

ERRATUM
ERRATUM

Santos HG, Nascimento CF, Izbicki R, Duarte YAO, Chiavegatto Filho ADP. *Machine learning* para análises preditivas em saúde: exemplo de aplicação para prever óbito em idosos de São Paulo, Brasil. Cad Saúde Pública 2019; 35(7):e00050818.

doi: 10.1590/0102-311XER050818

A revista foi informada sobre erros na Tabela 2. As correções seguem abaixo:

Onde se lê:

Tabela 2

Aprendizado (treinamento) e avaliação (teste) de modelos preditivos.

Algoritmos	Regressão logística	Regressão logística com penalização	Redes neurais	Gradient boosted trees	Random Forest
<i>TREINAMENTO</i> (resultados para o melhor modelo)					
Hiperparâmetros otimizados	-	$\alpha = 1$ $\lambda = 0,003$	$size = 3$ $decay = 2$	$nrounds = 100$ $maxdepth = 1$ $\eta = 0,3$	$mtry = 7$
AUC ROC	0,803 *	0,766 (0,07)	0,767 (0,06)	0,765 (0,07)	0,738 (0,05)
Média (dp) da validação cruzada					
<i>Ranking</i> de importância das variáveis					
1	Idade	Dificuldade para tomar banho	Idade	Idade	Idade
2	Consumo de tabaco	Idade	Dificuldade para tomar banho	Dificuldade para tomar banho	<i>Mini Exame do Estado Mental</i>
3	Dificuldade para tomar banho	Consumo de tabaco	Consumo de tabaco	Índice de massa corporal	Dificuldade para ir ao banheiro
4	Sexo	Dificuldade para comer	Dificuldade para ir ao banheiro	Sexo	Dificuldade para tomar banho
5	Diabetes mellitus	Sexo	Diabetes mellitus	<i>Mini Exame do Estado Mental</i>	Sexo
Teste					
AUC (IC95%)	0,773 (0,732; 0,814)	0,777 (0,735; 0,818)	0,779 (0,738; 0,820)	0,768 (0,724; 0,813)	0,744 (0,699; 0,789)
Pontos de corte para p (risco predito)					
$p = 0,5$					
S (VP)	0,144 (17)	0,130 (15)	0,08 (10)	0,144 (17)	0,110 (13)
$1 - E$ (FP)	0,03 (21)	0,026 (18)	0,020 (13)	0,026 (18)	0,019 (13)
p ótimo **					
S (VP)	0,653 (77)	0,670 (79)	0,703 (83)	0,712 (84)	0,700 (79)
$1 - E$ (FP)	0,251 (172)	0,253 (173)	0,300 (205)	0,304 (208)	0,328 (238)

(continua)



Tabela 2 (continuação)

Algoritmos	Regressão logística	Regressão logística com penalização	Redes neurais	<i>Gradient boosted trees</i>	<i>Random Forest</i>
<i>p</i> 10% com risco mais alto **					
<i>S</i> (VP)					
$1 - E$ (FP)					
Óbitos observados segundo o risco predito					
n (%) entre os 10% com risco mais alto	34 (28,814)	36 (30,508)	32 (27,119)	34 (28,814)	33 (27,966)
n (%) entre os 10% com risco mais baixo ***	-	-	-	-	-

dp: desvio padrão; E: especificidade; FP: falso positivo; IC95%: intervalo de 95% de confiança; S: sensibilidade; VP: verdadeiro positivo.

* AUC ROC do ajuste do modelo uma única vez aos dados de treinamento;

** Ponto de corte que maximiza a sensibilidade e a especificidade. Varia de acordo com o modelo preditivo;

*** Nenhum óbito observado entre os 10% de indivíduos com menor risco para o desfecho.

Leia-se:

Tabela 2

Algoritmos	Regressão logística	Regressão logística com penalização	Redes neurais	<i>Gradient boosted trees</i>	<i>Random Forest</i>
<i>TREINAMENTO</i> (resultados para o melhor modelo)					
Hiperparâmetros otimizados	-	$\alpha = 1$ $\lambda = 0,003$	$size = 3$ $decay = 2$	$nrounds = 100$ $maxdepth = 1$ $eta = 0,3$	$mtry = 7$
AUC ROC	0,803 *	0,766 (0,07)	0,767 (0,06)	0,765 (0,07)	0,738 (0,05)
Média (dp) da validação cruzada					
<i>Ranking</i> de importância das variáveis					
1	Idade	Dificuldade para tomar banho	Idade	Idade	Idade
2	Consumo de tabaco	Idade	Dificuldade para tomar banho	Dificuldade para tomar banho	<i>Mini Exame do Estado Mental</i>
3	Dificuldade para tomar banho	Consumo de tabaco	Consumo de tabaco	Índice de massa corporal	Dificuldade para ir ao banheiro
4	Sexo	Dificuldade para comer	Dificuldade para ir ao banheiro	Sexo	Dificuldade para tomar banho
5	Diabetes mellitus	Sexo	Diabetes mellitus	<i>Mini Exame do Estado Mental</i>	Sexo
Teste					
AUC (IC95%)	0,773 (0,732; 0,814)	0,777 (0,735; 0,818)	0,779 (0,738; 0,820)	0,768 (0,724; 0,813)	0,744 (0,699; 0,789)

(continua)

Tabela 2 (continuação)

Algoritmos	Regressão logística	Regressão logística com penalização	Redes neurais	Gradient boosted trees	Random Forest
Pontos de corte para p (risco predito)					
$p = 0,5$					
S (VP)	0,144 (17)	0,130 (15)	0,08 (10)	0,144 (17)	0,110 (13)
$1 - E$ (FP)	0,03 (21)	0,026 (18)	0,020 (13)	0,026 (18)	0,019 (13)
p ótimo **					
S (VP)	0,653 (77)	0,670 (79)	0,703 (83)	0,712 (84)	0,700 (79)
$1 - E$ (FP)	0,251 (172)	0,253 (173)	0,300 (205)	0,304 (208)	0,328 (238)
p 10% com risco mais alto					
S (VP)	0,288 (34)	0,305 (36)	0,271 (32)	0,288 (34)	0,280 (33)
$1 - E$ (FP)	0,067 (46)	0,064 (44)	0,070 (48)	0,066 (45)	0,066 (45)
Óbitos observados segundo o risco predito					
n (%) entre os 10% com risco mais alto	34 (28,814)	36 (30,508)	32 (27,119)	34 (28,814)	33 (27,966)
n (%) entre os 10% com risco mais baixo ***	-	-	-	-	-

dp: desvio padrão; E: especificidade; FP: falso positivo; IC95%: intervalo de 95% de confiança; S: sensibilidade; VP: verdadeiro positivo.

* AUC ROC do ajuste do modelo uma única vez aos dados de treinamento;

** Ponto de corte que maximiza a sensibilidade e especificidade. Varia de acordo com o modelo preditivo;

*** Nenhum óbito observado entre os 10% de indivíduos com menor risco para o desfecho.

Recebido em 25/Nov/2019
Aceito em 06/12/2019
Publicado em Jan/2020