

Medidas de desempenho em ...

REGRESSORES

RECONHECEDORES

DETECTORES

CLASSIFICADORES

© - Prof. Emilio Del Moral Hernandez

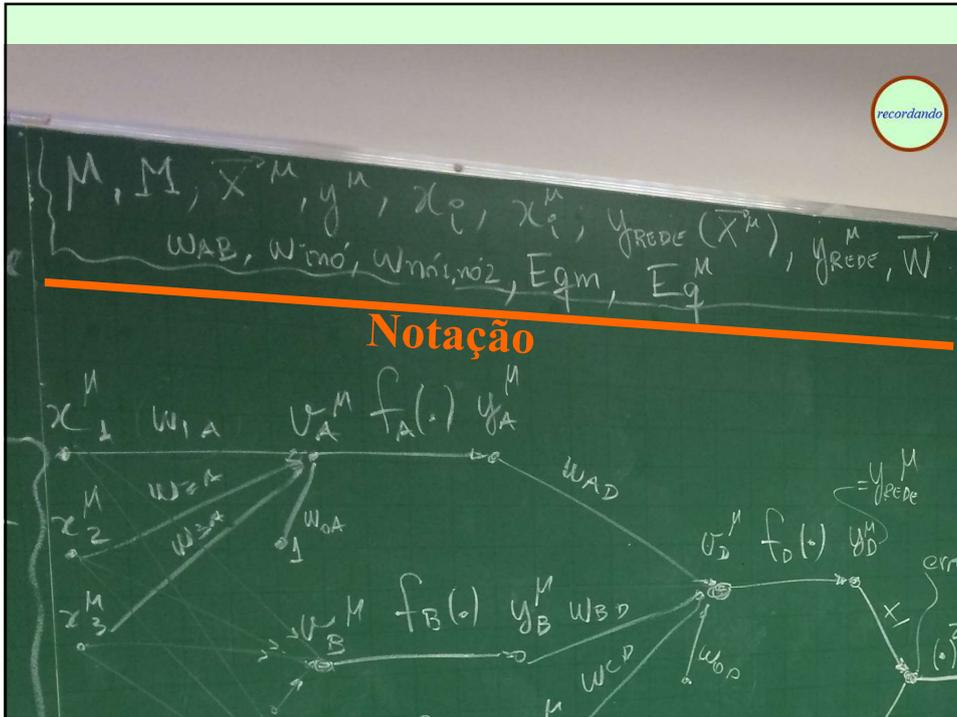
12

... 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?

© - Prof. Emilio Del Moral Hernandez

21



22

O treinamento mira minimizar o **Eqm** das amostras $(X ; y)$ de treino. A avaliação de qualidade de generalização é feita a partir de **Eqm** com novas amostras $(X ; y)$... Estas são chamadas amostras de teste :

Relação de amostragem ...

... E lembremos que as amostras sempre são uma representação parcial do comportamento mais geral do sistema que está sendo modelado.

Tabela II de observações $(X ; y)$

$$\sum_{\mu'} [Y_{RNA}(X^{\mu'}) - y_{sistema}^{\mu'}]^2 / M'Eqm \text{ p/ Cjto de Teste}$$

