

Tipos de Classes para Agrupamento por Similaridade

Objetivo: determinar a granularidade adequada, escolher $u_0 = -r_0$

1 - Classes esféricas com dimensão idêntica

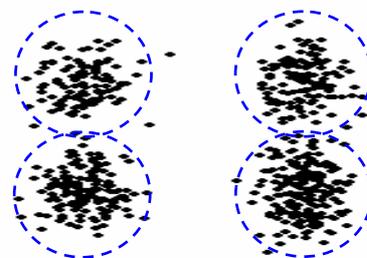
Considerando:

- que um caso bastante freqüente é as classes serem geradas a partir de padrões adicionados de ruído com média nula.
- que o valor rms (desvio padrão) do ruído de cada componente x_j é σ_j , dependente da componente mas independente da classe. Neste caso o uso do fator de escala $k_j = 1/\sigma_j$ torna o valor rms do ruído unitário para todas as componentes x_i



então

o separador ideal para as classes são esferas, que são também o separador natural dos processos não supervisionados.

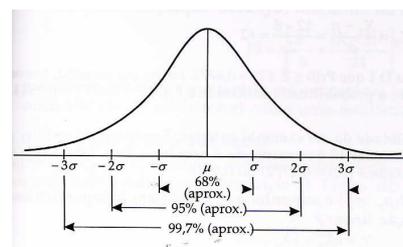


Se o ruído é gaussiano então as classes tem distribuição gaussiana multidimensional com desvio padrão unitário em todas as direções. Em conseqüência

$r_0 = 1$ captura 68 % dos elementos da classe,

$r_0 = 2$ captura 95 % dos elementos da classe,

$r_0 = 3$ captura 99,7 % dos elementos da classe, etc.



2 - Determinação do ruído σ_j em cada direção j

2.1 - Caso treinamento supervisionado (Kohonen ou ART)

As classes são conhecidas, determinar σ_i para qualquer classe C_j

Determinar para cada direção (componente) i

X_i – componente i de \underline{X} (entrada não normalizada)

$M_i = E_{\forall \underline{x}} X_i$ – componente i da média de todos os elementos
de todas as classes

$M_{i,j} = E_{\forall \underline{x} \in C_j} X_i$ – componente i da média dos elementos
da classe C_j

x_i – componente i de \underline{x} (entrada normalizada)

Ruído rms (desvio padrão) σ_j em cada direção j

$$\sigma_i = E_{\forall C_j} \left\{ \left[E_{\forall \underline{x} \in C_j} (X_i - M_{i,j})^2 \right]^{1/2} \right\} \quad \text{- ruído rms (desvio padrão) na direção j}$$

Escalamento por componente / direção

$$x_i = \frac{1}{\sigma_i} (X_i - M_i)$$

Todas as classes tem distribuição gaussiana com desvio padrão

$\sigma = 1$ em todas as direções

Como veremos posteriormente o escalamento utilizado para cada componente / direção influi na clusterização.

2.2 - Treinamento não supervisionado - determinação de σ

As classes não são conhecidas, como estimar σ_j ?

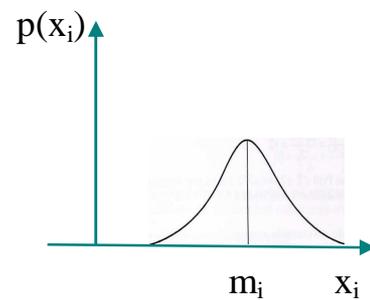
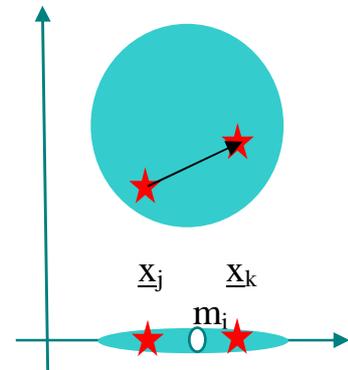
Caracterização simplificada do ruído:

A caracterização do ruído é feita em cada direção x_i de forma independente

2.2.1 - Caracterização do ruído gaussiano em uma direção x_j , caso de uma única classe.

$$p(x_i) = \frac{1}{\sqrt{2\pi}\sigma_i} \exp\left(-\frac{(x_i - m_i)^2}{2\sigma_i^2}\right)$$

onde $m_i = E x_i$



Histograma da distância entre entradas tomada na direção x_i

$$d_i = x_j - x_k$$

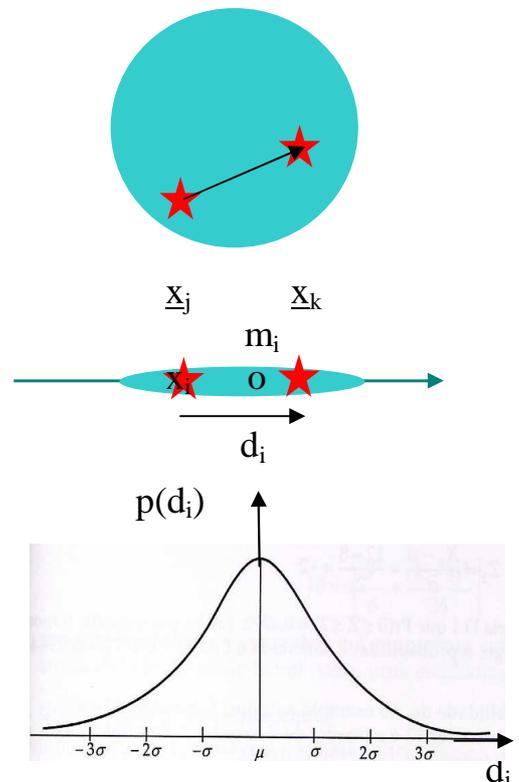
A distribuição da distância entre entradas é uma gaussiana com média nula e desvio σ_d

$$p(d_i) = \int_{-\infty}^{+\infty} p(x) p(x + d_i) dx =$$

$$= \frac{1}{\sqrt{2\pi}\sigma_d} \exp\left(-\frac{d_i^2}{2\sigma_d^2}\right)$$

$$\mu_d = 0$$

$$\sigma_d = \sqrt{2} \sigma_i$$



O desvio padrão da distribuição das diferenças entre os elementos de uma classe em uma direção x_i contém a informação do ruído σ_i naquela direção.

