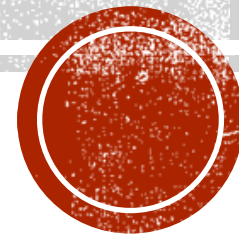


REDES NEURAIS ARTIFICIAIS

AULA 5 – REDE *PERCEPTRON MULTICAMADAS* *CLASSIFICAÇÃO DE PADRÕES*

Prof. Rodrigo Palácios

rodrigopalacios@utfpr.edu.br



PMC - PRINCIPAIS CLASSES DE PROBLEMAS

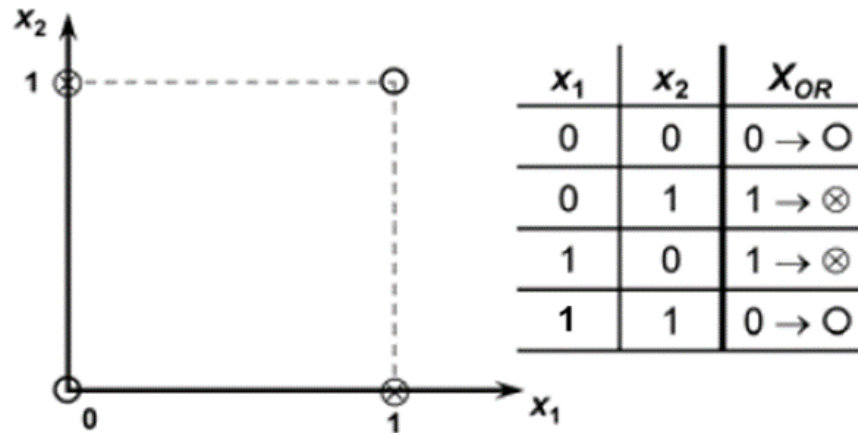
- Redes **PMC** podem ser consideradas as mais utilizadas na solução de problemas advindos das mais variadas áreas de conhecimento.
- Redes **PMC** são as mais amplamente empregadas em diferentes temáticas envolvendo as engenharias como um todo, em especial a Engenharia Elétrica.
- Existe ainda aplicações de redes **PMC** em medicina, biologia, química, física, economia, geologia, ecologia e psicologia.
- Considerando esses leques de aplicabilidades em que as redes **PMC** são passíveis de serem utilizadas, destacam-se três classes de problemas que acabam concentrando grande parte de suas aplicações, isto é:
 - **Problemas envolvendo aproximação funcional.**
 - **Problemas envolvendo classificação de padrões.**
 - **Problemas envolvendo sistemas variantes no tempo.**

PMC – PROBLEMAS DE CLASSIFICAÇÃO

- Problema de **Classificação de Padrões** consiste de associar um padrão de entrada (amostra) para uma daquelas classes que já foram previamente definidas.
 - Como exemplo, pode-se ter uma aplicação em que o **PMC** seja treinado para reconhecer sinais de vozes a fim de permitir acesso de pessoas à ambientes restritos.
 - Nesta situação, considerando-se que o treinamento foi realizado com vocábulos de apenas três pessoas, a resposta da rede, frente a um sinal de voz inserido em suas entradas, trataria então de informar de quem seria o referido sinal (dentre aquelas três pessoas).
- Diferentemente dos problemas envolvendo **Aproximação Funcional** (saídas reais/analógicas), as respostas associadas aos problemas de **Classificação de Padrões** estão sempre relacionadas com grandezas discretas (enumeráveis).
 - As situações mais elementares seriam aquelas das saídas binárias, em que se têm apenas duas classes como possíveis respostas, sendo que as mesmas poderiam estar representando, por exemplo, a “presença” ou “ausência” de determinado atributo em uma amostra apresentada à rede.
 - Como a saída da rede só fornece respostas numéricas, uma possível codificação seria assinalar o valor 0 ao atributo “presença”, ao passo que o valor 1 rotularia o atributo “ausência”.
 - Uma sistemática similar a esta poderia ser também utilizada para problemas multiclases (três ou mais classes).

PMC – PROBLEMAS DE CLASSIFICAÇÃO

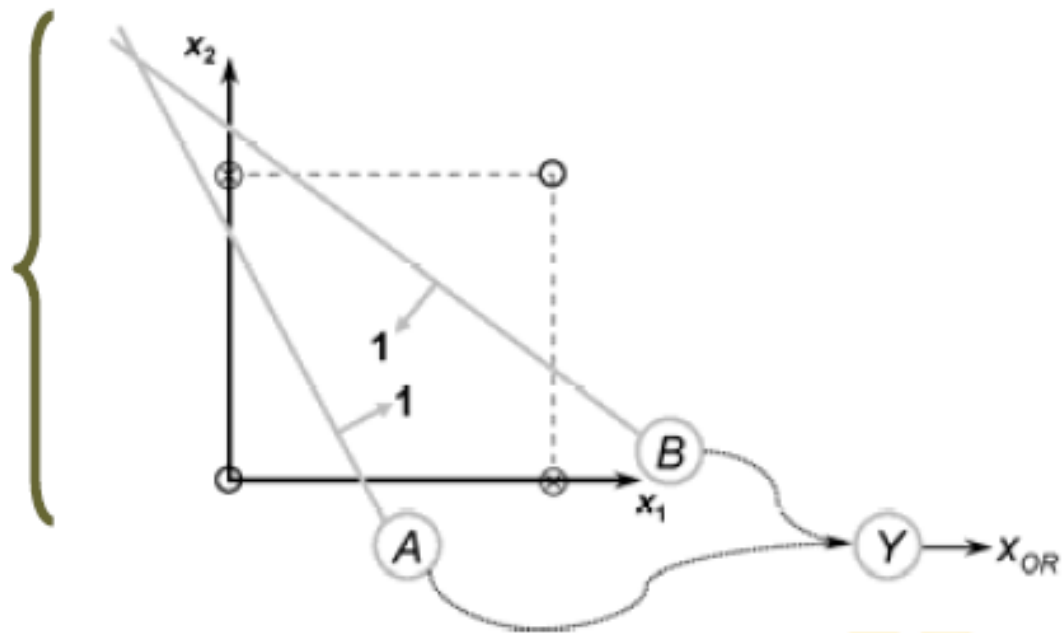
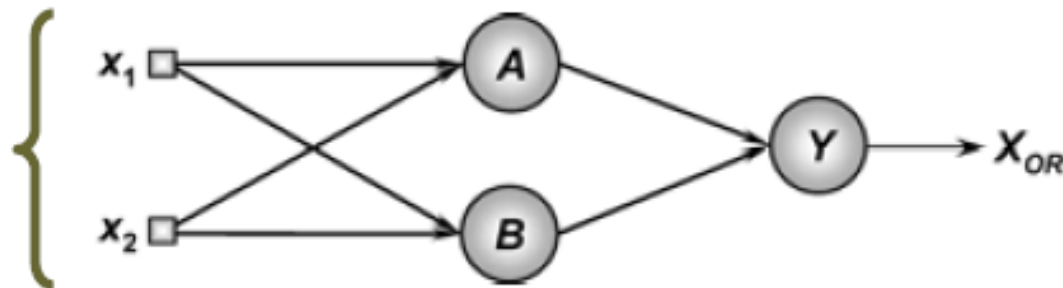
- O **Perceptron simples** (neurônio único) somente conseguiria convergir se as duas classes envolvidas com o problema fossem linearmente separáveis.
 - Caso contrário, o **Perceptron simples** jamais conseguiria convergir a fim de posicionar o seu hiperplano na faixa delimitada pela fronteira de separabilidade entre as classes.
 - Um caso clássico de tal fato é encontrado no problema do ou-exclusivo (porta *XOR*), envolvendo a lógica booleana, como ilustrado na figura seguinte:



