

# CPE 722 – Redes Neurais Não Supervisionadas e Agrupamentos

## 1- Introdução

### Redes Neurais Artificiais

#### Inteligência Artificial / Computacional

Cognitiva, Simbólica

IA “tradicional”

Evolucionista

AG, Vida Artificial, Enxame

Conexionista

Redes Neurais Artificiais

### Redes Neurais Artificiais

Redes Feedforward

*Aproximadores e classificadores,  
Redes “backpropagation”*

**Redes Não Supervisionadas**  
*Classificadores*

**Redes Realimentadas**  
*Otimizadores*

### Redes Neurais Naturais

## **Redes Neurais não Supervisionadas**

### **Redes Neurais com Treinamento não Supervisionado**

**uma família de redes:**

**Camada de Kohonen, SOM, ART,  
Counterpropagation, LQV, etc.**

### **Por que estudar isto ?**

**Propriedades importantes:**

**Aprendizado cego, não supervisionado (on line).**

**Plasticidade.**

### **Estas redes realizam:**

**Classificadores por similaridade**

### **Aplicações:**

**Reconhecimento de padrões;**

**Estatística de sinais;**

**Memórias associativas;**

**Filtragem não-linear;**

**Compressão de dados;**

**Mapas auto-organizáveis;**

**etc...**

## 2 - Noções iniciais sobre Classificadores

### Mapeadores



### Mapeadores lineares / não lineares

$$\underline{\tilde{y}} = \underline{M} \underline{x}$$

$$\underline{\tilde{y}} = \varphi(\underline{x})$$

### Tipos de mapeadores não lineares

#### Aproximadores

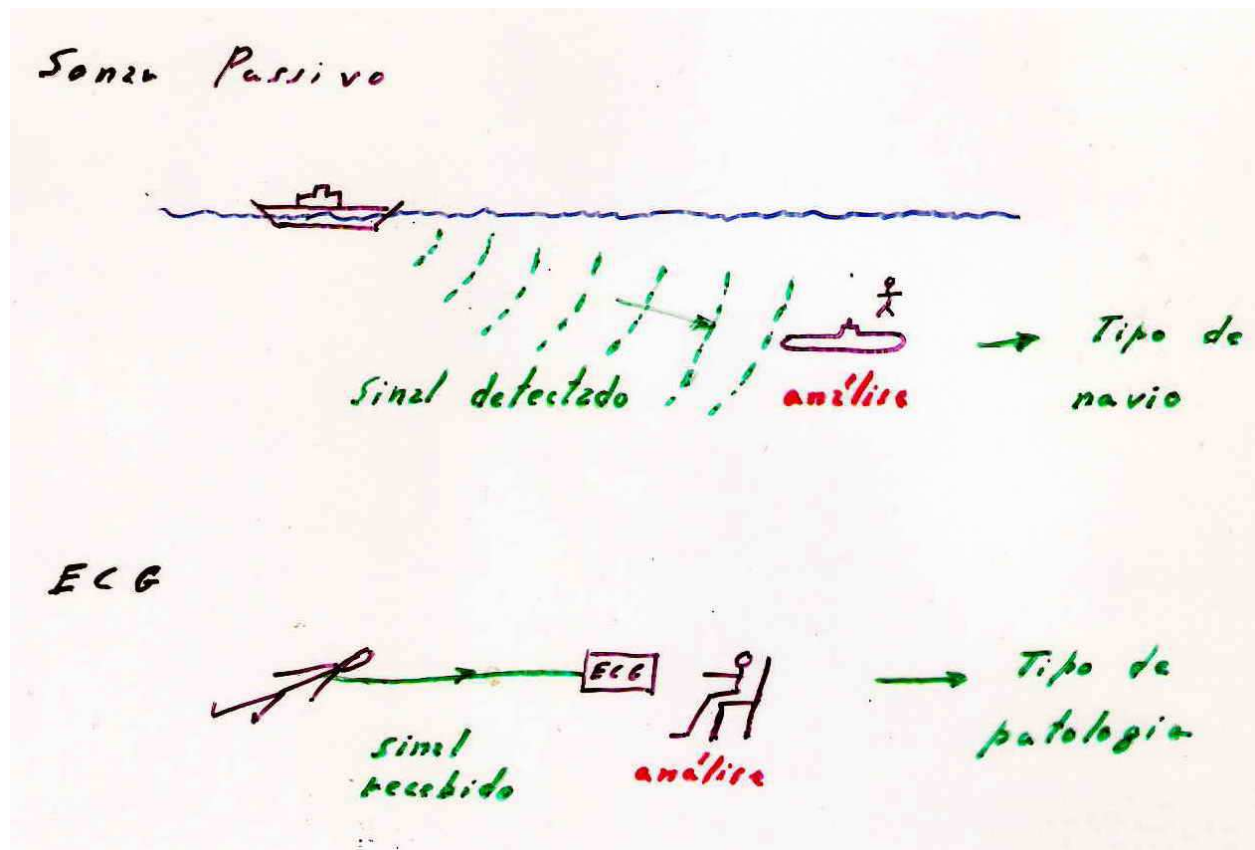
Associam saídas contínuas às entradas

Saídas contínuas  $y_i \in (-1, +1)$

#### Classificadores

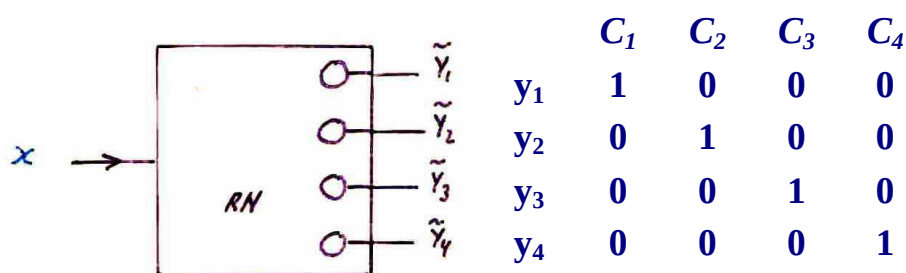
Associam saídas lógicas às entradas (classificam as entradas)

Saídas lógicas  $y_i \in \{0, 1\}$



## 2.1 Convenção / Estratégia:

**Codificação da Saída** (maximamente esparsa)



Cada saída separa a sua classe das demais, i.e.

classe versus não classe (um contra todos)

Separação de N classes

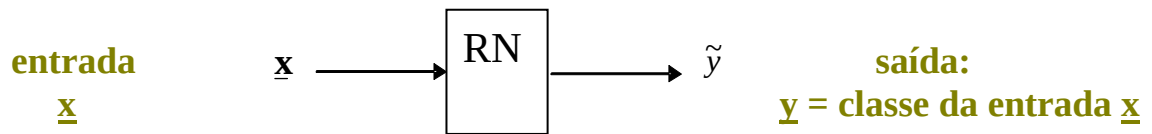


Separação de duas classes:

classe vs. não classe

### 3 - Tipos de classificação / Tipos de aprendizado

#### Classificadores:



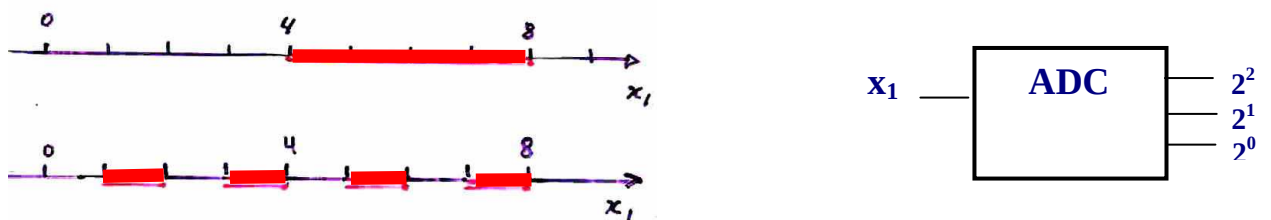
#### Critérios de classificação

Critérios “não conhecidos”, arbitrários

Critérios conhecidos, determinados

#### Classificadores por similaridade

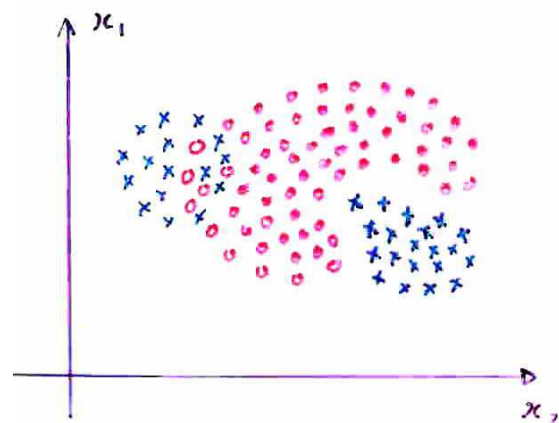
#### Classificações Arbitrárias

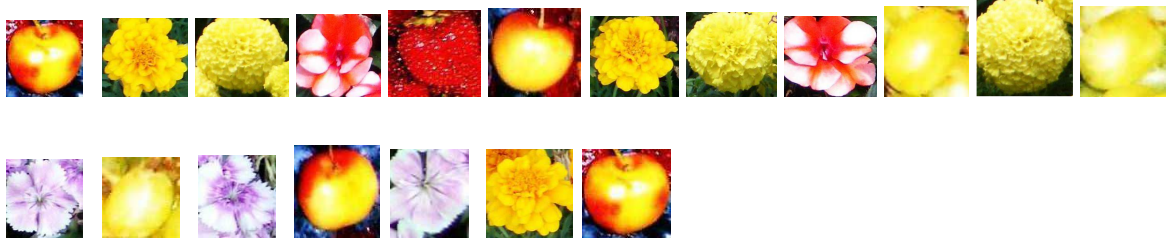


#### Regiões de fusão

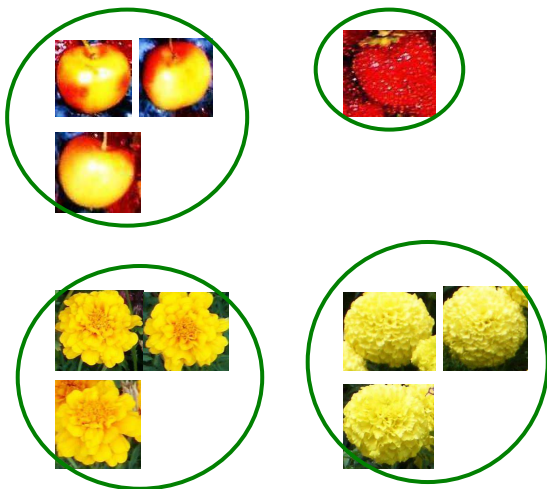
(confusão)

entre classes

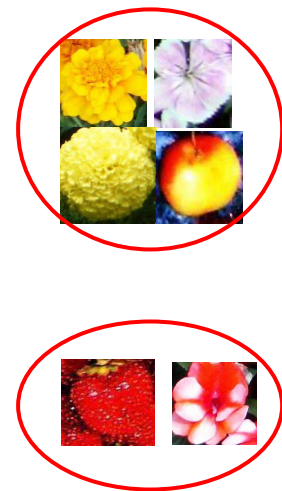




### Agrupamentos por similaridade arbitrários

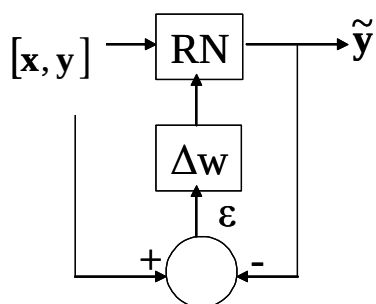


### Agrupamentos

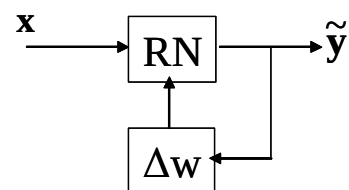


## 3.1 - Tipo de aprendizado

### Treinamento supervisionado / não-supervisionado



**Pode realizar  
classificações arbitrárias**



**Necessita algum critério:  
Similaridade**

## 4 - Classificação, Agrupamento ou Clusterização por Similaridade

Agrupar, “de forma natural”,  
conjuntos de dados com “similaridade interna”

Classes agrupam elementos `similares` entre si ‘de forma natural’ mas

**O que significa “de forma natural ?”**

**O que significa “similaridade interna ?”**

**Quais são as classes ?**

**Classificadores supervisionados**

**As classes são informadas**

**Ação: apenas classifica as entradas**

**Classificadores **não** supervisionados**

**As classes **não** são informadas**

**Ações:     **determina as classes e**  
  
              **classifica as entradas****

**Sobrevivência vs.****Classificação por similaridade****Aprendizado não supervisionado****Plasticidade****Animais e vegetais venenosos ou comestíveis ?****4.1 – Dados reais vs. Representação matemática na****Classificação por Similaridade**

objetos físicos  $O_1$   $O_2$   $\longrightarrow$  objetos matemáticos  
vetores  $\underline{x}_1$   $\underline{x}_2$

