

# Introdução às Redes Neurais Artificiais

## *Redes ART Modificada*

Prof. João Marcos Meirelles da Silva

<http://www.professores.uff.br/jmarcos>

Departamento de Engenharia de Telecomunicações  
Escola de Engenharia  
Universidade Federal Fluminense

# Créditos autorais

Este curso e estes slides são parcialmente adaptados da bibliografia citada e das aulas do professor Luiz Pereira Calôba - COPPE/UFRJ

[www.lps.ufrj.br/~caloba](http://www.lps.ufrj.br/~caloba)

# Sumário

- Introdução
- Arquitetura
- Treinamento
- Raio de Vigilância
- Caracterização do Ruído
- Esquecimento e Lembrança
- Classes Não-Esféricas

# Introdução

- A Rede ART modificada é uma versão da Rede de Kohonen aumentada;
- Não estabelece o número de classes “*a priori*” como a Rede de Kohonen, mas um número máximo de classes possíveis;
- Novas classes podem ser criadas para padrões que não são reconhecidos;
- Permite “descobrir” classes não-esféricas.

# Arquitetura

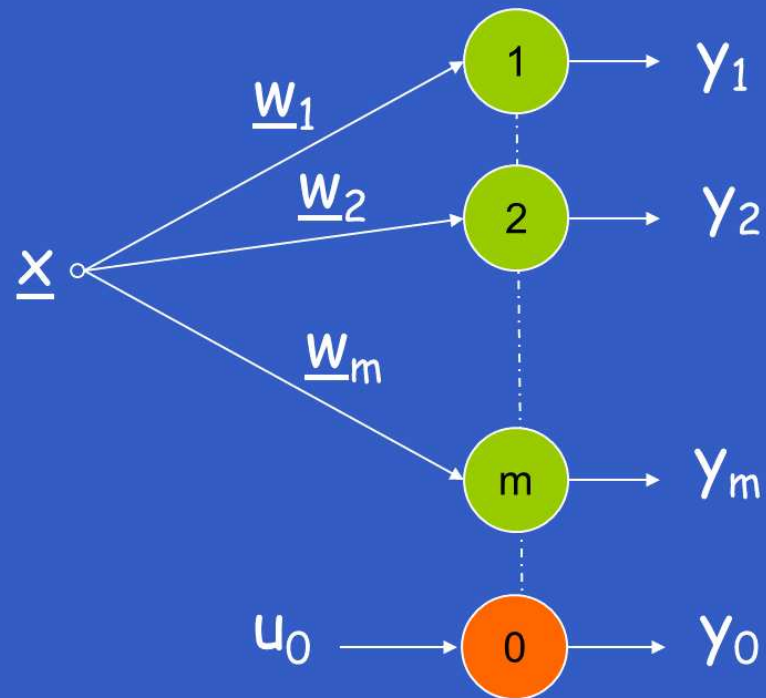


Figura 1: Arquitetura da Rede ART Modificada.

# Treinamento

## 1. Condições Iniciais

$$\underline{w}_1(0) = \underline{x} \quad (\text{aleatório})$$

$$\underline{w}_i(0) = 0; \quad i = 2, \dots, m$$

$m$  é o número máximo de classes desejáveis/possíveis.

# Treinamento

## 3. Treinamento

- Apresentar um padrão  $\underline{x}(n)$  qualquer à rede;
- Neurônio vencedor será aquele com menor valor  $u_j$

$$u_j = -||\underline{x} - \underline{w}_j||, \quad j = 0, \dots, m.$$

Se  $y_0(n) = 1 \rightarrow$  Ativar um neurônio  $i$  desativado

$$\begin{aligned} \underline{w}_i(n+1) &= \underline{x}(n) \\ \underline{w}_j(n+1) &= \underline{w}_j(n), \quad \forall j \neq i \end{aligned}$$

Se  $y_i(n) = 1 \rightarrow$  Treinar o neurônio vencedor  $i$

$$\begin{aligned} \underline{w}_i(n+1) &= \underline{w}_i(n) + \alpha[\underline{x}(n) - \underline{w}_i(n)] \\ \underline{w}_j(n+1) &= \underline{w}_j(n), \quad \forall j \neq i \end{aligned}$$

