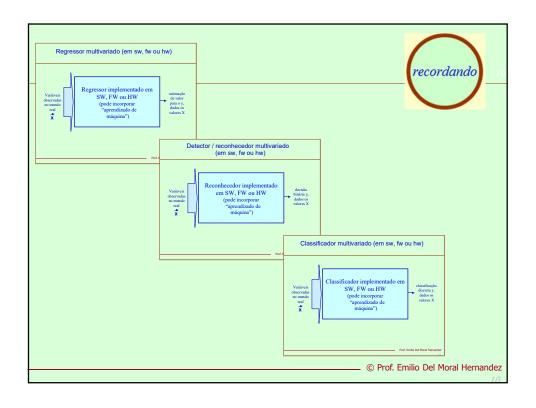
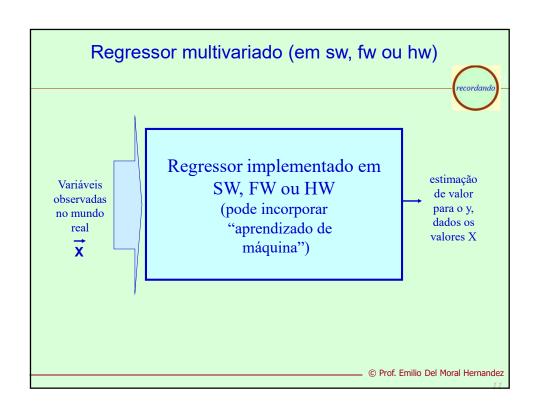
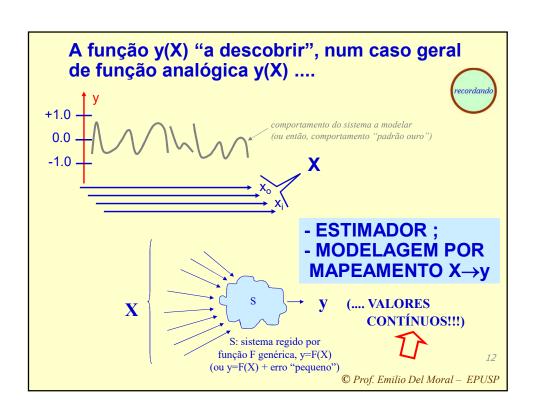
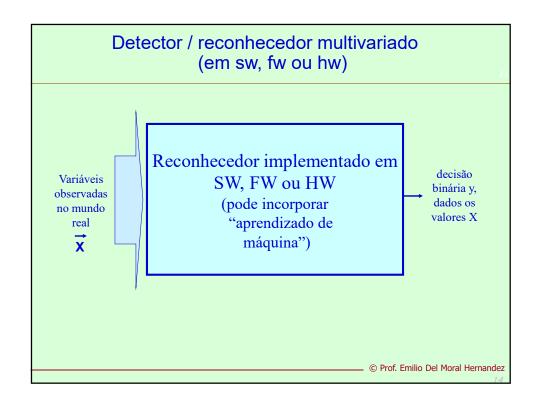
## Medidas de desempenho em ...

