Redes Neurais Artificiais

AULA 10 – Redes de Kohonen
– Arquitetura e Mapas Auto-Organizáveis –

Prof. Ivan Nunes da Silva

1. Redes de Kohonen
Aspectos introdutórios da arquitetura

- A maioria das estruturas de RNA já estudadas nas aulas anteriores necessita de um conjunto de padrões de **entradas/saídas** para que possam ser devidamente treinadas.
- Os ajustes de seus parâmetros livres (matrizes de pesos e limiares) são realizados a partir das apresentações sucessivas dessas amostras de treinamento, configurando-se assim numa **aprendizagem supervisionada**.
- Entretanto, em certas aplicações, somente o conjunto de **padrões de entradas** está disponível, inexistindo, para tanto, as respectivas saídas desejadas.
- Por outro lado, essas amostras possuem **informações relevantes** sobre o comportamento daquele sistema em que foram extraídas.

1. Redes de Kohonen

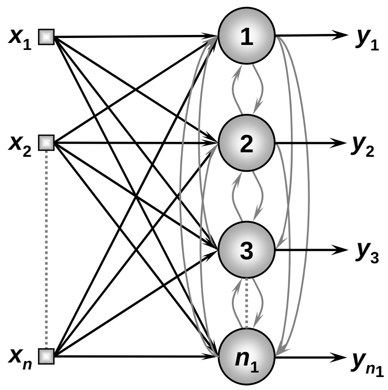
Aspectos de auto-organização e clusterização

- A maioria das redes utilizadas em problemas com essa configuração se auto-organizam por meio de **métodos de treinamento competitivos**.
- Tais métodos têm a capacidade de **detectar similaridades**, regularidades e correlações entre os padrões do conjunto de entrada, agrupando-os em classes (***clusters***).
- Cada uma dessas classes possui então **características particulares** que estão relacionadas com situações e condições que regem o funcionamento do processo.
- Assim, a identificação destes *clusters* é importante para **entendimento das relações** entre os seus elementos constituintes.
  - Permite ainda identificar as funções de um componente ou amostra com base nos atributos dos outros elementos que fazem parte do grupo (*cluster*).
- A principal arquitetura de rede auto-organizada é a **Rede de Kohonen**.

1. Redes de Kohonen

Estrutura neural básica

- Para descrever os passos envolvidos com o **processo de aprendizado competitivo** utilizado na rede de Kohonen, considera-se uma estrutura neural constituída de apenas uma camada neural.



- Como a rede aqui ilustrada é constituída de apenas uma camada, assume-se então a seguinte convenção para os seus vetores de pesos:

$$\mathbf{w}^{(1)} = [W_{1,1} \quad W_{1,2} \quad \ldots \quad W_{1,n}]^T$$

$$\mathbf{w}^{(2)} = [W_{2,1} \quad W_{2,2} \quad \ldots \quad W_{2,n}]^T$$

$$\mathbf{w}^{(3)} = [W_{3,1} \quad W_{3,2} \quad \ldots \quad W_{3,n}]^T$$

(...)

$$\mathbf{w}^{(n_1)} = [W_{n_1,1} \quad W_{n_1,2} \quad \ldots \quad W_{n_1,n}]^T$$

- As **conexões laterais** assumem aqui o papel de que um neurônio pode influenciar na resposta de saída produzida por outro neurônio.

2. Aprendizado Competitivo

Princípios básicos e objetivos

- O princípio básico para o processo de aprendizado competitivo é a **concorrência entre os neurônios**, tendo-se aqui o objetivo de se sair vencedor da prova de competição.
  - Ressalta novamente aqui que o processo é não-supervisionado (sem saída desejada).
- O prêmio para aquele que vence a competição será o **ajuste de seus pesos**, proporcionalmente aos valores do padrão de entrada apresentado, visando-se assim aperfeiçoar o seu estado para a próxima competição (próximo padrão a ser apresentado).
- Nesta circunstância, se todas as conexões laterais deste neurônio vencedor forem nulas (ausência de conexões laterais), então implica que somente os seus pesos serão ajustados, isto é, assume-se a estratégia do **"vencedor leva tudo"** (*winner-take-all*).
- Caso contrário, para a situação de haver valores para as conexões laterais com os seus vizinhos, então um **ajuste proporcional** a tais quantidades será também efetuado no vetor de pesos desses neurônios adjacentes.

