

**AA02  
Trabajo Práctico II  
Individual**

**Proyecto de clasificación empleando aprendizaje automático.**

**Objetivo:** El objetivo de este trabajo práctico es familiarizarse con el proceso de clasificación utilizando las técnicas de aprendizaje automático vistas, y construir un modelo que permita clasificar los billetes en verdaderos o falsos.

**1.- Conjunto de datos:** "data\_banknote\_authentication.txt".

URL: <https://archive.ics.uci.edu/ml/datasets/banknote+authentication#>

**2.- Información del conjunto de datos:**

Los datos se extrajeron de imágenes que se tomaron de especímenes similares a billetes genuinos y falsificados. Para la digitalización se utilizó una cámara industrial que se suele utilizar para la inspección de impresiones. Las imágenes finales tienen 400x 400 píxeles. Debido a la lente del objeto y la distancia al objeto investigado, se obtuvieron imágenes en escala de grises con una resolución de aproximadamente 660 dpi. La herramienta Wavelet Transform se utilizó para extraer características de las imágenes.

Información de atributos:

1. varianza de la imagen Wavelet transformada (continua)
2. asimetría de la imagen Wavelet transformada (continua)
3. curtosis de la imagen Wavelet Transformada (continua)
4. entropía de la imagen (continua)
5. clase (entero)

**3.- Se solicita que realice:**

**3.1 Preprocesamiento de datos:** Determinar si es necesario el preprocesamiento de datos y si es así generar los datos limpios.

**3.2 Visualización de datos.**

**3.3 Entrenamiento y evaluación del modelo.** Dividir el conjunto de datos en entrenamiento y test.

**3.4 Emplear uno o varios algoritmos y determinar cuál se posee un mejor rendimiento con nuestro problema según sus resultados.**

**3.5 Selección del modelo a emplear.**

**AA02  
Trabajo Práctico II  
Individual**

- 3.6 Ajuste y optimización del modelo.** Se pide que ajusten los hiperparámetros del modelo para mejorar su rendimiento. Pueden realizar una búsqueda de cuadrícula (grid search) o utilizar técnicas de optimización como RandomizedSearchCV para encontrar la mejor combinación de hiperparámetros.
- 3.7 Evaluación del modelo:** Evaluación del modelo entrenado utilizando métricas de evaluación de clasificación, como precisión, recall, F1-score y matriz de confusión. Se pide que se interpreten los resultados y analicen las fortalezas y debilidades del modelo.
- 3.8 Mejoras y reflexiones.** Proponga posibles mejoras o expansiones al modelo actual. Pueden discutir sobre técnicas avanzadas de aprendizaje automático y cómo podrían aplicarse a este problema.