Objetivo del proyecto: Predicción de cardiopatía dadas las características de los pacientes

Contexto: Data set con valores de pacientes registrados en hospitales de Cleveland, Hungría y Suiza.

**Fuente de los datos**

UCI Machine Learning Repository

**Datos clave según la OMS (Organización Mundial de la Salud)**

* Las enfermedades cardiovasculares (ECV) son la principal causa de muerte en el mundo.
* Se estima que 17,9 millones de personas murieron por ECV en 2019, lo que representa el 32% de todas las muertes mundiales. De estas muertes, el 85% se debieron a infarto de miocardio y accidente cerebrovascular.
* Más de tres cuartas partes de las muertes por ECV se producen en países de ingresos bajos y medios.
* De los 17 millones de muertes prematuras (menores de 70 años) debidas a enfermedades no transmisibles en 2019, el 38% fueron causadas por ECV.
* La mayoría de las enfermedades cardiovasculares pueden prevenirse abordando los factores de riesgo conductuales, como el consumo de tabaco, la dieta poco saludable y la obesidad, la inactividad física y el consumo nocivo de alcohol.
* Es importante detectar las enfermedades cardiovasculares lo antes posible para poder empezar a tratarlas con asesoramiento y medicamentos.

**Diccionario**

1. age: edad en años
2. sex: genero
   * 1 = hombre
   * 0 = mujer
3. cp: tipo de dolor toraxico
   * Value 0: angina típica
   * Value 1: angina no típica
   * Value 2: dolor no anginoso
   * Value 3: asintomatico
4. trestbps: tensión arterial en reposo (en mm Hg al ingreso en el hospital)
5. chol: colesterol sérico en mg/dl
6. fbs: (fasting blood sugar >(glucemia en ayunas > 120 mg/dl):

- 1 = verdadero

- 0 = falso

1. restecg: resultados electrocardiográficos en reposo

- Valor 0: normal

- Valor 1: con anomalía de la onda ST-T (inversiones de la onda T y/o elevación o depresión del ST de > 0,05 mV)

- Valor 2: hipertrofia ventricular izquierda probable o definida según los criterios de Estesthalach: maximum heart rate achieved

1. exang: angina inducida por el ejercicio

- 1 = sí

- 0 = no

1. oldpeak: = depresión del ST inducida por el ejercicio en relación con el reposo

La abreviatura "ST" se refiere a los segmentos ST en un electrocardiograma (ECG o EKG), que es una representación gráfica de la actividad eléctrica del corazón.

1. slope: la pendiente del segmento ST de ejercicio máximo

- Valor 0: pendiente ascendente

- Valor 1: plano

- Valor 2: pendiente descendente

1. ca: número de vasos principales (0-3) coloreados por flouroscopia
2. thal:

- 0 = error (en el conjunto de datos original 0 corresponde a NaN)

- 1 = defecto fijo

- 2 = normal

- 3 = defecto reversible

13. target (the lable):

- 0 = ninguna enfermedad,

- 1 = enfermedad

**Diagnóstico de cardiopatía (estado angiográfico de la enfermedad).**

**Valor 0: < 50% de estrechamiento del diámetro**

**Valor 1: > 50% de estrechamiento del diámetro**