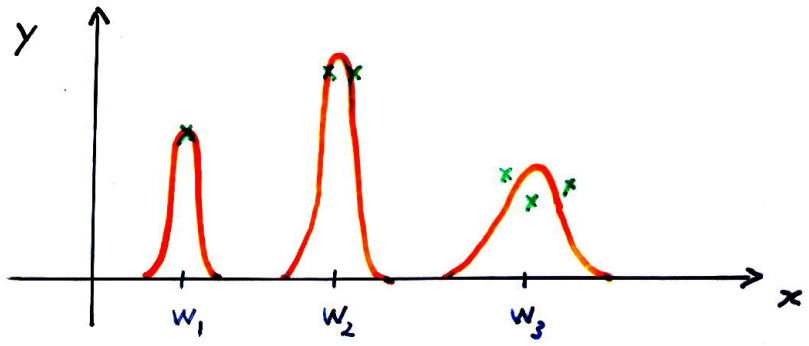
RBF - Radial Basis Function

$y = f(x)$ $\qquad (x_i, y_i)$



$f_i(x) = f_i(v_i)$ $\qquad v_i = |x - w_i|^2$

$f_i \qquad e^{-v_i / 2\sigma_i^2}$

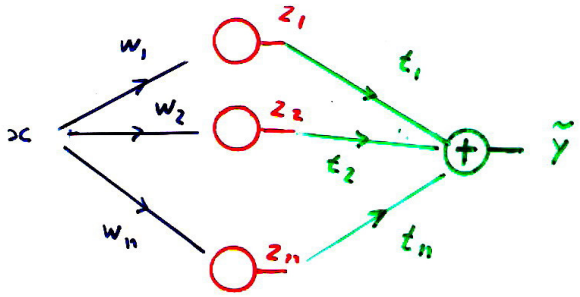
$$\frac{1}{1 + v_i / \sigma_i^2}$$

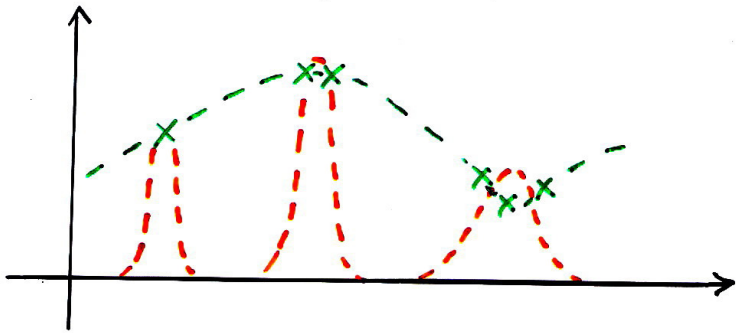
$f_i \in [0,1]$

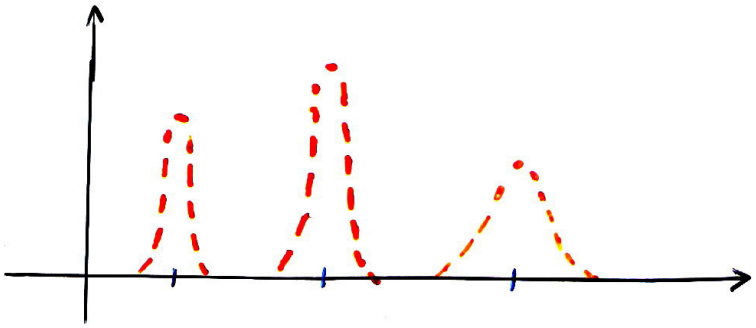
$$\begin{cases} x \simeq w_i & f_i \simeq 1 \\ x \neq w_i & f_i \simeq 0 \end{cases}$$

