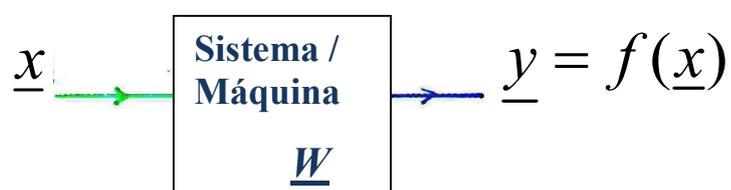


1 – Máquinas Matemáticas

Sistema / Máquina, Modelo

entrada \longrightarrow saída

mapeamento entrada / saída



Máquinas matemáticas - Mapeadores



Mapeadores lineares / não lineares

$$\underline{\tilde{y}} = \underline{M} \underline{x}$$

$$\underline{\tilde{y}} = \varphi(\underline{x})$$

2 - Tipos de máquinas

Aproximadores

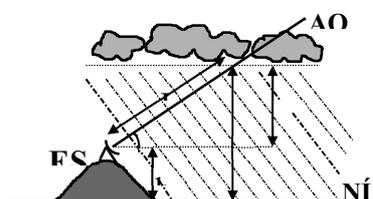
Associam entradas à saídas contínuas

Saídas contínuas $y_i \in (-1, +1)$

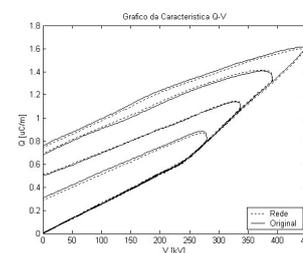
Conversor siderúrgico



Atenuação terra-satélite



Efeito corona



Classificadores

Associam a entrada à saídas lógicas (classificam as entradas)

Saídas lógicas $y_i \in \{0,1\}$

Iris

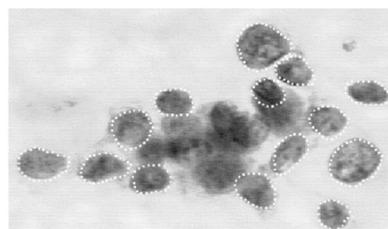


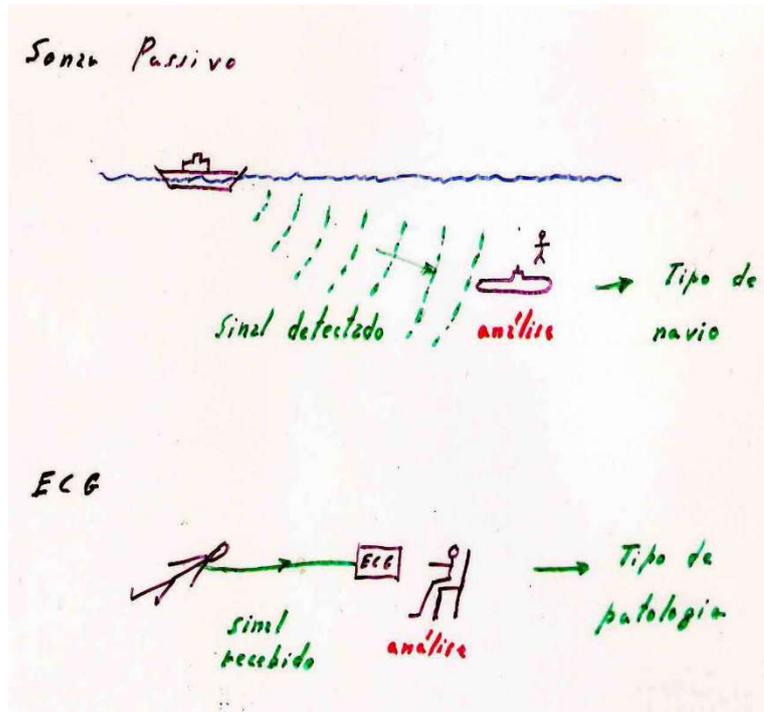
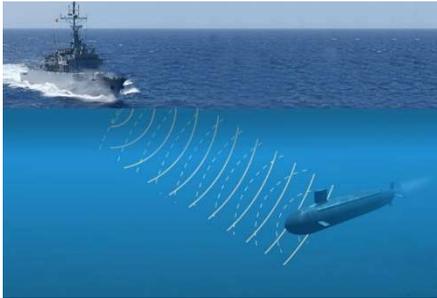
Iris Setosa

