



Classificador bayesiano (Naïve Bayes)

Agente Educacional

Sérgio M. Dias

Agenda

O classificador Naïve Bayes
Um exemplo didático
Teorema de Bayes



Classificador Naïve Bayes

O classificador Naïve Bayes é provavelmente **um dos classificadores mais utilizados** em aprendizado de máquina.

Também chamado “Classificador Bayesiano”, “Bayes Ingênuo”, “Bayes Simples”.

Trata-se de um classificador **estatístico**, baseado no **Teorema de Bayes**.



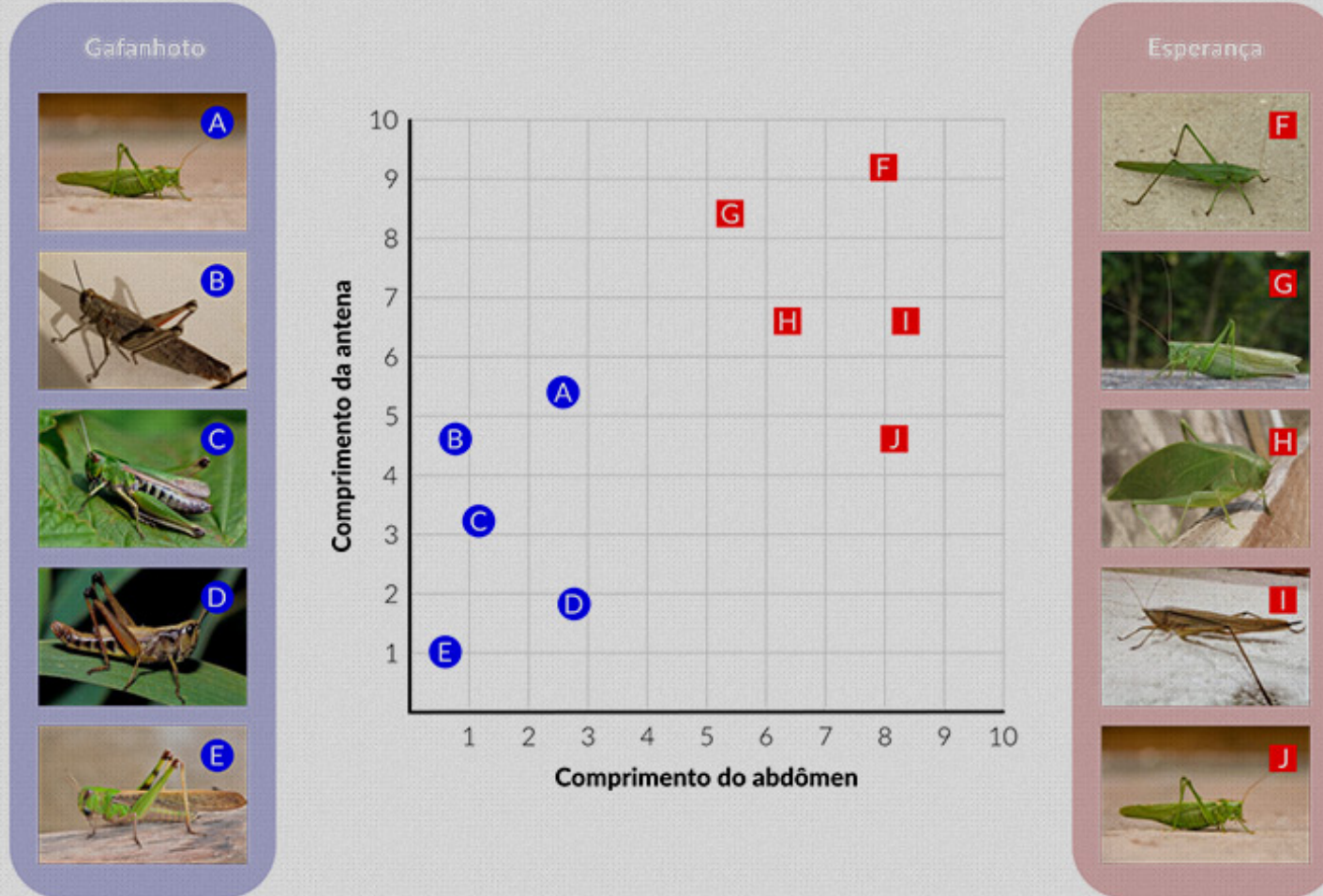
Classificador bayesiano (Naïve Bayes)

Recebe o nome de **ingênuo** (*naïve*)

- Presume que, se o valor de um atributo exerce algum efeito sobre a distribuição de classes existente no conjunto, esse efeito é independente dos valores assumidos por outros atributos e de seus respectivos efeitos sobre a mesma distribuição de classes.
- Em outras palavras, a informação de um evento não é informativa sobre nenhum outro (condicionalmente independentes).