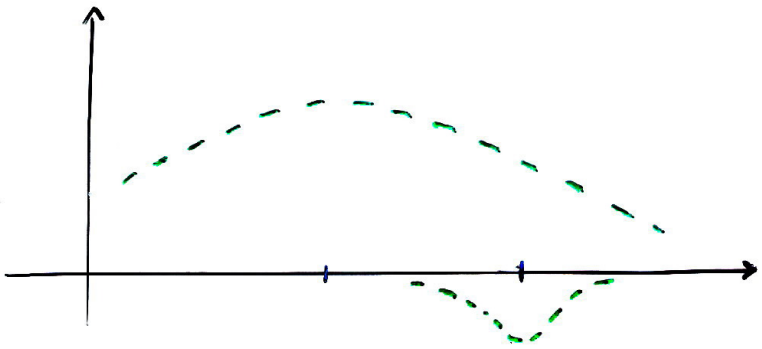
$$z_i = f_i(x) \qquad \begin{cases} \simeq 1 & x \simeq w_i \\ \simeq 0 & x \neq w_i \end{cases}$$

$$y_i = \sum_i t_i z_i \qquad \begin{cases} \simeq t_i & x \simeq w_i \\ \simeq 0 & x \neq w_i \quad \forall i \end{cases}$$

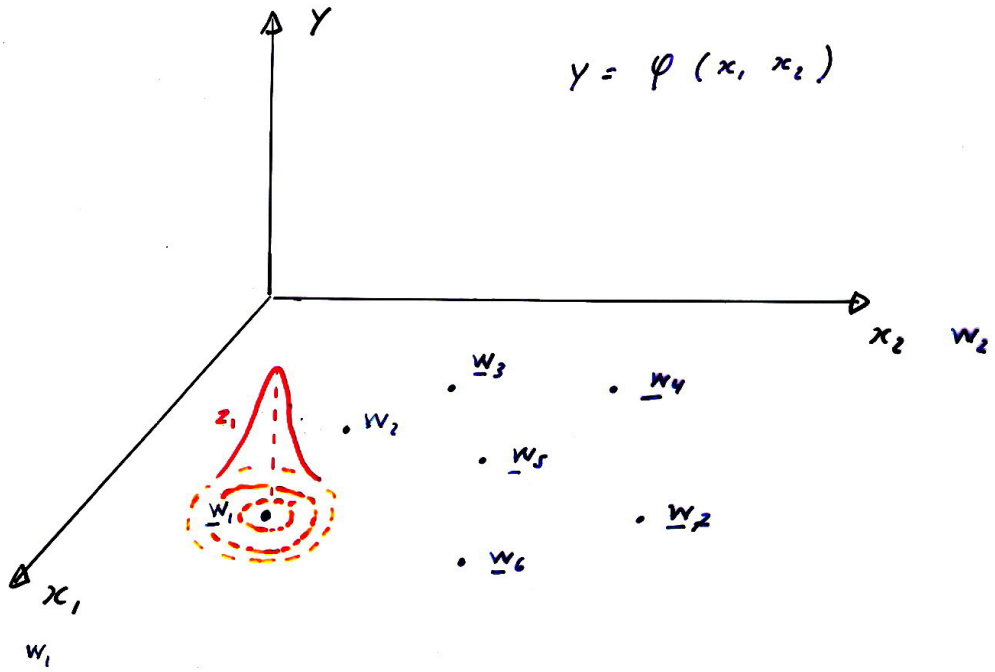


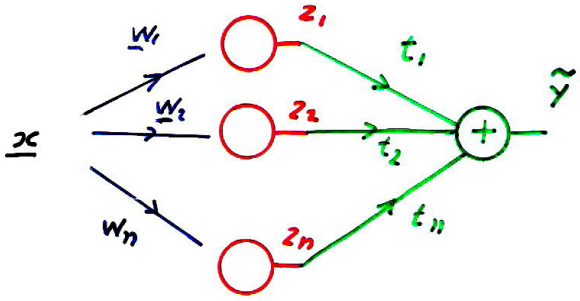






RBF - Radial Basis Function





$$v_i = d_i^2 = \| x - w_i \|^2$$

$$z_i = e^{-\frac{v_i}{2\sigma_i}}$$

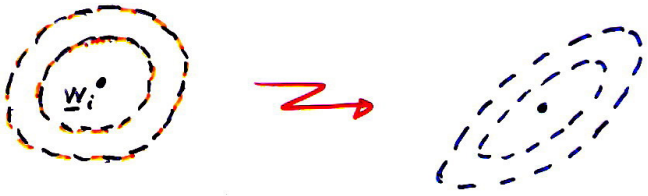
$$\tilde{y} = \sum t_i z_i$$

Generalização

$$v_i = \|\underline{x} - \underline{w}_i\|^2$$

$$v_i = (\underline{x} - \underline{w}_i)^t \underline{K}_i (\underline{x} - \underline{w}_i)$$

$$= \underline{d}_i^t \underline{K}_i \underline{d}_i$$



$$\underline{d}_i = \underline{x} - \underline{w}_i = d_\ell^i \,]$$

$\underline{K}_i = \underline{I}$ $\quad v_i = \underline{d}_i^t \underline{I} \underline{d}_i = \underline{d}_i^t \underline{d}_i = |\underline{d}_i|^2$

i.e. $v_i = \text{ctte} \Rightarrow$ esfera

$\underline{K}_i$ diagonal + $\quad v_i = \sum_\ell k_{\ell\ell} d_\ell^2$

elipsoide com eixos alinhados

