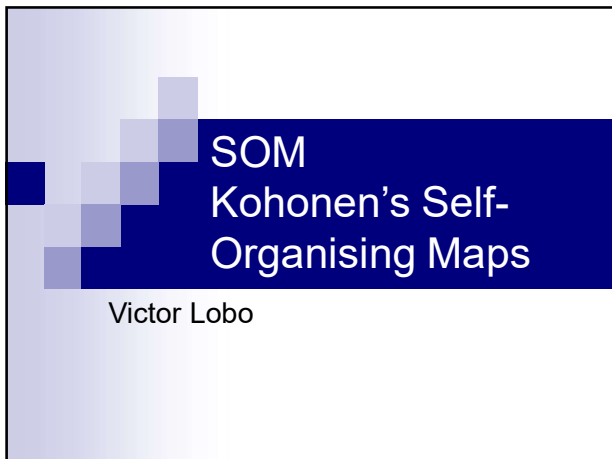


Mapas Auto-Organizados - SOM

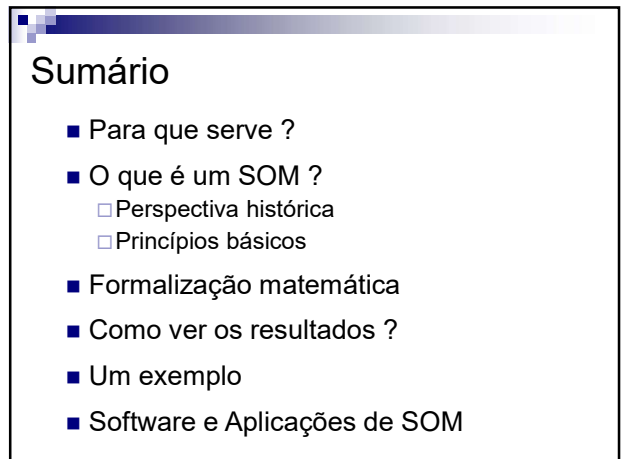
V 1.4 V.Lobo, EN 2021



SOM
Kohonen's Self-Organising Maps

Victor Lobo

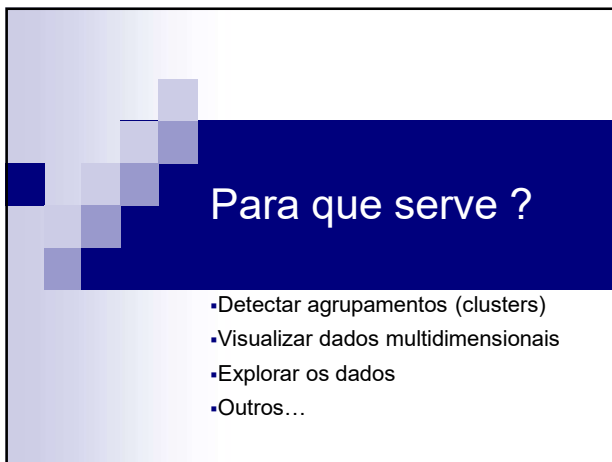
1



Sumário

- Para que serve ?
- O que é um SOM ?
 - Perspectiva histórica
 - Princípios básicos
- Formalização matemática
- Como ver os resultados ?
- Um exemplo
- Software e Aplicações de SOM

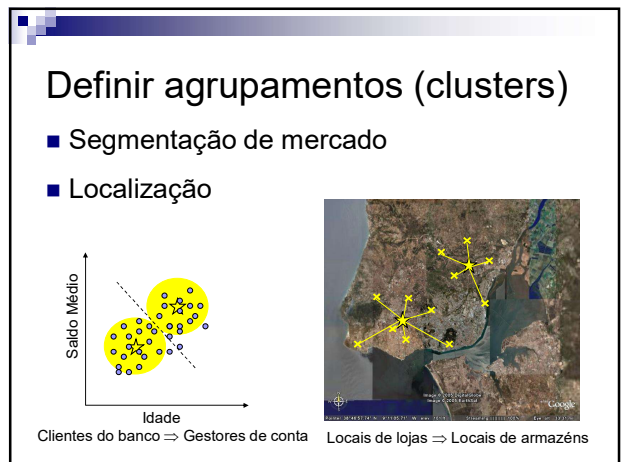
2



Para que serve ?

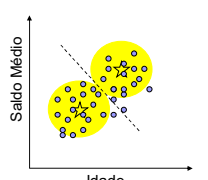
- Detectar agrupamentos (clusters)
- Visualizar dados multidimensionais
- Explorar os dados
- Outros...