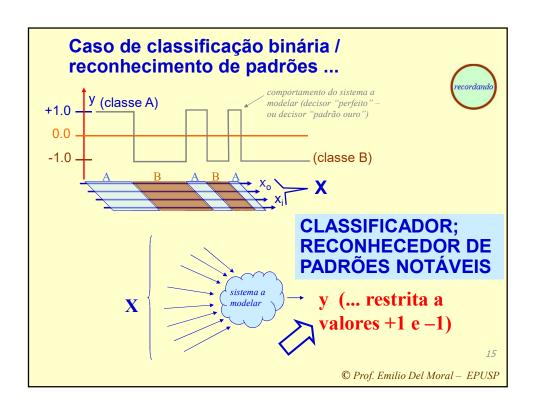
# REGRESSORES RECONHECEDORES DETECTORES CLASSIFICADORES

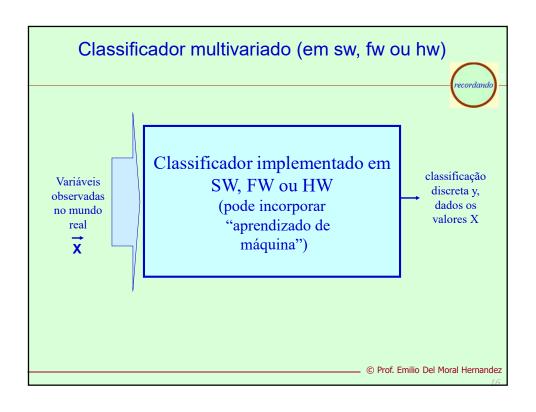


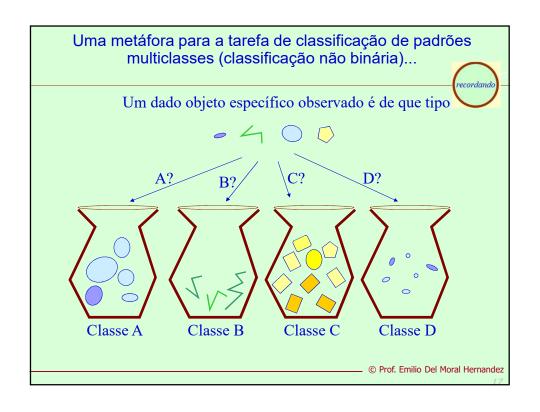






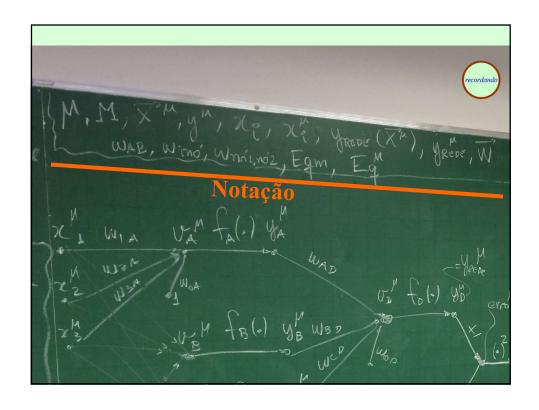


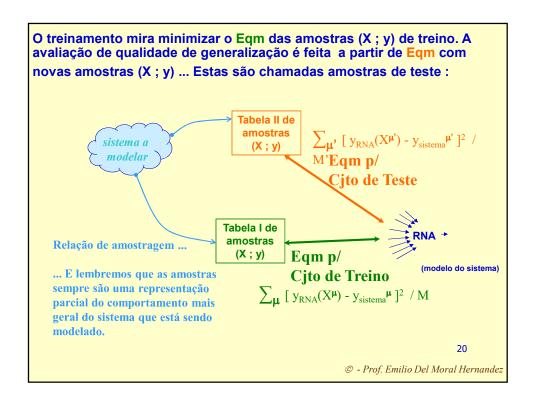


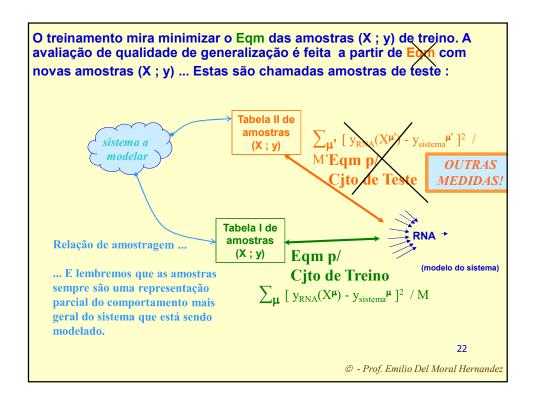


... A medida de aderência do modelo aos dados de treino clássica Eqm (ou RMS) é usada frequentemente tanto como indicador de qualidade da modelagem quanto como o "motor" do método do gradiente ...

Mas as medidas da caracterização final de qualidade do modelo – face aos dados empíricos de teste – têm mesmo que ser feitas com o Eqm (ou RMS), ou poderíamos usar outras medidas diferentes do Eqm?







# Pergunta: O Eqm (ou RMS) indica a qualidade do modelo? *Onde usar Eqm ou onde não?*

# Algumas potenciais medidas de qualidade em regressores (cada aplicação pode preferir umas ou outras ...)

- Eqm e RMS
- Módulo Médio do Erro
- Máximo Módulo do Erro
- Média dos Erros Positivos
- Máximo Módulo dos Erros Positivos
- Média dos Erros Negativos
- Máximo Módulo dos Erros Negativos
- Esses todos derivados dos erros acima, mas em suas versões normalizadas, com relação ao módulo de y
- Estes todos acima, mas sujeitos a conhecimento de X (local) qualidade dependente do valor do argumento X da regressão
- Histograma de erros (ou seja, a "densidade de probabilidade de erros empírica")
- Faixa de valores de erros que se enquadram num certo número de "deciles" ou terciles, ou quartiles, etc etc – da distribuição de erros (seja sobre a distribuição o erro com sinal + -, ou seja sobre o erro em módulo)
- Combinações específicas de vários acima ... Como bem percebido por colegas em sala, várias combinações fazem <u>muito</u> sentido, como informação de caracterização mais completa ao cliente / usuário do regressor!