$\underline{K}_i$ triangular superior $\quad v_i = \sum_\ell \sum_{m \geq \ell} k_{\ell m} d_\ell d_m$

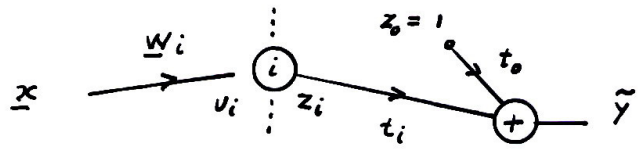
+

$k_{\ell m} = 0 \Leftarrow m < \ell$ $\quad$ elipsoide

Parametros

$t_i \qquad k^i_{lm} \qquad w^i_l$

Treinar por BP !



$F = E(\varepsilon^2) \qquad \varepsilon^2 = (y - \tilde{y})^2$

$\Delta p = - \alpha \frac{d\varepsilon^2}{dp} = 2 \alpha \varepsilon \frac{d\tilde{y}}{dp}$

$\tilde{y} = \sum_i t_i z_i$

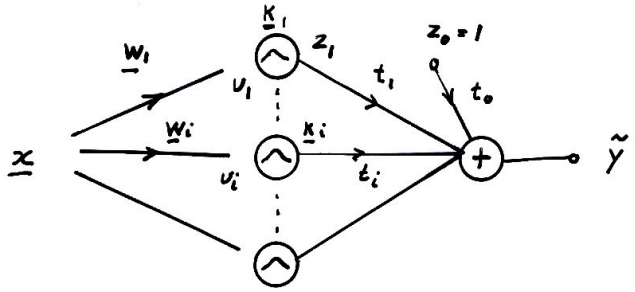
$z_i = e^{- v_i}$

$v_i = \underline{d}_i^t \underline{K}_i \underline{d}_i = \sum_l \sum_{m \geq l} k^i_{lm} d^i_l d^i_m$

$d^i_l = x_l - w^i_l$

Signal Feedforward &

Derivadas Parciais



$$\underline{d}_i = \underline{x} - \underline{W}_i$$

$$d_l^i = x_l - w_l^i \qquad \frac{d d_l^i}{d w_l^i} = -1$$

$$v_i = \underline{d}_i^t \, \underline{K}_i \, \underline{d}_i = \sum_l \sum_{m \geq l} k_{lm}^i \, d_l^i \, d_m^i$$

$$s_l^i = \sum_{m \geq l} k_{lm}^i \, d_m^i$$

$$v_i = \sum_l d_l^i \, s_l^i$$

$$\frac{d v_i}{d k_{lm}^i} = d_l^i \, d_m^i \qquad \frac{d v_i}{d d_l^i} = k_{ll}^i \, d_l^i + s_l^i$$