Tabela I de observações $(X ; y)$

Eqm p/ Cjto de Treino

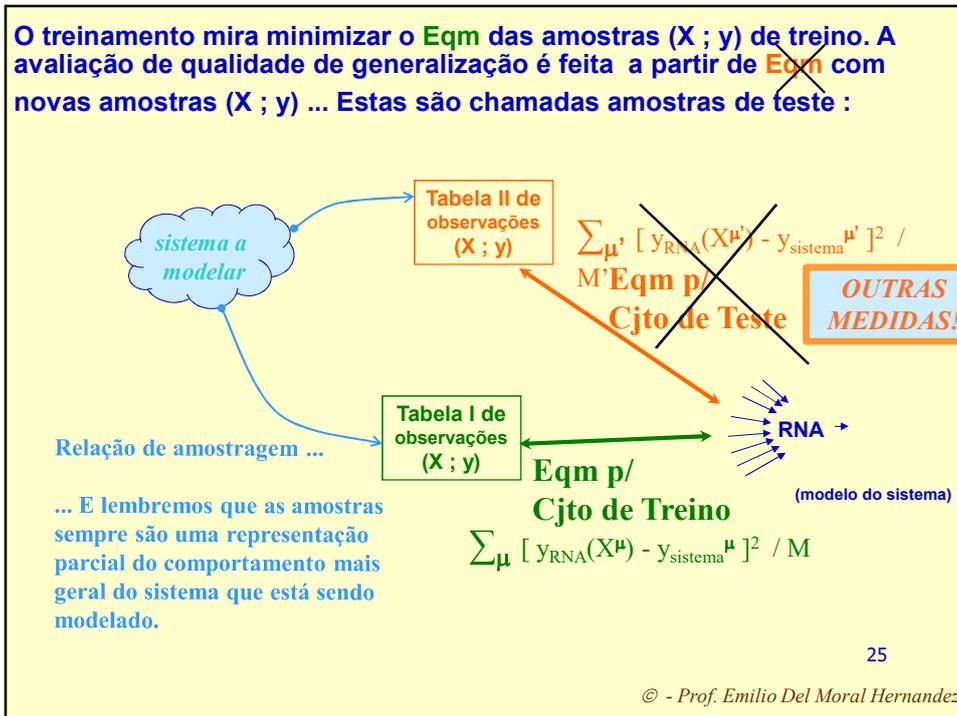
$$\sum_{\mu} [Y_{RNA}(X^{\mu}) - y_{sistema}^{\mu}]^2 / M$$

RNA (modelo do sistema)

23

© - Prof. Emilio Del Moral Hernandez

23



25

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 muito sentido, como informação de caracterização mais completa ao cliente / usuário do regressor!

11

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

mandez

26

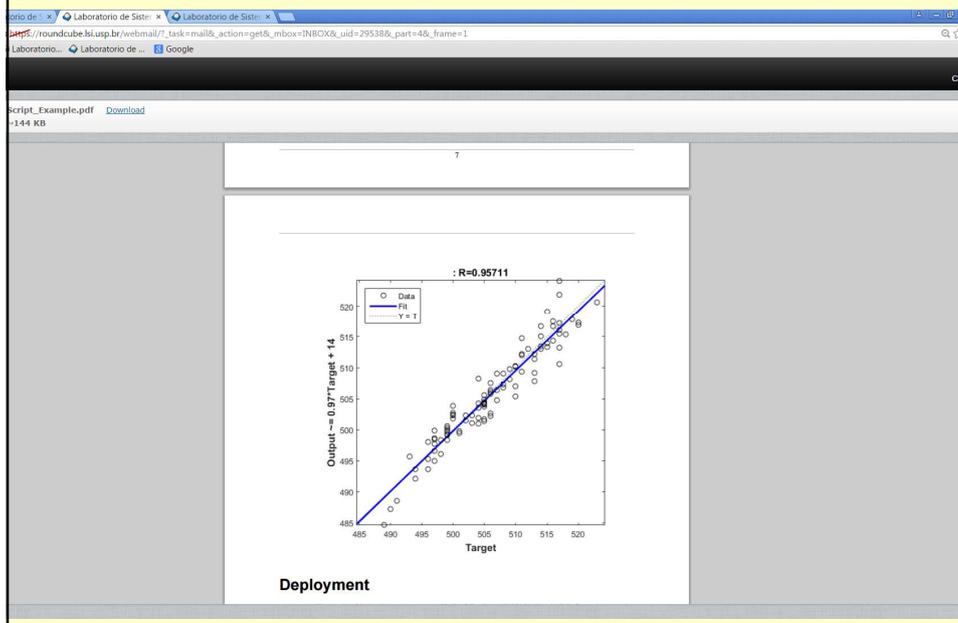
Algumas potenciais medidas de qualidade em regressores (cada aplicação pode preferir umas ou outras ...)

