Aprendizaje Automático 1

Trabajo Práctico

El dataset “credit\_japan\_ejercicio.csv” es real y cuenta con variables tanto numéricas como categóricas. En este caso sólo nos encargaremos de Imputar Nulos, Codificar las variables Categóricas y Normalizar las numéricas. La última variable, la 16 no debe modificarse porque es la variable target.

En este trabajaremos sólo sobre el dataset y no generaremos train y test sets.

Imputación:

* Las variables numéricas se imputarán con kNN usando 5 vecinos.
* Las variables categóricas se imputarán con la moda.

Luego

* A las variables categóricas aplicarles OneHotEncoder con una columna por valor
* las variables numéricas se deberán ser normalizadas al intervalo entre 0 y 1.

-Cree el ColumnTransformer que efectúe todas estas operaciones, entrénelo sobre todo el dataset y luego aplíquelo al mismo dataset.

-Al dataset obtenido guárdelo en un DataFrame de Pandas y denomínelo df\_transformado.