$$z_i = e^{-v_i} \qquad \frac{dz_i}{dv_i} = -z_i$$

$$\tilde{y} = \sum_{i=0} t_i z_i \qquad \frac{d\tilde{y}}{dt_i} = z_i$$

$$z_0 = 1$$

$$\frac{d\tilde{y}}{dz_i} = t_i$$

Error feedback

$$\Delta t_i = 2\alpha\varepsilon \frac{d\tilde{y}}{dt_i}$$

$$= 2\alpha\varepsilon z_i$$

$$\Delta k_{lm}^i = 2\alpha\varepsilon \frac{d\tilde{y}}{dk_{lm}^i} = 2\alpha\varepsilon \frac{d\tilde{y}}{dz_i}\frac{dz_i}{dv_i}\frac{dv^i}{dk_{lm}^i}$$

$$= 2\alpha\varepsilon (t_i)(-z_i)(d_l^i d_m^i)$$

$$= -2\alpha\varepsilon t_i z_i d_l^i d_m^i$$

$$\Delta w_l^i = 2\alpha\varepsilon \frac{d\tilde{y}}{dw_l^i} = 2\alpha\varepsilon \frac{d\tilde{y}}{dz_i}\frac{dz_i}{dv_i}\frac{dv_i}{dd_l^i}\frac{dd_l^i}{dw_l^i}$$

$$= 2\alpha\varepsilon (t_i)(-z_i)(d_l^i k_{ll}^i + s_l^i)(-1)$$

$$= 2\alpha\varepsilon t_i z_i (d_l^i k_{ll}^i + s_l^i)$$

Caso particular

$$\underline{K}_i = k_i \underline{I}$$

$$d_{ij} = x_j - w_{ij} \qquad \frac{d}{dw_{ij}} d_{ij} = -1$$

$$D_i^2 = \sum_j d_{ij}^2 \qquad \frac{d}{d_{ij}} D_i^2 = 2\, d_{ij}$$

$$v_i = k_i D_i^2 \qquad \frac{dv_i}{dk_i} = D_i^2 \qquad \frac{dv_i}{dD_i^2} = k_i$$

$$z_i = e^{-v_i} \qquad \frac{dz_i}{dv_i} = -z_i$$

$$\tilde{y} = \sum_{i=0} t_i z_i \qquad \frac{d\tilde{y}}{dt_i} = z_i \qquad \frac{d\tilde{y}}{dz_i} = t_i$$

$$\Delta p = -2 \alpha \varepsilon \frac{d\tilde{y}}{dp}$$

$$\frac{d\tilde{y}}{dt_i} = z_i$$

$$\frac{d\tilde{y}}{dk_i} = \frac{d\tilde{y}}{dz_i} \frac{dz_i}{dv_i} \frac{dv_i}{dk_i} = t_i (-z_i) D_i^2$$

$$\frac{d\tilde{y}}{dw_{ij}} = \frac{d\tilde{y}}{dz_i} \frac{dz_i}{dv_i} \frac{dv_i}{dD^2} \frac{dD^2}{dd_{ij}} \frac{dd_{ij}}{dw_{ij}} = t_i (-z_i) k_i 2 d_{ij} (-1)$$

Complexidade — # parâmetros { N entradas 10 ; M neurônios 10 }

$\underline{K}$ triangular superior $\quad M\left[N + \frac{N(N+1)}{2} + 1\right] + 1$ 661

$\underline{K}$ diagonal $\quad M\left[N + N + 1\right] + 1$ 211

$\underline{K}$ identidade $\quad M\left[N + 1 + 1\right] + 1$ 121

Comparando com feed forward clássica (BP)

$$M\left[N + 1 + 1\right] + 1 \qquad 121$$

Atuação dos neurônios { RBF – local ; FF – global }

Simplificações

(ou valores iniciais)