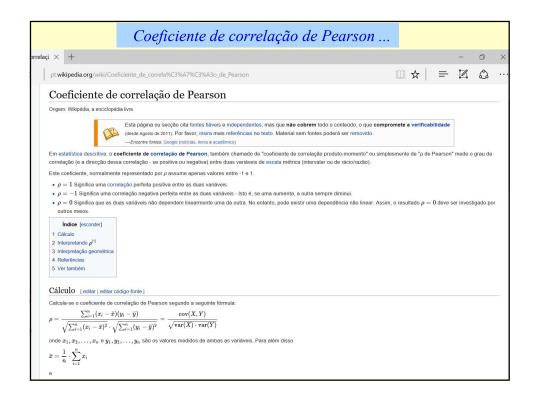
PSI2672 – Reconhec. de Padrões, Modelagem e Redes Neurais – Prof. Emilio Del Moral Hernandez – © 2015

rnandez

# Algumas potenciais medidas de qualidade em regressores (cada aplicação pode preferir umas ou outras ...)

- Eqm e RMS
- Módulo Médio do Erro
- Máximo Módulo do Erro
- Média dos Erros Positivos
- Máximo Módulo dos Erros Positivos
- Média dos Erros Negativos
- Máximo Módulo dos Erros Negativos
- Esses todos derivados dos erros acima, mas em suas versões normalizadas, com relação ao módulo de y
- Estes todos acima, mas sujeitos a conhecimento de X (local) qualidade dependente do valor do argumento X da regressão
- Histograma de erros (ou seja, a "densidade de probabilidade de erros empírica")
- Faixa de valores de erros que se enquadram num certo número de "deciles" ou terciles, ou quartiles, etc etc – da distribuição de erros (seja sobre a distribuição o erro com sinal + -, ou seja sobre o erro em módulo)
- Combinações específicas de vários acima ... Como bem percebido por colegas em sala, várias combinações fazem <u>muito</u> sentido, como informação de caracterização mais completa ao cliente / usuário do regressor!

L



2 Interpretando  $\rho$ 

3 Interpretação geométrica

4 Referências

5 Ver também

#### Cálculo [editar | editar código-fonte]

Calcula-se o coeficiente de correlação de Pearson segundo a seguinte fórmula:

$$\rho = \frac{\sum_{i=1}^{n} (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^{n} (x_i - \bar{x})^2} \cdot \sqrt{\sum_{i=1}^{n} (y_i - \bar{y})^2}} = \frac{\text{cov}(X, Y)}{\sqrt{\text{var}(X) \cdot \text{var}(Y)}}$$

onde  $x_1, x_2, \dots, x_n$  e  $y_1, y_2, \dots, y_n$  são os valores medidos de ambas as variávei

$$ar{x} = rac{1}{n} \cdot \sum_{i=1}^n x_i$$

е

Coeficiente de correlação de Pearson ...

2

 ${\it @}$  - Prof. Emilio Del Moral Hernandez

## Medidas de desempenho em reconhecedores de padrões / detectores e classificadores

© - Prof. Emilio Del Moral Hernandez

#### E quanto a reconhecedores de padrões?

- Taxa de acertos (%)
- Taxa de erros ( % complementar )
- Taxa de acertos em separado para as duas classes
   (A x B) ... Taxas de acerto / erro para A .. e para B ...

Sensibilidade, Especificidade, etc etc

- ... Sensibilidade = VP / Todos casos realmente Positivos
- ... Especificidade = VN / Todos casos realmente Negativos

29

 ${\it @}$  - Prof. Emilio Del Moral Hernandez

# Continuando com + conceitos: avaliando o desempenho de reconhecedores de padrões

- O Erro quadrático médio Eqm (ou o RMSE), que é fornecido automaticamente pelas bibliotecas neurais ... poderia ser usado, mas há indicadores mais interessantes, que são específicos para o contexto de reconhecimento e classificação de padrões
- Taxa de acertos em classificação / reconhecimento
- 4 Combinações possíveis (ou (K<sub>classes</sub>)<sup>2</sup> no caso geral multiclasses) envolvendo as classificações do classificador "padrão ouro" e as classificações do modelo neural

(diagramas de Venn / conjuntos nos ajudam a entender)

• Matriz de confusão: tamanho é K<sub>classes</sub> x K<sub>classes</sub>

3

© - Prof. Emilio Del Moral Hernandez

(I) Podemos ter <u>imperfeição</u> na separabilidade de classes gerada por limitações da fronteira separadora da classe de modelagem empregada:

fronteira linear

x

não linear ("curva") específica

x

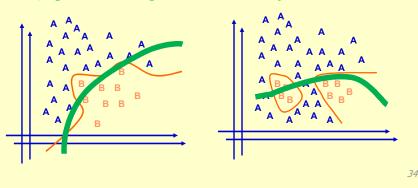
fronteira universal

33

# Capacidade de reconhecimento de padrões em casos complexos NÃO LINEARES

# Com as RNAs, a hipersuperfície de separação entre classes vai muito além dos hiperplanos

Em verde: fronteiras limitadas (implementadas com poucos nós neurais), que levam a alguns erros na classificação automática.

