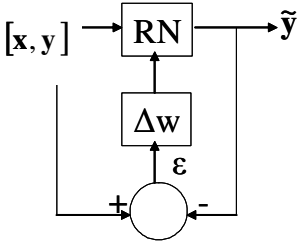
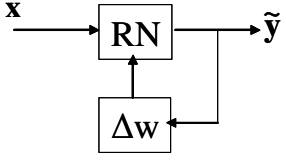
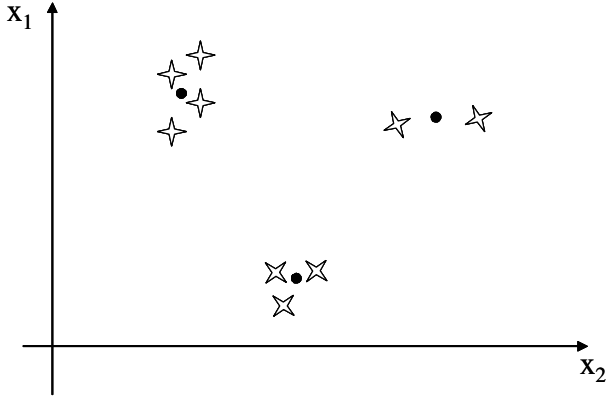
Camada de Kohonen – Treinamento cego

1 - Treinamento supervisionado e não-supervisionado



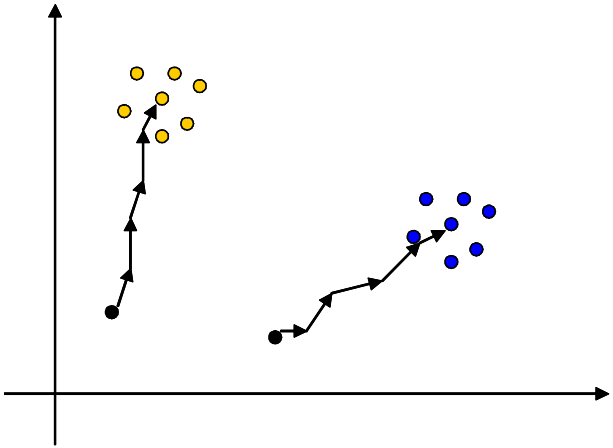


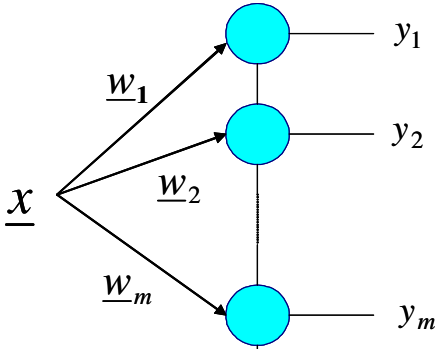


2 - Camada de Kohonen

Classificação por similaridade - Treinamento não supervisionado

-





Camada de Kohonen

Treinamento não supervisionado

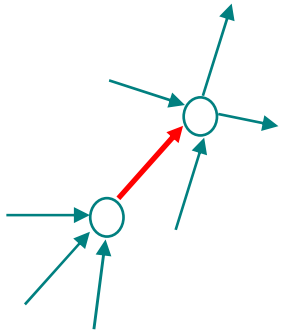
$$\underline{x}(n) \quad >>>> \quad y_i = 1$$

$$\begin{aligned} \underline{w}_i(n+1) &= \underline{w}_i(n) + \alpha[\,\underline{x}(n) - \underline{w}_i(n)\,] \\ &= (1-\alpha)\,\underline{w}_i(n) + \alpha\,\underline{x}(n) \end{aligned}$$

$$\underline{w}_j(n+1) = \underline{w}_j(n) \quad \forall \; j \neq i$$

O treinamento é dito competitivo, porque apenas o neurônio vencedor treina.

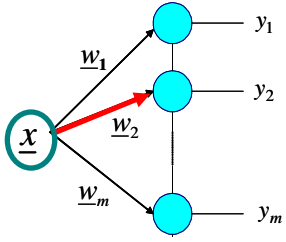
Obs: Este algorítmo pode ser visto como uma aplicação da Regra de Hebb, a mais antiga regra de aprendizado (1949), que pode ser escrita:



"Quando dois neurônios conectados por uma sinapse são ativados simultâneamente (em sincronismo) a sinapse entre eles tende a se fortalecer seletivamente."

