

Avaliação de Desempenho

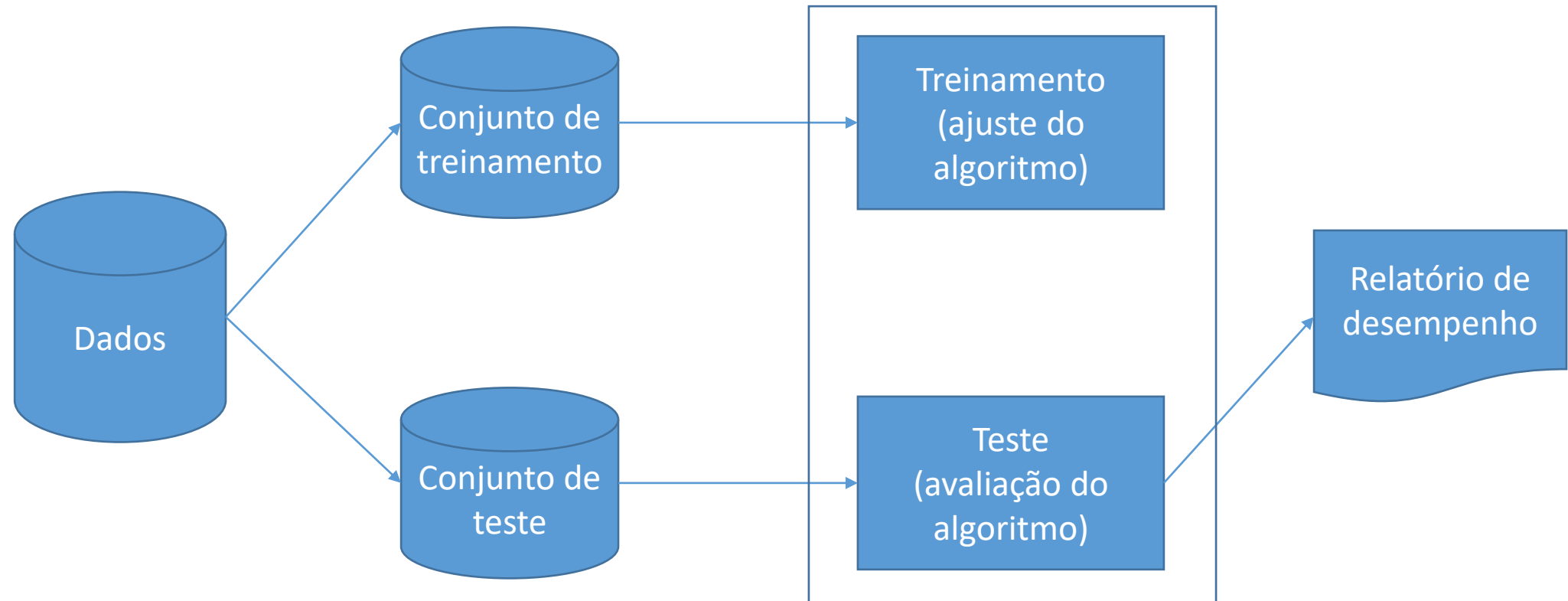
Classificador Bayesiano

Algoritmo de Classificação

- Algoritmo capaz de separar os exemplos observados em classes

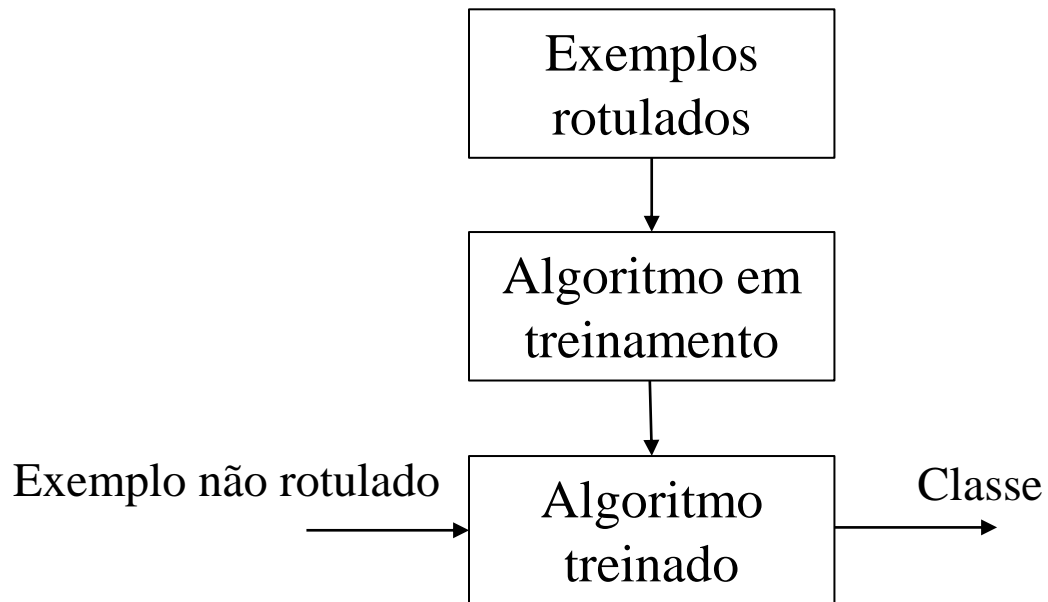


Algoritmo de classificação



Problema de Classificação

- Treinamento supervisionado



- Exemplos: valores dos atributos para os objetos observados
 - Palavras de um Email
- Rótulo: atributo que classifica objetos
 - Spam

Classificador Bayesiano

- Algoritmo capaz de separar exemplos observados em certo problema em um conjunto de classes com base na equação de Bayes

$$P[Y = y|X = x] = \frac{P[X = x|Y = y] \cdot P[Y = y]}{P[X = x]}$$

- Y classes
- $P[Y = y]$ distribuição de probabilidade das classes (Y)
- X atributos observados (evidências)
- $P[X = x]$ é a distribuição de probabilidade de X (dos atributos observados)
- $P[Y = y | X = x]$ é a distribuição condicional das classes (Y) dado os valores de atributos que foram observados
- $P[X=x | Y=y]$ é a distribuição condicional dos atributos dado que a classe é conhecida

Classificador Bayesiano

- Algoritmo de classificação: selecionar Y que maximiza a probabilidade condicional

$$h(X) = \operatorname{argmax}_{y \in \Omega_Y} P[Y = y | X = x]$$

- A partir da equação de Bayes

$$h(X) = \operatorname{argmax}_{y \in \Omega_Y} \frac{P[X = x | Y = y] \cdot P[Y = y]}{P[X = x]}$$

- Como $P[X = x]$ não contém a variável Y que está sendo maximizado

$$h(X) = \operatorname{argmax}_{y \in \Omega_Y} P[X = x | Y = y] \cdot P[Y = y]$$

Classificador Bayesiano

- Algoritmo de classificação

Selecionar a classe que maximize seguinte expressão:

$$P[X = x|Y = y] \cdot P[Y = y]$$

Classificador Bayesiano

- Assume-se a hipótese de independência entre as variáveis que representam os atributos (hipótese ingênua)
- Assume que a presença ou ausência de uma característica particular, não está relacionado com a presença ou ausência de qualquer outra característica
- Exemplo
 - Três atributos são usados para classificar um motor como defeituoso ou não: tipo, aquecimento, e rendimento
 - Os atributos são considerados independentes para cálculo da probabilidade de que um motor esteja com defeito

Classificador Bayesiano

- X representa a ocorrência simultânea dos seguintes eventos:

$$X=x_1, X=x_2, \dots, X=x_n$$

- $P[X = x|Y = y] = P[X_1 = x, X_2 = x, \dots, X_n = x|Y = y]$
- Pela hipótese de independência das variáveis de atributos
- $P[X = x|Y = y] = P[X_1 = x|Y = y] \cdot P[X_2 = x|Y = y] \dots P[X_n = x|Y = y]$

Classificador Bayesiano

- $P[X = x|Y = y] = P[X_1 = x|Y = y] \cdot P[X_2 = x|Y = y] \dots P[X_n = x|Y = y]$
- Cálculo de $P[X_i = x|Y = y]$
- Para X_i discreto
- $\#(X_i = x|Y = y)$ quantidade de amostras da classe y (no conjunto de treinamento) cujo valor do atributo X_i é igual a x
- $$P[X_i = x|Y = y] = \frac{\#(X_i = x|Y = y)}{\text{quantidade de amostras da classe } y}$$

Classificador Bayesiano

- $P[X = x|Y = y] = P[X_1 = x|Y = y] \cdot P[X_2 = x|Y = y] \dots P[X_n = x|Y = y]$
- Cálculo de $P[X_i = x|Y = y]$
- Para X_i contínuo
- $P[X_i = x|Y = y] = \text{normpdf}(z|y, 0, 1)$
- $\mu_i | y$ e $\sigma_i | y$ são a média e desvio padrão de X_i para as amostras da classe y
- $z|y = (x - \mu_i) / \sigma_i$

