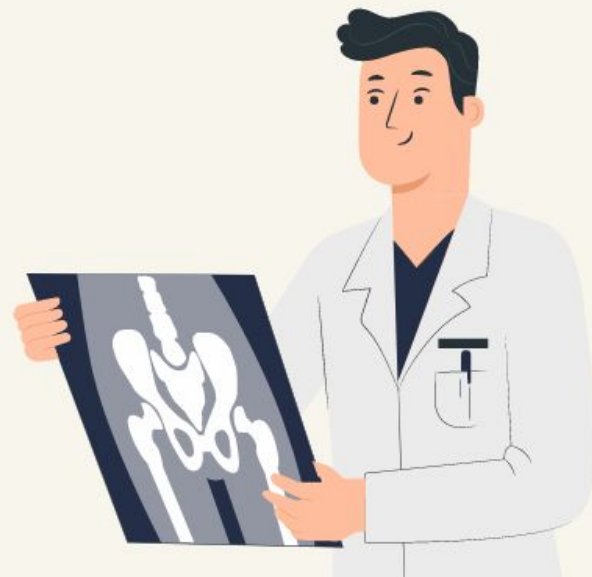


Proyecto Final Aprendizaje Máquina

Mónica Lara Pineda
Manuel Ortiz Hernández
Omar Alejandro Robledo Rodríguez
Diego Ramírez Levy
Raúl González Cardona



Contenido

01

Objetivo

Enfoque completo de modelación de problemas

02

Exploración de datos

Explicación del conjunto de datos a utilizar

03

Selección de modelos

Explicación de la metodología de aprendizaje automático

04

Conclusiones

Resolución sobre desarrollo de modelo ML



Enfermedades Cardiovasculares!



Afectación cardíaca que se manifiesta mediante vasos sanguíneos enfermos, problemas estructurales y coágulos sanguíneos.