$\underline{W}$ aleatorios

entradas

clusterização (ART)

$\underline{K}$ elipse (diagonal)

circular (identidade)

$\sigma = \frac{d_{max}}{\sqrt{2N}}$ ← entre centros $\underline{W}$

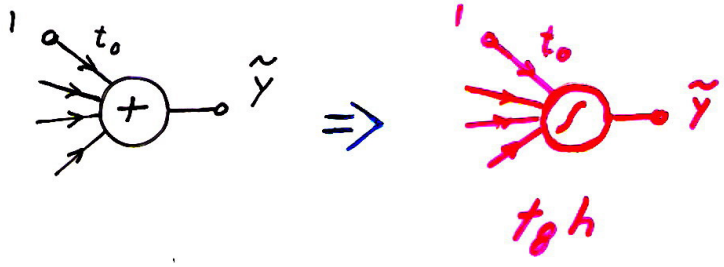
← nº clusters

local - global

⟹ Treinar !

RBF como classificador

Neuronio de saida

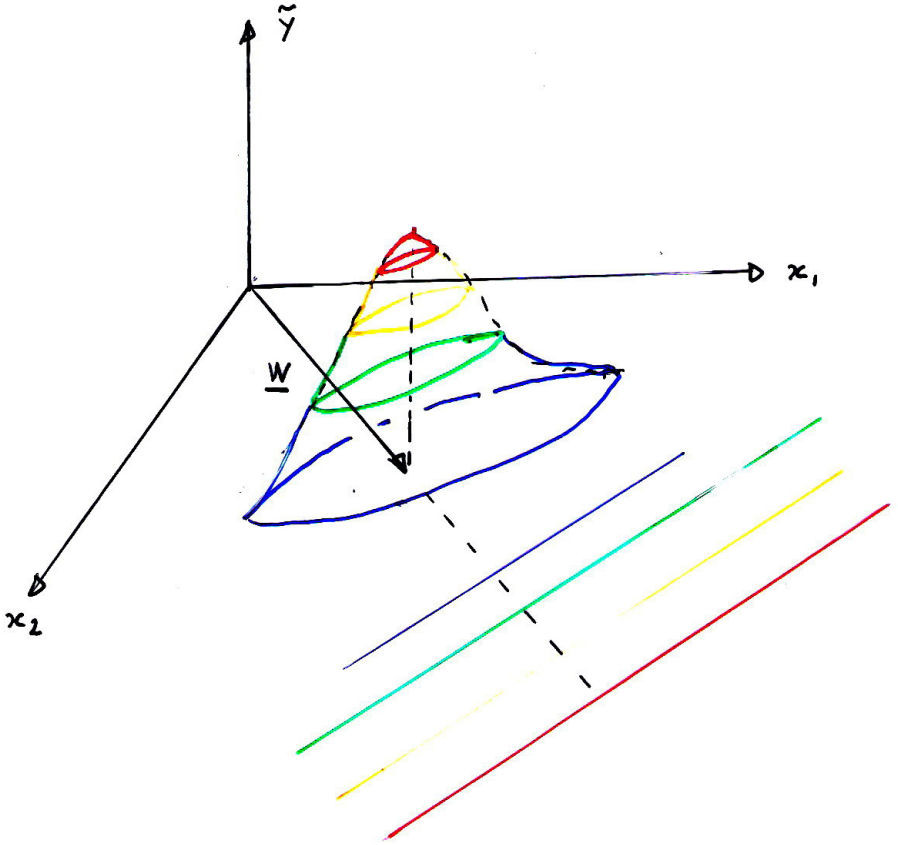


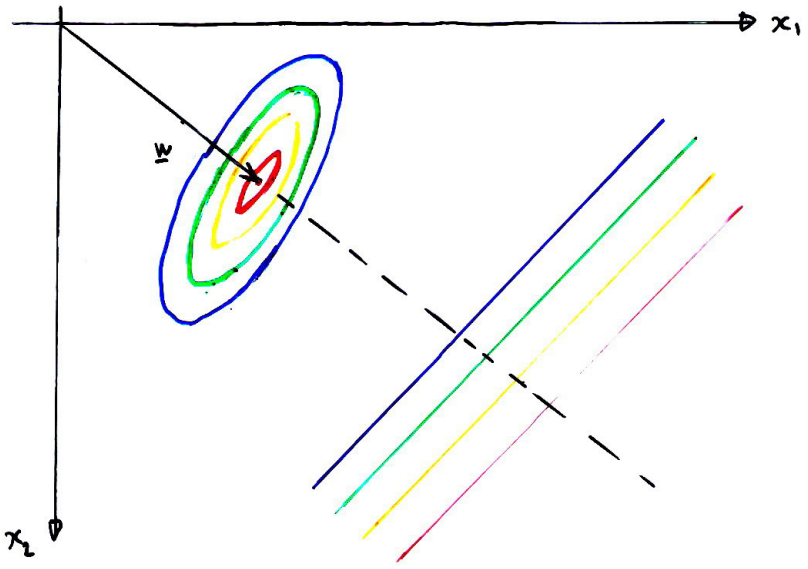
Separadores elementares