- 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 muito sentido, como informação de caracterização mais completa ao cliente / usuário do regressor!

11

27

Medindo qualidade de regressores ... Plots y corretos versus y estimados + valor de “R” é outra possibilidade



28

Coefficiente de correlação de Pearson ...

pt.wikipedia.org/wiki/Coefficiente_de_correla%C3%A7%C3%A3o_de_Pearson

Coeficiente de correlação de Pearson

Origem: Wikipédia, a enciclopédia livre.

Esta página ou secção cita fontes fiáveis e independentes, mas que **não cobrem** todo o conteúdo, o que **compromete a verificabilidade** (desde Agosto de 2011). Por favor, insira mais referências no texto. Material sem fontes poderá ser removido.
—Encontre fontes: [Google](#) (notícias, livros e académico)

Em estatística descritiva, o **coeficiente de correlação de Pearson**, também chamado de "coeficiente de correlação produto-momento" ou simplesmente de "rho de Pearson" mede o grau da correlação (e a direcção dessa correlação - se positiva ou negativa) entre duas variáveis de escala métrica (intervalar ou de rácio/razão).

Este coeficiente, normalmente representado por ρ assume apenas valores entre -1 e 1.

- $\rho = 1$ Significa uma correlação perfeita positiva entre as duas variáveis.
- $\rho = -1$ Significa uma correlação negativa perfeita entre as duas variáveis - Isto é, se uma aumenta, a outra sempre diminui.
- $\rho = 0$ Significa que as duas variáveis não dependem linearmente uma da outra. No entanto, pode existir uma dependência não linear. Assim, o resultado $\rho = 0$ deve ser investigado por outros meios.

Índice [esconder]

- Cálculo
- Interpretando ρ^{11}
- Interpretação geométrica
- Referências
- 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 variáveis. Para além disso

$$\bar{x} = \frac{1}{n} \cdot \sum_{i=1}^n x_i$$

29

- Interpretando ρ^{11}
- Interpretação geométrica**
- Referências
- 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 variáveis

$$\bar{x} = \frac{1}{n} \cdot \sum_{i=1}^n x_i$$

e

Coefficiente de correlação de Pearson ...

30

© - Prof. Emilio Del Moral Hernandez

30

Pergunta perturbadora ... e as medidas de aderência do modelo aos dados empíricos (para conjunto de treino) ou de qualidade de extrapolação do modelo (para o conjunto de teste) não são dependentes das específicas amostras estatísticas / dos dados empíricos coletados?

Se meus dados coletados fossem algo distintos dos que obtivemos, os valores das medidas não seriam diferentes?

31

© Prof. Emilio Del Moral – EPUSP

31

A resposta é SIM! Sim, se os dados empíricos coletados forem algo distintos, mesmo que vindos de observações do mesmo fenômeno, os valores das medidas de qualidade do regressor seriam diferentes!!!

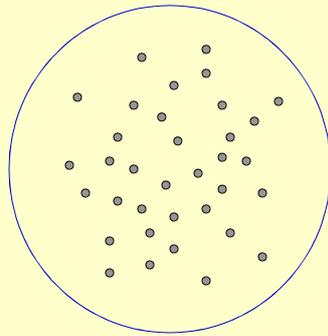
Com técnicas como a validação cruzada podemos avaliar a extensão da flutuação dessas medidas com as mudanças no conjunto de observações de treino e/ou teste (amostra estatística de dados empíricos para o treino e amostra estatística de dados empíricos para teste de generalização).