- Fazendo uso do gráfico, constata-se que seria impossível posicionar uma única reta que permitiria separar as duas classes resultantes do problema do ou-exclusivo.
- Outras situações similares somente podem ser resolvidas por intermédio de um PMC de duas camadas neurais (uma camada escondida) ou mais.

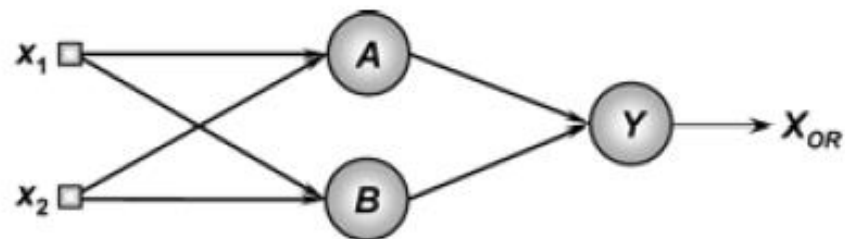
PMC – CLASSIFICAÇÃO DE PADRÕES

- Visando entender os mecanismos envolvidos em classificação de padrões por meio do **PMC**, volta-se novamente para o problema do ou-exclusivo.
- Topologia de **PMC** para resolver o problema do ou exclusivo. Nota-se aqui que o **PMC** tem apenas **Uma Camada Escondida**
- Configuração de retas (de separabilidade) que poderia ser implementada pelos neurônios *A* e *B* do **PMC**, após treinamento.

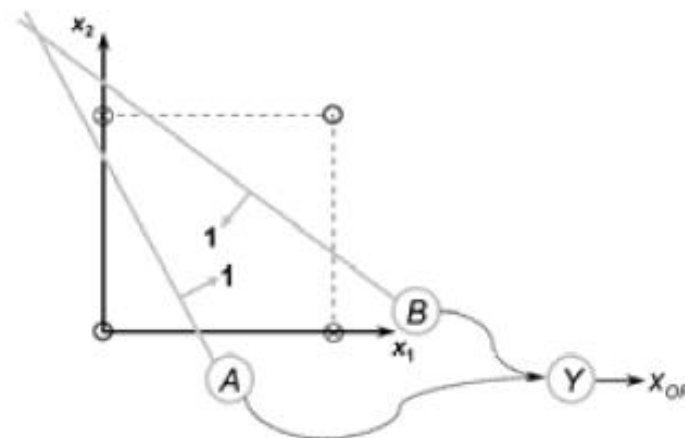


PMC – CLASSIFICAÇÃO DE PADRÕES

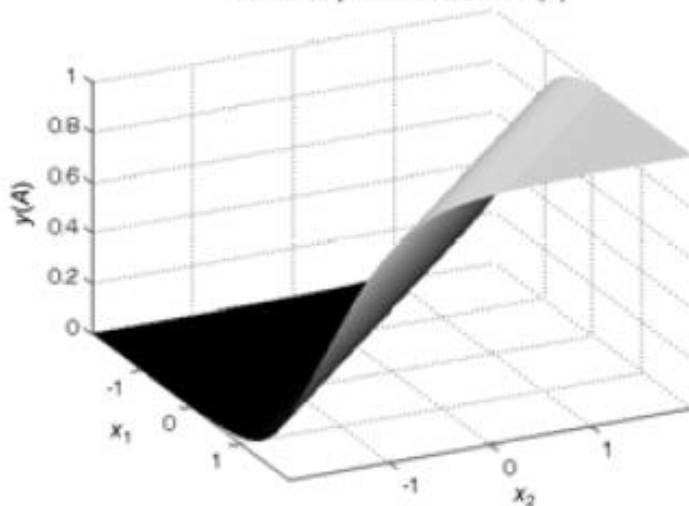
- Configuração topológica do *PMC*
(Problema do ou-exclusivo)



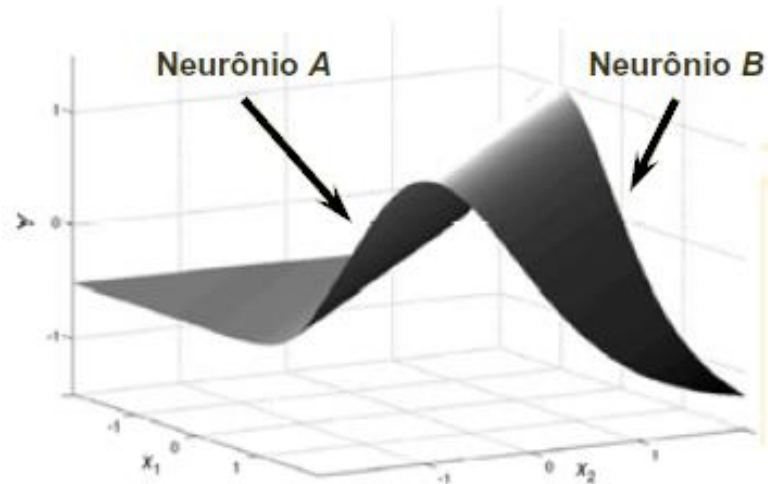
- Configuração de separabilidade



- Saída do neurônio A
Saída do primeiro neurônio (A)