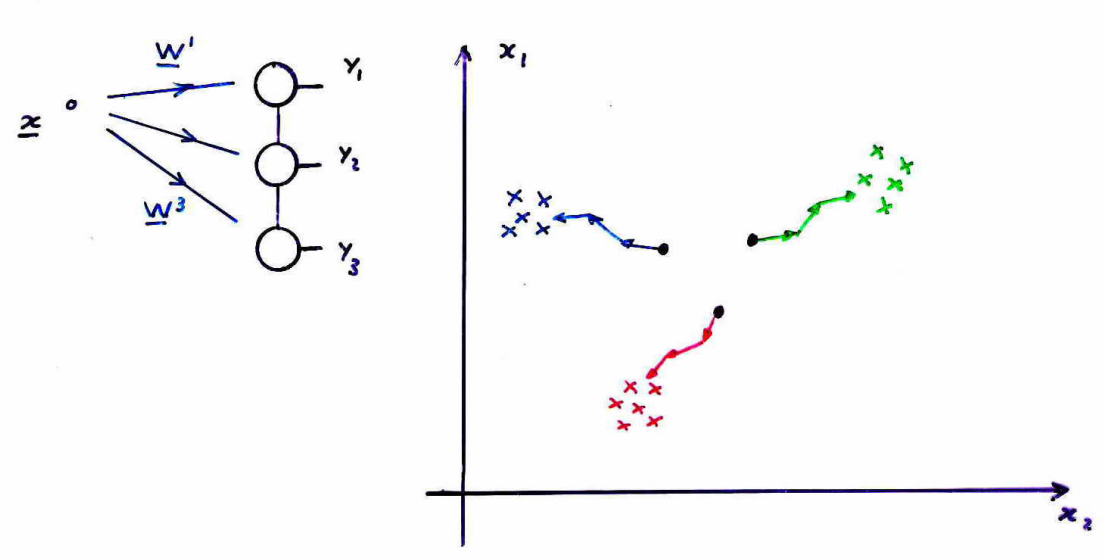
uma segunda parte foi posteriormente acrescentada a esta regra, e será utilizada futuramente para explicar o mecanismo de esquecimento:

"Quando dois neurônios conectados por uma sinapse não são ativados simultâneamente (são assíncronos) a sinapse entre eles tende a se enfraquecer seletivamente ou a ser eliminada."



em nosso caso a entrada atua como um dos neurônios, e a sinapse entre os dois é o vetor sinapse.

Evolução do Treinamento



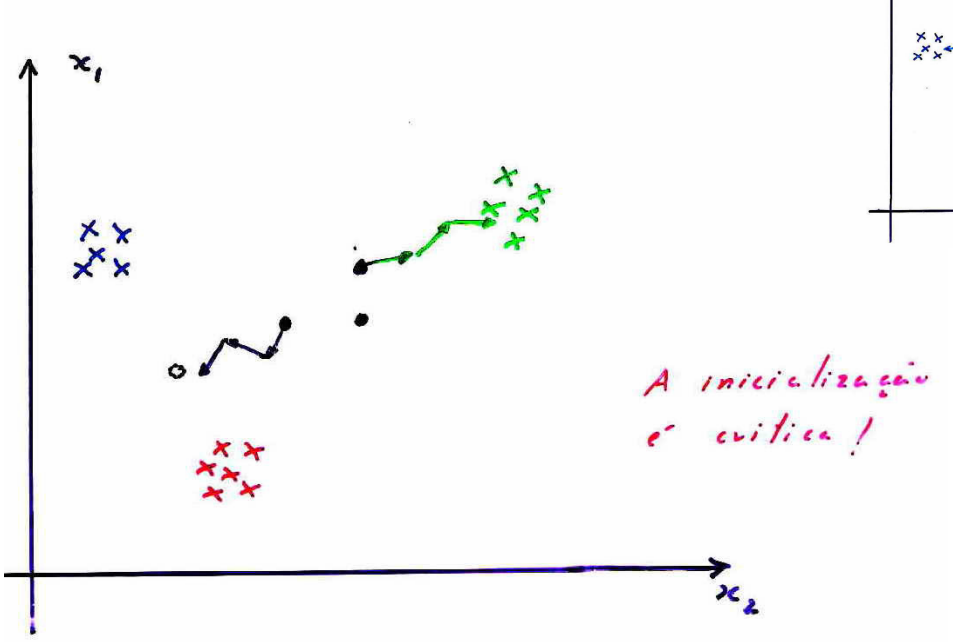
Fim do treinamento ?

