

**Aulas remotas de PSI3471-2020
com temáticas programadas para
as semana de 23/03 e 25/03
Prof. Emilio Del Moral Hernandez/**

Temas da quarta semana c/ Prof Emilio

#7 (23/março – 2ªf) Foco da semana: Reconhecedores – Circuitos lógicos genéricos e suas implementações através de computação neural; relação com discriminadores. Construindo reconhecedores de padrões com fronteiras de separação genéricas: separadores lineares; bancos de separadores lineares como aproximadores efetivos de fronteiras genéricas.

#8 (25/março – 4ªf) ... Técnicas de medida de qualidade em reconhecedores de padrões: matriz de confusão, conceitos de falsos positivo e negativo, especificidade e sensibilidade. Curvas ROC.

... nestes slides: primeira destas 2 aulas (#7 -23/03)

14

Reconhecimento de padrões

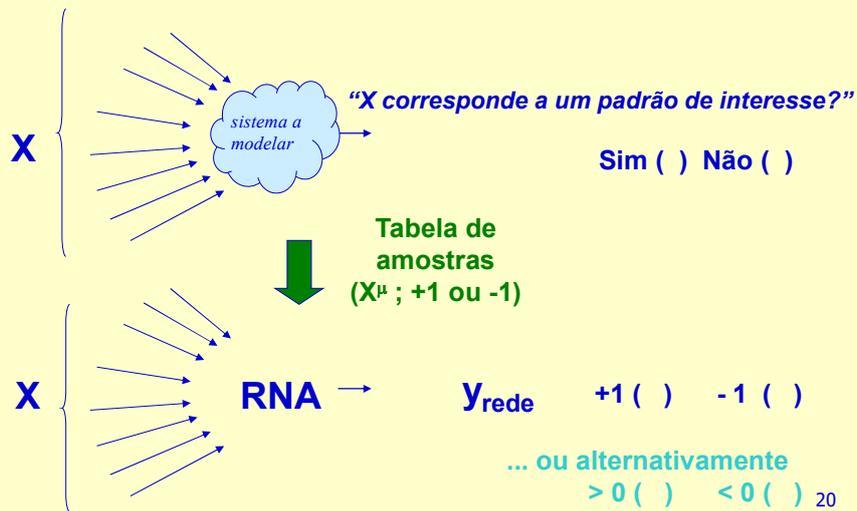
- A modelagem / emulação pela RNA pretende capturar, de maneira automática, a regularidade quantitativa supostamente existente entre os valores numéricos das variáveis componentes de X (medidas) e a pertinência às classes complementares $+1$ e -1 , regularidade essa expressa no dados empíricos disponíveis (O conjunto de casos (X,y))
- Os elementos do Conjunto de treinamento $\{(X^\mu; y^\mu_{\text{classe binária}})\}$, nos problemas de reconhecimento / detecção de padrões, serão sempre do tipo $(X ; +1)$ ou $(X ; -1)$
- O conjunto de treinamento ... $\{(X^\mu; y^\mu)\}$, tem M elementos
- O modelo obtido generalizará bem quando for usado na prática para prever “ y ” em novas situações? ... Conjunto de treinamento *versus* Conjunto de testes

19

© Emilio Del Moral Hernandez

19

RNAs como reconhecedor de padrões ...



© Emilio Del Moral Hernandez

20

Exemplo simples de Reconhecimento de padrões usando o MLP

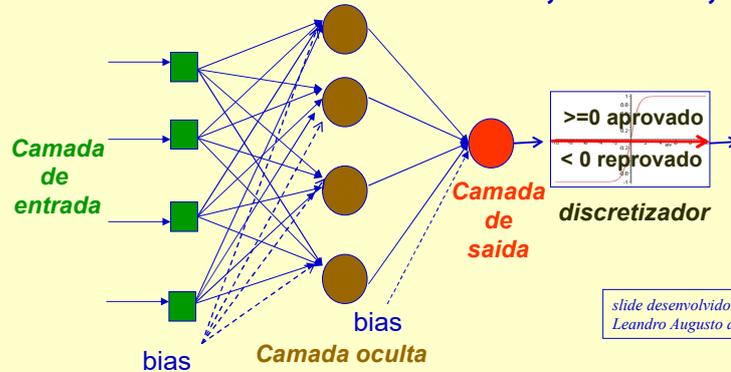
- Padrões para reconhecimento de bons pagadores, para concessão de crédito pessoal (crédito aprovado) ou não concessão de crédito pessoal (crédito reprovado).
- Quatro possíveis variáveis de entrada - x's:
 - Valor Credito Solicitado,
 - Saldo Conta Corrente,
 - Valor médio dos bens e
 - Renda Declarada.
- Problema: projetar uma MLP para reconhecer bons pagadores para concessão de crédito.

21

© Emilio Del Moral Hernandez

21

**Exemplo de Arquitetura MLP (ilustrativo).
Várias decisões ... N. de entradas, de nós, η ...**



Neurônios de Entrada	4	}	16 pesos + 4 bias
Neurônios na Camada Escondida	4		
Neurônio de Saída	1	}	4 pesos + 1 bias
Discretizador		}	>=0 crédito aprovado < 0 crédito reprovado

© Emilio Del Moral Hernandez

22

Representação gráfica muito usada em livros e artigos para a discussão de estratégias algorítmicas, cenários e desafios diversos, no universo de reconhecimento de padrões e discriminação / separação deles em 2 classes (ou mais)

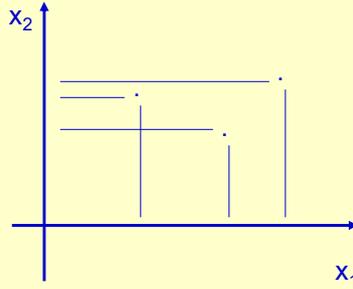
33

© Emilio Del Moral Hernandez

33

Representação de padrões / vetores a reconhecer como pontos num espaço de variáveis observadas aqui x_1 e x_2 apenas

- três padrões ... representados por três pontos no espaço bi-dimensional: X tem apenas 2 dimensões



... acima, 3 padrões bidimensionais = sinônimo de 3 vetores X^u

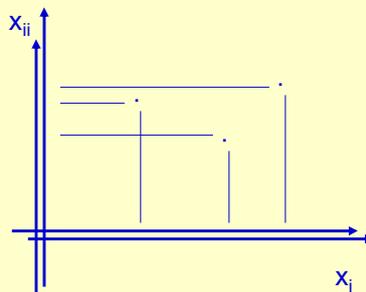
35

© Emilio Del Moral Hernandez

35

Lembrando que os padrões tX também podem ter mais que 2 dimensões $x_1, x_2, x_3, x_4, \text{etc}$

- três padrões ... representados por três pontos no espaço ~~bi-dimensional~~ ... multi-dimensional (dos X)



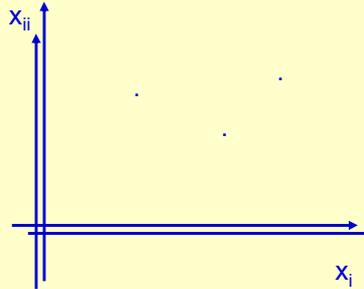
36

© Emilio Del Moral Hernandez

36

O conceito de padrões / vetores a reconhecer como pontos num espaço de variáveis observadas x_1, x_2, \dots e também x_3, x_4, \dots etc

