

# Mapas Auto-Organizáveis e Aprendizado Não-Supervisionado

1	Organização e Ordem .....	3
2	Auto-Organização .....	4
3	Complexidade de sistemas .....	9
4	Complexificação: evolução da complexidade .....	9
5	Exemplos de auto-organização na natureza .....	11
6	Motivação para treinamento não-supervisionado: clusterização .....	16
6.1	Modelo simples de classificação para dados rotulados.....	16
6.2	Modelo composto de classificação para dados rotulados.....	17
6.3	Modelo composto de classificação para dados não-rotulados .....	19
6.4	Comparação com o k-means .....	20
7	Agrupamento em aprendizado de máquina .....	21
8	Treinamento não-supervisionado.....	25
9	Mapas Auto-Organizáveis de Kohonen.....	28

9.1	Arranjo unidimensional .....	29
9.2	Arranjo bidimensional .....	33
9.3	Fase de competição .....	37
9.4	Fase de aprendizado não-supervisionado .....	38
9.5	Algoritmo de ajuste dos pesos .....	40
9.6	Ajuste de pesos considerando a vizinhança .....	41
9.7	Discriminação dos agrupamentos .....	46
9.8	Aplicação .....	47
9.9	Agrupamento de dados (Clusterização) .....	50
9.10	Mapa contextual .....	59
9.11	Ferramentas de visualização e discriminação .....	62
9.12	Comparação com PCA e Análise por Componentes Curvos (CCA) .....	67
9.13	Ordenamento de pontos em espaços multidimensionais.....	73
9.14	Roteamento de veículos (múltiplos mapas auto-organizáveis) .....	76
9.15	Mapas auto-organizáveis construtivos .....	77
9.16	Questões a serem melhor investigadas .....	78
9.17	Extensões .....	79
10	Referências.....	79

## 1 Organização e Ordem

- Um cristal tende a apresentar mais ordem entre seus componentes que uma célula. No entanto, a célula tende a apresentar mais organização que um cristal.
- Um papel de parede que apresenta um padrão de repetição tende a apresentar mais ordem espacial que uma pintura. No entanto, a pintura tende a apresentar mais organização espacial.
- Um som de alarme tende a apresentar mais ordem temporal que uma música. No entanto, a música tende a apresentar mais organização temporal.
- Objetos estão dispostos no tempo e/ou no espaço de forma ordenada se eles seguem uma regra específica de disposição temporal e/ou espacial.
- Objetos estão dispostos no tempo e/ou no espaço de forma organizada se eles contribuem em conjunto para produzir alguma funcionalidade.
- Ideias podem estar logicamente ordenadas (para efeito de apresentação) ou logicamente organizadas (para efeito de argumentação).

## 2 Auto-Organização

- O estudo de sistemas auto-organizados é recente, embora a humanidade tenha sempre se ocupado com questões vinculadas à origem de sistemas organizados.
- As formas que podem ser observadas no mundo à nossa volta representam apenas uma pequena parcela de todas as formas possíveis. Logo, por que não existe mais variedade?
- Para procurar respostas a questões como esta é que se estudam sistemas auto-organizados e teoria da complexidade (ADAMI, 2002).
- Exemplos de sistemas naturais que apresentam organização: galáxias, planetas, componentes químicos, células, organismos, sociedades.
- Existem campos de atuação científica que procuram explicar a auto-organização recorrendo a propriedades e leis aplicáveis às partes constituintes (componentes) de um sistema organizado específico.

- No entanto, a auto-organização pode ser abordada de modo completamente distinto, recorrendo-se às propriedades e leis comuns a todos os sistemas organizados, independente de suas particularidades.
- Neste caso, a atual disponibilidade de recursos computacionais é fundamental para viabilizar a investigação dos processos envolvidos, através de simulações que envolvem um grande número de etapas e uma grande variedade de parâmetros e condições iniciais.
- Mesmo assim, o estudo está restrito a fenômenos (concretos ou abstratos) que são facilmente reproduzíveis, os quais certamente representam um subconjunto de todos os fenômenos possíveis.
- A reprodução em computador de fenômenos auto-organizados tem levado à geração de teorias que procuram descrever sistemas complexos e sua organização espontânea (sistemas parcialmente decomponíveis).

- Um sistema complexo pode ser caracterizado como o resultado da auto-organização de componentes sob forte interação, produzindo estruturas sistêmicas cujas propriedades geralmente não estão presentes em nenhum de seus componentes, já que estas dependem de níveis mais elevados de organização.
- Um sistema, por sua vez, pode ser definido como um agrupamento coerente de componentes que operam como um todo e que apresentam uma individualidade, ou seja, se distinguem de outras entidades por fronteiras reconhecíveis. Há muitas variedades de sistemas, as quais podem ser classificadas em 3 grandes grupos:
  1. Quando as interações de seus componentes são fixas. Ex: máquina.
  2. Quando as interações de seus componentes são irrestritas. Ex: gás.
  3. Quando existem interações fixas e variáveis de seus componentes. Ex: célula.
- Os sistemas de maior interesse aqui são aqueles pertencentes à classe 3, já que dependem da natureza e forma das interações de seus componentes ao longo de sua existência (MCSHEA, 2000). Assim, o sistema vai apresentar um novo

comportamento sempre que componentes forem adicionados, removidos ou rearranjados, ou então sempre que houver modificação nas interações.

- A essência da auto-organização está no surgimento de estrutura (formas restritas) e organização sem que estas sejam impostas de fora do sistema. Isto implica que este fenômeno é interno ao sistema, resulta da interação de seus componentes e, em essência, não depende da natureza física destes componentes.
- A organização pode se dar no espaço, no tempo, ou em ambos.
- O que se busca são regras gerais para o crescimento e evolução de estruturas sistêmicas. Com isso, espera-se poder prever a organização futura que irá resultar de alterações promovidas junto aos componentes de um dado sistema, além de poder estender estes resultados a outros sistemas semelhantes.
- Em geral, os mecanismos estabelecidos pelos componentes de um dado sistema capaz de expressar auto-organização são: **realimentação positiva, realimentação negativa e interação local** (MITLETON-KELLY, 2003).

- É importante buscar respostas (direta ou indiretamente) para as seguintes questões:
  - Uma vez caracterizadas as regras de interação que irão definir o comportamento individual de cada componente de um sistema, ao longo do tempo, e que supostamente permitem desenvolver e manter um processo auto-organizado, ocorrerá efetivamente algum tipo de auto-organização do comportamento espaço-temporal?
  - Se a resposta à questão acima for afirmativa, qual será o padrão emergente, ou seja, o resultado final da auto-organização?
  - Dado que existe, na natureza, a formação de padrões espaço-temporais não devidos a processos auto-organizados, como diferenciá-los de um processo auto-organizado?
  - Qual é o nível de complexidade local que cada componente deve apresentar para permitir a emergência da complexidade global observada no sistema?
  - O quanto do comportamento observado no agregado é devido ao efeito da ação individual de cada componente e o quanto é devido à interação desses componentes?



### 3 Complexidade de sistemas

- “À medida que a complexidade de um sistema aumenta, nossa habilidade de fazer afirmações precisas e que sejam significativas acerca deste sistema diminui até que um limiar é atingido, além do qual precisão e significância (ou relevância) tornam-se quase que características mutuamente exclusivas.”

Lofti Zadeh (O pai da lógica nebulosa e da computação granular)

### 4 Complexificação: evolução da complexidade

- É possível estudar sistemas complexos em seus aspectos evolutivos, procurando responder a duas questões fundamentais (MALKIN, 2009):
  1. Quais os mecanismos que fundamentam a evolução de sistemas complexos?
  2. É possível influenciar a evolução futura de sistemas complexos?
- No entanto, a vasta maioria dos estudos que investigam o potencial evolutivo de sistemas complexos ignoram o fato de que todos os sistemas complexos naturais estão sujeitos a processos de complexificação.

- A complexificação se refere a alguma forma de mudança estrutural que faz um sistema complexo se tornar ainda mais complexo do que ele já é (STANLEY & MIIKKULAINEN, 2004).
- A complexificação se dá pelos seguintes fenômenos, em algum momento da evolução histórica do sistema:
  1. Adição de novos componentes;
  2. Intensificação das interações já existentes;
  3. Introdução de novas interações de componentes existentes.
- MALKIN (2009) mostrou que uma complexificação gradual amplia de forma significativa o potencial evolutivo de três sistemas complexos estudados por ele.
- Isso sugere que sistemas complexos naturais são significativamente mais “evolutíveis” do que se previa, o que pode implicar que os passos evolutivos para se atingir um certo grau de complexidade podem ser mais resumidos do que se pensava (KAUFFMAN, 1993; 1995; COHEN & STEWART, 1995).

## 5 Exemplos de auto-organização na natureza

Conchas (por que esses padrões e não outros padrões possíveis?)



## Processos Físicos

- ✓ Dunas
- ✓ Reagentes químicos
- ✓ Rachaduras na lama
- ✓ Células de convecção de Bénard
- ✓ Rachadura em tinta
- ✓ Rugas em verniz



## Em animais

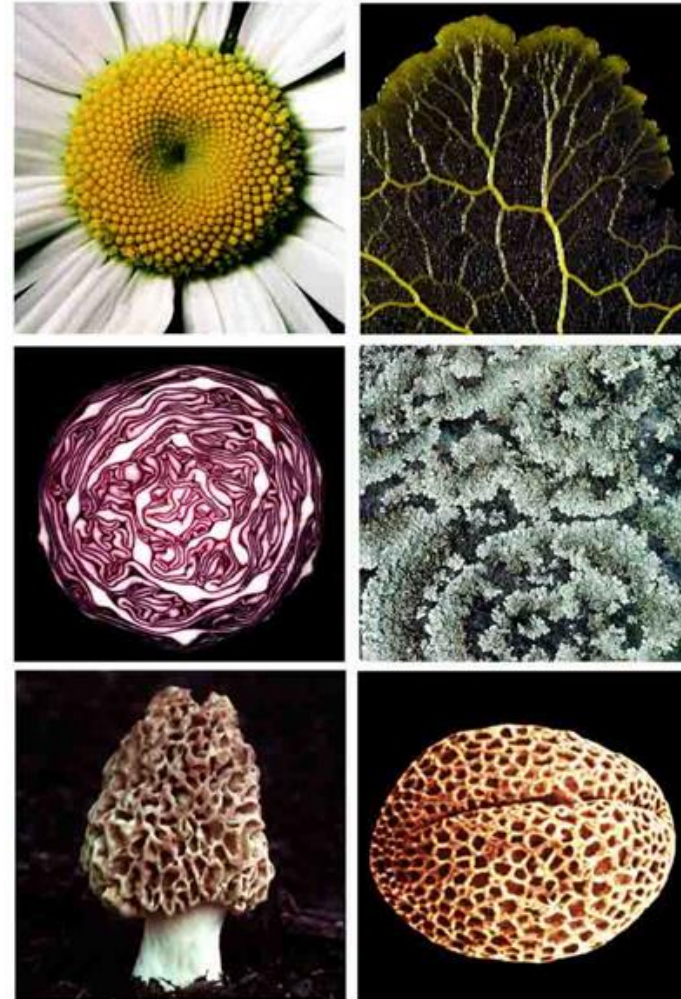
- ✓ Listras da zebra
- ✓ Manchas da girafa
- ✓ Listras do tigre
- ✓ Anfíbios
- ✓ Lagartos





## Em plantas

- ✓ Margarida
- ✓ *Slime mold*
- ✓ Repolho vermelho
- ✓ Líquen
- ✓ Cogumelo
- ✓ Grão de pólen



## Em insetos

- ✓ Asas de borboletas
- ✓ Exoesqueletos de besouros



## 6 Motivação para treinamento não-supervisionado: clusterização

- Dados rotulados são aqueles que assumem valores em um mesmo espaço vetorial multidimensional e que vêm acompanhados da classe a que cada um pertence (rótulo), podendo haver múltiplas classes, com variâncias e número de dados distintos ou não para cada classe.
- Dados não-rotulados são aqueles que assumem valores em um mesmo espaço vetorial multidimensional, e que não se conhece a priori a classe a que cada um pertence, embora cada um pertença a uma classe específica. O número de classes pode ser conhecido a priori ou não. A variância e o número de dados de cada classe podem diferir ou não.

### 6.1 Modelo simples de classificação para dados rotulados

- Hipótese: as classes apresentam propriedades distintas (seus elementos pertencem a regiões distintas do espaço vetorial multidimensional);
- Modelagem: um representante para cada classe;



- Objetivo: minimizar o somatório das distâncias entre os dados e o respectivo representante da classe a que pertencem;
- Aplicação: após finalizar o posicionamento de todos os representantes, definir o rótulo de cada novo dado não-rotulado como aquele associado ao representante que possuir a menor distância ao dado.
- Trata-se, portanto, de um problema de otimização, que pode ser resolvido por intermédio de técnicas de treinamento supervisionado.
- No caso, basta dividir o presente problema em  $C$  problemas distintos, sendo  $C$  o número de classes.
- Exemplos gráficos e análise topológica.
- Limitações.

## 6.2 Modelo composto de classificação para dados rotulados

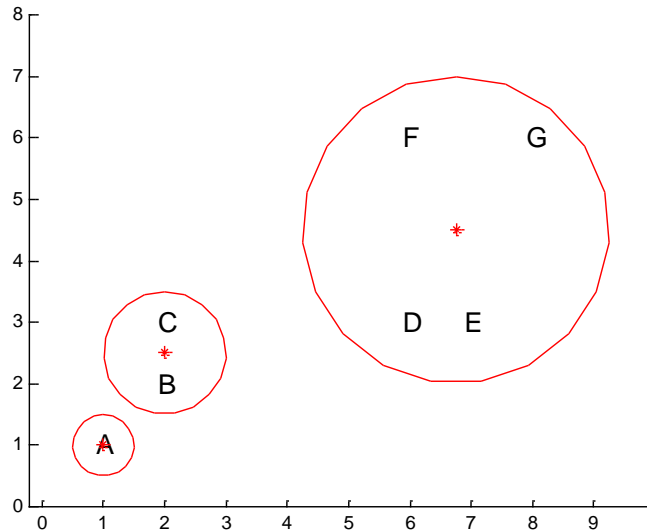
- Hipótese: as classes apresentam propriedades distintas (seus elementos pertencem a regiões distintas do espaço vetorial multidimensional);

- Modelagem: múltiplos representantes para cada classe;
- Objetivo: minimizar o somatório da distância entre cada dado e o representante mais próximo da classe a que pertence;
- Aplicação: após finalizar o posicionamento de todos os representantes, definir o rótulo de cada novo dado não-rotulado como aquele associado ao representante que possuir a menor distância ao dado.
- Trata-se também de um problema de otimização, mas neste caso técnicas de treinamento não-supervisionado devem ser empregadas, pois os representantes de cada classe devem se auto-organizar no espaço de acordo com a distribuição apresentada pelos dados da respectiva classe.
- Como no caso anterior, pode-se dividir o presente problema em  $C$  problemas distintos.
- Exemplos gráficos e análise topológica.

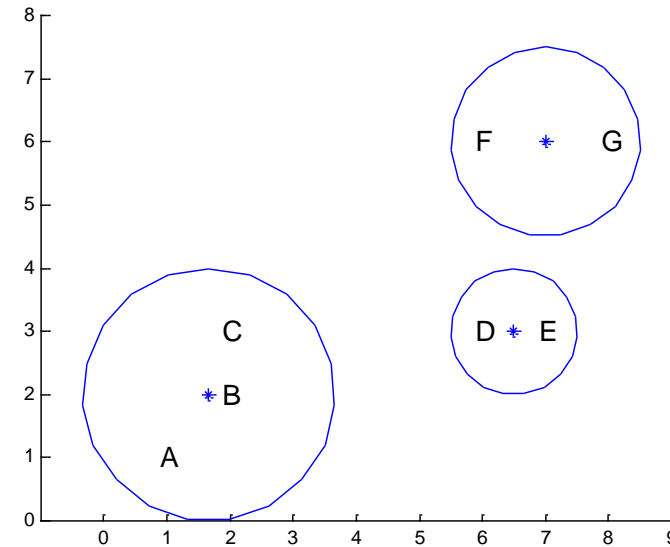
### 6.3 Modelo composto de classificação para dados não-rotulados

- Hipótese: não se conhece o número de classes, mas sabe-se que elas apresentam propriedades distintas (seus elementos pertencem a regiões distintas do espaço vetorial multidimensional);
  - Modelagem: múltiplos representantes não-rotulados;
  - Objetivo: minimizar o somatório da distância entre cada dado e o representante não-rotulado mais próximo.
  - Rotulagem: após finalizar o posicionamento de todos os representantes (também por auto-organização), aplicar alguma técnica de discriminação que agrupe representantes de acordo com as posições relativas entre eles. A seguir, atribuir um rótulo diferente a cada grupo de representantes.
  - Aplicação: definir o rótulo de cada novo dado não-rotulado como aquele associado ao representante que possuir a menor distância ao dado.
- Exemplos gráficos e análise topológica.

## 6.4 Comparação com o k-means



Erro total = 12,250

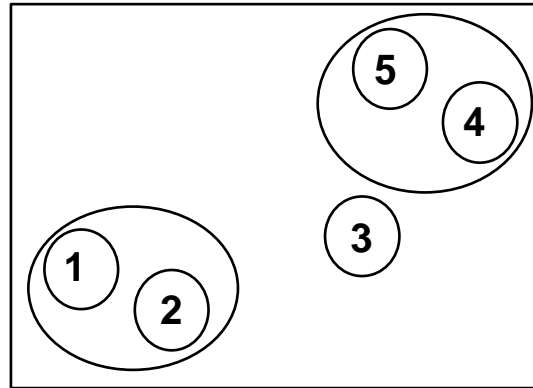


Erro total = 5,167

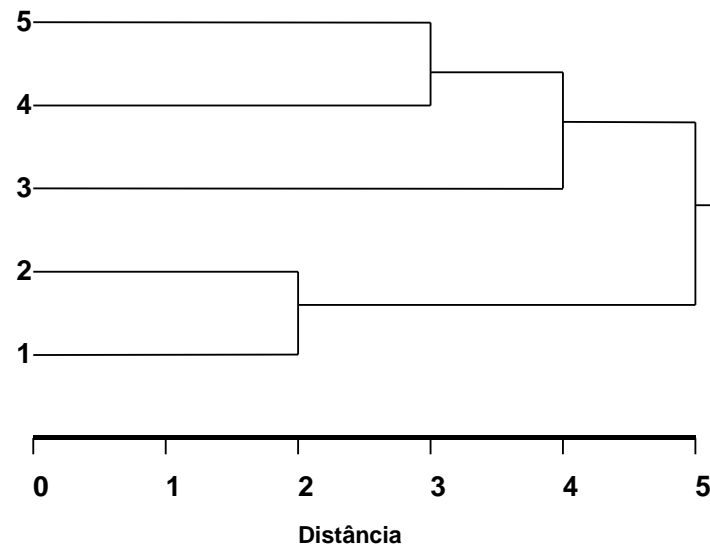
- Figuras extraídas de ZUCHINI (2003) e baseadas em JAIN *et al.* (1999)
- Sensibilidade à condição inicial
- Ausência de vizinhança topológica

## 7 Agrupamento em aprendizado de máquina

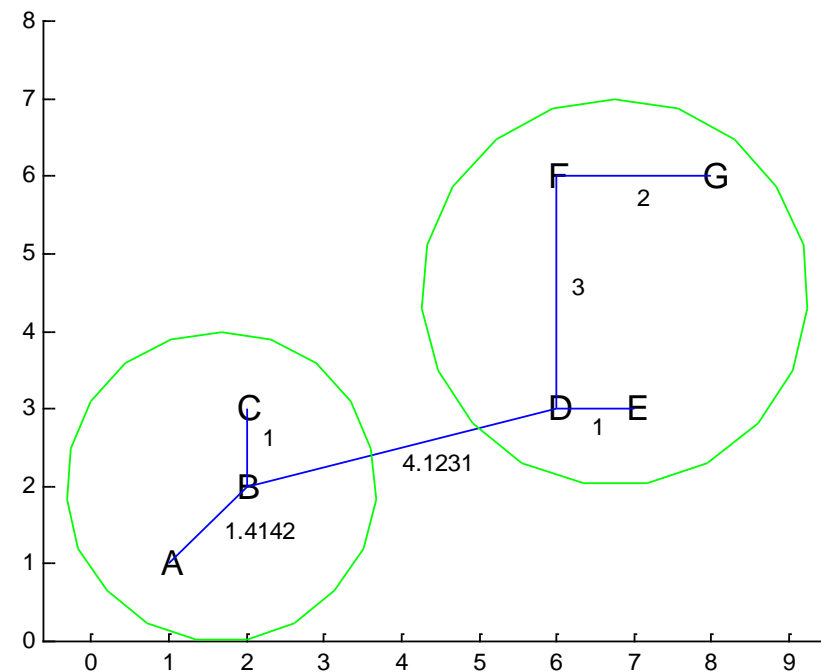
- Aprendizado de máquina envolve técnicas computacionais que visam ampliar a automatização e eficiência de processos de aquisição de conhecimento por um computador, a partir da exposição a eventos, manipulação de bases de dados ou qualquer outra forma de transmissão de informação ou processamento de dados.
- Aquisição de conhecimento: Geração de modelos do mundo, descoberta de regularidades e outras propriedades estatísticas presentes nos dados (associações espaciais e temporais), geração de leis de discriminação para efeito de classificação, realização de agrupamento de dados (previamente não-rotulados) em classes.
- Métodos de agrupamento (JAIN & DUBES, 1988; EVERITT, 1993; JAIN *et al.*, 1999):
  - ✓ Hierárquicos: aglomerativos, divisivos;
  - ✓ Particionais: k-means, *Minimum Spanning Tree* (MST);



Objetos



Agrupamento hierárquico (figuras extraídas de ZUCHINI, 2003)



*Minimum Spanning Tree (MST) (figura extraída de ZUCHINI, 2003)*

- Arcos com ‘comportamento inconsistente’ são eliminados, formando os grupos. Na figura acima, o arco BD pode ser considerado inconsistente para algum limiar.
- Detecção de comunidades (que podem ser interpretadas como grupos) em grafos tem evoluído muito, no estudo de redes sociais, e pode ser adotada aqui.

