

Taller #2: Bank-Marketing-Data

Instrucciones:

- Fecha de publicación: 05 de Abril de 2025
- Fecha de entrega: 26 de Abril 2025
- Medio de entrega: Sustentación del proyecto final plataforma.

Conjunto de datos: Contiene datos sobre campañas de marketing directo (llamadas telefónicas) de una entidad bancaria portuguesa.

<https://archive.ics.uci.edu/dataset/222/bank+marketing>

Ejercicio:

- A. Realizar un análisis exploratorio al conjunto de datos entregado, para ello debe crear diferentes preguntas e hipótesis a resolver en los datos partiendo de un problema que cada grupo debe plantear.
- B. Después de realizar el análisis exploratorio, debe hacer el preprocesamiento de los datos según como considere: limpieza, transformación, reducción de datos o discretización de los datos.
- C. Debe realizar un entrenamiento del modelo de machine learning.

Algunos algoritmos de clustering:

- K-Means-Clustering.
- DBSCAN (Density-Based Spatial Clustering of Applications with Noise).
- Hierarchical Clustering.
- Gaussian Mixture Model (GMM).
- Principal Component Analysis (PCA).

Entregables:

- Análisis exploratorio en notebook (10 puntos).
- Preprocesamiento de los datos (15 puntos).
- Entrenamiento del Modelo de Machine Learning con redes neuronales (10 puntos).
- Evaluación del modelo 75% Accuracy (10 puntos).

Taller #2: Bank Marketing Data.

Bank Marketing Data: <https://github.com/dg2c4/Bank-Marketing-Big-Data>

Problema:

El objetivo de las campañas es convencer a los clientes en suscribirse al producto “depósito a plazo fijo”. De este modo, lo que se quiere resolver es:

Pregunta Problema:

¿Cuáles son los patrones de comportamiento de los clientes que se suscriben, con el fin de mejorar las estrategias de marketing bancario?

Objetivo:

Desarrollar un análisis completo que permita extraer datos sobre campañas de marketing directo (llamadas telefónicas) de una entidad bancaria portuguesa, utilizando herramientas de visualización de datos para presentar los resultados de manera efectiva.

Análisis Exploratorio:

- Formulación de preguntas de investigación.
- Creación de hipótesis basadas en un problema específico.
- Análisis inicial del conjunto de datos seleccionado.

Preprocesamiento de Datos:

- Limpieza de datos.
- Transformación de variables.
- Reducción de datos cuando sea necesario.
- Discretización de datos según corresponda.

Modelado y Evaluación:

- Implementación de modelo de regresión lineal o logística.
- Selección y cálculo de métricas de evaluación.
- Análisis del rendimiento del modelo machine learning.

Systems engineering:

User: David Gutierrez Chaves

Code: 506222728

Subject: Big Data Electiva-I

Institution: Fundación Universitaria Konrad Lorenz

Institutional email: david.gutierrec@konradlorenz.edu.co

Created By: <https://github.com/dg2c4>