3




Definir agrupamentos (clusters)

- Segmentação de mercado
- Localização

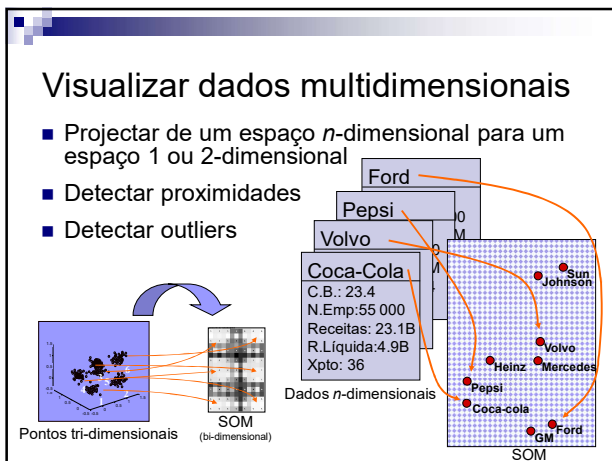


Clientes do banco ⇒ Gestores de conta



Locais de lojas ⇒ Locais de armazéns

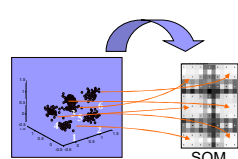
4



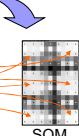
Visualizar dados multidimensionais

- Projectar de um espaço n -dimensional para um espaço 1 ou 2-dimensional
- Detectar proximidades
- Detectar outliers

Pontos tri-dimensionais

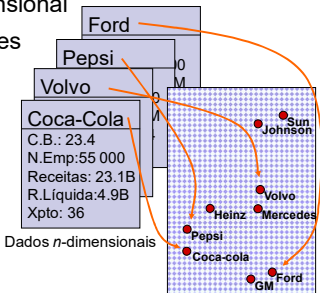


SOM (bi-dimensional)



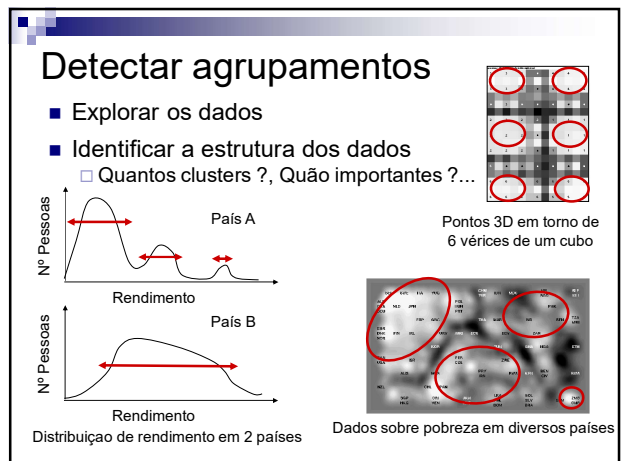
Dados n -dimensionais

Coca-Cola
C.B.: 23.4
N.Emp:55 000
Receitas: 23.1B
R.Líquida:4.9B
Xpto: 36



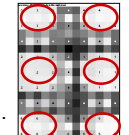
SOM

5

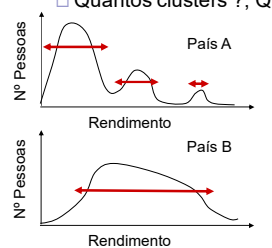


Detectar agrupamentos

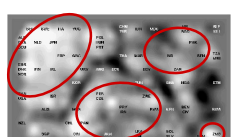
- Explorar os dados
- Identificar a estrutura dos dados
 - Quantos clusters ?, Quão importantes ?...



Pontos 3D em torno de 6 vértices de um cubo



Distribuição de rendimento em 2 países



Dados sobre pobreza em diversos países

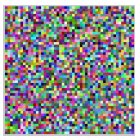
6

Mapas Auto-Organizados - SOM

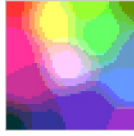
V 1.4 V.Lobo, EN 2021

Outros tipos de problemas...

- Resolver o TSP, controlar Robots, ordenar dados, interpolar pontos, classificar dados, extrair características, amostrar dados, alarmes, etc,etc,etc



Ordenar cores



Problema do caixeiro-viajante

7

O que é um SOM ?

- Perspectiva histórica
- Visão geral
- A matemática

8

Perspectiva histórica

Perspectiva histórica

- Prof. Tuevo Kohonen (Universidade Técnica de Helsinquia)
 - 1970s - Memórias associativas
 - 1982 - Primeiros artigos sobre SOM
 - 1988 - Livro sobre SOM, artigos sobre SOM no IEEE
 - 1990s - Grande divulgação
 - 1995,1997,2001 – Livro "Self Organizing Maps"
- Inspiração
 - Códigos para quantização de vectores
 - Memórias associativas
 - Preservar a topologia nos mapeamentos: padrões vizinhos devem ser mapeados para neurónios vizinhos

9

10

Inspiração biológica (Apenas interessante...)

- Sistemas biológicos têm auto-organização

- Evidência de:

- Estrutura de camadas no cérebro
- Cérebro organiza espacialmente a informação
- "Conceitos" similares são mapeados para áreas adjacentes
- Trabalho experimental com visão em animais sugere uma organização similar ao SOM no córtex



11

Visão geral do SOM

12

Mapas Auto-Organizados - SOM

V 1.4 V.Lobo, EN 2021

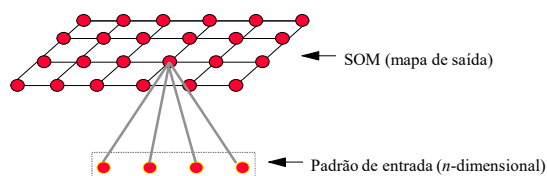
Algumas ideias gerais

- Rede neuronal
 - Conjunto de neurónios, ou UNIDADES
 - Aprendizagem não supervisionada (contrário do usual)
- TREINO da rede
 - A rede é construída, ou seja os seus parâmetros são ajustados, com base nas características dos dados disponíveis
- UTILIZAÇÃO da rede
 - Com a rede já treinada, podemos usá-la para muitas coisas...

