

## CPE 722 - 1ª série de exercícios

1 – Os oito elementos à classificar apresentados a seguir são uni-dimensionais:

-3,2 -3,0 -0,7 0,0 0,8 2,0 2,4 2,9

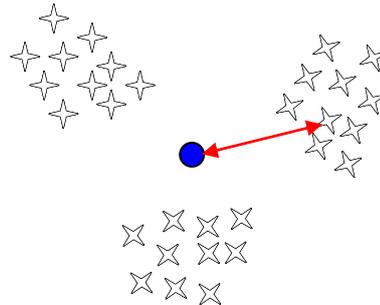
1.1 - Calcule a média  $m$  e a dissimilaridade total  $F_0$  do conjunto de elementos.

### Média

$$\vec{m} = E_{\forall i} \vec{x}_i = \frac{1}{N} \sum_{i=1}^N \vec{x}_i$$

### Dispersão, dissimilaridade total

$$F_0 = \sum |\vec{x}_i - \vec{m}|^2$$



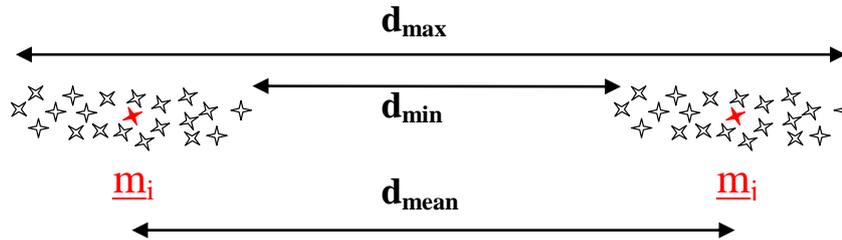
1.2 - Classifique os elementos usando um classificador k-means cujos três padrões iniciais são  $-3,0$   $-0,7$  e  $2,0$ . Calcule o  $F_{in}$  e  $F_{out}$  obtidos a partir de suas fórmulas de definição, comprovando que sua soma é igual ao  $F_0$  calculado anteriormente.

$$F_{in} = \sum_{\forall C_j} F_j \geq 0 \quad F_j = n_j \sigma_j^2 = \sum_{\forall \vec{x}_i \in C_j} \|\vec{x}_i - \vec{m}_j\|^2$$

$$F_{out} = \sum_{\forall j} n_j \|\vec{m}_j - \vec{m}\|^2 \geq 0$$

1.3 – Para a classificação obtida em 1.2 calcule a dispersão inter-classes pelo critério

a) do vizinho mais próximo; b) do vizinho mais distante; c) da distância média entre elementos e d) da distância entre os baricentros.



**vizinho + próximo**

$$d_{\min}(\mathcal{X}_i, \mathcal{X}_j) = \min_{\mathbf{x} \in \mathcal{X}_i, \mathbf{x}' \in \mathcal{X}_j} \|\mathbf{x} - \mathbf{x}'\|$$

**vizinho + distante**

$$d_{\max}(\mathcal{X}_i, \mathcal{X}_j) = \max_{\mathbf{x} \in \mathcal{X}_i, \mathbf{x}' \in \mathcal{X}_j} \|\mathbf{x} - \mathbf{x}'\|$$

**distância média**

$$d_{\text{avg}}(\mathcal{X}_i, \mathcal{X}_j) = \frac{1}{n_i n_j} \sum_{\mathbf{x} \in \mathcal{X}_i} \sum_{\mathbf{x}' \in \mathcal{X}_j} \|\mathbf{x} - \mathbf{x}'\|$$

**distância entre médias**

$$d_{\text{mean}}(\mathcal{X}_i, \mathcal{X}_j) = \|\mathbf{m}_i - \mathbf{m}_j\|.$$

1.4 - A partir do classificador obtido em 1.2 reduza o número de classes de uma unidade usando um algoritmo associativo. Que classes você escolhe para associar, e qual os  $F_{in}$  e  $F_{out}$  após a associação ?

**Junte os dois clusters que aumentam  $F_{in}$  o mínimo possível**

$$\Delta F = \frac{n_j n_k}{n_j + n_k} |\vec{m}_j - \vec{m}_k|^2$$

**escolher centros próximos**

$$\vec{m}_j \approx \vec{m}_k$$

**e de preferência com populações desiguais**

1.5 – A partir do classificador obtido em 1.2 aumente o número de classes de uma unidade usando um algoritmo dissociativo. Que classe você escolhe para dissociar, e qual os  $F_{in}$  e  $F_{out}$  após a dissociação ?

**Critério: maior  $F_j$  , por exemplo**

1.6 – A partir dos resultados acima é possível concluir algo a respeito do nível do “agrupamento natural ótimo” ?