- três padrões ... representados por três pontos no espaço **bi-dimensional** ... multi-dimensional (dos X)



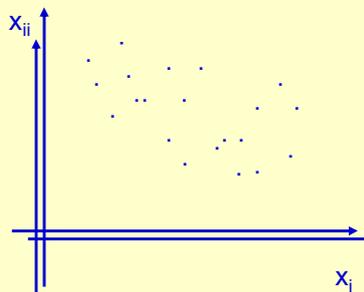
37

© Emilio Del Moral Hernandez

37

O conceito de padrões / vetores a reconhecer como pontos num espaço de variáveis observadas $x_1, x_2, x_3, x_4, \dots$ etc

- uma população com grande número de padrões



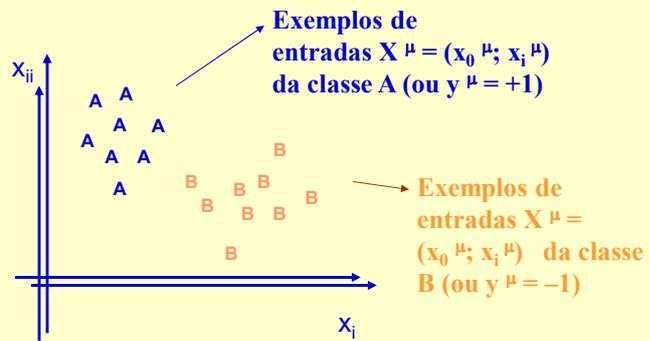
.... aqui, temos muitos padrões = muitos vetores X^p

38

© Emilio Del Moral Hernandez

38

O conceito de representação das classes a que pertence cada vetor de treino ($X^\mu; y^\mu$)



*Junto a cada padrão X^μ , podemos ter seu rótulo / label:
A versus B, ou algum homólogo numérico, como +1 versus -1*

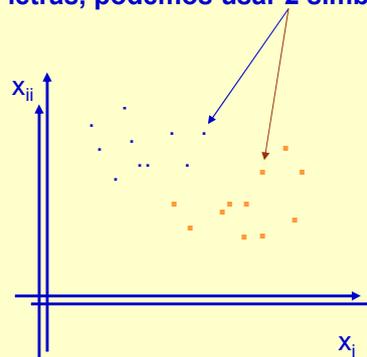
39

© Emilio Del Moral Hernandez

39

O conceito de representação das classes a que pertence cada vetor de treino ($X^\mu; y^\mu$)

- ... em vez de 2 letras, podemos usar 2 símbolos distintos



... podemos também ter o rótulo / label representado mais compactamente por símbolos pequenos ou mesmo por cores

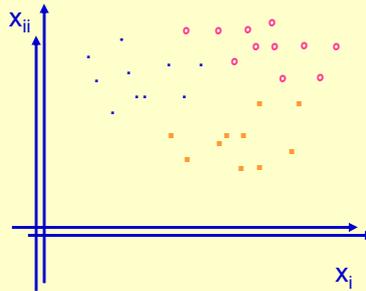
40

© Emilio Del Moral Hernandez

40

O conceito de representação das classes a que pertence cada vetor de treino ($X^{\mu}; y^{\mu}$)

- ... ou 3 (ou mais) símbolos, se tivermos 3 (ou mais) classes



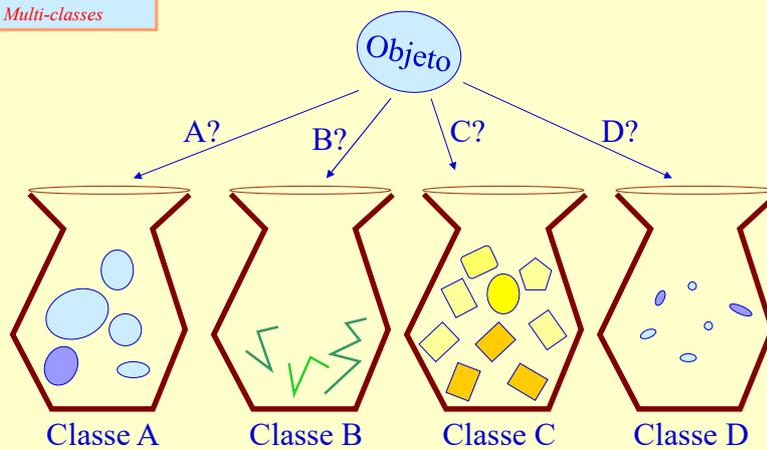
41

© Emilio Del Moral Hernandez

41

Uma metáfora para a tarefa de classificação genérica ...

*Lembrando que já
Falamos antes em
Multi-classes*



42

© Emilio Del Moral Hernandez

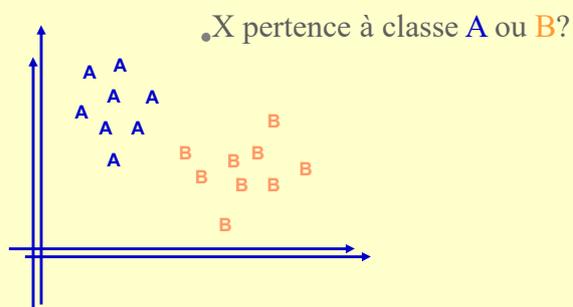
42

Agora fica bastante fácil falarmos do conceito de **fronteira de separação** entre duas classes A e B

(conceito já antecipado brevemente em aulas anteriores)

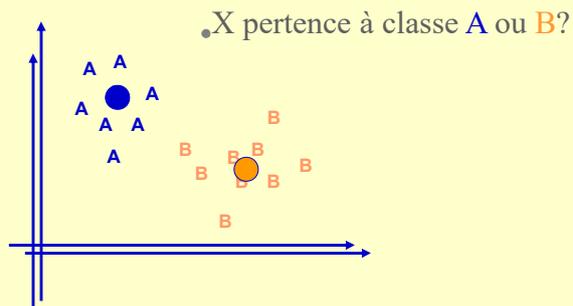
Uma estratégia de classificação automática (não neural) simples (e com falhas em alguns casos, como veremos) ...

... Decisão entre A e B por distância menor aos dois centróides dos conjuntos de exemplares de treino das classes A e B



Uma estratégia de classificação automática (não neural) simples (e com falhas em alguns casos, como veremos) ...

... Decisão entre A e B por distância menor aos dois centróides dos conjuntos de exemplares de treino das classes A e B

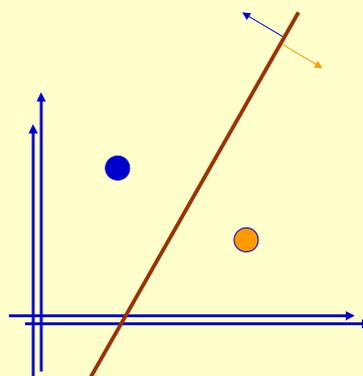


48

© Emilio Del Moral Hernandez

48

Um caso em que o método dos centróides não funciona bem, mesmo havendo separabilidade linear de classes ...

