



UANL

AMA a tu CORAZÓN

EQUIPO 5: Quiroz García V. Martínez Loera M. Rodríguez Charles K.



1

Introducción



Las **enfermedades cardíacas** son la principal causa de defunción en todo el mundo.

Cada año mueren más personas por alguna de estas enfermedades que por cualquier otra causa.

Este indicador nos servirá para obtener información sobre la importancia que tienen los diferentes síntomas o pruebas clínicas para el diagnóstico de este tipo de enfermedades.

2

Objetivos

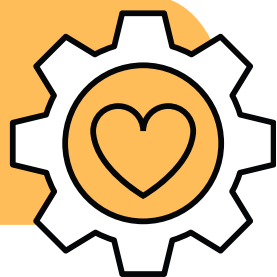


Principal. Determinar cuales son los indicadores que tienen mayor importancia para el diagnóstico de una enfermedad cardíaca

Secundario. Saber si alguno de estos indicadores tienen mayor importancia en su diagnóstico a comparación de los restantes.

3

Recursos



Librerías utilizadas



Base de datos

Enlace código

Enlace imágenes



4

Metodología



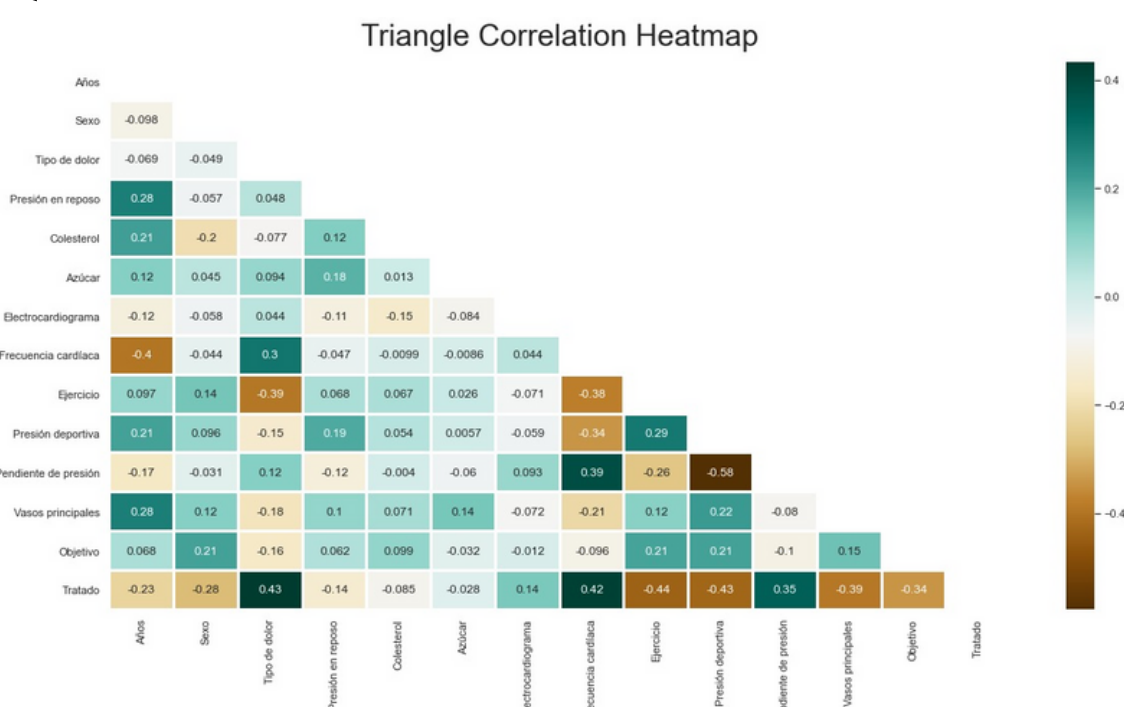
La **base de datos** representa a pacientes que fueron a los hospitales de Cleveland por algún síntoma o creencia de padecimiento de una enfermedad cardiovascular.