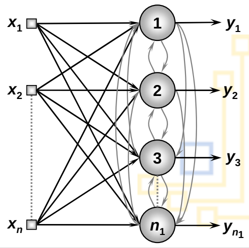
2. Aprendizado Competitivo

Regra para seleção do vencedor

- Para o aprendizado competitivo, há então a necessidade de se estabelecer uma **regra** que defina quem vai ser o neurônio vencedor.
- Uma das medidas mais utilizadas consiste em determinar o **nível de proximidade** existente entre o vetor de pesos de cada neurônio, frente ao vetor de entrada contendo os elementos da *k*-ésima amostra $\{\mathbf{x}^{(k)}\}$, a qual será apresentada nas entradas da rede.
- Uma métrica de proximidade normalmente usada é a distância **(norma euclidiana)** entre esses dois parâmetros, ou seja:

$$dist_j^{(k)} = \sqrt{\sum_{i=1}^{n} \left(x_i^{(k)} - w_i^{(j)}\right)^2}, \quad \text{com } j = 1, \ldots, n_1$$

  - onde $dist_j^{(k)}$ quantifica a distância entre o vetor de entrada, representando a *k*-ésima amostra $\{\mathbf{x}^{(k)}\}$, em relação ao vetor de pesos do *j*-ésimo neurônio $\{\mathbf{w}^{(j)}\}$.



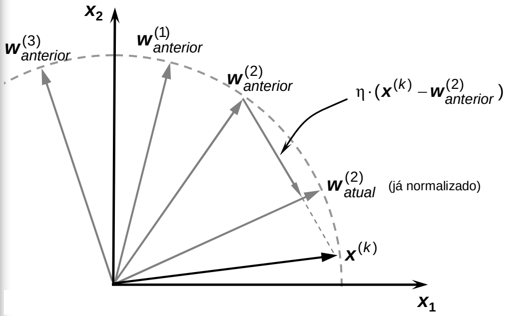
2. Aprendizado Competitivo

Atualização dos pesos do neurônio vencedor

- Como prêmio pela vitória, os pesos $\{\boldsymbol{w}^{(v)}\}$ do neurônio vencedor será ajustado de modo que ele se aproxime ainda mais daquela amostra. Para tanto, utiliza-se o seguinte método de adaptação:

$$\boldsymbol{w}^{(v)}_{atual} = \boldsymbol{w}^{(v)}_{anterior} + \eta \cdot (\boldsymbol{x}^{(k)} - \boldsymbol{w}^{(v)}_{anterior}) \Leftrightarrow \boxed{\boldsymbol{w}^{(v)} \leftarrow \boldsymbol{w}^{(v)} + \eta \cdot (\boldsymbol{x}^{(k)} - \boldsymbol{w}^{(v)})}$$

- Para maior eficiência, é desejável que se normalize unitariamente (divisão pelo módulo) todos os vetores de pesos dos neurônios, assim como os vetores de amostras.



1) Três neurônios {representados por $\boldsymbol{w}^{(1)}$, $\boldsymbol{w}^{(2)}$ e $\boldsymbol{w}^{(3)}$} disputam o torneio frente à apresentação do padrão $\boldsymbol{x}^{(k)}$, sendo o mesmo composto por duas entradas x_1 e x_2.
2) Para a ilustração, o neurônio vencedor da competição foi o $\boldsymbol{w}^{(2)}$.
3) como recompensa, o seu vetor de pesos será ajustado.
4) O ajuste acabou rotacionando o vetor de pesos do neurônio vencedor em direção ao vetor da amostra.

7

2. Aprendizado Competitivo

Algoritmo competitivo (Fase de treinamento)

Início {Algoritmo Competitivo – Fase de Treinamento}

<1> Obter o conjunto de amostras de treinamento $\{\boldsymbol{x}^{(k)}\}$;
<2> Iniciar o vetor de pesos de cada neurônio considerando os valores das n_1 primeiras amostras de treinamento;
<3> Normalizar os vetores de amostras e de pesos;
<4> Especificar a taxa de aprendizagem $\{\eta\}$;
<5> Iniciar o contador de número de épocas $\{$ *época* $\leftarrow 0\}$;
<6> Repetir as instruções:
  <6.1> Para todas as amostras de treinamento $\{\boldsymbol{x}^{(k)}\}$, fazer:
    <6.1.1> Calcular as distâncias euclidianas entre $\boldsymbol{x}^{(k)}$ e $\boldsymbol{w}^{(j)}$, conforme a expressão (8.1).
    <6.1.2> Declarar como vencedor o neurônio j que contenha a menor distância euclidiana;
    <6.1.3> Ajustar o vetor de pesos do neurônio vencedor conforme a expressão (8.3);
    <6.1.4> Normalizar o vetor de pesos que foi ajustado na instrução precedente;
  <6.2> *época* $\leftarrow$ *época* + 1;
 Até que: não haja mudanças significativas nos pesos.

