

# Árvores de Decisão

Prof. Me. Alexandre Henrick

Sistemas de Informação - 8º P

# Regras de Decisão

- Forma de representação do conhecimento.  
Relaciona informações ou fatos a alguma ação ou resultado
- Definem um conjunto de condições que devem ocorrer para que uma determinada ação seja executada
- Emula a estratégia de tomada de decisão de especialistas humanos (Sistema Especialista)

# Regras de Decisão

- Conforme já estudamos, sistemas especialistas possuem dois elementos importantes:
  - Conjunto de regras: Base de conhecimento
  - Inferência: Infere conclusão a partir de dados de entrada e da base de conhecimento. Por exemplo a atribuição de uma classe em uma observação

Registro	<i>sepal length</i>	<i>sepal width</i>	<i>petal length</i>	<i>petal width</i>	Classe
1	5,1	3,5	1,4	0,2	Iris-setosa
2	5,0	3,3	1,4	0,2	Iris-setosa
3	7,0	3,2	4,7	1,4	Iris-versicolor
4	5,7	2,8	4,1	1,3	Iris-versicolor
5	6,3	3,3	6,0	2,5	Iris-virginica
6	5,9	3,0	5,1	1,8	Iris-virginica

Regra que classifica Iris-virginica

SE *petal length* > 5 E *petal width* > 2 ENTÃO *Iris-virginica*

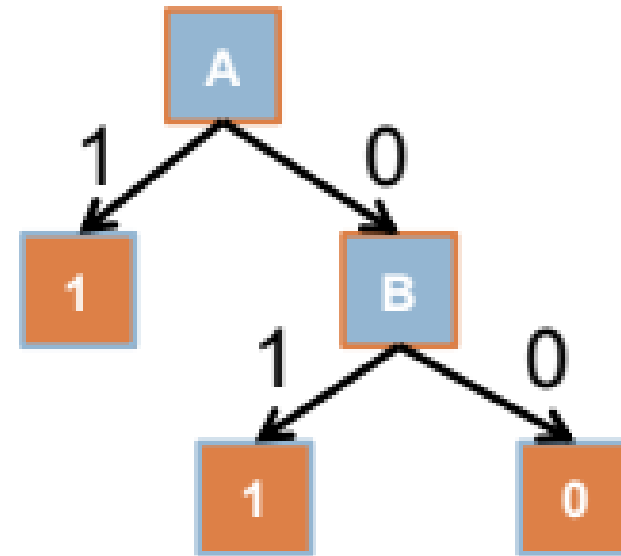
# Árvores de Decisão

- Classe de algoritmos que utilizam a estratégia de "dividir para conquistar"
- Divide em problemas menores para chegar a um resultado final
- Imita o processo de raciocínio humano, assim como nas regras de decisão

# Árvores de Decisão

- A ideia central é dividir um problema grande em problemas menores (ramos) de forma que esses problemas menores tenham complexidade menor
- Essa estratégia é replicada a cada sub-problema
- É um classificador baseado em regras SE-ENTÃO
- É uma representação de uma tabela de decisão sob forma de árvore

A	B	A ou B
0	0	0
0	1	1
1	0	1
1	1	1



# Árvores de Decisão

- É uma das técnicas mais utilizadas
- Possui uma grande vantagem em relação a outros métodos mais sofisticados: Interpretabilidade (White box)
- As condições de uma regra envolvem intervalos para os atributos