# Raio de Vigilância

A escolha do valor de  $r_0$  é crítico, pois:

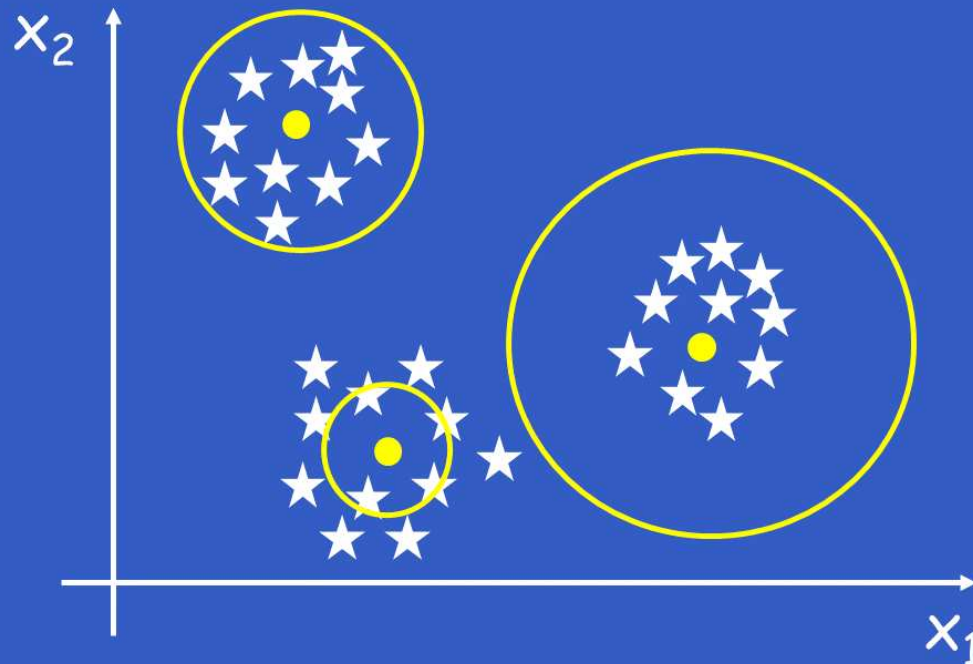


Figura 2: Diferentes escolhas de  $r_0$ .



# Raio de Vigilância

Frequentemente, as classes são geradas a partir de padrões contaminados com ruído aditivo Gaussiano.

$$\underline{x} \in C_k \rightarrow \underline{x} = \underline{w}_k + \underline{r}$$

Logo,

$r_0$  depende do ruído !

# Caracterização do Ruído

- Distância entre entradas: Uma classe

$$\dim x = 1, \quad d_{ij}^2 = |x_i - x_j|^2$$

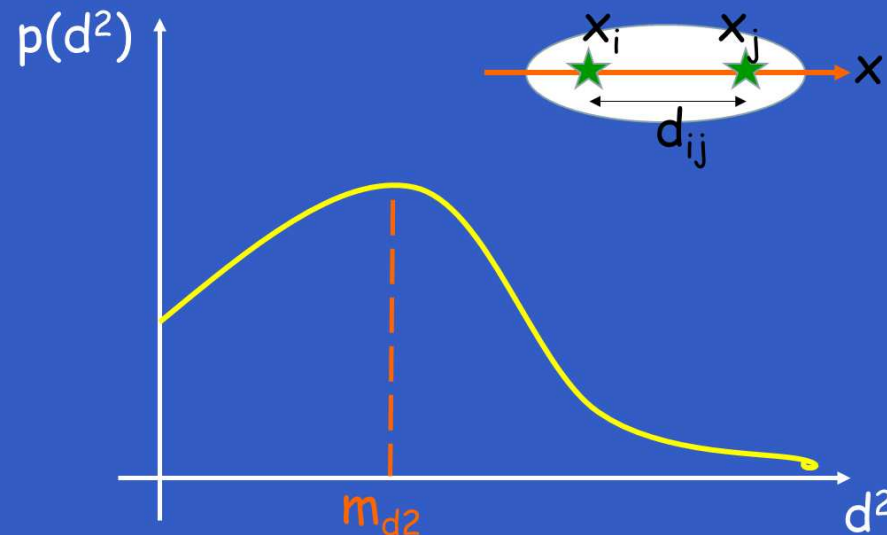


Figura 3: Moda de  $d^2$ .