Iris Virginia

Iris Versicolor

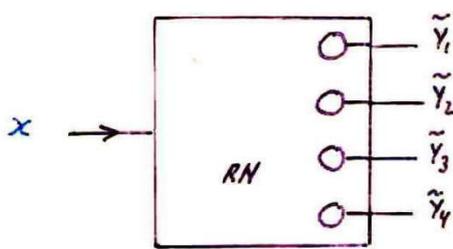
Cancer de mama





2.1 Classificadores: Convenção / Estratégia:

Codificação da Saída (maximamente esparsa)



	C_1	C_2	C_3	C_4
y_1	1	0	0	0
y_2	0	1	0	0
y_3	0	0	1	0
y_4	0	0	0	1

Cada saída separa a sua classe das demais, i.e.

classe vs não classe (um contra todos)

Separação de N classes



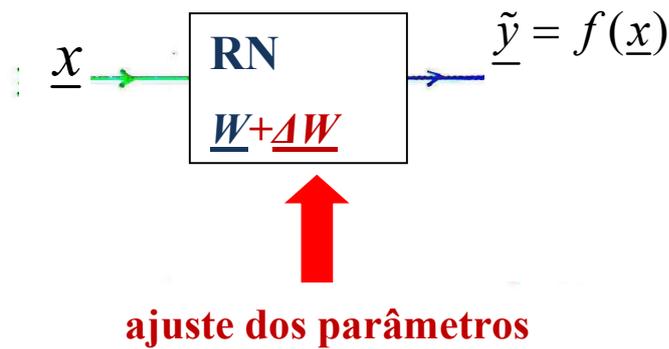
Separação de duas classes:

classe vs. não classe

3 - Treinamento (Aprendizagem) da máquina

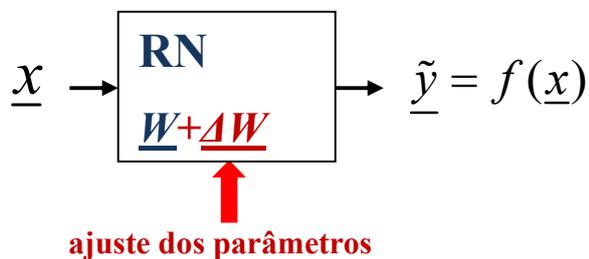
(será revisto em mais detalhes na disciplina CPE 721)

entrada → saída desejada

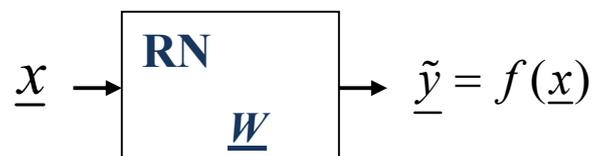


Treinamento vs Operação da máquina

Treinamento



Operação



Meta:

Otimizar (minimizar) uma função objetivo $F(\underline{w}) \geq 0$

tal que a saída \tilde{y} **seja a mais próxima possível da**

desejada y **para todas as entradas** x **consideradas.**

$$F(\underline{w}) = \underset{\forall x}{E} f[\underline{w}, \underline{x}, \tilde{y}(x), y(x)]$$

$$\text{p.ex. } F(\underline{w}) = \underset{\forall x}{E} |y(x) - \tilde{y}(x)|^2$$

Encontrar \underline{w}^*

$$\underline{w}^* \quad | \quad F(\underline{w}^*) \leq F(\underline{w}) \quad \forall \quad \underline{w} \neq \underline{w}^*$$

3.1 - Solução Analítica

$$\underline{w}^* = \begin{bmatrix} w_1^* \\ \dots \\ w_n^* \end{bmatrix} \quad | \quad \left. \frac{\partial F(\underline{w})}{\partial w_i} \right|_{w_i = w_i^*} = 0 \quad \forall i = 1, \dots, n$$

usualmente muito complexa para calcular analiticamente

3.2 - Solução numérica recursiva

$$\underline{w} \rightarrow \underline{w} + \underline{\Delta w} \quad | \quad F(\underline{w} + \underline{\Delta w}) < F(\underline{w})$$