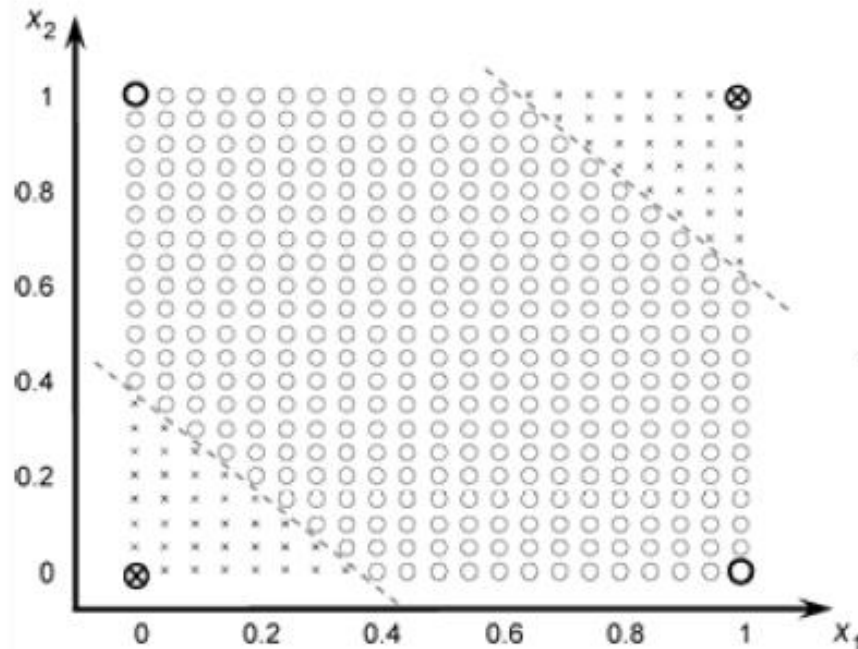
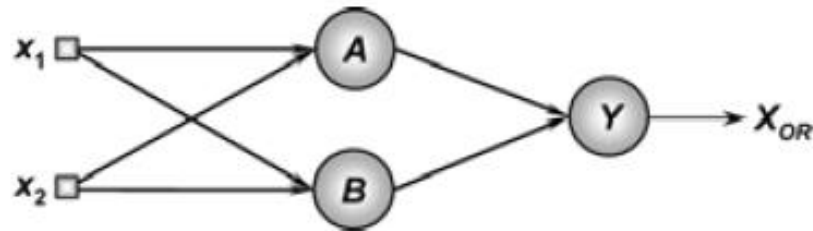


- Saída do neurônio Y
– Neurônio B inverteu a função logística
Resultado do neurônio de saída (Y)



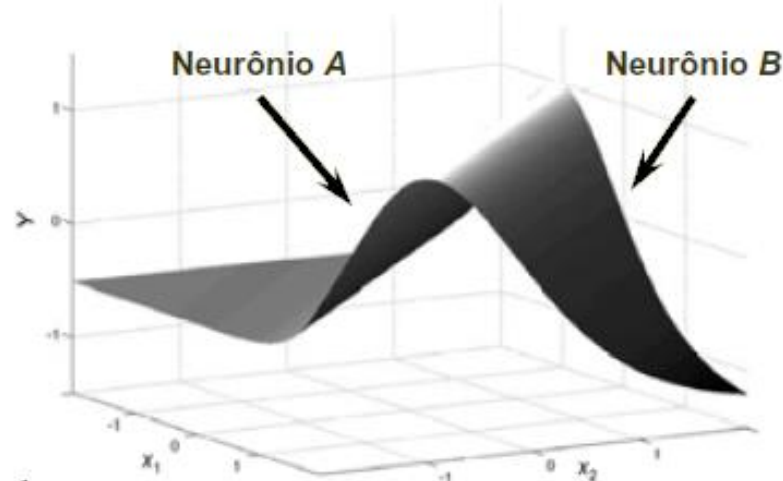
PMC – CLASSIFICAÇÃO DE PADRÕES

- Configuração topológica do *PMC*
(Problema do ou-exclusivo)



- Saída do neurônio Y

– Neurônio B inverteu a função logística
Resultado do neurônio de saída (Y)



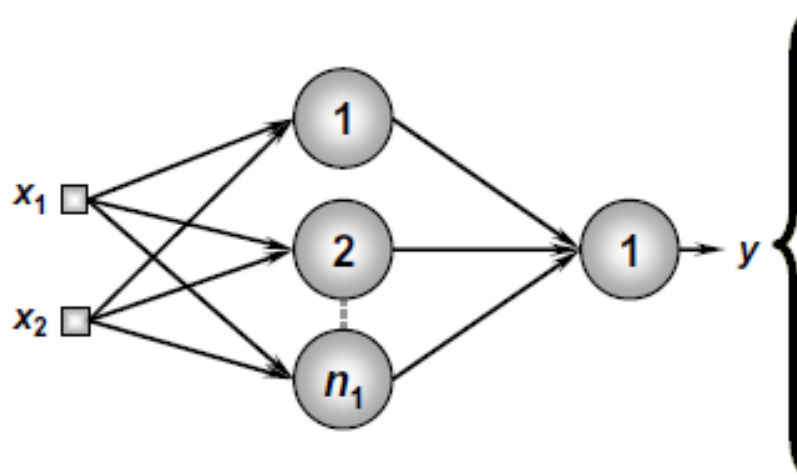
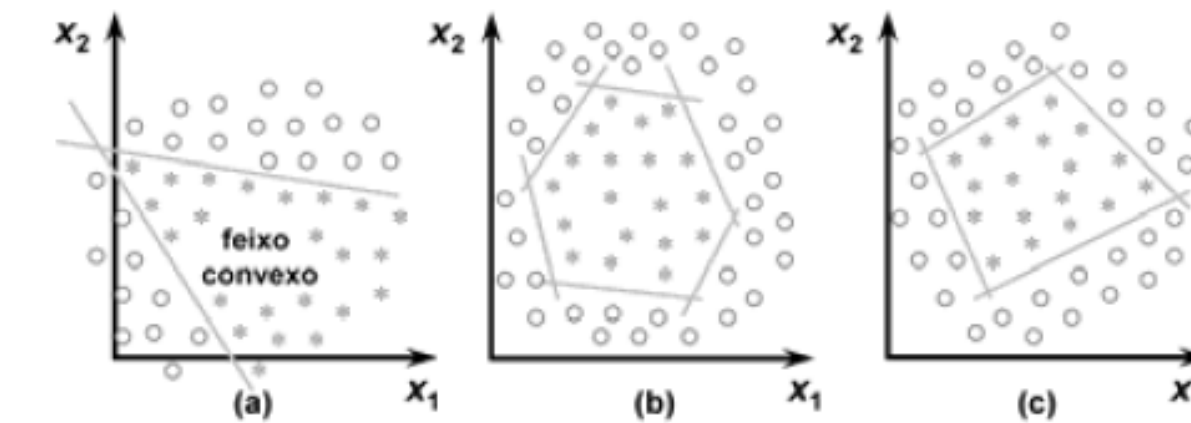
A figura mostra as fronteiras de classificação associadas ao problema do ou-exclusivo.