Fim {Algoritmo Competitivo – Fase de Treinamento}

8

2. Aprendizado Competitivo

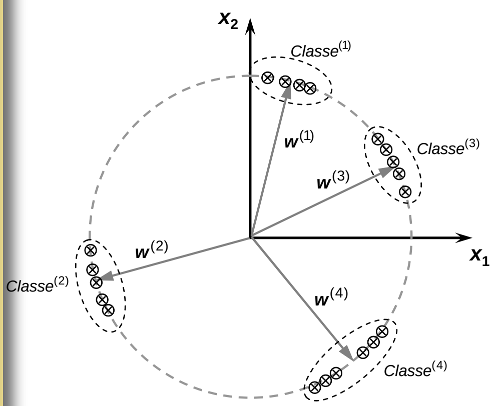
Aspectos de estabelecimento de classes

- Após a convergência, cada um dos vetores de pesos estará posicionado nos **centros dos aglomerados** (*clusters*) que possuem características em comum.
- Tais clusters representarão as **respectivas classes** associadas ao problema de classificação de padrões, sendo que a quantidade destas ficará automaticamente vinculada ao número de neurônios utilizados.
- Vale ressaltar que a **quantidade inicial** de neurônios a ser utilizado, visando representar as classes com características em comum, é de antemão desconhecido, pois o aprendizado é não-supervisionado.
- Assim, torna-se também de suma importância a obtenção de outras **informações adicionais** a respeito do problema a ser mapeado, podendo ser obtidas por:
  a) Especialistas no assunto;
  b) Informações registradas na literatura correlata;
  c) Aplicação de métodos estatísticos que leve para uma estimativa inicial sobre a quantidade de possíveis classes associadas ao problema.
- Para a aplicação formulada no Capítulo 15 (classificação de café), o número de neurônios foi definido em função dos **possíveis tipos** de adulteração que poderiam ser identificados pela rede.

2. Aprendizado Competitivo

Interpretação geométrica da convergência (I)

- Como exemplo, para o caso de se ter 4 neurônios disponíveis, a figura seguinte ilustra como ficaria uma eventual distribuição dos vetores de pesos da rede, após as suas devidas estabilizações.

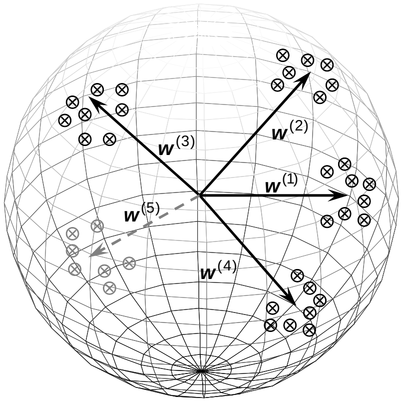


- Cada amostra "⊗" é constituída por **dois sinais** de entrada {x_1 e x_2}.
- Todos os vetores acabam ficando localizados dentro do **círculo unitário**, à medida que são bidimensionais e estão unitariamente normalizados.
- Verifica-se, ainda, que os quatro vetores de pesos se posicionaram nos **centros dos aglomerados** que representam as amostras (identificadas pelo símbolo "⊗").
- Se a estrutura neural fosse constituída por 5 neurônios, ao invés de 4, as amostra pertencentes à *Classe*$^{(4)}$ seriam provavelmente divididas em duas classes.

2. Aprendizado Competitivo

Interpretação geométrica da convergência (II)

- Se cada amostra fosse constituída de três sinais de entrada $\{x_1, x_2, x_3\}$, todos os vetores de amostras e de pesos ficariam confinados numa esfera de raio unitário.
- Como exemplo, os 5 vetores de pesos do exemplo seguinte se posicionaram nos centros dos aglomerados que representam as amostras (identificadas pelo símbolo "⊗").