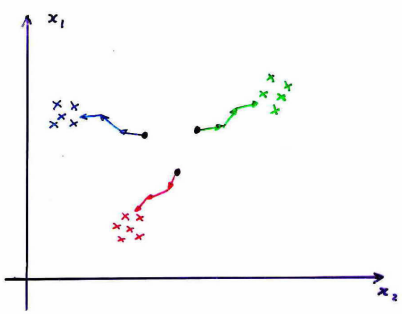
$$E\left[\underline{\Delta w}\right] = \underline{0}$$

Usar um α por vetor sinapse (populações díspares por neurônio)

Realizar o treinamento com decaimento do passo por vetor sinapse

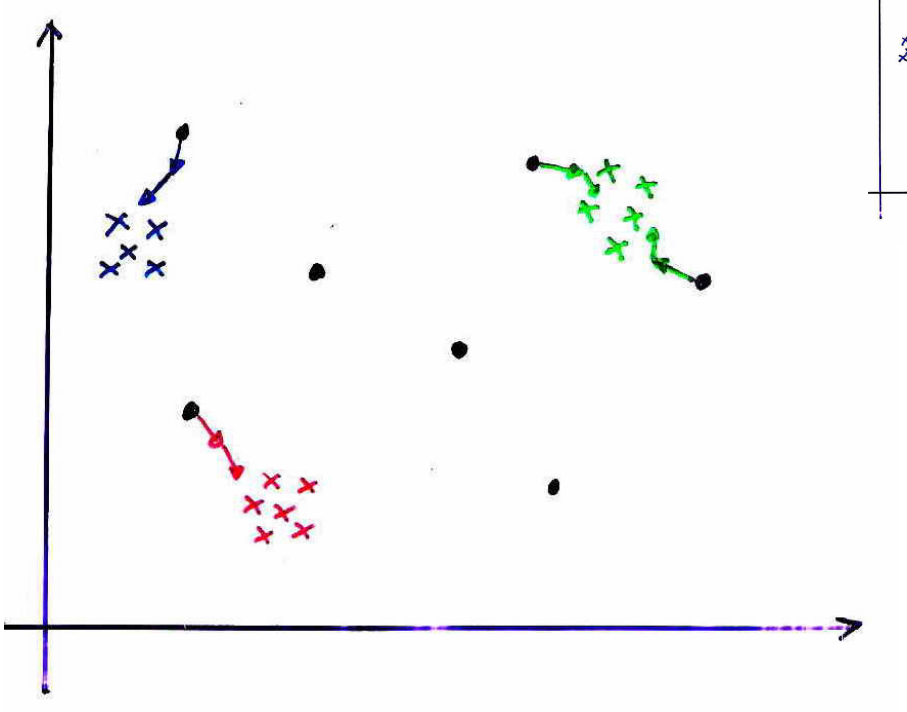
3 - Problemas no treinamento:

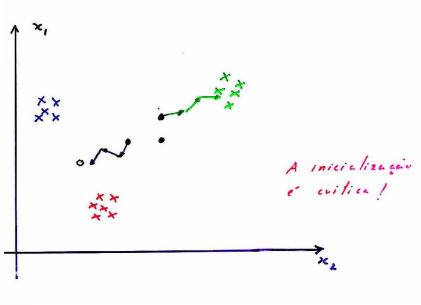




Aumentar o número de sinapses

→ aumentar soft ou hardware





uma classe é representada por mais de um padrão

3.1 - Inicialização

A inicialização é crítica $\underline{w_i}(0)$

- randômicos

- orientados pela população

primeiras entradas $\underline{w_i}(0) = x\ (i)$

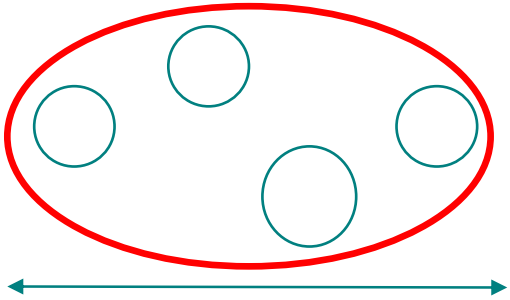
primeiras entradas não muito próximas $\underline{w_i}(0) = \underline{x}(i)$

O que significa "não muito próximas" ?

