

Caso Credit Scoring

CREDSCO

Karina Gibert Oliveras
Dep. Estadística e Investigación Operativa
Knowledge Engineering and Machine Learning group
Universitat Politècnica de Catalunya

Descripción:

Los datos responden a información sobre solicitudes de crédito a un banco.

Se aporta una muestra de 4455 clientes de los que se dispone de información sociodemográfica y económica, las características del crédito solicitado y el resultado de la evaluación.

Existen dos categorías de dictamen: positivo indica que el crédito se puede conceder, y negativo, que de va a denegar.

Se quieren entender las diferencias entre aquéllos que tienen dictamen positivo y los que no. Existen varias alternativas, pero para cualquiera de ellas, el primer paso es realizar una buena descriptiva y depuración de los datos.

Diccionario:

El valor 99999999 en las variables continuas indica valor faltante. El 0 lo indica en la categóricas.

Codificaciones:

dictamen final (variable de respuesta)

- 1 positivo
- 2 negativo

Antigüedad en el trabajo (años) (how many years in the current work?)

vivienda (kind of house)

- 1 alquiler (renting)
- 2 escritura publica (registered)
- 3 contrato privado (private contract)
- 4 ignora contrato (ignore contract)
- 5 padres (parents)
- 6 otros (other)

Plazo del préstamo (meses) (timing for credit)

Edad (age)

estado civil (civil status)

- 1 soltero (bachelor)
- 2 casado (married)

- 3 viudo (widow)
- 4 separado (separated)
- 5 divorciado (divorced)

registros (registers)

- 1 no
- 2 si

tipo de trabajo (work)

- 1 empleado fijo (indefinite)
- 2 empleado temporal (temporal)
- 3 autonomo (autonomous)
- 4 otros (other)

Gastos (Kpts) (Expenditures)

Ingresos (Kpts) (Incomes)

Patrimonio (Kpts)

Patrimonio por amortizar (Kpts) (part of the patrimonium financed and pending to pay)

Préstamo solicitado (Kpts) (Credit required)

Precio del bien financiado (Kpts) (Price of the target good)