13

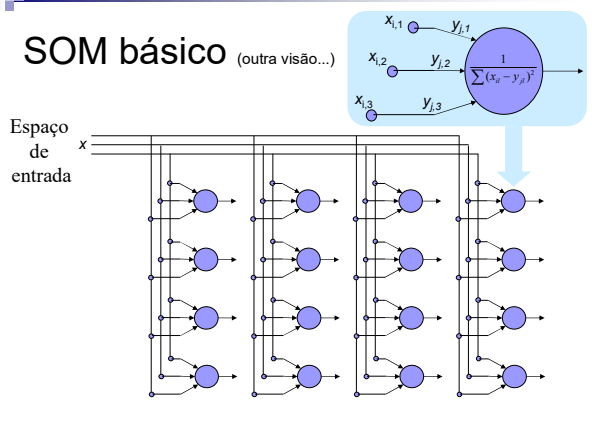
SOM básico

- Neurónios (**unidades**) dispostos numa grelha bi-dimensional
 - Pode ser uma grelha 1-dimensional (linha) ou m -dimensional ...
- Uma única camada
- Aprendizagem competitiva (quase "winner-take all")



14

SOM básico (outra visão...)



15

Espaço de entrada vs saída

- **Espaço de entrada** = espaço n -dimensional onde estão os dados
- **Espaço de saída** = espaço definido pela grelha das unidades
- Cada unidade é um ponto no espaço de saída e um **vector no espaço de entrada**, tal como os padrões de dados.

16

Treino da rede

- As unidades são **puxadas** para as posições dos dados, **arrastando** consigo as suas vizinhas no espaço de saída
- SOM \approx **superfície de borracha**, esticada e torcida de modo a passar pelos padrões de dados (ou pelo menos a ficar perto)

17

BMU- Best Matching Unit (Unidade vencedora)

- Padrões de entrada são comparados com todas as unidades; a mais próxima é considerada a **BMU**.
- Consideramos que o padrão de entrada é **mapeado** para a BMU.
- **A BMU actualiza-se** (de modo a aproximar-se mais do padrão de dados que representa), e os seus vizinhos actualizam-se também um pouco
- Há sempre uma ligeira diferença entre os dados e as BMUs que as representam. Essa diferença é o **erro de quantização**.

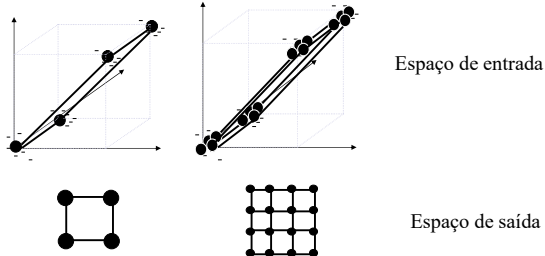
18

Mapas Auto-Organizados - SOM

V 1.4 V.Lobo, EN 2021

Exemplo 1: mapeamento de 3D para 2D

- Pontos agrupados em 4 cantos de um cubo



19

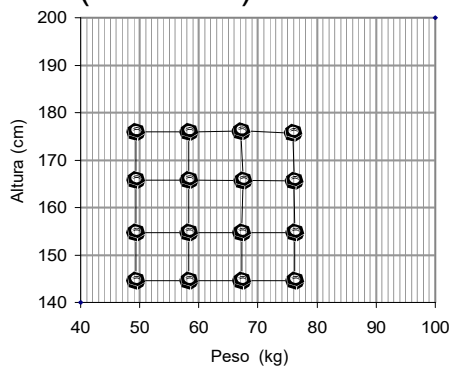
Exemplo 2:

- Uma rede "manual" de porcas...
□ (ver para crer...)

20

Exemplo 2 (simulado):

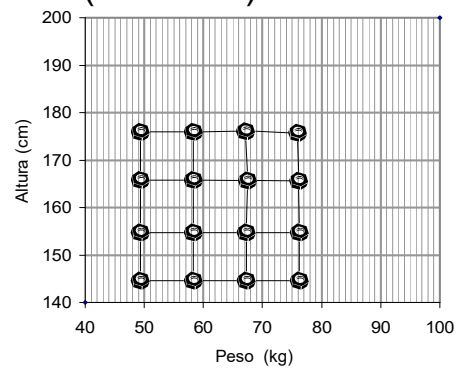
- Exemplo "físico"
 - Neurónios= porcas= unidades
 - Folha = espaço de input
- Problema:
- Analizar altura e peso das pessoas da turma



21

Exemplo 2 (simulado):

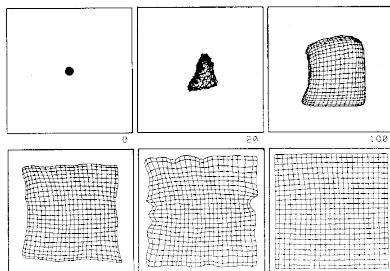
- Exemplo "físico"
 - Neurónios= porcas= unidades
 - Folha = espaço de input
- Problema:
- Analizar altura e peso das pessoas da turma