Domínio de todos os elementos

distância entre entradas

$$d_{ij} = \left| \vec{x}_i - \vec{x}_j \right|$$



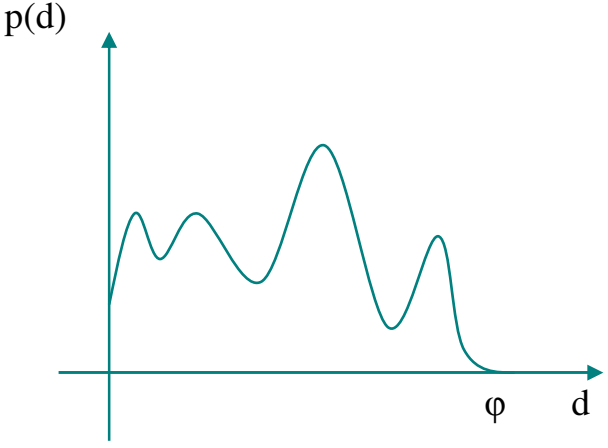
diâmetro da classe única

φ = max (d_{ij})

N classes

Distância aceitável entre sinapses w

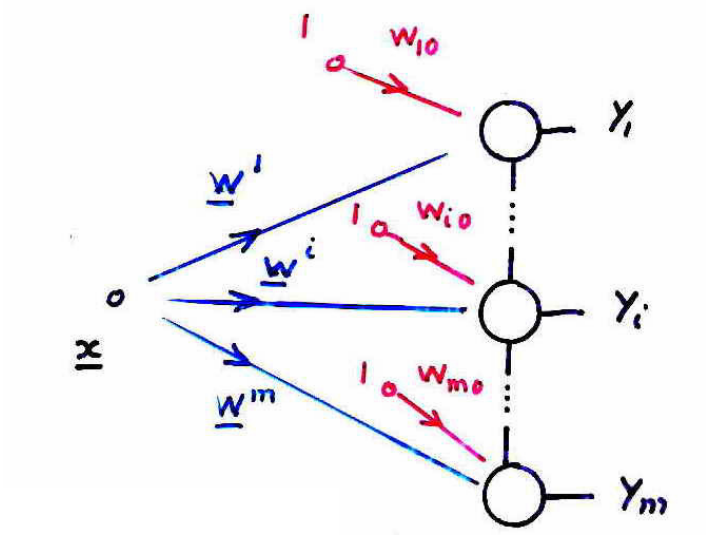
$$d_{\vec{w}} \approx \frac{\varphi}{N}$$



4 - Consciência

O neurônio que treinou muitas vezes abre mão do treinamento para o segundo ganhador.

É adicionado um "handicap" no(s) neurônio(s) muito treinados



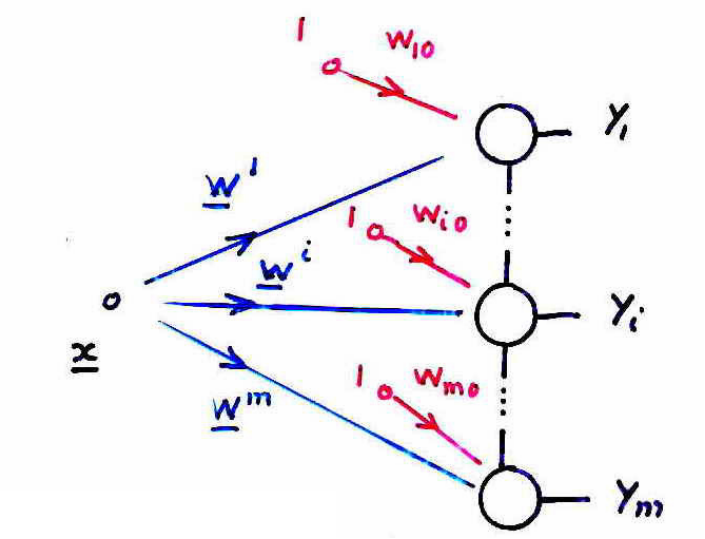
$u_j = - d_j^2 + w_{j0}$ $\quad w_{j0} < 0$