Caracterização simplificada do ruído:

2.2.2 - Caracterização em uma dimensão, caso duas classes

Histograma da distância entre entradas

$$d_i = x_j - x_k$$

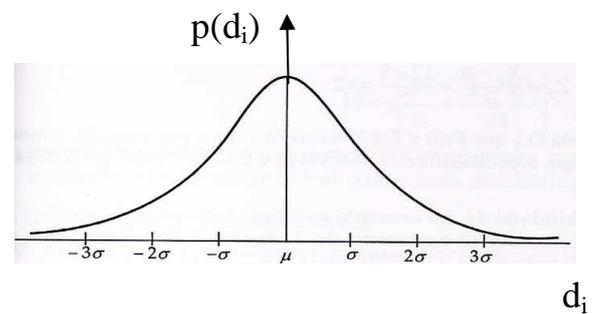
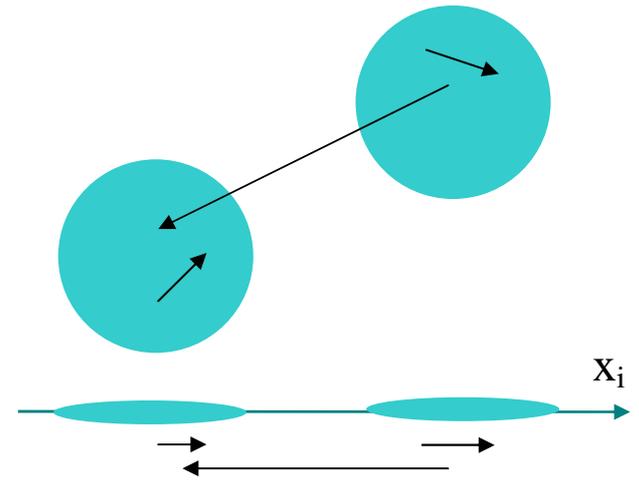
Soma de distribuições

Intra classe: x_j e x_k pertencem à mesma classe

gaussiana com média nula e desvio padrão

$$\mu_d = 0$$

$$\sigma_d = \sqrt{2} \sigma_i$$



Inter classes

x_j e x_k pertencem à classes diferentes

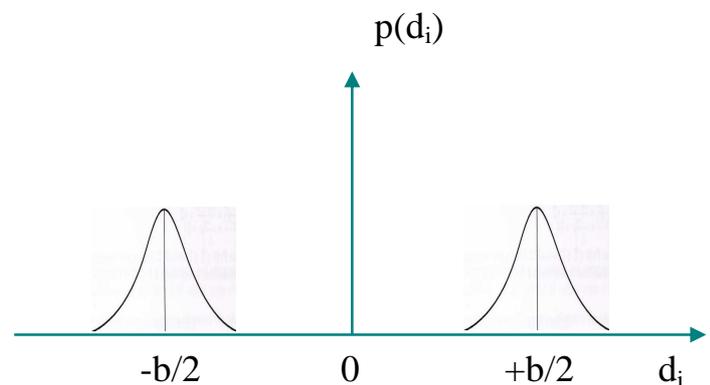
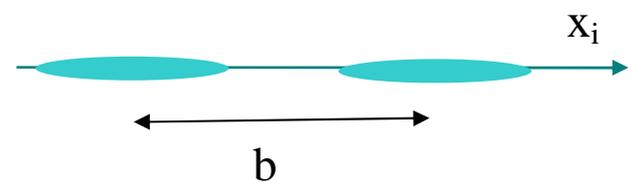
gaussianas com

médias: $\mu_1 = +b$ e $\mu_2 = -b$

onde b é a distância entre os baricentros

e desvios padrão: $\sigma_d = \sqrt{2} \sigma_i$

Os histogramas inter classes são simétricos

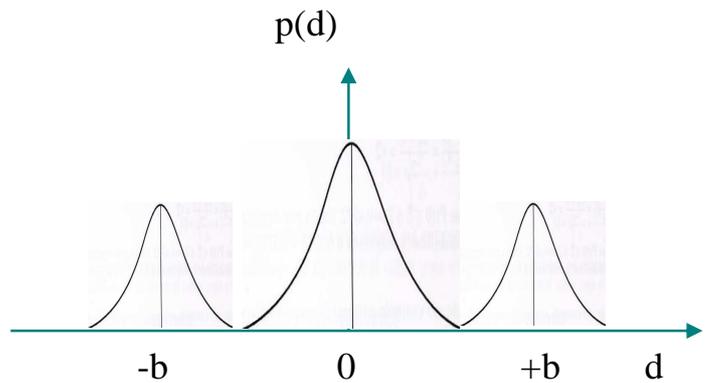


$$p(d_i) = \frac{1}{2} \frac{1}{\sqrt{2\pi}\sigma_d} \left[\exp\left(-\frac{(d_i - b)^2}{2\sigma_d^2}\right) + \exp\left(-\frac{(d_i + b)^2}{2\sigma_d^2}\right) \right]$$

Histograma completo:

$$p(x_i) = \frac{1}{\sqrt{2\pi}\sigma_i} \exp\left(-\frac{x_i^2}{2\sigma_i^2}\right)$$

$$\sigma_d = \sqrt{2} \sigma_i$$



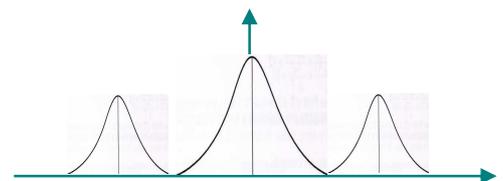
$$p(d) = \frac{1}{\sigma_d} \left[k_0 \exp\left(-\frac{d^2}{2\sigma_d^2}\right) + k_1 \left(\exp\left(-\frac{(d+b)^2}{2\sigma_d^2}\right) + \exp\left(-\frac{(d-b)^2}{2\sigma_d^2}\right) \right) \right]$$

onde k_0 e k_1 dependem da população das classes.

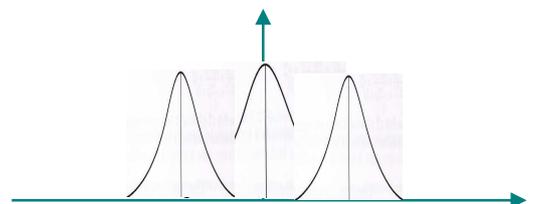
devido à simetria dos histogramas das distâncias interclasses no entorno da origem

$$p(d_i) \cong \frac{k}{\sqrt{2\pi}\sigma_d} \exp\left(-\frac{d_i^2}{2\sigma_d^2}\right) \quad \text{válido se } b \gg \sqrt{2} \sigma_i$$

Se as classes não forem demasiado próximas a forma da distribuição de d_i no entorno da origem pode então ser utilizada para calcular $\sigma_i \cong \sigma_d / \sqrt{2}$, que será utilizado para determinar o fator de escala k_i na direção x_i .



Caso contrário tomando $p(d)$ para diversos valores de d cria-se um sistema de equações transcendentais possível de ser resolvido



Caracterização simplificada do ruído:

2.2.3 - Caracterização em uma dimensão, caso múltiplas classes

Histograma da distância entre entradas

$$d_i = x_j - x_k$$

Soma de distribuições

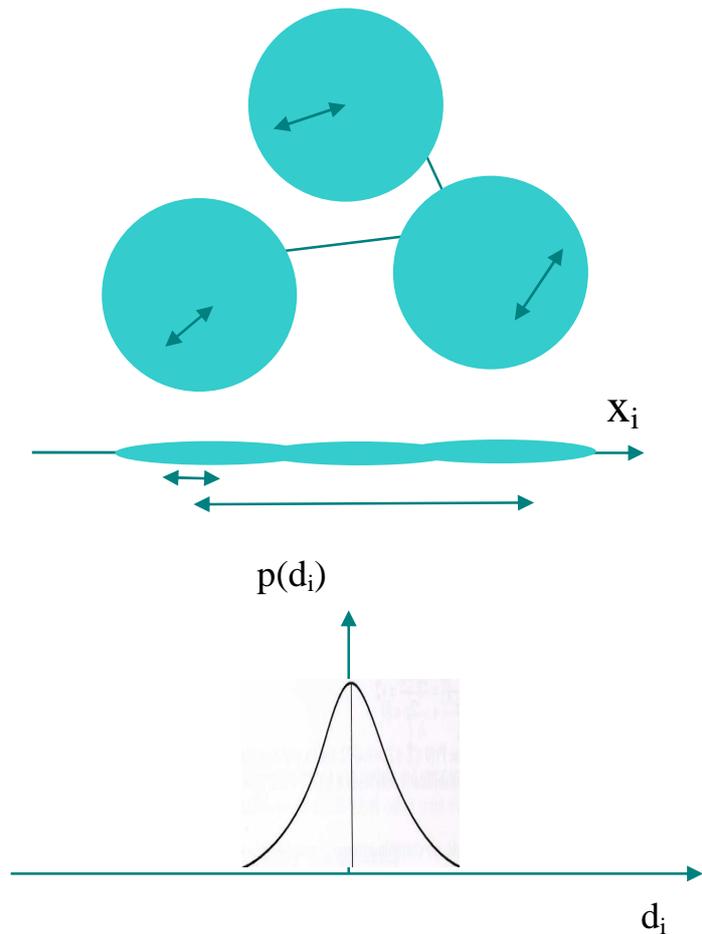
Intra classe:

x_j e x_k pertencem à mesma classe

gaussiana com média nula e desvio padrão

$$\mu_d = 0$$

$$\sigma_d = \sqrt{2} \sigma_i$$



Inter classes

x_j e x_k pertencem à classes diferentes

gaussianas com

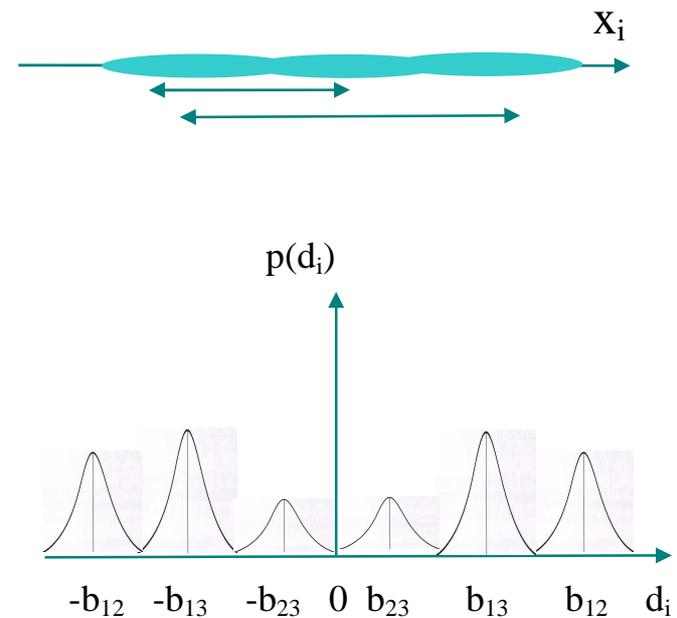
médias: +/- distâncias inter baricentros

$$\mu_{k11} = + b_{kl} \quad \text{e} \quad \mu_{k12} = - b_{kl}$$

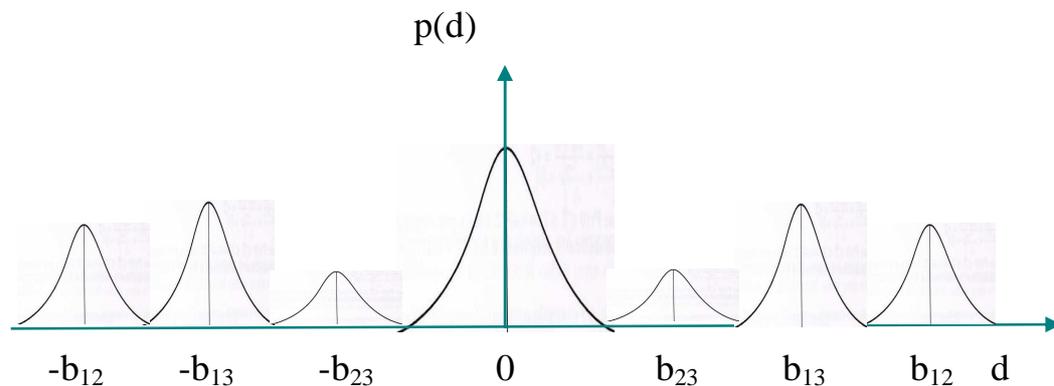
onde $b_{kl} = |w_k - w_l|$

onde w_k e w_l são as projeções sobre a direção x_i de dois centros de classe

e desvios padrão: $\sigma_d = \sqrt{2} \sigma_i$



Os histogramas inter classes são simétricos

Histograma completo:

Novamente,

se as classes não forem demasiado próximas a forma da distribuição e d_i no entorno da origem pode então ser utilizada para calcular $\sigma_i \cong \sigma_d / \sqrt{2}$, que será utilizado para determinar o fator de escala k_i .

Caso contrário tomando $p(d)$ para diversos valores de d cria-se um sistema de equações transcendentais possível de ser resolvido

$$p(d) = \frac{1}{\sigma_d} \sum_i k_i \left(\exp\left(-\frac{(d + b_i)^2}{2\sigma_d^2}\right) + \exp\left(-\frac{(d - b_i)^2}{2\sigma_d^2}\right) \right)$$

na maioria dos casos apenas três gaussianas devem ser suficientes fazer uma aproximação aceitável no entorno da origem

$$p(d) = \frac{1}{\sigma_d} \left[k_0 \exp\left(-\frac{d^2}{2\sigma_d^2}\right) + k_1 \left(\exp\left(-\frac{(d + b)^2}{2\sigma_d^2}\right) + \exp\left(-\frac{(d - b)^2}{2\sigma_d^2}\right) \right) \right]$$

Para tornar o valor rms do ruído (desvio padrão intraclasse) de cada componente unitário normalizamos

$$x_i = \frac{1}{\sigma_i} (X_i - M_i) = \frac{\sqrt{2}}{\sigma_d} (X_i - M_i)$$

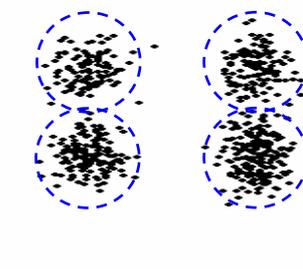
Com esta escolha o ruído rms (desvio padrão do ruído intraclasse) em todas as direções x_i será aproximadamente unitário

$$\sigma_i \cong 1 \quad \forall i = 1, 2, \dots, n$$

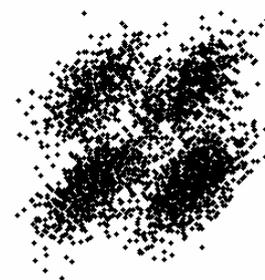
Novamente é bom destacar que, como veremos posteriormente, o escalamento utilizado para cada componente / direção influi na clusterização.

3 - Ruídos correlatos

Seo os ruídos das componentes não tiverem correlação significativa os domínios das classes serão delimitados por esferas



Entretanto, se os ruídos das componentes forem correlatos os domínios das classes serão elipsóides, mas estarão contidos nas esferas determinadas.