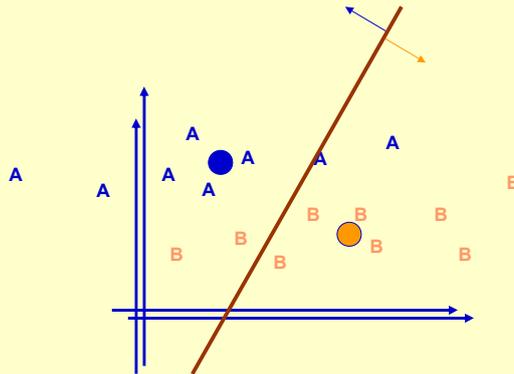


49

© Emilio Del Moral Hernandez

49

Um caso em que o método dos centróides não funciona bem, mesmo havendo separabilidade linear de classes ...

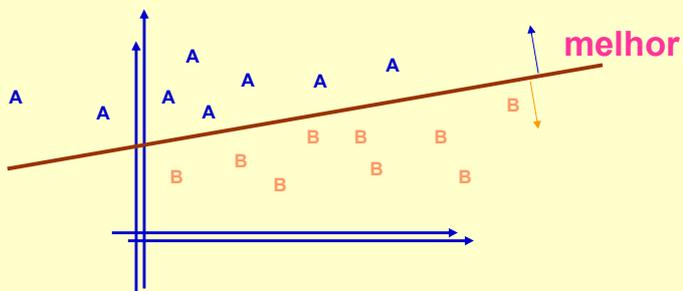


50

© Emilio Del Moral Hernandez

50

Neste caso, **melhor usarmos outra fronteira linear**, não aquela definida por equidistância aos centróides (... é o discriminador linear / neural de Rosemblat!)



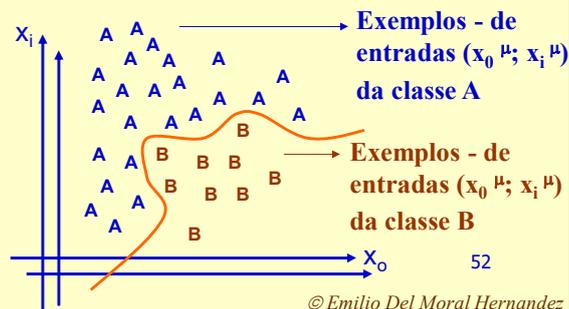
51

© Emilio Del Moral Hernandez

51

Em cenários mais complicados (e com “exigência” do máximo desempenho) ...

- Não há separabilidade linear (através de hiperplano)
- A linha / ou a hipersuperfície de separação tem que ser algo mais elaborado que um hiperplano



52

53

Começemos pelo entendimento da matemática das fronteiras mais simples ...

a) Primeiro, a fronteira linear sem viés (fronteira passa pela origem do espaço dos padrões X), usando um equacionamento orientado por vetores / álgebra linear

b) Depois, a fronteira linear com viés (fronteira “fora / longe” da origem), usando novamente vetores / álgebra linear (e também usando o neurônio isolado de Roseblatt)

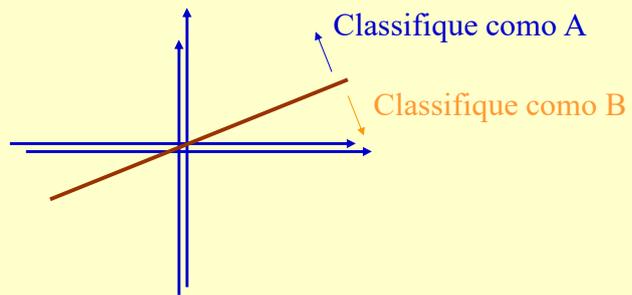
c) Finalmente, a não linear com qualquer formato

© Prof. Emilio Del Moral Hernandez

53

53

Primeiro estudemos o poder de separação de um hiperplano sem viés (ele “passa na origem”)

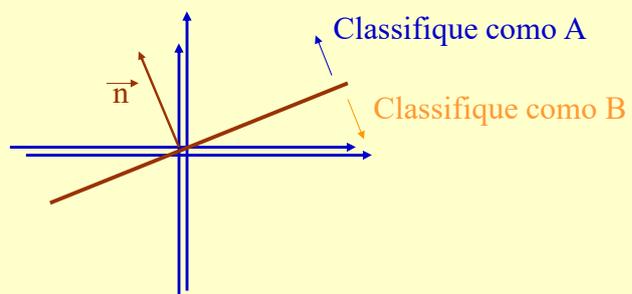


55

© Prof. Emilio Del Moral Hernandez

55

Primeiro estudemos o poder de separação de um hiperplano sem viés (ele “passa na origem”)

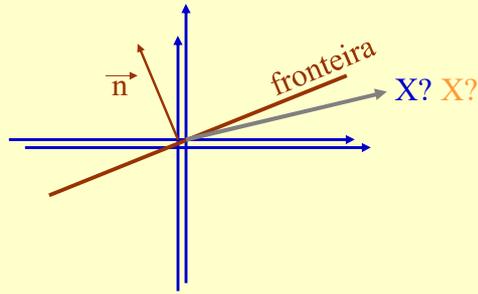


56

© Prof. Emilio Del Moral Hernandez

56

Primeiro estudemos o poder de separação de um hiperplano sem viés (ele “passa na origem”)

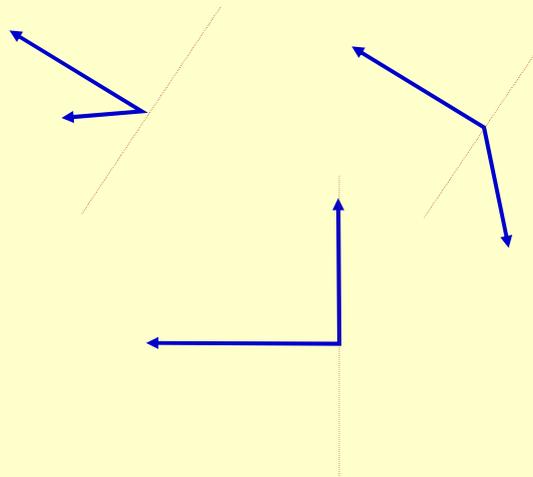


57

© Prof. Emilio Del Moral Hernandez

57

Aparte: interpretação geométrica do produto escalar em 2D e 3D: valores positivos indicam alinhamento dos dois vetores; negativos, o anti-alinhamento; valores nulos, indicam a ortogonalidade

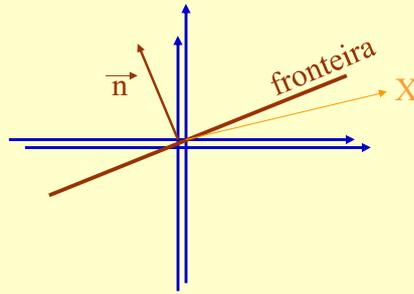


58

© Prof. Emilio Del Moral Hernandez

58

Primeiro estudemos o poder de separação de um hiperplano sem viés (ele “passa na origem”)



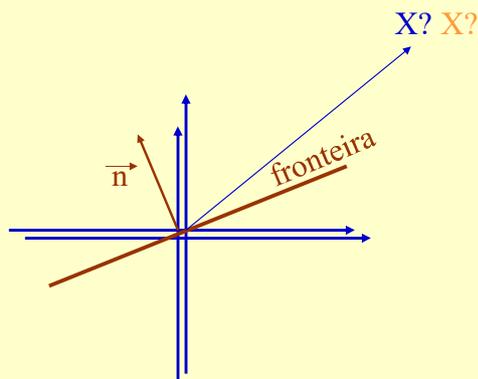
... neste caso, $X \cdot N < 0$: esse sinal negativo no valor do produto escalar $X \cdot N$ indica que X está no “semi-espaço oposto” à direção do vetor normal ao plano.