- Para um mesmo conjunto de dados, dadas múltiplas propostas de agrupamento, existem formas automáticas de ordená-las de acordo com a qualidade do agrupamento?
- Com a ressalva de que agrupar pode envolver subjetividade e de que não existe nenhum índice que seja um consenso na literatura, a resposta é afirmativa.
- Um índice muito popular é conhecido como Davies–Bouldin, tendo sido proposto em: Davies, D.L.; Bouldin, D.W. “A Cluster Separation Measure”, IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 1, no. 2, pp. 224-227, 1979.
- Também é muito utilizado o índice Silhouette, proposto em: Rousseeuw, P.J. “Silhouettes: a Graphical Aid to the Interpretation and Validation of Cluster Analysis”, Computational and Applied Mathematics, vol. 20, pp. 53-65, 1987.
- Silhouette tende a produzir resultados mais acurados, enquanto que Davies-Bouldin requer menos recursos computacionais.

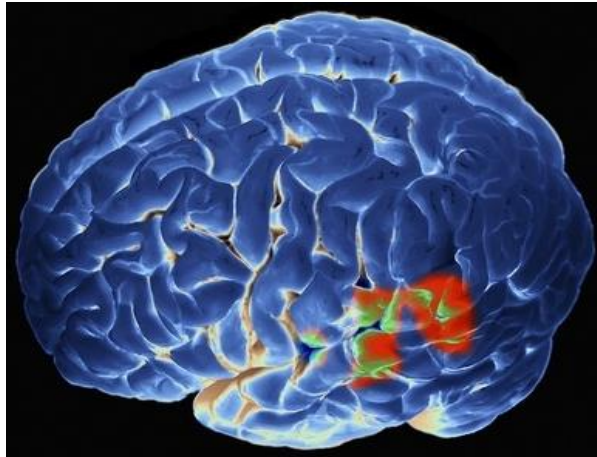


## 8 Treinamento não-supervisionado

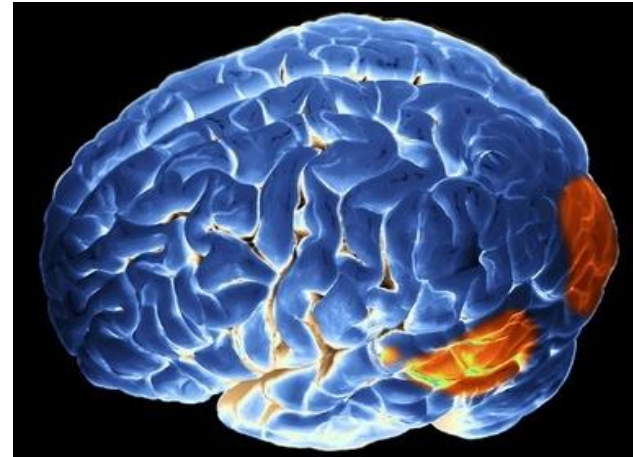
- Como aprender a representar padrões de entrada de modo a refletir a estrutura estatística de toda a coleção de dados de entrada? Que aspectos da entrada devem ser reproduzidos na saída?
- Em contraposição ao treinamento supervisionado e ao treinamento por reforço, não há aqui nenhuma saída desejada explícita ou avaliação externa da saída produzida para cada dado de entrada.
- O treinamento não-supervisionado é predominante no cérebro humano. É sabido que as propriedades estruturais e fisiológicas das sinapses no córtex cerebral são influenciadas pelos padrões de atividade que ocorrem nos neurônios sensoriais. No entanto, em essência, nenhuma informação prévia acerca do conteúdo ou significado do fenômeno sensorial está disponível.
- Sendo assim, a implementação de modelos computacionais para ajuste de pesos sinápticos via treinamento não-supervisionado deve recorrer apenas aos dados de

entrada, tomados como amostras independentes de uma distribuição de probabilidade desconhecida.

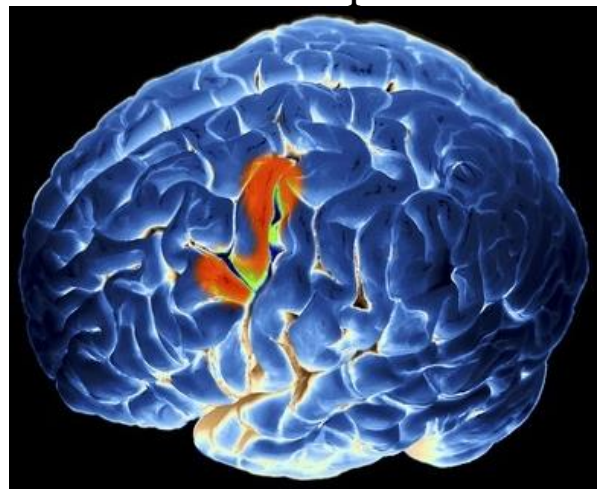
- Duas abordagens têm sido propostas para aprendizado não-supervisionado:
  1. Técnicas para estimação de densidades de probabilidade, que produzem modelos estatísticos explícitos para descrever os fenômenos responsáveis pela produção dos dados de entrada. Ex: redes bayesianas.
  2. Técnicas de extração de regularidades estatísticas diretamente dos dados de entrada. Ex: redes de Kohonen.
- A história do aprendizado não-supervisionado é longa e diversificada:
  - BARLOW (1989), HEBB (1949), HINTON & SEJNOWSKI (1986), MACKAY (1956), MARR (1970);
  - BECKER & PLUMBLEY (1996), HINTON (1989), KOHONEN (1989), KOHONEN (1997) e LINSKER (1988).
- Para mais detalhes, consultar DUDA *et al.* (2001) e HINTON & SEJNOWSKI (1999).



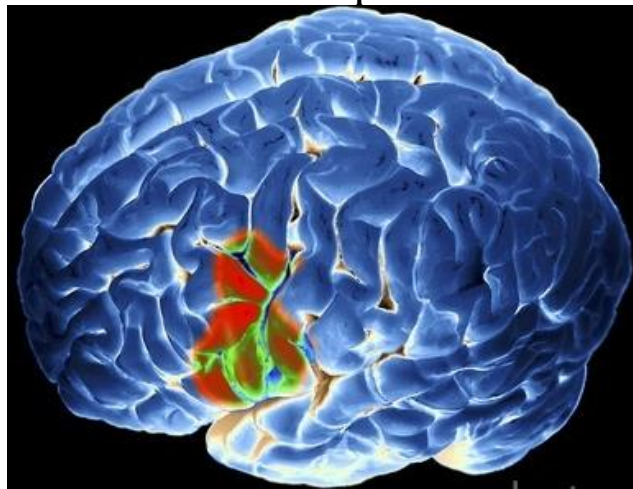
Ouvindo palavras



Lendo palavras



Falando palavras



Pensando em palavras

*Functional magnetic resonance imaging or functional MRI (fMRI)*

## 9 Mapas Auto-Organizáveis de Kohonen

- Um mapa de Kohonen é um arranjo de neurônios, geralmente restrito a espaços de dimensão 1 ou 2, que procura estabelecer e preservar noções de vizinhança (preservação topológica).
- Se estes mapas apresentarem propriedades de auto-organização, então eles podem ser aplicados a problemas de clusterização e ordenação espacial de dados.
- Neste caso, vai existir um mapeamento do espaço original (em que os dados se encontram) para o espaço em que está definido o arranjo de neurônios.
- Como geralmente o arranjo de neurônios ocorre em espaços de dimensão reduzida (1 ou 2), vai existir uma redução de dimensão sempre que o espaço original (em que os dados se encontram) apresentar uma dimensão mais elevada.
- Toda redução de dimensão (relativa à dimensão intrínseca dos dados) pode implicar na perda de informação (por exemplo, violação topológica). Sendo assim, este mapeamento deve ser tal que minimiza a perda de informação.

- A informação é uma medida da redução da incerteza, sobre um determinado estado de coisas. Neste sentido, a informação não deve ser confundida com o dado ou seu significado, e apresenta-se como função direta do grau de originalidade, imprevisibilidade ou valor-surpresa do dado (ou conjunto de dados).
- Espaços normados são aqueles que permitem o estabelecimento de propriedades topológicas entre seus elementos. Ex: medida de distância (fundamental para a definição do conceito de vizinhança), ordenamento.

### **9.1 Arranjo unidimensional**

- Um mapa de Kohonen unidimensional é dado por uma sequência ordenada de neurônios lineares, sendo que o número de pesos de cada neurônio é igual ao número de entradas.
- Há uma relação de vizinhança entre os neurônios (no espaço unidimensional vinculado ao arranjo), mas há também uma relação entre os pesos dos neurônios

no espaço de dimensão igual ao número de entradas. Para entender a funcionalidade dos mapas de Kohonen, é necessário considerar ambas as relações.

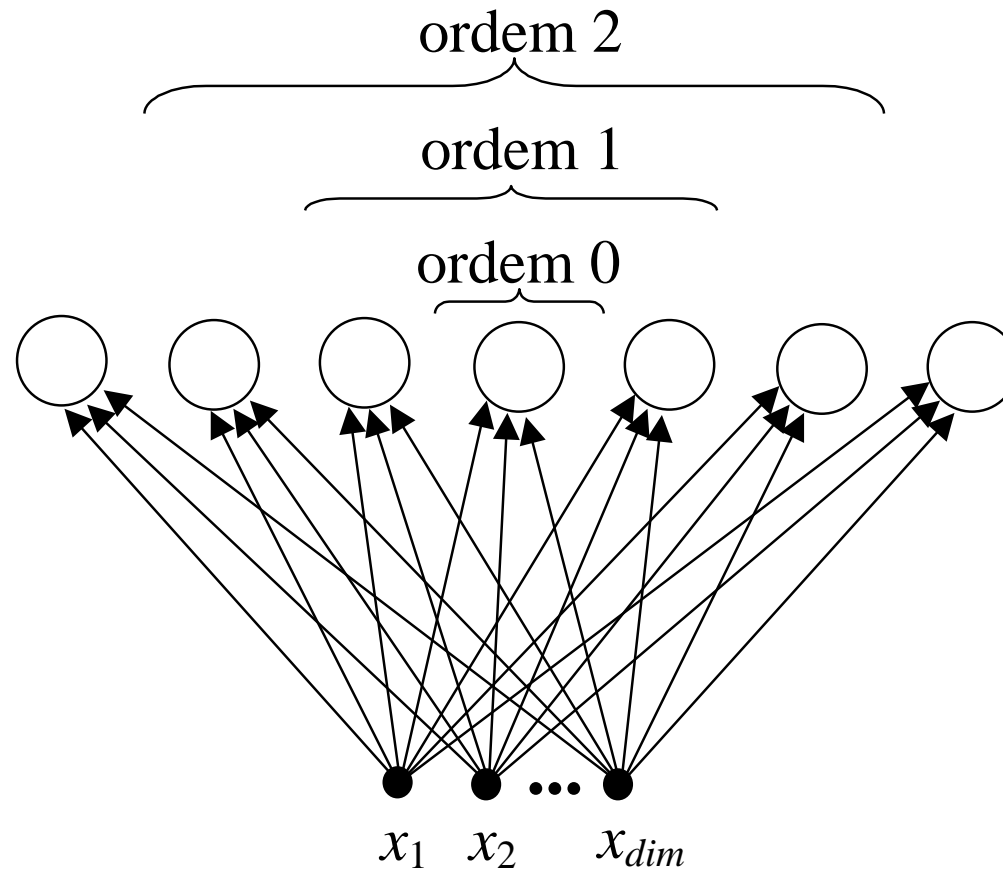


Figura 1 – Rede de Kohonen em arranjo unidimensional: ênfase na vizinhança

- $t_{il} = \begin{cases} +1 & \text{se } i = l \\ -\varepsilon & \text{se } i \neq l, \quad \varepsilon < \frac{1}{n} \end{cases}$
- $\bar{y}(0) = \mathbf{W}\bar{x}$
- $y_i(k+1) = \max \left\{ 0, \sum_{l=1}^n t_{il} y_l(k) \right\}$

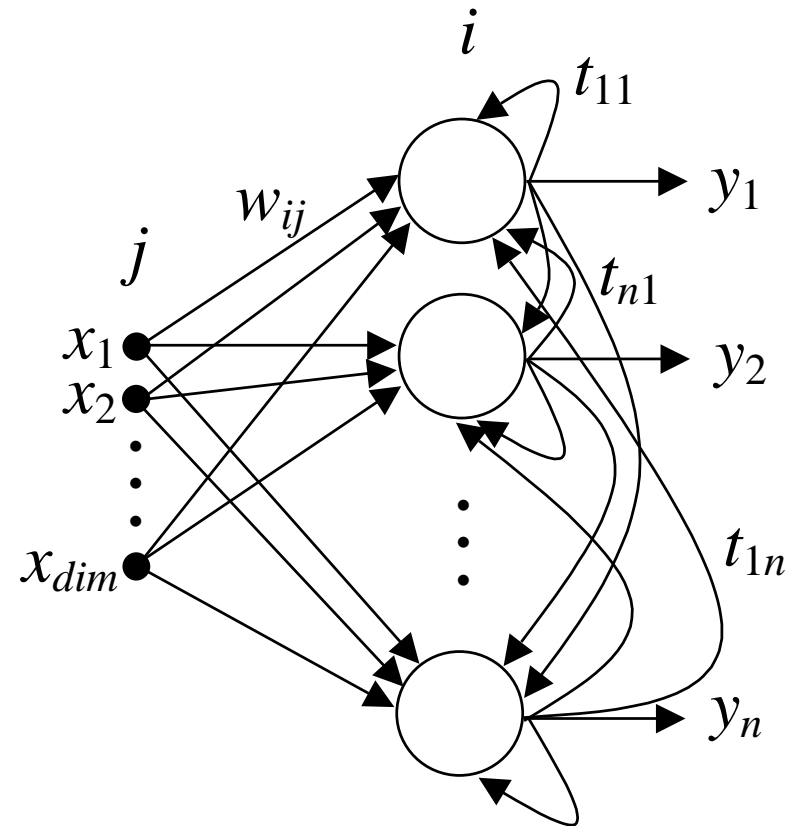


Figura 2 – Rede de Kohonen em arranjo unidimensional: ênfase nas conexões

Exemplo: Para  $dim = 2$ , considere  $n = 4$ , sendo que os vetores de pesos são dados na forma:

$$\mathbf{w}_1 = \begin{bmatrix} 2 \\ 3 \end{bmatrix}, \quad \mathbf{w}_2 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \quad \mathbf{w}_3 = \begin{bmatrix} -3 \\ 0 \end{bmatrix} \quad \text{e} \quad \mathbf{w}_4 = \begin{bmatrix} 3 \\ -2 \end{bmatrix}$$

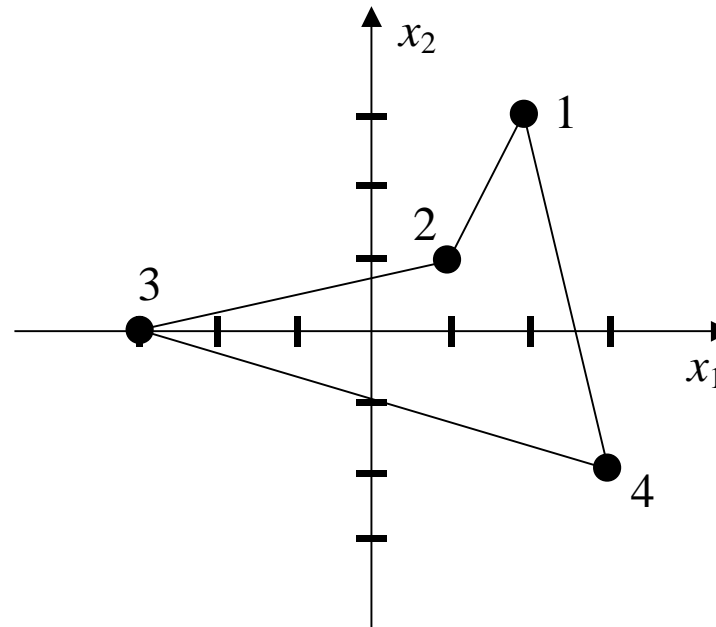


Figura 3 – Mapa de Kohonen em arranjo unidimensional (neste caso, existe vizinhança entre primeiro e último neurônios)



## 9.2 Arranjo bidimensional

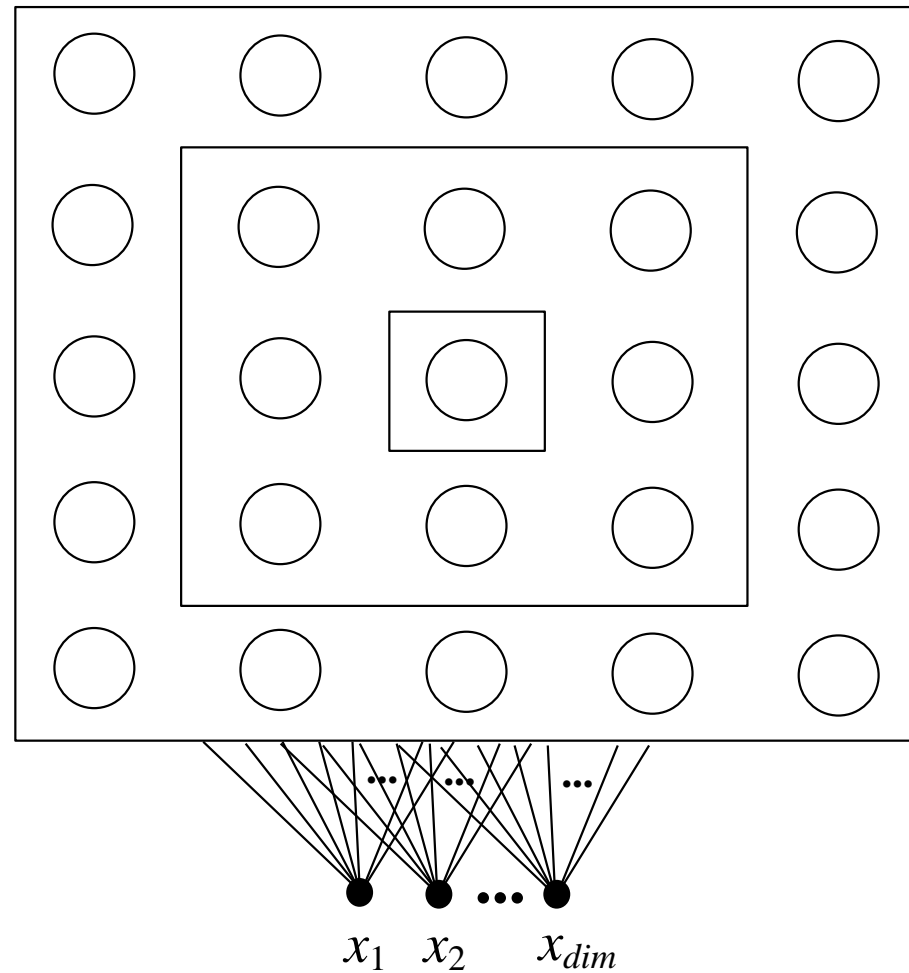


Figura 4 – Rede de Kohonen em arranjo bidimensional: ênfase na vizinhança

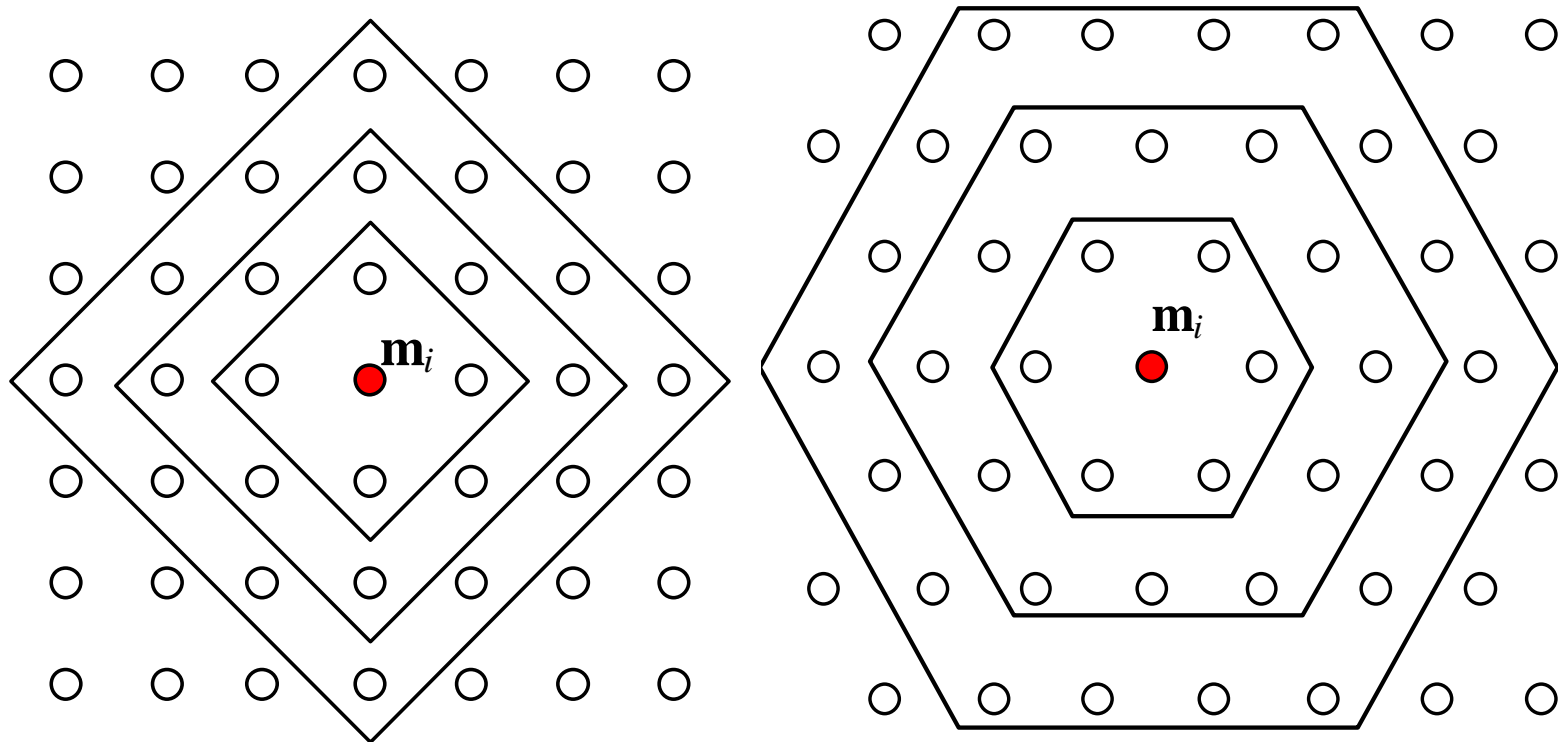


Figura 5 – Outras configurações de mapas e de vizinhança (figuras extraídas de ZUCHINI, 2003)