# Caracterização do Ruído

Moda de  $d^2$ :

$$m_{d^2} \cong \sigma_x^2$$

Logo, para

$$r_0 \approx 2,5\sigma_x, \quad r_0 \approx 2,5\sqrt{m_{d^2}}$$

OBS: Na prática, a Figura 3 assumirá a forma de um histograma devido ao número finito de dados para a pdf.

# Caracterização do Ruído

- Distância entre entradas: Uma classe multidimensional e dispersões iguais por atributo

$$\dim \underline{x} = n, \quad \sigma_{x_j}^2 = \sigma_x^2, \quad \forall j = 1, \dots, n, \quad d_{ij}^2 = |\underline{x}_i - \underline{x}_j|^2$$

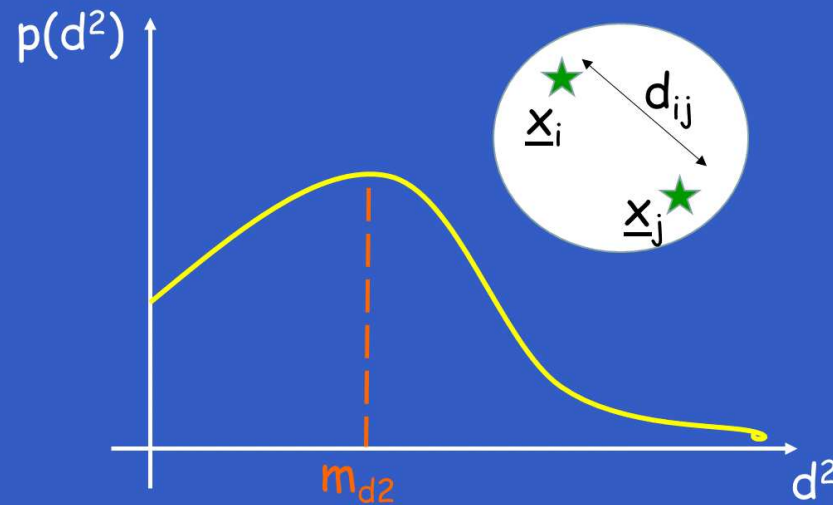


Figura 4: Moda de  $d^2$ .

# Caracterização do Ruído

Moda de  $d^2$ :

$$m_{d^2} \cong \sigma_x^2 n$$

Logo, para

$$r_0 \approx 2,5\sigma_x, \quad r_0 \approx 2,5\sqrt{\frac{m_{d^2}}{n}}$$

# Caracterização do Ruído

- Distância entre entradas: Uma classe multidimensional e dispersões diferentes por atributo

$$\dim \underline{x} = n, \quad \sigma_{x_j}^2 \quad \forall j = 1, \dots, n, \quad d_{ij}^2 = |x_i - x_j|^2$$

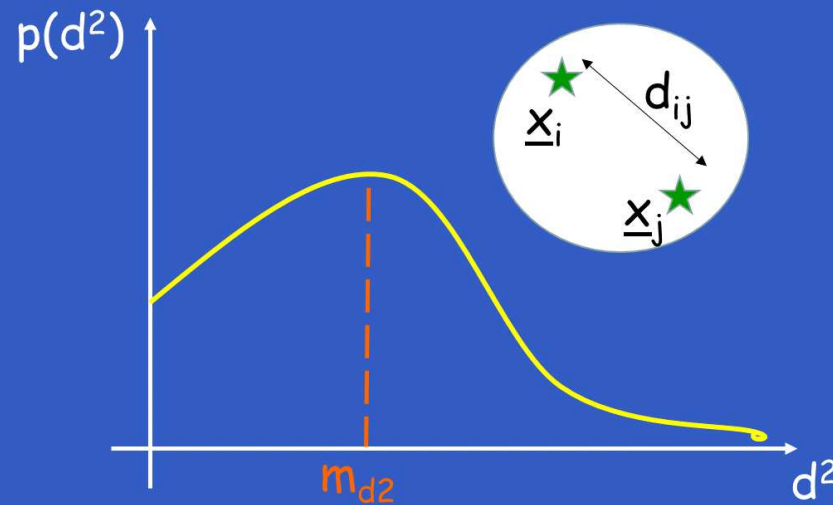


Figura 5: Moda de  $d^2$ .