En la limpieza de datos se revisaron las columnas para saber si existían datos nulos; hicimos un cambio de nombre en las columnas, ya que se encontraron ciertas abreviaciones en ellas.

Años	Sexo	Tipo de dolor	Presión en reposo	Colesterol	Azúcar	Electrocardiograma	Frecuencia cardíaca	Ejercicio	Presión deportiva	Pendiente de presión	Vasos principales	Objetivo	Tratado
0	63	1	3	145	233	1	0	150	0	2.3	0	0	1
1	37	1	2	130	250	0	1	187	0	3.5	0	0	2
2	41	0	1	130	204	0	0	172	0	1.4	2	0	2
3	56	1	1	120	236	0	1	178	0	0.8	2	0	2
4	57	0	0	120	354	0	1	163	1	0.6	2	0	2

Tabla 1. Base de datos.

Se utilizó la gráfica 1 para revisar la relación que tienen las variables entre sí.



Gráfica 1. Triángulo de correlación

5

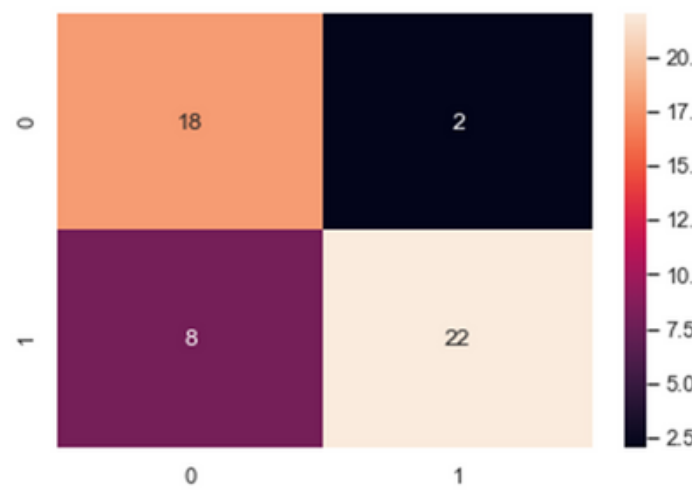
Resultados



Para elaborar la técnica de **árbol de decisión** se verificaron los siguientes componentes:

Métricas de evaluación

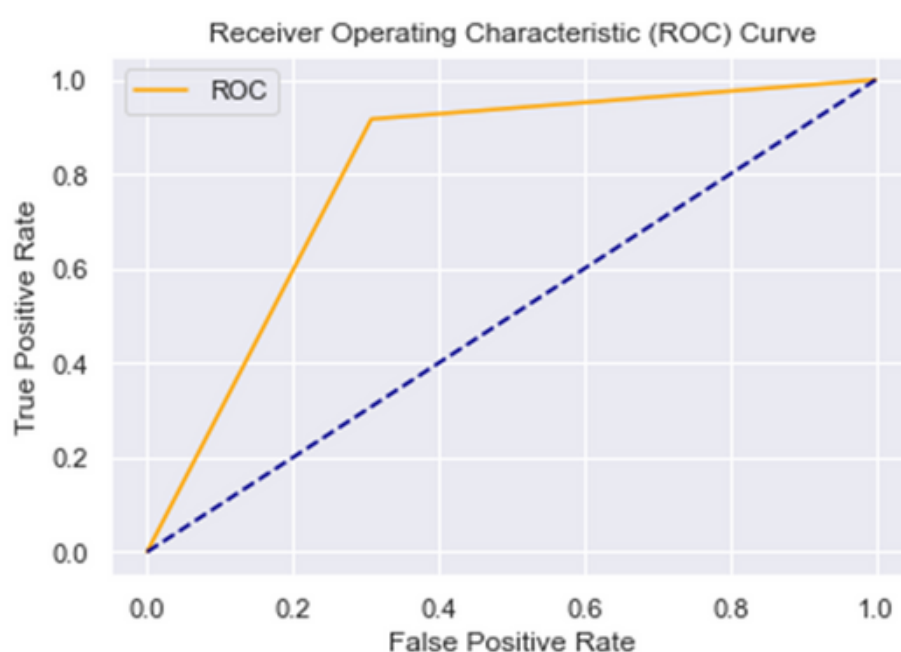
La gráfica 2 es una herramienta de medición que proporciona información sobre la precisión de las predicciones.



