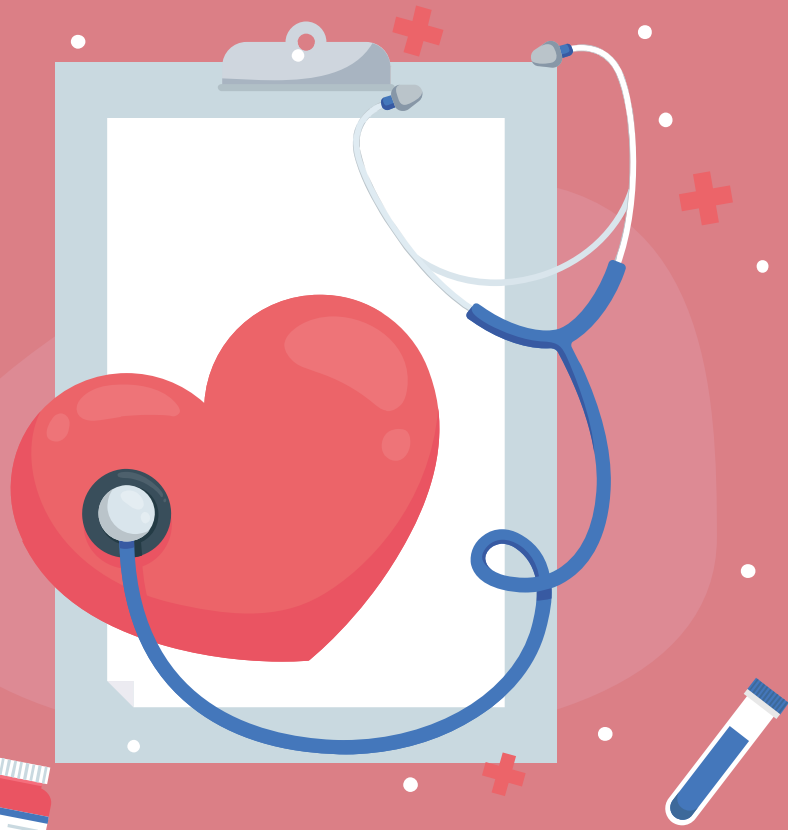


H-FAILURE

Predicciones sobre la insuficiencia cardíaca

Gisela criollo suarez



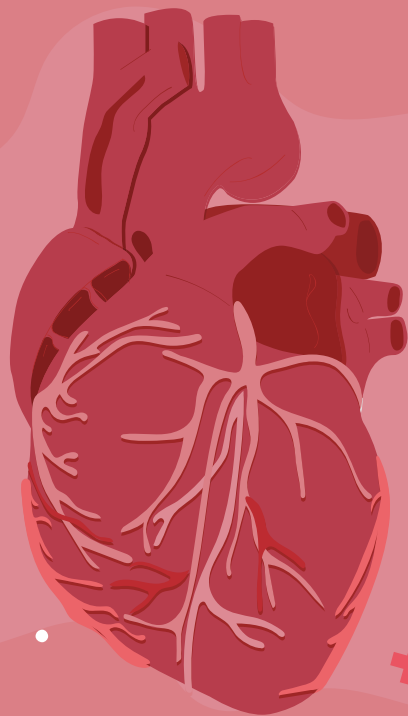


La insuficiencia cardíaca es una enfermedad progresiva crónica en la que el músculo cardíaco no puede bombear suficiente sangre para satisfacer las necesidades de glóbulos rojos y oxígeno del cuerpo.

Básicamente, el corazón no puede seguir el ritmo de trabajo, debido a la presencia de uno o más factores de riesgo como hipertensión, diabetes o enfermedad ya establecida, necesitan una detección y manejo precoces donde un modelo de aprendizaje automático puede ser de gran ayuda.

Por ello se ha desarrollado H-FAILURE, un proyecto cuyo objetivo es determinar la probabilidad de padecer insuficiencia cardíaca.





• **Objetivo general**

- Comprender y analizar los datos presentados para realizar una buena predicción de la insuficiencia cardiaca.

Objetivos específicos

- Llevar a cabo el análisis y lectura del DataSet.
- Realizar diagramas que permitan un mejor entendimiento del DataSet.
- Aplicar los conocimientos adquiridos durante el bootcamp.



