

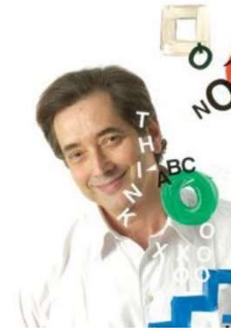
Rede ART (Adaptive Resonance Theory) (modificada)

Condições iniciais:

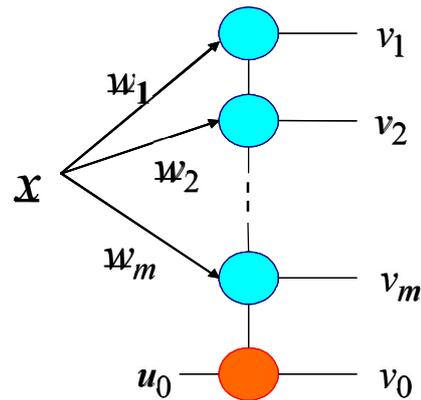
$$u_0 = -r_0^2 \quad e$$

todos os neurônios $\underline{w}_i(0)$
estão desativados

Obs: neurônio desativado não
tem condição de ganhar a competição



Stephen Grossberg
(neurocientista)



Neurônio \underline{w}_i desativado

(não tem condição de ganhar a competição) :

Não existe na rede (melhor)

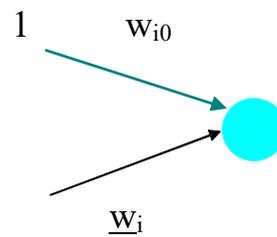
ou

$$|\vec{w}_i|^2 > \text{Max}|\vec{x}|^2 + r_0^2$$

ou

usar bias $w_{i0} = -r_0^2$

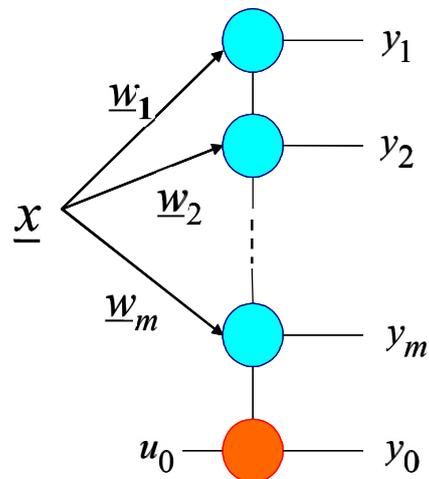
$$w_{i0} = \begin{cases} 0 & \text{neurônio ativo} \\ u_0 = -r_0^2 & \text{neurônio desativado} \end{cases}$$



Treinamento:

Apresentar entrada $\underline{x}(n)$

Neurônio vencedor y_0 ou y_i



Se $y_0(n) = 1$

Ativar um neurônio i desativado

$$\underline{w}_i(n) \text{ desativado} \Rightarrow \underline{w}_i(n+1) = \underline{x}(n)$$

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

Se $y_i(n) = 1$

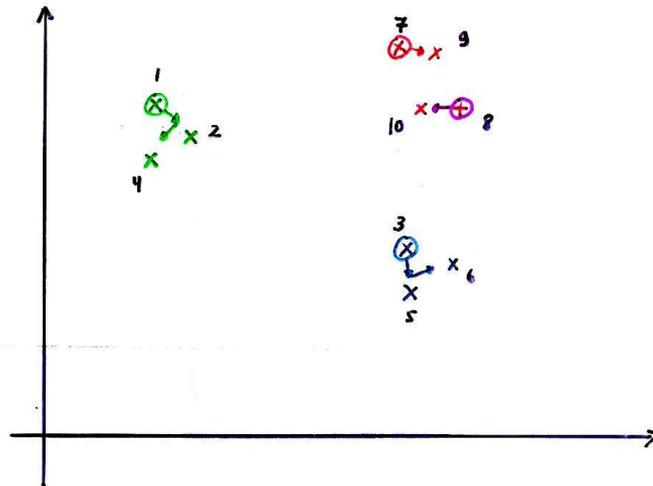
Treinar o neurônio vencedor, i

$$\underline{w}_i(n+1) = \underline{w}_i(n) + \alpha [\underline{x}(n) - \underline{w}_i(n)]$$

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

Plasticidade, Treinamento competitivo

Treinamento ART



este processo garante que o valor inicial da sinapse está dentro do cluster. Não há clusters sem padrão, mas um cluster pode ter mais de um padrão representando-o

