

Marketing bancario

I dati sono relativi a campagne di marketing diretto (telefonate) di un istituto bancario portoghese. L'obiettivo della classificazione è di prevedere se il cliente sottoscriverà un deposito a termine (variabile y). Descrizione degli attributi:

Dati clienti bancari:

Età (numerico)

Lavoro: tipo di lavoro (categoriale: 'amministratore', 'operaio', 'imprenditore', 'cameriera', 'dirigente', 'pensionato', 'lavoratore autonomo', 'servizi', 'studente', 'tecnico', 'disoccupato', 'sconosciuto')

Stato civile: stato civile (categoria: 'divorziato', 'sposato', 'single', 'sconosciuto'; nota: 'divorziato' significa divorziato o vedovo)

Istruzione (categoria: 'basic.4y', 'basic.6y', 'basic.9y', 'high.school', 'analfabeta', 'professional.course', 'university.degree', 'sconosciuto')

Default: ha credito in default? (categoria: 'no', 'sì', 'sconosciuto')

Housing: ha mutuo per la casa? (categoria: 'no', 'sì', 'sconosciuto')

Prestito: ha prestito personale? (categoria: 'no', 'sì', 'sconosciuto')

Relativo all'ultimo contatto della campagna in corso:

Contatto: tipo di comunicazione di contatto (categoria: 'telefono cellulare')

Mese: ultimo mese di contatto dell'anno (categoria: 'jan', 'feb', 'mar', ..., 'novembre', 'dicembre')

Dayofweek: ultimo giorno di contatto della settimana (categoria: 'lun', 'mar', 'mer', 'gio', 'ven')

Durata: durata dell'ultimo contatto, in secondi (numerico).

Altri attributi:

Campagna: numero di contatti effettuati durante questa campagna e per questo cliente (numerico, include l'ultimo contatto)

Pdays: numero di giorni trascorsi dall'ultimo cliente contattato da una campagna precedente (numerico; 999 significa che il cliente non lo era precedentemente contattato)

Precedente: numero di contatti effettuati prima di questa campagna e per questo cliente (numerico)

Poutcome: esito della precedente campagna di marketing (categoria: 'fallimento', 'inesistente', 'successo')

Attributi del contesto sociale ed economico

Emp.var.rate: tasso di variazione dell'occupazione - indicatore trimestrale (numerico)

Cons.price.idx: indice dei prezzi al consumo - indicatore mensile (numerico)

Cons.conf.idx: indice di fiducia dei consumatori - indicatore mensile (numerico)

Euribor3m: tasso euribor 3 mesi - indicatore giornaliero (numerico)

Nr.occupati: numero dipendenti - indicatore trimestrale (numerico)

In particolare:

1. Analizzare i dati a vostra disposizione e correggere eventuali errori presenti nell'insieme di dati (dati mancanti, dati palesemente errati, etc...)
2. Individuare quali sono gli attributi che sembrano maggiormente correlati alla variabile target
3. Costruire almeno un paio di modelli previsivi, appartenenti a tipologie diverse (ad esempio alberi di classificazione e classificatori a regole)
4. Studiare l'accuratezza dei modelli ottenuti