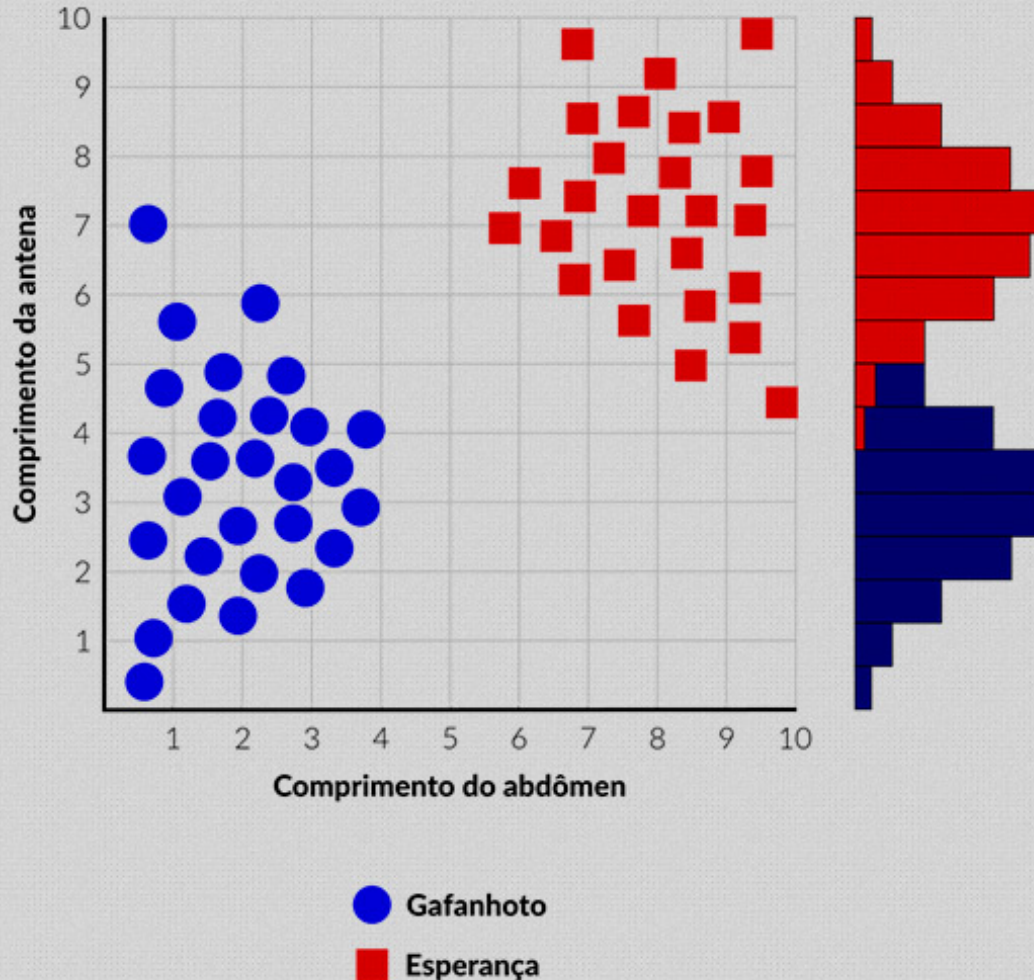
*“Apesar desta premissa “ingênua” e simplista,
o classificador reporta bom desempenho em várias
tarefas de classificação”*

Classificador bayesiano (Naïve Bayes)