similaridade física  $\longrightarrow$  similaridade matemática

$$O_1 \approx O_2$$

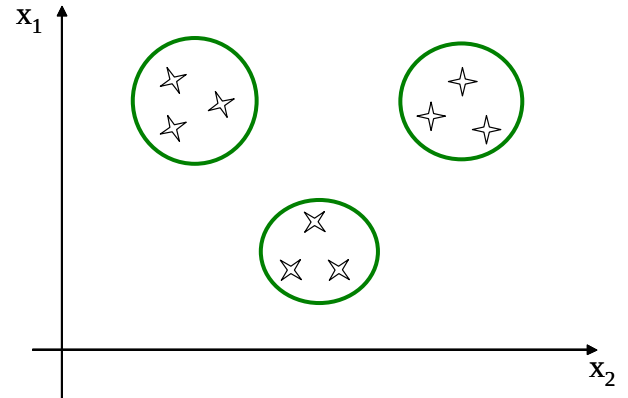
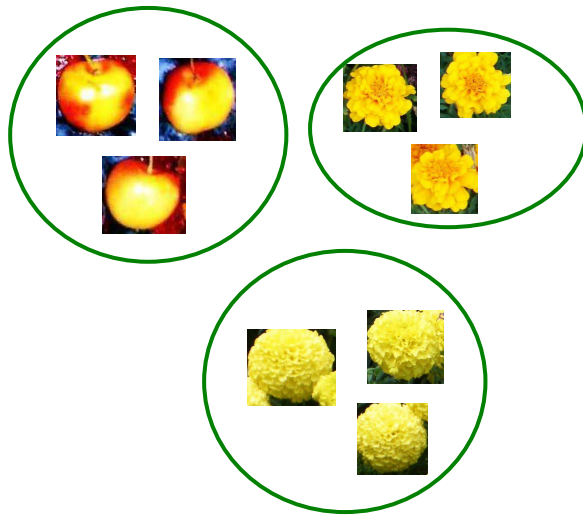
$$\underline{x}_1 \cong \underline{x}_2 \quad \text{ou} \quad |\underline{x}_1 - \underline{x}_2| \ll \ll$$

$$O_1 \approx O_2$$



$$\underline{x}_1 \cong \underline{x}_2$$





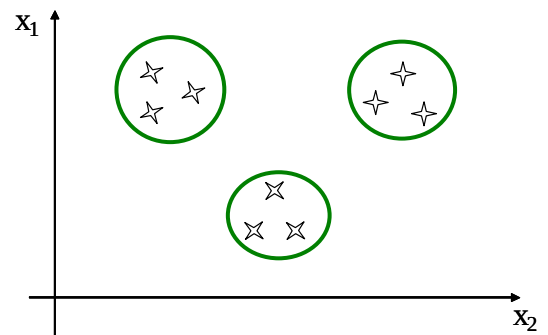
**Classificar por similaridade : alocar na mesma classe  
vetores que estão relativamente próximos entre si.**

**Como traduzir matematicamente**

**“relativamente próximos entre si ?”**

**Como determinar as classes ?**

**Como realizar a alocação ?**



## Noções iniciais sobre

### Clusterização

### Agrupamento ou

### Classificação

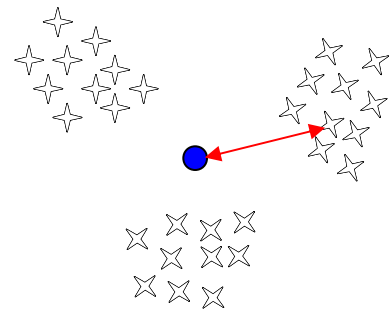
### por similaridade

## 5 - Características de um Agrupamento

### 5.1 - Características do conjunto de todos os elementos

#### Média

$$\vec{m} = E_{\forall i} \vec{x}_i = \frac{1}{N} \sum_{i=1}^N \vec{x}_i$$



#### Dispersão, dissimilaridade total

$$F_0 = \sum |\vec{x}_i - \vec{m}|^2$$

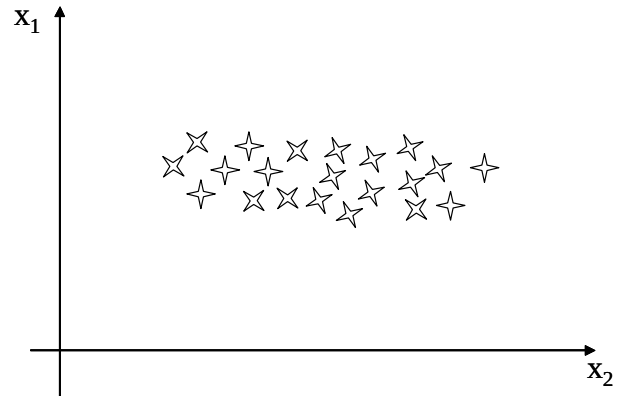
$\underline{m}$  e  $F_0$  são independentes da clusterização