22

Exemplo 3: mapeamento de 2D para 2D

- Dados distribuídos uniformemente num quadrado
- Usado na demo do Matlab

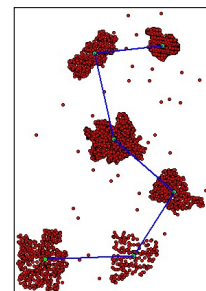


[Kohonen 95]

23

Exemplo 4: mapeamento de 2D para 1D

- [Animação](#)



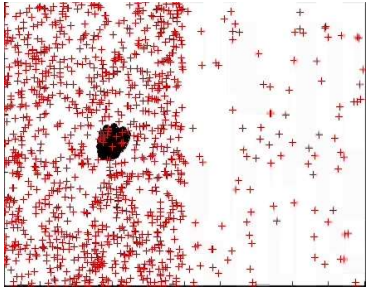
24

Mapas Auto-Organizados - SOM

V 1.4 V.Lobo, EN 2021

Exemplo 5: mapeamento de 2D para 2D

- Densidades diferentes



(animação)

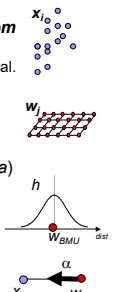
25

Formalização matemática

26

Os dados, a rede e a inicialização

- Seja:
 - $X = \{x_1, x_2, \dots, x_n\}$ o conjunto de dados de treino, com dimensão m .
 - $x_i = [x_i^1, x_i^2, \dots, x_i^m]^T$, onde x_{ij} são escalares de valor real.
 - W uma grelha de $p \times q$ unidades w_j
 - $w_j = [w_j^1, w_j^2, \dots, w_j^m]^T$
 - Valores de w_j escolhidos aleatoriamente na zona dos dados
 - $h(w_j, w_j, r)$ uma função real (denominada de vizinhança)
 - Quando $\|w_r - w_j\|_{(na\ grelha)} \rightarrow \infty$, $h(w_j, w_j, r) \rightarrow 0$
 - r determina o raio (zona de influência)
 - α a taxa de aprendizagem
 - $0 \leq \alpha \leq 1$
 - A começa grande e vai diminuindo ao longo do treino



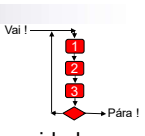
27

Algoritmo de treino

- Para todos os $x_i \in X$:
 - Calcular** a distância entre x_i e todas as unidades w ($d_{ij} = \|x_i - w_j\|$)
 - Escolher** a unidade vencedora (BMU)

$$w_{bmu} : d_{i,bmu} = \min(d_{ij})$$
 - Atualizar** cada unidade de acordo com a regra

$$w_j = w_j + \alpha \cdot h(w_{bmu}, w_j, r) \cdot \|x_i - w_j\|$$
- Repetir** o processo, diminuindo o valor de α e r , e percorrendo várias vezes todos os dados de treino, até atingir um critério de paragem.



28

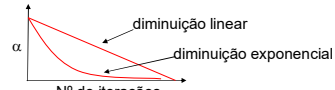
Principais decisões a tomar:

- Quantas unidades? Que tipo de grelha? Que dimensão?
- Quantas iterações?
- Que tipo de função de vizinhança? Que valor inicial para r ? Que valor final?
- Que valor para α inicial?

29

Taxa de aprendizagem α

- $0 \leq \alpha \leq 1$
- Regula a **plasticidade** da rede
 - Valores altos \Rightarrow A rede move-se muito, adapta-se "depressa"
 - Valores baixos \Rightarrow A rede move-se pouco e estabiliza
- Começa com valores altos e vai diminuindo até chegar a 0




30


Mapas Auto-Organizados - SOM

V 1.4 V.Lobo, EN 2021

Função de vizinhança

- A forma
 - Gaussiana
 - Rectangular (bolha)
 - Rampa
 - Outras
- Responsável pela ordenação topológica
 - Faz com que haja "ligações laterais" entre unidade próximas na grelha

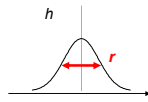
$$h_g(w_{pq}, w_{mn}, r) = e^{-\frac{\sqrt{(p-n)^2 + (q-m)^2}}{r}}$$


$$h_g(w_{pq}, w_{mn}) = \begin{cases} 1 & \text{if } \sqrt{(p-n)^2 + (q-m)^2} \leq r \\ 0 & \text{if } \sqrt{(p-n)^2 + (q-m)^2} > r \end{cases}$$


31

Raio r da função de vizinhança

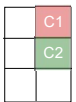
- Função to tempo (ou nº de iterações) $r(t)$
 - Raio grande \Rightarrow Muitas unidades actualizadas \Rightarrow Permite desdobragem
 - Raio pequeno \Rightarrow Só vizinhos próximos são actualizados \Rightarrow Ajuste fino
- Valores iniciais de r
 - 1ª fase – semelhante ao tamanho da rede
 - 2ª fase – semelhante ao raio dos clusters que esperamos obter
- Valor final de r
 - 0 \Rightarrow Ajuste bom (k -médias)
 - 1 \Rightarrow Manter a ordenação, efeito de fronteira

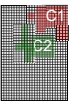


32

Tamanho do SOM

- SOM de k -médias
 - Tem poucas unidades
 - 1 unidade para cada cluster esperado
- SOM emergente
 - Muitas unidades para cada cluster esperado
 - Permite representar clusters "complicados", e "variados"
 - Permite entender a estrutura dos dados, detectar o número de clusters, etc...