As fronteiras são delimitadas por duas retas, as quais são formadas quando se executa a interseção da superfície ilustrada acima (Saída do neurônio Y) próximo ao plano $y = 0$.

PMC – CLASSIFICAÇÃO DE PADRÕES - TOPOLOGIA

■ PMC de Uma Camada Escondida

- De forma similar ao problema do ou-exclusivo, pode-se então deduzir que um **PMC de Uma Camada Escondida** (duas camadas neurais) é capaz de mapear qualquer problema de classificação de padrões cujos elementos estejam dentro de uma **Região Convexa**.
- Alguns exemplos de regiões convexas são dadas a seguir.



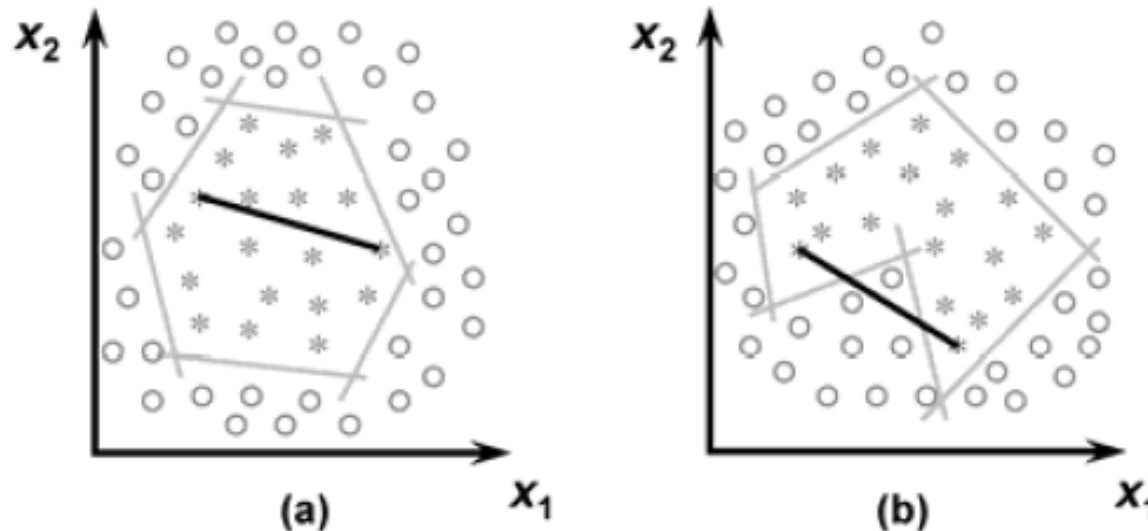
Como exemplo, visa-se aplicar este **PMC** (uma camada escondida) aos problemas acima. Para tanto, a quantidade mínima de neurônios para cada uma das situações seriam as seguintes:

- ❖ Exemplo (a) → 2 neurônios ($n_1 = 2$)
- ❖ Exemplo (b) → 6 neurônios ($n_1 = 6$)
- ❖ Exemplo (c) → 4 neurônios ($n_1 = 4$)

PMC – CLASSIFICAÇÃO DE PADRÕES - TOPOLOGIA

■ Conceito de Região Convexa

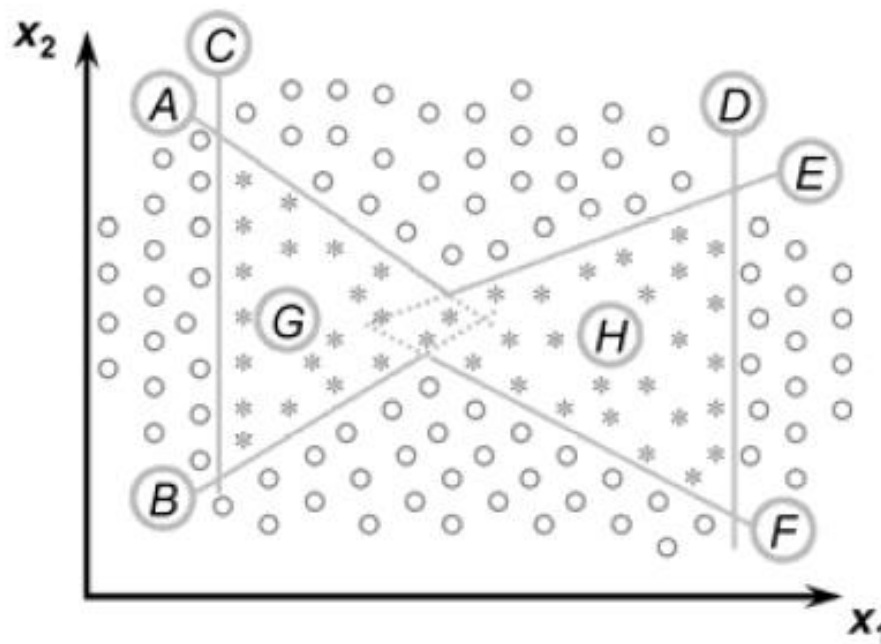
- **Definição:** Do ponto de vista geométrico, uma região é considerada **Convexa** se, e somente se, todos os pontos contidos em quaisquer segmentos de reta, os quais estão também definidos por quaisquer dois pontos delimitados pelo respectivo domínio, estiverem ainda dentro da mesma.
- As figuras seguintes mostram ilustrações de região convexa (a) e região não-convexa (b).