Metodologia



El dataSet cuenta con un total de 299 filas que corresponden a las personas que hicieron parte de la recolección de los datos.

age	anaemia	creatinine_phosphokinase	diabetes	ejection_fraction	high_blood_pressure	platelets	serum_creatinine	serum_sodium	sex	smoking	time
75.0	0	582	0	20	1	265000.00	1.9	130	1	0	4
55.0	0	7861	0	38	0	263358.03	1.1	136	1	0	6
65.0	0	146	0	20	0	162000.00	1.3	129	1	1	7
50.0	1	111	0	20	0	210000.00	1.9	137	1	0	7
65.0	1	160	1	20	0	327000.00	2.7	116	0	0	8
90.0	1	47	0	40	1	204000.00	2.1	132	1	1	8
75.0	1	246	0	15	0	127000.00	1.2	137	1	0	10
60.0	1	315	1	60	0	454000.00	1.1	131	1	1	10
65.0	0	157	0	65	0	263358.03	1.5	138	0	0	10
80.0	1	123	0	35	1	388000.00	9.4	133	1	1	10

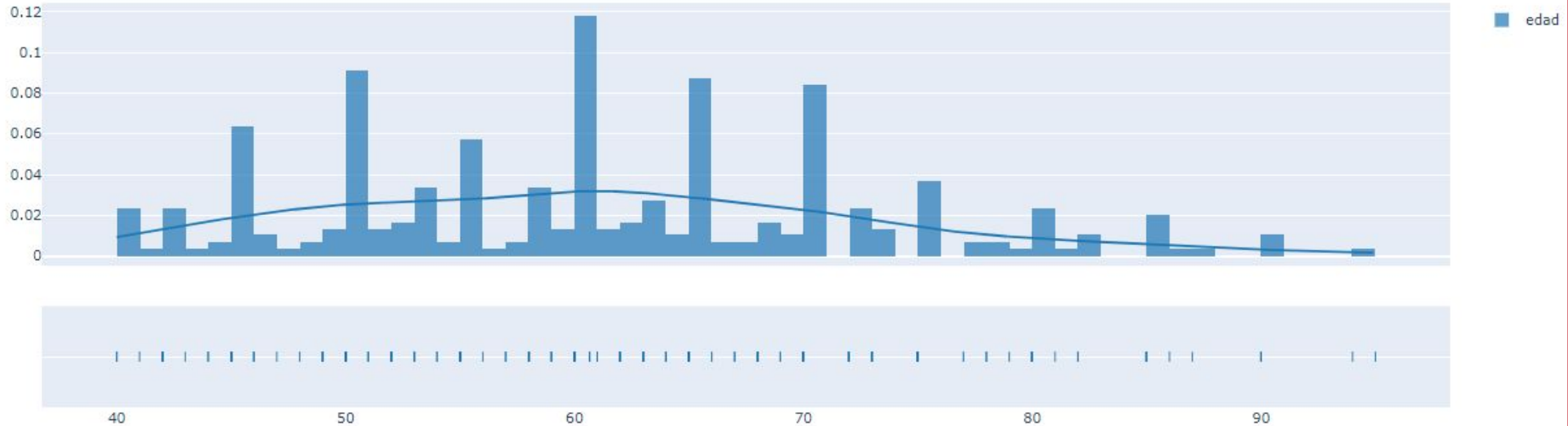
Para H-FAILURE se utilizaron un total de 12 características para predecir la mortalidad por insuficiencia cardiaca.

El dataSet se dividió en 80 % entrenamiento y 20% para prueba.