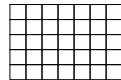


33

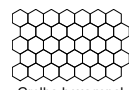
Dimensão e tipo de grelha

- Grelha unidimensional (linha)
 - Substitui k -médias
 - Permite estabelecer uma ordenação dos dados
- Grelha 3D ou n -dimensional
 - Difícil de visualizar
- Grelha bidimensional
 - A mais usada
 - Grelha quadrada
 - Grelha hexagonal ou triangular
 - Induz menos distorção





Grelha rectangular

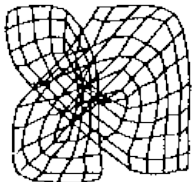


Grelha hexagonal

34

Nº Iterações e desdobramento

- Nº Iterações
 - Épocas vs indivíduos
 - Na dúvida... escolher mais!
- Problemas de desdobramento
 - Existem mínimos locais
 - Medidas de erro topológico
- Soluções
 - Várias inicializações
 - 2 fases (unfolding+fine tuning)
 - Observar erro topológico e erro de quantização



[Ritter 92]

35

Aspectos teóricos

- Função de energia minimizada: [Hertz 91]

$$V(w) = \frac{1}{2} \sum_x \sum_i \Lambda(i, i^*) |\bar{x} - \bar{w}|^2 = \frac{1}{2} \sum_x \sum_k M_{xk} \sum_i \sum_j \Lambda(i, k) (x_j - w_i)^2$$
 - Altamente não-linear, e não é uma função global devido ao conceito de unidade vencedora.
 - Resultados mais completos para dimensão 1 [Cottrell]
 - Algumas boas aproximações para 2D [Ritter]
 - Há regras de aprendizagem "bem comportadas" [Heskes]
 - Há outras alternativas "bem fundamentadas" [Bishop]
- Factor de magnificação
 - Densidade de neurónios \propto (densidade dos dados) k , com $k < 1$
 - Existe um factor de "ampliação" das zonas com menor densidade

} Pouco usadas

36

Mapas Auto-Organizados - SOM


V 1.4 V.Lobo, EN 2021



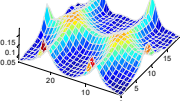
Como ver os resultados ?

- Matrizes U (U-MAT)
- Calibração (ou etiquetagem)
- Planos de componentes
- Outros

37

Matrizes U (U-MAT) [Utsch 93]

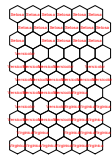
- Permite-nos identificar clusters 
- Calcula a distância, no espaço de entrada, entre vizinhos no espaço de saída
 - Distâncias codificadas em cores
 - Valores baixos \Rightarrow Neurónios próximos \Rightarrow cluster
 - Valores altos \Rightarrow Neurónios longe \Rightarrow Espaço vazio*

Ideal  U-mat real  

38

Calibração (ou etiquetagem)

- Objectivo
 - Identificar *o que são* os clusters
 - Fazer classificação supervisionada
 - LVQ seria melhor...
- Como ?
 - Se os dados de treino tiverem classes associadas...
 - ...as suas BMU podem herdar essas classes



39

Planos de componentes

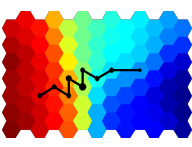
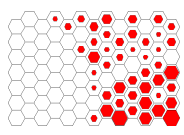
- Permite ver como é que uma dada variável (ou componente) varia ao longo do mapa
- Permite perceber o que define os clusters ou zonas, ou o que as forma



40

Outras visualizações

- Hits
 - Identificar quantos dados são mapeados para cada unidade
- Trajectórias
 - Observar como varia o BMU



41

Exemplo

- Dados "artificiais"

42

Mapas Auto-Organizados - SOM

V 1.4 V.Lobo, EN 2021

Dados de partida

- Pontos num espaço tri-dimensional
- Gerados com alguma dispersão em torno de 6 cantos de um cubo
- Código MATLAB
 - Com SOMTOOLBOX

43

Treino de um mapa de 3x2

- Parâmetros
 - Grelha rectangular (3x2)
 - Inicialização linear
 - r inicial = 2
 - r final = 0
 - α inicial = 0.1
 - N^oIterac. = 900 (1,5 épocas)

44

Visualização no espaço de output

0.730/0.44/0.61	0.800/0.66/0.62	5	6
0.67/0.49/0.43	0.73/0.71/0.41	3	4
0.56/0.52/0.29	0.67/0.74/0.25	1	2

- Durante o treino
- Depois de calibrado com os dados originais

45

Treino de um mapa de 30x20

- Parâmetros
 - Grelha rectangular (30x20)
 - Inicialização linear
 - r inicial = 15
 - r final = 0
 - α inicial = 0.1
 - N^oIterac. = 900 (1,5 épocas)

46

U-MAT e planos de componentes

Depois de calibrado com os dados originais

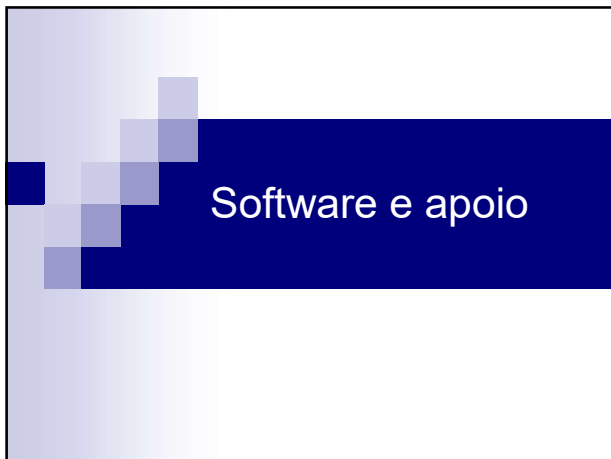
47

Maior variância...