PMC – CLASSIFICAÇÃO DE PADRÕES - TOPOLOGIA

■ PMC de Duas Camadas Escondidas

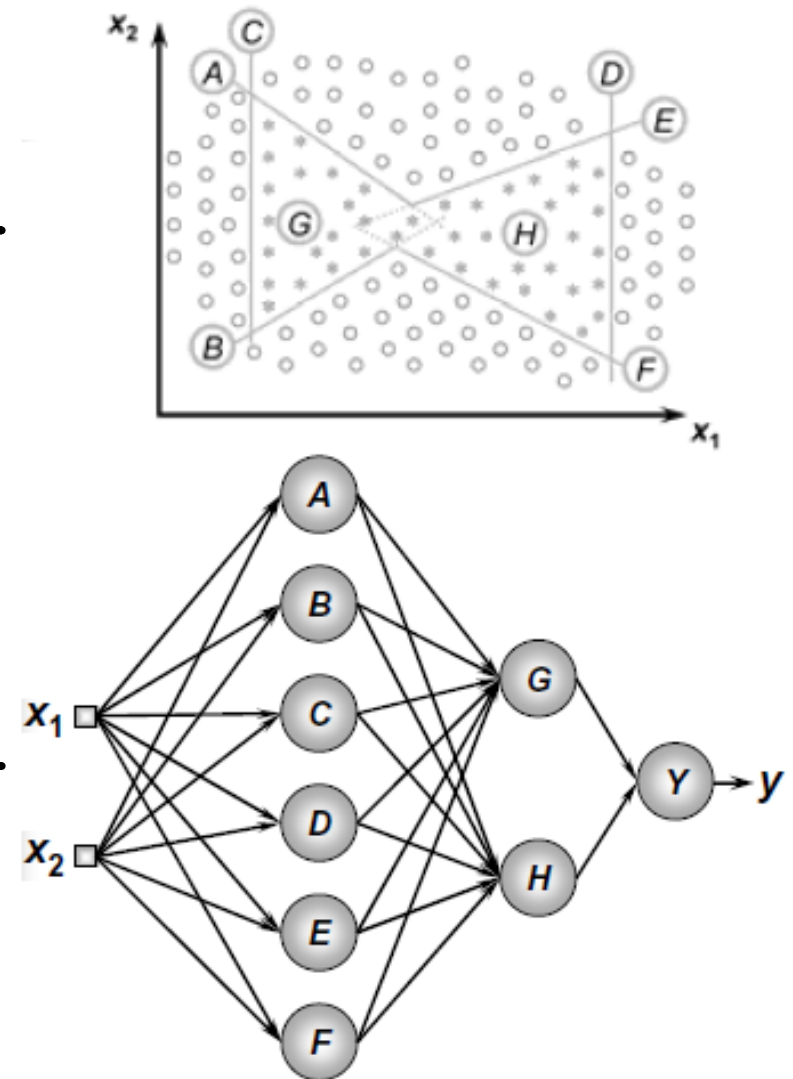
- Do slide anterior, conclui-se que **PMC de uma Camada Escondida** consegue classificar padrões dispostos em **Regiões Convexas**.
- Consequentemente, pode-se também deduzir que redes **PMC de Duas Camadas Escondidas** são capazes de classificar padrões que estejam em quaisquer tipos de regiões geométricas.
- Para tanto, a figura seguinte elucida tal afirmativa.



PMC – CLASSIFICAÇÃO DE PADRÕES - TOPOLOGIA

■ PMC de Duas Camadas Escondidas (Exemplo)

