

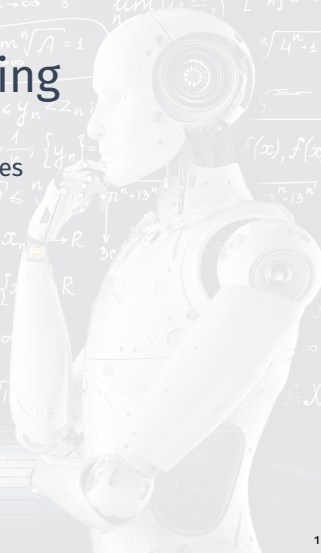
# Machine Learning

Aula 10

Comitês de Classificadores

Evandro J.R. Silva

Uninassau Teresina



# Sumário

- 1 Introdução
- 2 Seleção de Classificadores
- 3 Fusão de Classificadores
- 4 Modelos Clássicos
- 5 FIM



# 1 Introdução

## 2 Seleção de Classificadores

## 3 Fusão de Classificadores

## 4 Modelos Clássicos

## 5 FIM



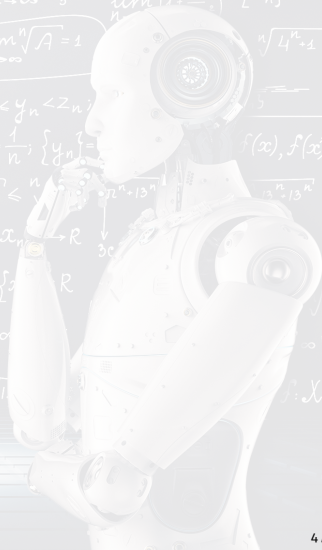
# Introdução

- Comitês de Classificadores (*Ensemble of Classifiers*) — combinação de dois ou mais classificadores a fim de que se obtenham melhores resultados na classificação de padrões.



# Introdução

## • Motivações principais:



# Introdução

## • Motivações principais:

1 Conjuntos de diferentes classificadores pode oferecer informações complementares sobre os padrões a serem classificados.

• Isso aumenta a efetividade geral do processo de reconhecimento de padrões.

# Introdução

## • Motivações principais:

- 2 A construção de um único classificador **forte** (ou seja, que tenha alto desempenho) pode ser extremamente difícil.
  - Construir vários classificadores **fracos** é mais fácil e pode retornar o desempenho desejado.

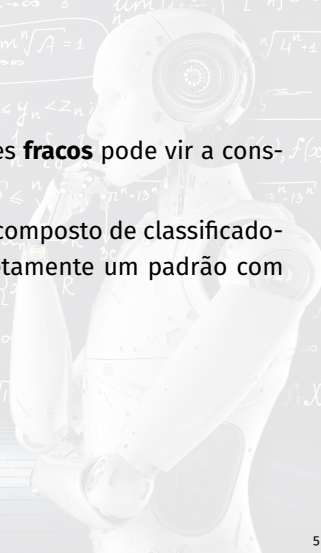
# Introdução

- Teoricamente, um comitê de classificadores **fracos** pode vir a consistir um classificador **forte**.



# Introdução

- Teoricamente, um comitê de classificadores **fracos** pode vir a substituir um classificador **forte**.
- Para que isso ocorra, o comitê precisa ser composto de classificadores independentes que classifiquem corretamente um padrão com uma probabilidade maior que 50%.



# Introdução

- Um comitê de classificadores típico contém normalmente três componentes:

1 **Indutor base:** ou classificador base, é um algoritmo de indução que para um dado conjunto de treinamento forma um classificador. Uma Árvore de Decisão é um indutor base de um *Random Forest*.

# Introdução

- Um comitê de classificadores típico contém normalmente três componentes:

- 2 **Gerador de diversidade:** o qual é responsável por gerar classificadores diversificados.

# Introdução

- Um comitê de classificadores típico contém normalmente três componentes:

③ **Combinador:** é responsável pela combinação dos classificadores.

# Introdução

- Para se alcançar uma maior **diversidade**, é possível escolher dentre várias abordagens:

- 1 **Manipular o conjunto de treinamento:** cada membro do comitê é treinado em um conjunto diferente de amostras, ou subespaço diferente.

# Introdução

- Para se alcançar uma maior **diversidade**, é possível escolher dentre várias abordagens:

- 2 **Manipular o indutor:** cada membro do comitê é treinado com um indutor que é diferentemente manipulado.

# Introdução

- Para se alcançar uma maior **diversidade**, é possível escolher dentre várias abordagens:

- 3 **Modificando a representação do atributo desejado:** cada classificador lida com um conceito diferente do resultado desejado.

# Introdução

- Para se alcançar uma maior **diversidade**, é possível escolher dentre várias abordagens:

- 4 **Hibridização:** a diversidade é obtida ao se usar vários indutores base ou estratégias de construção de comitês.