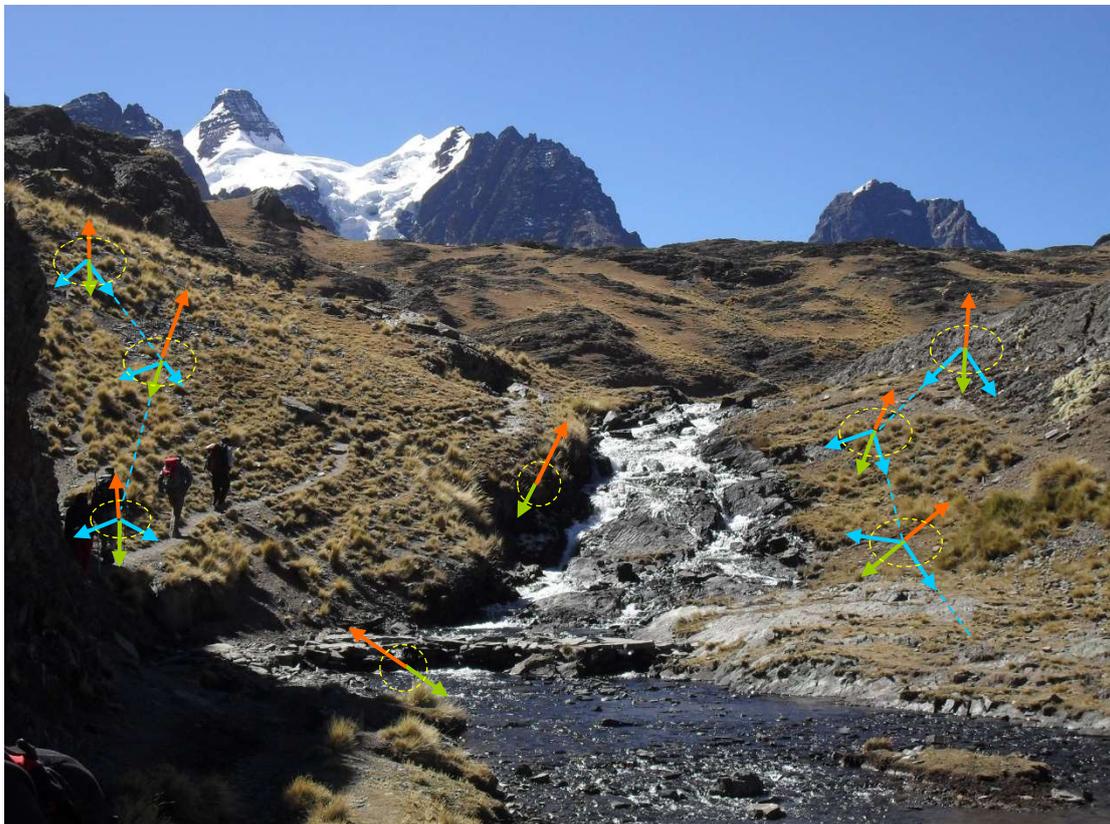
$\underline{\Delta w}$??

$$\underline{\Delta w} = \alpha \underline{d}$$

Passo
controla $|\underline{\Delta w}|$
 $\alpha > 0$, pequeno

controla
direção $\underline{\Delta w}$

Descida continuada, permanente - **decisão local**



Em que direção mover o ponto de operação para reduzir F_{in} ?

Gradiente:

$$F_0(w_1, w_2, \dots) = F_0(\underline{w})$$

$$\underline{w} = \begin{bmatrix} w_1 \\ w_2 \\ \vdots \end{bmatrix} \quad \nabla_{\underline{w}} F_0 = \underline{\nabla} = \begin{bmatrix} \frac{\partial F_0}{\partial w_1} \\ \frac{\partial F_0}{\partial w_2} \\ \vdots \end{bmatrix}$$

$$\underline{w} \Rightarrow \underline{w} + \Delta \underline{w} = \underline{w} + \alpha \underline{d}$$