# Árvores de Decisão

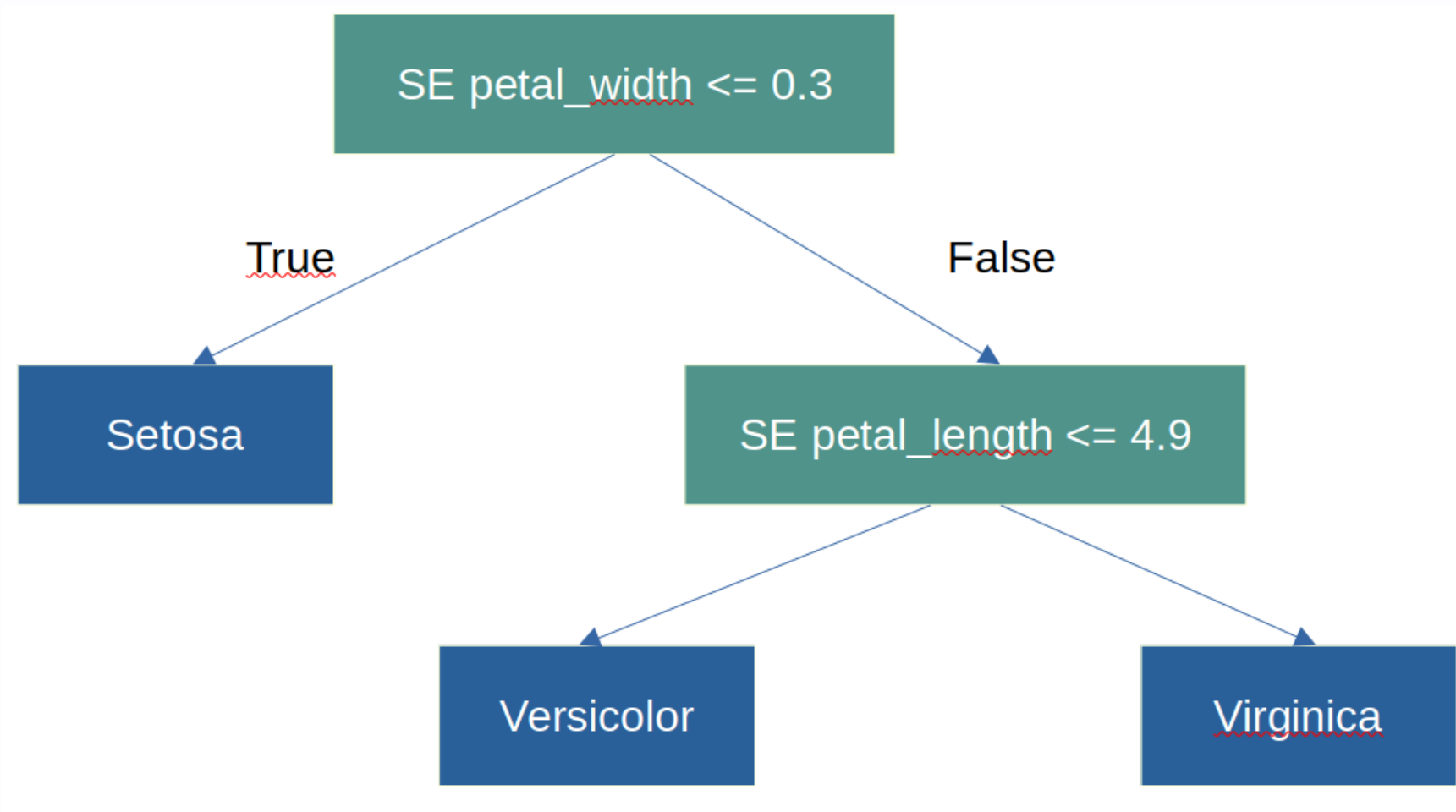
- É aplicado uma sequencia de regras, onde a observação vai sendo classificada em classes cada vez menores até chegar na classe final
- Ou seja, vai descendo na ramificação da árvore até chegar nas folhas, que representam as classes

# Árvores de Decisão

- É composta por 3 elementos principais:
  - Nós intermediários ou raiz: definem o parâmetro que será avaliado
  - Arestas: definem a transição entre os nós de acordo com os valores do atributo
  - Folhas: a classificação final da observação

# Árvores de Decisão

- Podemos usar 1 ou mais atributos para montar a árvore de decisão
- A escolha dos atributos é parte da criação da árvore de decisão



**Como a árvore de decisão escolhe os atributos e condições para montar sua estrutura?**

# Entropia

- Precisamos entender que na classificação queremos **segregar/separar/generalizar as observações**
- Para qualquer classificador é de extrema importancia saber **quais características mais dividem os dados**, ou seja, **quais atributos melhor separa os dados**
- Além de saber os atributos, precisamos saber também **quais valores desses atributos melhor generaliza os dados**

# Entropia

- A entropia é uma grandeza utilizada para **medir o grau de impureza ou desorganização dos dados**
- Foi adotada em Árvores de Decisão para construir as condições de cada nó

$$H(Y) = - \sum_{i=1}^c p_i \log_2 p_i,$$

onde  $p_i$  é a proporção de registros pertencentes a uma classe  $i$  e  $c$  a quantidade de classes