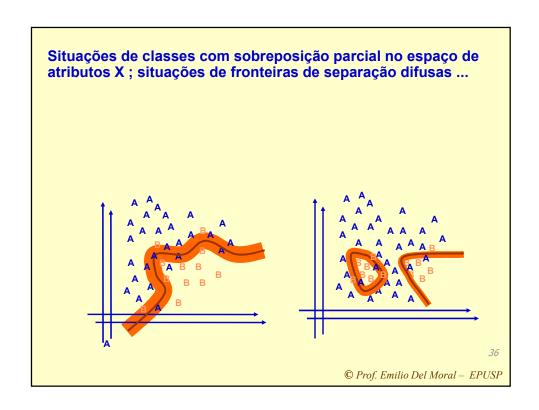


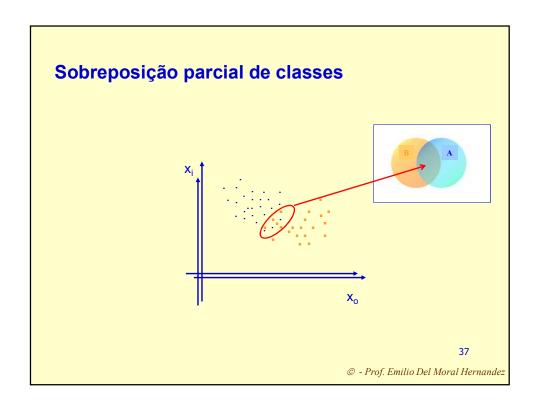
© Prof. Emilio Del Moral – EPUSP

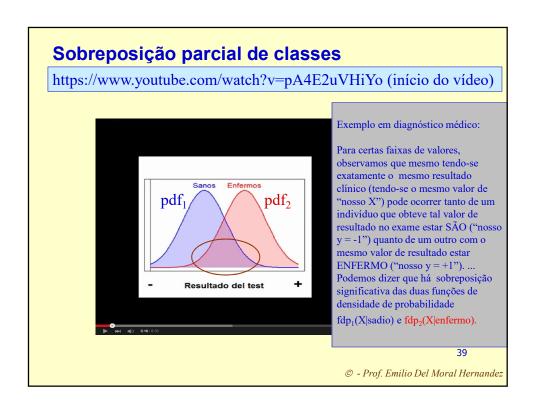
(II) Mas ... Pode também haver imperfeição na separabilidade de classes mesmo com uma fronteira genérica, se existir no fenômeno modelado uma sobreposição parcial de classes (B) no espaço de medidas X:

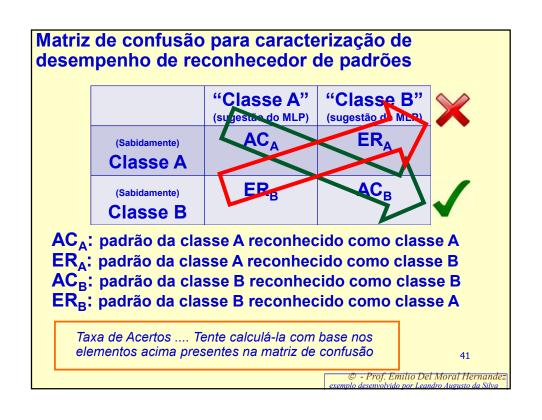
... nesse caso, nenhuma fronteira de separação escolhida garantirá a separação perfeita

35









Matriz de confusão avaliada sobre resultados com o o conjunto de dados empíricos (treino e/ou teste) ... Há inúmeros "estilos" envolvendo porcentages ...