59

© Prof. Emilio Del Moral Hernandez

59

Primeiro estudemos o poder de separação de um hiperplano sem viés (ele “passa na origem”)

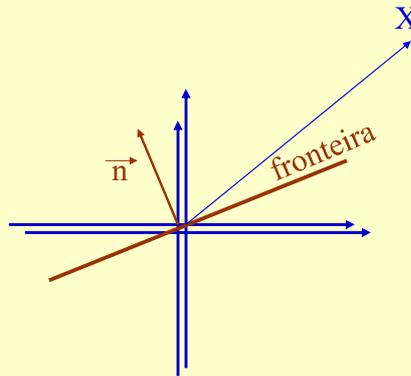


60

© Prof. Emilio Del Moral Hernandez

60

Primeiro estudemos o poder de separação de um hiperplano sem viés (ele “passa na origem”)



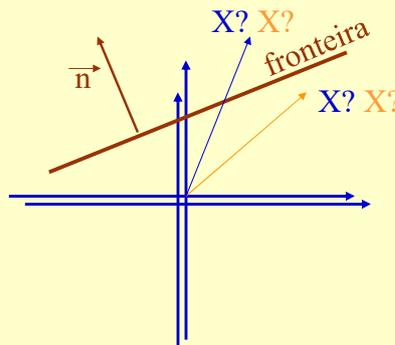
... neste caso, $X \cdot N > 0$: esse sinal positivo no valor do produto escalar $X \cdot N$ indica que X está no “semi-espaco alinhado” à direção do vetor normal ao plano.

61

© Prof. Emilio Del Moral Hernandez

61

Agora, o poder de separação de um hiperplano COM viés (ele não necessariamente “passa na origem”)



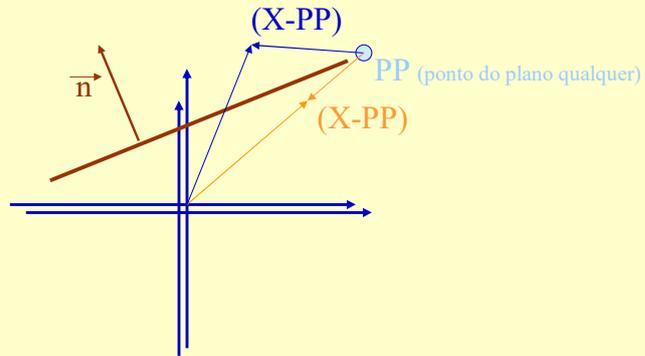
Agora, infelizmente temos $X \cdot N > 0$ nos dois casos de X!!!
A regra de separação anterior não nos serve mais!

62

© Prof. Emilio Del Moral Hernandez

62

Agora, foquemos no maior poder do hiperplano COM viés (ele não necessariamente “passa na origem”)

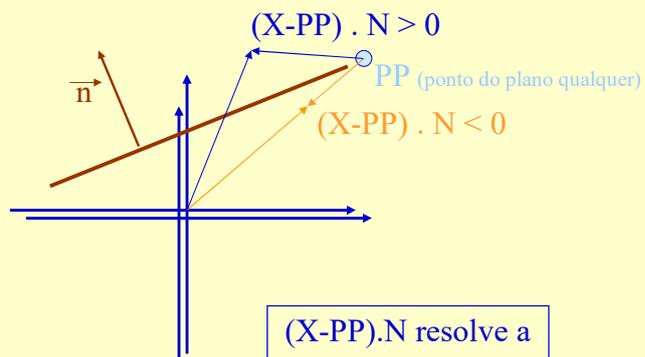


63

© Prof. Emilio Del Moral Hernandez

63

Agora, foquemos no maior poder do hiperplano COM viés (ele não necessariamente “passa na origem”)



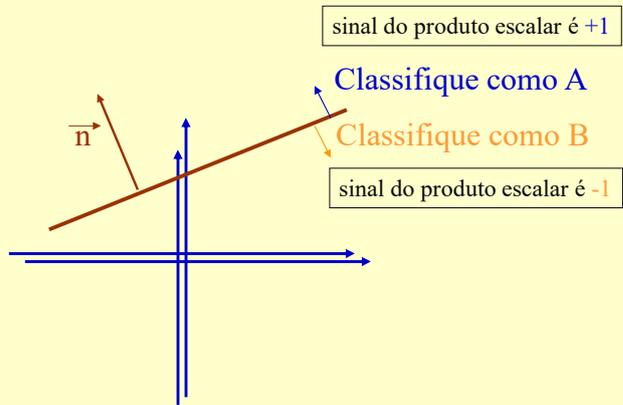
$(X-PP) \cdot N$ resolve a decisão!
 Nota: PP pode ser qualquer ponto do plano

64

© Prof. Emilio Del Moral Hernandez

64

Agora, foquemos no maior poder do hiperplano COM viés (ele não necessariamente “passa na origem”)

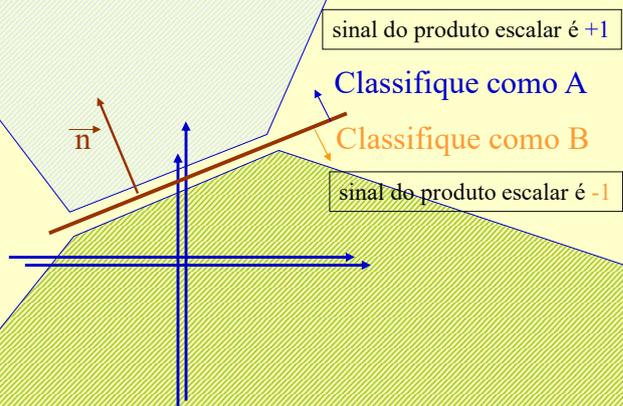


65

© Prof. Emilio Del Moral Hernandez

65

Agora, foquemos no maior poder do hiperplano COM viés (ele não necessariamente “passa na origem”)



66

Parênteses / Desafio a vocês: conciliem ...

