
2	1	0	1	1	1
3	0	1	0	0	0
4	1	0	1	1	1
5	1	0	1	1	1
6	0	0	0	0	1
7	0	1	1	0	0
8	1	0	1	1	1
9	0	1	0	0	0
10	0	1	1	0	0
11	?	0	1	0	1



Classificador Naive Bayes

Solução de Engenharia

Instâncias	Classe	A1	A2	A3	A4
1	1	0	1	1	1
2	1	0	1	1	1
3	0	1	0	0	0
	1	0	1	1	1
5	1	0	1	1	1
6	0	0	0	0	1
7	0	1	1	0	0
8	1	0	1	1	1
9	0	1	0	0	0
10	0	1	1	0	0
11	?	0	1	0	1

$$P(\text{Classe}=0|A1=0, A2=1, A3=0, A4=1) =$$

$$(\prod P(A_i=x_i | \text{Classe}=0)) * P(\text{Classe}=0)=$$

$$(P(A1=0|C=0) * P(A2=1|C=0) * P(A3=0|C=0) * P(A4=1|C=0)) * P(\text{Classe} = 0)$$

$$(1+0.5)/(5+1) * (2+0.5)/(5+1) * (5+0.5)/(5+1) * (1+0.5)/(5+1) *$$

$$(5/10) = 0.25 * 0.41 * 0.91 * 0.25 * 0.5 = 0.011$$

Agora fazemos o mesmo processo para a classe=1:

$$P(\text{Classe}=1|A1=0, A2=1, A3=0, A4=1) =$$

$$(\prod P(A_i=x_i | \text{Classe}=1)) * P(\text{Classe}=1)=$$

$$(P(A1=0|C=1) * P(A2=1|C=1) * P(A3=0|C=1) * P(A4=1|C=1)) * P(\text{Classe} = 1)$$

$$(5+0.5)/(5+1) * (5+0.5)/(5+1) * (0+0.5)/(5+1) * (5+0.5)/(5+1) *$$

$$(5/10) = 0.91 * 0.91 * \mathbf{0.09} * 0.91 * 0.5 = \mathbf{0.033}$$

Classificador Naive Bayes



Solução de Engenharia

$$P(\text{Classe}=0|A1=0, A2=1, A3=0, A4=1) = 0.011$$

$$P(\text{Classe}=1|A1=0, A2=1, A3=0, A4=1) = \mathbf{0.033}$$

Normalizando, temos:

$$P(\text{Classe}=0|A1=0, A2=1, A3=0, A4=1) = 0.011/(0.011+\mathbf{0.033}) = 0.25$$

$$P(\text{Classe}=1|A1=0, A2=1, A3=0, A4=1) = \mathbf{0.033}/(0.011+\mathbf{0.033}) = 0.75$$

A classe mais provável é a classe 1.

Instâncias	Classe	A1	A2	A3	A4
1	1	0	1	1	1
2	1	0	1	1	1
3	0	1	0	0	0
4	1	0	1	1	1
5	1	0	1	1	1
6	0	0	0	0	1
7	0	1	1	0	0
8	1	0	1	1	1
9	0	1	0	0	0
10	0	1	1	0	0
11	?	0	1	0	1



Rede Bayesiana

Solução de
Problemas de
impressão no
Windows 95

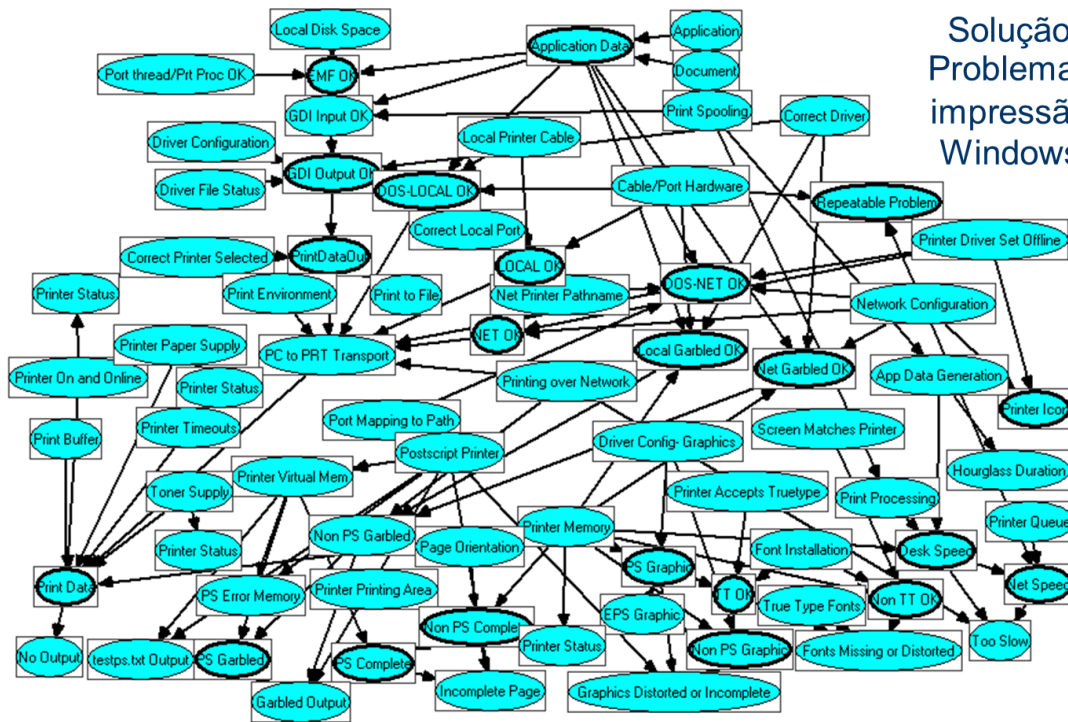


Figura gerada com o software MSBN – Microsoft Research (<http://research.microsoft.com>)

Questions and Feedback



[Thank you!](#)

Obrigado !



Vinicius Fernandes Caridá
vfcarida@gmail.com



@Vinicius Caridá



@Vinicius Caridá
@machine learning Brasil



@vfcarida

MBA⁺

Copyright © 2018 **Prof. Vinicius Fernandes Caridá**
 Todos direitos reservados. Reprodução ou divulgação
 total ou parcial deste documento é expressamente
 proibido sem o consentimento formal, por escrito, do
 Professor (autor).