

Inteligência Artificial

Aula 9 - vídeo 2 - Aprendizado de Máquina

João C. P. da Silva

Dept. Ciência da Computação - UFRJ

21 de outubro de 2020

Conceitos Básicos

- **Conceito Alvo:** O que deve ser aprendido.

Conceitos Básicos

- **Conceito Alvo:** O que deve ser aprendido.
- **Atributos:** Conjunto de características com valores fixos que identificam as instâncias.

Conceitos Básicos

- **Conceito Alvo:** O que deve ser aprendido.
- **Atributos:** Conjunto de características com valores fixos que identificam as instâncias.
 - **Categóricos:** quando o atributo assume valores discretos.
 - **Numéricos:** valores inteiros ou reais.

Conceitos Básicos

- **Conceito Alvo:** O que deve ser aprendido.
- **Atributos:** Conjunto de características com valores fixos que identificam as instâncias.
 - **Categóricos:** quando o atributo assume valores discretos.
 - **Numéricos:** valores inteiros ou reais.
- **Instâncias:** Elementos de um conceito.

Conceitos Básicos

- **Conceito Alvo:** O que deve ser aprendido.
- **Atributos:** Conjunto de características com valores fixos que identificam as instâncias.
 - **Categóricos:** quando o atributo assume valores discretos.
 - **Numéricos:** valores inteiros ou reais.
- **Instâncias:** Elementos de um conceito.
- **Exemplos:** Conjunto de instâncias utilizado para aprender o conceito.

Conceitos Básicos

- **Conceito Alvo:** O que deve ser aprendido.
- **Atributos:** Conjunto de características com valores fixos que identificam as instâncias.
 - **Catégoricos:** quando o atributo assume valores discretos.
 - **Numéricos:** valores inteiros ou reais.
- **Instâncias:** Elementos de um conceito.
- **Exemplos:** Conjunto de instâncias utilizado para aprender o conceito.
- **Descrição do Conceito:** Resultado produzido por um processo de aprendizado.

Conceitos Básicos - Exemplo

- **Conceito Alvo:** 'Dias em que Ana gosta de praticar seu esporte favorito.'
Vamos chamá-lo de **EnjoySpt**

Conceitos Básicos - Exemplo

- **Conceito Alvo:** 'Dias em que Ana gosta de praticar seu esporte favorito.'
Vamos chamá-lo de **EnjoySpt**
- **Atributos e seus Valores**
 - **Sky:** sunny, rainy, cloudy
 - **Temp:** warm, cold
 - **Humid:** normal, high
 - **Wind:** strong, weak
 - **Water:** warm, cool
 - **Forecast:** same, change

Conceitos Básicos - Exemplo

- **Conceito Alvo:** 'Dias em que Ana gosta de praticar seu esporte favorito.'
Vamos chamá-lo de **EnjoySpt**
- **Atributos e seus Valores**
 - **Sky:** sunny, rainy, cloudy
 - **Temp:** warm, cold
 - **Humid:** normal, high
 - **Wind:** strong, weak
 - **Water:** warm, cool
 - **Forecast:** same, change
- **Instâncias:** $3 \times 2 \times 2 \times 2 \times 2 \times 2 = 3 \times 2^5 = 96$ instâncias

Conceitos Básicos - Exemplo

- **Exemplos:** Conjunto de instâncias utilizado para aprender o conceito.

Sky	Temp	Humid	Wind	Water	Forecast	EnjoySpt
Sunny	Warm	Normal	Strong	Warm	Same	Yes
Sunny	Warm	High	Strong	Warm	Same	Yes
Rainy	Cold	High	Strong	Warm	Change	No
Sunny	Warm	High	Strong	Cool	Change	Yes
Rainy	Cold	High	Strong	Cool	Change	No
Sunny	Warm	High	Strong	Cool	Same	Yes
Cloudy	Cold	High	Strong	Cool	Change	No

Conceitos Básicos - Exemplo

- Exemplos:** Conjunto de instâncias utilizado para aprender o conceito.

Sky	Temp	Humid	Wind	Water	Forecast	EnjoySpt
Sunny	Warm	Normal	Strong	Warm	Same	Yes
Sunny	Warm	High	Strong	Warm	Same	Yes
Rainy	Cold	High	Strong	Warm	Change	No
Sunny	Warm	High	Strong	Cool	Change	Yes
Rainy	Cold	High	Strong	Cool	Change	No
Sunny	Warm	High	Strong	Cool	Same	Yes
Cloudy	Cold	High	Strong	Cool	Change	No

96 instâncias \Rightarrow 7 exemplos

Conceitos Básicos - Exemplo

- **Exemplos:** Conjunto de instâncias utilizado para aprender o conceito.