## 5.2 Características de uma classe $C_j$

### Características intra classe

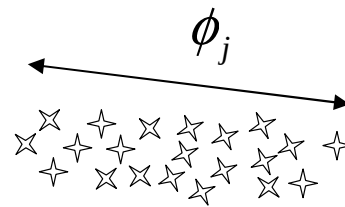
Classe  $C_j$

número de elementos  $n_j$



Diâmetro da classe:

$$\phi_j = \underset{\forall \vec{x}_i, \vec{x}_k \in C_j}{\text{Max}} (|\vec{x}_i - \vec{x}_k|)$$



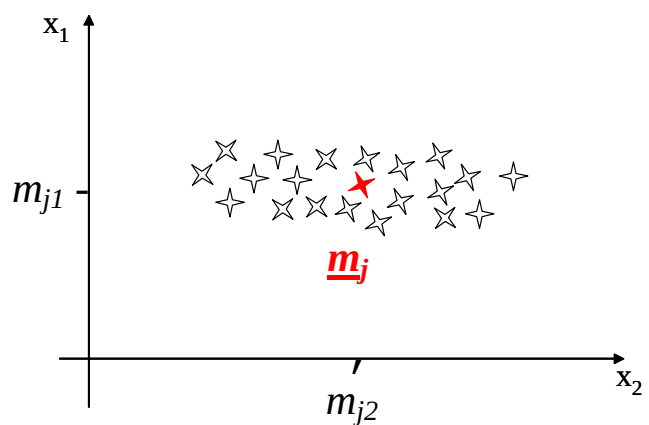
Centro, Baricentro ou Padrão da classe

$$\vec{m}_j = \underset{\forall \vec{x}_i \in C_j}{E} \vec{x}_i = \frac{1}{n_j} \sum_{\forall \vec{x}_i \in C_j} \vec{x}_i$$

$$\vec{m}_j = [m_{j1} \dots m_{jk} \dots]^t$$

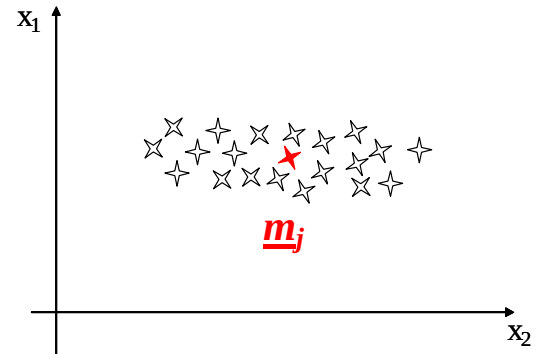
por componente k

$$m_{jk} = \underset{\forall \vec{x}_i \in C_j}{E} x_{ik} = \frac{1}{n_j} \sum_{\forall \vec{x}_i \in C_j} x_{ik}$$



## Dispersão média intra classe, Erro (médio quadrático) de representação

$$\sigma_j^2 = E_{\forall \vec{x}_i \in C_j} \|\vec{x}_i - \vec{m}_j\|^2 = \frac{1}{n_j} \sum_{\forall \vec{x}_i \in C_j} \|\vec{x}_i - \vec{m}_j\|^2$$



por componente k

$$\sigma_{jk}^2 = E_{\forall \vec{x}_i \in C_j} (x_{ik} - m_{jk})^2 = \frac{1}{n_j} \sum_{\forall \vec{x}_i \in C_j} (x_{ik} - m_{jk})^2$$

**e** 
$$\sigma_j^2 = \sum_{\forall k} \sigma_{jk}^2$$

**Obs: Valor do padrão  $p_j$  que minimiza a dispersão intra classe da classe  $j$ :**

$$\sigma_j^2 = \sum_{\forall k} \sigma_{jk}^2 \quad \sigma_{jk}^2 = \frac{1}{n_j} \sum_{\forall \vec{x} \in C_j} (x_{jk} - p_{jk})^2$$