2. Aprendizado Competitivo

Algoritmo competitivo (Fase de operação)

- Quando uma nova amostra for apresentada à rede para propósitos de classificação, basta então verificar qual neurônio será o vencedor.
- O neurônio vencedor estará assim indicando a classe em que a amostra possuirá o maior nível de proximidade.

Início {Algoritmo Competitivo – Fase de Operação}

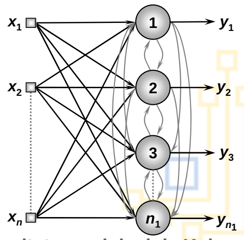
<1> Apresentar a amostra $\{ \boldsymbol{x} \}$ a ser classificada;
<2> Normalizar o vetor $\{ \boldsymbol{x} \}$ correspondente à amostra;
<3> Assumir os vetores de pesos $\{\boldsymbol{w}^{(j)}\}$ já ajustados durante a fase de treinamento;
<4> Execute as seguintes instruções:
  <4.1> Calcular as distâncias euclidianas entre $\boldsymbol{x}$ e $\boldsymbol{w}^{(j)}$ conforme a expressão (8.1);
  <4.2> Declarar como vencedor o neurônio j que contenha a menor distância euclidiana;
  <4.3> Associar a amostra àquela classe que está sendo representada pelo neurônio vencedor;
<5> Disponibilizar a classe em que a amostra foi associada.

Fim {Algoritmo Competitivo – Fase de Operação}

3. Mapas Auto-Organizáveis

Aspectos introdutórios

- Os mapas auto-organizáveis de Kohonen, também denominados de **SOM (*Self-Organization Maps*)**, são considerados uma arquitetura de RNA de estrutura reticulada, com aprendizado competitivo.
- Os mapas auto-organizáveis, em sua essência, são estruturas neurais competitivas como aquela original de **Kohonen** (figura abaixo).
- Adicionalmente, as conexões laterais, representando como a saída do neurônio vencedor influenciará os demais neurônios da rede, são fornecidas por **mapas topológicos de vizinhança**.
- O potencial dessas **conexões laterais** serão maiores (excitatórias) para os neurônios que estão mais próximos aos vencedores.
- Estes mapas topológicos informam como estarão **organizados espacialmente** os neurônios da rede frente ao comportamento de seus vizinhos
  - O mapas são normalmente formados por uma dimensão (*array*) ou duas dimensões (grade).



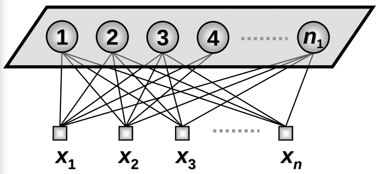
Arquitetura original de Kohonen

13

3. Mapas Auto-Organizáveis

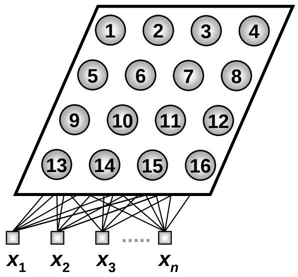
Topologias dos mapas auto-organizáveis

- A figuras seguintes mostram as topologias de mapas auto-organizáveis mais utilizadas.



Mapa Unidimensional

- O vetor de entradas {*x*}, associado às amostras, será apresentado a todos os neurônios do mapa.
- Todos os neurônios do mapa estão espacialmente organizados em uma única linha.



Mapa Bidimensional

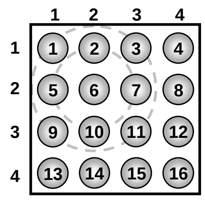
- Mapa bidimensional constituído por 16 neurônios.
- Para tal circunstância, os neurônios estão espacialmente arranjados em 4 linhas e 4 colunas.
- O número de linhas e colunas não precisam ser iguais.

14

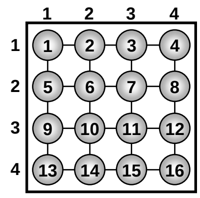
3. Mapas Auto-Organizáveis

Especificação de critérios de vizinhança

- Após a definição do arranjo espacial, deve-se então agora especificar o **critério de vizinhança** entre os neurônios.
- Um dos principais critérios de vizinhança consiste de definir um **raio *R* de abrangência**, que será utilizado pelos neurônios da rede visando definir seus respectivos vizinhos.
- Como exemplo, para o mapa topológico bidimensional anterior, tem-se



- Para um determinado neurônio *j*, seus vizinhos serão todos aqueles que estarão a uma **distância menor ou igual a *R***.
- Considerando o neurônio de número 6, tem-se que seus vizinhos, para raio de vizinhança igual a 1 (circunferência tracejada mais interna), seriam os neurônios 2, 5, 7 e 10.
- Caso se assumisse um raio de 1,75 (circunferência tracejada mais externa), sua vizinhança seria então composta pelos neurônios 1, 2, 3, 5, 7, 9, 10 e 11.