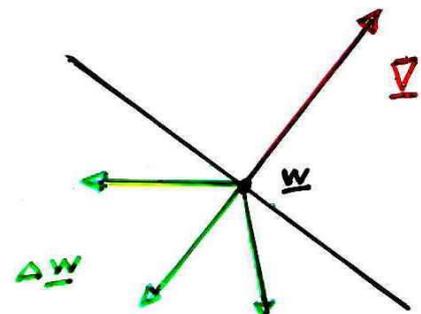
$$\Rightarrow F \Rightarrow F + \Delta F = F + \Delta \underline{w}^t \underline{\nabla} = F + \alpha \underline{d}^t \underline{\nabla}$$

Condições necessárias para reduzir F

α pequeno, **usualmente** $.1 \geq \alpha \geq .01$

direção \underline{d} condição necessária
 $\angle \underline{d}, \underline{\nabla} \in (90^\circ, 270^\circ)$

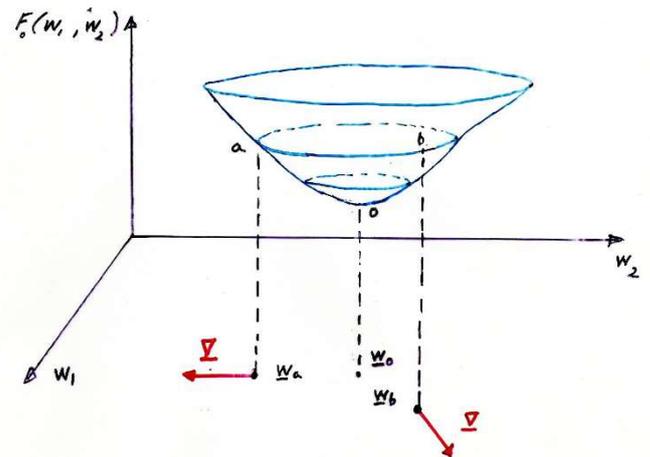
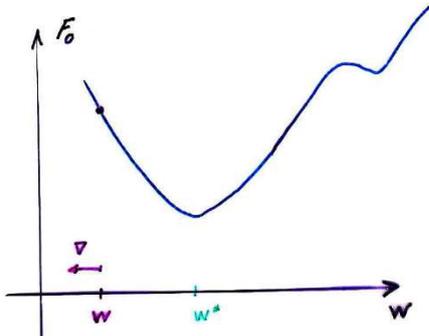
usualmente $\underline{d} = -\underline{\nabla}$



Mover o ponto de operação na mesma direção e no sentido contrário do gradiente reduz F

Gradiente descendente ou descida por gradiente

$$\underline{w} \longrightarrow \Delta \underline{w} = -\alpha \nabla(\underline{w}) \longrightarrow \underline{w} = \underline{w} + \Delta \underline{w}$$



3.3 - Aplicação dos passos de treinamento (solução numérica)

Regra delta

Para cada entrada \underline{x} calcular $\underline{\Delta w}$

Aplicar $\underline{\Delta w}$

Próxima entrada

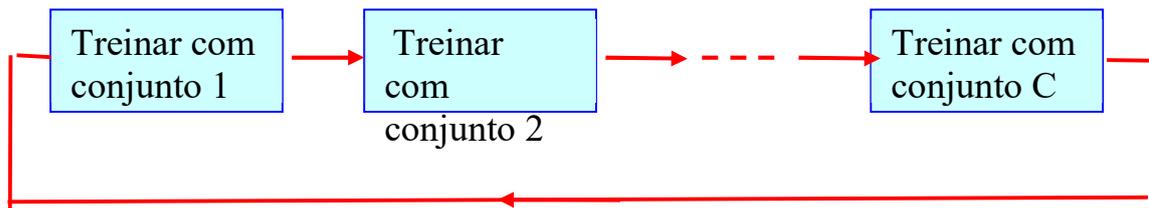
Próximo passo

Batelada

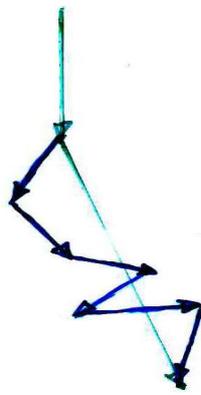
Para cada entrada \underline{x} calcular $\underline{\Delta w}$