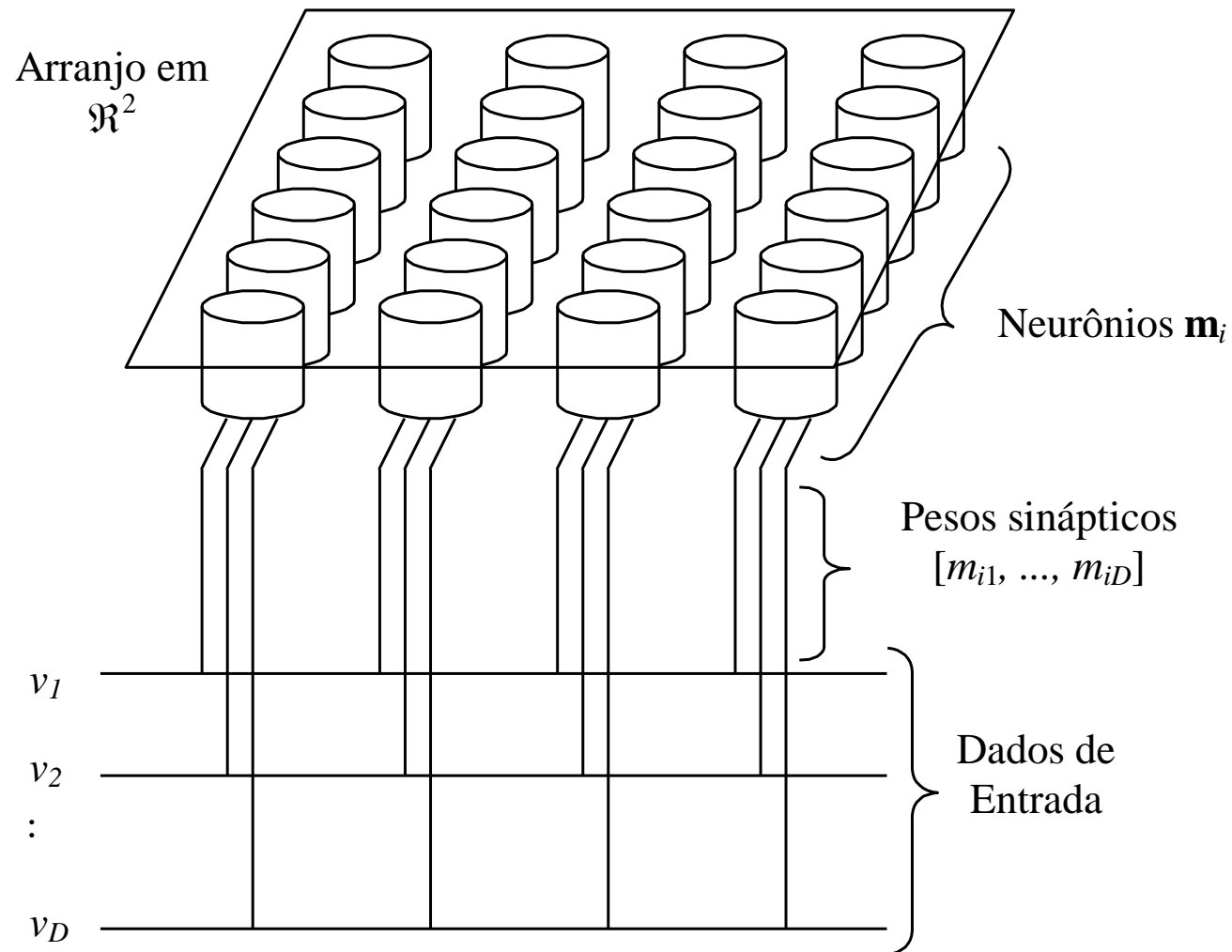


Figura 6 – Outra perspectiva para arranjo 2D (figura extraída de ZUCHINI, 2003)

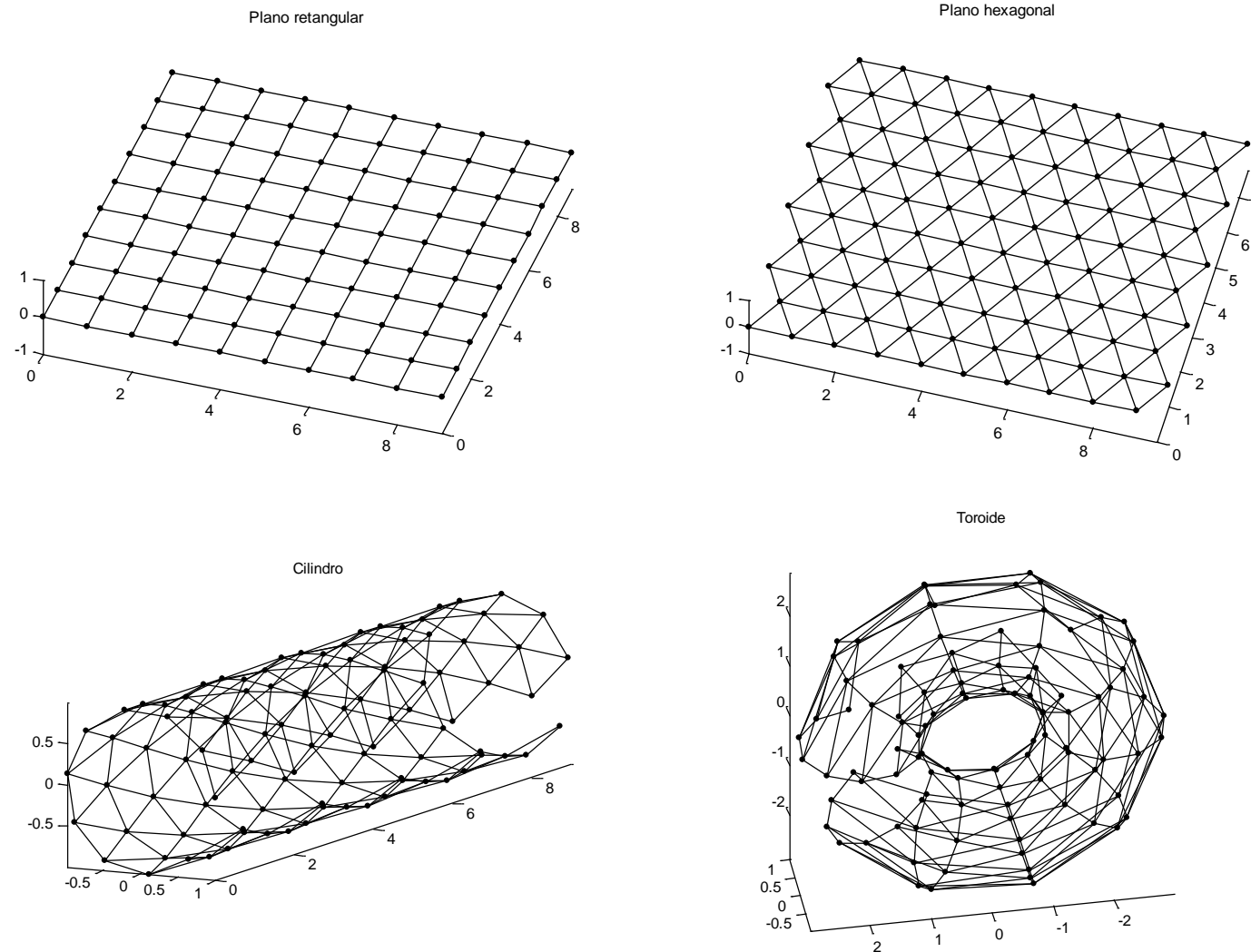
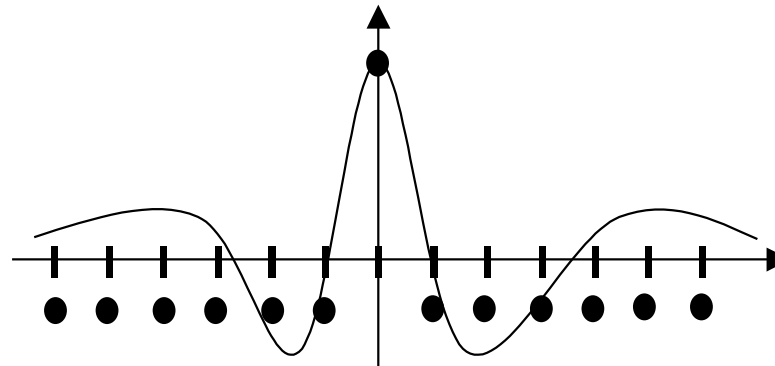


Figura 7 – Arranjos com e sem vizinhança nos extremos (figuras extraídas de ZUCHINI, 2003)

### 9.3 Fase de competição

- Três elementos básicos em uma lei de aprendizado competitivo:
  1. Um conjunto de neurônios similares, exceto pelo valor de seus pesos;
  2. Um limite imposto ao valor que os pesos podem atingir;
  3. Um mecanismo de competição.
- A competição vai produzir um único neurônio ativo para cada entrada (winner-takes-all).
- Como implementar? → Definindo uma regra de influência da ativação de um neurônio junto a todos os demais.



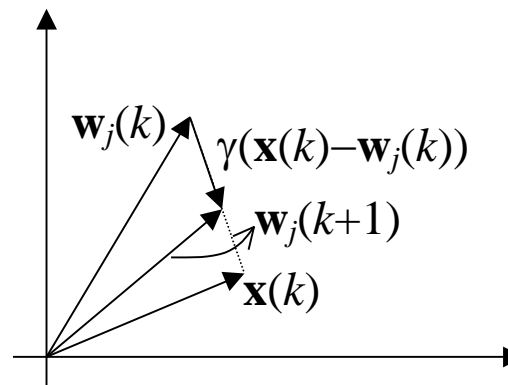
## 9.4 Fase de aprendizado não-supervisionado

- Se os pesos se mantiverem com norma unitária, então a definição do neurônio vencedor pode se dar de duas formas:
  1. Produto escalar;
  2. Norma euclidiana (é a forma adotada também no caso de pesos não-normalizados).
- Seja  $j$  o neurônio vencedor:

### Alternativa 1

- Somente o neurônio  $j$  é ajustado na forma:

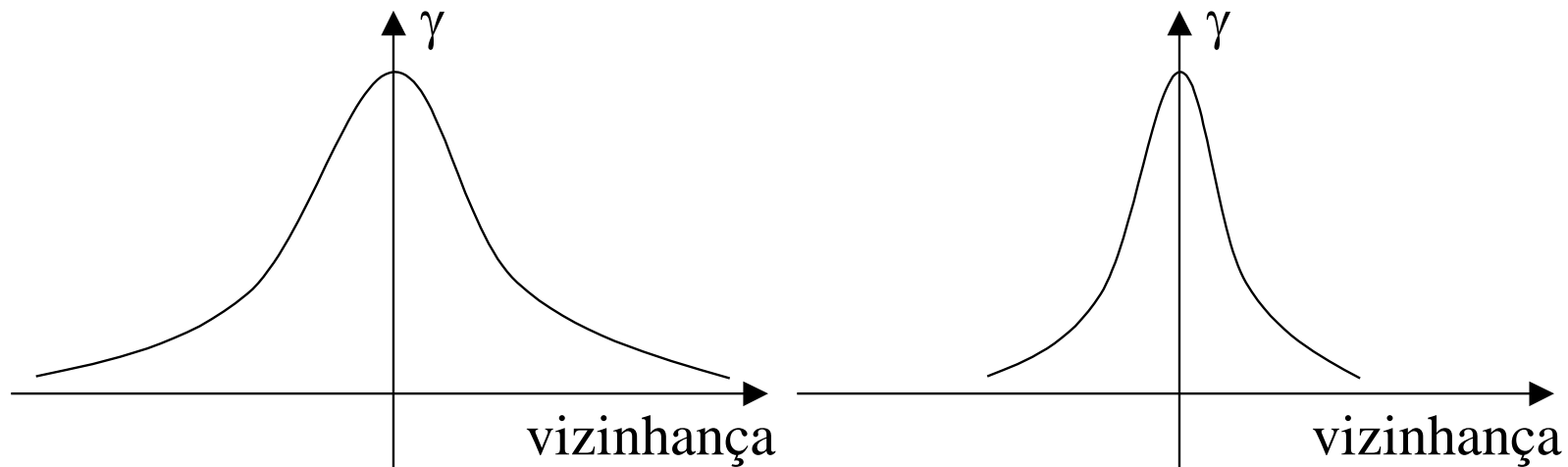
$$\mathbf{w}_j(k+1) = \mathbf{w}_j(k) + \gamma(\mathbf{x}(k) - \mathbf{w}_j(k))$$



## **Alternativa 2**

- Caso existam múltiplos representantes para cada agrupamento de dados, então é interessante ajustar o neurônio vencedor e seus vizinhos mais próximos.

## **Implementação**



- É importante que a influência de cada neurônio vencedor seja ampla no início do processo e sofra uma redução continuada com o decorrer das iterações.

## 9.5 Algoritmo de ajuste dos pesos

```
while <condição de parada> é falso,  
    Ordene aleatoriamente os  $N$  padrões de entrada;  
    for  $i=1$  até  $N$ ,  
         $j = \arg \min_j \|\mathbf{x}_i - \mathbf{w}_j\|$   
         $\forall J \in \text{Viz}(j)$  do:  
             $\mathbf{w}_J = \mathbf{w}_J + \gamma(\text{dist}(j, J))(\mathbf{x}_i - \mathbf{w}_J);$   
        end do  
    end for  
    Atualize a taxa de aprendizado  $\gamma$ ;  
    Atualize a vizinhança;  
    Avalie a condição de parada;  
end while
```



## 9.6 Ajuste de pesos considerando a vizinhança

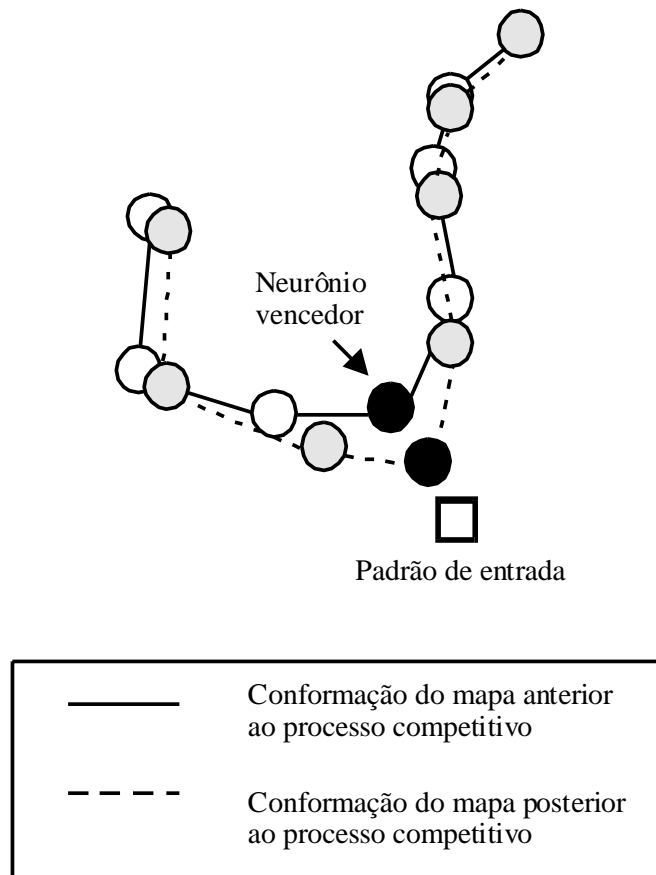


Figura 15 – Ajuste do neurônio vencedor e de seus vizinhos mais próximos: um passo de ajuste num arranjo unidimensional