A decorrelação (beranqueamento) dos ruídos permitirá uma melhor determinação dos domínios das classes, que passarão a ser esféricos.

Branqueamento do Ruído

1 - Realizar uma clusterização inicial

1. 1 - Normalizar a entrada \underline{X} . Para cada dimensão X_j da entrada

$$x_j = \frac{1}{\sigma_j} (X_j - M_j)$$

1. 2 - Realizar a clusterização, determinar a classe de cada entrada e os centros de classe M_{ij} .

2 - Determinar o ruído r_j em cada componente de cada entrada. Para um elemento da classe i

$$\text{como } x_j = M_{ij} + r_j \quad \text{então} \quad r_j = x_j - M_{ij}$$

3 - Verificar a correlação entre o ruído das diversas direções (componentes)

4 - Caso haja correlação, calcular as direções das componentes principais (PCA) do ruído \underline{r} (e não do sinal \underline{x}).

5 - Representar as entradas $\underline{x} + \underline{r}$ nesta nova base. O ruído passa agora a ser não correlato. Escalar as componentes da nova base para obter o valor rms do ruído idêntico e cada direção (componente).

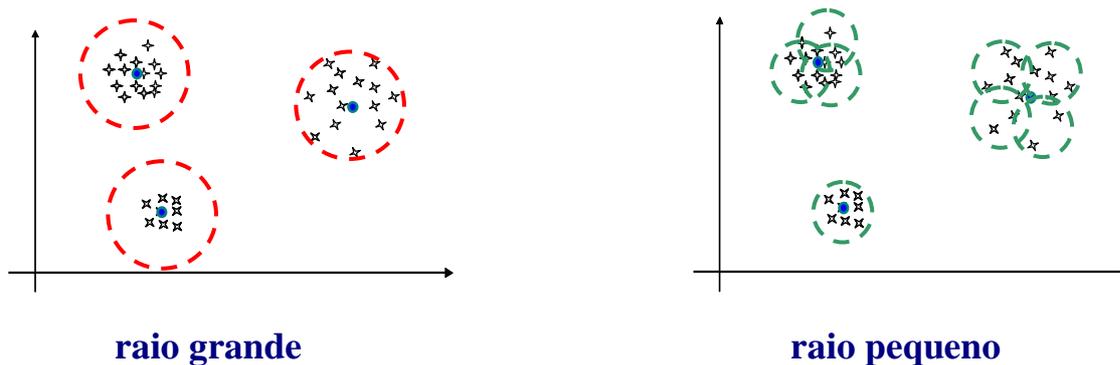
6 - As classes serão agora delimitadas por esferas.

4 - Classes Esféricas com dimensões variadas

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

muito pequeno: boa definição, mas muitos neurônios

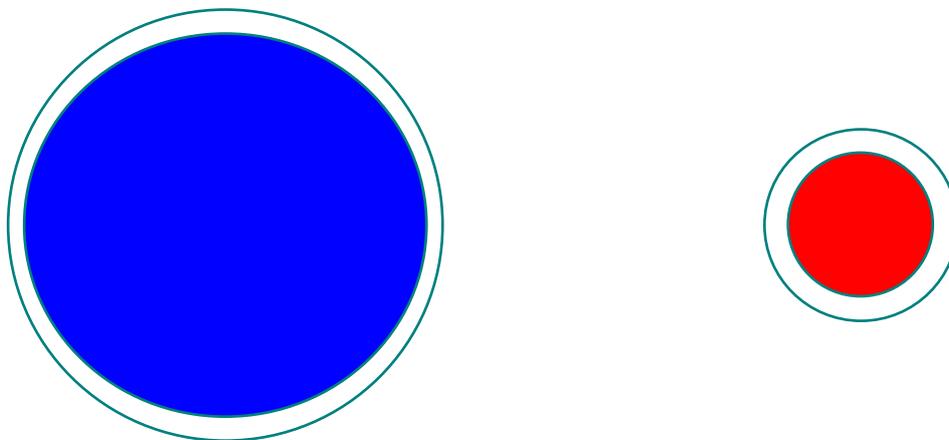
muito grande: poucos neurônios, mas definição pobre



Classes esféricas com dimensões variáveis



Neurônios com raio variável - idêntico à Camada de Kohonen



4.1 - Raio de vigilância variável por neurônio

Condição de pertinência à classe C_i

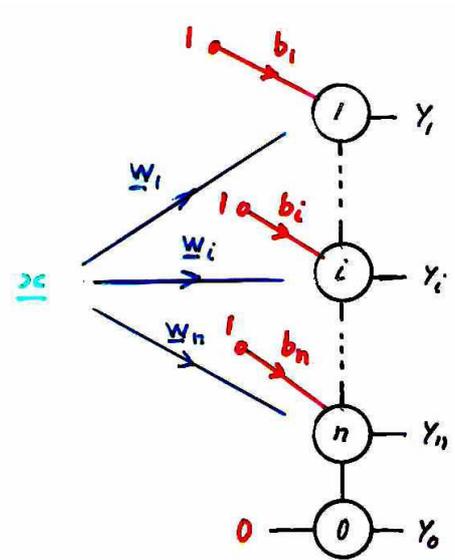
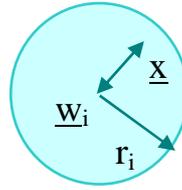
$$|\vec{x} - \vec{w}_i|^2 = d_i^2 < r_i^2$$

$$-d_i^2 > -r_i^2$$

$$-d_i^2 + r_i^2 > 0$$

$$b_i = r_i^2$$

$$u_i = -d_i^2 + b_i = -d_i^2 + r_i^2 > 0$$

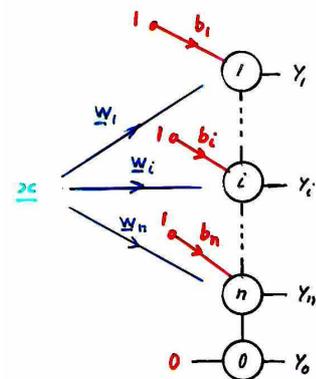
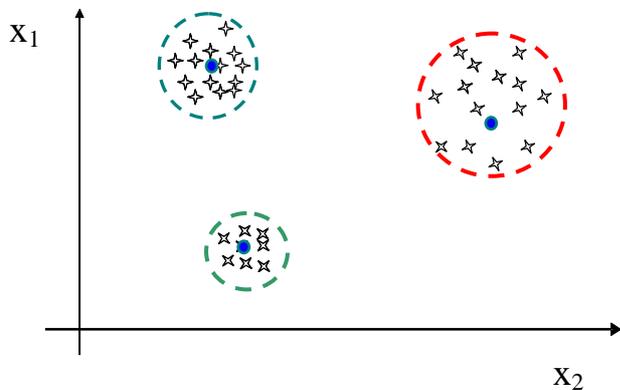


4.1 - ART com similaridade mínima

1 - Treinar a ART

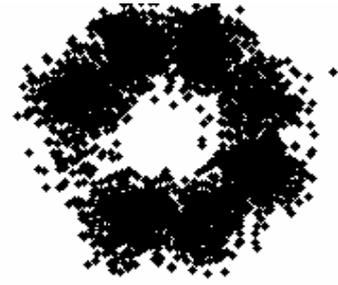
2 – Calcular o raio de cada classe como o máximo afastamento dos elementos aos centros das respectivas classes [$r_i = \text{Max} (-u_i)$]