$\underline{K} = I$ → esferas

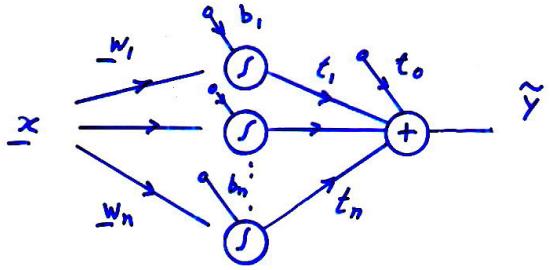
$\underline{K} = diag$ → elipsoides eixos alinhados

$\underline{K} = tricng\ sup$ → elipsoides

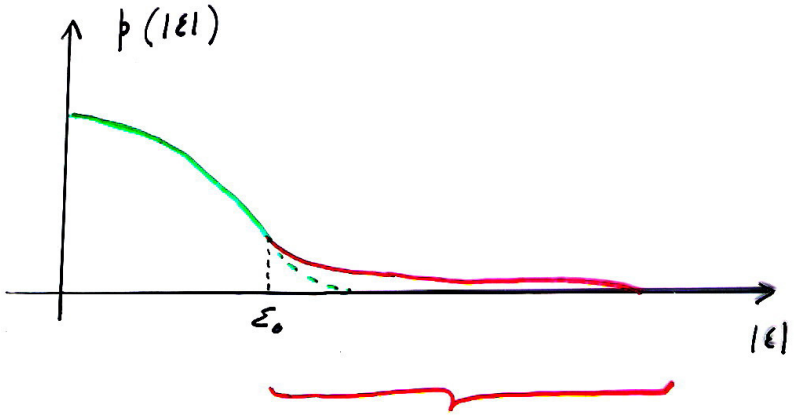

$\tilde{y}$
x_1
x_2
$\underline{w}$


x_1
x_2
$\underline{w}$

Auxílio à BP



$(\underline{x}, y) \Rightarrow (\underline{x}, y, \tilde{y}, \varepsilon) \Rightarrow (\underline{x}, \varepsilon)$