48

Mapas Auto-Organizados - SOM

V 1.4 V.Lobo, EN 2021



49

Software disponível

- SOM-PAK
 - (http://www.cis.hut.fi/research/som_lvq_pak.shtml)
 - Código C, compilável em UNIX ou MS-DOS
 - Rápido e fiável, fácil de utilizar
- Somtoolbox para MATLAB
 - (www.cis.hut.fi/projects/somtoolbox)
 - Boa visualização, facilmente alterável, "ideal" para I&D
- Muitos outros
 - SAS Enterprise Miner, SPSS-Clementine, IBM Intelligent Miner, Weka, etc...

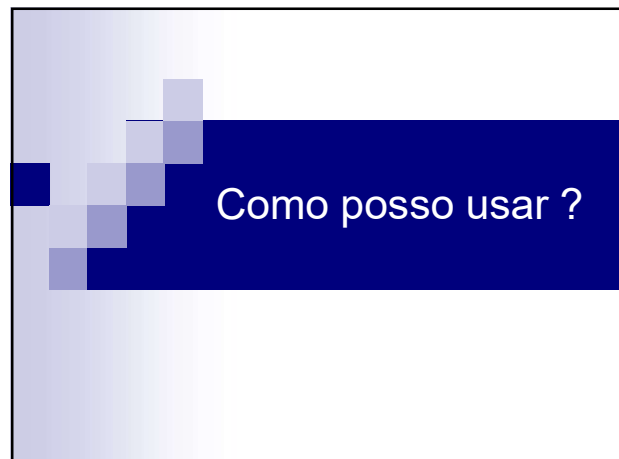
© DSOM: Our very own software! ©

50

Bibliografia e apoio

- "Self-Organizing Maps", Prof. Tuevo Kohonen
 - Springer-Verlag 2001
- Site da Universidade Técnica de Helsínquia
 - (www.cis.hut.fi/projects/somtoolbox/links)
 - **Public-domain software**
 - SOM-PAK para DOS, SOM Toolbox para MATLAB
 - **Manuais, guias, e documentação**
 - **Bibliografia extensa**
 - "www.cis.hut.fi/research/som-bibli"
 - 5384 referências em Abril de 2005

51



52

SOM-PAK 3.1

- Colecção de programas que recebem parâmetros através do comando-linha:
 - randinit- Inicializa um SOM
 - vsom - Treina um SOM
 - qerror - Erro de quantização
 - visual - Calcular o neurónio vencedor para cada dado
 - vcal - Calibrar um SOM (dar nomes aos neurónios)
 - umat - Calcular a U-Matrix de um SOM
 - plane - Visualizar uma das dimensões de um SOM

53

Formato dos dados

- Texto simples
 - Primeira linha tem informação de controlo
 - N° de atributos de cada padrão
 - Outros: tamanho da rede, função de vizinhança, etc
 - Linhas com dados
 - Valores dos atributos, opcionalmente com nomes
 - Comentários
 - #

```
3
#Isto é um comentário
10 10 5 label_1
3 10 4
7 8 2 label_2
```

54

Mapas Auto-Organizados - SOM

V 1.4 V.Lobo, EN 2021

Sessão típica

1) Inicialização do SOM

- `randinit` : Inicializa com valores aleatórios

2) Treino do SOM

- `vsom` : Implementa o algoritmo de treino
- Usado 2 vezes:
 - Primeiro com vizinhanças e ritmos de aprendizagem grandes, para permitir o **desdobramento**
 - Depois, com pequenas vizinhanças e ritmos de aprendizagem para ajustes finos

55

Sessão típica

3) Medição do erro de quantização (opcional)

- `qerror` : O mapa está a representar bem os dados ?

4) Visualização

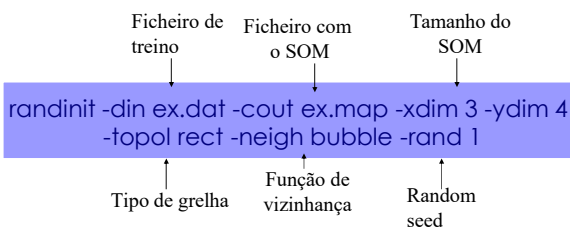
- `vcal` : Põe nomes (ou classes, ou labels) nos neurónios (só para problemas supervisionados)
- `visual`: Encontra o vencedor para cada padrão
 - Usar outro programa para visualizar de facto...
- `umat` : desenha a U-Matrix em ps
- `plane` : visualiza um dos planos

56

Inicialização do mapa

■ `randinit`

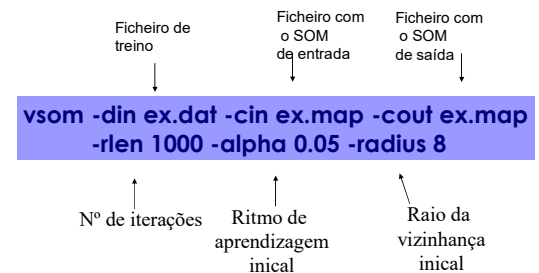
- Usa um ficheiro de dados como protótipo



57

Treino do SOM

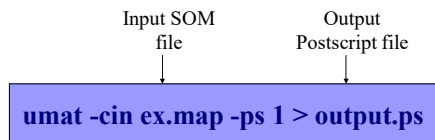
■ `Vsom` – Correr duas vezes !



