



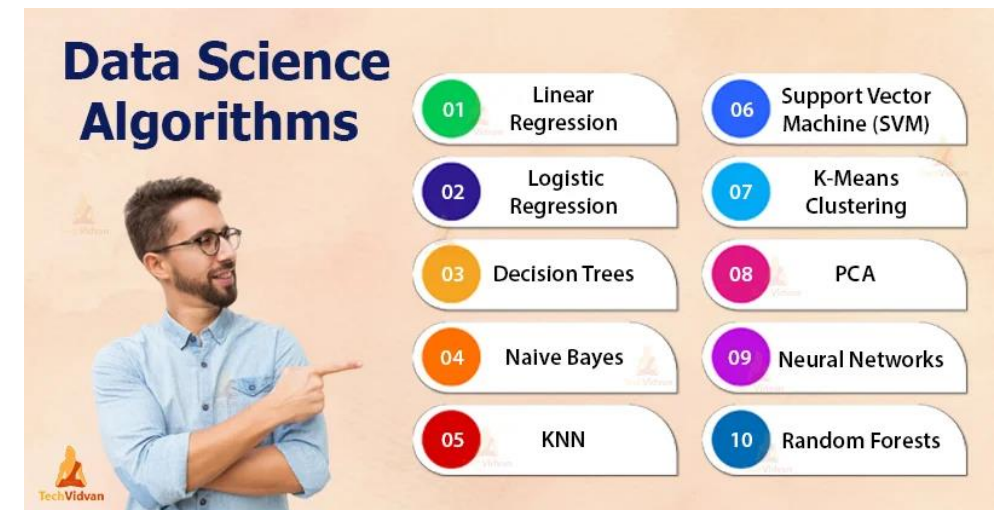
# MACHINE LEARNING APLICADO A MARKETING BANCARIO

- Carlos Garrido
- Iván García
- Adriana Pérez
- Lucas Fernández

# 1. Descripción del Caso

- Crisis de 2010
- Producto de alto valor bidireccional (depósito bancario)
- Gran impacto económico

- Caso transversal
- Algoritmos comunes
- Aplicable a muchos tipos de situaciones (ventas, supervivencia, fuga de talento)



## 2. Dataset

- Caso real, muy documentado y estudiado.
- Alternativas en la resolución. Ejemplos resueltos.
- Tuvo trascendencia académica como “Data mining”.
- 50.000 registros. Didáctico en el dimensionado y muestreo.

Journal of Service Science and Management > Vol.13 No.3, June 2020

### Bank Telemarketing Forecasting Model Based on t-SNE-SVM

Jianguo Che<sup>1</sup>, Sai Zhao<sup>1</sup>, Yongfan Li<sup>2</sup>, Kai Li<sup>1</sup>

<sup>1</sup>Business School, Nankai University, Tianjin, China.

<sup>2</sup>Mechanical Electrical Engineering School, Beijing Information Science & Technology University, Beijing, China.

DOI: 10.4236/jssm.2020.133029 PDF HTML XML 564 Downloads 2.446 Views Citations

Abstract

#### Variables de entrada:

```
# datos del cliente del banco:
1 - edad (numérico)
2 - trabajo: tipo de trabajo(categorico:"administrador","desconocido","desempleado", "gerencia", "criada",
"empresario"," estudiante", "cuello azul", "trabajador por cuenta propia", "jubilado", "técnico", "servicios").
3 - civil: estado civil (categorico: "casado", "divorciado", "soltero"; nota: "divorciado" significa divorciado o
viudo)
4 - educación (categorica: "desconocida", "secundaria", "primaria", "terciaria")
5 - mora: ¿tiene crédito en mora? (binario: "sí","no")
6 - saldo: saldo medio anual, en euros (numérico)
7 - vivienda: ¿tiene préstamo vivienda? (binario: "sí","no")
8 - préstamo: ¿tiene préstamo personal? (binario: "si","no")
```

#### Relacionado con el último contacto de la campaña actual:

```
9 - contacto: tipo de comunicación del contacto (categorico: "desconocido","teléfono","celular")
10 - día: último contacto día del mes (numérico)
11 - mes: último contacto mes del año (categorico: "ene", "feb", "mar", ..., "nov", "dec")
12 - duración: duración del último contacto , en segundos (numérico)

#### Otros atributos:
13 - campaña: número de contactos realizados durante esta campaña y para este cliente (numérico, incluye
último contacto)
14 - pdays: número de días que pasaron después de que el cliente fue contactado por última vez desde un
campaña anterior (numérico, -1 significa que el cliente no fue contactado previamente)
15 - anterior: número de contactos realizados antes de esta campaña y para este cliente (numérico)
16 - resultado: resultado de la campaña de marketing anterior (categorico: "desconocido",
otro","fracaso","éxito")
```

#### Variable de salida (objetivo deseado):

17 - y - ¿el cliente ha suscrito un depósito a plazo? (binario: "sí","no")

# 3. Estructura

## 0. PASOS PREVIOS

### 0.1 Librerías

### 0.2 Funciones

## 1. INTRODUCCIÓN

### 1.1. Descripción del caso

### 1.2. Descripción del DATASET

### 1.3. Objetivos

#### 1.3.1. Conceptuales y Estratégicos

#### 1.3.2. Técnicos

### 1.4. Autores del estudio

## 2. CARGA DE LOS DATOS

## 3. EXPLORATORY DATA ANALYSIS (EDA)

### 3.1 Análisis de nulos

### 3.2 Análisis de duplicados

### 3.3 Análisis de Variables

#### 3.2.1 Variables categóricas

#### 3.2.3 Variables numéricas (Tratamiento Outliers)

### 3.4 Conclusiones

## 4. TRANSFORMACIÓN DE DATOS

### 4.1 Variables categóricas a numéricas

### 4.2 Variables numéricas

### 4.3 Unión de variables

## 5. MODELO DE MACHINE LEARNING (ARBOL)

### 5.1 Selección de variables (Reducción de Dimensionalidad)

#### 5.1.2 Variables de mayor importancia