	Aprovar Crédito (sugestão MLP)	Reprovar Crédito (sugestão MLP)
Aprovar Crédito (registro da base de dados)	80.0%	20.0%
Reprovar Crédito (registro da base de dados)	5.0%	95.0%

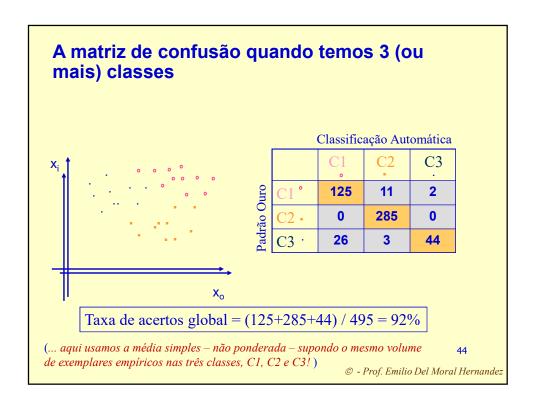
Acerto<sub>MLP</sub> = 
$$\frac{80.0\% + 95.0\%}{2} = 87.5\%$$

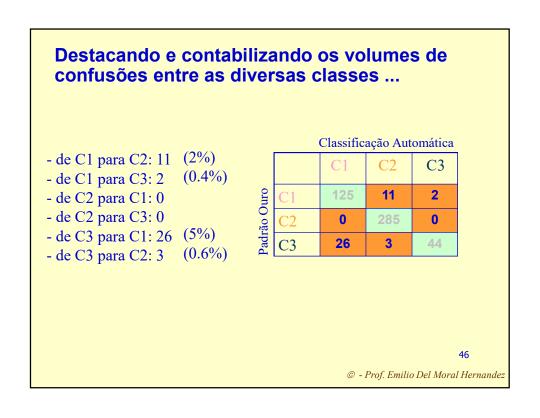
(... e aqui usamos a média simples, não ponderada, supondo mesmo volume de exemplares nas duas classes, positivos e falsos!)

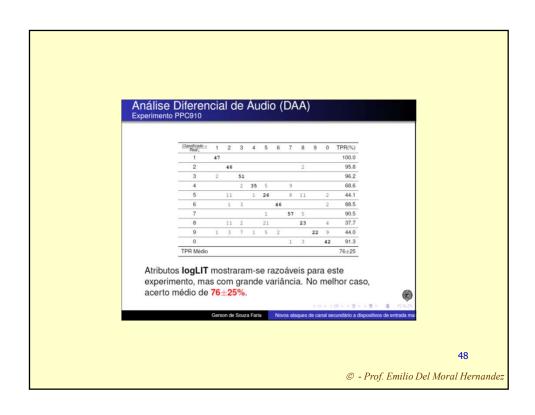
exemplo desenvolvido por Leandro Augusto da Silva

Um parênteses ... Como você generalizaria o conceito de matriz de confusão para o caso de 3 classes em vez de duas? E para 10 classes?

43









## Casos em que o classificador binário é encarado como um detector / identificador

- Uma das duas classes representa a detecção / o reconhecimento / a ocorrência de um dado cenário relevante e que está sendo alvo de detecção.
  - Por ex: deteção de provável doença num diagnóstico, detecção de uma potencial turbulência no mercado, detecção de uma provável fraude em curso, detecção de um potencial consumidor de um produto, etc etc etc.
- A outra classe, complementar à anterior, indica a NÃO ocorrência de tal cenário

#### relevante e que está sendo alvo de detecção.

- complementarmente: não há doença provável, não há potencial turbulência, não há provável fraude em curso, não é consumidor provável, etc etc etc.
- Nesse claso, uma classe é denominada positiva (= sim, temos detecção do cenário alvo) e a outra negativa (= não, o cenário alvo não foi detectado)

51

© - Prof. Emilio Del Moral Hernandez

# Casos em que o classificador binário é encarado como um detector / identificador

- Assim, em lugar da saída binária do classificador automático ser encarada não mais como sendo a indicação de uma de duas classes abstratas A e B, mas sim como sendo a indicação Positiva (P) do detector ou a indicação Negativa (N) do detector, para a ocorrência de um alvo específico de interesse na detecção
- E temos expressões específicas associadas aos casos de acertos ou erros na classificação automática realizada, respectivamente Classificações Verdadeiras (V) e Classificações Falsas (F)