# Caracterização do Ruído

Moda de  $d^2$ :

$$m_{d^2} \cong \sum_{j=1}^n \sigma_{x_j}^2$$

Calcular para cada atributo  $j$  do vetor  $\underline{x}$  a moda da dispersão de  $d_j^2$ .

$$\text{moda de } d_j^2 = \text{moda de } d_{jkl}^2$$

$$m_{d_j^2} = m_{d_{jkl}^2} = |x_{j_k} - x_{j_l}|^2$$

onde  $\underline{x}_{j_k}$  é a  $j$ -ésima componente de  $\underline{x}_k$ .

$$m_{d_j^2} \cong \sigma_{x_j}^2$$

# Caracterização do Ruído

Logo, para o  $r_0 \approx 2,5\sigma_x$  mais restritivo, na direção de menor dispersão, escolher:

$$r_0 \approx 2,5 \sqrt{\min_{\forall j} \{m_{d_j^2}\}}$$



# Caracterização do Ruído

- Múltiplas classes, dispersão igual por atributo

menor moda  $\Leftrightarrow$  medida intra classe

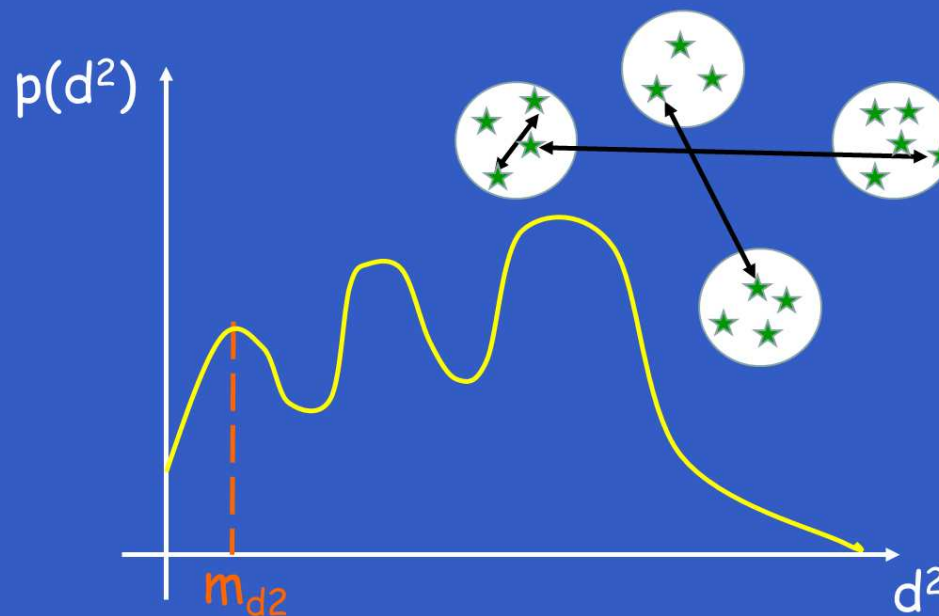


Figura 6: Moda de  $d^2$ .

# Caracterização do Ruído

- Múltiplas classes, dispersões diferentes por atributo

1. Fazer o histograma das distâncias<sup>2</sup> por atributo  $j$ ;

$$d_{jkl}^2 = |x_{jk} - x_{jl}|^2$$

2. Selecionar a menor moda de cada atributo  $m_{d_j^2}$ ;
3. Selecionar a menor de TODAS as modas:

$$\min_{\forall j} \{m_{d_j^2}\}$$

4. Usar:

$$r_0 \approx 2,5 \sqrt{\min_{\forall j} \{m_{d_j^2}\}}$$

# Esquecimento e Lembrança

Se  $y_i = 0$  por  $N$  entradas consecutivas para neurônio ativado  $i$ :

1. Guarde a informação para o futuro como uma lembrança  $\underline{m}_k$ ,  
 $\underline{m}_k = \underline{w}_i$ ;
2. Desative o neurônio  $i$ ,  $\underline{w}_i = 0$ .

