

Pràctica resposta binària – Enunciat

Descripció

Poder anticipar l'aparició d'una malaltia d'un pacient és un dels objectius de la medicina: el pronòstic. En aquest cas, el nostre objectiu és ajudar a un grup d'investigació mèdic a identificar quins són els factors que poden anticipar l'aparició d'una malaltia coronària en base a dades observacionals recollides en el seu hospital.

Dades: S'entregarà un conjunt de dades amb registres de 270 pacients i 14 variables.

Variables explicatives

1. *age* (numeric)
2. *sex* (factor)
3. *chest_pain*: chest pain type (4 values) (factor)
4. *resting_BP*: resting blood pressure (numeric)
5. *serum_cholest*: serum cholesterol in mg/dl (numeric)
6. *blood_sugar*: fasting blood sugar > 120 mg/dl (factor)
7. *electro*: resting electrocardiographic results (values 0,1,2)
8. *HR*: maximum heart rate achieved (numeric)
9. *exercise*: exercise induced angina (factor)
10. *oldpeak*: ST depression induced by exercise relative to rest (numeric)
11. *ST*: the slope of the peak exercise ST segment (factor)
12. *major_vessels*: number of major vessels (0-3) colored by fluoroscopy (numeric)
13. *thal*: 3 = normal; 6 = fixed defect; 7 = reversible defect (factor)

Variable resposta

disease: Absence (1) or presence (2) of heart disease

Tasques

1. Fer una descriptiva univariada de les variables predictores i bivariada de cadascuna d'elles amb la variable resposta.
2. Fer depuració de les dades que es consideri adient.
3. Ajustar un model de resposta binària manualment amb la funció link LOGIT: començar des del model nul i anar afegint variables una a una (originals o transformades) de forma argumentada.
4. Ajustar un model amb un procediment automàtic (funció *step*). Compara'l amb el model anterior. Fes servir el criteri AIC o BIC i argumenta-ho perquè ho has fet servir.
5. Escull un dels dos models (pas 3 o 4) i avalua les possibles interaccions de primer ordre (entre dos factors o entre un factor i una covariable) i inclou aquelles rellevants en el model (si n'hi hagués)

6. Avaluar la bondat del model mitjançant algun dels estadístics explicats a teoria i de forma gràfica mitjançant un calibration plot. (podeu usar la funció [calibration.plot](#)). Si el model no es prou bo, es poden fer modificacions per intentar millorar-lo.
7. Un cop seleccionat un model, avaluar la seva capacitat predictiva emprant l'AUC.
8. Interpretar almenys 1 coeficient d'una variable numèrica i 1 d'una variable categòrica.
9. Al final de la pràctica, presentar:
 - a. Una taula resum que contingui el model final complet (sortida del R) i els principals indicadors del model: AIC, BIC, Deviança residual, Pseudo- R^2 , AUC i tots aquells que considereu rellevants.
 - b. Un paràgraf en que s'expliqui als metges les principals conclusions extretes de l'estudi.

Avaluació

Es valoraran els següents aspectes:

1. Criteris de selecció del model.
2. Criteris en la validació de la bondat de l'ajustament.
3. Capacitat predictiva assolida (AUC).
4. Interpretació dels resultats.
5. Presentació i claredat de l'informe.

Data d'entrega: La data límit per realitzar l'entrega serà el dia **20/11/18** a través del campus virtual.

Documentació i format de la documentació a entregar

1. **Una descripció del treball realitzat.** En format *word*, *pdf* o *html* (es pot fer amb *Rmarkdown*). S'ha de descriure els aspectes més rellevants dins de l'informe deixant el material suplementari pels annexos. Extensió de l'informe: 5 a 10 pàgines. Extensió màxima annexos: il·limitada.
2. **El codi** utilitzat per a realitzar aquesta part amb comentaris explicatius. Format: **.R** o **.Rmd**

Normes de format

1. Els noms complets dels components del grup han d'estar tant dins del fitxer que conté la descripció del treball com dins del codi d'anàlisi.
2. El format del nom dels dos fitxers com de la carpeta comprimida que els contingui ha de ser cognom1_cognom2.ext. P.ex: garcia_lopez.html, garcia_lopez.Rmd i garcia_lopez.zip
3. Només es farà **una entrega per grup** en una carpeta comprimida que contingui els dos fitxers.

L'incompliment d'aquestes normes comportarà una penalització a la nota.