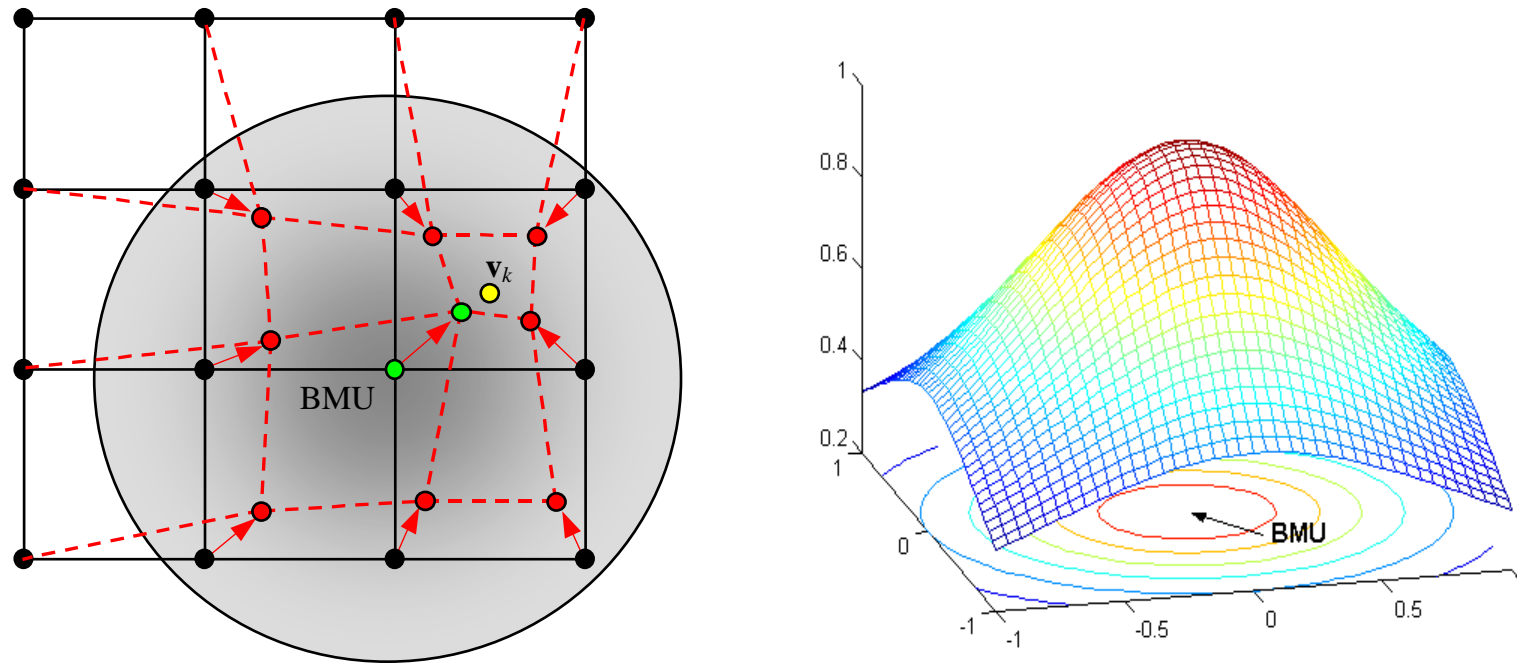
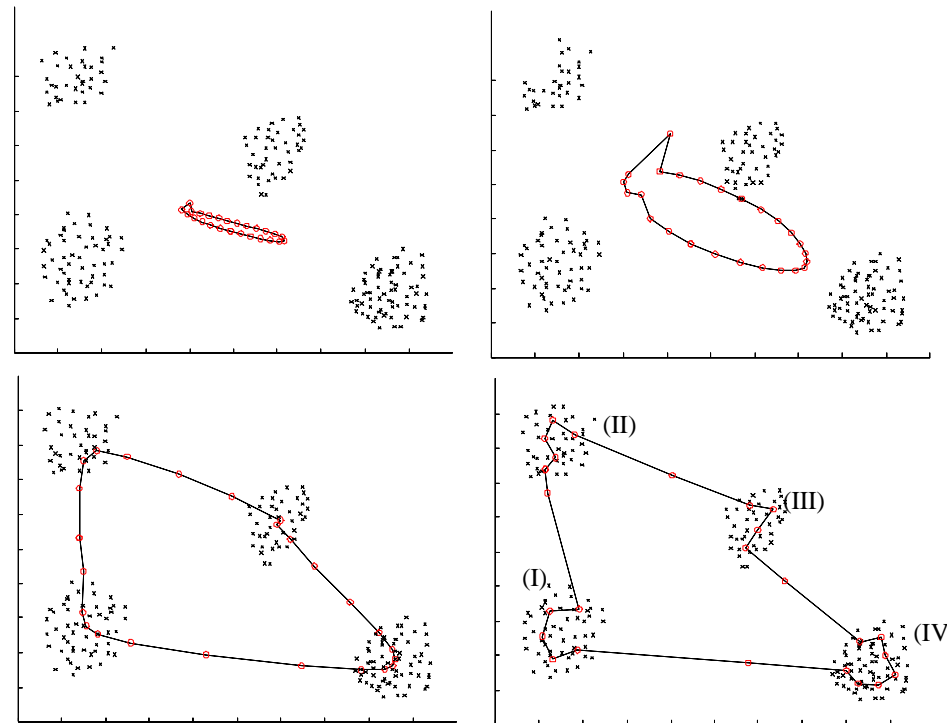
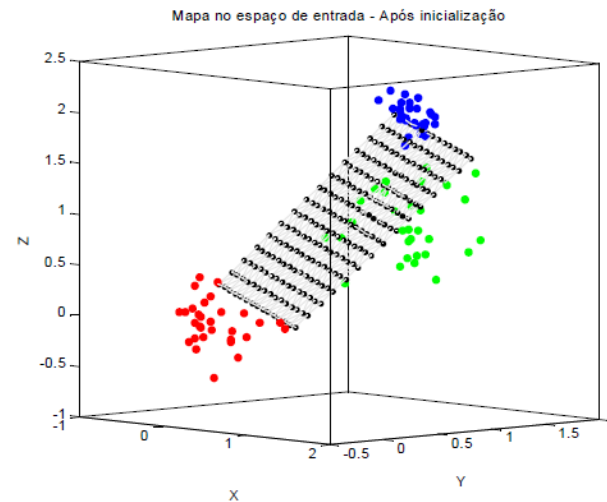
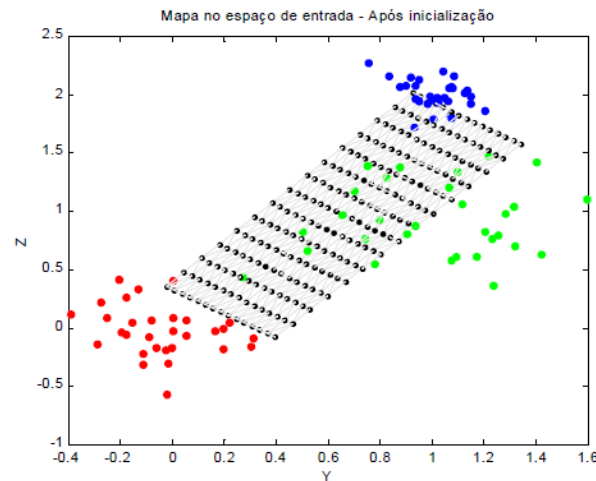
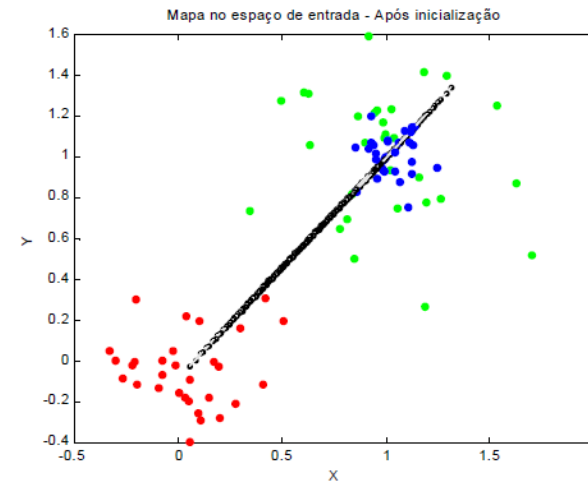
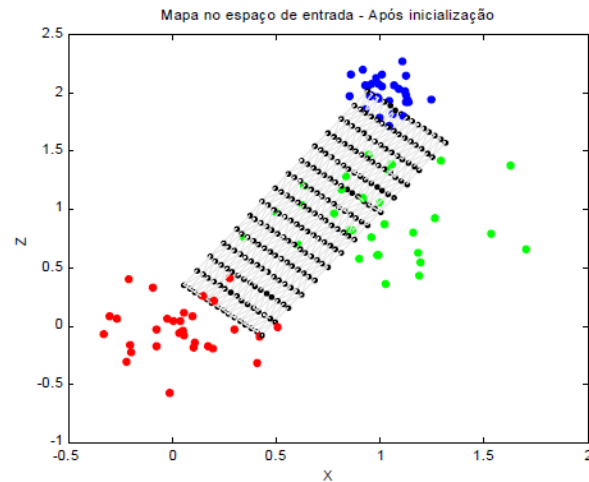


Figura 8 – BMU (*Best Matching Unit*) e seus vizinhos (figuras extraídas de ZUCHINI, 2003)

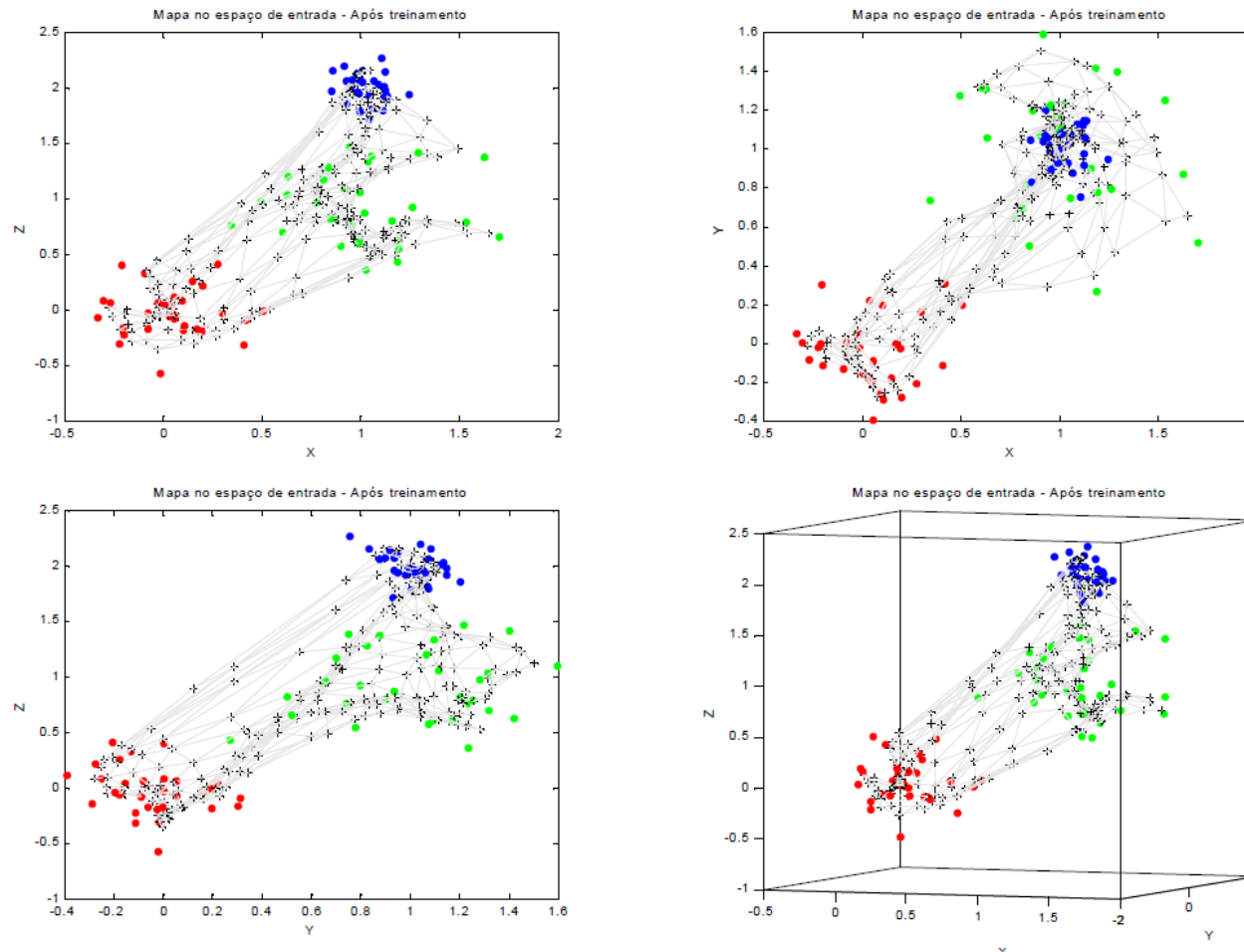
- O neurônio que venceu para uma dada amostra é o que sofre o maior ajuste. No entanto, dentro de uma vizinhança, todos os neurônios vizinhos também sofrerão um ajuste de pesos, embora de menor intensidade.

- A figura a seguir apresenta etapas da auto-organização para um mapa de Kohonen 1D (com vizinhança “toroidal”) e dados em 2D. Os neurônios estão posicionados de acordo com a localização de seus respectivos vetores de pesos, e estão ligados de acordo com a sua localização no mapa. O objetivo é ilustrar o efeito da “atração” exercida pelos dados sobre os vetores de pesos dos neurônios, levando a um “estado de equilíbrio de forças” ao final da auto-organização.





Recomenda-se que a inicialização dos pesos leve o mapa a “atravessar” os dados. Pode-se adotar, por exemplo, as direções indicadas por PCA (figuras extraídas de ZUCHINI, 2003).



Conformação do mapa de Kohonen após o processo de auto-organização (figuras extraídas de ZUCHINI, 2003).

## 9.7 Discriminação dos agrupamentos

- Dada a conformação final de neurônios (não rotulados), como realizar a clusterização, ou seja, definição de agrupamentos e atribuição do mesmo rótulo a todos os neurônios pertencentes a um dado agrupamento?
- Solução: matriz(vetor)-U (ULTSCH, 1993; COSTA, 1999)
- Aspecto a ser explorado: após o processo de auto-organização, dados de entrada com características semelhantes passam a promover reações semelhantes da rede neural.
- Assim, comparando-se as reações da rede neural treinada, é possível agrupar os dados pela análise do efeito produzido pela apresentação de cada um à rede.
- A matriz-U é uma ferramenta que permite realizar a discriminação dos agrupamentos, a partir de uma medida do grau de similaridade entre os pesos de neurônios adjacentes na rede. O perfil apresentado pelas distâncias relativas entre neurônios vizinhos representa uma forma de visualização de agrupamentos.

- Recebeu a denominação de matriz por ter sido proposta no caso de mapas bidimensionais, sendo que o grau de similaridade é plotado na terceira dimensão gerando uma superfície em relevo em 3D.
- Para o caso de mapas unidimensionais, tem-se o vetor-U.
- Topologicamente, as distâncias entre neurônios vizinhos refletem os agrupamentos, pois uma “depressão” ou um “vale” da superfície de relevo representa neurônios pertencentes a um mesmo agrupamento. Neurônios que têm uma distância grande em relação ao neurônio adjacente, a qual é representada por um pico da superfície de relevo, são neurônios discriminantes de agrupamentos.

## 9.8 Aplicação

- Como utilizar o mapa de Kohonen, após a fase de treinamento não-supervisionado e depois de ter as classes devidamente discriminadas, para classificação de padrões?
- Exemplos de aplicação

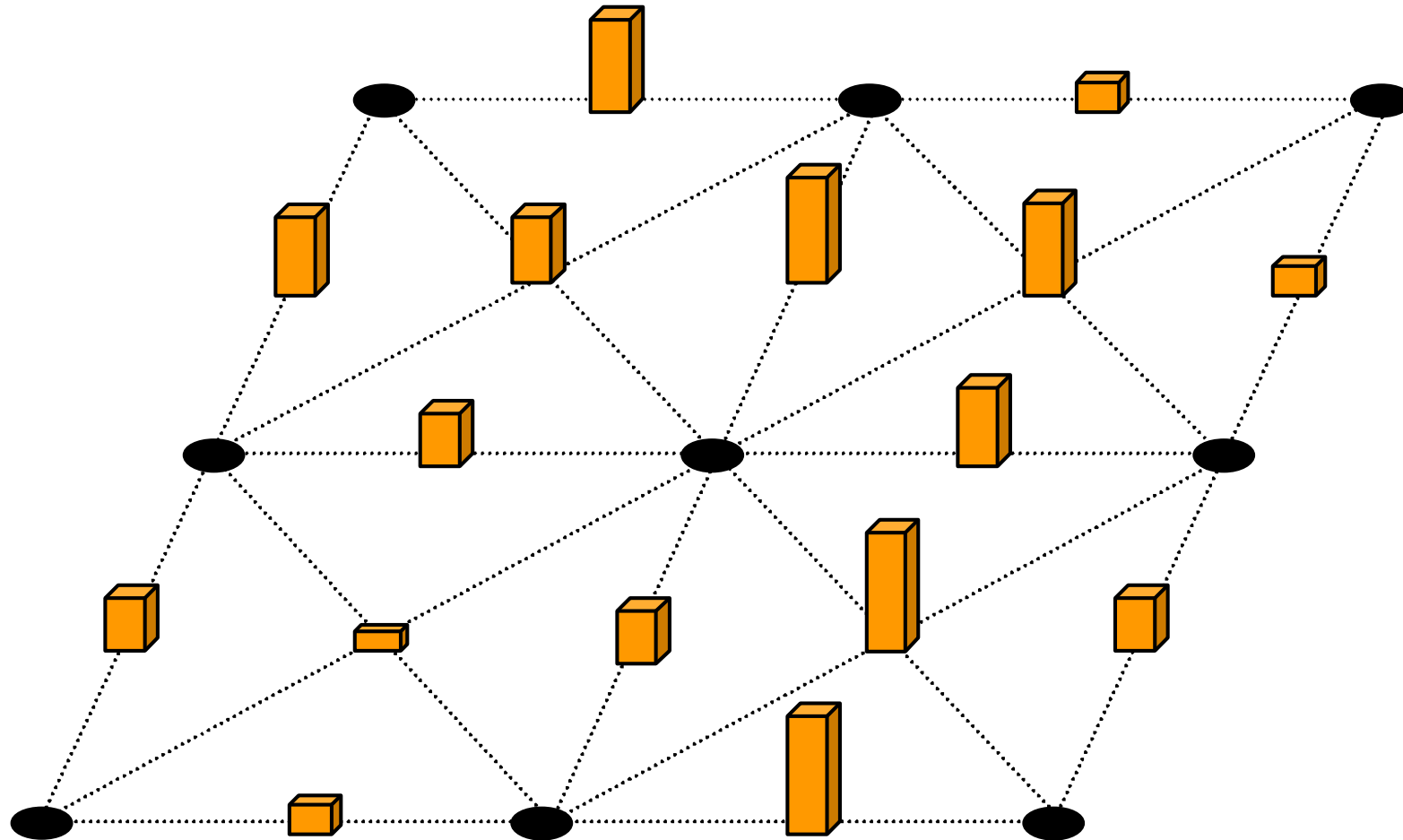


Figura 9 – Exemplo de matriz-U para arranjo retangular (figura extraída de ZUCHINI, 2003)



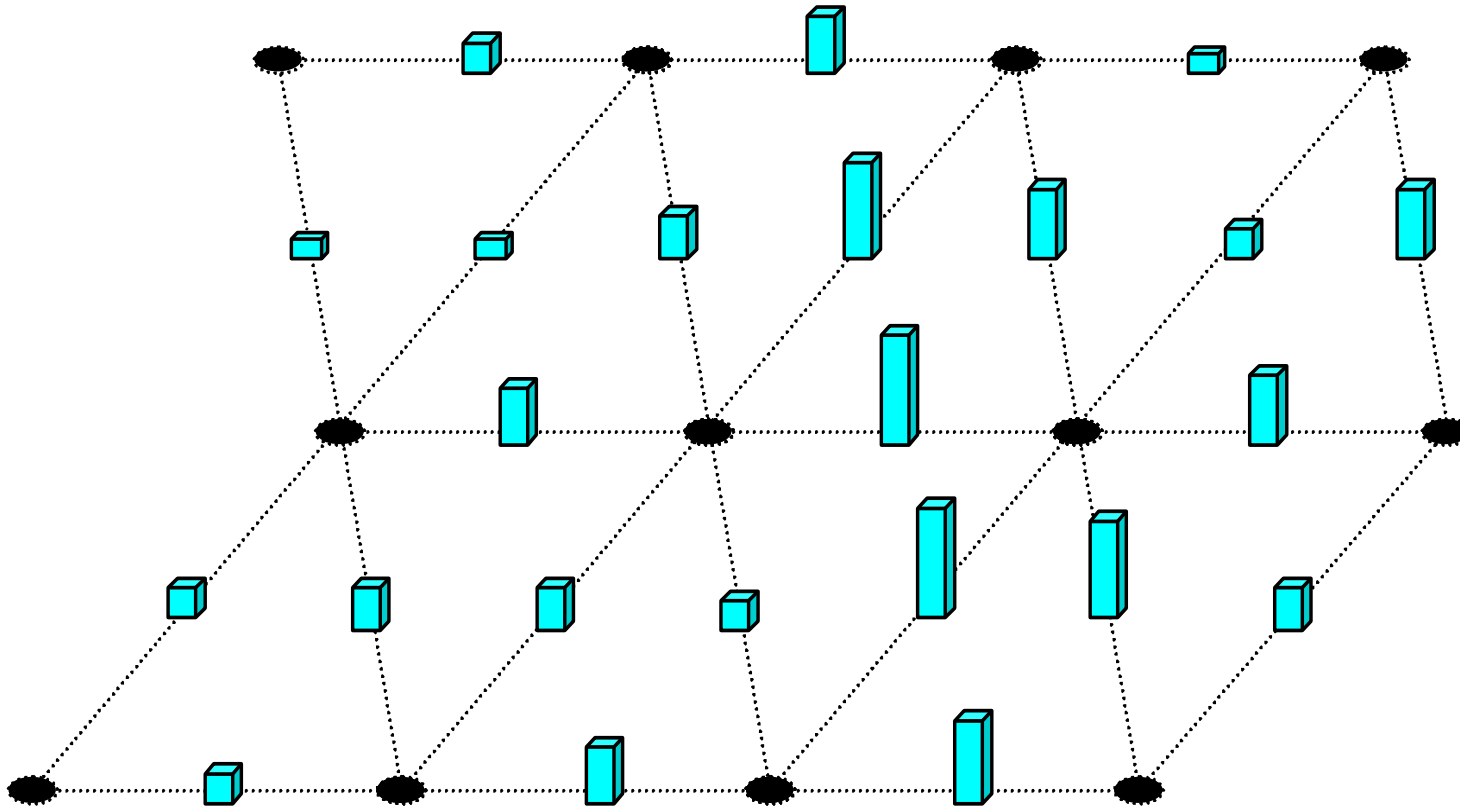


Figura 10 – Exemplo de matriz-U para arranjo hexagonal (figura extraída de ZUCHINI, 2003)