Conjuntos de vizinhança

(*R* = 1)

$$\begin{cases} \Omega_1^{(1)} = \{2,\ 5\} \\ \Omega_2^{(1)} = \{1,\ 3,\ 6\} \\ \Omega_3^{(1)} = \{2,\ 4,\ 7\} \\ \Omega_4^{(1)} = \{3,\ 8\} \\ \Omega_5^{(1)} = \{1,\ 6,\ 9\} \\ (\ldots) \end{cases}$$

$$(R = 1{,}75) \begin{cases} \Omega_1^{(1{,}75)} = \{2, 5, 6\} \\ \Omega_2^{(1{,}75)} = \{1, 3, 5, 6, 7\} \\ \Omega_3^{(1{,}75)} = \{2, 4, 6, 7, 8\} \\ \Omega_4^{(1{,}75)} = \{3, 7, 8\} \\ \Omega_5^{(1{,}75)} = \{1, 2, 6, 9, 10\} \\ (\ldots) \\ \Omega_{16}^{(1{,}75)} = \{11, 12, 15\} \end{cases}$$

15

3. Mapas Auto-Organizáveis

Regras de ajuste dos pesos dos neurônios

- Quando um neurônio *j* vence a competição frente à apresentação de uma amostra, tanto o seu vetor de pesos como de seus vizinhos serão ajustados.
- Todavia, os neurônios que estão na vizinhança do vencedor serão ajustados com **taxas menores** àquelas usadas para o ajuste do neurônio vencedor.
- Uma tática razoável, quando da utilização de raio de vizinhança unitário, tem sido assumir um ajuste para os vetores de pesos dos neurônios vizinhos ao vencedor igual à metade da taxa de aprendizagem, a qual é representada por duas regras:

$$\begin{cases} \text{Regra 1}: \boldsymbol{w}^{(v)} \leftarrow \boldsymbol{w}^{(v)} + \eta \cdot (\boldsymbol{x}^{(k)} - \boldsymbol{w}^{(v)}), \text{ para o neurônio vencedor} \\ \text{Regra 2}: \boldsymbol{w}^{(\Omega)} \leftarrow \boldsymbol{w}^{(\Omega)} + \dfrac{\eta}{2} \cdot (\boldsymbol{x}^{(k)} - \boldsymbol{w}^{(\Omega)}), \text{ para os neurônios vizinhos} \end{cases}$$

16

3. Mapas Auto-Organizáveis

Algoritmo Kohonen (Fase de treinamento)

Início {Algoritmo Kohonen – Fase de Treinamento}

<1> Definir o mapa topológico da rede;
<2> Montar os conjuntos de vizinhança $\{\Omega^{(j)}\}$;
<3> Iniciar o vetor de pesos de cada neurônio $\{\boldsymbol{w}^{(j)}\}$ considerando os valores das n_1 primeiras amostras de treinamento;
<4> Obter o conjunto de amostras de treinamento $\{\boldsymbol{x}^{(k)}\}$;
<5> Normalizar os vetores de amostras e de pesos;
<6> Especificar a taxa de aprendizagem $\{\eta\}$;
<7> Iniciar o contador de número de épocas $\{\,época \leftarrow 0\,\}$;
<8> Repetir as instruções:
  <8.1> Para todas as amostras de treinamento $\{\boldsymbol{x}^{(k)}\}$, fazer:
    <8.1.1> Calcular as distâncias euclidianas entre $\boldsymbol{x}^{(k)}$ e $\boldsymbol{w}^{(j)}$, conforme a expressão (8.1);
    <8.1.2> Declarar como vencedor o neurônio j que contenha a menor distância euclidiana:
    $$vencedor = \arg \min_j \{\| \boldsymbol{x}^{(k)} - \boldsymbol{w}^{(j)} \|\}$$
    <8.1.3> Ajustar o vetor de pesos do vencedor, conforme a regra 1 da expressão (8.4);
    <8.1.4> Ajustar o vetor de pesos dos neurônios vizinhos ao vencedor, definidos em $\Omega^{(j)}$, conforme regra 2 da expressão (8.4) ou (8.5);
    <8.1.5> Normalizar o vetor de pesos que foi ajustado na instrução anterior;
  <8.2> $época \leftarrow época + 1$;
  Até que: não haja mudanças significativas nos vetores de pesos;
<9> Analisar o mapa visando extração de características;
<10> Identificar regiões que possibilitem a definição de classes.