1) o Hiperplano separador (com viés)

e

2) o Perceptron de Rosenblatt

68

© Prof. Emilio Del Moral Hernandez

68

Mostre que matematicamente eles são exatamente equivalentes:

(e que portanto são equivalentes, do ponto de vista das classes que conseguem ou não separar)

1) A (in)equação para a separação por hiperplano (com viés):

$$(X-PP).N > 0?$$

e

2) O Perceptron de Rosenblatt (neurônio isolado):

$$y = \text{sinal}(W.X - \text{viés})$$

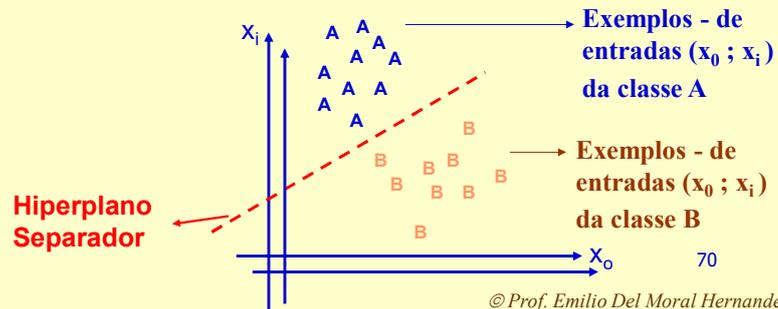
69

© Prof. Emilio Del Moral Hernandez

69

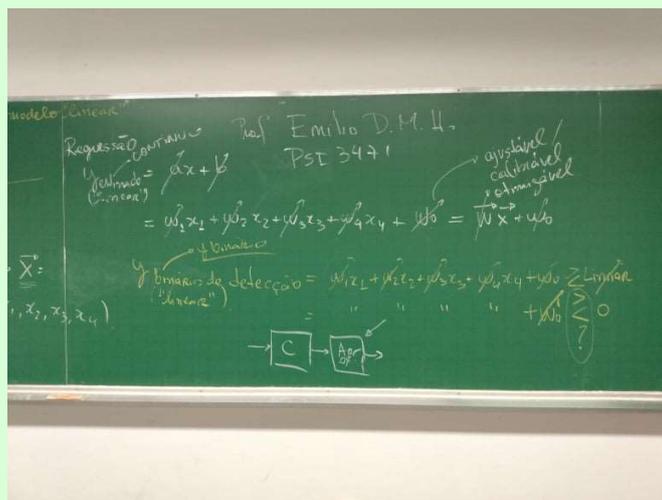
O Perceptron de Frank Rosenblatt ("o original"): $y = \text{sinal}(\sum w_i x_i - \theta)$

- Viabiliza a classificação de padrões com separabilidade linear
- O algoritmo de aprendizado adapta os Ws de forma a encontrar o hiperplano de separação adequado
- Aprendizado por conjunto de treinamento



70

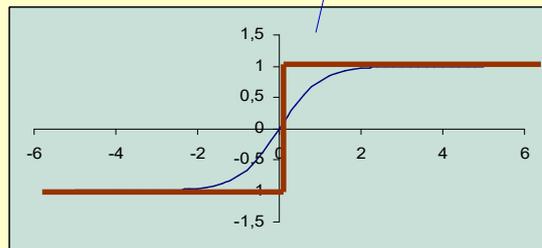
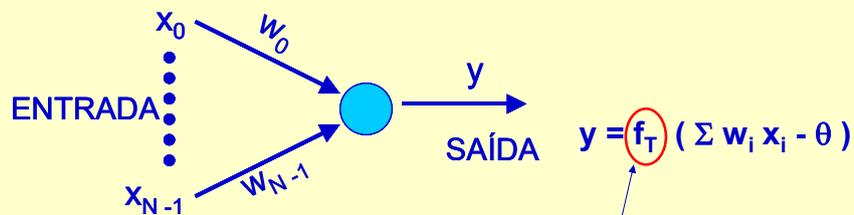
... nas aulas iniciais, já adiantamos algo sobre discriminadores lineares em lousa (em giz amarelo)



© Prof. Emilio Del Moral Hernandez

71

$f_T \text{ Sign}(\cdot) \dots \text{versus} \dots f_T \text{ Sigmoide}(\cdot)$



© Prof. Emilio Del Moral Hernandez

73

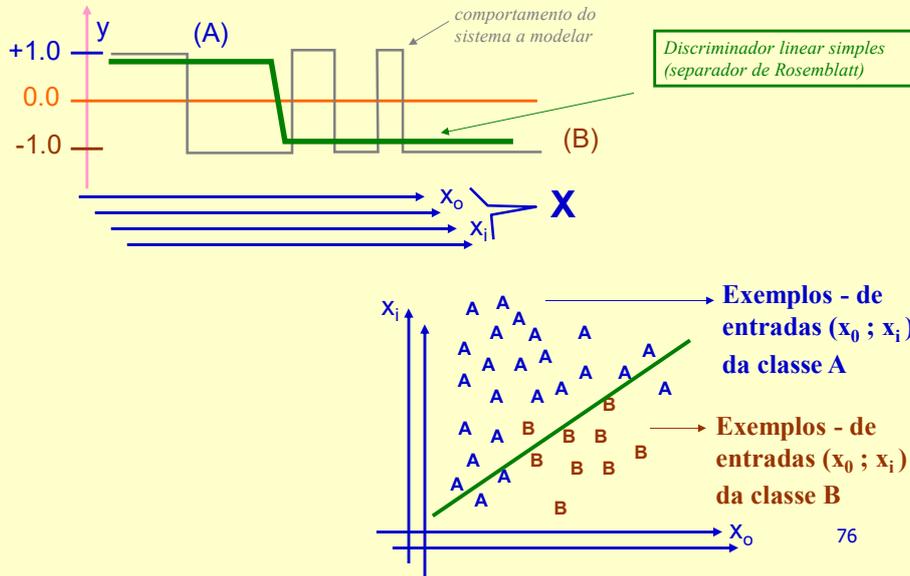
E os separadores com fronteiras lineares sempre vão resolver adequadamente os problema de discriminação de duas classes complementares A e B?

75

© Prof. Emilio Del Moral Hernandez

75

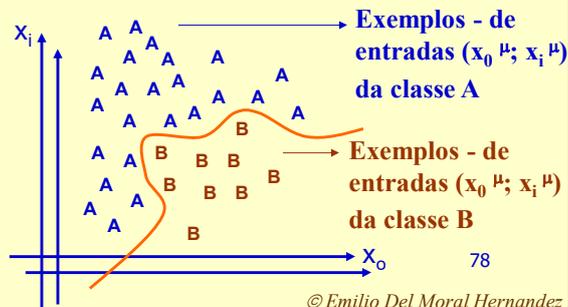
Classificação binária / reconhecimento de padrões, em casos mais complexos de classes A e B ...



76

Em cenários mais complicados ...

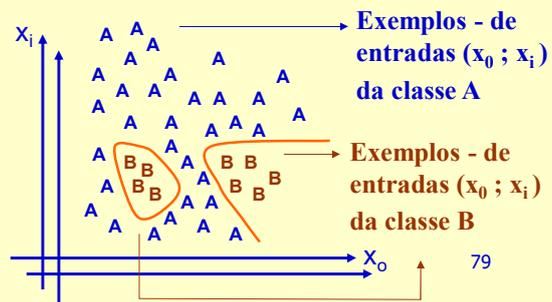
- **Não há separabilidade linear entre as 2 classes**
(nenhum hiperplano é capaz de deixar uma classe integralmente de cada um de seus "lados")
- **O "muro" / ou a hipersuperfície de separação tem que ser algo mais elaborado que um hiperplano**
(claro, assumindo que você precisa um desempenho maior que aquele do separador linear, que já faz uma discriminação, mas com alguns erros)



78

Seguindo com casos complicados de configurações de classes A e B

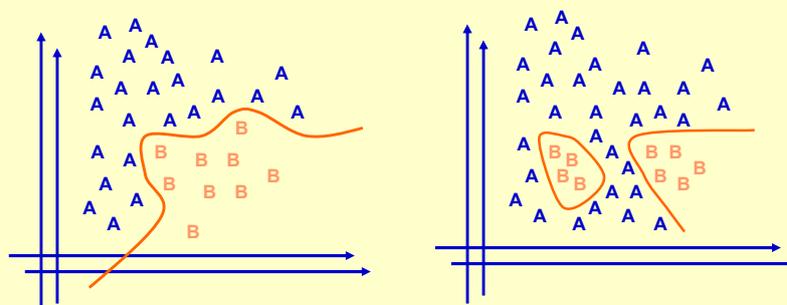
- Neste caso adiante, não há sequer contiguidade / há fragmentação das regiões do espaço associadas a cada classe



79

Precisamos empregar estratégias NÃO LINEARES para o reconhecimento de padrões complexos

Com as **Redes Neurais Artificiais**, a hipersuperfície de separação entre classes pode ser genérica



81

© Prof. Emilio Del Moral – EPUSP

81

Entendamos porque MLPs são capazes de implementar ...