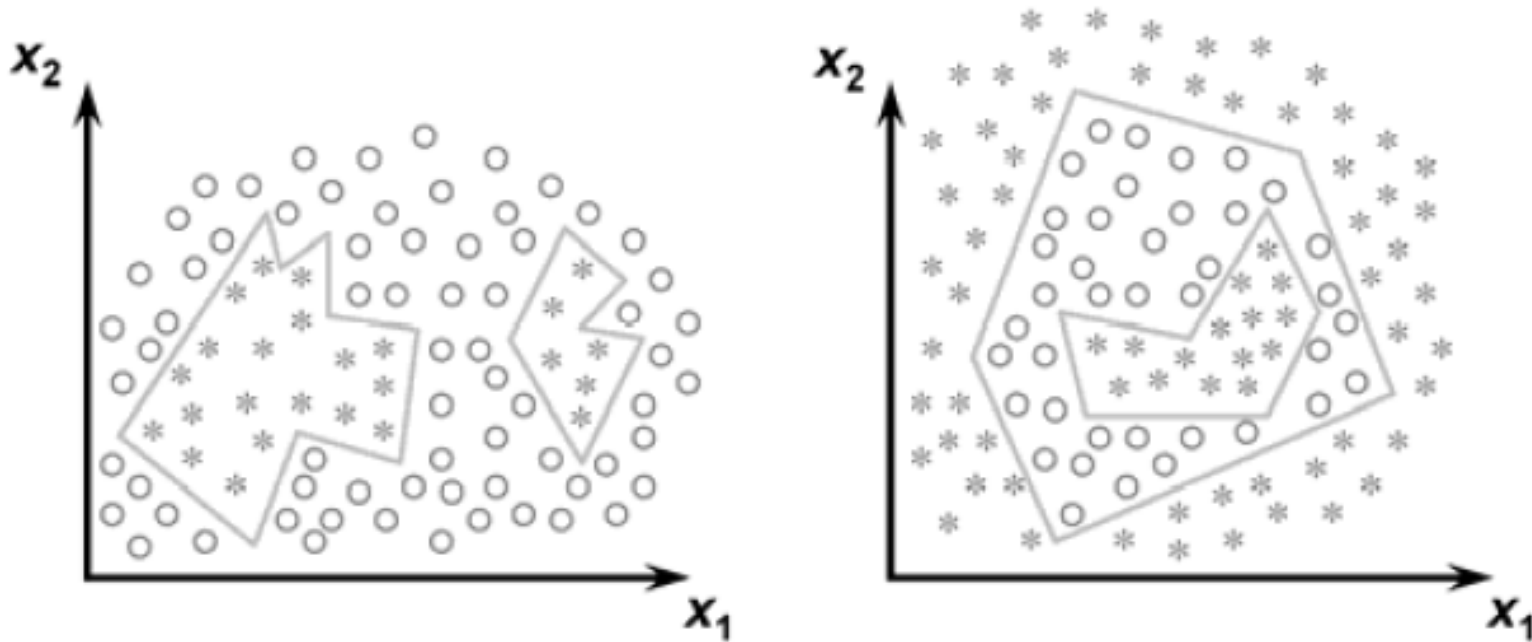
- Neste caso, a região não-convexa acaba sendo formada por duas **Regiões Convexas**.
- **1ª Região Convexa (1ª RC):** Formada pelos segmentos de reta A , B e C , compondo o triângulo da esquerda.
- **2ª Região Convexa (2ª RC):** Formada pelos segmentos de reta D , E e F , compondo o triângulo da direita.
- Nesta situação, tem-se ações específicas para os neurônios das camadas do **PMC**.
- Neurônios da 1ª Camada Escondida:
 - Os neurônios A , B e C são responsáveis pela delimitação da 1ª RC (triângulo da esquerda).
 - Os neurônios D , E e F são responsáveis pela delimitação da 2ª RC (triângulo da direita).
- Neurônios da 2ª Camada Escondida:
 - Neurônio G combina saídas dos neurônios A , B e C , a fim de enquadrar a parcela das amostras que pertenceriam à 1ª RC, sendo que nesta condição a sua saída seria igual a 1.
 - Neurônio H combina saídas dos neurônios D , E e F , a fim de enquadrar a 2ª RC, produzindo então valor 1.
- Neurônio de Saída:
 - Neurônio Y estaria então incumbido de efetuar uma operação de disjunção booleana (porta OR), pois se qualquer uma das saídas produzidas pelos neurônios G ou H for igual a 1, sua resposta final y seria também igual a 1.



PMC – CLASSIFICAÇÃO DE PADRÕES - TOPOLOGIA

- **PMC de Duas Camadas Escondidas (Disjuntos)**

- Adicionalmente, o **PMC** com duas camadas escondidas poderia ainda mapear outros tipos de regiões geométricas, tais como aquelas formadas por conjuntos disjuntos (desconexos), conforme ilustradas nas figuras seguintes.



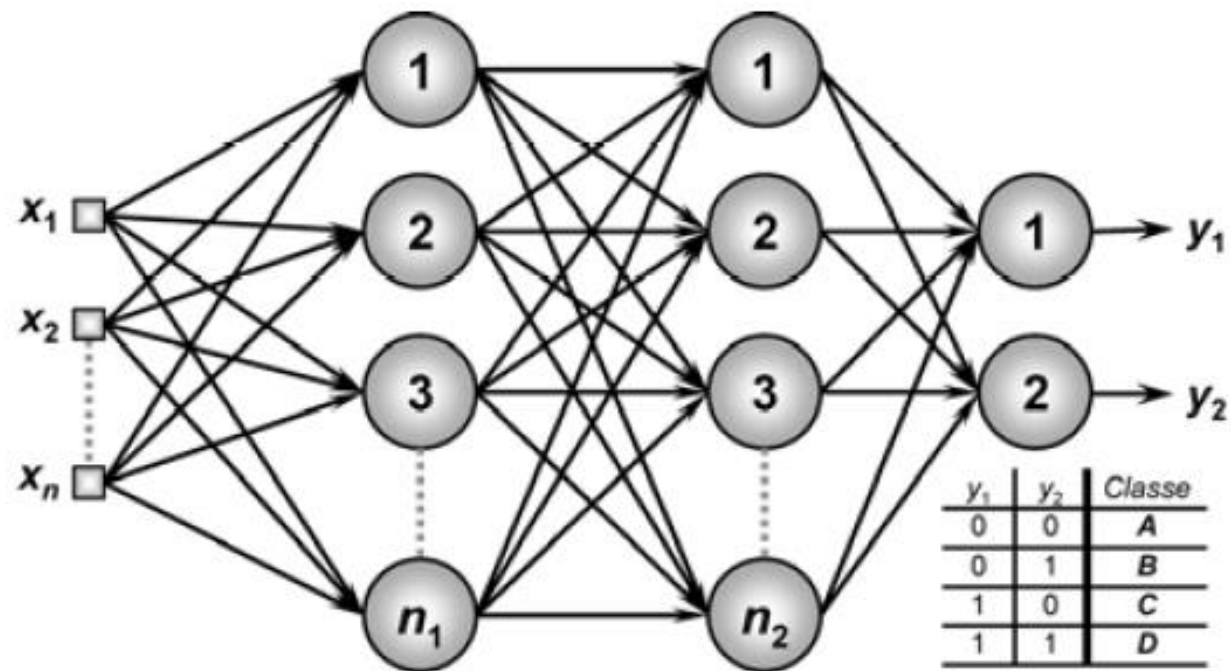
PMC - CLASSIFICAÇÃO DE PADRÕES - CODIFICAÇÃO DE CLASSES

- **Aspectos de atribuição de classes - Codificação em Problemas Multiclasses**
 - Para problemas de classificação de padrões **com mais de duas classes**, há então a necessidade de se inserir mais neurônios na camada de saída da rede.
 - **PMC** com apenas **Um** neurônio em sua camada de saída é capaz de distinguir somente duas classes.
 - **PMC** composto de **Dois** neurônios em sua camada de saída poderia representar, no máximo, quatro classes possíveis.
 - **PMC** composto de **Três** neurônios em sua camada de saída poderia diferenciar, no máximo, oito classes no total.
 - Generalizando, **PMC** com m neurônios em sua camada de saída seria capaz de classificar, teoricamente, até 2^m classes possíveis.