3 – Inserir os bias $b_i = -r_i$ e o neurônio de controle y_0



5 - Classes Não-Esféricas

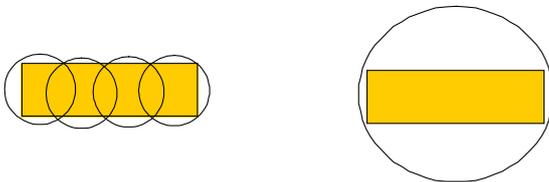
Representar como a união de classes esféricas



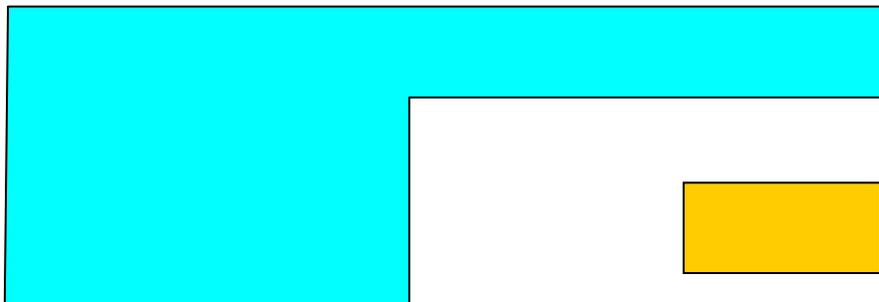
Que raio de similaridade a usar ?

muito pequeno: boa definição, mas muitos neurônios

muito grande: poucos neurônios, mas definição pobre

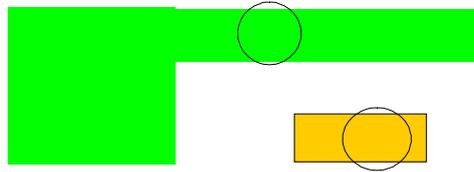


5.1 - Treinamento Supervisionado

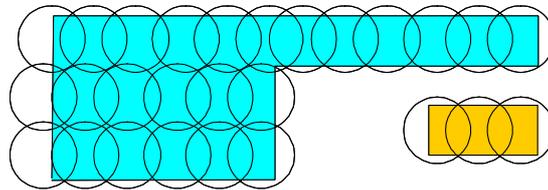


- **1a. Etapa – estabelecer o raio de similaridade a ser usado**
- **2a. Etapa – ART - determinar os domínios esféricos**
- **3a. Etapa – OU – conecta os domínios pertencentes a mesma classe**

**1a. Etapa -
escolher um raio de similaridade r_0 adequado**

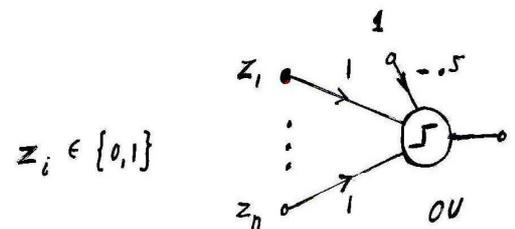
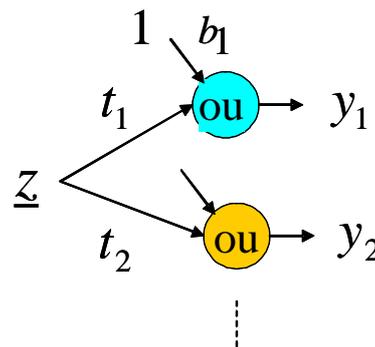
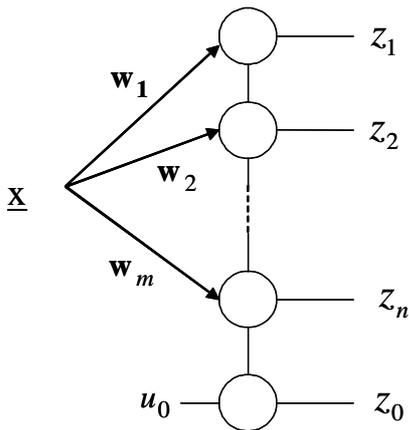


• 2a. Etapa – ART - determinar os domínios esféricos

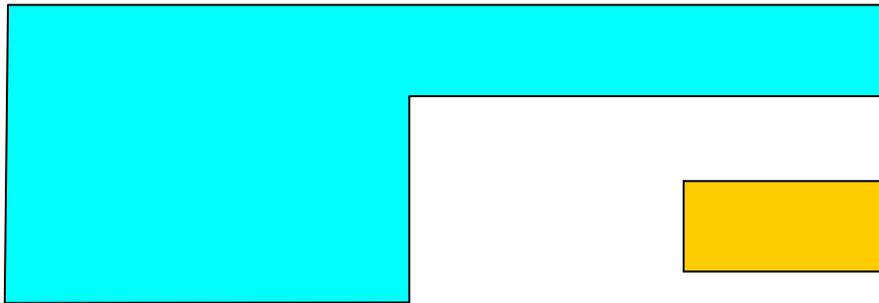


3a. Etapa – OU – conecta os domínios pertencentes a mesma classe

Construir a segunda camada com um neurônio “ou” para cada classe



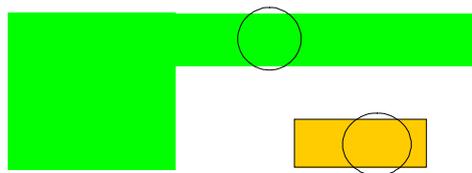
5.2 - Treinamento Não-Supervisionado



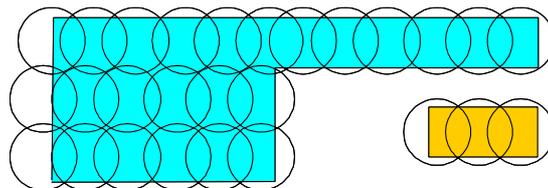
- 1a. Etapa – estabelecer o raio de similaridade a ser usado
- 2a. Etapa – ART - determinar os domínios esféricos
- 3a. Etapa – determinar os domínios vizinhos
- 4a. Etapa – OU – conecta os domínios vizinhos