82

- A) Qualquer fronteira de separação entre duas classes, através de um banco de separadores lineares adequado – temos pois um separador genérico
- B) Qualquer cálculo lógico combinatório / qualquer circuito digital combinatório – veremos que isto (B) é também um dos ingredientes para implementar a conjunção lógica genérica de muitos elementos de separação linear, levando à emulação de um separador genérico, acima mencionado no item A

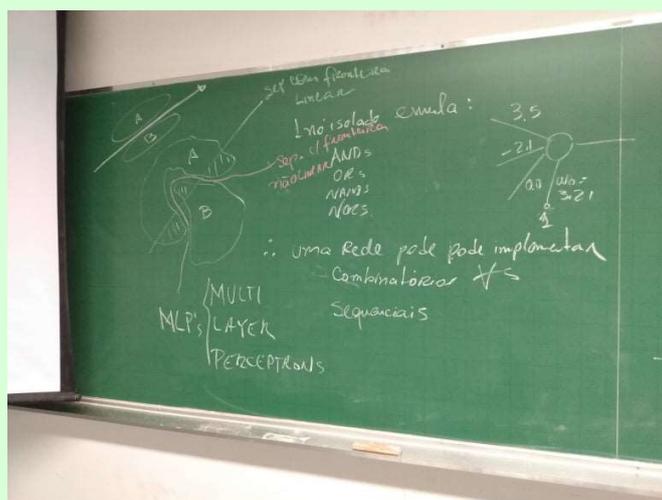
Prof. Emilio Del Moral Hernandez

82

82

Em lousa de aula anterior, já mencionamos brevemente estas possibilidades; hoje vamos entendê-las

83



© Prof. Emilio Del Moral Hernandez

83

83

Nos próximos slides, construiremos um banco de separadores lineares diferentes, cujos veredictos individuais “+1 ou -1” acerca das suas classificações binárias sobre o pertencimento de X a uma das duas classes (A versus B), serão depois combinados com cálculos lógicos, gerando um resultado binário global do comitê, adequado à separação A versus B.

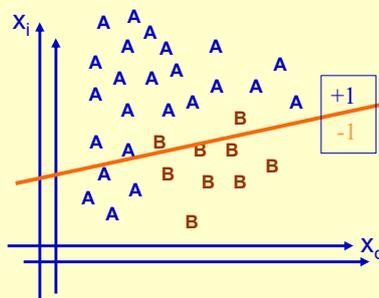
Vamos ilustrar o método com um caso simples ...

84

© Emilio Del Moral Hernandez

84

Qualquer separador linear tem algumas falhas

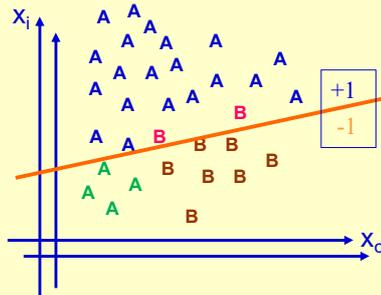


86

© Emilio Del Moral Hernandez

86

Qualquer separador linear tem algumas falhas

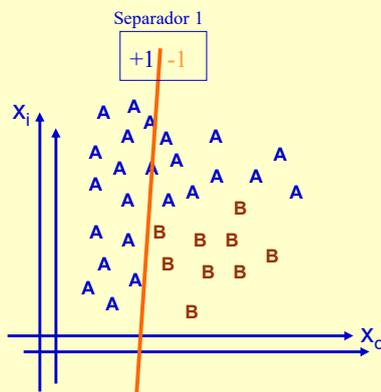


87

© Emilio Del Moral Hernandez

87

Este separador linear tem falhas, ... Mas ele NÃO FALHA com os exemplares B

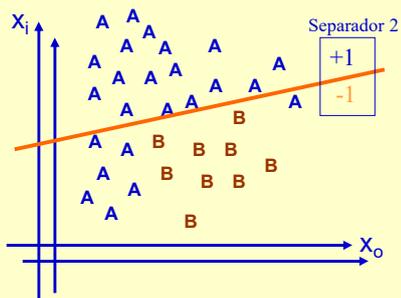


88

© Emilio Del Moral Hernandez

88

Este outro tem falhas diferentes; mas ele tampouco falha com B

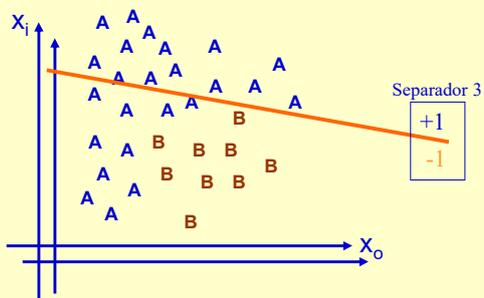


89

© Emilio Del Moral Hernandez

89

Este terceiro separador tem novas falhas; mas tampouco falha com B

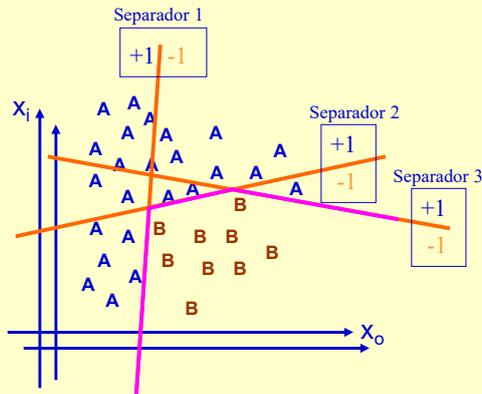


90

© Emilio Del Moral Hernandez

90

Cada um dos 3 separadores falha; mas os 3 conjugados não!



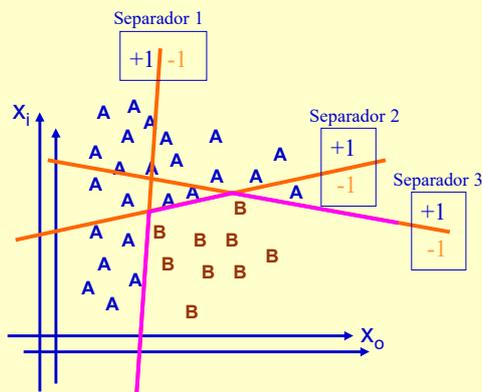
91

© Emilio Del Moral Hernandez

91

Cybenko “pelas bordas” ... Conjugando vários separadores lineares para uma fronteira genérica

Cada um dos 3 separadores falha,
Mas os 3 conjugados não!