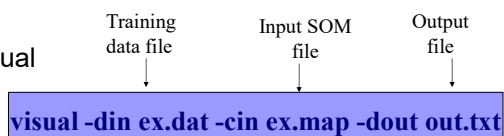
58

Visualização dos resultados

■ U-Mat



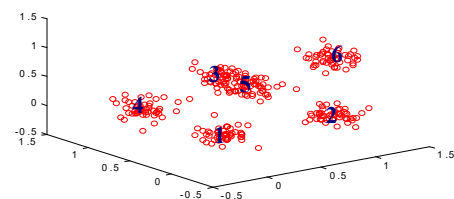
■ `visual`



59

Exemplo prático

- Vamos mapear um conjunto de pontos de um espaço 3-dimensional, que estão em vértices de um cubo



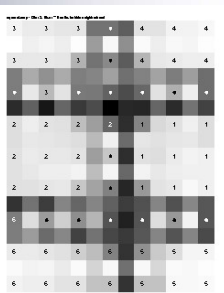
60

Mapas Auto-Organizados - SOM

V 1.4 V.Lobo, EN 2021

Exemplo prático

- 306 pontos centrados em 6 vértices com $\sigma=0.1$
 - Áreas escuras indicam fronteiras entre clusters



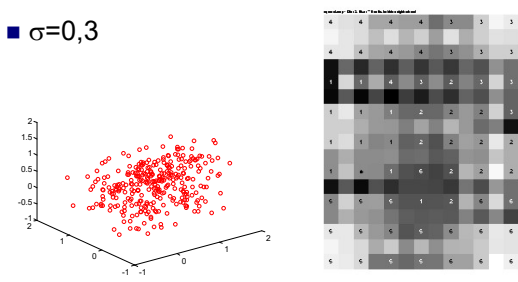
```

randinit -xdim 7 -ydim 9 -din square.dat -cout square.map -topol rect -neigh bubble
vsom -din square.dat -cin square.map -cout square.map -rlen 1000 -alpha 0.1 -radius 7
vsom -din square.dat -cin square.map -cout square.map -rlen 10000 -alpha 0.02 -radius 3
vcal -din square.dat -cin square.map -cout square.map
umat -cin square.map -ps 1 > square.ps
    
```

61

Outros testes

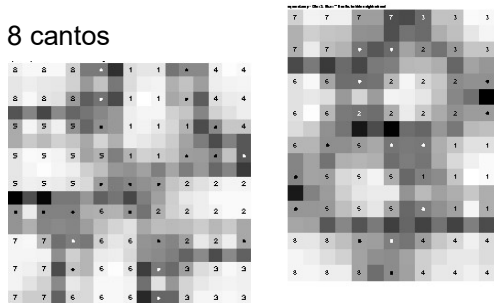
- $\sigma=0,3$



62

Outros testes

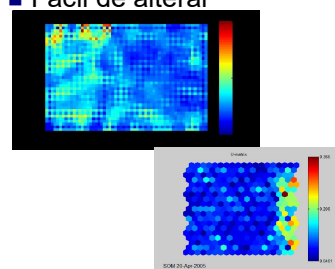
- 8 cantos



63

Testes com o SOMToolbox

- Excelentes gráficos
- Fácil de alterar



```

load conchos2
sDnsom_data_struct(data);
sM = som_randinit_geod(sD,'msize',[5 10] 'rect','sheer');
sM neigh = 'bubble';

iterations_1 = 1000; iterations_2 = 2000;
radius_in1 = 5; radius_in2 = 2;
alpha_in1 = 0.7; alpha_in2 = 0.1;

% Standard SOM
sM1 = som_septrain(sM,sD,...
    'radius_in1',radius_in1,...
    'radius_in2',...
    'alpha_in1',alpha_in1,...
    'trainlen', iterations_1, 'epochs');
sM1 = som_septrain(sM1,sD,...
    'radius_in1',radius_in1,...
    'radius_in2',...
    'alpha_in1',alpha_in2,...
    'trainlen', iterations_2, 'epochs');
save conchos_som sM1
[get 1,te] [srom_quality(sM1,sD);
srom_show(sM1,'umat','all','footnode','Standard SOM')
sDnsom_umat(sM1);
    
```

64

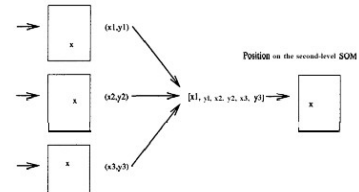
Outros trabalhos com SOM

65

SOM multicamada e Redução da dimensionalidade

- Análise de relatórios Finaceiros [Kiviluoto 98]
 - Coordenadas do neurónio vencedor são os dados para a camada seguinte

