Entregable: Definición de la Base de Datos.

1. Clasificación de billetes falsos.

Se extrajo la información de fotografias tomadas de billetes reales y falsificaciones. Para digitalizarlas se utilizó una cámara industrial. Las imagenes finales son de 400 x 400 pixeles; debido a la lente usada y la distancia al objetivo se obtuvieron imagenes en escala de grises con una resolución de 660 dpi. Para obtener los atributos de las imagenes se usaron herramientas de transformación ondícula.

Atributos:

1. Variance of Wavelet Transformed image (continuous) – Varianza de la imagen transformada.
2. Skewness of Wavelet Transformed image (continuous) – Sesgo o asimetria\* de la imagen transformada.
3. Curtosis of Wavelet Transformed image (continuous) – Curtosis\*\* de la imagen transformada.
4. Entropy of image (continuous) – Entropia de la imagen.
5. Class (integer) – Clase.

\*: Asimetría: La asimetría es la medida que indica la simetría de la distribución de una variable respecto a la media aritmética, sin necesidad de hacer la representación gráfica.

\*\*: Curtosis: La curtosis o medida de apuntamiento es una medida estadística que determina el grado de concentración que presentan los valores de una variable alrededor de la zona central de la distribución de frecuencias.

URL: [https://archive.ics.uci.edu/ml/datasets/banknote+authentication#](https://archive.ics.uci.edu/ml/datasets/banknote+authentication)

1. Clasificación de clientes de banco.

Debido a una baja en las ganancias causado por la baja de inversiones en depósitos a largo plazo, un banco portugués se propuso identificar los clientes que tienen chances más altas de pedir depósitos a largo plazo de forma de focalizar sus esfuerzos de marketing en ellos.

Atributos:

1. Age (numeric) - Edad.
2. Job (categorical,nominal) - Trabajo.
3. Marital (categorical,nominal) - Estado civil.
4. Education (categorical,nominal) - Educación.
5. Default (categorical,nominal) - Crédito en default.
6. Housing (categorical,nominal) – Prestamo hipotecario.
7. Loan (categorical,nomina) – Prestamos personales.
8. Contact (categorical,nomina) – Forma de contacto.
9. Month (categorical,nomina) – Ultimo mes de contacto.
10. Dayofweek (categorical,nominal) – Ultimo día de contacto.
11. Duration (numeric) – Duración del último contacto.
12. Campaign (numeric) – Número de contactos durante una campaña.
13. Pdays (numeric) – Número de días desde el último contacto.
14. Previous (numeric) – Número de contactos durante la campaña anterior.
15. Poutcome (categorical,nominal) – Rédito de la última campaña.

Output deseado:

1. Y (binario) – Se suscribió el cliente a un préstamo a largo plazo?

URL: <https://www.kaggle.com/datasets/rashmiranu/banking-dataset-classification>

1. Base de Datos de Diabetes – India.

Esta base de datos corresponde al Instituto Nacional de Diabetes y Enfermedades Digestivas y de Riñón. El objetivo es predecir si el paciente tiene o no diabetes basandose en ciertas medidas de diagnóstico.

Atributos:

1. Pregnancies – Número de embarazos.
2. Glucose – Concentración de glucose plasmática.
3. Diastolic blood presure – Presón sanguínea diastólica.
4. Triceps skin fold thickness – Grosor de la piel del triceps.
5. 2-Hour serum insulin – Nivel de insulina en suero a las 2 horas.
6. Body mass index – Indice de masa corporal.
7. Diabetes pedigree function – Función de pedigree de diabetes.
8. Age – Edad.
9. Outcome – Resultado.

https://www.kaggle.com/datasets/uciml/pima-indians-diabetes-database