$\varepsilon \Rightarrow \underline{x}$

$|\varepsilon| > \varepsilon_0 \qquad \underline{x} \in X_0$

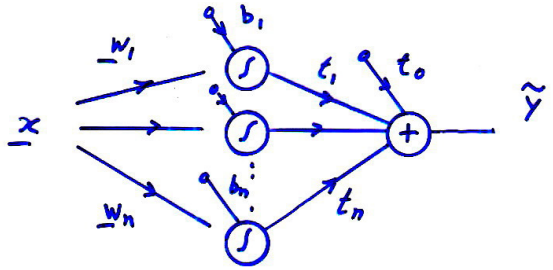
Clusterizer $\underline{x} \in X_0$

ART

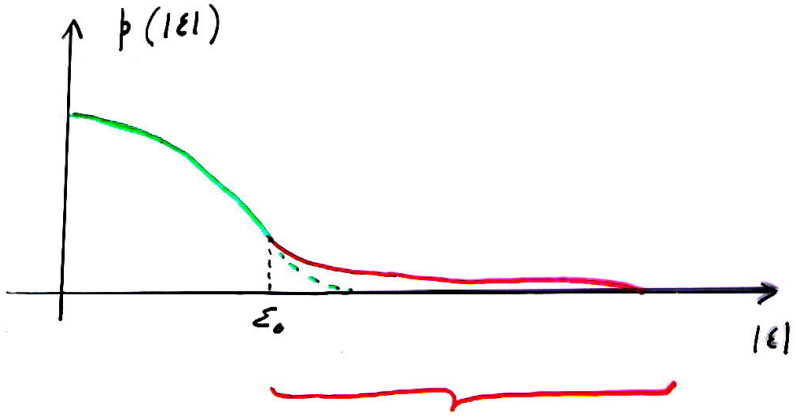
histograma $\Rightarrow r_0$

$\Rightarrow$ centros $\underline{m}_i$

Auxílio à BP



$(\underline{x}, y) \Rightarrow (\underline{x}, y, \tilde{y}, \varepsilon) \Rightarrow (\underline{x}, \varepsilon)$



$\varepsilon \Rightarrow \underline{x}$

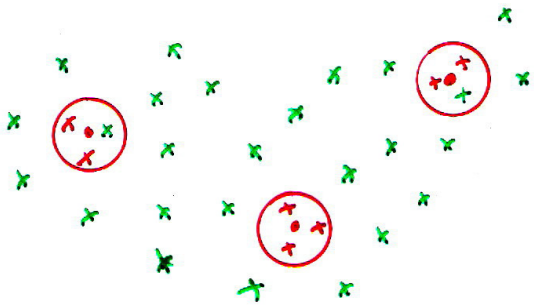
$|\varepsilon| > \varepsilon_0 \qquad \underline{x} \in X_0$

Clusterizer $\underline{x} \in X_0$

ART

histograma $\Rightarrow r_0$

$\Rightarrow$ centros $\underline{m}_i$



Classificador: LC dos grandes erros

Classe dos grandes erros, X_1

Verificar população – identificar outliers

Classificar todos os $\underline{x}$ por região

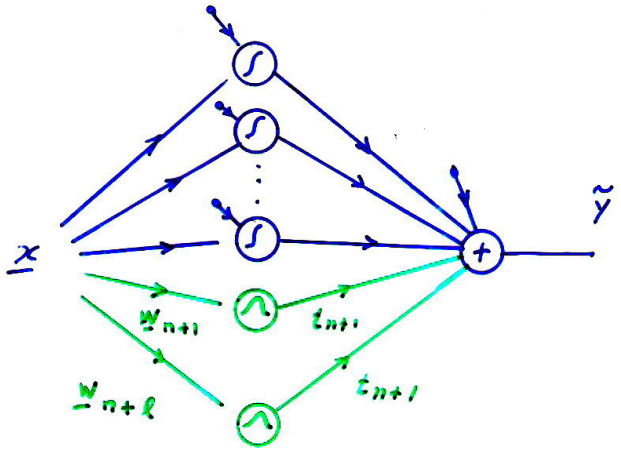
$\underline{x} \in X_2$ mesmos centros $r_2 > r_0$

Montar novo conjunto de treinamento

(apenas a região de grandes erros aumentada)

$(\underline{x}, y) \quad \forall\, \underline{x} \in X_2$

Rede Aumentada



Neurônios de base radial

início $\quad \underline{w}_{n+i} = m_i$

um ou + por centro

esféricos, $\sigma \simeq r_0$

Treinar por BP

apenas os novos neurônios, esféricos

apenas com $(\underline{x}, y)$!

$\underline{x} \in X_2$