Treinamento ART

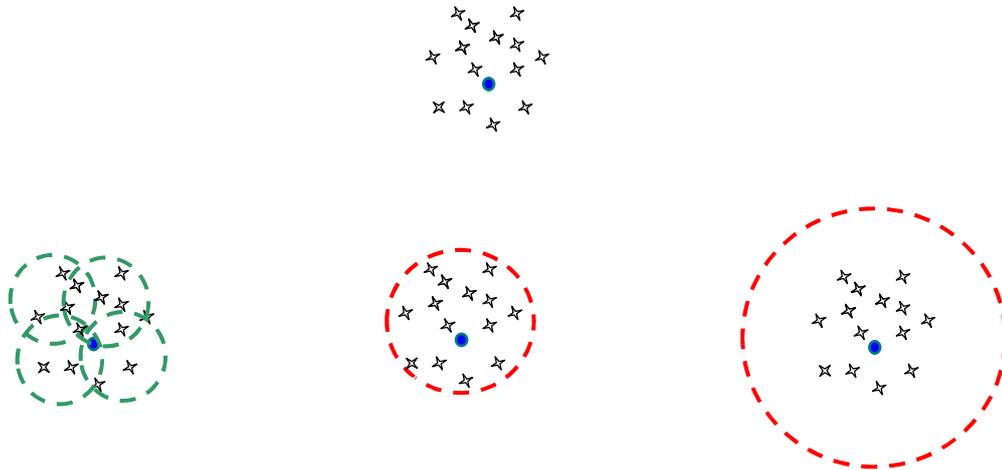
INPUT	EXEMPLARS AFTER EACH INPUT
C	C
E	C E
F	C E F
F	C E F
F	C E F F

Figure 11. An example of the behavior of the Carpenter Grossberg net for letter patterns. Binary input patterns on the left were applied sequentially starting with the upper "C" pattern. Exemplars formed by top-down connection weights after each input was presented are shown at the right.

Ponto crítico: Granularidade, escolha de $u_0 = -r_0$

**muito pequeno: boa definição do domínio da classe,
 mais de um neurônio por classe**

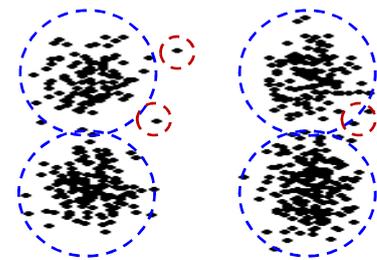
muito grande: definição pobre do domínio da classe



A escolha de r_0 é crítica e depende das características do ruído

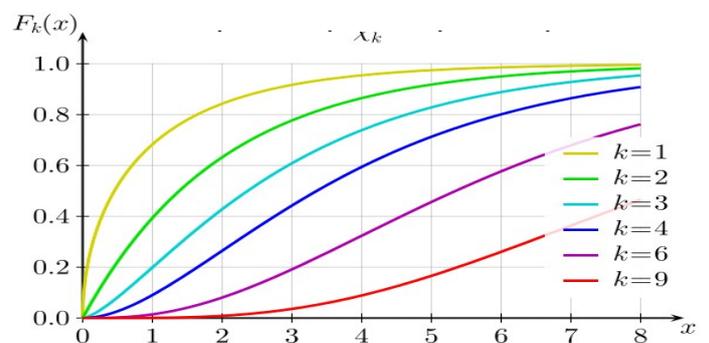
A escolha de r_0 é crítica e depende das características do ruído

Se o ruído em cada dimensão é não correlato e gaussiano com desvio padrão unitário a distribuição de $-u = d^2$ é χ_k^2 e em consequência o raio r_0 deve ser escolhido baseado nesta distribuição



Problema:

criação de padrões espúrios



ART

Esquecimento

pode ser usado na fase de operação e/ou na de treinamento.

Esquecimento (abrupto)

Se $y_i = 0$ por N entradas consecutivas

Guarde a informação para o futuro como
uma lembrança $\underline{\mathbf{m}}_k$, $\underline{\mathbf{m}}_k = \underline{\mathbf{w}}_i$ (opcional)

Desative o neurônio i, $\underline{\mathbf{w}}_i$

Esquecer elimina padrões não mais usados, diminui o número de sinapses e reduz o trabalho computacional

No treinamento o esquecimento pode ser usado para eliminar automaticamente padrões espúrios, porque inicialmente estes são muito pouco ativados.

