Bank Marketing

Donated on 2/13/2012

The data is related with direct marketing campaigns (phone calls) of a Portuguese banking institution. The classification goal is to predict if the client will subscribe a term deposit (variab...

Dataset Characteristics	Subject Area	Associated Tasks
Multivariate	Business	Classification

Feature Type	# Instances	# Features
Categorical, Integer	45211	16

Dataset Information

Additional Information

The data is related with direct marketing campaigns of a Portuguese banking institution. The marketing campaigns were based on phone calls. Often, more than one contact to the same client was required, in order to access if the product (bank term deposit) would be ('yes') or not ('no') subscribed.

There are four datasets:
1) bank-additional-full.csv with all examples (41188) and 20 inputs, ordered by date (from May 2008 to November 2010), very close to the data analyzed in [Moro et al., 2014]
2) bank-additional.csv with 10% of the examples (4119), randomly selected from 1), and 20

Creators

S. Moro

P. Rita

P. Cortez

DOI

10.24432/C5K306

License

inputs.
3) bank-full.csv with all examples and 17 inputs, ordered by date (older version of this dataset with less inputs).
4) bank.csv with 10% of the examples and 17 inputs, randomly selected from 3 (older version of this dataset with less inputs).
The smallest datasets are provided to test more computationally demanding machine learning algorithms (e.g., SVM).

The classification goal is to predict if the client will subscribe (yes/no) a term deposit (variable y).

SHOW LESS

Has Missing Values?
No

Introductory Paper

A data-driven approach to predict the success of bank telemarketing
By Sérgio Moro, P. Cortez, P. Rita. 2014
Published in Decision Support Systems

Variables Table

Variable Name	Role	Type	Demographic	Description
age	Feature	Integer	Age	
job	Feature	Categorical	Occupation	type of job (categorical: 'admin.','blue-coll

Variable Name	Role	Type	Demographic	Description
				employed','services','student','technician','
marital	Feature	Categorical	Marital Status	marital status (categorical: 'divorced','marr widowed)
education	Feature	Categorical	Education Level	(categorical: 'basic.4y','basic.6y','basic.9y','high.school','
default	Feature	Binary		has credit in default?
balance	Feature	Integer		average yearly balance
housing	Feature	Binary		has housing loan?
loan	Feature	Binary		has personal loan?
contact	Feature	Categorical		contact communication type (categorical:
day_of_week	Feature	Date		last contact day of the week

Rows per page 10 0 to 10 of 17

Additional Variable Information

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")

SHOW LESS

Papers Citing this Dataset

SORT BY YEAR, DESC

[Fair Algorithms for Clustering]
By Suman Bera, Deeparnab Chakrabarty, Nicolas Flores, Maryam Negahbani. 2019
Published in ArXiv.

[Clustering with Fairness Constraints: A Flexible and Scalable Approach]
By Imtiaz Ziko, Eric Granger, Jing Yuan, Ismail Ayed. 2019
Published in ArXiv.

[Noise-tolerant fair classification]
By Alexandre Lamy, Ziyuan Zhong, Aditya Menon, Nakul Verma. 2019
Published in ArXiv.

[AdaFair: Cumulative Fairness Adaptive Boosting]
By Vasileios Iosifidis, Eirini Ntoutsi. 2019
Published in

[Quantification under prior probability shift: the ratio estimator and its extensions]
By Afonso Vaz, Rafael Izbicki, Rafael Stern. 2018
Published in ArXiv.

Rows per page 5 0 to 5 of 9

Reviews

4 (1 rating)

LOGIN TO WRITE A REVIEW

甯文駿 智慧商務系

Nice

3/1/2024