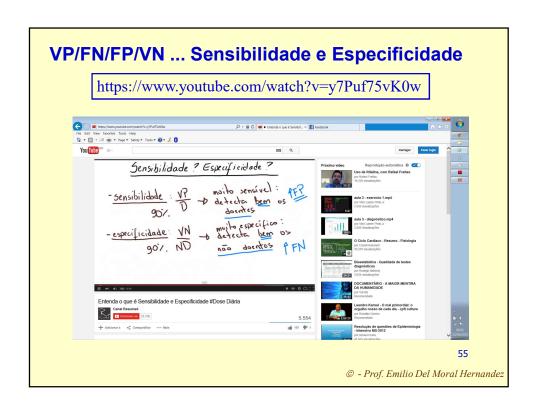
52

# Estendendo termos técnicos relacionados à matriz de confusão e suas componentes ...

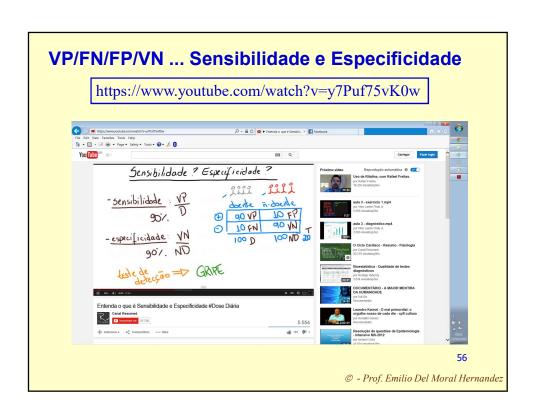
- Dando destaque a uma das duas classes binárias:
   "Diagnóstico Positivo" / "Diagnóstico Negativo"
- Interpretações em outros contextos que não o médico:
  - Alarme de reconhecimento disparado x Alarme em silêncio
  - Risco provável previsto x Sem risco previsto
  - Necessidade de providência x Normalidade (sem necessidade)
  - Oportunidade de ação x Sem destaques que mereçam ação
- Contrastando o diagnóstico perfeito com o diagnóstico imperfeito de um exame ou de um autômato:
   Diagnóstico Verdadeiro / Diagnóstico Falso
- Combinações dos diversos cenários VP / FP / VN / FN
  - Verdadeiro Positivo; Falso Positivo;
  - Verdadeiro Negativo; Falso Negativo.

54

© - Prof. Emilio Del Moral Hernandez



18



# VP/FN/FP/VN ... Sensibilidade e Especificidade (& ROC) https://www.youtube.com/watch?v=y7Puf75vK0w

# Resumo ... Medidas de qualidade em reconhecedores de padrões / detectores

- Taxa de acertos (%)
- Taxa de erros ( % complementar )
- Taxa de acertos em separado para as duas classes:
  - (AxB)...Taxas de acerto e de erro para A e p / B
- Em situações em que uma das classes tem o significado de diagnóstico positivo e a outra de diagnóstico negativo, temos outras taxas consideradas relevantes ...
- Sensibilidade, Especificidade, etc etc
- ... Sensibilidade = VP / Todos casos realmente Positivos
- ... Especificidade = VN / Todos casos realmente Negativos

5

© - Prof. Emilio Del Moral Hernandez

Extra — o que segue nos slides restantes deste bloco de slides (slides deste) não foram ainda abordados em sala e não serão objeto da P1!!!

Mais medidas de desempenho específicas para detectores

 ${\it @}$  - Prof. Emilio Del Moral Hernandez

Extra — o que segue nos slides restantes deste bloco de slides (slides restantes deste arquivo) não foram ainda abordados em sala e não serão objeto da P1!!!

Mais medidas de desempenho específicas para detectores

© - Prof. Emilio Del Moral Hernandez

O conceito de limiar ajustável em detectores

62

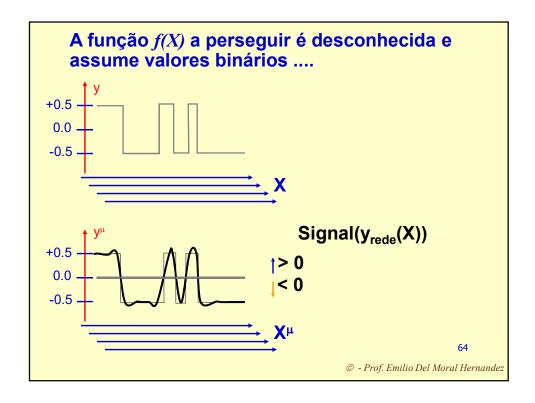
 $\oslash$  - Prof. Emilio Del Moral Hernandez

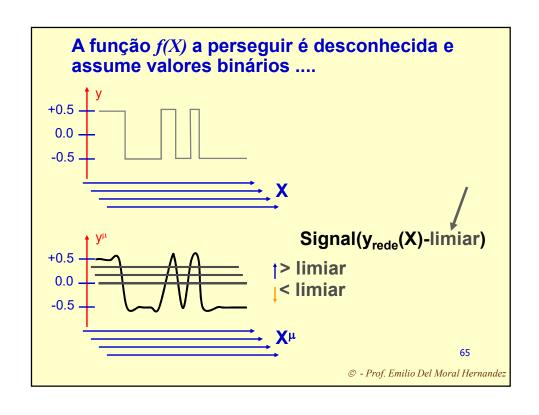
# Estendendo conceitos relacionados à matriz de confusão e suas componentes VP,VN, FP, FN ...