01

Ejemplo de enfermedades

Una enfermedad cardiovascular engloba varios problemas del corazón



Ejemplos



Arterias coronarias

Lesión en los principales vasos sanguíneos



Ritmo cardíaco

Impulsos eléctricos que no funcionan adecuadamente



Defectos cardíacos de nacimiento

Defectos congénitos



Válvulas cardíacas

Válvulas no funcionan adecuadamente



Músculo cardíaco

Dificultad del corazón para transportar sangre



Infección del corazón

Infección del revestimiento interior del corazón

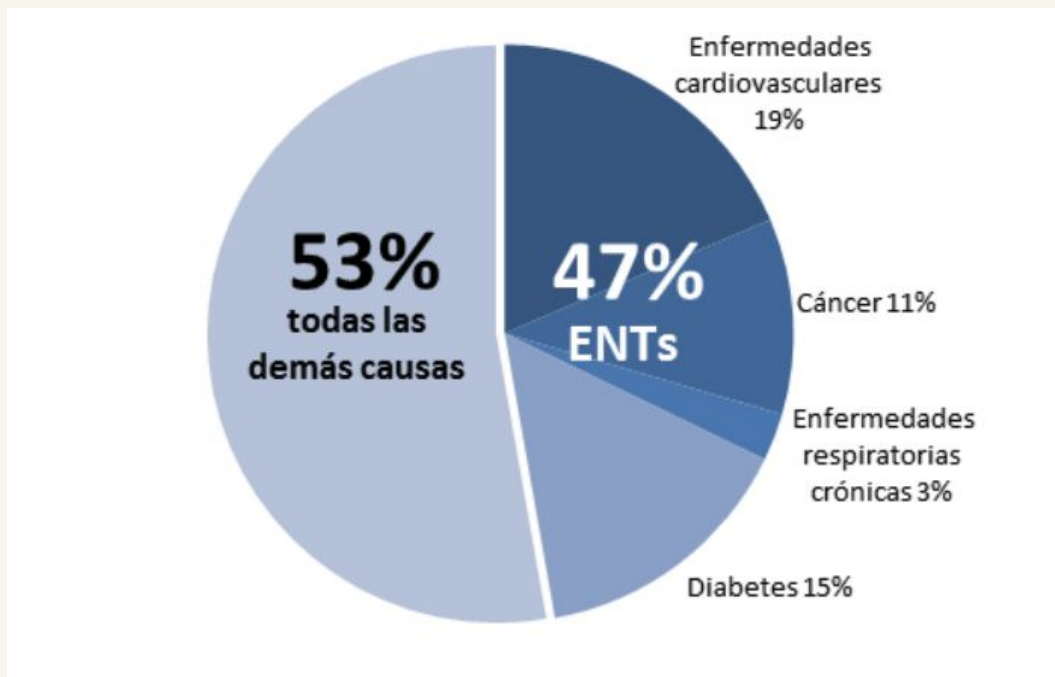


02

Datos en México

Las enfermedades del corazón son la 1era causa de muerte en nuestro país

Causas de muerte en México (2010-2020)



Datos enfermedades relacionadas al corazón





105,485

personas que fallecen al año en México por enfermedades cardiovasculares

Objetivo modelo



El objetivo es poder establecer una **opción viable** para aquellos individuos que lleguen a ser diagnosticados con cualquier enfermedad relacionada al corazón

Exploración de datos



	Age	Sex	ChP nTy	Res BP	Cho lest.	Fast .BS	Rest .EC G	Max HR	Exe r.An.	Old pea k	ST_ Slop e	Hea rtDi s.
0	40	M	ATA	140	289	0	Norm al	172	N	0.0	Up	0
1	49	F	NAP	160	180	0	Norm al	156	N	1.0	Flat	1
2	37	M	ATA	130	283	0	ST	98	N	0.0	Up	0
3	3	48	F	ASY	138	214	0	Norm al	108	Y	1.5	Flat
4	4	54	M	NAP	150	195	0	Norm al	122	N	0.0	Up



Limpieza de datos

Se ejecutó una limpieza debido a que algunos valores no permiten trabajar con el modelo de forma correcta

Variables afectadas



Sex



Chest
Pain Type



RestingE
CG



ExerciseA
ngina



ST_Slope

03

Selección de modelos

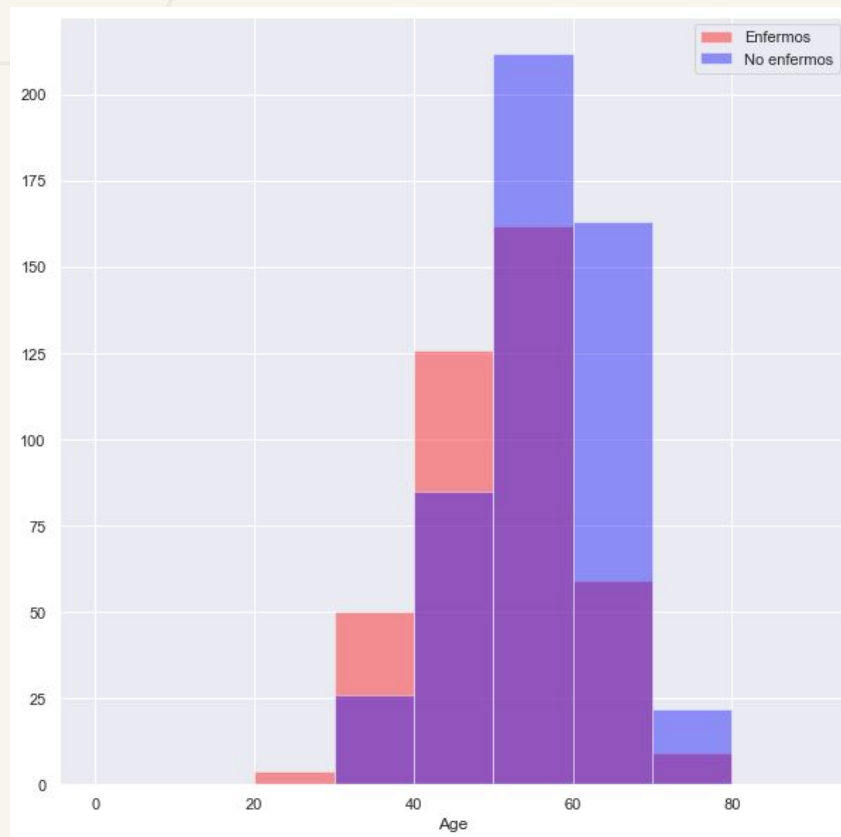
En esta sección se muestra la metodología usada para determinar el modelo que se va a usar.