$u_k = - d_k^2 + w_{k0}$

Para $u_j > u_k$

$d_j^2 - w_{j0} < d_k^2 - w_{k0}$

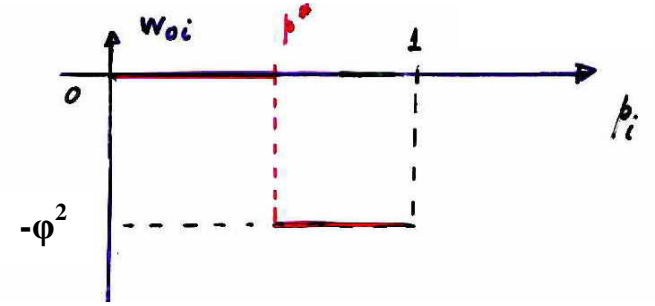
$d_j^2 < d_k^2 - (w_{k0} - w_{j0})$



Formas de consciência:

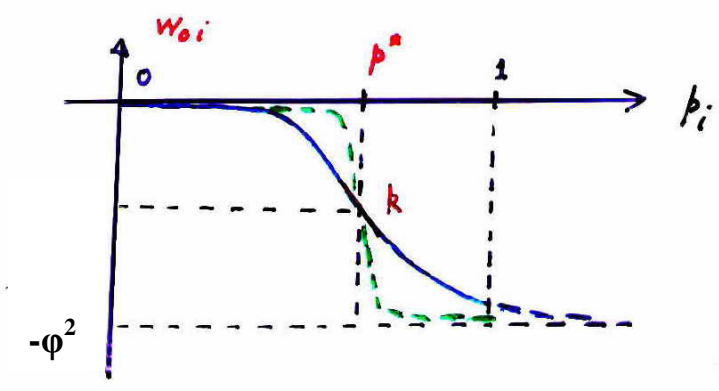
abrupta:

$$p_i \begin{cases} < p^* & w_{i0} = 0 \\ > p^* & w_{i0} = -\varphi^2 \end{cases}$$



suave

$$w_{i0} = -\frac{\varphi^2}{2}\{1 - tgh[k(p_i - p^*)]\}$$



Obs:

1 – Critério para φ: o neurônio i perde sempre

$$\varphi \mid Max(u_i) < Min(u_j) \;\; \forall \;\; i \neq j$$

logo φ pode ser maior ou igual ao diâmetro da classe única que inclui todas as entradas

$$\varphi \geq Max\left|x_k - x_l\right| \quad \forall \;\; k,l$$

2 – Critério para p_i^* : se as populações das classes tem a mesma ordem de grandeza

$$p_i^* \approx {}^{1}\!/_{número\_de\_\sin apses}$$

3 - Obtenção de p_i

$$y_i \in \{0,1\}$$

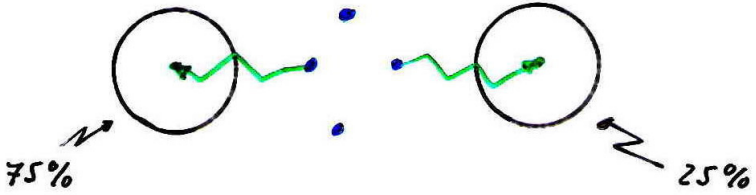
$$p_i(0) = 0$$

$$p_i(n+1) = (1-\alpha)\, p_i(n) + \alpha y_i(n)$$

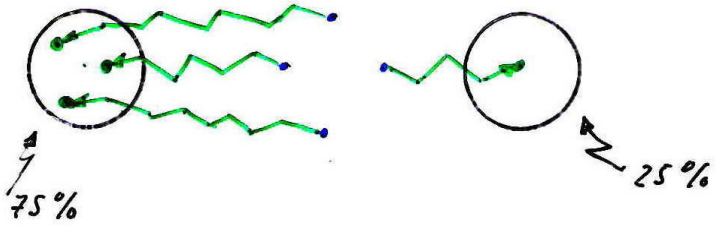
$$p_i(n) \cong \frac{\sum_{j=1}^{n} (1-\alpha)^{n-j}\, y(j)}{\sum_{j=1}^{n} (1-\alpha)^{n-j}} \qquad N \approx \frac{\alpha}{4}$$

