

# Reto: Predicción de Enfermedades del Corazón con TensorFlow/Keras

Este reto tiene como objetivo que los estudiantes construyan, entrenen y evalúen una red neuronal para predecir la presencia de enfermedades cardíacas utilizando el dataset **Heart Disease (OpenML id=61)**. Los participantes deberán desarrollar el código, analizar los resultados y explicar los conceptos teóricos en cada paso.

- 1 **Instalación y carga de librerías:** Instalar y cargar TensorFlow, scikit-learn, pandas y matplotlib. Explicar: ¿Para qué sirve cada librería en este proyecto? ¿Por qué necesitamos TensorFlow y Keras?
- 2 **Carga del dataset Heart Disease desde OpenML:** Usar `fetch_openml` para descargar el dataset 'heart-disease'. Convertir la variable objetivo a valores binarios (0 = sano, 1 = enfermedad). Explicar: ¿Qué variables contiene este dataset? ¿Cuál es la variable objetivo?
- 3 **Preprocesamiento de datos:** Dividir los datos en entrenamiento (80%) y prueba (20%). Escalar características con `StandardScaler`. Explicar: ¿Por qué es importante escalar los datos? ¿Qué significa hacer fit en train y transform en test?
- 4 **Construcción de la Red Neuronal:** Crear un modelo `Sequential` con la arquitectura: `Dense(32, relu) → Dense(16, relu) → Dense(1, sigmoid)`. Explicar: ¿Por qué usamos relu en capas ocultas y sigmoid en la salida? ¿Qué representan los parámetros entrenables?
- 5 **Entrenamiento del modelo:** Usar `binary_crossentropy` como pérdida y `Adam` como optimizador. Entrenar 100 épocas con `batch_size=16`. Incluir `EarlyStopping`. Explicar: ¿Qué significa una época? ¿Qué hace `EarlyStopping` y por qué es útil?
- 6 **Evaluación del modelo:** Calcular Accuracy, AUC y Matriz de confusión. Explicar: ¿Qué significan TP, TN, FP y FN? ¿Por qué AUC es importante en problemas médicos?
- 7 **Predicciones con un nuevo paciente:** Simular un registro de paciente y predecir su probabilidad de enfermedad. Explicar: ¿Qué significa que el modelo devuelva valores cercanos a 0.9 o 0.1? ¿Cómo interpretar este resultado en un contexto clínico?
- 8 **Visualización de resultados:** Graficar curva ROC, curva de pérdida y curva de AUC. Explicar: ¿Qué indica que la curva de validación se separe mucho de la de entrenamiento? ¿Qué significa un AUC cercano a 0.5 vs cercano a 1?

**Entregable:** Un notebook o script .py que siga los 8 pasos, incluya explicaciones teóricas en cada punto y muestre tablas y gráficas (matriz de confusión, curva ROC, curvas de entrenamiento).