# Introdução

- A **combinação** dos classificadores é feita normalmente através de duas abordagens: **Seleção** de Classificadores e **Fusão** de Classificadores.

## 1 Introdução

## 2 Seleção de Classificadores

## 3 Fusão de Classificadores

## 4 Modelos Clássicos

## 5 FIM



# Seleção de Classificadores

- A seleção pode ocorrer de forma **estática** ou **dinâmica**.



# Seleção de Classificadores

- A seleção pode ocorrer de forma **estática** ou **dinâmica**.
- A partir de um conjunto de classificadores treinados, um subgrupo é escolhido para classificar o conjunto de teste.

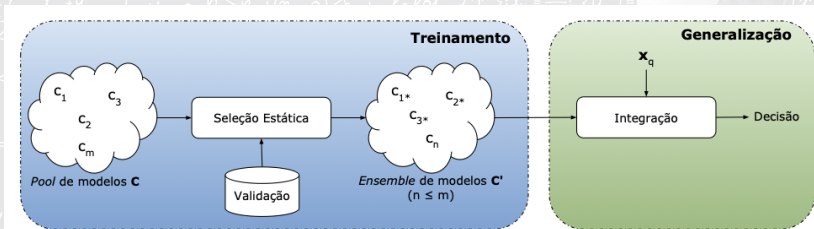


Figura 1: Fonte: [Máquinas que aprendem](#).

# Seleção de Classificadores

- A seleção pode ocorrer de forma **estática** ou **dinâmica**.
  - Para cada instância de teste, o melhor classificador, ou um conjunto dos melhores classificadores, é selecionado para a classificação

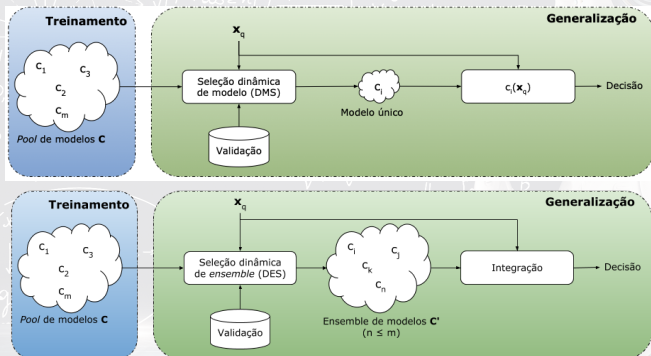


Figura 2: Fonte: [Máquinas que aprendem.](#)

# Seleção de Classificadores

- Existe também o método **clustering-and-selection**, onde é primeiro aplicado um algoritmo de *clustering*.
- Então, para cada *cluster* classificadores são treinados e selecionados estática ou dinamicamente.

## 1 Introdução

## 2 Seleção de Classificadores

## 3 Fusão de Classificadores

## 4 Modelos Clássicos

## 5 FIM



# Fusão de Classificadores

- A Fusão de Classificadores é uma mímica ao processo humano de buscar várias opiniões de especialistas antes de tomar uma decisão importante.



# Fusão de Classificadores

- A opinião de vários especialistas sobre determinado assunto é observada em conjunto, e através de alguma função sobre tais opiniões, uma decisão final é tomada.

# Fusão de Classificadores

- Seja  $D = \{D_1, D_2, \dots, D_L\}$  um conjunto de  $L$  classificadores e  $\Omega = \{\omega_1, \omega_2, \dots, \omega_C\}$  um conjunto de  $C$  classes.
- Uma matriz  $D_{ij}$ , onde  $i = 1, \dots, L$  e  $j = 1, \dots, C$  é uma representação do resultado da classificação e mostra os graus de suporte do classificador  $i$  para a classe  $j$ .
- Vejamos um exemplo (uma instância sendo classificada por 3 classificadores):

	$\omega_1$	$\omega_2$	$\omega_3$	$\omega_4$
$D_1$	0,1	0,5	0,4	0,0
$D_2$	0,3	0,4	0,2	0,1
$D_3$	0,2	0,0	0,8	0,0

# Fusão de Classificadores

- Funções comumente usadas para a fusão dos classificadores: **voto majoritário, voto ponderado, regra do mínimo, regra do máximo, regra do produto, regra da média e regra da mediana.**

# Fusão de Classificadores

- Mais comuns!
- Funções comumente usadas para a fusão dos classificadores: **voto majoritário**, **voto ponderado**, **regra do mínimo**, **regra do máximo**, **regra do produto**, **regra da média** e **regra da mediana**.