Calcular e aplicar a média $\underline{E} \frac{\Delta \underline{w}}{\nabla \underline{x}}$

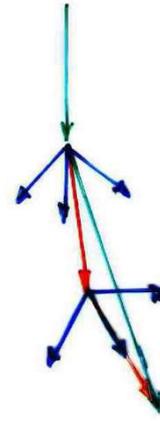
Próximo passo

Lotes - pequenas bateladas

batelada



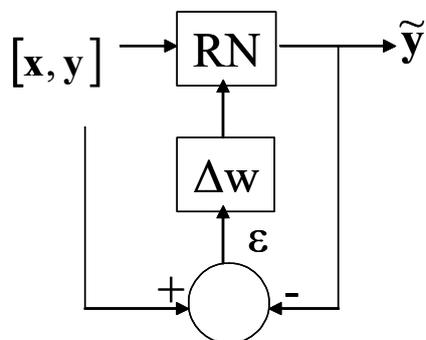
regra delta



lotes

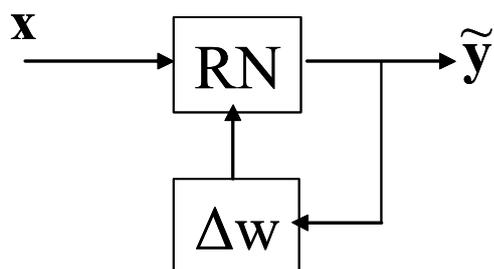
3.4 - Tipo de treinamento

- Treinamento supervisionado (com professor)

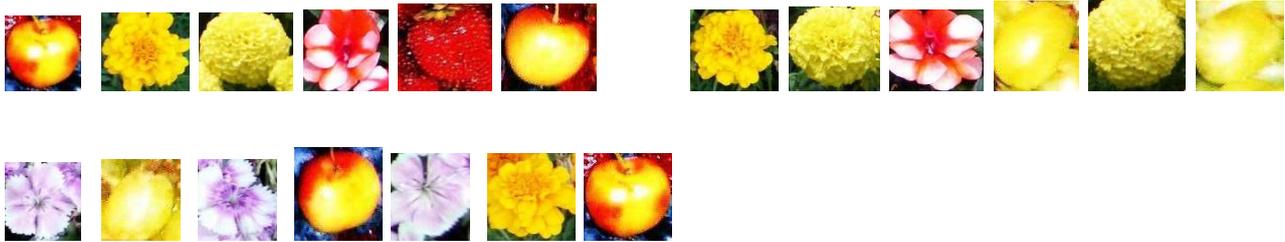


**Pode realizar mapeamentos arbitrários,
por critérios não evidentes**

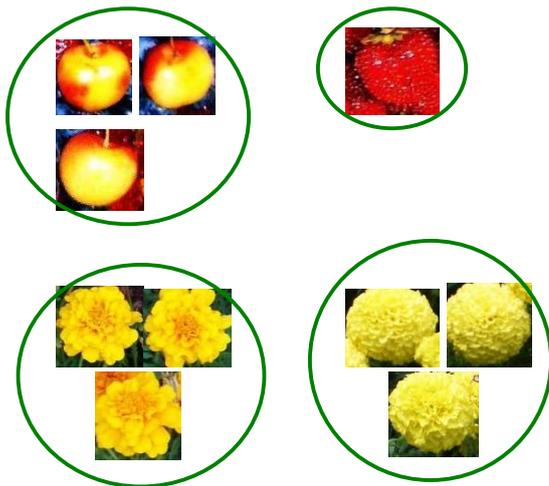
- Treinamento não-supervisionado (sem professor)



Necessita algum critério: Similaridade



Agrupamentos por similaridade



Agrupamentos arbitrários



5 - Classificação, Agrupamento ou Clusterização por Similaridade

Objetivo:

Agrupar, “de forma natural”,
conjuntos de dados com “similaridade interna”

Classes agrupam elementos `similares` entre si `de forma natural` mas

O que significa “de forma natural ?”

O que significa “similaridade interna ?”

Quais são as classes ?

Treinamento supervisionado

As classes são informadas

Ação: apenas classifica as entradas

Treinamento **não** supervisionado

As classes **não** são informadas

Ações: **determina as classes e**

classifica as entradas

Sobrevivência vs.