32

© Prof. Emilio Del Moral – EPUSP

32

Conjunto total de observações empíricas

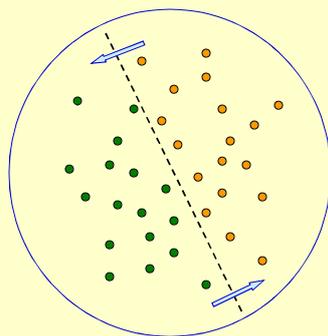


33

© Prof. Emilio Del Moral Hernandez

33

Cross validation / Validação cruzada ... e teremos muitos mais ensaios, se tivermos mudança na partição de observações



2 fold cross
validation:
50% treino e
50% teste

34

© Prof. Emilio Del Moral Hernandez

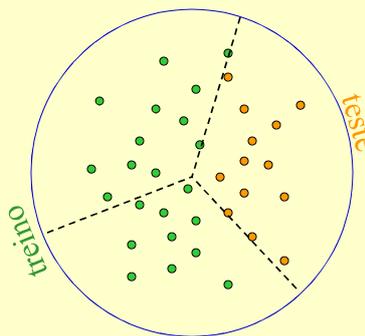
34

k-fold Cross Validation:
O conjunto total é “retalhado”
em k partes, e uma delas apenas
é usada para teste, com k
ensaios distintos

© Prof. Emilio Del Moral Hernandez

35

**Cross validation / Validação cruzada – k fold
cross validation**



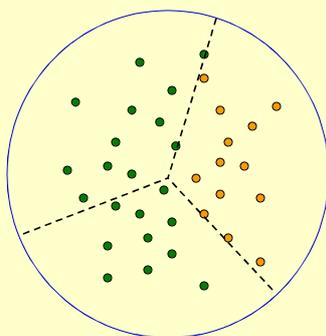
3 fold cross
validation:
66% treino e
33% teste

36

© Prof. Emilio Del Moral Hernandez

36

Cross validation / Validação cruzada e flutuação estatística na partição de amostras



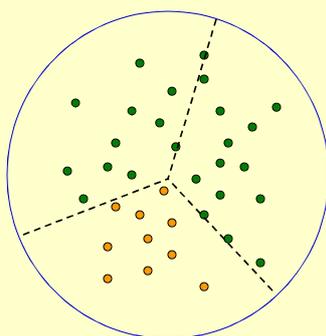
3 fold cross
validation:
66% treino e
33% teste

37

© Prof. Emilio Del Moral Hernandez

37

Cross validation / Validação cruzada e flutuação estatística na partição de amostras



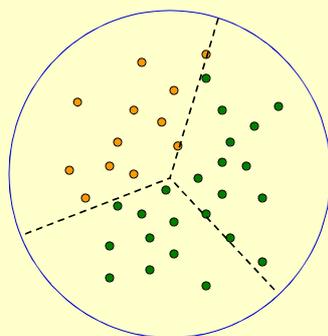
3 fold cross
validation:
66% treino e
33% teste

38

© Prof. Emilio Del Moral Hernandez

38

Cross validation / Validação cruzada e flutuação estatística na partição de amostras



3 fold cross
validation:
66% treino e
33% teste

39

© Prof. Emilio Del Moral Hernandez

39

Caso extremo do k fold cross validation ...

$$k = M_{\text{empíricos}}$$

(número de “partes da pizza” = número M de observações empíricas disponíveis, ou seja, cada parte da pizza tem um único exemplar / uma única observação)

Nesse caso, chamamos o método de “Leave one Out” – ou seja, usamos $(M_{\text{empíricos}} - 1)$ observações para treino do modelo e apenas UM exemplar empírico é deixado de fora do treino, para ser usado para o teste

40

© Prof. Emilio Del Moral Hernandez

40

*Conceito geral que engloba
as discussões anteriores*

...

*Reamostragem /
Data Resampling*

Pergunta ... Que impacto isto tem nas medidas de
qualidade de regressores? Que impacto isto tem nas
medidas de classificadores? Que informação ao cliente /
usuário podemos fornecer com base neste conceito?

41

© Prof. Emilio Del Moral Hernandez

41

42

Aula 13
(de 15)
do Prof. Emilio
de PSI3471 de
08-abr-2019
se encerrou aqui

© Prof. Emilio Del Moral Hernandez

42

42