

Bases de datos para practicar

Jorge Luis Villalba Acevedo

2021-02-18

Pulse

La tabla de datos pulse.txt que encontraréis en <http://aprender.uib.es/Rdir/pulse.txt> recoge una serie de informaciones de tipo general (altura, peso, sexo, si corren, si fuman y su nivel de actividad física: 1 si es bajo, 2 si es moderado y 3 si es alto) sobre algunos estudiantes matriculados en un curso de estadística de la Universidad Estatal de Pensilvania hace unos años. A estos estudiantes se les pidió que lanzasen una moneda al aire: a los que sacaron cara, se les hizo correr un minuto sin moverse del sitio, y los que sacaron cruz, descansaron un minuto. Todos los estudiantes (tanto los que corrieron como los que no) midieron sus pulsaciones por minuto antes y después de este minuto de ejercicio o descanso, y estas medidas también aparecen en esta tabla (en las variables PuBefor y PuAfter, respectivamente).

Bank

Los datos están relacionados con campañas de marketing directo de una institución bancaria portuguesa. Las campañas de marketing se basaron en llamadas telefónicas. A menudo, se requería más de un contacto con el mismo cliente, para acceder a si el producto (depósito a plazo bancario) estaría (o no) suscrito.

- El objetivo de la clasificación es predecir si el cliente suscribirá un depósito a plazo (variable y).

Input variables:

bank client data:

1 - age (numeric)

2 - job : type of job (categorical: “admin.”, “unknown”, “unemployed”, “management”, “housemaid”, “entrepreneur”, “student”, “blue-collar”, “self-employed”, “retired”, “technician”, “services”)

3 - marital : marital status (categorical: “married”, “divorced”, “single”; note: “divorced” means divorced or widowed)

4 - education (categorical: “unknown”, “secondary”, “primary”, “tertiary”)

5 - default: has credit in default? (binary: “yes”, “no”)

6 - balance: average yearly balance, in euros (numeric)

7 - housing: has housing loan? (binary: “yes”, “no”)

8 - loan: has personal loan? (binary: “yes”, “no”)

related with the last contact of the current campaign:

9 - contact: contact communication type (categorical: “unknown”, “telephone”, “cellular”)

10 - day: last contact day of the month (numeric)

11 - month: last contact month of year (categorical: “jan”, “feb”, “mar”, ..., “nov”, “dec”)

12 - duration: last contact duration, in seconds (numeric)

other attributes:

13 - campaign: number of contacts performed during this campaign and for this client (numeric, includes last contact)

14 - pdays: number of days that passed by after the client was last contacted from a previous campaign (numeric, -1 means client was not previously contacted)

15 - previous: number of contacts performed before this campaign and for this client (numeric)

16 - poutcome: outcome of the previous marketing campaign (categorical: “unknown”, “other”, “failure”, “success”)

Output variable (desired target):

17 - y - has the client subscribed a term deposit? (binary: “yes”, “no”)

Missing Attribute Values: None