## 9.9 Agrupamento de dados (Clusterização)

### Caso 1: Ausência de agrupamentos

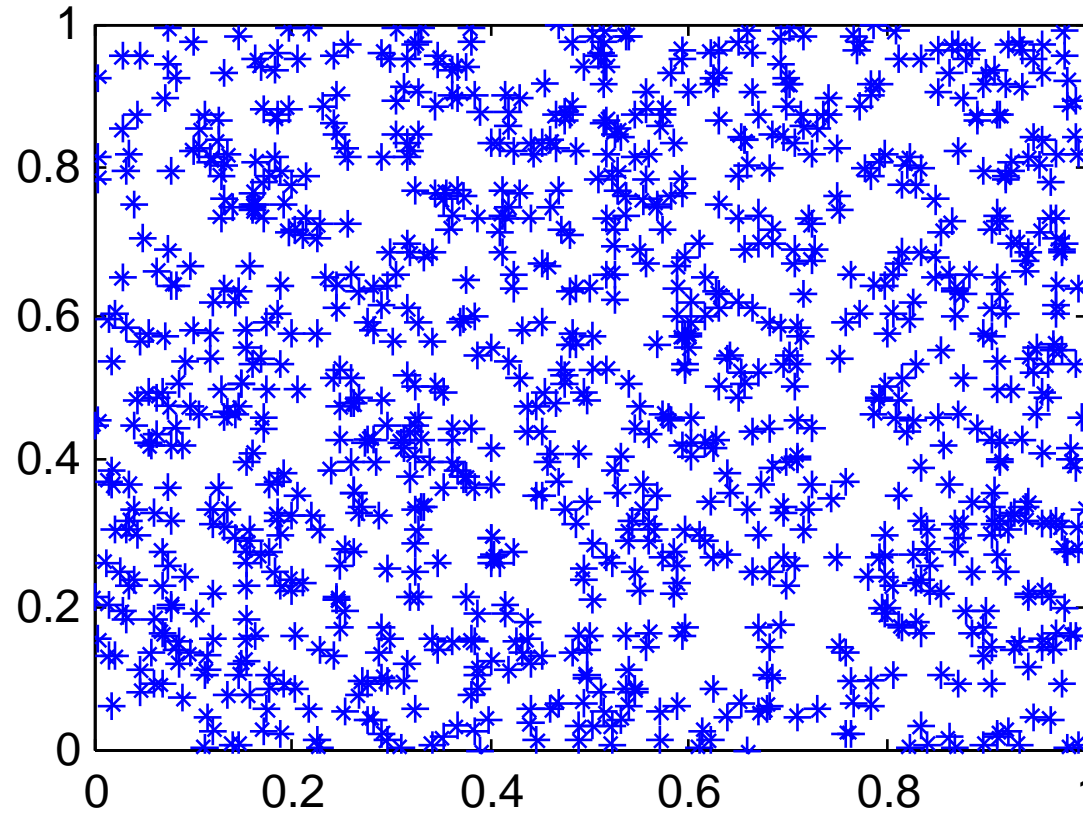


Figura 11 – Padrões de entrada com distribuição uniforme



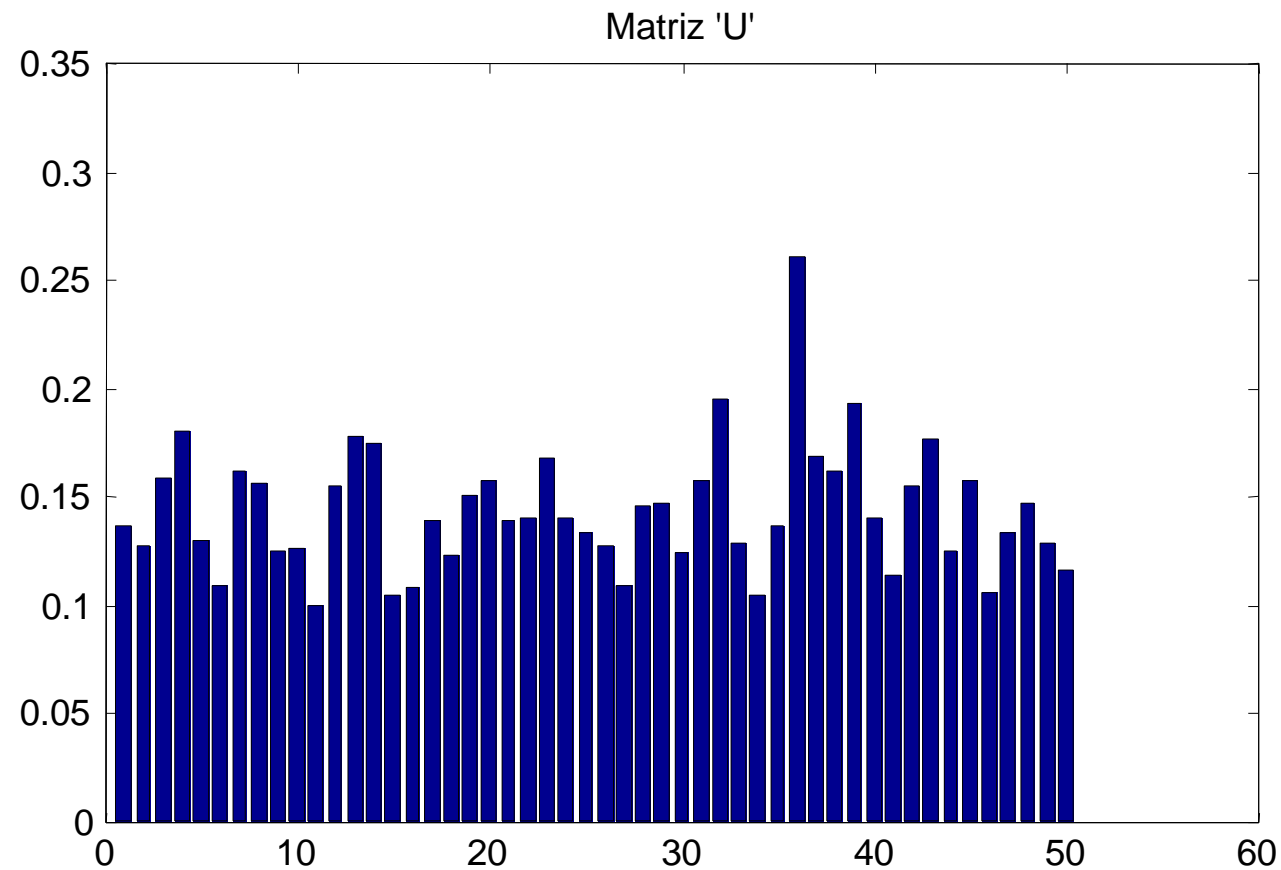


Figura 13 – Matriz-U (na verdade, vetor-U) para o mapa da Figura 12. Como esperado, não há nenhum agrupamento evidente.

Caso 2: Presença de agrupamentos bem distintos

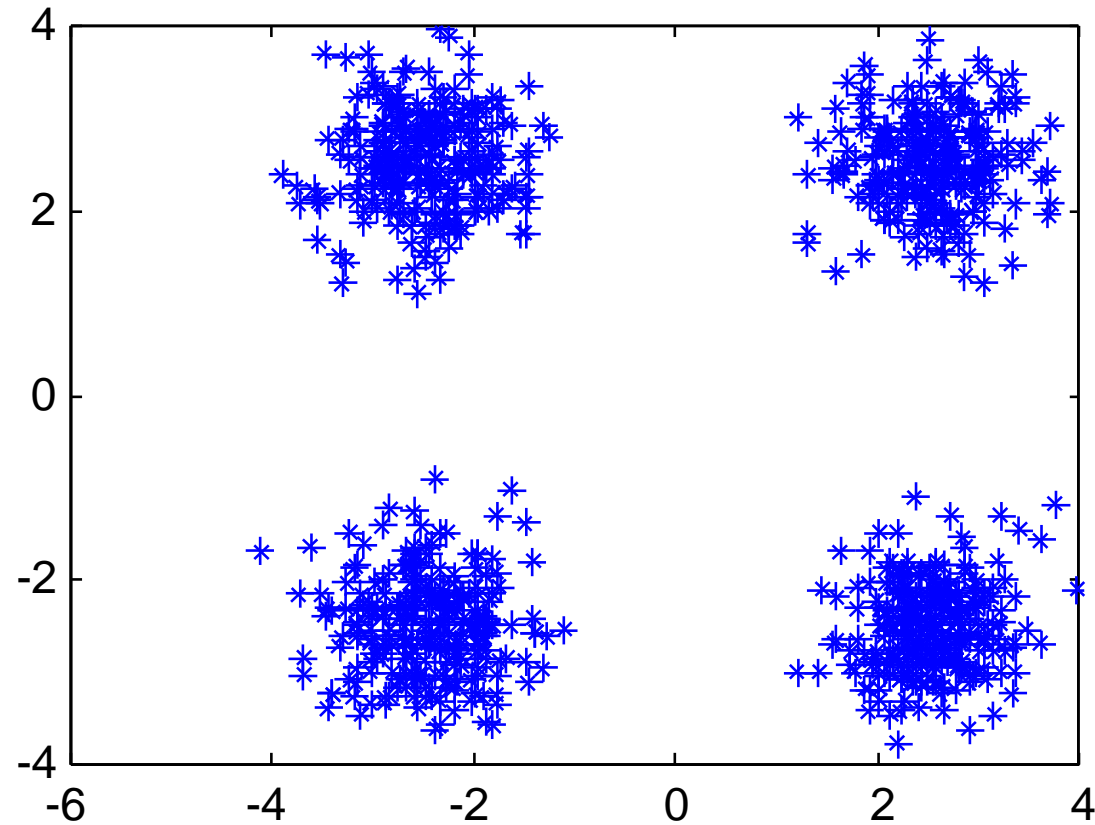


Figura 14 – Padrões de entrada com 4 agrupamentos

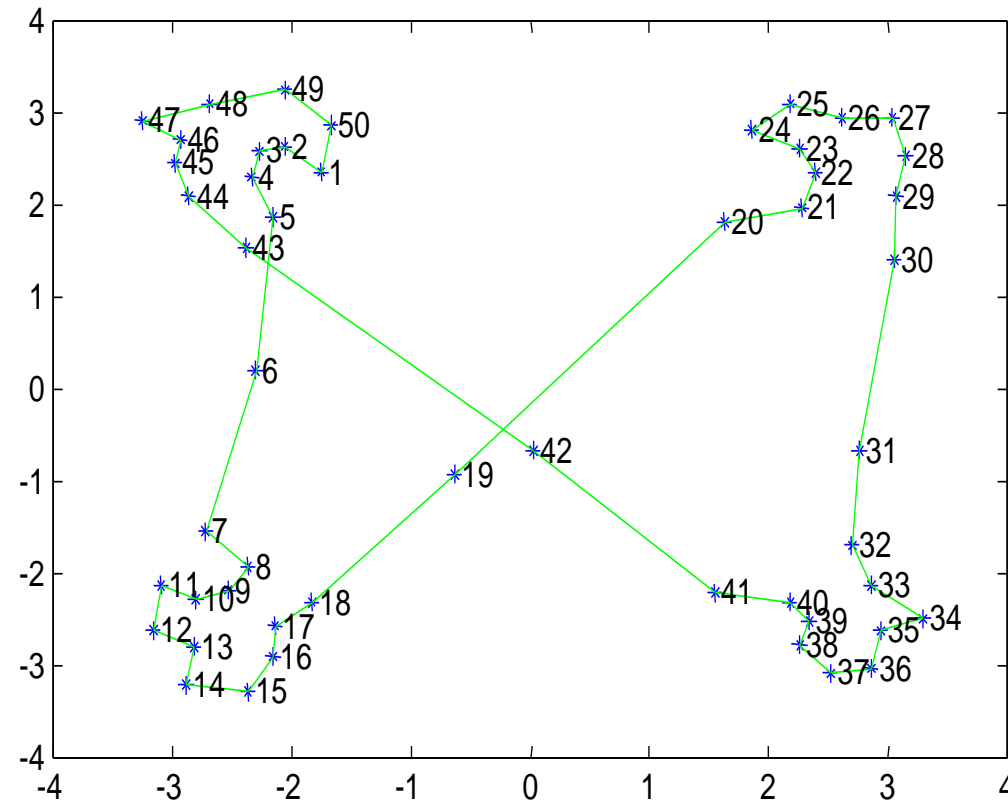


Figura 15 – Configuração final de 50 neurônios em um arranjo unidimensional

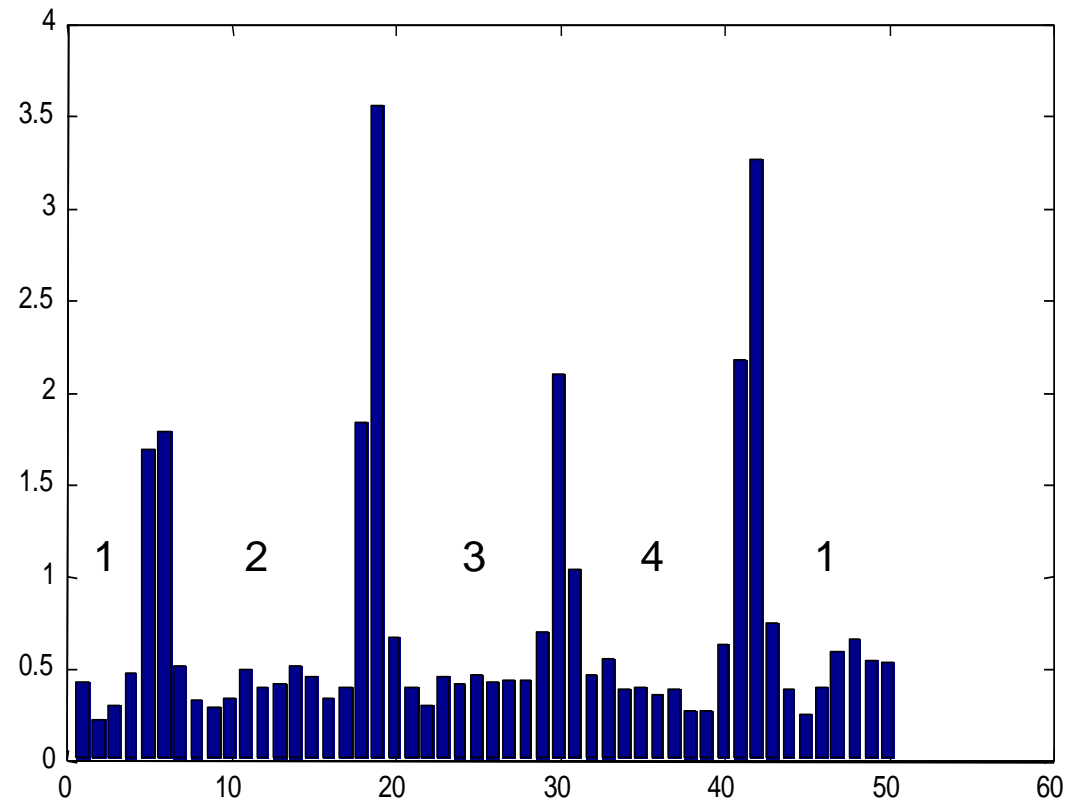


Figura 16 – Matriz-U (na verdade, vetor-U) para o mapa da Figura 15.

### Caso 3: Presença de agrupamentos pouco distintos

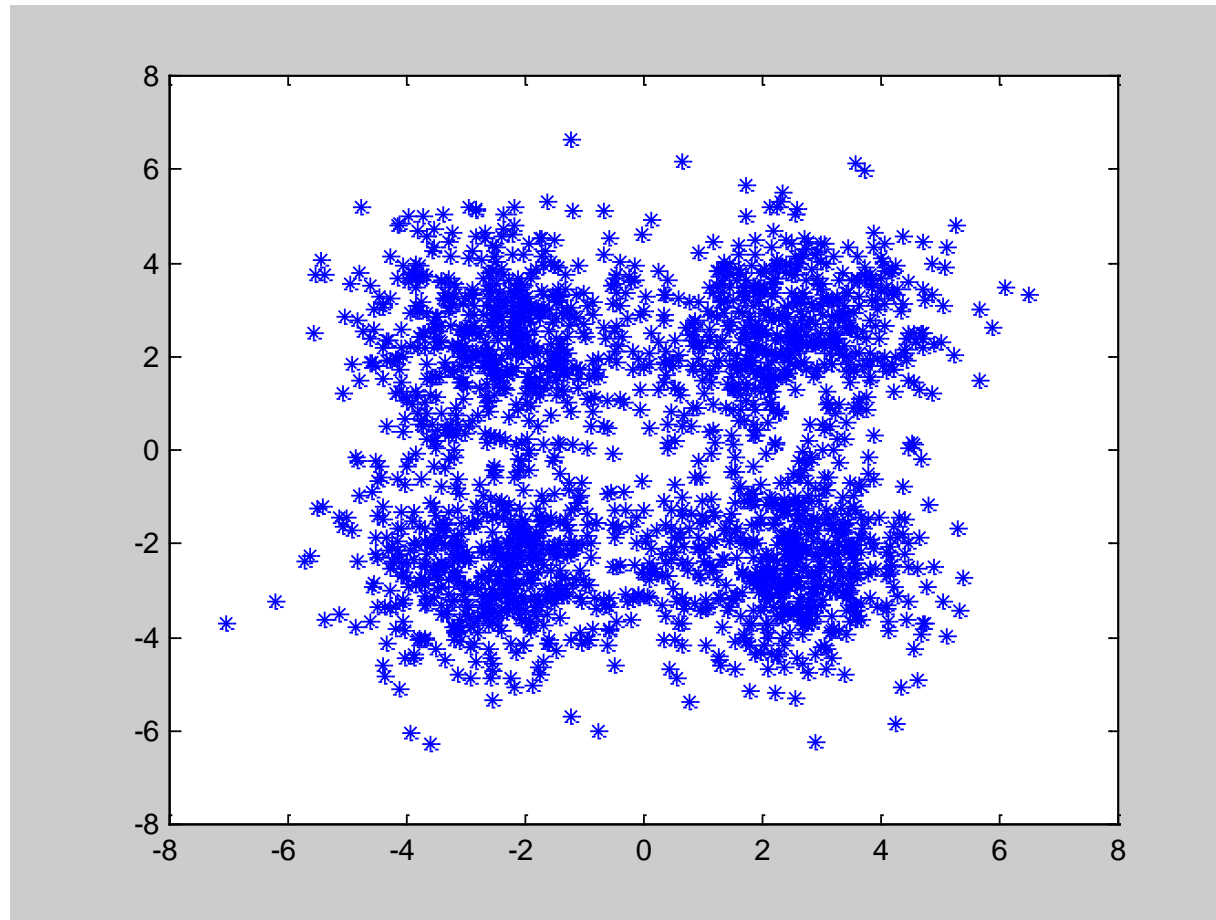


Figura 17 – Quatro agrupamentos de dados não-rotulados distribuídos no  $\mathbb{R}^2$



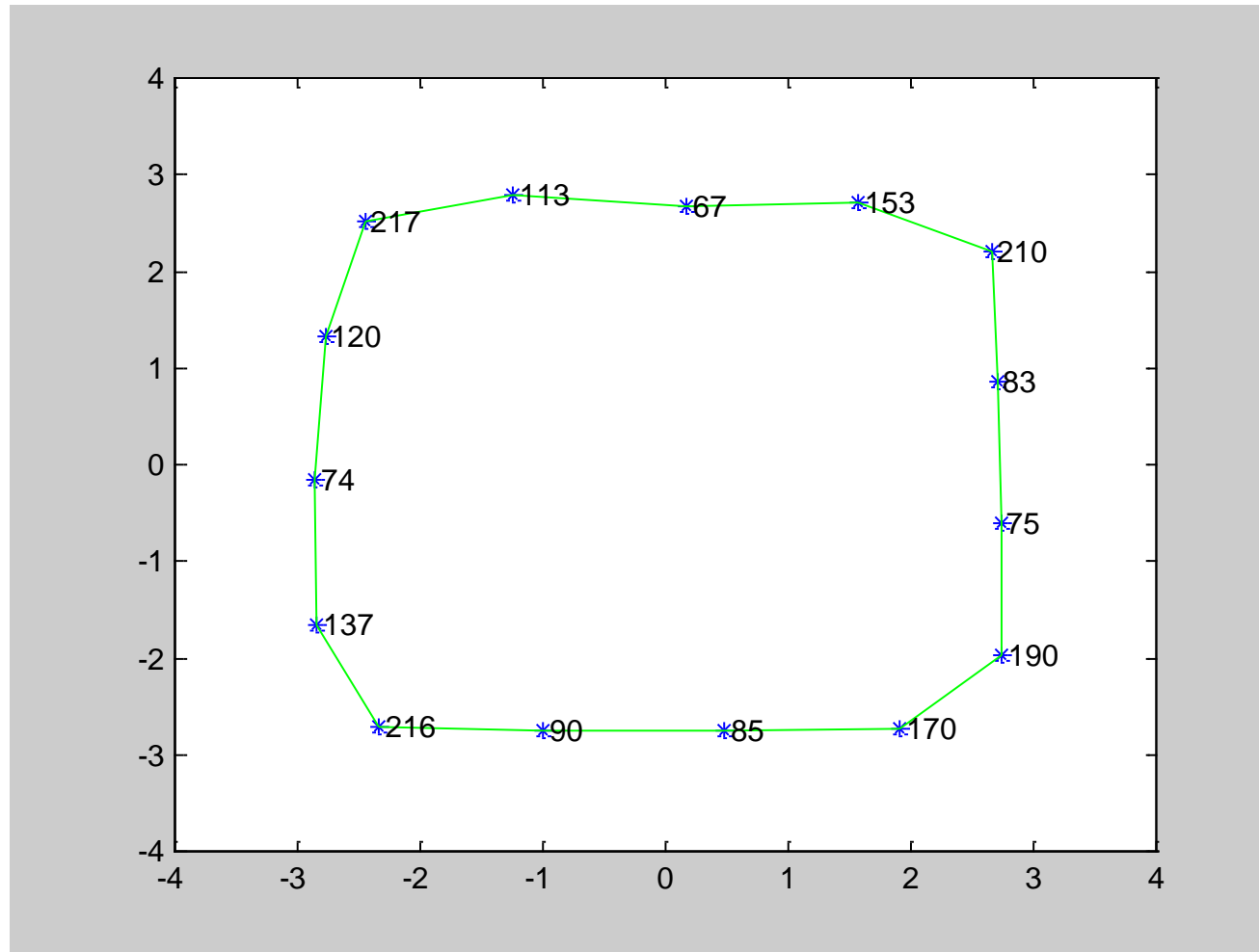


Figura 18 – Configuração final de 15 neurônios em um arranjo unidimensional e estatística dos vencedores