# Fusão de Classificadores

## • Voto majoritário: $\omega_2$

- $\omega_1$  — 0 votos
- $\omega_2$  — 2 votos [ $D_1, D_2$ ]
- $\omega_3$  — 1 voto [ $D_3$ ]
- $\omega_4$  — 0 votos

	$\omega_1$	$\omega_2$	$\omega_3$	$\omega_4$
$D_1$	0,1	0,5	0,4	0,0
$D_2$	0,3	0,4	0,2	0,1
$D_3$	0,2	0,0	0,8	0,0

# Fusão de Classificadores

## • Voto ponderado: $\omega_3$

- $\omega_1$  — 0 votos
- $\omega_2$  — 2 votos [ $D_1 \times 1$ ,  $D_2 \times 1$ ]
- $\omega_3$  — 3 votos [ $D_3 \times 3$ ]
- $\omega_4$  — 0 votos

	$\omega_1$	$\omega_2$	$\omega_3$	$\omega_4$
$D_1$	0,1	0,5	0,4	0,0
$D_2$	0,3	0,4	0,2	0,1
$D_3$	0,2	0,0	0,8	0,0

# Fusão de Classificadores

## • Regra do mínimo: $\omega_3$

- 1 Selecionar o menor valor de suporte dado a cada classe.
- 2 Escolher a classe com o maior valor de suporte, dentre os mínimos.

	$\omega_1$	$\omega_2$	$\omega_3$	$\omega_4$
$D_1$	0,1	0,5	0,4	0,0
$D_2$	0,3	0,4	0,2	0,1
$D_3$	0,2	0,0	0,8	0,0

# Fusão de Classificadores

- Regra do máximo:  $\omega_3$**

- 1 Selecionar o maior valor de suporte dado a cada classe.
- 2 Escolher a classe com o maior valor de suporte, dentre os máximos.

	$\omega_1$	$\omega_2$	$\omega_3$	$\omega_4$
$D_1$	0,1	0,5	0,4	0,0
$D_2$	0,3	0,4	0,2	0,1
$D_3$	0,2	0,0	0,8	0,0



# Fusão de Classificadores

- **Regra do produto:**  $\omega_3$ 
  - 1 Produto do suporte de cada classe.
  - 2 Escolher a classe com o maior valor de suporte, após o produto.

	$\omega_1$	$\omega_2$	$\omega_3$	$\omega_4$
$D_1$	0,1	0,5	0,4	0,0
$D_2$	0,3	0,4	0,2	0,1
$D_3$	0,2	0,0	0,8	0,0
	0,006	0,0	0,064	0,0

# Fusão de Classificadores

- **Regra da média:**  $\omega_3$ 
  - 1 Média aritmética sobre os valores de suporte de cada classe.
  - 2 Escolher a classe com a maior média.

	$\omega_1$	$\omega_2$	$\omega_3$	$\omega_4$
$D_1$	0,1	0,5	0,4	0,0
$D_2$	0,3	0,4	0,2	0,1
$D_3$	0,2	0,0	0,8	0,0
	0,2	0,3	0,47	0,03

# Fusão de Classificadores

- **Regra da mediana:**  $\omega_2$  e  $\omega_3$ 
  - 1 Mediana sobre os valores de suporte de cada classe.
  - 2 Escolher a classe com a maior mediana.

	$\omega_1$	$\omega_2$	$\omega_3$	$\omega_4$
$D_1$	0,1	0,5	0,4	0,0
$D_2$	0,3	0,4	0,2	0,1
$D_3$	0,2	0,0	0,8	0,0
	0,2	0,4	0,4	0,0



# 1 Introdução

## 2 Seleção de Classificadores

### 3 Fusão de Classificadores

#### 4 Modelos Clássicos

##### 5 FIM

# Modelos Clássicos

- Dois dos modelos mais clássicos são: **Bagging** e **AdaBoost**.

# Modelos Clássicos

- Dois dos modelos mais clássicos são: **Bagging** e **AdaBoost**.
  - Seu nome vem de **Bootstrap aggregating**.
  - A partir de um conjunto de treinamento  $m$  cópias são feitas com o uso do método *bootstrap* e, para cada cópia, um classificador é treinado.
  - As instâncias que não foram selecionadas para qualquer uma das cópias (*out-of-bag*) são utilizadas para verificar a qualidade do comitê de classificadores.
  - Instâncias de teste são classificadas a partir da fusão dos classificadores treinados.

# Modelos Clássicos

- Dois dos modelos mais clássicos são: **Bagging** e **AdaBoost**.
  - A ideia principal é dar um maior foco a padrões que são mais difíceis de classificar.
  - O AdaBoost mantém um conjunto de pesos sobre os exemplos de treinamento. Em cada iteração  $i$ , o classificador  $C_i$  é treinado para minimizar o erro ponderado sobre o conjunto de treinamento.
  - As instâncias que são erroneamente classificadas têm seus pesos aumentados, e as que foram corretamente classificadas têm seus pesos diminuídos. O próximo classificador é treinado com os pesos atualizados.
  - Após o treinamento, a fusão dos classificadores é feita através do voto ponderado. O peso de cada classificador é calculado de acordo com sua precisão durante o treinamento.



# 1 Introdução

## 2 Seleção de Classificadores

### 3 Fusão de Classificadores

#### 4 Modelos Clássicos

##### 5 FIM



Antes de terminar, vamos ver um vídeo:

[Random Forests](#)