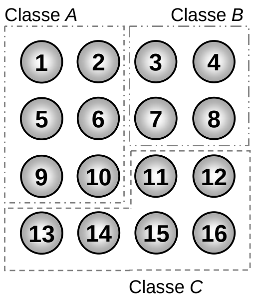
Fim {Algoritmo Kohonen – Fase de Treinamento}

- Os mapas auto-organizáveis não possuem de antemão as saídas (classes) desejadas.
- Assim, os passos <9> e <10> da fase de treinamento podem requerer o uso de ferramentas estatísticas e de conhecimento especialista, a fim de identificar as possíveis classes que agrupam dados com características em comum.
- Por conseguinte, o mapa topológico pode ser particionado em regiões que definem as respectivas classes, sendo tal arranjo denominado de **MAPA DE CONTEXTO.**

3. Mapas Auto-Organizáveis

Exemplo de aplicabilidade

- Como exemplo, mediante análises efetuadas depois da fase de treinamento, um eventual resultado possível de mapa de contexto para uma topologia bidimensional 4x4 poderia ser representada pela figura seguinte.



Mapa de Contexto

- Classe *A* ⇒ {1, 2, 5, 6, 9, 10}
- Classe *B* ⇒ {3, 4, 7, 8}
- Classe *C* ⇒ {11, 12, 13, 14, 15, 16}

- Desta forma, quando da apresentação de uma **nova amostra** a ser classificada, basta-se, em seqüência, obter quem foi o neurônio vencedor da competição.
- Após a definição do vencedor, recorre-se então ao **mapa de contexto** a fim de resgatar a classe em que o mesmo estará incumbido de representar.
- Finalmente, a amostra será então classificada como pertencente à respectiva classe daquele neurônio declarado como vencedor.

3. Mapas Auto-Organizáveis

Algoritmo Kohonen (Fase de operação)

Início {Algoritmo Kohonen – Fase de Operação}

<1> Apresentar a amostra { $\boldsymbol{x}$ } a ser classificada e normalizar;

<2> Assumir os vetores de pesos { $\boldsymbol{w}^{(j)}$ } já ajustados durante a fase de treinamento;

<3> Executar as seguintes instruções:

- <3.1> Calcular as distâncias euclidianas entre $\boldsymbol{x}$ e $\boldsymbol{w}^{(j)}$, conforme a expressão (8.1);
- <3.2> Declarar como vencedor o neurônio j que contenha a menor distância euclidiana;
- <3.3> Localizar o neurônio vencedor dentro do mapa auto-organizável;
- <3.4> Associar a amostra à classe que foi identificada a partir da confecção do mapa de contexto;

<4> Disponibilizar a eventual classe em que a amostra foi associada.

Fim {Algoritmo Kohonen – Fase de Operação}

3. Mapas Auto-Organizáveis

Aspectos de configuração

- Três aspectos para configurar um mapa auto-organizável:
  - Definição da organização espacial (*grid*) dos neurônios.
  - Delimitação dos conjuntos de vizinhança de cada neurônio.
  - Especificação do critério de ajuste do vetor de pesos do neurônio vencedor e de seus vizinhos.
- Resumo da dinâmica envolvida com o processo de treinamento:
  1) Cada neurônio da rede computa o **nível de proximidade** de seu vetor de pesos em relação a cada padrão de entrada.
  2) Um **mecanismo de competição** entre os neurônios é aplicado com o objetivo de escolher o vencedor.
  3) A partir da definição do neurônio vencedor, resgata-se, por meio do **mapa topológico**, o conjunto de vizinhança que informa quais são os neurônios vizinhos ao vencedor.
  4) Os pesos do neurônio vencedor e de seus vizinhos são incrementados com o objetivo de **aumentar o nível de proximidade** com a respectiva entrada.
  5) Após a convergência, é possível identificar **regiões do mapa** (contexto) que correspondem às classes do problema.