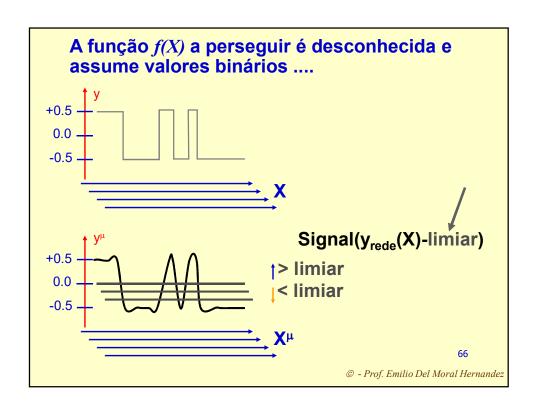
- Conceito de limiar ajustável como parte do reconhecedor (sensibilidade ajustável pelo usuário / variável de acordo com a circunstância)
- Indo do Reconhecedor / Detector super estressado detector super sensível ao Reconhecedor / Detector super dorminhoco detector super específico ...
   (vamos de de 100% diagnósticos P a 100% diagnósticos N)
- Sensibilidade e Especificidade combinadas com o conceito de limiar ajustável leva ao conceito da curva ROC ...

Curva ROC de um reconhecedor com limiar ajustável:
Mapeamos a evolução dos valores de Sensibilidade e de
Especificidade para todos os limiares, desde o extremo de
detecção super sensível (baixo limiar) até o extremo oposto de
detecção super específica (alto limiar), passando por todos os
cenários intermediários de limiar.

63







#### Podemos ter limiar variável em ...

- Redes neurais
- SVMs
- Comitês com composição (Adaboost e outros)
- Decisores estatísticos
- ...
- ...
- Sempre que houver uma grandeza de "saída" comparada com um limiar no estágio final, para a geração a classificação da decisão

Positivo versus Negativo

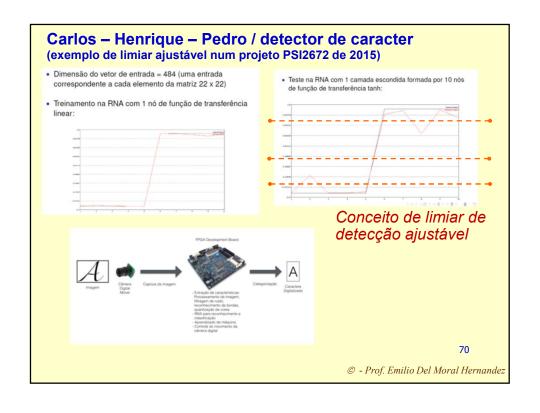
67

© - Prof. Emilio Del Moral Hernandez

# Aplicando o conceito de limiar ajustável a projetos concretos

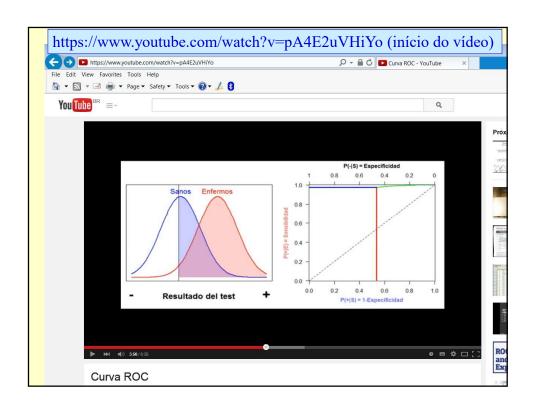
68

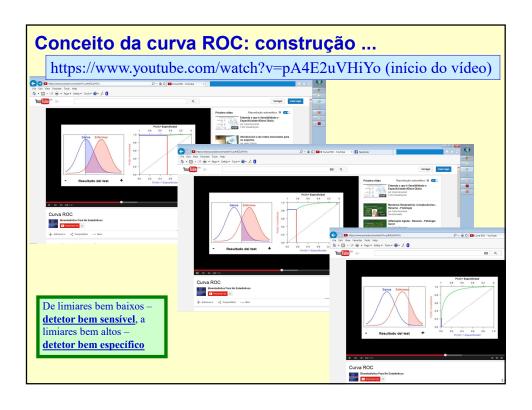
 $\oslash$  - Prof. Emilio Del Moral Hernandez



O conceito de curva ROC associada ao exercício completo do leque de valores do limiar ajustável em um detector

7





A comparação geral de qualidades de detecção associadas as duas ROCs distintas é muitas vezes feita pelo contraste dos valores da área sob a curva ROC: AUC (área under the curve)