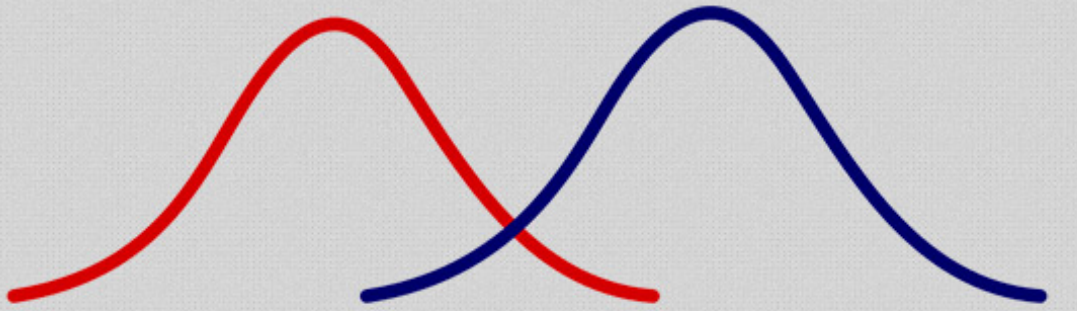
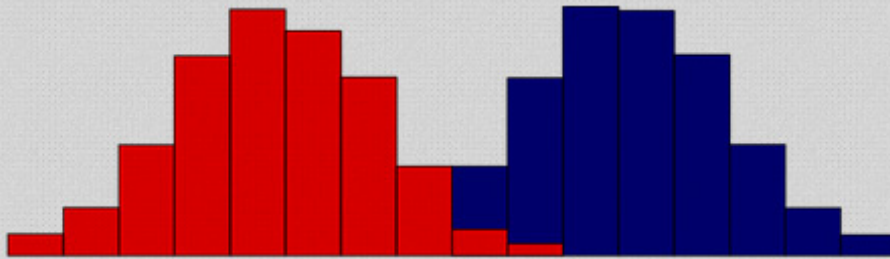
Classificador bayesiano (Naïve Bayes)



Juntando muitos dados podemos construir um **histograma**.

A princípio, vamos construir apenas um para “**Comprimento da antena**”.

Classificador bayesiano (Naïve Bayes)



Para facilitar a visualização, vamos **sumarizar** os histogramas com duas **distribuições normais**.

Nos próximos slides vamos utilizar as **distribuições normais**.

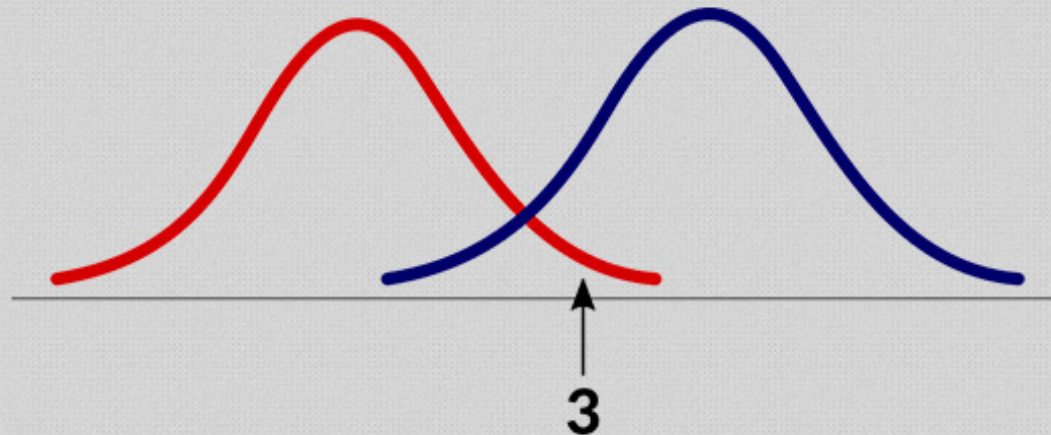
Classificador bayesiano (Naïve Bayes)

Queremos classificar um inseto que encontramos. Sua antena tem **3 unidades** de comprimento. **Como classificá-lo?**

Podemos apenas nos perguntar, **dada a distribuição de comprimento** de antenas que vimos, é mais **provável** que nosso inseto seja um “**Gafanhoto**” ou um “**Esperança**”?