Solução coletiva, sem falhas:

Decisão “A / não A” =

(A segundo Sep. 1)

OR

(A segundo Sep. 2)

OR

(A segundo Sep. 3)

Em outros casos, o cálculo lógico pode ser mais complexo, mas com ORs, NORs, NOTs, ANDs e NANDs, sempre daremos conta do recado!

93

© Emilio Del Moral Hernandez

93

Você consegue imaginar a generalização desta estratégia aqui discutida e delineada, para problemas em que X tem mais de 2 dimensões?

Como seria se X tivesse 3 dimensões?

E se X tiver $N > 3$ dimensões?

94

© Prof. Emilio Del Moral Hernandez

94

Discutindo o neurônio camaleão e como diferentes combinações de seus w 's levam a diferentes mapeamentos entre entrada e saída ...

- *Neurônios lineares ... será que podemos especular sobre modelagens variáveis??*

- *E sobre discriminadores Lin*

Façamos agora ...

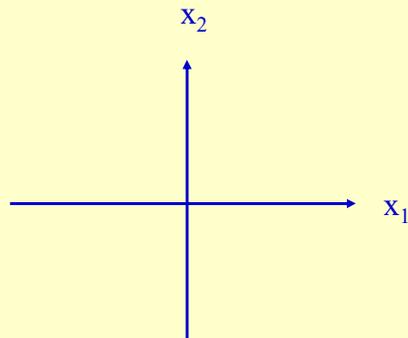
ORs, NORs, NOTs, ANDs e NANDS, de 2 entradas.

- *E sobre algumas portas lógicas que citamos em várias ocasiões? ... análise de como implementá-las (com o uso prático importante em classificação, discutido mais adiante)*

99

**Implementemos algumas portas digitais de 2 entradas,
 x_1 e x_2 , com o neurônio camaleão ...**

$$y_{\text{neurônio}} = \text{tgh}(w_1 \cdot x_1 + w_2 \cdot x_2 + w_0)$$



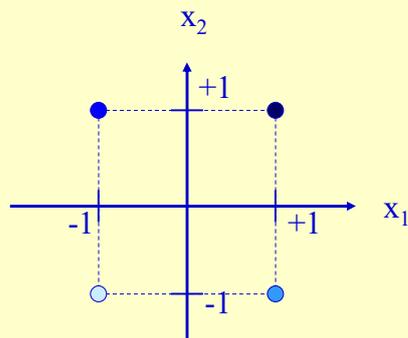
100

© Emilio Del Moral Hernandez

100

**Implementemos algumas portas digitais de 2 entradas,
 x_1 e x_2 , com o neurônio camaleão ...**

$$y_{\text{neurônio}} = \text{tgh}(w_1 \cdot x_1 + w_2 \cdot x_2 + w_0)$$



**Domínio dos X possíveis ... 4 valores:
 $(-1,-1)$, $(-1,+1)$, $(+1,-1)$ e $(+1,+1)$**

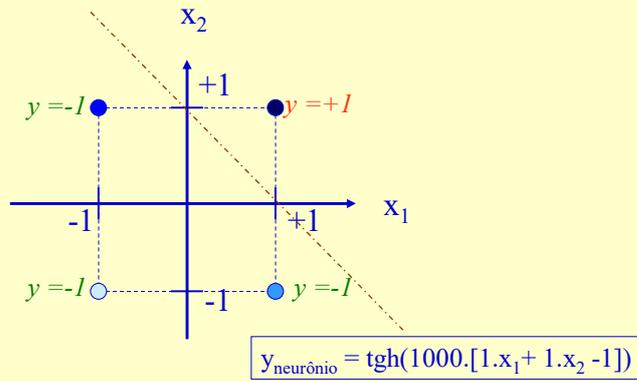
101

© Emilio Del Moral Hernandez

101

Implementemos algumas portas digitais de 2 entradas, x_1 e x_2 , com o neurônio camaleão ...

façamos uma porta AND: $y_{\text{neurônio}} = \text{tgh}(w_1 \cdot x_1 + w_2 \cdot x_2 + w_0)$



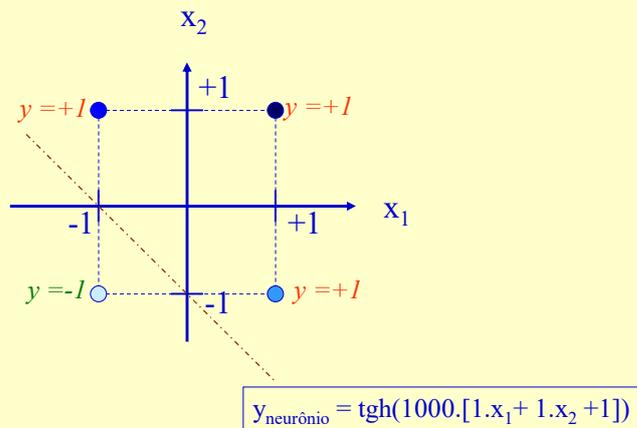
102

© Emilio Del Moral Hernandez

102

Implementemos algumas portas digitais de 2 entradas, x_1 e x_2 , com o neurônio camaleão ...

façamos uma porta OR: $y_{\text{neurônio}} = \text{tgh}(w_1 \cdot x_1 + w_2 \cdot x_2 + w_0)$



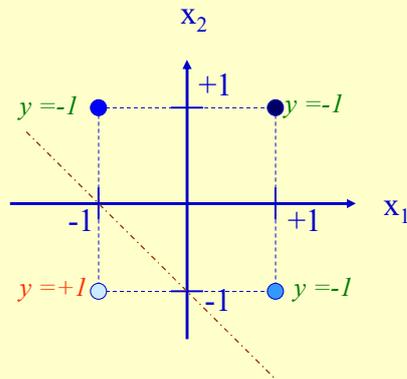
103

© Emilio Del Moral Hernandez

103

Implementemos algumas portas digitais de 2 entradas, x_1 e x_2 , com o neurônio camaleão ...

façamos uma porta NOR: $y_{\text{neurônio}} = \text{tgh}(w_1 \cdot x_1 + w_2 \cdot x_2 + w_0)$



$$y_{\text{neurônio}} = \text{tgh}(1000 \cdot [-1 \cdot x_1 - 1 \cdot x_2 - 1])$$

105

© Emilio Del Moral Hernandez

105

Discutindo o neurônio camaleão e como diferentes combinações de seus w's levam a diferentes mapeamentos entre entrada e saída ...

- Neurônios
podemos
de 2 variáveis

- E sobre

- E sobre
em várias
impleme
em class

Vocês fariam as portas que faltam? ...
NOTs e NANDS de 2 entradas?

E também conseguiriam refazer as portas
AND / OR / NOR de 2 entradas aqui já
feitas, mas agora com 3 entradas? ...

Conseguem vislumbrar como?

E também vislumbram como essas portas
seriam feitas com N entradas?