5 - Geometria da Classificação:

Sem consciência **por localização geométrica**



Com consciência **por população**

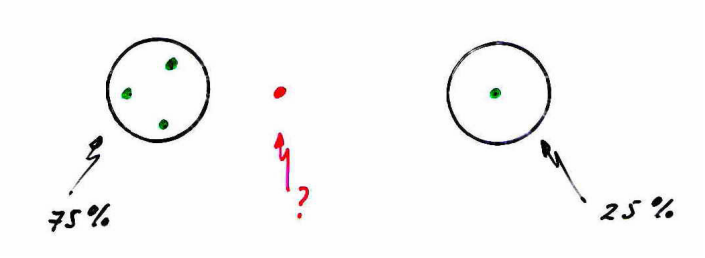


Representação da geometria das classes (sem consciência)

versus

Representação da população (com consciência)

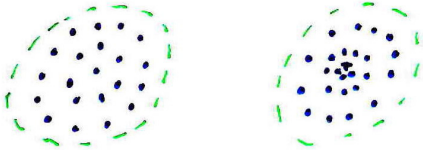
Com poucos neurônios pode ser imprecisa:



Definição de domínios

Caracterização da estatística de $\underline{x}$

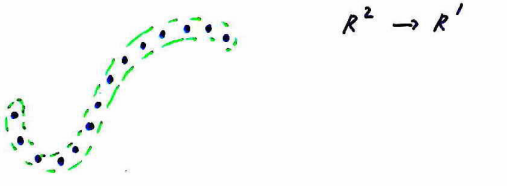
$$p(\underline{x}) = p(\underline{w})$$



Mapas Auto Organizáveis de Kohonen

SOM

Representação em dimensão reduzida



Classificadores

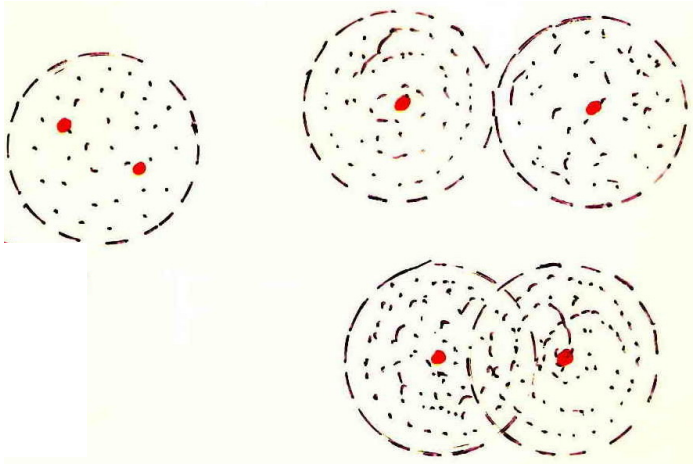
6 - Crítica pós treinamento (fundamental !)

a - Neurônios não (ou pouco) treinados - eliminar

b - Dois neurônios partilham a mesma classe

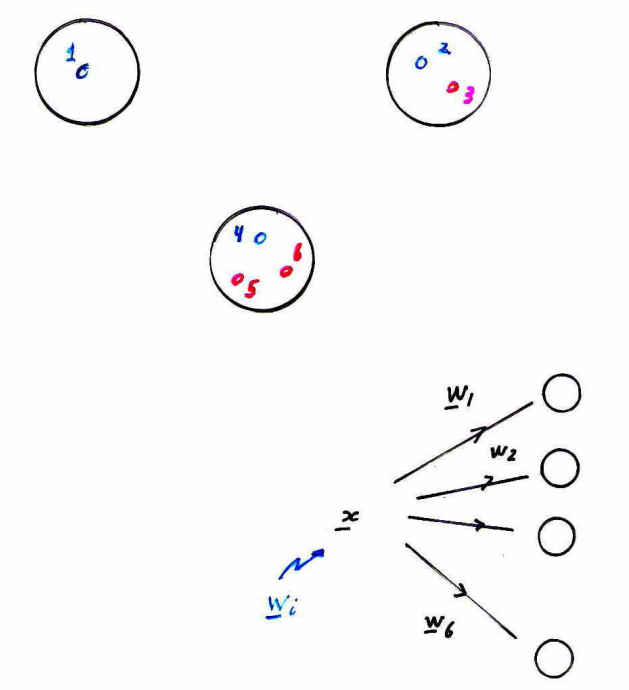
Detecção:

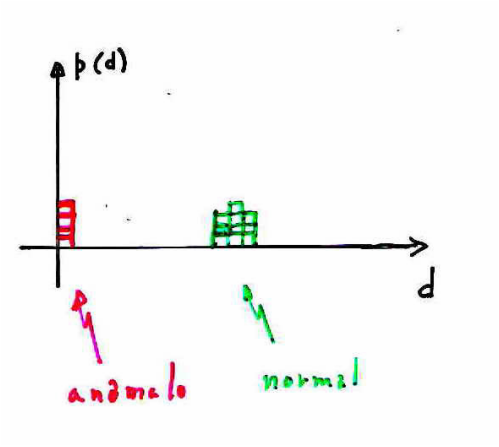
$$\left|\vec{w}_i - \vec{w}_j\right| < 2\,r_0$$



Mais de um neurônio por classe mas r_0 é desconhecido ?

Detecção:





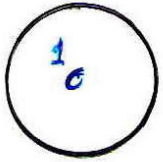
Correção

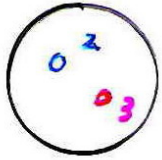
Retreinar ? não obrigatoriamente !

$$\vec{w}_1' = \vec{w}_1$$

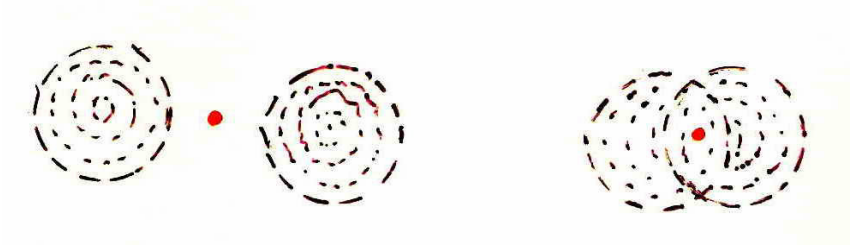
$$\vec{w}_2' = \frac{1}{2}\left(\vec{w}_2 + \vec{w}_3\right)$$

$$\vec{w}_4' = \frac{1}{3}\left(\vec{w}_4 + \vec{w}_5 + \vec{w}_6\right)$$

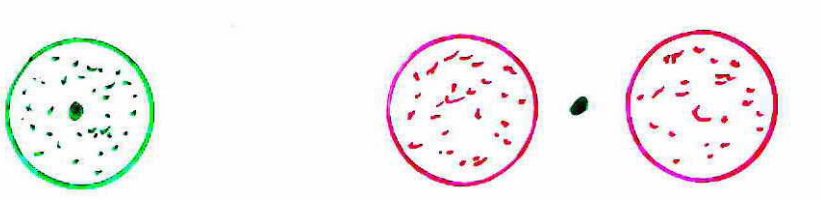




c – Um neurônio atende duas classes

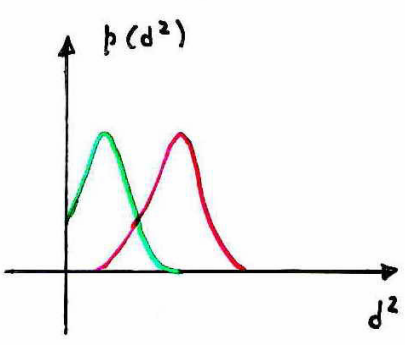


r_0 desconhecido ?