- Por vezes todos os dados do nerónio vencedor são usados



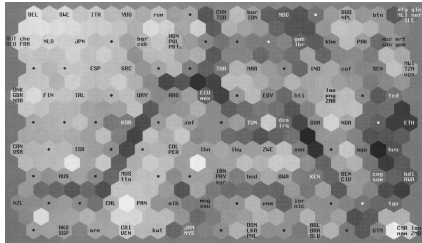
66

Mapas Auto-Organizados - SOM

V 1.4 V.Lobo, EN 2021

Som Para clustering

- Pobreza no mundo [Kohonen 95]
 - Agrupa países de acordo com uma série de indicadores económicos

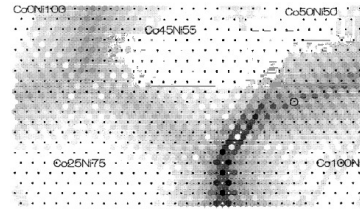


[Kohonen 95]

67

Som para clustering/classificação

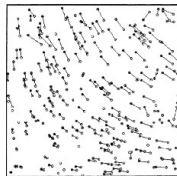
- Análise química (dados de espectroscopia) [Tokutaka 98]



68

SOM para seguimento

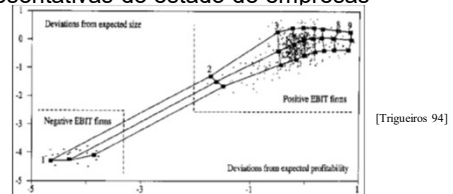
- Análise do movimento de fluidos [Labonté 98]
 - Segue a posição de partículas num fluido
 - Analisa a trajectória dos neurónios durante a aprendizagem



69

SOM para amostragem

- Avaliação de empresas (contabilidade) [Trigueiros 94]
 - Mapeamento de 2D para 2D
 - Usado para seleccionar situações representativas do estado de empresas

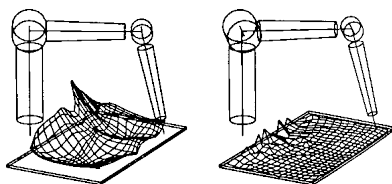


[Trigueiros 94]

70

SOM em robótica

- Controlo de um braço robot [Ritter 91]
 - SOM contém os sinais de controlo necessários para atingir uma dada localização



[Ritter 91]

71

SOM para monitorização

- Trajectórias em SOM [Shu-Ching Kuo 2004]
 - Shu-Ching Kuo 2004

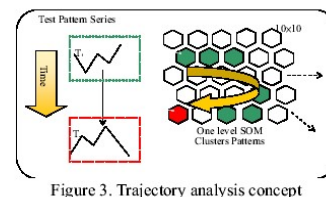


Figure 3. Trajectory analysis concept

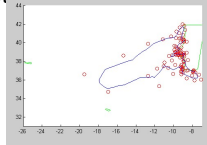
72

Mapas Auto-Organizados - SOM

V 1.4 V.Lobo, EN 2021

SOM para o problema do caixeiro viajante

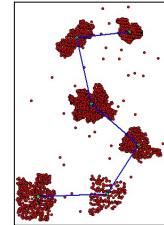
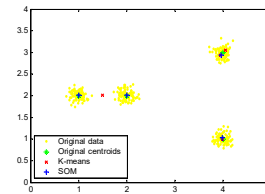
- Problema exacto
 - Passar exactamente por n locais, fazer o planeamento de C.I. [Hueter 88][Choy 95]
- Problema relaxado
 - Determinação de rotas para patrulhas [Lobo 2005]



73

SOM para k-médias

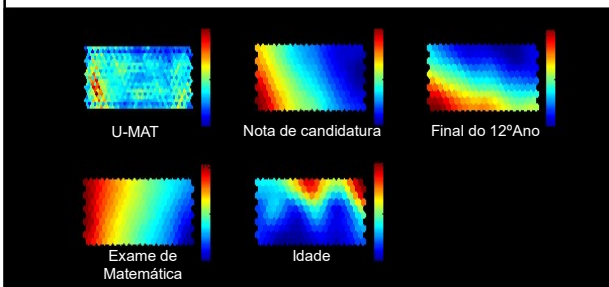
- SOM 1-dimensional, com k -neurónios
- Serve de inicialização robusta [Bação 05]



74

SOMs e planos de componentes

- Permitem ver as diferentes variáveis



75

Outras aplicações de SOM

- Previsão
 - Consumos energéticos [Osowski 98]
- CIM
 - Agrupamento de ferramentas [Guerrero 98]
- Monitorização de processos (análise de trajectórias no espaço de saída), do estado de condição de máquinas, reconhecimento de voz, análise de imagem, estudos musicais, desenho de circuitos, análise geopolítica, linguística, pesquisa na Web, Economia, biologia, química, ...

76

Problema para casa: Iris

- A flor Iris tem várias variantes, 3 das quais são:
 - 1 -Iris Setosa
 - 2 -Iris Versicolour
 - 3 -Iris Virginica
- Para 50 flores de cada uma das variantes foram medidas 4 características (medidas em cm)
 - Largura da pétala
 - Comprimento da pétala
 - Largura da Sépala
 - Comprimento da Sépala



Iris Setosa

- Questões:
 - É possível determinar a variante a partir desses 4 parâmetros usando um SOM ?
 - Quão parecidas são as 3 espécies ?

77



78