- 1a. Etapa -

escolher um raio de similaridade r_0 adequado

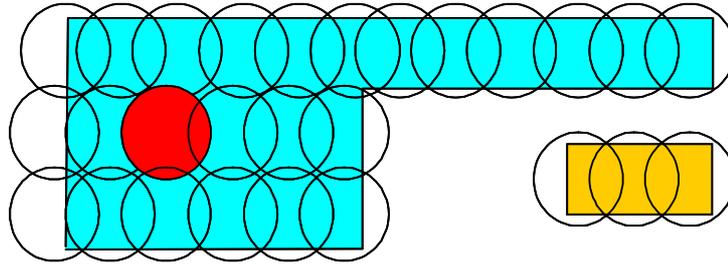


- 2a. Etapa – ART - determinar os domínios esféricos

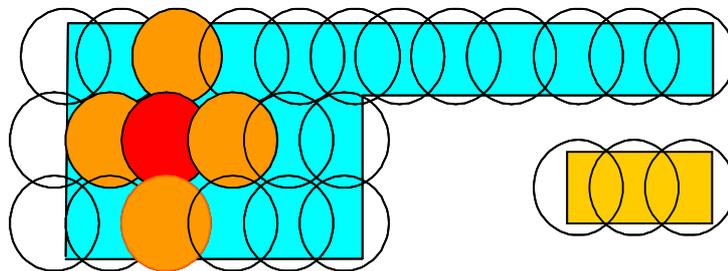


- **3a. Etapa – determinar os domínios vizinhos**

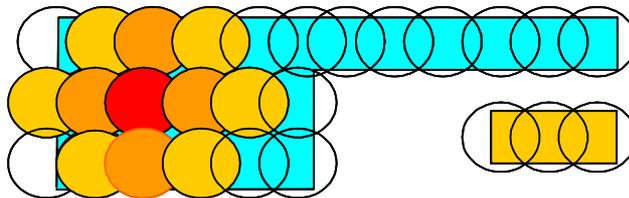
Centro de classe \underline{w}_i escolhido arbitrariamente



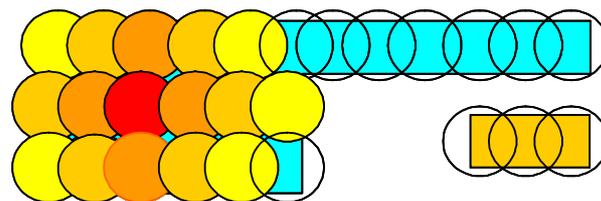
vizinhos $|\underline{w}_i - \underline{w}_j| \leq 2r_0$



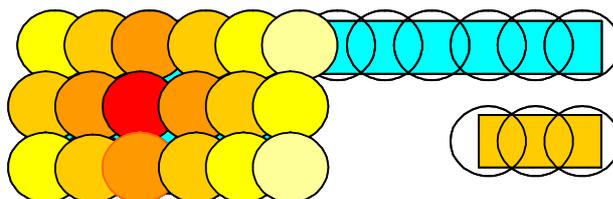
Vizinhos de vizinhos



Vizinhos de vizinhos de vizinhos



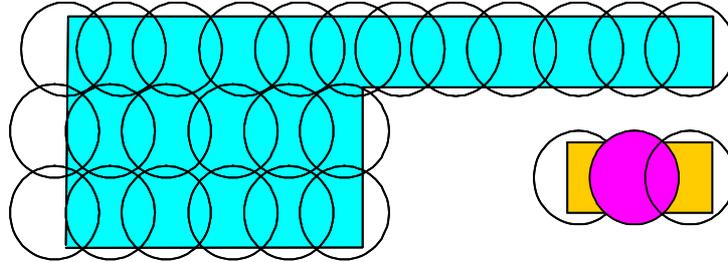
Vizinhos de vizinhos de ...



Todos pertencem à uma mesma classe.

Sobraram neurônios ?

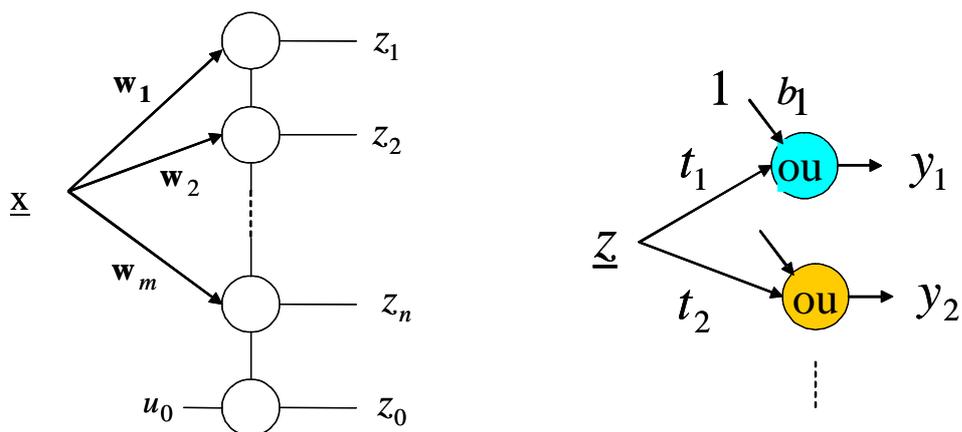
Recomeçar o processo, outra classe.



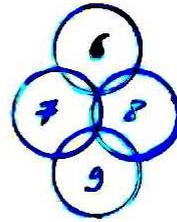
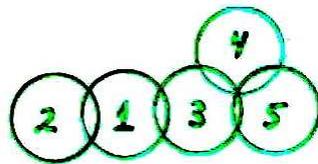
Não há mais neurônios ?

- **4a. Etapa – OU – conecta os domínios vizinhos**

Construir a segunda camada com um neurônio “ou” para cada classe



Exemplo:



1	2	-	
	3	4	-
		5	-

6	7	9	-
	8	-	

10	-

$C_1 = \{1,2,3,4,5\}$

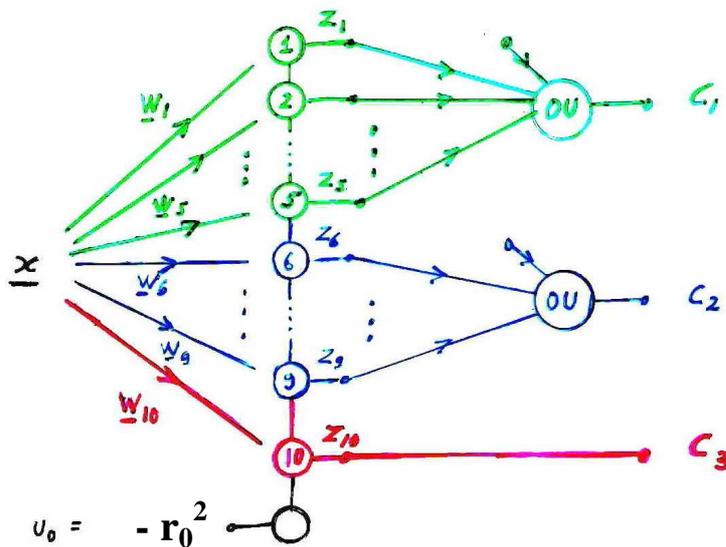
$C_2 = \{6,7,8,9\}$

$C_3 = \{10\}$

$C_1 = \{1,2,3,4,5\}$

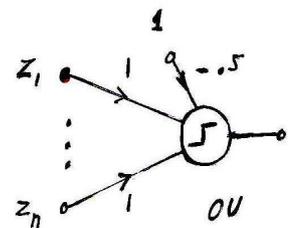
$C_2 = \{6,7,8,9\}$

$C_3 = \{10\}$



$u_0 = -r_0^2$

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



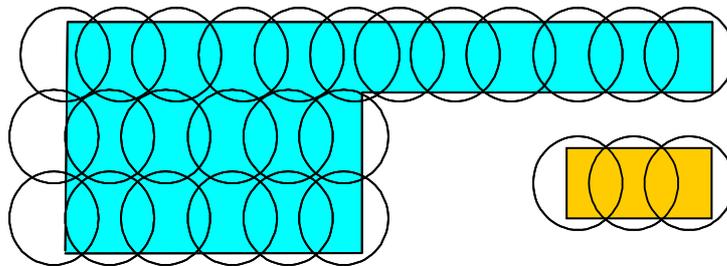
Obs: Sendo um processo não supervisionado este método não consegue diferenciar classes cuja separação (a) é muito menor que a dimensão das classes ou (b) é da ordem de grandeza de r_0 .