Comprimento da antena = 3



Há uma maneira formal de discutir a classificação mais “provável”

$$p(c_j | d)$$

(probabilidade da classe c_j , dado que observamos d)

Classificador bayesiano (Naïve Bayes)

$$p(c_j | d)$$

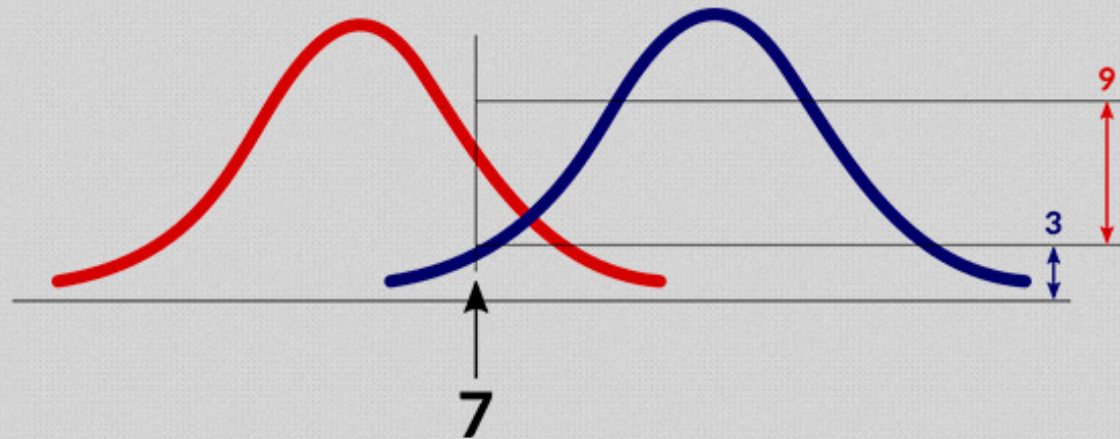
(probabilidade da classe c_j , dado que observamos d)

$$p(\text{Gafanhoto} | 7) = 3 / (3 + 9) = 0.250$$

$$p(\text{Esperança} | 7) = 9 / (3 + 9) = 0.750$$



Comprimento da antena = 7



Classificador bayesiano (Naïve Bayes)

Classificador Bayesiano usa o *teorema de Bayes*, que diz $p(c_j | d) = \frac{p(d | c_j) p(c_j)}{p(d)}$

$p(c_j | d)$ = probabilidade do exemplo d ser da classe c_j

É isto que estamos tentando calcular.

$p(d | c_j)$ = probabilidade de gerar exemplo d dada a classe c_j

Nós podemos imaginar que ser da classe leva a ter a característica d com alguma probabilidade.

$p(c_j)$ = probabilidade de ocorrência da classe c_j

Representa apenas quão frequente a classe c_j é na nossa base.

$p(d)$ = probabilidade de exemplo d ocorrer

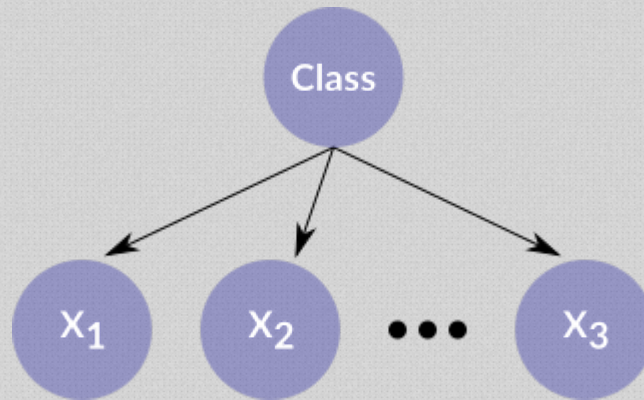
Isto pode ser ignorado, por ser igual para todas as classes.

Nos exemplos consideramos a classificação **Bayesiana** quando temos um **atributo**, por exemplo “**Comp. da antena**”.

“Podemos ter **muitas** características.
Como usamos **todas** as características?”

Classificador bayesiano (Naïve Bayes)

Os classificadores **Naïve Bayes** são geralmente representados com este tipo de grafo



A direção das flechas indica que **cada classe** está relacionada a **certas características**, com uma **certa probabilidade**.

Vantagens

- Treinamento rápido
- Rápido para classificar
- Não sensível a características irrelevantes
- Lida com dados reais e discretos
- Lida bem com dados contínuos

Desvantagens

- Assume independência das características (*atributos*)

[Exercício Naïve Bayes – KNIME]





Obrigado!

Agente Educacional

Sérgio M. Dias

sergio.dias@serpro.gov.br | #31 6539

Demais agentes educacionais sobre o assunto

Marcelo Pita | marcelo.pita@serpro.gov.br | #81 8794

Gustavo Torres | gustavo.gamatorres@serpro.gov.br | #31 6950