Sky	Temp	Humid	Wind	Water	Forecast	EnjoySpt
Sunny	Warm	Normal	Strong	Warm	Same	Yes
Sunny	Warm	High	Strong	Warm	Same	Yes
Rainy	Cold	High	Strong	Warm	Change	No
Sunny	Warm	High	Strong	Cool	Change	Yes
Rainy	Cold	High	Strong	Cool	Change	No
Sunny	Warm	High	Strong	Cool	Same	Yes
Cloudy	Cold	High	Strong	Cool	Change	No

96 instâncias \Rightarrow 7 exemplos

Aprender: generalizar a classificação das 89 instâncias não conhecidas a partir dos 7 exemplos.

Conceitos Básicos - Exemplo

- **Exemplos:** Conjunto de instâncias utilizado para aprender o conceito.

Sky	Temp	Humid	Wind	Water	Forecast	EnjoySpt
Sunny	Warm	Normal	Strong	Warm	Same	Yes
Sunny	Warm	High	Strong	Warm	Same	Yes
Rainy	Cold	High	Strong	Warm	Change	No
Sunny	Warm	High	Strong	Cool	Change	Yes
Rainy	Cold	High	Strong	Cool	Change	No
Sunny	Warm	High	Strong	Cool	Same	Yes
Cloudy	Cold	High	Strong	Cool	Change	No

96 instâncias \Rightarrow 7 exemplos

Aprender: generalizar a classificação das 89 instâncias não conhecidas a partir dos 7 exemplos.

Como aprender ?

Separar os exemplos

- Separe os exemplos conhecidos em dois conjuntos: **treinamento** e **teste**.

TREINAMENTO

Sky	Temp	Humid	Wind	Water	Forecast	EnjoySpt
Sunny	Warm	Normal	Strong	Warm	Same	Yes
Sunny	Warm	High	Strong	Warm	Same	Yes
Rainy	Cold	High	Strong	Warm	Change	No
Sunny	Warm	High	Strong	Cool	Change	Yes
Rainy	Cold	High	Strong	Cool	Change	No
Sunny	Warm	High	Strong	Cool	Same	Yes

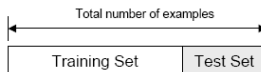
TESTE

Cloudy	Cold	High	Strong	Cool	Change	No
--------	------	------	--------	------	--------	----

Conjuntos de Treinamento e Teste

- **Holdout**

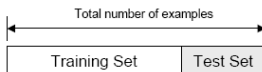
- $2/3$ para treinamento
- $1/3$ para teste
- **Estimativa Pessimista:** usa somente parte dos dados.



Conjuntos de Treinamento e Teste

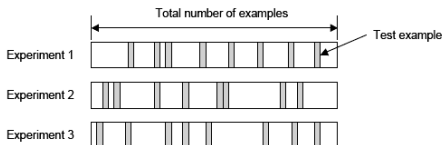
- **Holdout**

- 2/3 para treinamento
- 1/3 para teste
- **Estimativa Pessimista:** usa somente parte dos dados.



- **Amostragem Aleatória**

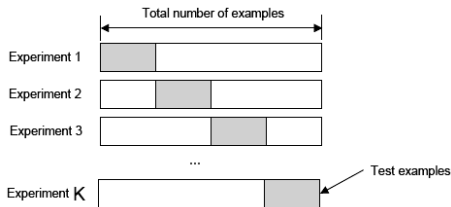
- Cada divisão seleciona um número fixo de exemplos sem reposição.
- Para cada conjunto de dados dividido, o conjunto de treinamento é usado para treinar e depois o de teste usado para avaliar o erro obtido em cada conjunto.
- O erro real estimado é a média dos erros obtidos.



Conjuntos de Treinamento e Teste

- **K- Fold Cross-validation**

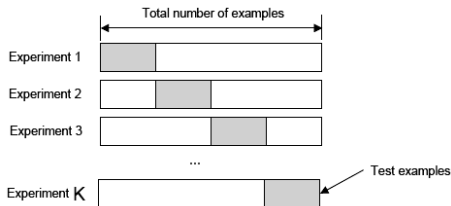
- Para cada um dos k experimentos: $k - 1$ folds para treinar e 1 para testar.
- Todos os elementos são usados para treinamento e teste.



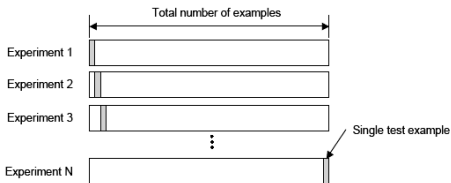
Conjuntos de Treinamento e Teste

- **K- Fold Cross-validation**

- Para cada um dos k experimentos: $k - 1$ folds para treinar e 1 para testar.
- Todos os elementos são usados para treinamento e teste.