**2** – Em uma camada de Kohonen o padrão  $\vec{w}^i$  da classe  $C_i$  deve ser tal que minimize o valor esperado da distância dos elementos pertencentes a  $C_i$  ao seu centro  $\vec{w}^i$ , e maximize o valor esperado da distância à  $\vec{w}^i$  das entradas não pertencentes a  $C_i$ . Para tanto é proposto minimizar a função objetivo:

$$F_i = \gamma \mathbf{E}_{\forall \vec{x}^j \in C_i} \|\vec{x}^j - \vec{w}^i\|^2 - (1 - \gamma) \mathbf{E}_{\forall \vec{x}^j \notin C_i} \|\vec{x}^j - \vec{w}^i\|^2, \quad 0 < \gamma < 1$$

Calcule  $\vec{w}^i$  minimante de  $F_i$  em função de

$$\vec{m}^k = \mathbf{E}_{\forall \vec{x}^j \in C_k} \vec{x}^j, \quad k = 1, 2, \dots \text{ e } \vec{m} = \mathbf{E}_{\forall \vec{x}^j} \vec{x}^j$$

Obs: Se quiser, mostre que é possível trabalhar com cada componente dos vetores separadamente, e use esta propriedade.

3 - Considere um conjunto de classes  $C_i$ ,  $i = 1, 2, \dots$  com baricentros  $\underline{w}_i$  e raios de similaridade  $r_i$  respectivamente.

3.1 - Mostre que a equação do plano separador entre duas classes cujos centros são  $\underline{w}_j$  e  $\underline{w}_k$  pelo critério do padrão mais próximo (critério 1) é

$$2 \underline{x}^t (\underline{w}_j - \underline{w}_k) + |\underline{w}_k|^2 - |\underline{w}_j|^2 = 0$$

3.2 - Mostre que a equação da esfera separadora da classe  $k$  pelo critério de similaridade mínima (critério 2) é

$$\underline{x}^t (\underline{x} - 2\underline{w}_k) + |\underline{w}_k|^2 - r_k^2 = 0$$

3.3 - Mostre que o domínio da classe  $C_k$  com baricentro  $\underline{w}_k$  e raio de similaridade  $r_k$  pelos critérios de centro mais próximo (critério 1) e de similaridade mínima (critério 2) é definido por todo  $\underline{x}$  que satisfaz simultaneamente as inequações

$$\underline{x}^t (\underline{x} - 2\underline{w}_k) \leq r_k^2 - \left| \underline{w}_k \right|^2$$
$$2\underline{x}^t (\underline{w}_i - \underline{w}_k) \leq \left| \underline{w}_i \right|^2 - \left| \underline{w}_k \right|^2 \quad \forall i \neq k$$

Observe que a pertinência a uma classe independe do raio de similaridade das demais classes.

4 – Mostre que a dispersão média intra classe da classe  $C_i$  classificada por uma camada de Kohonen pode ser calculada por

$$\sigma_i^2 = \underset{\forall \underline{x} \in C_i}{E} |\underline{x} - \underline{w}_i|^2 = - \frac{\sum_{\forall \tilde{x}} y_i u_i}{\sum_{\forall \tilde{x}} y_i}$$

note que os somatórios do segundo termo são sobre todos os elementos à classificar (e não apenas sobre os da classe  $C_i$ ).

**5** - Considere um classificador em que todas as entradas  $\underline{x}$  e todos os padrões de classe  $\underline{w}_i$  tem módulo unitário. Para este caso mostre que:

a -  $u_i = \underline{x}^t \underline{w}_i$  é uma medida de similaridade entre  $\underline{x}$  e  $\underline{w}_i$

b - a condição para que uma entrada  $\underline{x}$  pertença à uma classe com centro em  $\underline{w}$  e raio de similaridade  $r_0$  é que  $u_i = \underline{x}^t \underline{w}_i > 1 - (r_0^2 / 2)$ .

6 - Em uma camada de Kohonen o vetor sinapse do  $i$ -ésimo neurônio é  $\underline{w}_i = [w_{i1} \ w_{i2} \ \dots \ w_{in}]^t$ . Após o treinamento verificou-se que as sinapses  $w_{i2}$  eram iguais para todo  $i$ . Estas sinapses foram então zeradas, e todas as demais mantidas com seus valores inalterados. Explique:

a - Isto modifica o funcionamento da rede como classificador (isto é, alguma entrada  $\underline{x}$  pode vir a ser classificada diferentemente) ?

b - Se estava sendo usado um neurônio com excitação constante  $u_0$  para delimitar o raio das classes, isto continua a valer ? Interprete.