ANÁLISIS DE EDAD

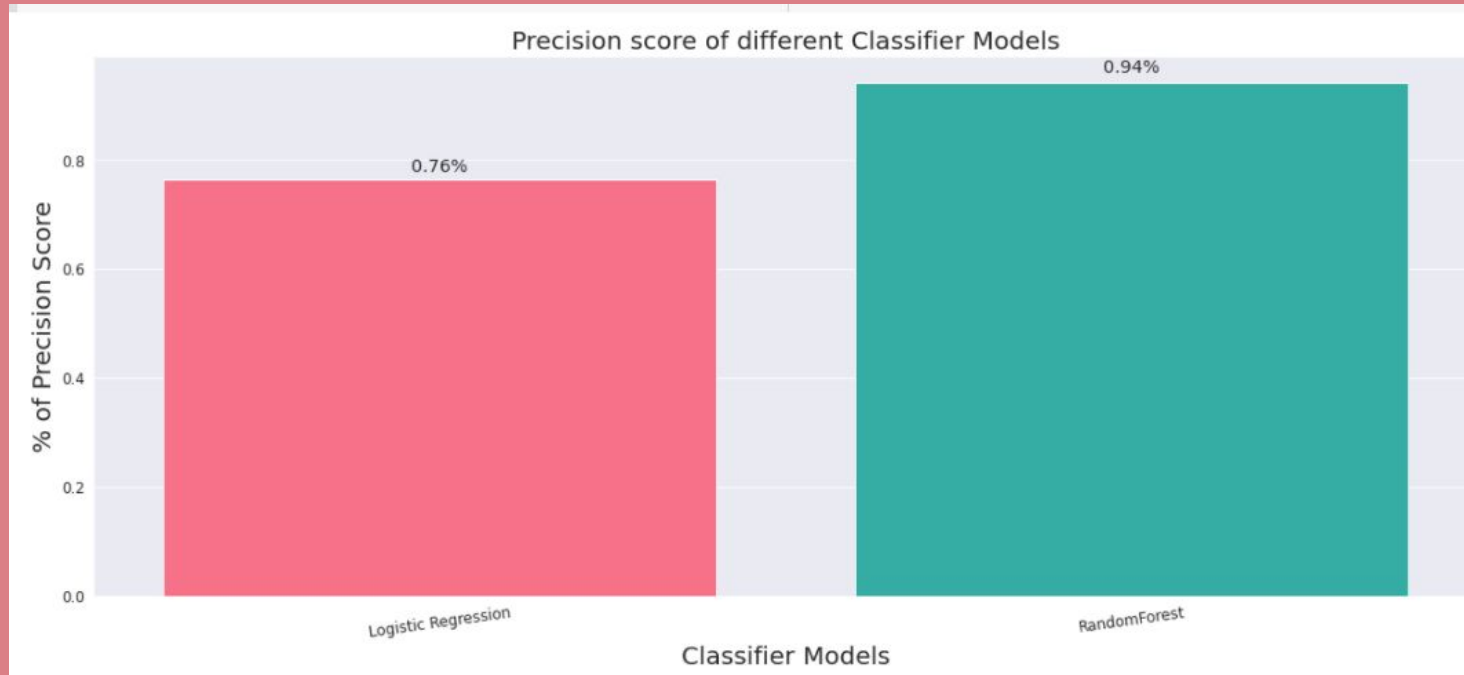
En este caso, podemos observar que en un rango de 40 a 80 años el indicativo de un evento de muerte es más alto, en cambio las personas menores a 40 años y mayores a 80 tienen un indicativo de muerte más bajo.

Grafico de distribucion de edad



Resultados

Realizamos el entrenamiento de dos modelos, de los cuales, “Logistic Regression” logró un 76% de precisión en la predicción y “RandomForest” logró un 94% de precisión ante el evento de muerte.



Conclusiones



1

En nuestro análisis de datos encontramos que en el informe de edad y género, la propagación de supervivencia es alta en la edad de 40 a 70, también que la supervivencia es alta tanto para hombres de 50 a 60 años como para mujeres de 60 a 70 años respectivamente.

2

Los modelos de entrenamiento nos indican que 14 personas mueren por insuficiencia cardiaca, mientras que 46 no.

3

Durante todo el bootcamp se adquieren habilidades que nos permiten realizar el análisis de los datos, a comprender y a llevar a cabo el entrenamiento de modelos para la predicción de sucesos.

Referencias

- Dataset tomado de: <https://www.kaggle.com/andrewmvd/heart-failure-clinical-data>

- Código fuente en:

https://github.com/GiselaCS/Mujeres_Digitales/blob/1946448fa4915f99ba4a5a06908b47bf4ff06133/H_FAILURE,_Gisela_Criollo.ipynb