Evolução de padrões múltiplos por classe

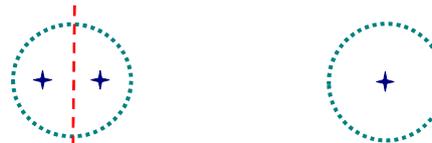
Posições iniciais



Provável posição final
(sem esquecimento)



Provável posição final
(com esquecimento)



ART

Lembrança

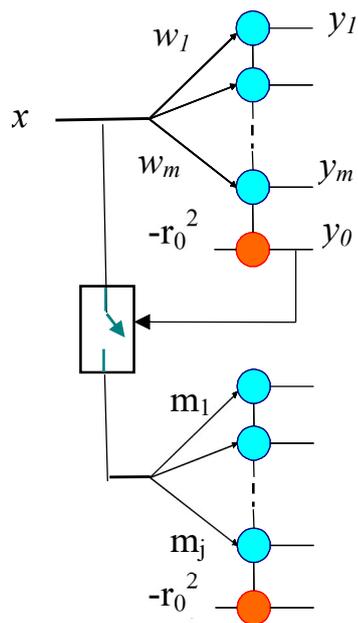
Se $y_0 = 0$ Ativar neurônio, mas

Verifique se algum padrão $\underline{\mathbf{m}}_k$ estocado serve

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

sim = existe lembrança $\underline{\mathbf{m}}_k$, ative $\underline{\mathbf{w}}_j = \underline{\mathbf{m}}_k$

não = não existe lembrança, ative $\underline{\mathbf{w}}_j = \underline{\mathbf{x}}$



Memória Principal

Lembranças

Esquecimento / Lembrança graduais

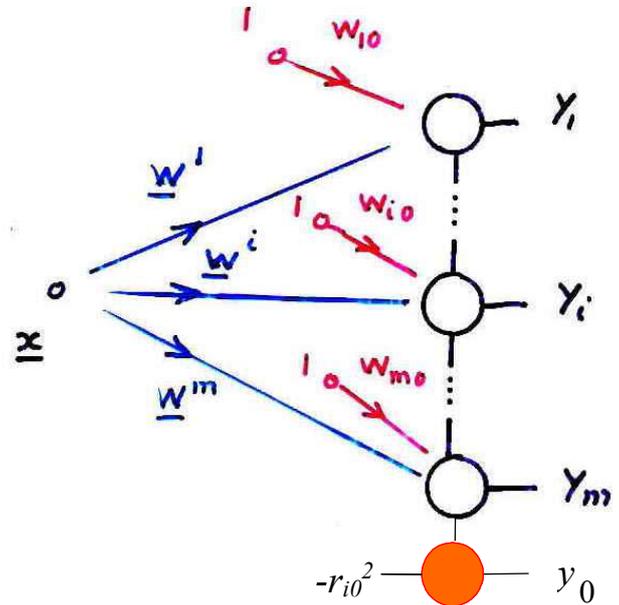
funciona como uma “consciência invertida”, neurônios pouco ativados vão tendo sua possibilidade de vencer reduzida e os muito ativados aumentada.

É adicionado um bias w_{i0} á cada neurônio, que representará um “handicap” para o(s) neurônio(s) pouco ativados e uma vantagem para os muito ativados

Isto que corresponde a alterar o raio de vigilância do neurônio para r_{i0}

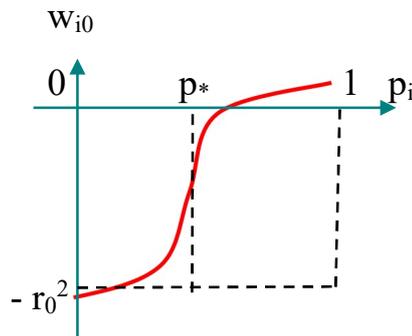
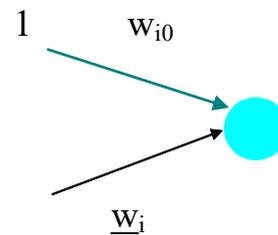
$$r_{i0}^2 = r_0^2 + w_{i0}$$

$w_{i0} < 0$ reduz o raio, $w_{i0} > 0$ aumenta o raio



w_{i0} deve variar segundo uma sigmóide em função do percentual de ativações (vitórias) do neurônio i

$$w_{i0} = a + b \operatorname{tgh} [c(p_i - p_*)]$$



ainda necessita maior desenvolvimento