

Lista 01

Exercícios sobre árvores de decisão

Prof. Marcelo Keese Albertini
Universidade Federal de Uberlândia

Exercício 1

Em que tipo de problemas um algoritmo de classificação baseado em árvores é útil?

Exercício 2

O que são medidas de impureza e como ela são utilizadas para a construção de árvores de decisão? Porque medidas de erros não são normalmente utilizadas nessa situação?

Exercício 3

Descreva os três principais componentes de um algoritmo de aprendizado de máquina baseado em árvores de decisão.

Exercício 4

Com o objetivo de identificar o sexo de penguins usando o dataset `data(penguins)`, escreva um código em R para medir a acurácia de um modelo baseado em árvore **CART** em um conjunto de treino.

Exercício 5

Com o objetivo de identificar o sexo de penguins usando o dataset `data(penguins)`, escreva um código em R para medir a acurácia de um modelo baseado em árvore **C5.0** em um conjunto de teste.

Exercício 6

Em qual situação existe risco de overfitting quando do aprendizado de árvores de decisão?

Exercício 7

Como evitar o risco de overfitting quando do aprendizado de árvores de decisão?

Exercício 8

Com o objetivo de identificar a espécie de flores usando o dataset `data(iris)`, escreva um código em R, pacote `ggplot2`, para visualizar as distribuições de cada atributo desse conjunto usando caixas de bigodes (`boxplots`).

Exercício 9

Qual é a taxa baseline de acurácia esperada para o seguinte dataset ?

Exemplo	a1	a2	a3	a4	classe
1	0	0	1	0	0
2	0	1	0	0	0
3	1	0	0	1	1
4	0	1	1	0	0
5	1	1	0	0	0
6	0	1	0	1	0

Exercício 10

O que é uma árvore de decisão?

Exercício 11

O conjunto padrão `iris` do R é armazenado originalmente em um formato chamado `data.frame`. Escreva um código em R para obter uma árvore do conjunto `iris` usando 2/3 da tabela para o processo de treino e 1/3 para teste.

Exercício 12

Escreva um algoritmo para construir uma árvore de decisão utilizando atributos categóricos e a medida de acurácia para escolha de um atributo para ramificação.

Exercício 13

Aplique o algoritmo para construção de árvore de decisão anteriormente escrito para o seguinte conjunto de exemplo. Escreva passo a passo a comparação de atributos para escolha e ramificação.

Exemplo	a1	a2	a3	a4	classe
1	0	0	1	0	0
2	0	1	0	0	0
3	1	0	0	1	1
4	0	1	1	0	0
5	1	1	0	0	0
6	0	1	0	1	0

Exercício 14

O que é e para que serve poda em árvores de decisão?

Exercício 15

Escreva um algoritmo de poda e aplique na árvore de decisão criada anteriormente.

Exercício 16

Em que tipo de problemas um algoritmo de classificação baseado em árvores **não** é útil?

Exercício 17

Qual é a relação entre regras de classificação e árvores de decisão?