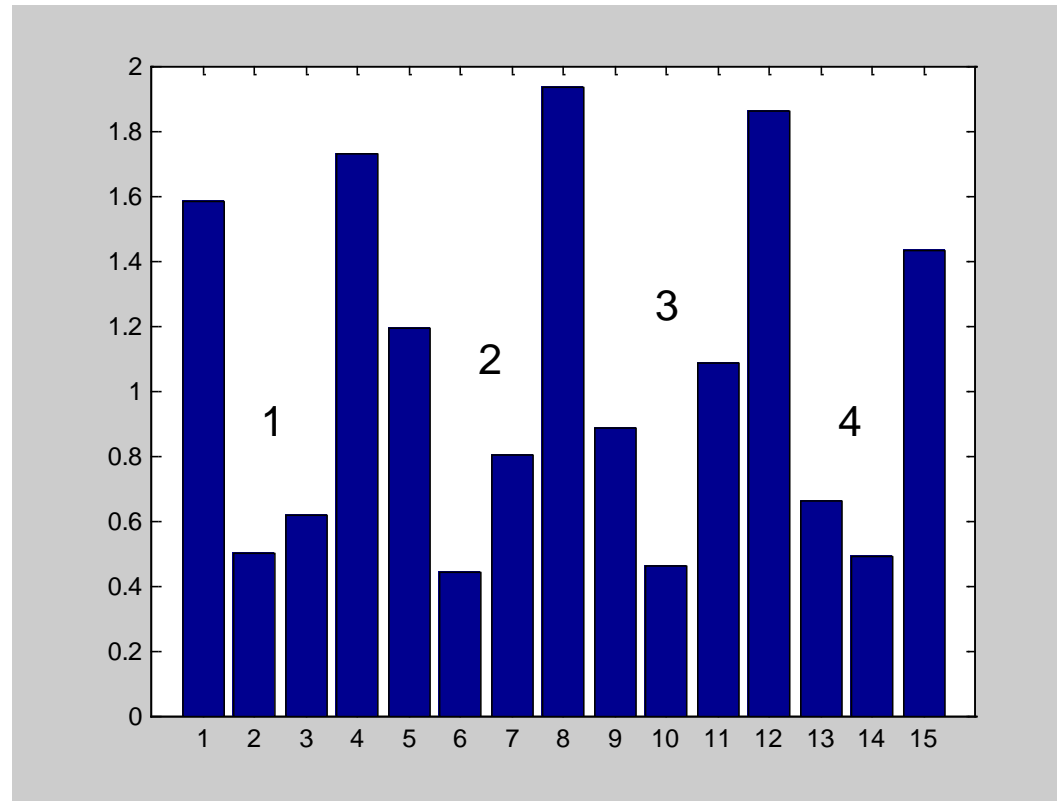


Figura 19 – Vetor-U levando em conta a estatística dos vencedores

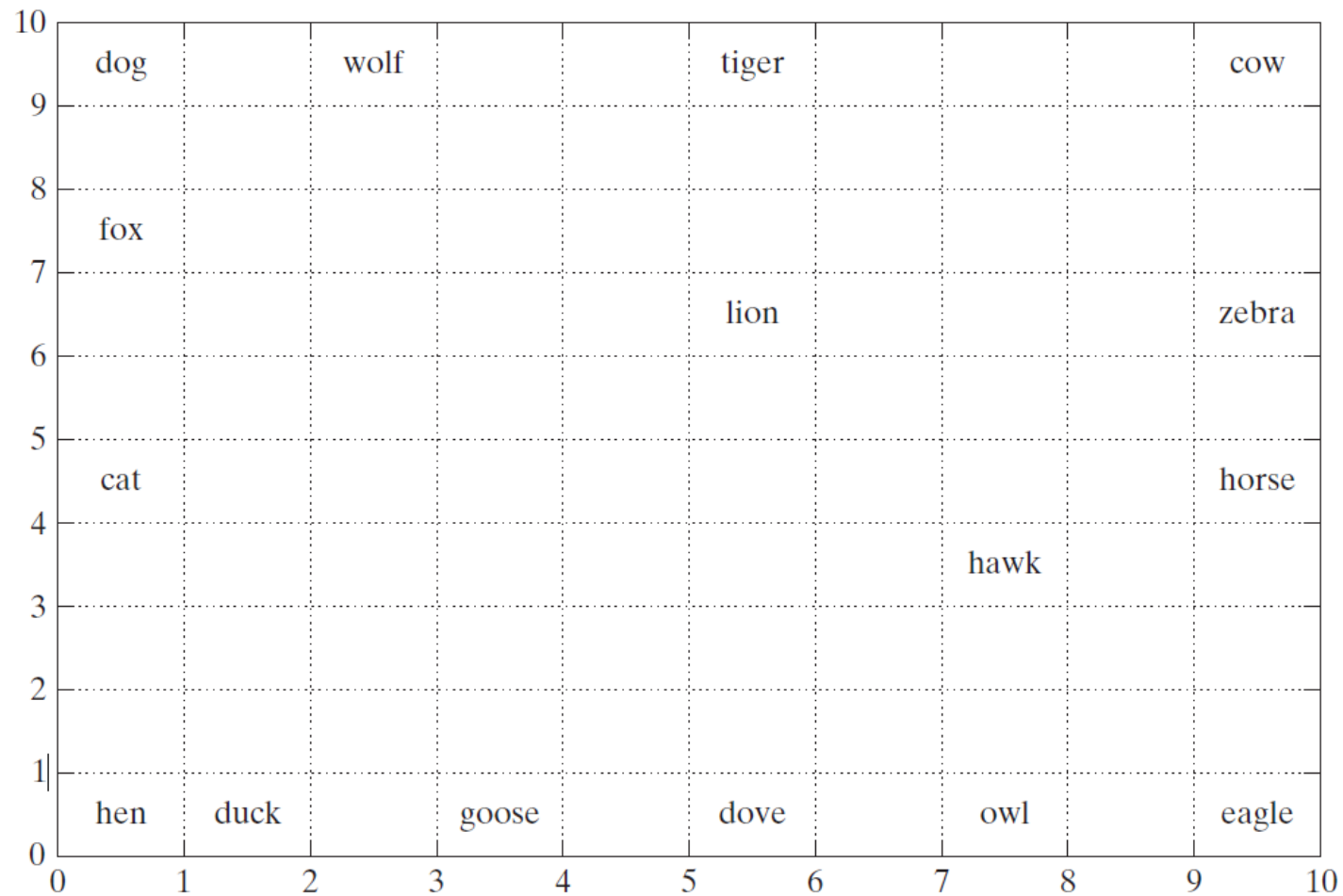
Material a consultar: Silva, A.E.A.; Cunha, K.E.B.K.; Von Zuben, F.J.; Martins, L.E.G. “Uma Análise da Representatividade da Matriz\_U para o Mapa Auto-Organizável de Kohonen, III Encontro Nacional de Inteligência Artificial (ENIA), 2001, Fortaleza, CE. Anais do XXI Congresso da Sociedade Brasileira de Computação (SBC’2001), 2001.

## 9.10 Mapa contextual

TABLE 9.2 Animal Names and Their Attributes

Animal		Dove	Hen	Duck	Goose	Owl	Hawk	Eagle	Fox	Dog	Wolf	Cat	Tiger	Lion	Horse	Zebra	Cow
is	small	1	1	1	1	1	1	0	0	0	0	1	0	0	0	0	0
	medium	0	0	0	0	0	0	1	1	1	1	0	0	0	0	0	0
	big	0	0	0	0	0	0	0	0	0	0	0	1	1	1	1	1
has	2 legs	1	1	1	1	1	1	1	0	0	0	0	0	0	0	0	0
	4 legs	0	0	0	0	0	0	0	1	1	1	1	1	1	1	1	1
	hair	0	0	0	0	0	0	0	1	1	1	1	1	1	1	1	1
	hooves	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	1
	mane	0	0	0	0	0	0	0	0	0	1	0	0	1	1	1	0
	feathers	1	1	1	1	1	1	1	0	0	0	0	0	0	0	0	0
likes to	hunt	0	0	0	0	1	1	1	1	0	1	1	1	1	0	0	0
	run	0	0	0	0	0	0	0	0	1	1	0	1	1	1	1	0
	fly	1	0	0	1	1	1	1	0	0	0	0	0	0	0	0	0
	swim	0	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0

Figura extraída de HAYKIN (2008), pg. 448.



**FIGURE 9.10** Feature map containing labeled neurons with strongest responses to their respective inputs.

Figura extraída de HAYKIN (2008), pg. 449.

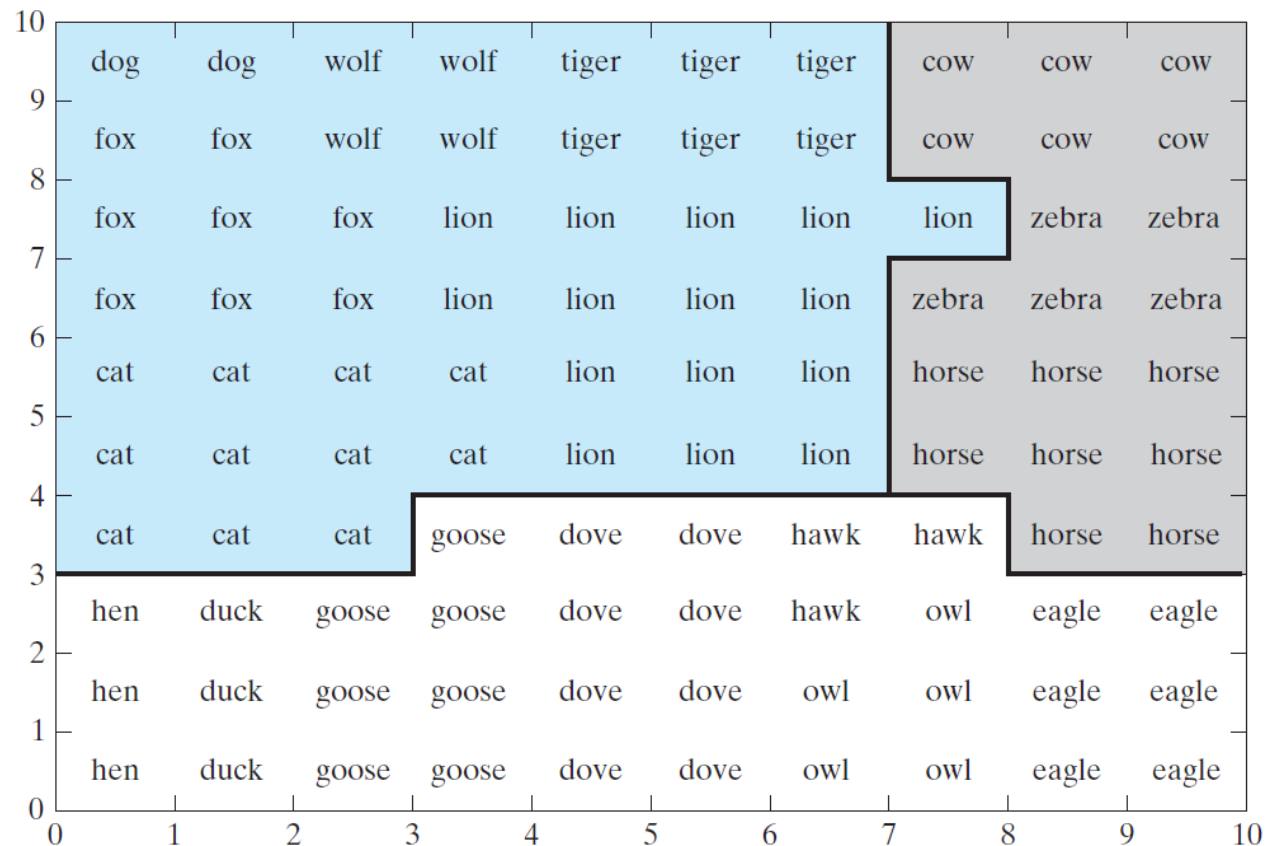


Figura extraída de HAYKIN (2008), pg. 450.

Cada neurônio foi rotulado de acordo com o animal para o qual há uma maior ativação. Repare que o mapa contextual separou as espécies em 3 conjuntos: pássaros (em branco), espécies pacíficas (em cinza) e predadores (em azul).

## 9.11 Ferramentas de visualização e discriminação

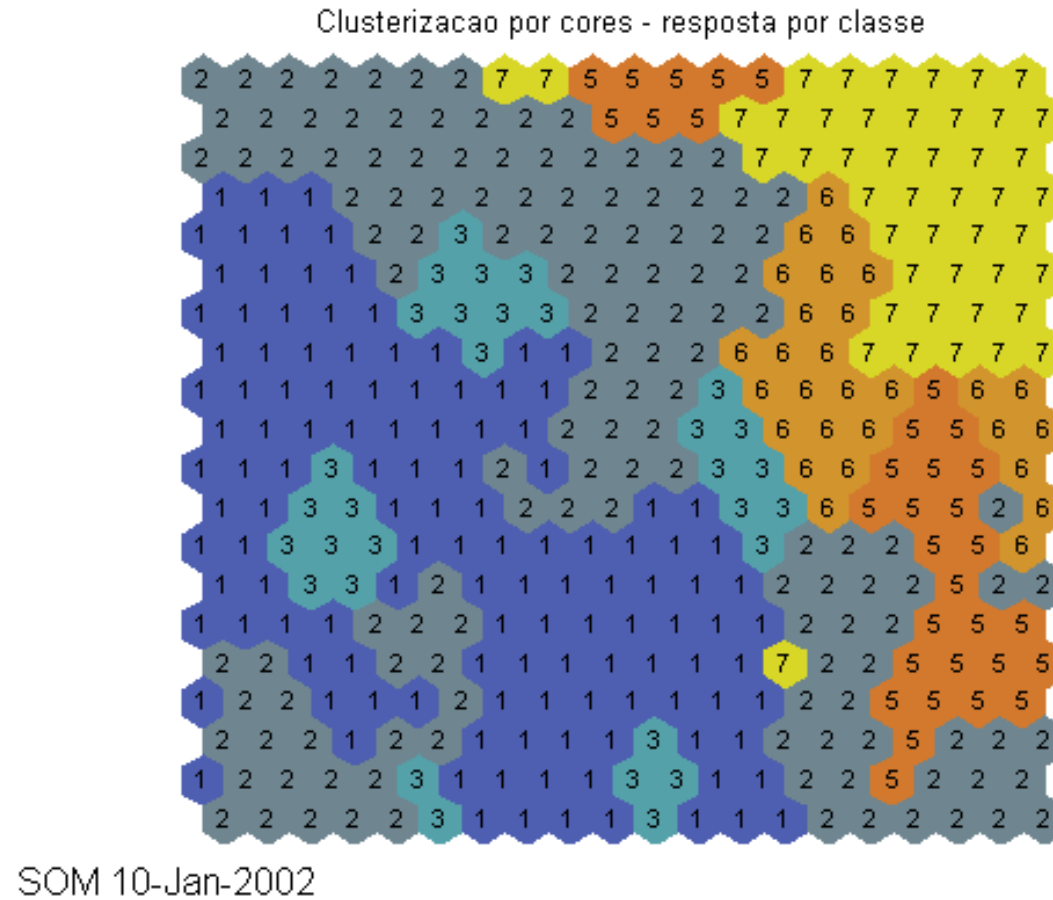
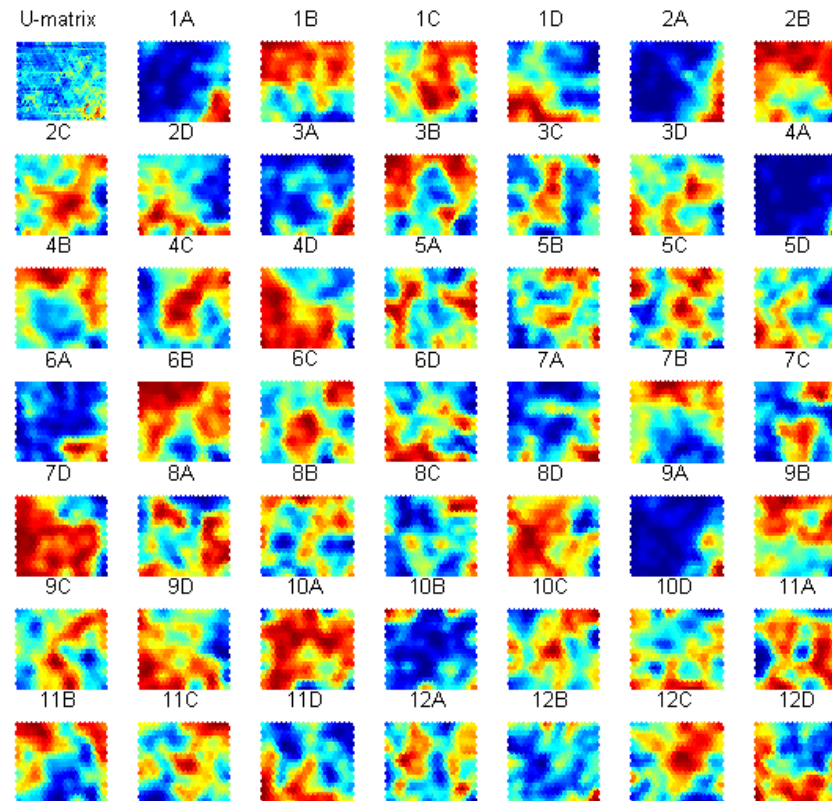


Figura 20 – Interpretação do mapa após auto-organização (figura extraída de ZUCHINI, 2003)

- No mapa da Figura 20, cada neurônio é representante (neurônio com maior ativação) de um subconjunto dos dados de treinamento e recebe o rótulo da classe da maioria dos dados por ele representados.
- Algumas informações podem ser extraídas do mapa da Figura 20, mesmo sem saber qual é a natureza dos dados de treinamento que levaram a este resultado:
  - ✓ Por requerer muitos neurônios para representá-la, quando comparada, por exemplo, à Classe 6, pode-se afirmar que a Classe 1 é bastante diversa em termos de perfil característico, embora haja uma continuidade de perfis (o subconjunto de neurônios que a representa é conexo).
  - ✓ Níveis elevados de diversidade também são verificados junto ao perfil da Classe 2, mas esta não exibe uma continuidade de perfis característicos (é representada por subconjuntos desconexos de neurônios), o que indica que a Classe 2 poderia ser dividida em subclasses mais coerentes internamente.
  - ✓ As Classes 3 e 5 também poderiam ser divididas em subclasses (4 ou 5 subclasses para a Classe 3 e 2 subclasses para a Classe 5) com maior coerência interna.

- ✓ Há exemplares da Classe 2 parecidos com o perfil característico de qualquer uma das outras classes existentes, pois neurônios alocados como representantes da Classe 2 são vizinhos de neurônios alocados como representantes de todas as demais classes. Já neurônios da Classe 5 são vizinhos apenas de neurônios das Classes 2, 6 e 7. Neurônios da Classe 1, por sua vez, são vizinhos apenas de neurônios das Classes 2 e 3 (tirando os possíveis *outliers* da Classe 7).
- ✓ O perfil característico da Classe 6 é mais similar ao da Classe 7 do que o perfil característico da Classe 1 o é para a mesma Classe 7 (os neurônios representantes das Classes 1 e 7 estão mais distantes no mapa do que os representantes das Classes 6 e 7).
- ✓ A Classe 7 e também as Classes 1, 2 e 3, possivelmente, têm alguns *outliers*, representados por um ou dois neurônios isolados dos demais da mesma classe.





SOM 04-Apr-2002

Figura 21 – Busca por correlações entre atributos (matriz-U por atributo de entrada do mapa) após a auto-organização (figura extraída de ZUCHINI, 2003) (VESANTO & AHOLA, 1999)

- Trata-se de um caso de estudo em que os dados de entrada têm 48 atributos. A matriz-U original (alto à esquerda) considera essas 48 dimensões para calcular a distância entre neurônios vizinhos, enquanto que as demais matrizes-U só consideram um atributo por vez.

