Classificação

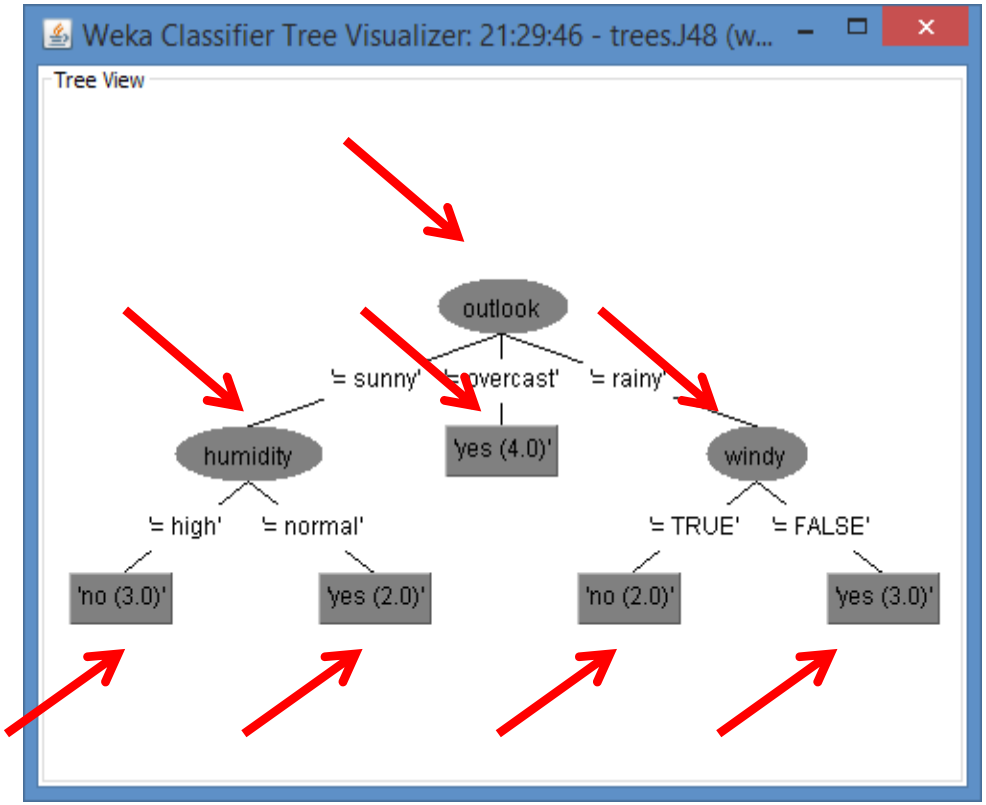
Formação Cientista de Dados

Tipos de Algoritmos

- Árvores de Decisão
- Regras
- Naïve Bayes
- Redes Bayesianas
- Redes Neurais Artificiais e aprendizado Profundo (Próxima Seção)
- Maquina de Vetor de Suporte
- Métodos de Grupos
- Aprendizado Baseado em Instância

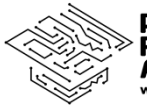
Árvores de Decisão

- Nodo raiz
- Nodos Internos
- Nodos Terminais
- Algoritmo de Partição: grau de pureza



Árvores de Decisão

```
=== Classifier model (full training set) ===
J48 pruned tree
------------------
outlook = sunny
|   humidity = high: no (3.0)
|   humidity = normal: yes (2.0)
outlook = overcast: yes (4.0)
outlook = rainy
|   windy = TRUE: no (2.0)
|   windy = FALSE: yes (3.0)
Number of Leaves  :         5
Size of the tree :          8
```



PROF. FERNANDO AMARAL
www.datascientist.com.br

Regras

sintaxe do tipo if/then/else

NNGE classifier

Rules generated :

class no IF : outlook in {rainy} ^ temperature in {mild,cool} ^ humidity in {high,normal} ^ windy in {TRUE} (2)

class yes IF : outlook in {overcast,rainy} ^ temperature in {hot,mild,cool} ^ humidity in {high,normal} ^ windy in {FALSE} (5)

class yes IF : outlook in {overcast} ^ temperature in {mild,cool} ^ humidity in {high,normal} ^ windy in {TRUE} (2)

class yes IF : outlook in {sunny} ^ temperature in {mild,cool} ^ humidity in {normal} ^ windy in {TRUE,FALSE} (2)

class no IF : outlook in {sunny} ^ temperature in {hot,mild} ^ humidity in {high} ^ windy in {TRUE,FALSE} (3)

Stat :

class yes : 3 exemplar(s) including 3 Hyperrectangle(s) and 0 Single(s).

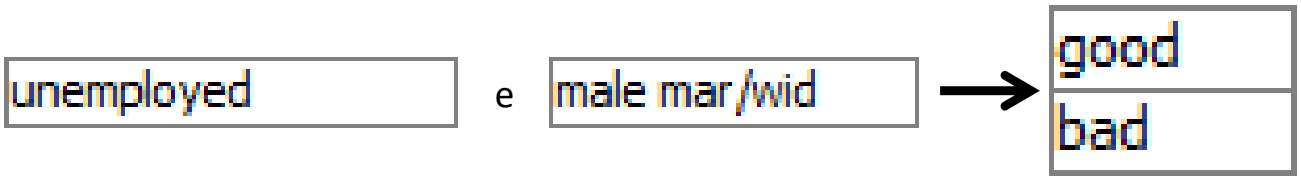
class no : 2 exemplar(s) including 2 Hyperrectangle(s) and 0 Single(s).

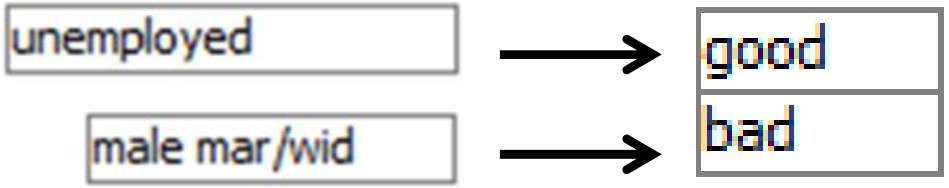
Total : 5 exemplars(s) including 5 Hyperrectangle(s) and 0 Single(s).

Feature weights : [0.24674981977443894 0.029222565658954577 0.15183550136234153 0.04812703040826924]

Naïve Bayes

Baseado na teoria das probabilidades e que supõe que os atributos vão influenciar a classe de forma independente





Naïve Bayes

Attribute	yes	no
	(0.63)	(0.38)
outlook		
sunny	3.0	4.0
overcast	5.0	1.0
rainy	4.0	3.0
[total]	12.0	8.0
temperature		
hot	3.0	3.0
mild	5.0	3.0
cool	4.0	2.0
[total]	12.0	8.0
humidity		
high	4.0	5.0
normal	7.0	2.0
[total]	11.0	7.0
windy		
TRUE	4.0	4.0
FALSE	7.0	3.0
[total]	11.0	7.0

Redes Bayesianas

Uma Rede Bayesiana pode mostrar eventual dependência entre os atributos através de probabilidade condicional