Se  $y_0 = 1$ , ativar neurônio, mas verificar primeiro se algum padrão  $\underline{m}_k$  guardado (em disco) serve:

$$-|\underline{x} - \underline{m}_k|^2 > -r_0^2?$$

- SIM  $\rightarrow$  existe lembrança, ativar  $\underline{w}_j = \underline{m}_k$ ;
- NÃO  $\rightarrow$  não existe lembrança, ativar  $\underline{w}_j = \underline{x}$ .

# Classes Não-Esféricas

- Passo 1: Estabelecer raio de similaridade  $r_0$  adequado

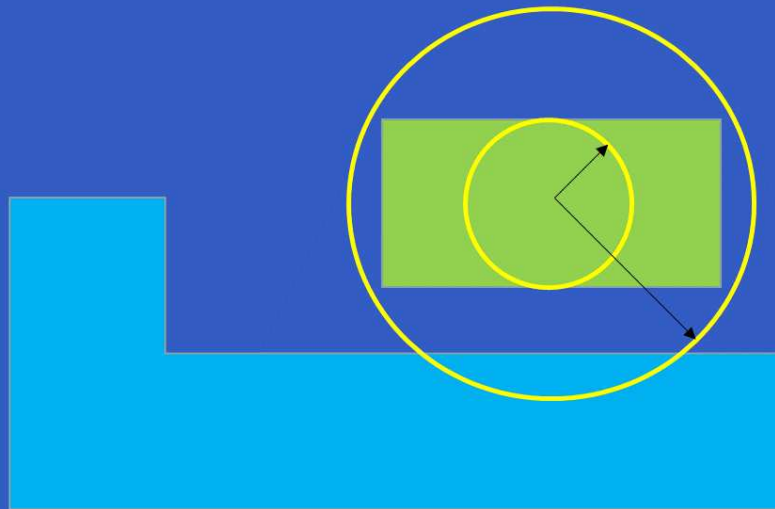


Figura 7: Escolhendo raio de similaridade adequado.

# Classes Não-Esféricas

- Passo 2: Determinar os domínios esféricos

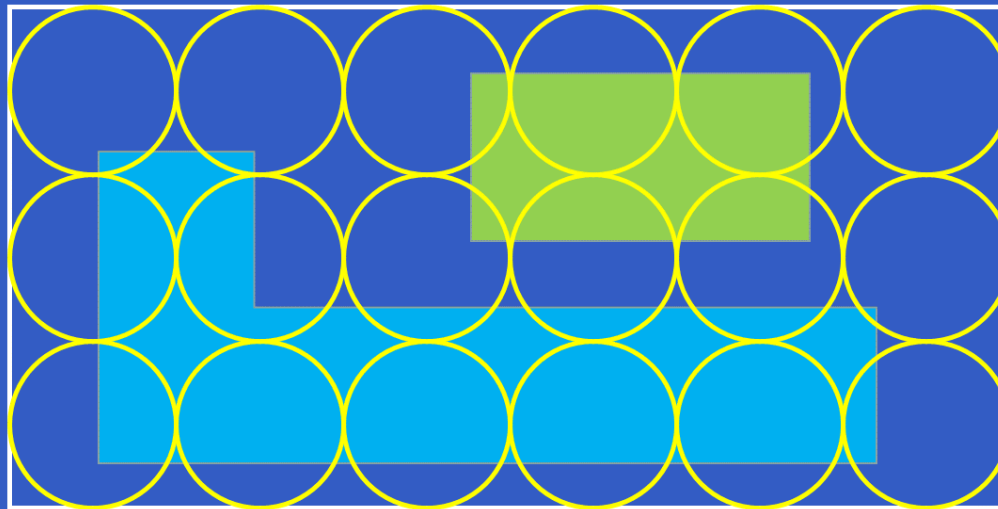


Figura 7: Escolhendo raio de similaridade adequado.