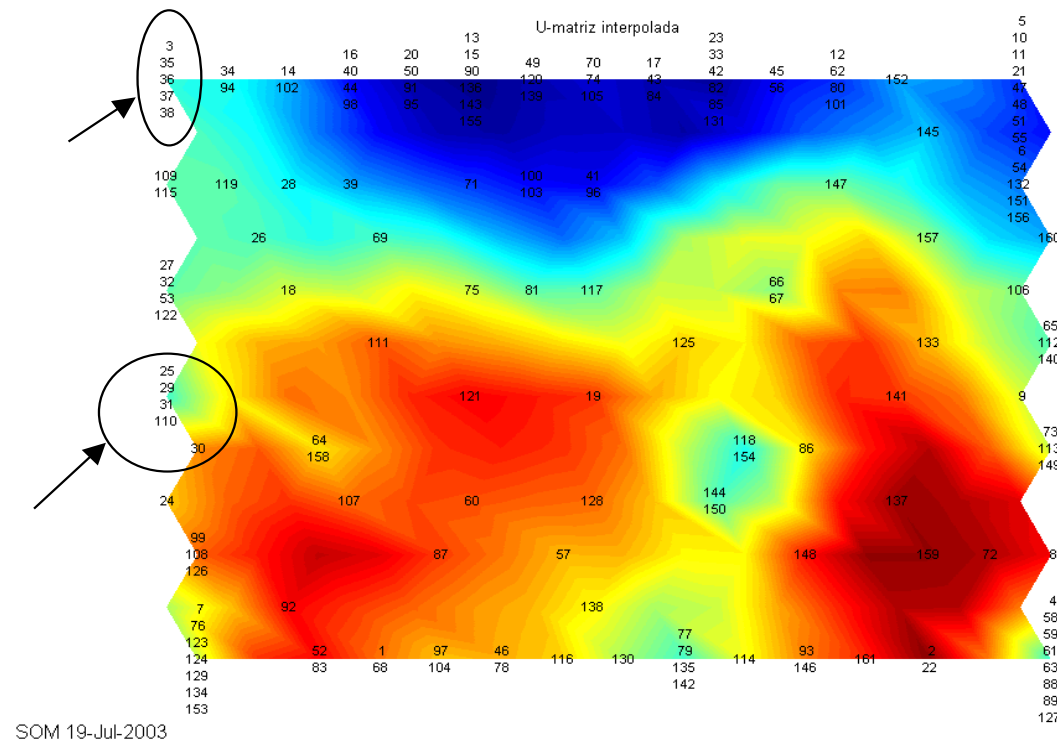
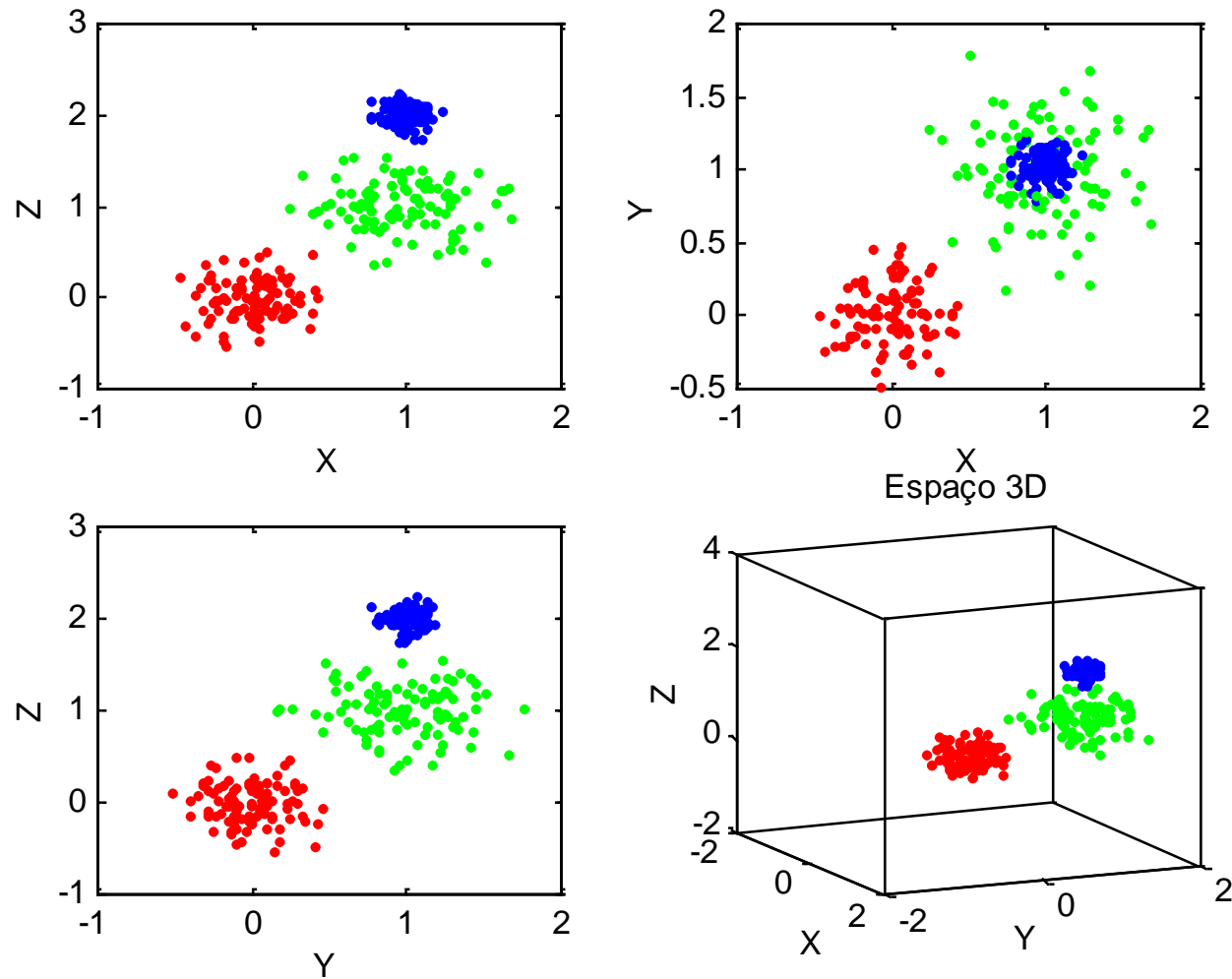


Figura 22 – Matriz-U após auto-organização de textos (figura extraída de ZUCHINI, 2003)

- Este mapa 2D foi treinado com textos referentes a centenas de resumos de um congresso interno de uma faculdade brasileira, que envolvia as áreas de (1) Exatas e Tecnológicas (55 textos); (2) Humanas e Sociais (69 textos); e (3) Biológicas e Saúde (37 textos). A matriz-U da Figura 22 não possui 3 grupos característicos, mas muitos grupos de resumos sabidamente correlacionados em seu conteúdo foram mapeados num mesmo neurônio (ou em neurônios vizinhos), conforme indicado pelas setas e círculos.

## 9.12 Comparação com PCA e Análise por Componentes Curvos (CCA)



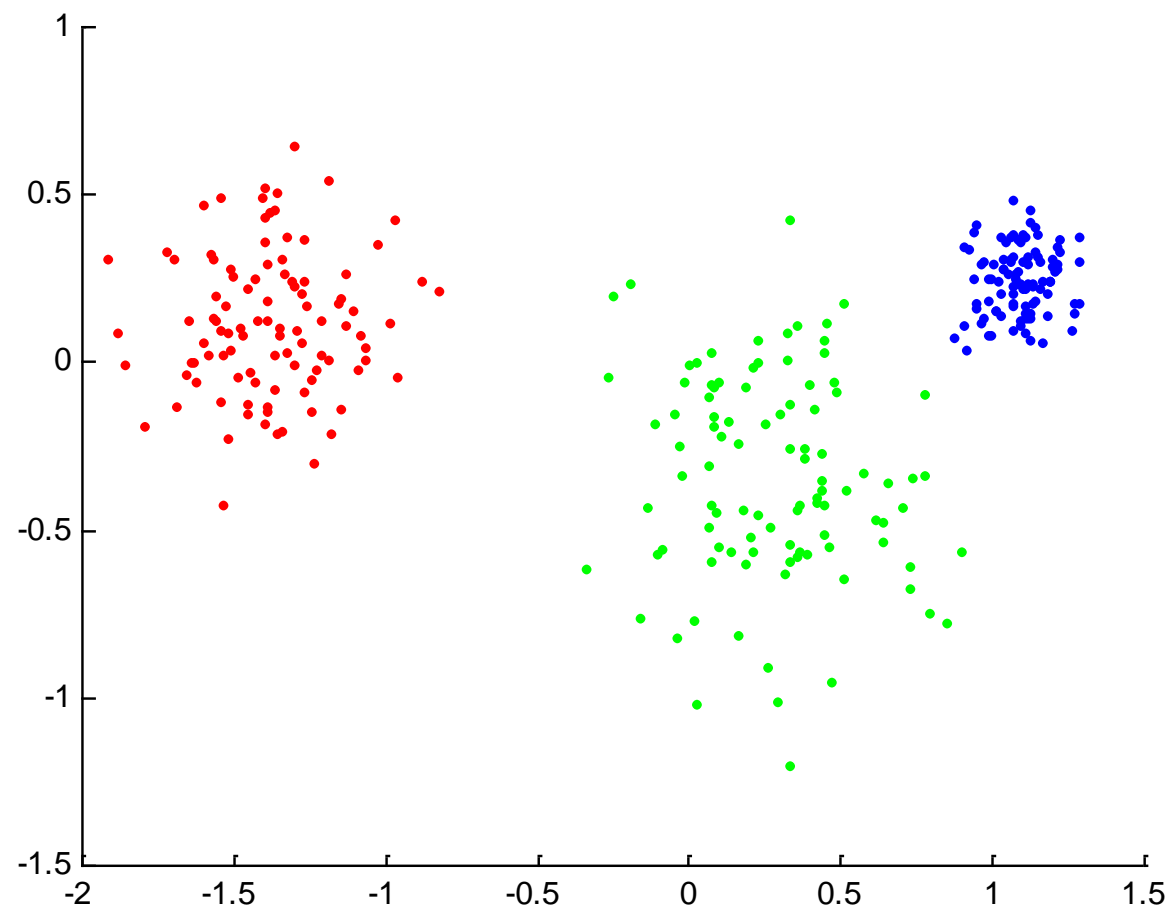


Figura 21 – Projeção PCA (figuras extraídas de ZUCHINI, 2003)

Matriz de Distancia

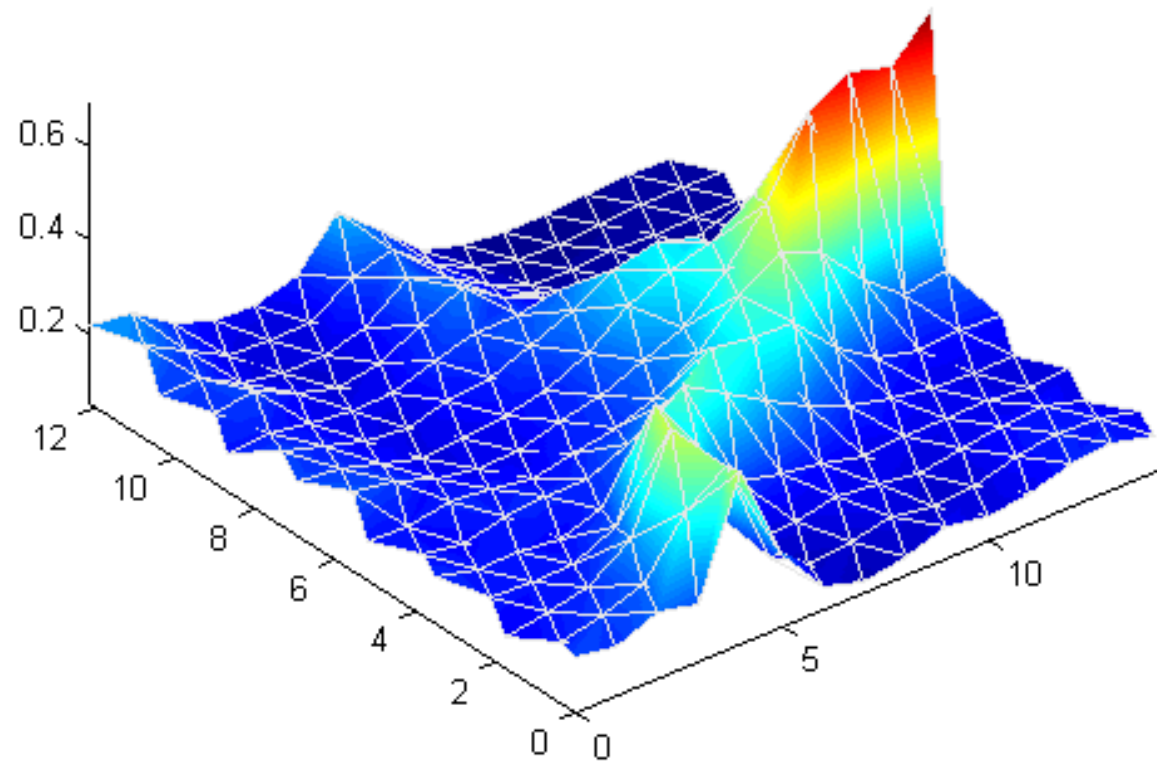


Figura 22 – Matriz-U para grid hexagonal (figura extraída de ZUCHINI, 2003)

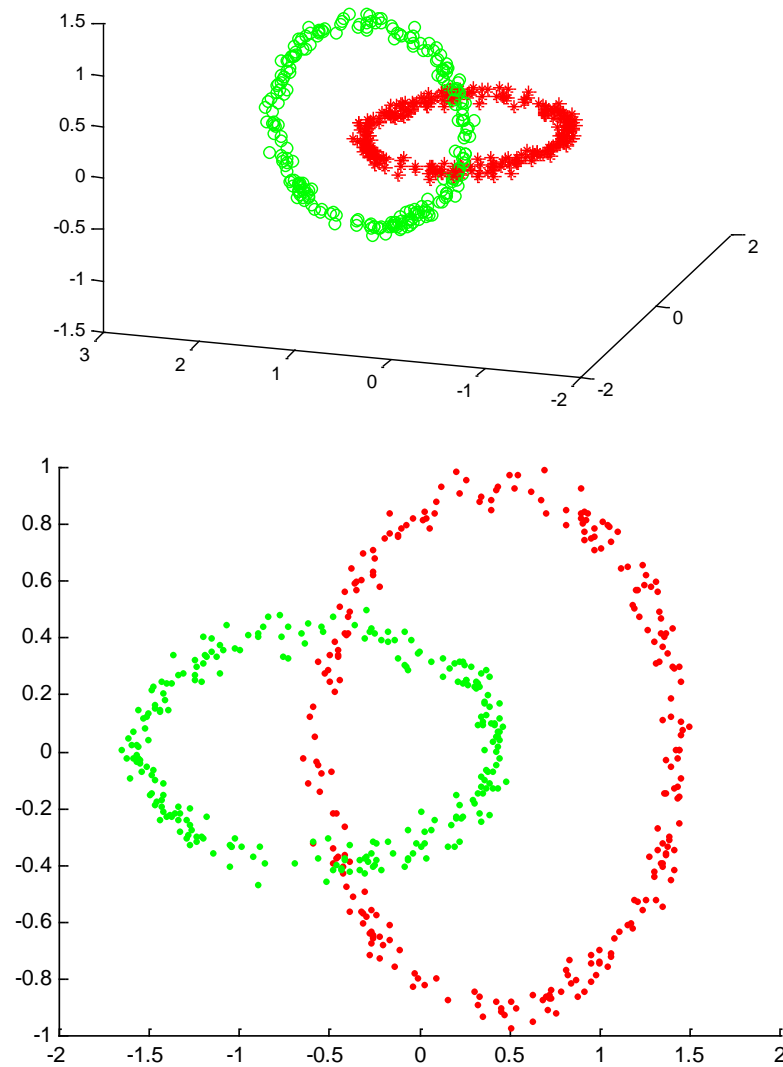


Figura 23 – Projeção PCA (figuras extraídas de ZUCHINI, 2003)

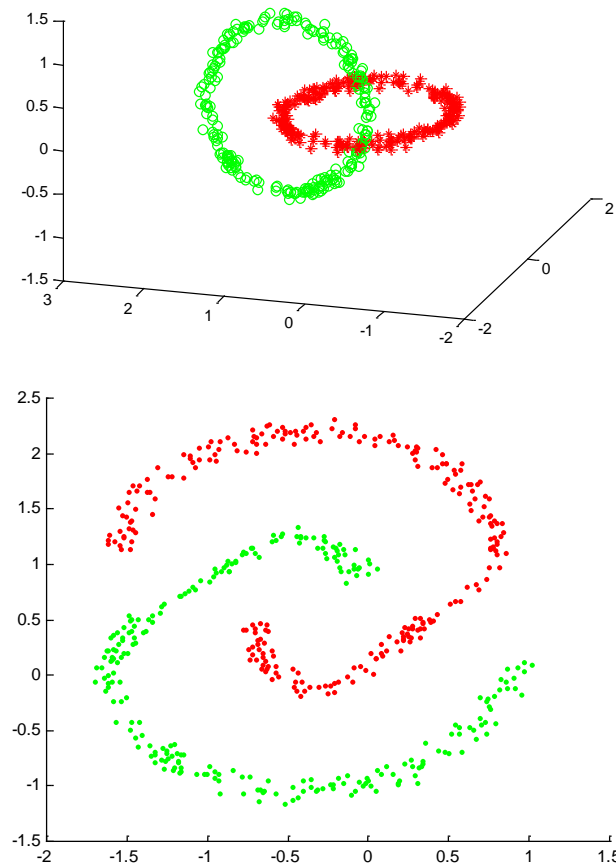


Figura 24 – CCA (figuras extraídas de ZUCHINI, 2003)

Referência para CCA: Hastie, T. & Stuetzle, W. “Principle curves”, Journal of the American Statistical Association, vol. 84, no. 406, pp. 502-516, 1989.

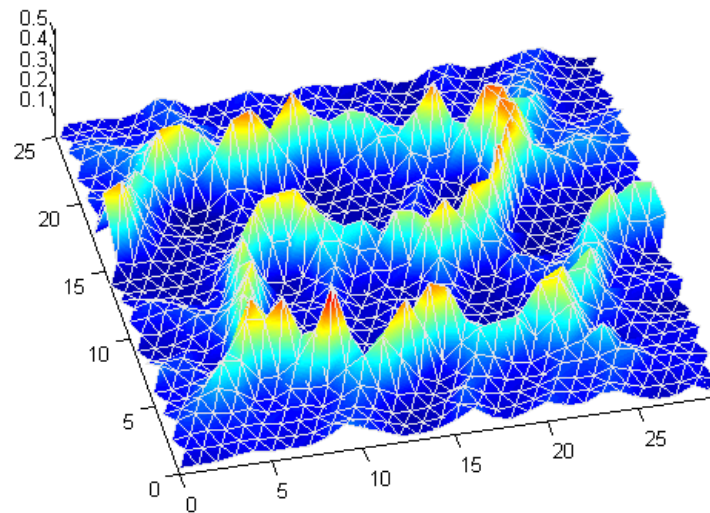
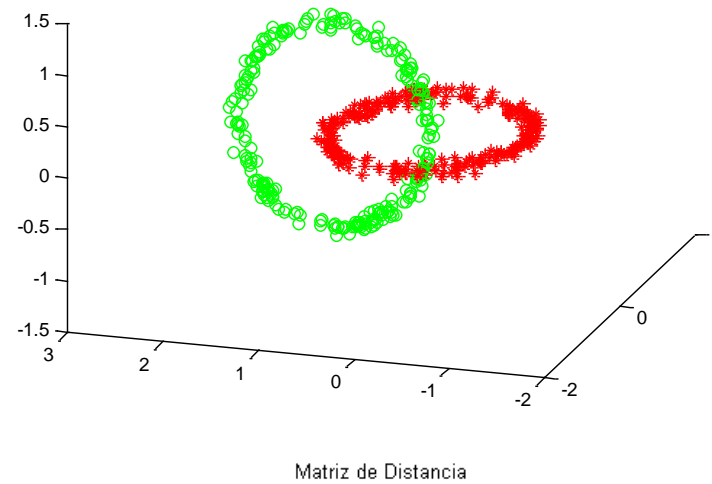


Figura 25 – Matriz-U para grid hexagonal (figuras extraídas de ZUCHINI, 2003)



### 9.13 Ordenamento de pontos em espaços multidimensionais

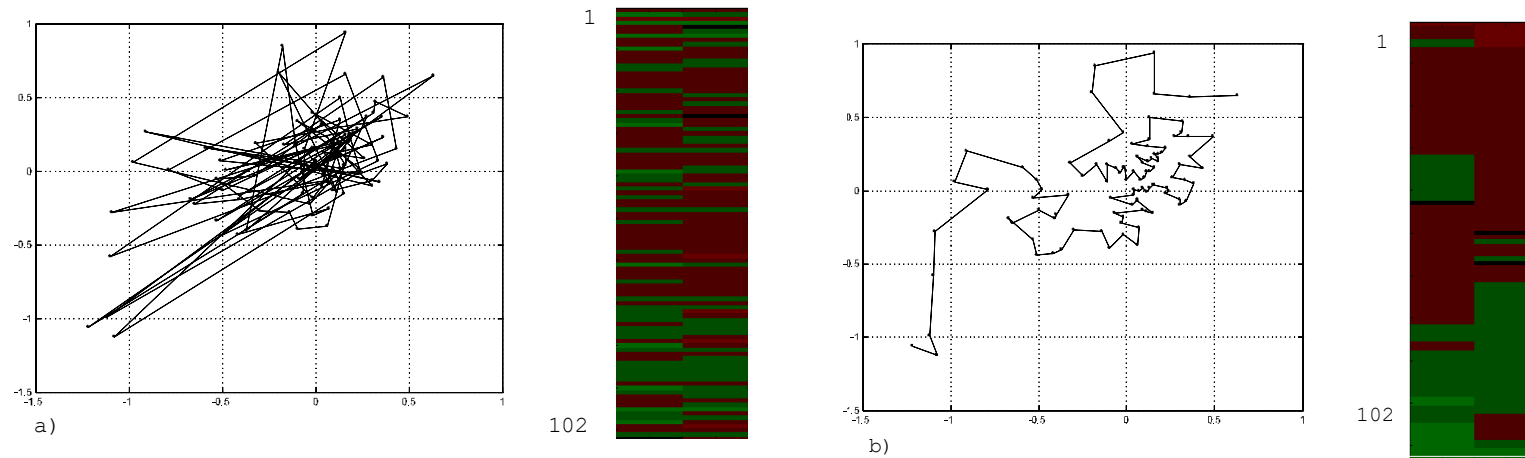


Figura 26 – Situação anterior (esquerda) e situação posterior (direita) ao ordenamento de pontos no  $\mathbb{R}^2$  (GOMES *et al.*, 2004).