Exemplo 1

- Atributos de categoria

Nome	Identificador	Categorias	Codificação
Idade	X1	Adulto	0
		Idoso	1
		Jovem	2
Estudante	X3	Não	0
		Sim	1
Crédito	X4	Bom	0
		Excelente	1

Exemplo 1

- Atributo numérico

Nome	Identificador	Codificação
Renda	X2	Float

- Classes

Nome	Identificador	Categorias	Codificação
Compra	Y	Não	0
		Sim	1

Exemplo 1

Idade (X_1)	Renda (X_2)	Estudante (X_3)	Crédito (X_4)	Compra (Y)
Jovem (2)	6300	Não (0)	Bom (0)	Não (0)
Jovem (2)	5800	Não (0)	Bom (0)	Não (0)
Adulto (0)	6100	Não (0)	Bom (0)	Sim (1)
Idoso (1)	5200	Não (0)	Bom (0)	Sim (1)
Idoso (1)	2300	Sim (1)	Bom (0)	Sim (1)
Idoso (1)	2800	Sim (1)	Excelente (1)	Não (0)
Adulto (0)	3100	Sim (1)	Excelente (1)	Sim (1)
Jovem (2)	4500	Não (0)	Bom (0)	Não (0)
Jovem (2)	3300	Sim (1)	Bom (0)	Sim (1)
Idoso (1)	4800	Sim (1)	Bom (0)	Sim (1)
Jovem (2)	4400	Sim (1)	Excelente (1)	Sim (1)
Adulto (0)	4900	Não (0)	Excelente (1)	Sim (1)
Adulto (0)	6000	Sim (1)	Bom (0)	Sim (1)
Idoso (1)	4600	Não (0)	Excelente (1)	Não (0)

4 características ($n=4$)

Todas discretas

2 classes ($k = 2$)

14 amostras ($m = 14$)

Classe $Y = 0$

corresponde a não
comprar computador

Classe $Y = 1$

corresponde a comprar
computador

Queremos classificar
uma pessoa com idade
< 30, renda 4560,
estudante, crédito bom
 $X = (0, 4650, 0, 0)$

Exemplo 1

- Valor de Y que maximiza $P[X = x|Y = y] \cdot P[Y = y]$
- $P[Y = 0] = 9/14 = 0,6429$ $P[Y = 1] = 5/14 = 0,3571$
- Calcular $P[X = x|Y = y]$ para o objeto com atributos X , para as duas classes
- $X = (2, 4650, 1, 0)$
- $x_1 = 2$ $x_2 = 4650$ $x_3 = 1$ $x_4 = 0$

Exemplo 1

- Atributos discretos

$$P[X_1 = 2|Y = 1] = 2/9 = 0,2222$$

$$P[X_3 = 1|y = 1] = 6/9 = 0,6667$$

$$P[X_4 = 0|Y = 1] = 6/9 = 0,6667$$

$$P[X_1 = 2|Y = 0] = 3/5 = 0,6$$

$$P[X_3 = 1|Y = 0] = 1/5 = 0,2$$

$$P[X_4 = 0|y = 0] = 3/5 = 0,6$$

- Atributos contínuos

$$(\mu_2 | Y=1) = 4455,6 \quad (\mu_2 | Y=0) = 4800,0 \quad (\sigma_2 | Y=1) = 1237,5 \quad (\sigma_2 | Y=0) = 1214,9$$

$$P[X_2 = 4650|Y = 1] = \text{normpdf}(4650, 4455.6, 1237.5) = 0,394$$

$$P[X_2 = 4650|Y = 0] = \text{normpdf}(4650, 4800.0, 1214.9) = 0,396$$

Exemplo 1

$$P[X = x|Y = 0] = 0,22 \cdot 0,394 \cdot 0,667 \cdot 0,667 = 0,0386$$

$$P[X = x|Y = 1] = 0,6 \cdot 0,396 \cdot 0,2 \cdot 0,6 = 0,0285$$

$$P[X = x|Y = 0] \cdot P[Y = 0] = 0,0386 \cdot 0,6429 = 0,0248$$

$$P[X = x|Y = 1] \cdot P[Y = 1] = 0,0285 \cdot 0,3571 = 0,0102$$

O classificador prevê a classe $Y = 1$ (compra) para $X = (2, 4650, 1, 0)$

Avaliação de Desempenho

Classificador Bayesiano