Classificação por similaridade

Aprendizado não supervisionado

Plasticidade

Animais e vegetais venenosos ou comestíveis ?



5 – Dados reais vs. Representação matemática na

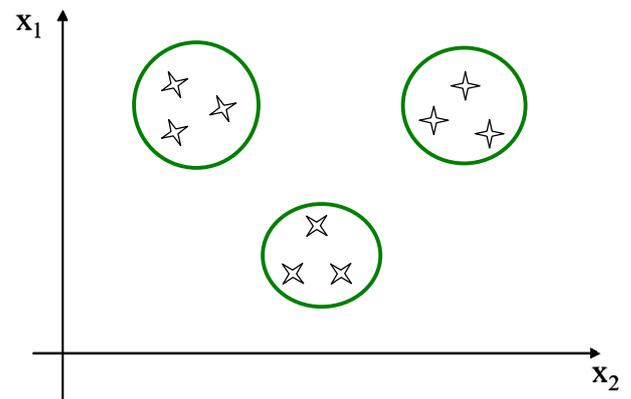
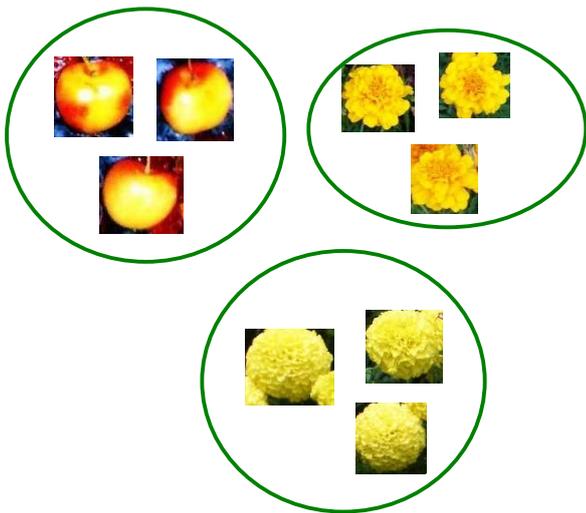
Classificação por Similaridade

objetos físicos \mathcal{O}_1 \mathcal{O}_2 \rightarrow objetos matemáticos
 vetores \underline{x}_1 \underline{x}_2

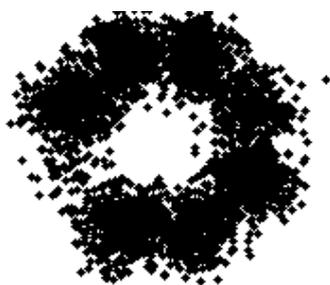
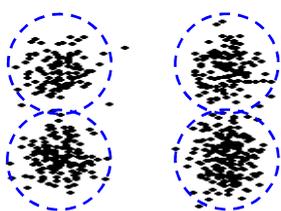
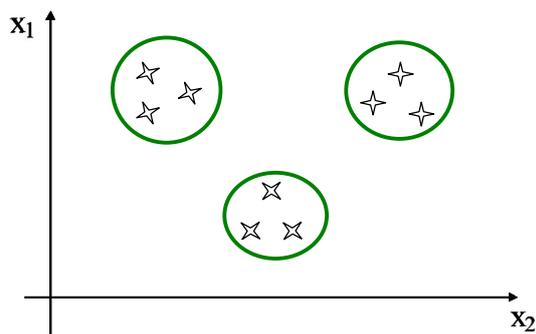
similaridade física \rightarrow similaridade matemática

$$\mathcal{O}_1 \approx \mathcal{O}_2 \quad \underline{x}_1 \cong \underline{x}_2 \quad \text{ou} \quad |\underline{x}_1 - \underline{x}_2| \ll \ll$$

$$\mathcal{O}_1 \approx \mathcal{O}_2 \quad \rightarrow \quad \underline{x}_1 \cong \underline{x}_2$$



Casos simples / complexos



**Classificar por similaridade : alocar na mesma classe
vetores que estão relativamente próximos entre si.**

Como traduzir matematicamente

“relativamente próximos entre si ?”

Como determinar as classes ?

Como realizar a alocação ?