Se realizaron los siguientes pasos:

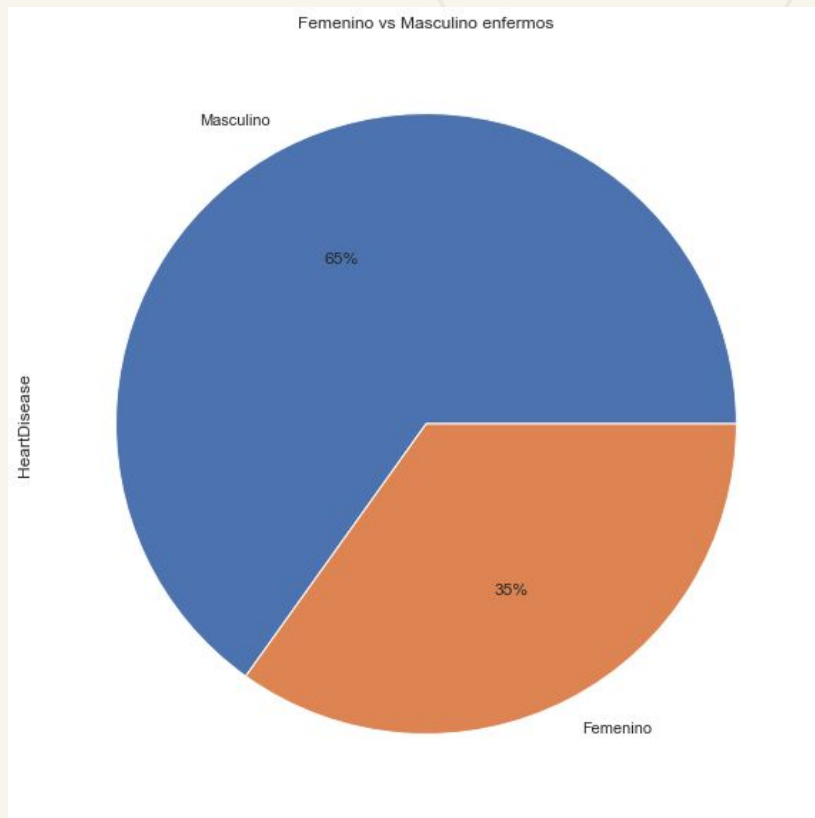
- Separación de variables independientes
- Gráfica relación Enfermos vs Edad
- Gráfica relación Sexo vs Enfermos
- Gráfica relación nivel Colesterol
- División de datos 80/20