PMC – CLASSIFICAÇÃO DE PADRÕES – CODIFICAÇÃO DE CLASSES

- **Método sequencial para atribuição de classes - Codificação Multiclasses (Sequencial)**

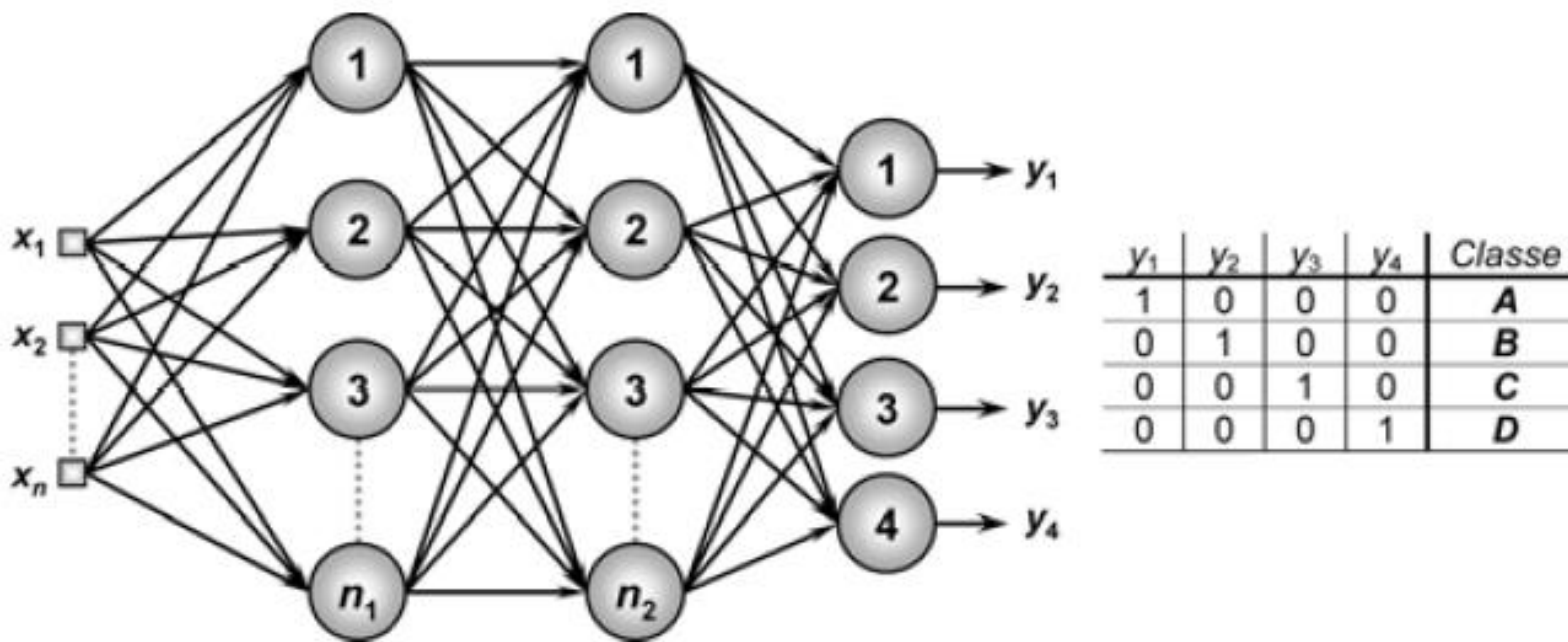
- A figura seguinte mostra um exemplo de codificação sequencial para um **PMC** com duas saídas.
- Em termos práticos, devido à eventual complexidade do problema, a adoção desta codificação pode tornar o treinamento bem mais difícil.
- Neste caso, as classes estariam sendo representadas por pontos que estão espacialmente bem próximos entre si.
- Tal situação poderia então demandar um incremento substancial no número de neurônios de suas camadas intermediárias.



PMC – CLASSIFICAÇÃO DE PADRÕES – CODIFICAÇÃO DE CLASSES

- **Codificação Multiclasses (Método *One of C-Class*)**

- É um dos métodos mais utilizados para codificação multiclasses.
- Consiste de associar a saída de cada neurônio diretamente à classe correspondente.
- Neste caso, a quantidade de neurônios na camada de saída do **PMC** será igual ao número de classes do problema.
- Como exemplo, se houver 4 classes possíveis, as saídas do **PMC** seria também composta de 4 neurônios, conforme figura seguinte:



PMC – PÓS-PROCESSAMENTO DE DADOS

- **Critérios de arredondamento de saída - Processo de Pós-Processamento**

- Os valores produzidos pelos neurônios da camada de saída da rede **PMC** são números reais.
- Considerando os problemas de classificação de padrões, as respostas devem ser então pós-processadas, pois o que importa são aqueles valores discretos usados para rotular as classes.
- Para o caso da função logística, estes valores podem estar próximos de 1 ou próximos de 0.
- Face a esta circunstância, dependendo da precisão requerida, os valores advindos desta operação podem ser obtidos pela aplicação da seguinte sistemática:

$$Y_i^{pós} = \begin{cases} 1, & \text{se } Y_i^{saída} \geq \lim^{sup} \\ 0, & \text{se } Y_j^{saída} \leq \lim^{inf} \end{cases}$$

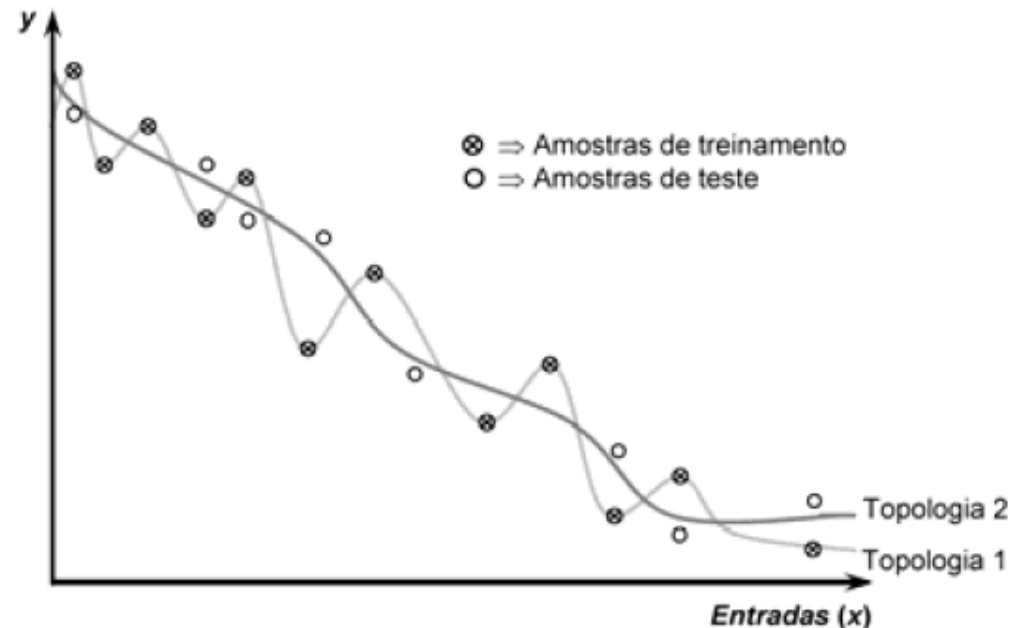
- Os valores tipicamente adotados para os limitantes (\lim^{sup} e \lim^{inf}) são definidos por:

$$\begin{cases} \lim^{sup} \in [0,5 \quad 0,9] \\ \lim^{inf} \in [0,1 \quad 0,5] \end{cases}$$

OVERFITTING

■ Aspectos de ocorrência de *overfitting* (*sobre-treinamento*)

- O aumento indiscriminado de neurônios, assim como de camadas intermediárias, não assegura a generalização apropriada do **PMC** frente às amostras pertencentes aos subconjuntos de teste.
- Esse aumento indiscriminado tende a levar a saída do **PMC** para a circunstância de memorização excessiva (***overfitting***), em que o mesmo acaba decorando as suas respostas frente aos estímulos introduzidos em suas entradas. Aqui, verifica-se os seguintes aspectos:
 - Durante a fase de aprendizado → Erro quadrático tende a ser bem baixo.
 - Durante a fase de teste (**generalização**) → Erro quadrático já tende a assumir valores bem elevados frente às amostras do subconjunto de teste.

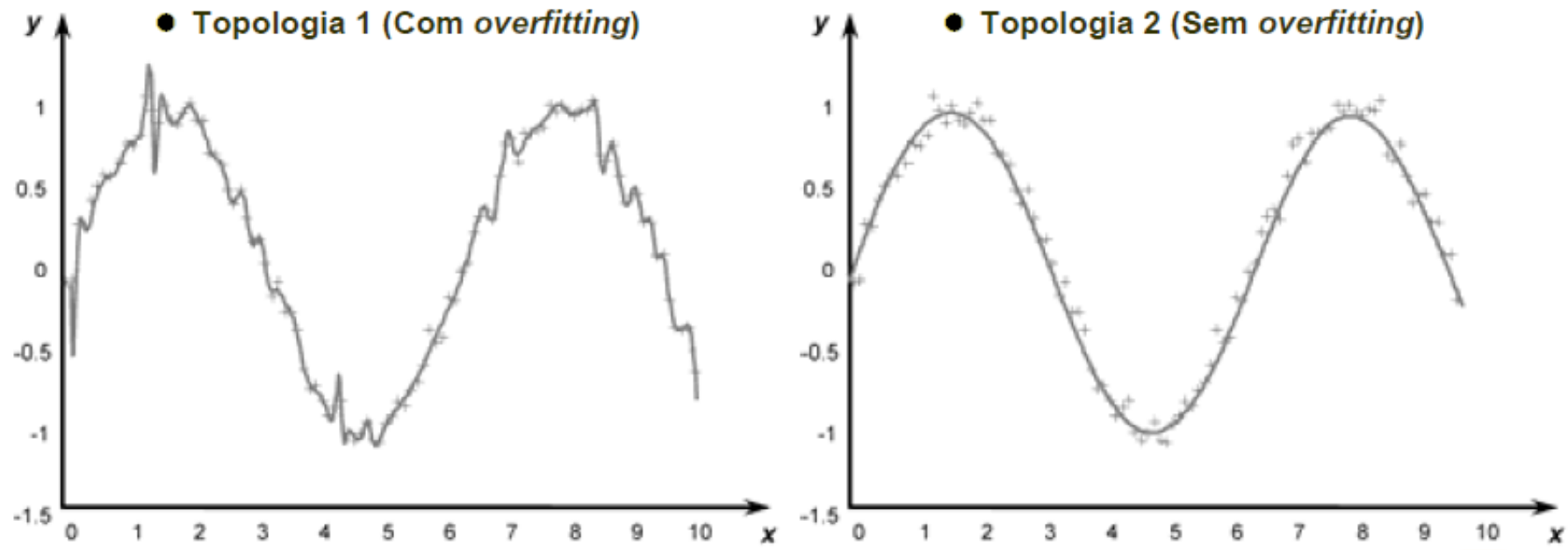


- **Topologia 1 (Com *overfitting*)**
 - Composta de uma camada escondida.
 - 20 neurônios nesta camada.
 - Apresenta menor Erro frente às amostras de treinamento.
 - Apresenta maior Erro frente às amostras de teste.
- **Topologia 2 (Sem *overfitting*)**
 - Composta de uma camada escondida.
 - 10 neurônios nesta camada.
 - Apresenta maior Erro frente às amostras de treinamento.
 - Apresenta menor Erro frente às amostras de teste.

OVERFITTING

- **Ilustração de ocorrência de *overfitting* (sobre-treinamento)**

- Mapeamento da função seno (que foi afetada por ruídos)



- **Aspectos de ocorrência de *underfitting* (sub-treinamento):**

- Em contrapartida, frente à precisão requerida, uma topologia de **PMC** com número muito reduzido de neurônios pode ser insuficiente para a extração e armazenamento de características que permitam à rede implementar as hipóteses a respeito do comportamento do processo.
- Nesses casos, por sua vez, o erro quadrático tanto na fase de aprendizado como na fase de teste serão bem significativos.



REFERÊNCIA

- SILVA, Ivan Nunes da e SPATTI, Danilo Hernane e FLAUZINO, Rogério Andrade. Redes neurais artificiais para engenharia e ciências aplicadas. . São Paulo: Artliber Editora. . Acesso em: 23 dez. 2023. , 2010