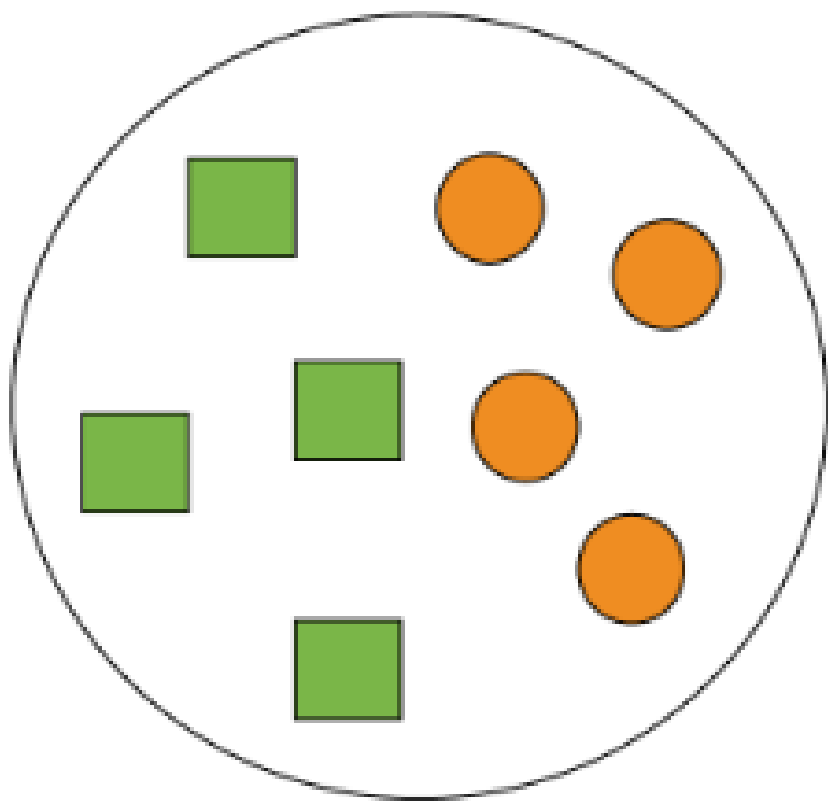


$$H(Y) = -\frac{8}{8} \log_2 \left( \frac{8}{8} \right) = 0$$

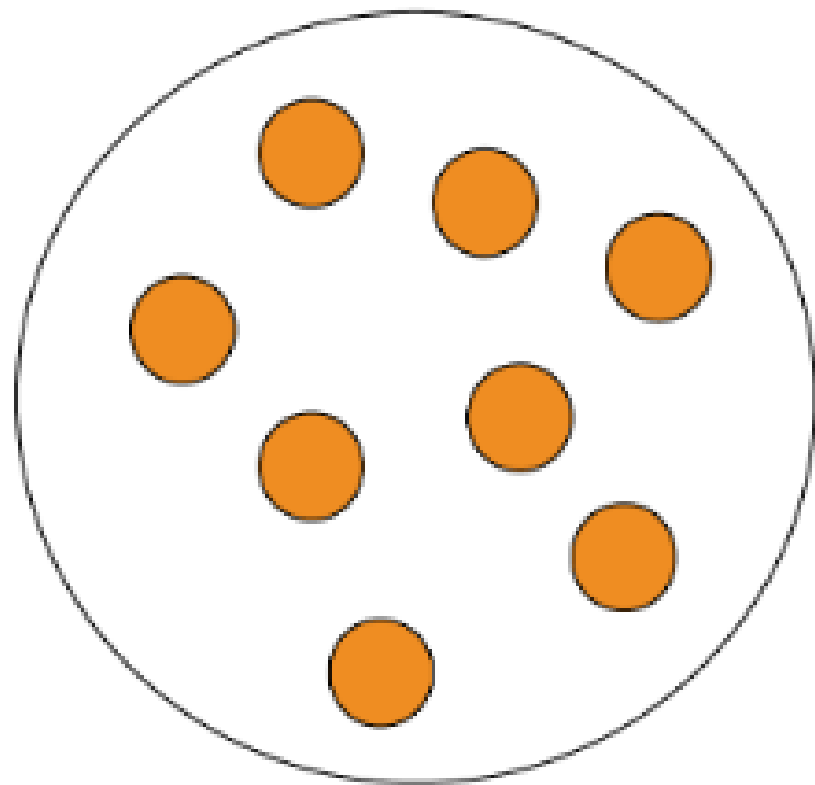
Exemplo 1: se temos 8 registros de apenas uma classe, temos uma proporção de 8/8 e seu valor de entropia será 0.

$$H(Y) = -\frac{4}{8}\log_2\left(\frac{4}{8}\right) - \frac{4}{8}\log_2\left(\frac{4}{8}\right) = 1$$

Exemplo 2: se temos 8 registros onde 4 deles pertencem a uma classe qualquer e os outros 4 pertencem a uma classe diferente, cada uma dessas classes possui uma proporção 4/8. Portanto a entropia do conjunto é 1 (desorganização máxima)



(a) Entropia = 1



(b) Entropia = 0

# Entropia

- Quando definimos uma condição qualquer, por exemplo: `SE petal_length >= 2`, o conjunto de registros que atendem essa condição possui uma entropia e podemos medi-la

# Entropia

- Portanto, a árvore de decisão vai calcular a entropia de várias condições e selecionar aquelas que melhor generalizam os dados, ou seja, aquelas com menor entropia
- Isso porque **se uma condição separa bem os dados, nosso conjunto fica mais organizado**
- Esse processo é repetido até chegar nas folhas

# Referências

- Baseado nos materiais disponibilizados pelo professor Dr. André Backes. Disciplina de Reconhecimento de Padrões - UFU-FACOM.  
Disponível em:  
<https://www.facom.ufu.br/~backes/pgc204.php>
- Russell, S. J. 1., & Norvig, P. (1995). Artificial intelligence: a modern approach. Englewood Cliffs, N.J., Prentice Hall.