Gráfica 2. Matriz de confusión

Cuanto mayor sea el área cubierta bajo la curva ROC, mejores serán los modelos.

En la gráfica 3 representa el área de nuestros datos con un AUC (Área bajo la curva) de 80.44871.



Gráfica 3. AUC-ROC

Árbol de decisión

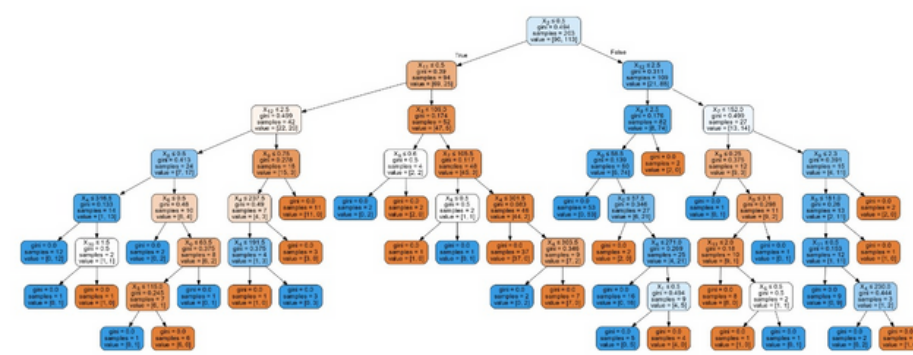
Condición: nodo donde se toma alguna decisión.

Gini: medida de impureza.

Samples: número de muestras que satisfacen las condiciones necesarias para llegar a este nodo.

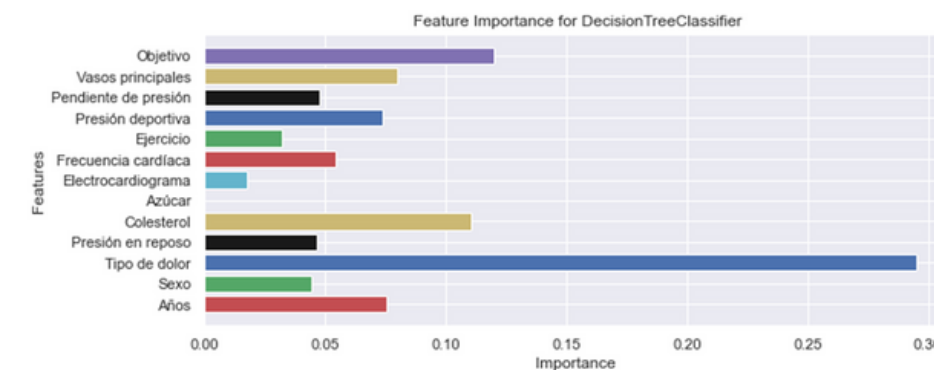
Value: cuántas muestras de cada clase llegan a este nodo.

Class: qué clase se le asigna a las muestras que llegan a este nodo.



Gráfica 4. Árbol de decisión

La gráfica 5 representa la importancia de cada indicador de las pruebas en los hospitales de Cleveland y también observamos el indicador de mayor relevancia.



Gráfica 5. Importancia de factores

6

Conclusión



- Las enfermedades cardíacas son la principal causa de defunción en todo el mundo. Cada año mueren más personas por alguna de estas enfermedades que por cualquier otra causa.
- Las enfermedades cardiovasculares afectan en mucha mayor medida a los países de ingresos bajos y medios y afectan casi por igual a hombres y mujeres.
- Se pronóstica que de aquí a 2030, casi 23.6 millones de personas morirán por alguna enfermedad cardiovascular.
- Con esta información nos puede servir principalmente para crear un indicador que nos de información de la importancia que tienen los diferentes síntomas o pruebas clínicas para el diagnóstico de una enfermedad cardiovascular.