u_i do neurônio vencedor

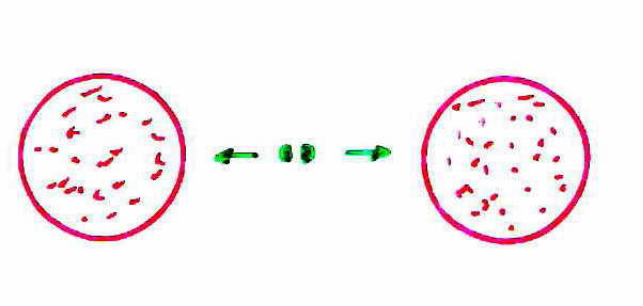
é muito pequeno



Correção:

Duplicar w's anômalos

Retreinar os w's anômalos (duplicados)



7 – Resumo: O que usar ?

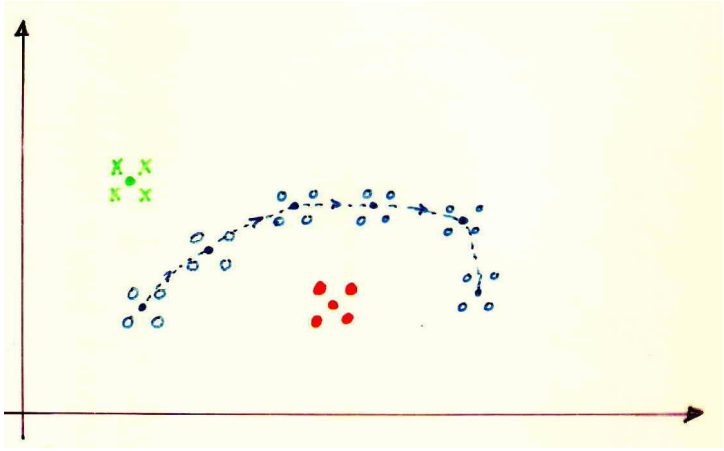
Número sinapses > número esperado de classes

Consciência

Crítica pós treinamento

8 - Sistemas Variantes no tempo:

Variações lentas dos baricentros ? OK



Variações rápidas do baricentro ?

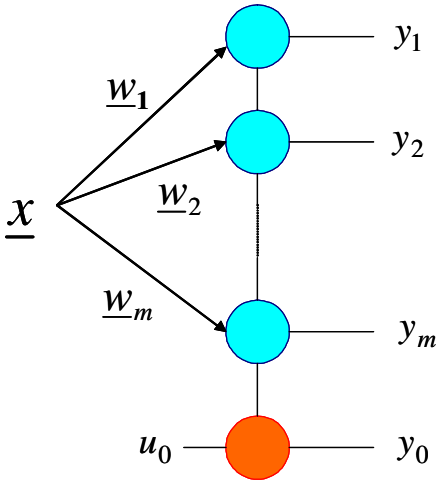
→ novas classes

reinicializar, ao menos parcialmente

9 - E o segundo critério (similaridade mínima) ?

Idêntico ao caso supervisionado:

Camada de Kohonen (aumentada)



$$u_0 = - r_0^2$$

Se $y_i = 1$ entao

$$\underline{x} \in C_i$$

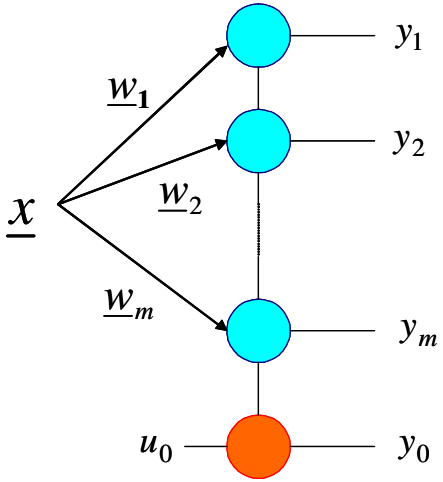
pelos critérios 1 (padrao mais similar a entrada) e

2 (satisfaz a similaridade minima exigida).

Se $y_i = 1$ então

$$\underline{x} \in C_i$$

pelos critérios
1 (centro de classe mais similar à entrada) e
2 (satisfaz à similaridade mínima)



Se $y_0 = 1$ então

$$\underline{x} \notin C_i \quad \forall\, i$$

$\underline{x}$ não satisfaz ao critério 2
para nenhuma classe