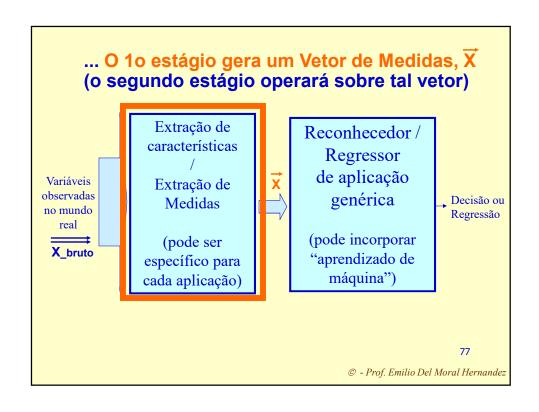
7

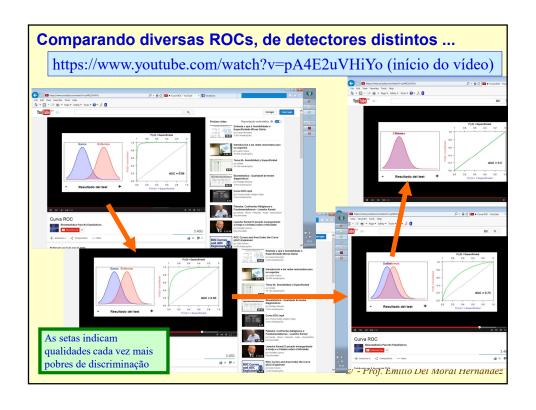
© - Prof. Emilio Del Moral Hernandez

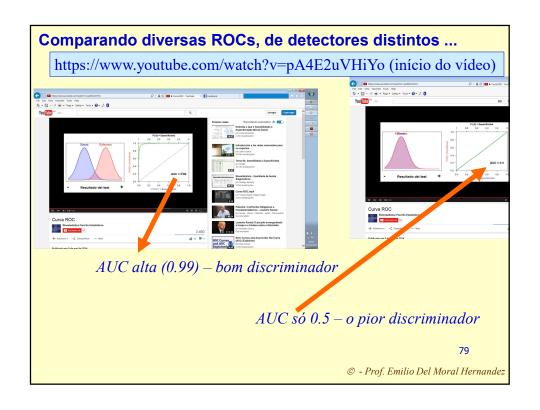
#### Importante lembrete:

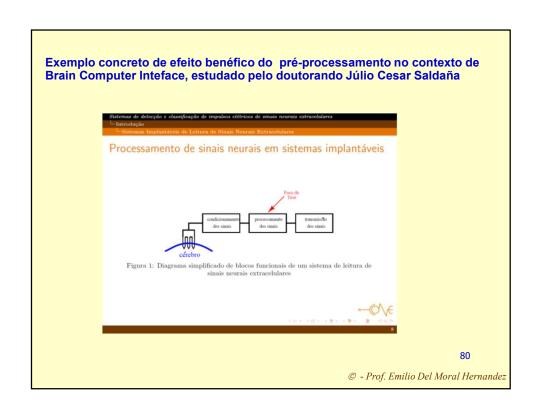
diferentes estratégias de pré –
processamento e mesmo de
reconhecimento podem levar a
diferentes desempenhos de detecção
e portanto a diferentes ROCs

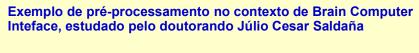
76

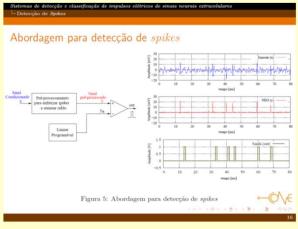








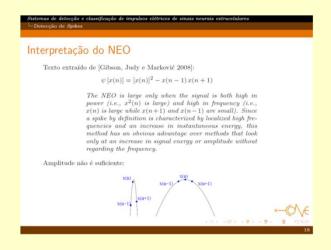




81

© - Prof. Emilio Del Moral Hernandez

### Exemplo de pré-processamento no contexto de Brain Computer Inteface, estudado pelo doutorando Júlio Cesar Saldaña



ดว

# Contexto: BCI. Duas ROCs contrastando pre-processamento distintos em um detetor de pulsos corticais (BCI) com limiar variável

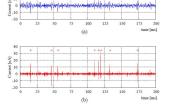


Fig. 4: (a) Synthetic data applied to NEO circuit. Noise standard deviation is equal to 20% the maximum neural spike amplitude. (b) Response of the proposed NEO core circuit to the synthetic input signal.

Trabalho do doutorando Júlio Cesar Saldaña

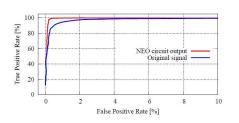


Fig. 5: ROC curve for the input signal (line in blue) and the NEO circuit output signal (red line).

83