6 - Uso de esferas de raios variados

1. Determine os domínios utilizando esferas suficientemente pequenas.

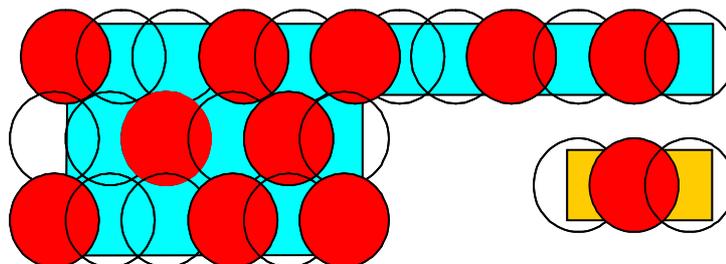
Como no processo inicial para classes não esféricas.

2. Determine as classes e seus elementos



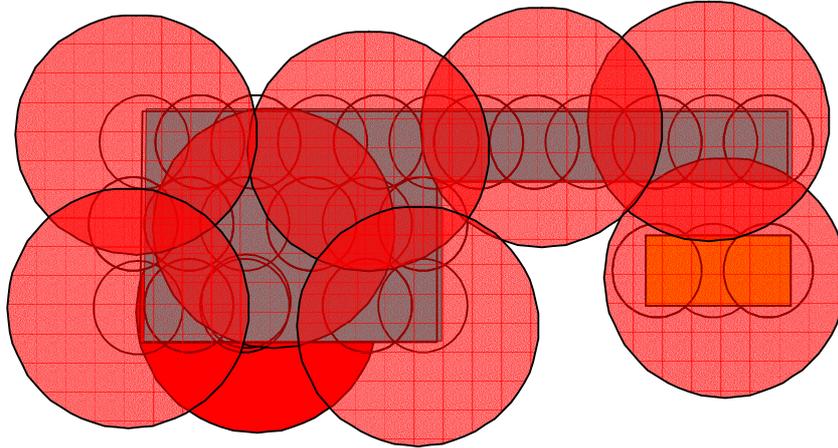
6.1 - Processo associativo

a, Escolha um conjunto de neurônios não vizinhos entre si tais que todos os neurônios são vizinhos de ao menos um neurônio do conjunto selecionado. Comece com os neurônios com maior número de vizinhos.



b. Para cada neurônio do conjunto selecionado e seus vizinhos recalculamos o baricentro e o raio ou

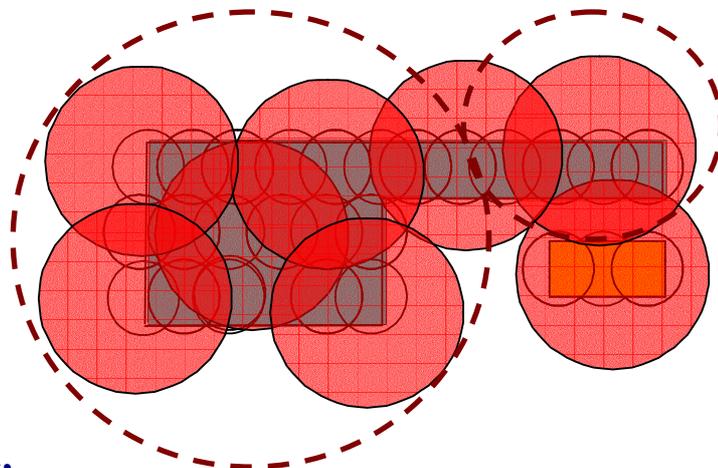
b.1. Alternativa mais simples: use o mesmo centro e triplique o raio



c. Verificar se não houve classificações erradas. Se sim, reduzir o raio do neurônio ganhador errado até que o neurônio correto ganhe.

d. Verificar se não houve elementos não classificados. Se sim, criar novos neurônios para classificarlos corretamente

e. Repetir o processo com os novos neurônios até que o número de classes esteja adequado.



f. Construir a outstar de OU's

6.2 - Processo divisivo

a. Estimar o raio de cada classe

b. Treinamento ART para cada classe usando seu raio

c. Verificar se houve “invasões”

SIM neurônios não invadidos - congelar \underline{w}_i e r_i

neurônios invadidos - reduzir r_i , congelar \underline{w}_i e r_i

verificar se há áreas “descobertas”

SIM - retornar ao passo b

NÃO- passo d

NÃO passo d

d. Construir a “outstar” de OU’s

Exemplo:

Classes

b – Treinamento ART

c – Há “invasões”

c – Redução de raio

Há áreas descobertas

b – ART novamente

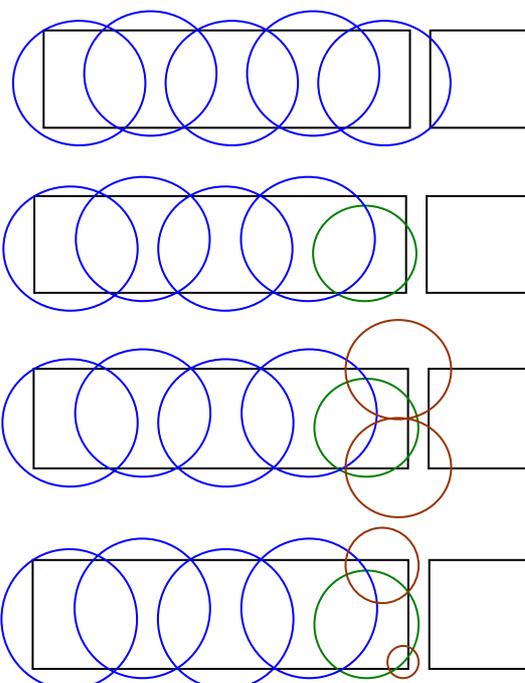
São criados novos neurônios

c - Há novas invasões

c – Redução de raio

Não há áreas descobertas

Fim



A Speaker Verification Method Using LPC Singularity Location

Hardy L. C. P. Pinto

Rafael G. C. P. Pinto

Luiz P. Calôba

Representação LPC dos fonemas /a/ e /i/

Cinco locutores masculinos

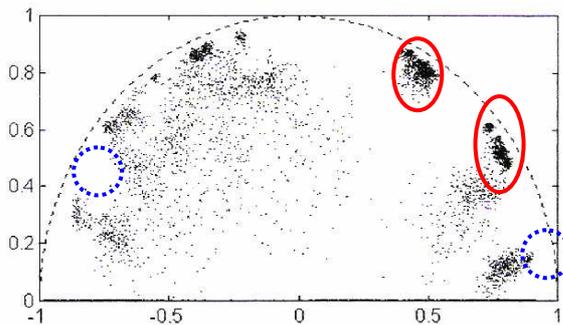


Figure 1- Roots of the predictor polynomial for segments of the sustained Portuguese phoneme /a/ for five different people of the same sex.

