

# Exercícios

Árvores de Decisão

e

Estratégias de Avaliação de Classificadores

- Construa uma árvore de decisão com base no critério de seleção de atributos *ganho de informação*, considerando o seguinte conjunto de dados para treinamento

Name	Body Temperature	Skin Cover	Gives Birth	Aquatic Creature	Aerial Creature	Has Legs	Hibernates	Class Label
human	warm-blooded	hair	yes	no	no	yes	no	mammal
python	cold-blooded	scales	no	no	no	no	yes	non-mammal
salmon	cold-blooded	scales	no	yes	no	no	no	non-mammal
whale	warm-blooded	hair	yes	yes	no	no	no	mammal
frog	cold-blooded	none	no	semi	no	yes	yes	non-mammal
komodo dragon	cold-blooded	scales	no	no	no	yes	no	non-mammal
bat	warm-blooded	hair	yes	no	yes	yes	yes	mammal
pigeon	warm-blooded	feathers	no	no	yes	yes	no	non-mammal
cat	warm-blooded	fur	yes	no	no	yes	no	mammal
leopard shark	cold-blooded	scales	yes	yes	no	no	no	non-mammal
turtle	cold-blooded	scales	no	semi	no	yes	no	non-mammal
penguin	warm-blooded	feathers	no	semi	no	yes	no	non-mammal
porcupine	warm-blooded	quills	yes	no	no	yes	yes	mammal
eel	cold-blooded	scales	no	yes	no	no	no	non-mammal
salamander	cold-blooded	none	no	semi	no	yes	yes	non-mammal

- Calcule o *ganho de informação* de cada atributo considerando o seguinte conjunto de dados

Name	Body Temperature	Skin Cover	Gives Birth	Aquatic Creature	Aerial Creature	Has Legs	Hibernates	Class Label
human	warm-blooded	hair	yes	no	no	yes	no	mammal
python	cold-blooded	scales	no	no	no	no	yes	reptile
salmon	cold-blooded	scales	no	yes	no	no	no	fish
whale	warm-blooded	hair	yes	yes	no	no	no	mammal
frog	cold-blooded	none	no	semi	no	yes	yes	amphibian
komodo dragon	cold-blooded	scales	no	no	no	yes	no	reptile
bat	warm-blooded	hair	yes	no	yes	yes	yes	mammal
pigeon	warm-blooded	feathers	no	no	yes	yes	no	bird
cat	warm-blooded	fur	yes	no	no	yes	no	mammal
leopard shark	cold-blooded	scales	yes	yes	no	no	no	fish
turtle	cold-blooded	scales	no	semi	no	yes	no	reptile
penguin	warm-blooded	feathers	no	semi	no	yes	no	bird
porcupine	warm-blooded	quills	yes	no	no	yes	yes	mammal
eel	cold-blooded	scales	no	yes	no	no	no	fish
salamander	cold-blooded	none	no	semi	no	yes	yes	amphibian

- Considerando o contexto abaixo, indique quais as medidas que podem ser extraídas da matriz de confusão que podem trazer as melhores avaliações sobre o contexto do classificador. Justifique com base no contexto fornecido.
- um restaurante indiano está inserindo novos itens bem apimentados em seu cardápio, e o gerente do restaurante decidiu fazer uma degustação às cegas. Então, ele recorre a um classificador binário para tomar a decisão sobre servir (classe positiva) ou não servir (classe negativa) os novos pratos para determinados clientes. Dentre os 22.400 pedidos existentes no subconjunto de teste usado na aferição de desempenho do classificador, apenas 2.900 não incluem pratos apimentados, sendo que, desses, 90% referem-se a clientes que possuem alergia à pimenta – o gerente não tem acesso a essas informações e deposita total confiança no software classificador à sua disposição.

		<i>Classe predita (<math>f(x)</math>)</i>	
		<i>Pedidos com pratos apimentados</i>	<i>Pedidos sem pratos apimentados</i>
<i>Classe esperada (<math>y</math>)</i>	<i>Pedidos com pratos apimentados (servir)</i>	17.312	2.188
	<i>Pedidos sem pratos apimentados (não servir)</i>	1.340	1.560