# Classes Não-Esféricas

- Passo 3: Determinar os domínios vizinhos

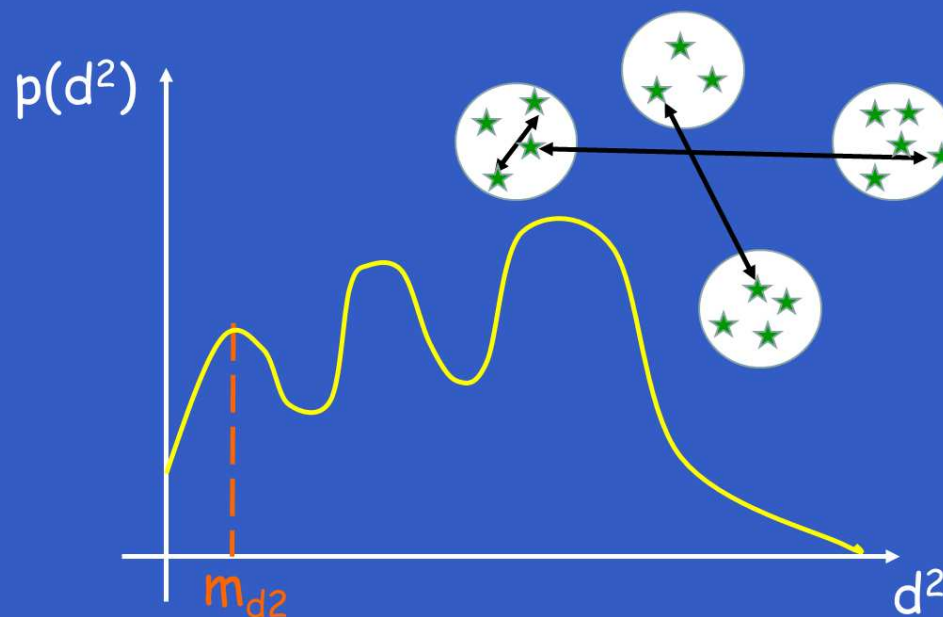


Figura 7: Escolhendo raio de similaridade adequado.

# Classes Não-Esféricas

- Passo 4: Agrupar os domínios vizinhos

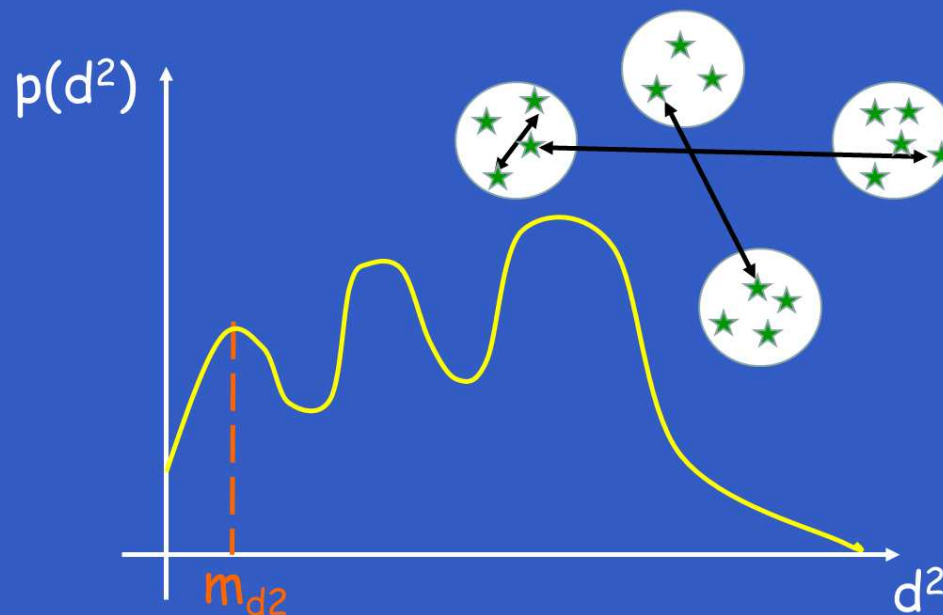


Figura 7: Escolhendo raio de similaridade adequado.