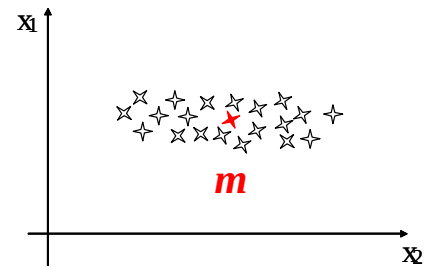
após pequena álgebra

$$\frac{\partial \sigma_{jk}^2}{\partial p_{jk}} = 0 \quad \Rightarrow \quad p_{jk} = m_{jk} \quad \Rightarrow \quad \vec{p}_j = \vec{m}_j$$

**o padrão que minimiza a dispersão intra classe (ou erro de representação) de uma classe é o seu baricentro**

### Dispersão total intra classe da classe $C_j$ :

$$F_j = n_j \sigma_j^2 = \sum_{\forall \vec{x}_i \in C_j} \left\| \vec{x}_i - \vec{m}_j \right\|^2$$

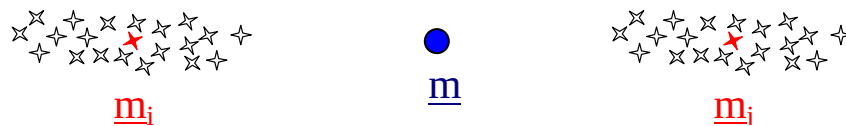


### Dispersão total intra classe para todas as classes:

$$F_{in} = \sum_{\forall C_j} F_j \geq 0$$

**$F_{in}$  - parâmetro à minimizar**

### 5.3 – Características (medidas) Inter Classes



**M – número de classes**

### Dissimilaridade total inter classes

$$F_{out} = \sum_{\forall j} n_j \left\| \vec{m}_j - \vec{m} \right\|^2 \geq 0$$

**$F_{out}$  - parâmetro à maximizar**

**Para um bom agrupamento  
escolher as classes de forma a**

**Minimizar a dispersão intra classe total  $F_{in}$**

**Maximizar a dissimilaridade inter classes total  $F_{out}$**

**Com alguma álgebra é possível mostrar que**

$$F_{in} + F_{out} = F_0 = \text{constante, independente da clusterização}$$

**Logo a clusterização que minimiza  $F_{in}$   
ao mesmo tempo maximiza  $F_{out}$ , i.e.,**

**basta que o processo de clusterização minimize  $F_{in}$**

**Obs 1:**

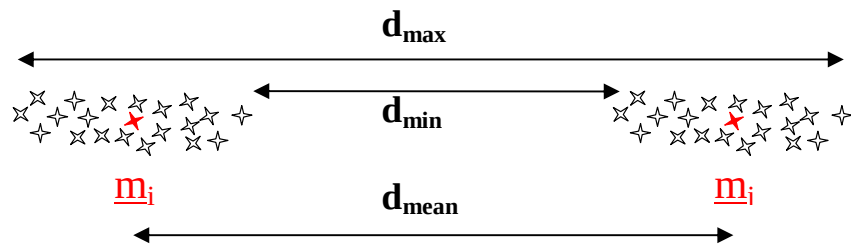
**Usamos como medida inter classes a maximizar a**

**Dissimilaridade total inter classes**

$$F_{out} = \sum_{\forall j} n_j \left\| \vec{m}_j - \vec{m} \right\|^2 \geq 0$$

**mas existem outras medidas de dissimilaridade (distância, separação) inter classes:**

## Outras medidas de dissimilaridade (distância, separação) inter classes:



**vizinho + próximo**

$$d_{\min}(\mathcal{X}_i, \mathcal{X}_j) = \min_{\mathbf{x} \in \mathcal{X}_i, \mathbf{x}' \in \mathcal{X}_j} \|\mathbf{x} - \mathbf{x}'\|$$

**vizinho + distante**

$$d_{\max}(\mathcal{X}_i, \mathcal{X}_j) = \max_{\mathbf{x} \in \mathcal{X}_i, \mathbf{x}' \in \mathcal{X}_j} \|\mathbf{x} - \mathbf{x}'\|$$

**distância média**

$$d_{\text{avg}}(\mathcal{X}_i, \mathcal{X}_j) = \frac{1}{n_i n_j} \sum_{\mathbf{x} \in \mathcal{X}_i} \sum_{\mathbf{x}' \in \mathcal{X}_j} \|\mathbf{x} - \mathbf{x}'\|$$

**distância entre médias**

$$d_{\text{mean}}(\mathcal{X}_i, \mathcal{X}_j) = \|\mathbf{m}_i - \mathbf{m}_j\|.$$

(assim como existem também outras medidas de dispersão intra classe)