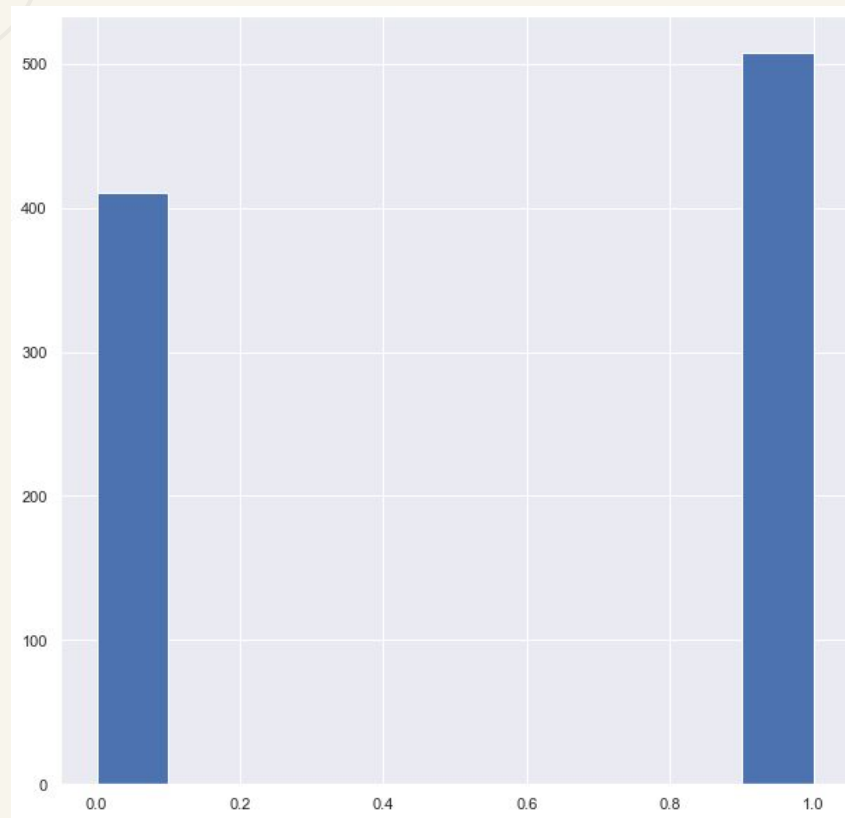
Gráfica relación Enfermos vs Edad



Gráfica relación Sexo vs Enfermos



Gráfica relación nivel Colesterol



Random Forest

Precisión del modelo

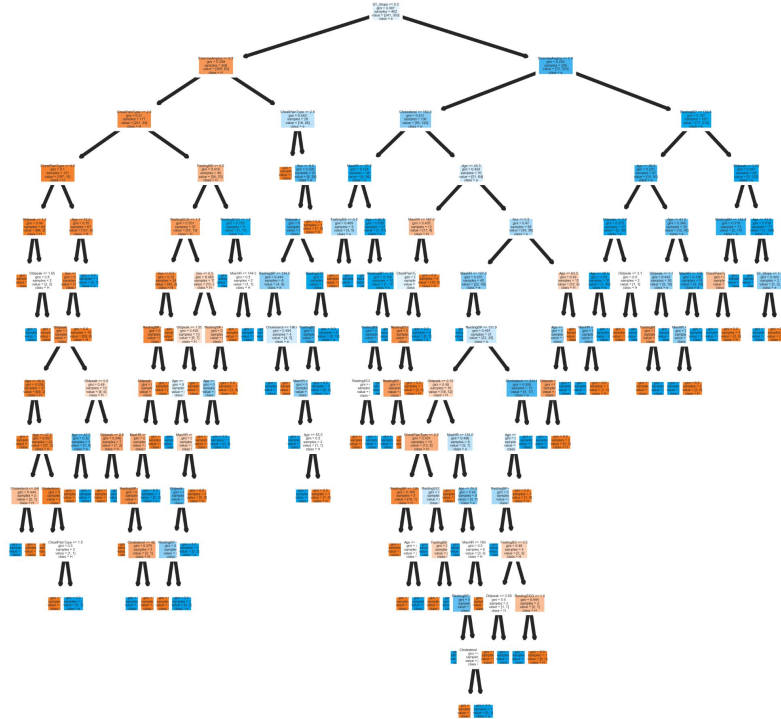
	Accuracy on Train	Accuracy on Text
0	1.0	0.880435

Matriz de confusión

399	0
0	335

Para la primera parte del notebook, se decidió utilizar Random Forest. Se tienen 399 pacientes en total que tenían la enfermedad y la falla del corazón, en donde 335 tenían la enfermedad y en realidad no contaban con falla del corazón.

Demostración visual



Se presenta el recorrido que hace el modelo por las variables iniciales que denotan si el paciente tiene una falla del corazón o no,

04

Conclusiones

Al finalizar este proyecto nos dimos cuenta que existen muchas herramientas de aprendizaje de máquina que nos pueden ayudar que nos faciliten tareas importantes como lo son los diagnósticos médicos.