106

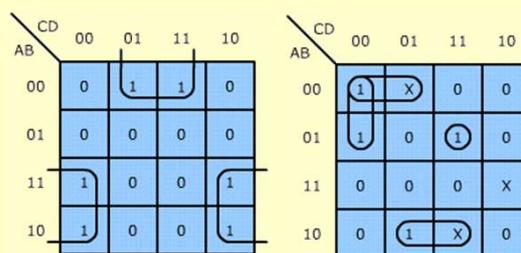
*E conseguiríamos a
implementação de
tabelas da verdade
genéricas com
múltiplas entradas
???*

108

© Emilio Del Moral Hernandez

108

Sabemos de sistemas digitais como implementar qualquer tabela da verdade a partir de AND's das variáveis de entrada (ou de suas negações) e um OR dessas and's ...



Exemplo de cálculo digital envolvendo 4 variáveis binárias ABCD correspondentes às saídas de 4 separadores lineares (S_1 , S_2 , S_3 , e S_4) ...

$$F = S_1' \cdot S_2' \cdot S_4' + S_1' \cdot S_2 \cdot S_3 \cdot S_4 + S_1 \cdot S_2' \cdot S_4$$

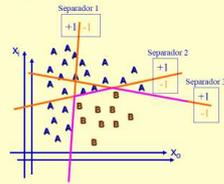
109

© Prof. Emilio Del Moral Hernandez

109

Cybenko "pelas bordas" ... Conjugando vários separadores lineares para uma fronteira genérica

Cada um dos 3 separadores falha, Mas os 3 conjugados não!

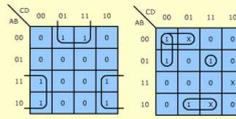


Solução coletiva, sem falhas:
 Decisão "A / não A" =
 (A segundo Sep. 1)
 OR
 (A segundo Sep. 2)
 OR
 (A segundo Sep. 3)

Em outros casos, o cálculo lógico pode ser mais complexo, mas com ORs, NOTs, ANDs e NANDs, sempre daremos conta do recado!

PS12672 - Rec. de Padrões, Modelagem e Redes Neurais - Prof. Emilio Del Moral Hernandez - © 2015-16

Sabemos de sistemas digitais como implementar qualquer tabela da verdade a partir de AND's das variáveis de entrada (ou de suas negações) e um OR dessas and's ...



Exemplo de cálculo digital envolvendo 4 variáveis binárias ABCD correspondentes às saídas de 4 separadores lineares (S_1, S_2, S_3, S_4) ...

$$F = S_1' \cdot S_2' \cdot S_4' + S_1' \cdot S_2 \cdot S_3 \cdot S_4 + S_1 \cdot S_2' \cdot S_4$$

PS12672 - Rec. de Padrões, Modelagem e Redes Neurais - Prof. Emilio Del Moral Hernandez - © 2015-16

Discutindo o neurônio camaleão e como diferentes combinações de seus w's levam a diferentes mapeamentos entre entrada e saída ...

- Neurônios podemos de 2 vari

- E sobre

- E sobre em várias implementações em class

Vocês fariam as portas que faltam? ... NOTs e NANDs de 2 entradas?

E também conseguiriam refazer as portas AND / OR / NOR de 2 entradas aqui já feitas, mas agora com 3 entradas? ...

Conseguem vislumbrar como?

E também vislumbram como essas portas seriam feitas com N entradas?

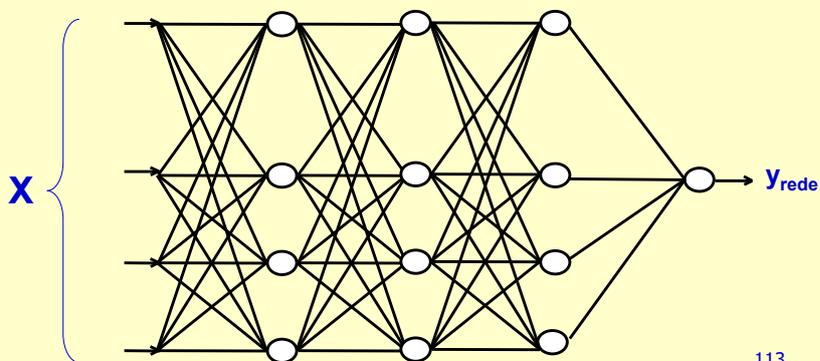
Com a implementação de diversos separadores lineares (na primeira camada de neurônios) com portas lógicas realizando operações lógicas sobre as saídas binárias desses separadores (portas nas demais camadas de neurônios), temos uma separação genérica entre os conjuntos de amostras de A e B 110

© Prof. Emilio Del Moral Hernandez

110

O grande poder do MLP como classificador/detector universal

- Múltiplas entradas
- Saídas digitais (+1 e -1)
- Não há a restrição de separabilidade linear

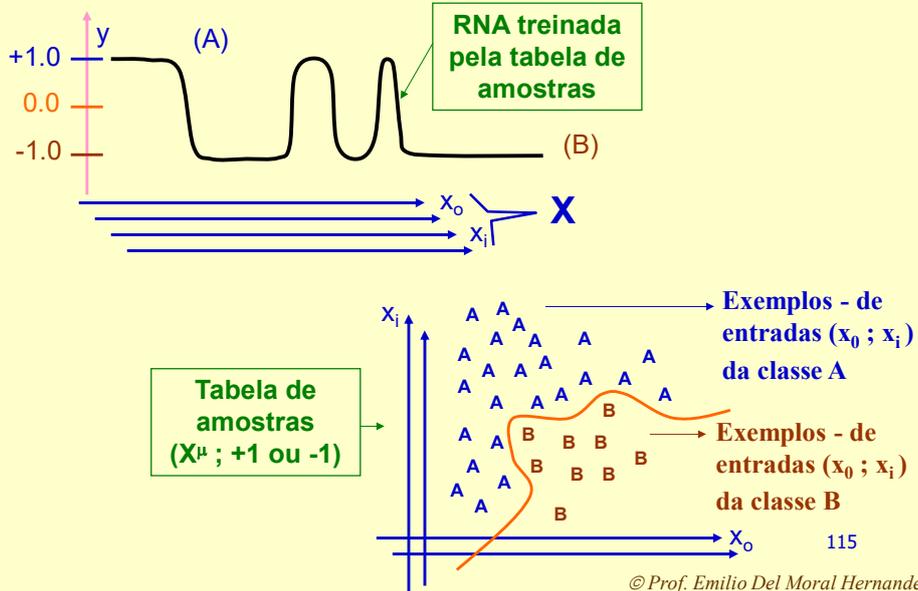


113

© Prof. Emilio Del Moral Hernandez

113

A função $y_{rede}(X)$ implementada pela RNA / MLP neste caso



© Prof. Emilio Del Moral Hernandez

115

E os hiperplanos separadores que compõem o banco e as portas lógicas que conjugam as separações realizadas por eles aparecem explicitamente em nós neurais específicos de uma rede neural após treinada (por Error Back Propagation, por exemplo) para implementar uma superfície de separação complexa e necessária a um conjunto de dados das duas classes de nosso interesse ??

Geralmente não, pois a escolha de pesos é feita mirando a otimização do resultado final, ou seja, minimizando o Eqm do conjunto de pares empíricos de treino. Se esta minimização for obtida de forma mais eficiente com outra escolha de pesos distinta daquela que acabamos de construir, essa outra escolha será priorizada.

116

© Prof. Emilio Del Moral Hernandez

116