- **Leave One Out Cross-Validation**



Acurácia

- Como saber se o que foi aprendido, de fato, ajuda quando novas instâncias são examinadas?

Acurácia

- Como saber se o que foi aprendido, de fato, ajuda quando novas instâncias são examinadas?
- Como medir a acurácia do aprendizado?

Acurácia

- Como saber se o que foi aprendido, de fato, ajuda quando novas instâncias são examinadas?
- Como medir a acurácia do aprendizado?
- **Acurácia:** percentagem dos elementos que são corretamente classificadas (taxa de reconhecimento)

Acurácia

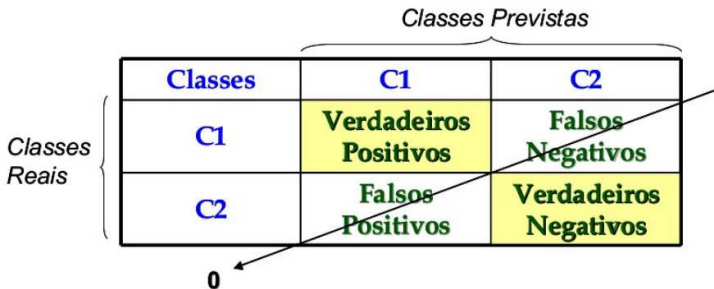
- Como saber se o que foi aprendido, de fato, ajuda quando novas instâncias são examinadas?
- Como medir a acurácia do aprendizado?
- **Acurácia:** percentagem dos elementos que são corretamente classificadas (taxa de reconhecimento)

		Classes Previstas			
	Classes	A = yes	A = no	Total	Reconhecimento (%)
Classes Reais	A = yes	6954	46	7000	99.34
	A = no	412	2588	3000	86.27
	Total	7366	2634	10000	95.42

Verdadeiros Positivos : *Número de elementos da classe A=yes, classificados como A=yes*

Acurácia

- **Acurácia:** percentagem dos elementos que são corretamente classificadas (taxa de reconhecimento)

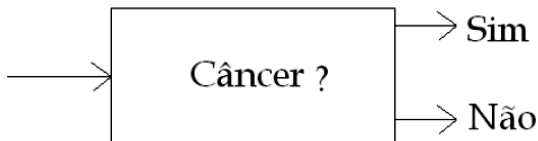


		Classes Previstas	
		Classes	
Classes Reais	C1	Verdadeiros Positivos	Falsos Negativos
	C2	Falsos Positivos	Verdadeiros Negativos

0

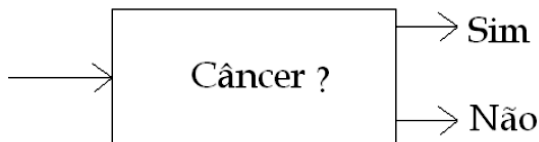
Acurácia

- Considere um conjunto de exemplos com 1000 exemplos, sendo 30 positivos e 970 negativos (3% de casos de câncer)
- Suponha que treinamos um classificador e obtemos acurácia de 90%.
- O classificador é bom?



Acurácia

- Considere um conjunto de exemplos com 1000 exemplos, sendo 30 positivos e 970 negativos (3% de casos de câncer)
- Suponha que treinamos um classificador e obtemos acurácia de 90%.
- O classificador é bom?



Acurácia: 90% (caso com 5+ e 45-)

Medidas Alternativas

- **Sensibilidade:** proporção de instâncias positivas corretamente classificadas.

$$\frac{\#verdadeiros_positivos}{\#positivos}$$

- **Especificidade:** proporção de instâncias negativas corretamente classificadas.

$$\frac{\#verdadeiros_negativos}{\#negativos}$$

- **Precisão:** proporção de instâncias corretamente classificadas.

$$\frac{\#verdadeiros_positivos}{\#verdadeiros_positivos + \#falsos_positivos}$$

Medidas Alternativas

		Classes Previstas			
Classes Reais	Classes	A = yes	A = no	Total	Reconhecimento (%)
	A = yes	6954	46	7000	99.34
	A = no	412	2588	3000	86.27
	Total	7366	2634	10000	95.42

- **Sensibilidade:** proporção de instâncias positivas corretamente classificadas.

$$\frac{6954}{7000} = 0.9934$$

- **Especificidade:** proporção de instâncias negativas corretamente classificadas.

$$\frac{2588}{3000} = 0.8627$$

- **Precisão:** proporção de instâncias corretamente classificadas.

$$\frac{6954}{7366} = 0.9441$$

As medidas devem ser feitas no treinamento e no teste

Inteligência Artificial

Aula 9 - vídeo 2 - Aprendizado de Máquina

João C. P. da Silva

Dept. Ciência da Computação - UFRJ

21 de outubro de 2020