- A extensão para pontos em espaços de maior dimensão é imediata (generalização do problema do caixeiro viajante)

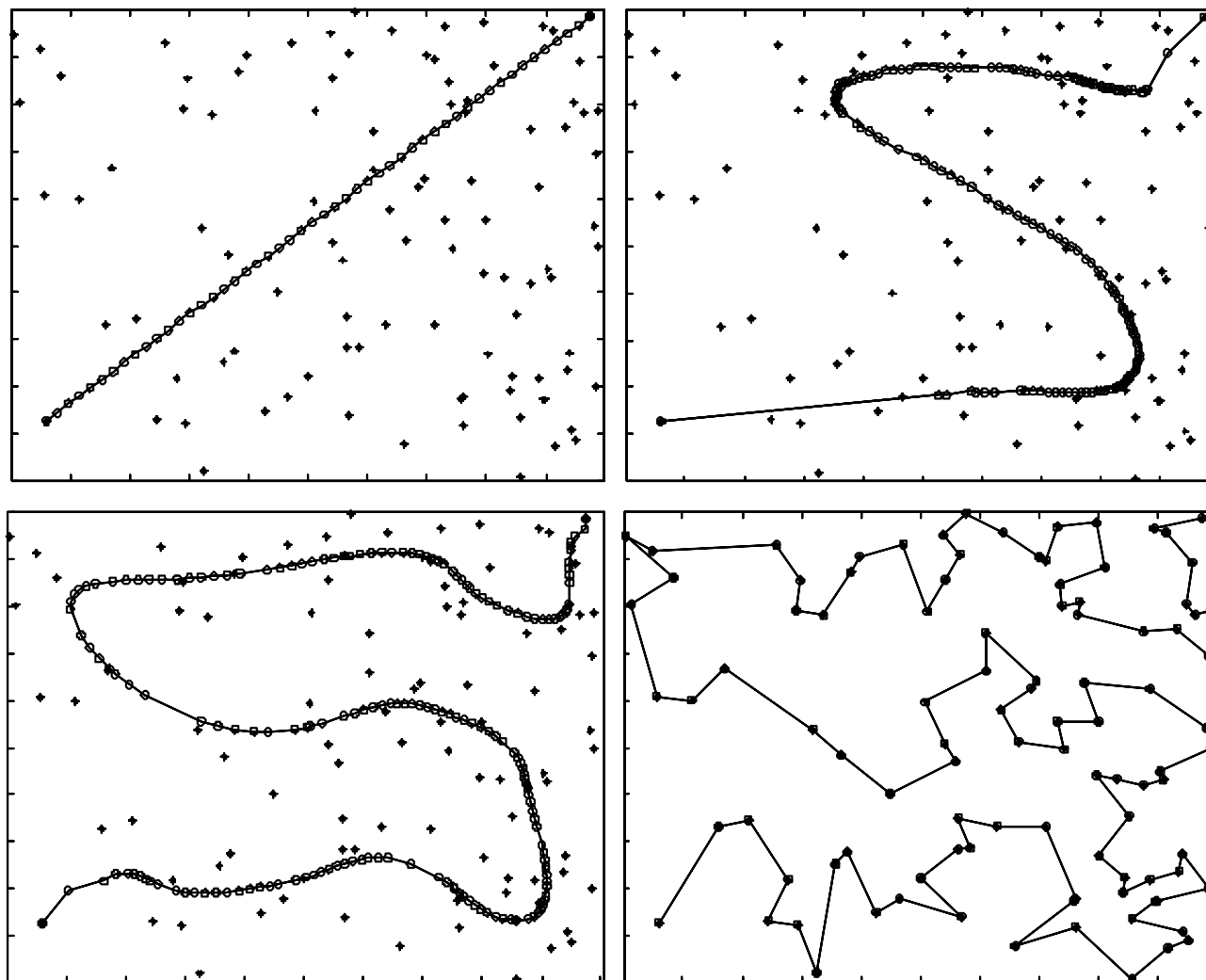


Figura 27 – Modo de operação (GOMES *et al.*, 2004)

- O mapa da Figura 29 é construtivo, com etapas de poda.
- Quando e onde inserir novos neurônios?
- Resposta: Neurônios que vencem muito (conceito que precisa de uma definição matemática) devem ser “clonados”, gerando um vizinho bem próximo a ele.
- Quando e onde podar neurônios presentes no mapa?
- Resposta: Neurônios que nunca vencem ou vencem muito pouco (conceito que precisa de uma definição matemática) devem ser podados e a vizinhança deve ser adequadamente redefinida.
- Esta técnica pode ser prontamente aplicada a problemas de caixeiro viajante. De fato, o problema da Figura 29 é um problema de caixeiro viajante, mas com a particularidade de que não é necessário retornar da última cidade à primeira cidade.
- Para mais detalhes, consultar: ANGÉNIOL *et al.* (1988), ARAS *et al.* (1999), GOMES *et al.* (2004), MATSUYAMA (1991) e MODARES *et al.* (1999).

## 9.14 Roteamento de veículos (múltiplos mapas auto-organizáveis)

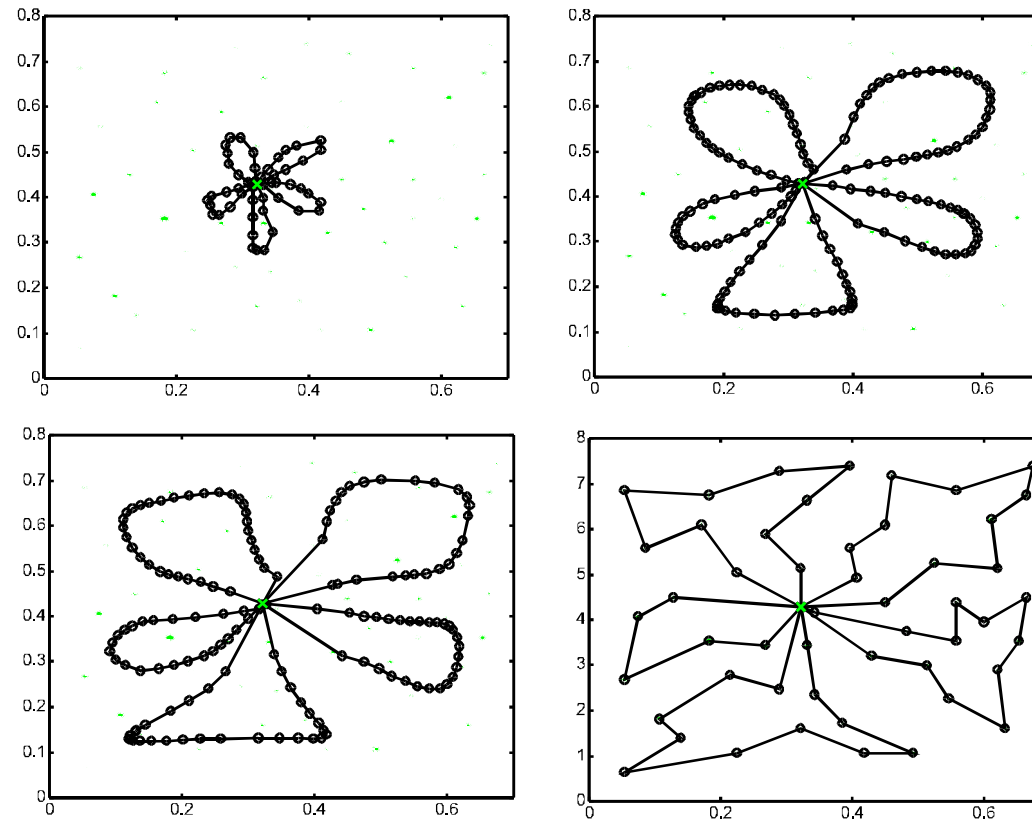


Figura 28 – Várias etapas do processo de auto-organização

- Por existirem múltiplos mapas competindo, o desafio aqui é definir o neurônio vencedor, pois deve ser levada em conta a competição entre os mapas (GOMES & VON ZUBEN, 2002).

## 9.15 Mapas auto-organizáveis construtivos

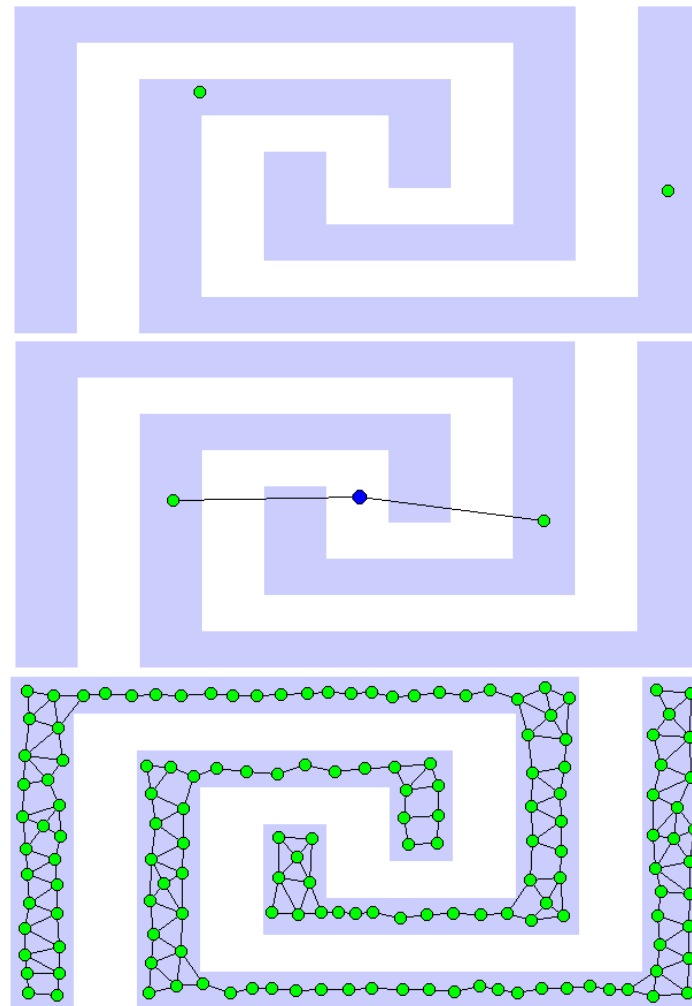


Figura 29 – *Growing Neural Gas* (FRITZKE, 1995)

## 9.16 Questões a serem melhor investigadas

- Sintonia de parâmetros
- Neurônios que não vencem nunca (devem ser podados para aumentar eficiência)
- Neurônios que vencem sempre
- Dimensão do arranjo para uma dada aplicação
- Número de neurônios, uma vez definido o arranjo
- Inicialização dos pesos
- Apresentação dos dados à rede (padrão-a-padrão ou em batelada?)
- Interpretação do mapa resultante (análise discriminante)
- Métodos construtivos e de poda
- Outras aplicações e múltiplos mapeamentos simultâneos
- Comparações com ferramentas similares

## 9.17 Extensões

- Learning Vector Quantization (LVQ) (KOHONEN, 1995)

$$\Delta \mathbf{w}_j = \begin{cases} \gamma(\mathbf{x}_i - \mathbf{w}_j) & \text{se a classe for correta} \\ -\gamma(\mathbf{x}_i - \mathbf{w}_j) & \text{se a classe for incorreta} \end{cases}$$

## 10 Referências

- ADAMI, C. “What is complexity?”, *BioEssays*, vol. 24, no. 12, pp. 1085-1094, 2002.
- ANGÉNIOL, B., VAUBOIS, C. & LE TEXIER, J.Y. “Self-organizing feature maps and the travelling salesman problem”, *Neural Networks*, 1, 289-293, 1988.
- ARAS, N., OOMMEN, B.J. & ALTINEL, I.K. “The Kohonen Network Incorporating Explicit Statistics and its Application to The Travelling Salesman Problem”, *Neural Networks*, vol. 12, pp. 1273-1284, 1999.
- BARLOW, H. B. “Unsupervised learning”, *Neural Computation*, vol. 1, pp. 295-311, 1989.
- BAK, P. “How Nature Works: The Science of Self-Organized Criticality”, Copernicus, 1996.
- BECKER, S. & PLUMBLEY, M. “Unsupervised neural network learning procedures for feature extraction and classification”, *International Journal of Applied Intelligence*, 6: 185-203, 1996.
- BUSHEV, M. “Synergetics: Chaos, Order, Self-Organization”, World Scientific, 1994.

- COHEN, J. & STEWART, I. “The collapse of chaos: discovering simplicity in a complex world”, Penguin Books, 1995.
- COSTA, J.A.F. “Classificação Automática e Análise de Dados por Redes Neurais Auto-Organizáveis”, Tese de Doutorado, Faculdade de Engenharia Elétrica e de Computação (FEEC/Unicamp), Dezembro 1999.
- COVENEY, P. & HIGHFIELD, R. “Frontiers of Complexity: The Search for Order in a Chaotic World”, Fawcett Columbine, 1995.
- COWAN, G.A., PINES, D. & MELTZER, D. (eds.) “Complexity: Metaphors, Models, and Reality”, Proceedings Volume XIX, Santa Fe Institute Studies in the Sciences of Complexity, Perseus Books, 1994.
- DE CASTRO, L.N. & VON ZUBEN, F.J. An Improving Pruning Technique with Restart for the Kohonen Self-Organizing Feature Map. *Proceedings of the IEEE International Joint Conference on Neural Networks (IJCNN'99)*, vol. 3, pp. 1916-1919, July 1999.
- DE CASTRO, L.N. & VON ZUBEN, F.J. A Pruning Self-Organizing Algorithm to Select Centers of Radial Basis Function Neural Networks. *Proceedings of the International Conference of Artificial Neural Networks and Genetic Algorithms (ICANNGA'01)*, pp. 114-117, 2001.
- DUDA, R.O.; HART, P.E.; STORK, D.G. “Unsupervised Learning and Clustering”, Book chapter, in “Pattern Classification”, 2nd edition, Wiley, 2001.
- EVERITT, B. “Cluster Analysis”, 3rd. edition, John Wiley, 1993.
- FAQ: The self-organized systems (<http://www.calresco.org/sos/sosfaq.htm>)



- FAVATA, F. & WALKER, R. “A Study of the Application of Kohonen-Type Neural Networks to the Traveling Salesman Problem”, *Biological Cybernetics* 64, 463-468, 1991.
- FORT, J.C. “Solving a Combinatorial Problem via Self-Organizing Maps”, *Biological Cybernetics*, 59, 33-40, 1988.
- FRITZKE, B. “A Growing Neural Gas Network Learns Topologies”, in Tesauro, G., Touretzky, D.S., and Leen, T.K. (eds.). *Advances in Neural Information Processing Systems 7*, The MIT Press, pp. 625-632, 1995.
- GOMES, L.C.T. & VON ZUBEN, F.J. A Neuro-Fuzzy Approach to the Capacitated Vehicle Routing Problem. *Proceedings of the IEEE International Joint Conference on Neural Networks (IJCNN'2002)*, vol. 2, pp. 1930-1935, Honolulu, Hawaii, May 12-17, 2002.
- GOMES, L.C.T. & VON ZUBEN, F.J. Vehicle Routing Based on Self-Organization With and Without Fuzzy Inference. *Proceedings of the IEEE International Conference on Fuzzy Systems (FUZZ-IEEE'2002)*, vol. 2, pp. 1310-1315, Honolulu, Hawaii, May 12-17, 2002.
- GOMES, L.C.T., VON ZUBEN, F.J. & MOSCATO, P.A. “A Proposal for Direct-Ordering Gene Expression Data by Self-Organising Maps”, *International Journal of Applied Soft Computing*, vol. 5, pp. 11-21, 2004.
- HAYKIN, S. *Neural Networks and Learning Machines*, 3rd edition, Prentice Hall, 2008.
- HEBB, D. O. “The Organization of Behavior”, Wiley, 1949.
- HINTON, G.E. “Connectionist learning procedures”, *Artificial Intelligence*, 40: 185-234, 1989.

- HINTON, G.E. & SEJNOWSKI, T.J. “Learning and relearning in Boltzmann machines”, in D. E. Rumelhart, J. L. McClelland & The PDP Research Group (eds.) *Parallel Distributed Processing: Explorations in the Microstructure of Cognition*, MIT Press, vol. 1, pp. 282-317, 1986.
- HINTON, G.E. & SEJNOWSKI, T.J. (Eds.) “Unsupervised Learning: Foundations of Neural Computation”, MIT Press, 1999.
- JAIN, A.K. & DUBES, R.C. “Algorithms for Clustering Data”, Prentice-Hall, 1988.
- JAIN, A.K., MURTY, M.N. & FLYNN, P.J. “Data Clustering: A Review”, *ACM Computing Surveys*, vol. 31, no. 3, pp. 264-323, 1999.
- KASKI, S. “Data Exploration Using Self-Organizing Maps”, Ph.D. Thesis, Helsinki University of Technology, Neural Networks Research Centre, 1997.
- KAUFFMAN, S. “At Home in the Universe: The Search for the Laws of Self-Organization and Complexity”, Oxford University Press, 1995.
- KAUFFMAN, S. “The Origins of Order: Self-Organization and Selection in Evolution”, Oxford University Press, 1993.
- KOHONEN, T. “Self-organized formation of topologically correct feature maps”, *Biological Cybernetics*, vol. 43, pp. 59-69, 1982.
- KOHONEN, T. “Self-Organization and Associative Memory”, 3rd. edition, Springer, 1989 (1st. edition, 1984).
- KOHONEN, T. “The Self-Organizing Map”, *Proceedings of the IEEE*, 78:1464-1480, 1990.
- KOHONEN, T. “Learning vector quantization”, in M.A. Arbib (ed.) *The Handbook of Brain Theory and Neural Networks*, pp. 537-540, MIT Press, 1995.

- KOHONEN, T., OJA, E., SIMULA, O., VISA, A. & KANGAS, J. “Engineering applications of the self-organizing map”, *Proceedings of the IEEE*, 84:1358-1384, 1996.
- KOHONEN, T. “Self-Organizing Maps”, 2nd. edition, Springer, 1997.
- LINSKER, R. “Self-organization in a perceptual network”, *Computer*, 21: 105-128, 1988.
- MACKAY, D. M. “The epistemological problem for automata”, in C. E. Shannon & J. McCarthy (eds.) *Automata Studies*, Princeton University Press, pp. 235-251, 1956.
- MALKIN, D. “The Evolutionary Impact of Gradual Complexification on Complex Systems”, Engineering Doctorate Thesis, University College London, 2009.
- MARR, D. “A theory for cerebral neocortex”, *Proceedings of the Royal Society of London, Series B*, vol. 176, pp. 161-234, 1970.
- MATSUYAMA, Y. “Self-Organization via Competition, Cooperation and Categorization Applied to Extended Vehicle Routing Problems”, *Proc. International Joint Conference on Neural Networks*, 1, 385-390, 1991.
- MCSHEA, D. W. “Functional Complexity in Organisms: Parts as Proxies”, *Biology and Philosophy*, vol. 15, no. 5, pp. 641-668, 2000.
- MITLETON-KELLY, E. “Ten Principles of Complexity and Enabling Infrastructures”, in *Complex Systems and Evolutionary Perspectives on Organisations: The Application of Complexity Theory to Organisations*. Elsevier Science, Series in Advanced Management, Chapter 2, 2003.
- MODARES, A., SOMHOM, S. & ENKAWA, T. “A Self-Organizing Neural Network Approach for Multiple Traveling Salesman and Vehicle Routing Problems”, *Int. Transactions in Operational Research*, 6, 591-606, 1999.

- NICOLIS, G. & PRIGOGINE, I. “Exploring complexity: An introduction”, W H Freeman, 1989.
- POTVIN, J.-I. & ROBILLARD, C. “Clustering for Vehicle Routing with a Competitive Neural Network”, *Neurocomputing*, 8, 125-139, 1995.
- SIMON, H.A. “The Sciences of the Artificial”, The MIT Press, 3rd. Edition, 1996.
- SMITH, K.A. “Neural Networks for Combinatorial Optimization: A Review of More than a Decade of Research”, *INFORMS Journal on Computing*, 11, 1, 15-34, 1999.
- STANLEY, K. O. & MIKKULAINEN, R. “Competitive coevolution through evolutionary complexification”, *J. of Artificial Intelligence Research*, vol. 21, no. 1, pp. 63-100, 2004.
- TOOLBOX: <http://www.cis.hut.fi/projects/somtoolbox/>
- ULTSCH, A. “Knowledge Extraction from Self-Organizing Neural Networks”, in O. Opitz *et al.* (eds.) *Information and Classification*, Springer, pp. 301-306, 1993.
- VAKHUTINSKY, A. I. & GOLDEN, B. L. “Solving Vehicle Routing Problems Using Elastic Nets”, *Proc. IEEE International Conference on Neural Networks*, 7, 4535-4540, 1994.
- VESANTO, J. & AHOLA, J. “Hunting for Correlations in Data Using the Self-Organizing Map”, in International ICSC Congress on Computational Intelligence Methods and Applications (CIMA'99), ICSC Academic Press, pp. 279-285, 1999.
- WILSON, R.A. & KEIL, F.C. (eds.) “The MIT Encyclopedia of the Cognitive Sciences”, MIT Press, 1999.
- ZUCHINI, M.H. “Aplicações de Mapas Auto-Organizáveis em Mineração de Dados e Recuperação de Informação”, Tese de Mestrado, Faculdade de Engenharia Elétrica e de Computação (FEEC/Unicamp), Setembro 2003.