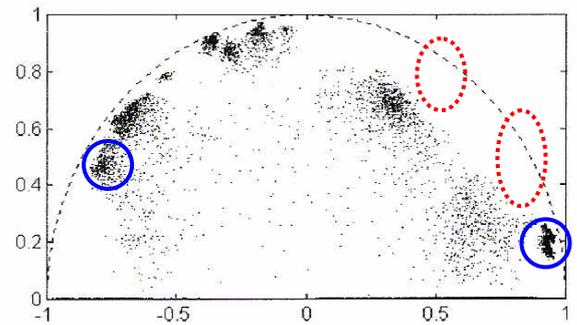


Figure 2- Same as fig. 1 for the phoneme /i/.

Diferenciação dos locutores

Fonema /a/ por dois locutores diferentes

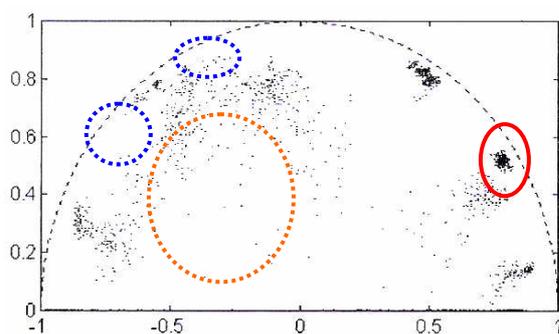


Figure 3 - Phoneme /a/ spoken by one speaker.

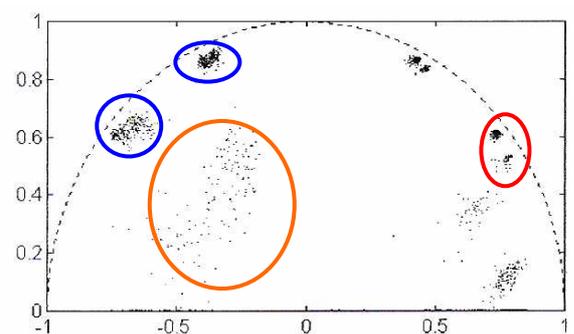


Figure 4 - Phoneme /a/ spoken by another speaker.

Rede ART modificada:

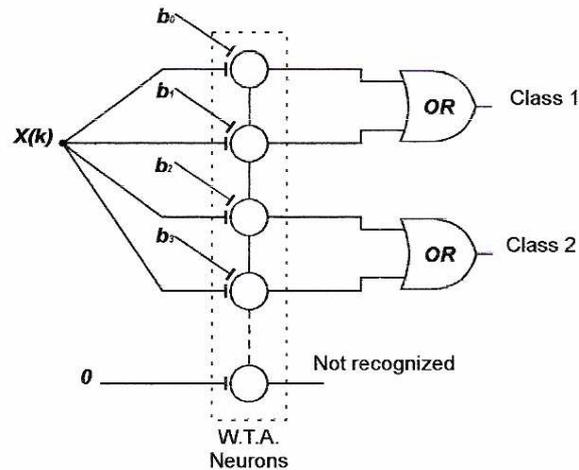


Figure 5 - Modified ART structure

Taxas de acerto de Fonemas:

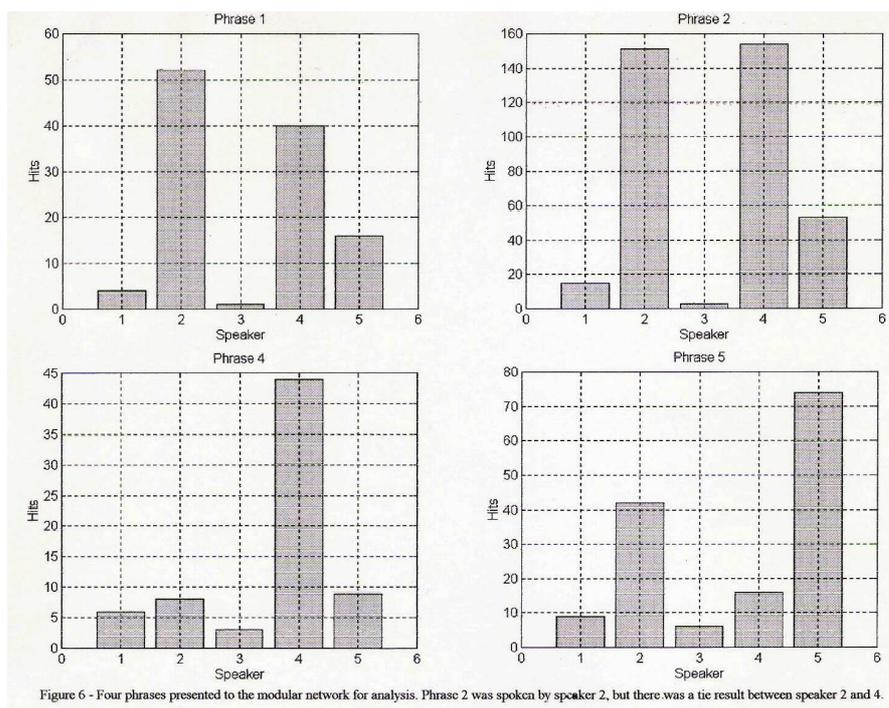
TABLE I A - PHONEME TEST SET

Phoneme	Total segments	Correct	Not recognized	Wrong classif.
/a/	548	98.91 %	0.00 %	1.09 %
/e/	600	99.17 %	0.33 %	0.50 %
/i/	551	99.64 %	0.18 %	0.18 %
/o/	601	97.84 %	0.67 %	1.49 %
/u/	549	97.45 %	1.09 %	1.46 %

TABLE I B - CONFUSION TABLE FOR THE PHONEMES

Spoken	Recognized				
	/a/	/e/	/i/	/o/	/u/
/a/	98.90 %	0.37 %	0.00 %	0.73 %	0.00 %
/e/	0.17 %	99.17 %	0.00 %	0.00 %	0.33 %
/i/	0.00 %	0.00 %	99.63 %	0.00 %	0.18 %
/o/	1.50 %	0.00 %	0.00 %	97.84 %	0.00 %
/u/	0.00 %	0.18 %	1.09 %	0.18 %	97.45 %

Votação para Locutores:



Taxa de acerto de locutores

TABLE II A - SPEAKER TEST SET (MODULAR TOPOLOGY)

Speaker	Total segments	Correct	Not recognized	Wrong classif.
2	750	75.20 %	7.87 %	16.93 %
3	749	91.32 %	3.20 %	5.48 %
4	598	70.23 %	11.71 %	18.06 %
5	752	82.71 %	6.52 %	10.77 %

TABLE II B - CONFUSION TABLE FOR SPEAKERS (MODULAR)

Speaker	Recognized Speaker				
	1	2	3	4	5
2	0.93 %	75.20 %	7.33 %	7.20 %	1.47 %
3	0.00 %	2.80 %	91.32 %	0.88 %	1.87 %
4	1.17 %	13.88 %	1.00 %	70.23 %	2.01 %
5	0.00 %	2.79 %	6.25 %	1.73 %	82.71 %

Os autores do trabalho:

