



# Clasificación de ondas de electrocardiograma

Cynthia Cristal Quijas Flores

A01655996

Alejandro Sánchez Flores

A01662783

Carlos Adrián Palmieri Álvarez

A01635776

Dabria Camila Carrillo Meneses

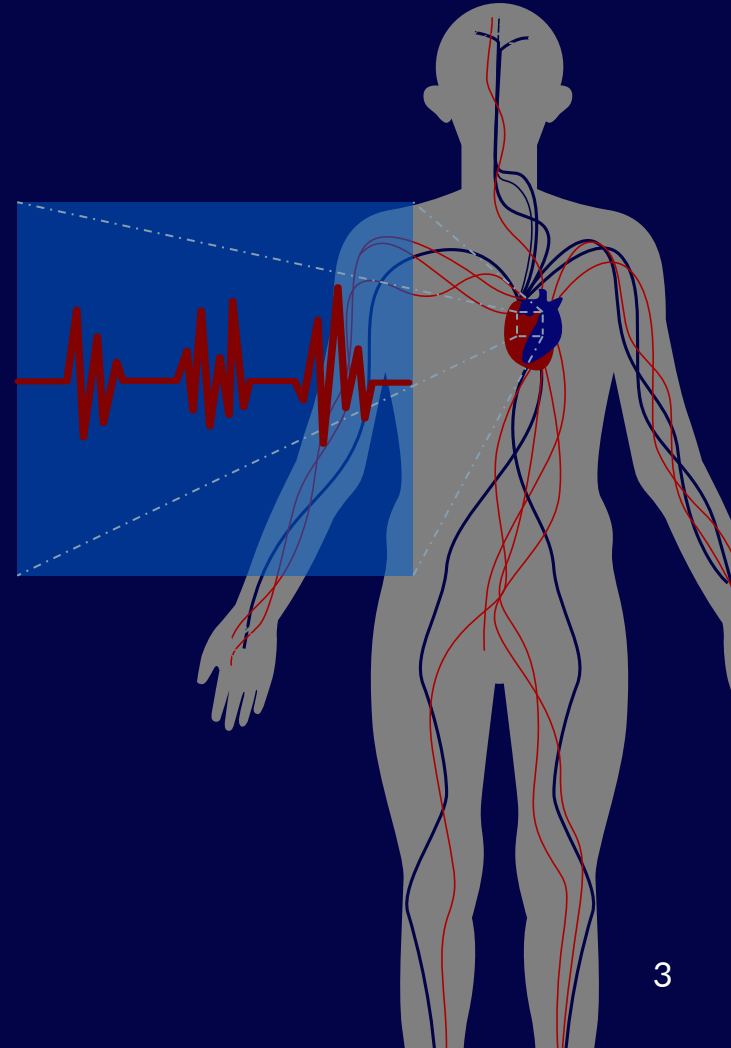
A01656716



01

# Entendimiento del negocio

Detección de Anomalías Cardíacas a  
través de Señales de Electrocardiogramas



# Tipos de ondas cardiacas

00

## Normal

Latidos cardíacos sanos.

01

## SUPRAVENTRICULAR

Latidos cardíacos ectópico  
supraventriculares.

02

## VENTRICULAR

Latidos cardíacos ectópicos  
ventriculares.

03

## FUSION

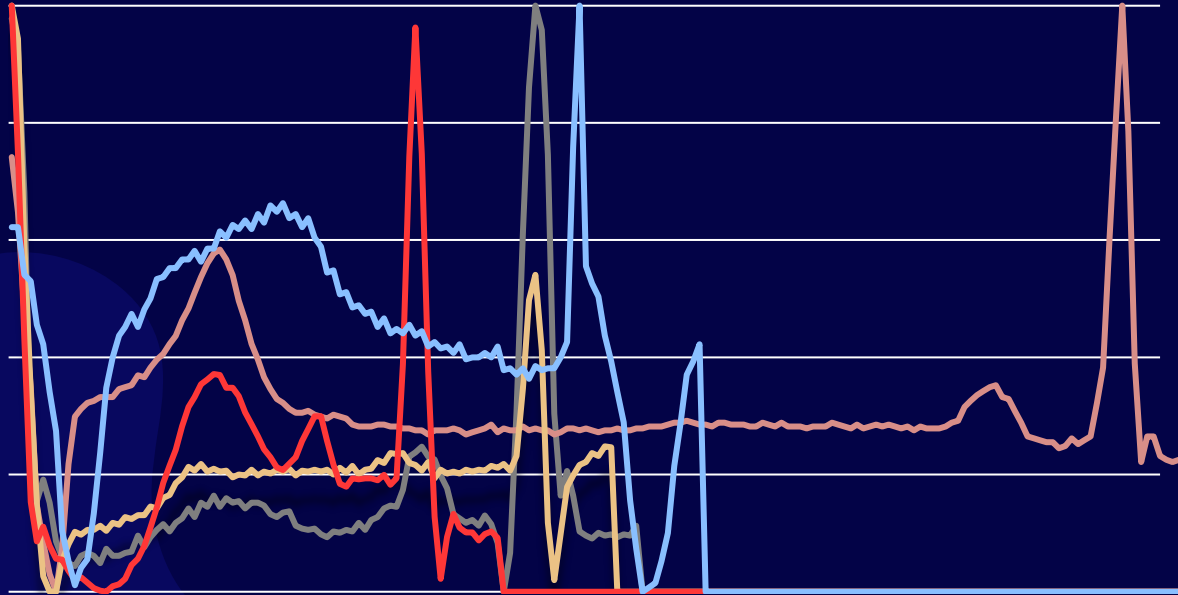
Latidos cardíacos de fusión.






04

## DESCONOCIDO

Latidos cardíacos que no pertenecen a  
ninguna de las categorías anteriores.

## Electrocardiograma de las clases



Categoría	Anotaciones
 N	<ul style="list-style-type: none"><li>• Normal</li><li>• Bloqueo de rama izquierda/derecha</li><li>• Escape auricular</li><li>• Escape nodal</li></ul>
 S	<ul style="list-style-type: none"><li>• Auricular prematuro</li><li>• Auricular aberrante prematuro</li><li>• Prematuro ganglionar</li><li>• Prematuro supraventricular</li></ul>
 V	<ul style="list-style-type: none"><li>• Contracción ventricular prematura</li><li>• Escape ventricular</li></ul>
 F	<ul style="list-style-type: none"><li>• Fusión de la función ventricular y normal</li></ul>
 Q	<ul style="list-style-type: none"><li>• Ritmo</li><li>• Fusión de ritmo y normalidad</li><li>• Inclasificable</li></ul>

Se grafica un electrocardiograma de una persona perteneciente a cada clase (onda) del dataset.

A stylized, thick red line representing an ECG (heart rate) waveform, positioned behind the text. The background is a dark blue gradient with abstract, wavy shapes.

¿Por qué es importante el  
análisis de los  
**Electrocardiogramas?**

# Situación en México

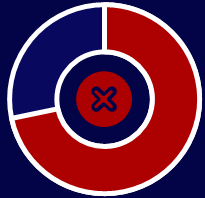


**19%**

## Fallecidos

Mujeres y Hombres  
de 30 a 69 años

(Sánchez & Bobadilla  
& Dimas, 2016)



**70.3%**

## Riesgo

Al menos un factor  
de riesgo  
cardiovascular

(Lara & Rosas & Pastelín & Aguilar  
& Attie & Velázquez, 2004)



# CAUSAS DE ESTE PADECIMIENTO

(Medina, 2010)

1,300

## TABAQUISMO

Casos anuales de anomalías atribuidas

1 de cada 3

## ESTRÉS

Casos reportados por centros urbanos con alto estrés laboral

(UNAM Global, 2023).

(Instituto de Seguridad y Servicios Sociales de los Trabajadores del Estado, 2016)

70%

## OBESIDAD

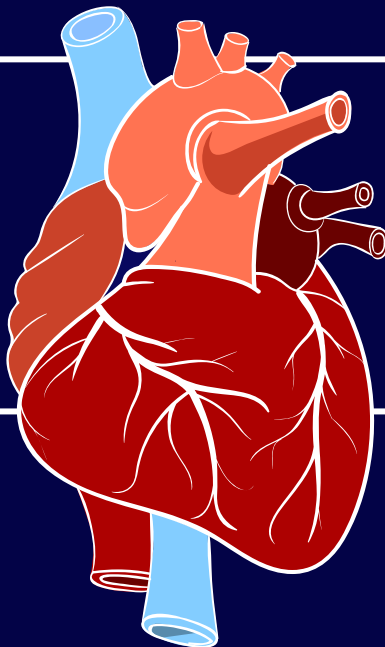
Mayor riesgo de latidos ventriculares irregulares

14%

## DIABETES

Desarrollan arritmias cardíacas

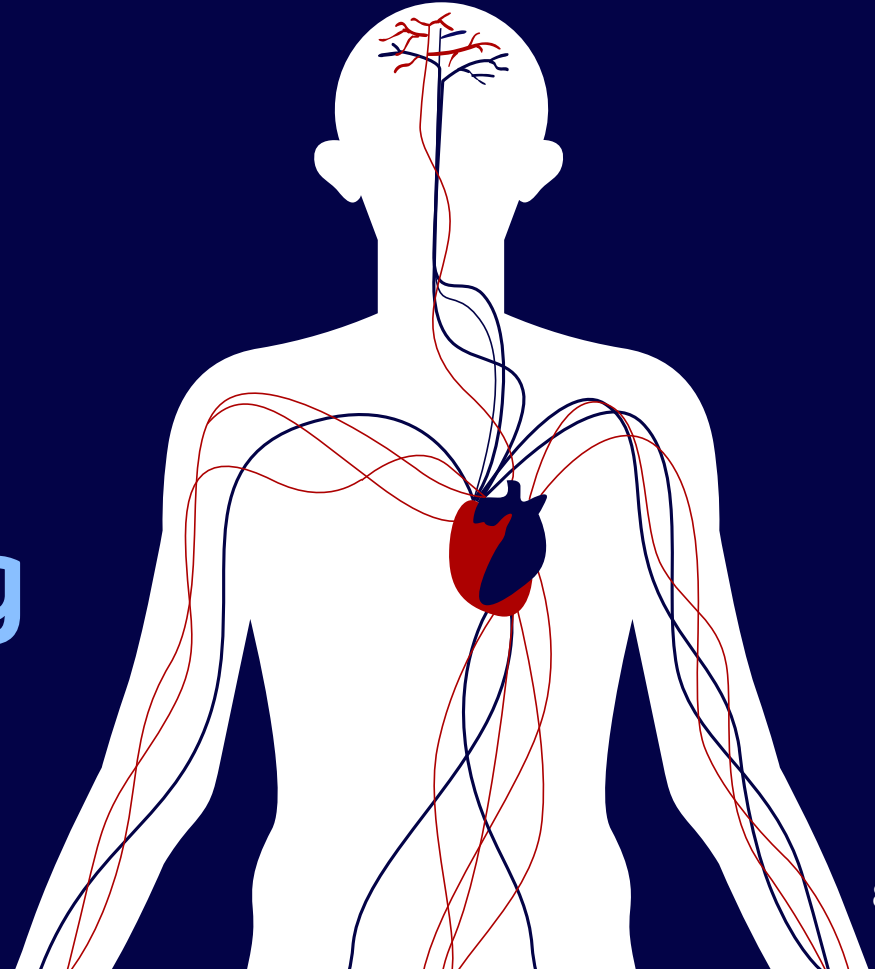
(Secretaría de Salud, 2022)





02

# Data understanding



# Características del dataset

109,446

Muestras de latidos  
cardiacos



Representados en una  
amplitud de

187

valores de tiempo

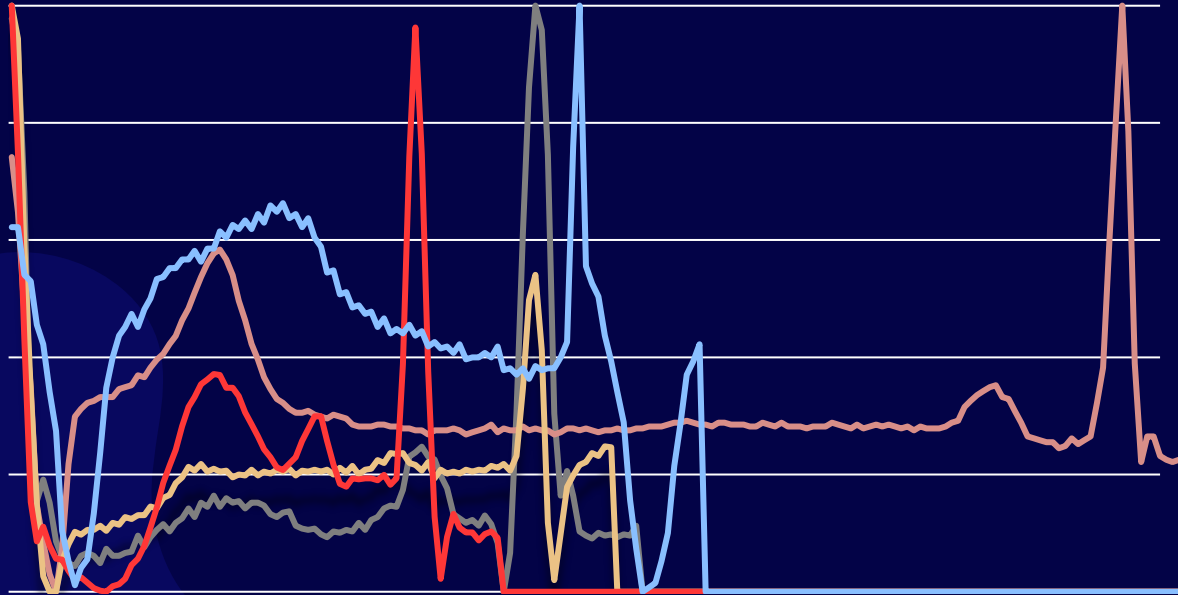
Categorizados en






5

clases diferentes  
(N, S, V, F, Q)



## Electrocardiograma de las clases



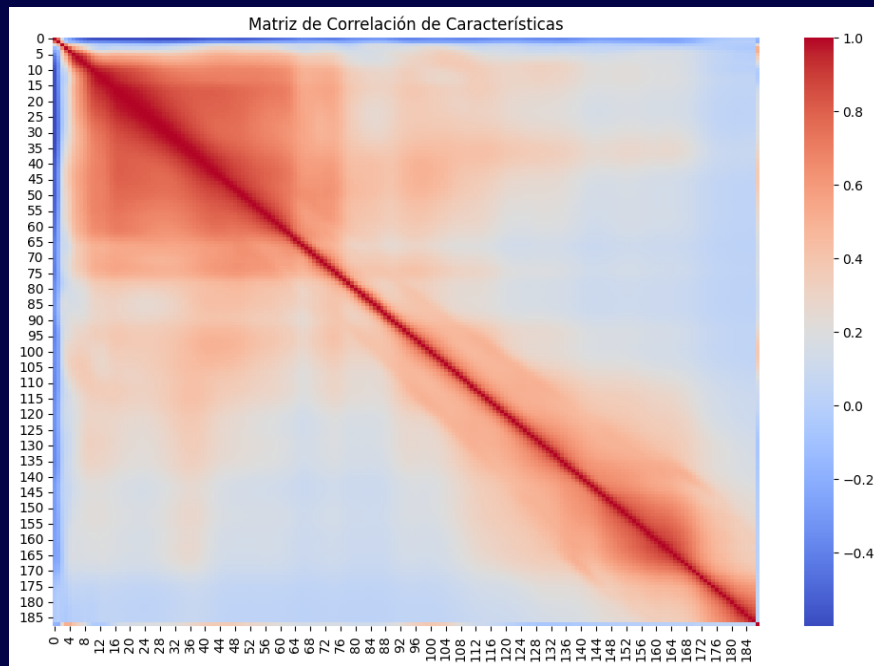
Categoría	Anotaciones
 N	<ul style="list-style-type: none"><li>• Normal</li><li>• Bloqueo de rama izquierda/derecha</li><li>• Escape auricular</li><li>• Escape nodal</li></ul>
 S	<ul style="list-style-type: none"><li>• Auricular prematuro</li><li>• Auricular aberrante prematuro</li><li>• Prematuro ganglionar</li><li>• Prematuro supraventricular</li></ul>
 V	<ul style="list-style-type: none"><li>• Contracción ventricular prematura</li><li>• Escape ventricular</li></ul>
 F	<ul style="list-style-type: none"><li>• Fusión de la función ventricular y normal</li></ul>
 Q	<ul style="list-style-type: none"><li>• Ritmo</li><li>• Fusión de ritmo y normalidad</li><li>• Inclasificable</li></ul>

Se grafica un electrocardiograma de una persona perteneciente a cada clase (onda) del dataset.

# Pruebas con los datos

## Análisis de correlación

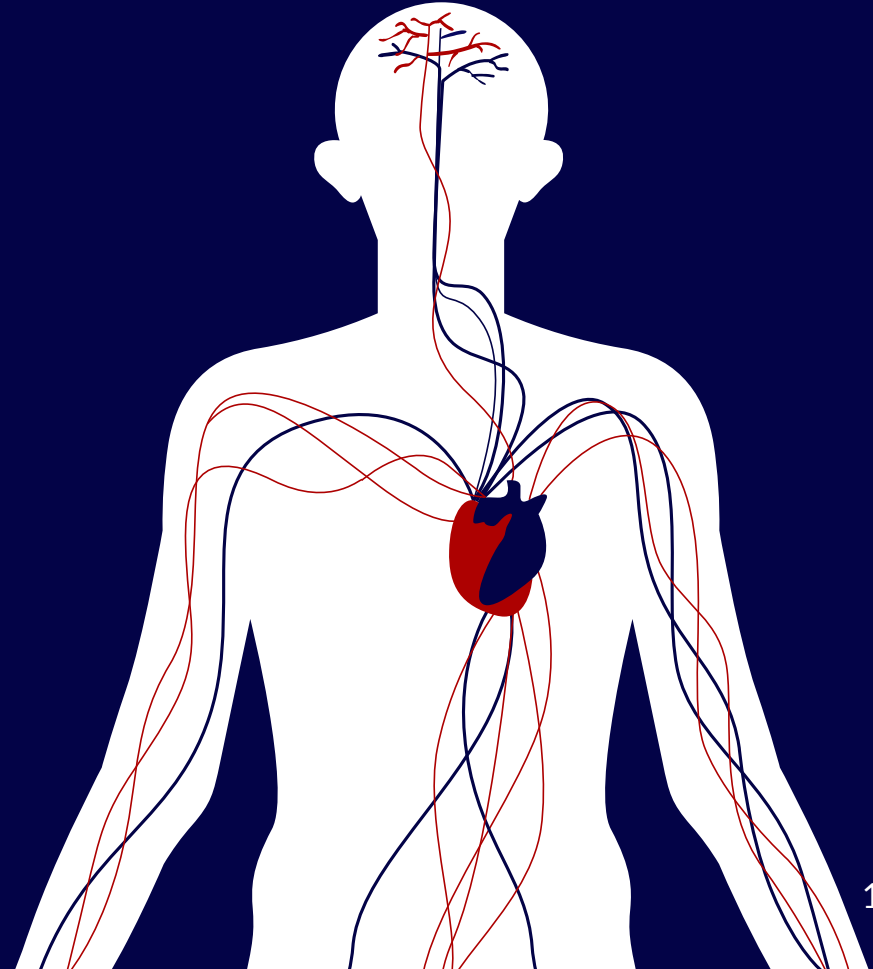
- Creación de matriz de correlación.
- Se grafican las características del dataset excepto la columna de etiquetas.
- Heatmap para facilitar visualización de correlaciones fuertes y débiles.
- Selección de mejor características.



03

# Data preparation

Exploración de los datos y divisiones



# División del dataset

## Conjunto de entrenamiento

- Usado para entrenar los modelos de clasificación
- 87,554 muestras
- 188 características

## Conjunto de prueba

- Evalúa el rendimiento y comportamiento del modelo
- ¿Qué posibilidad tiene el modelo de crear nuevos datos?
- 21,892 muestras
- 188 características

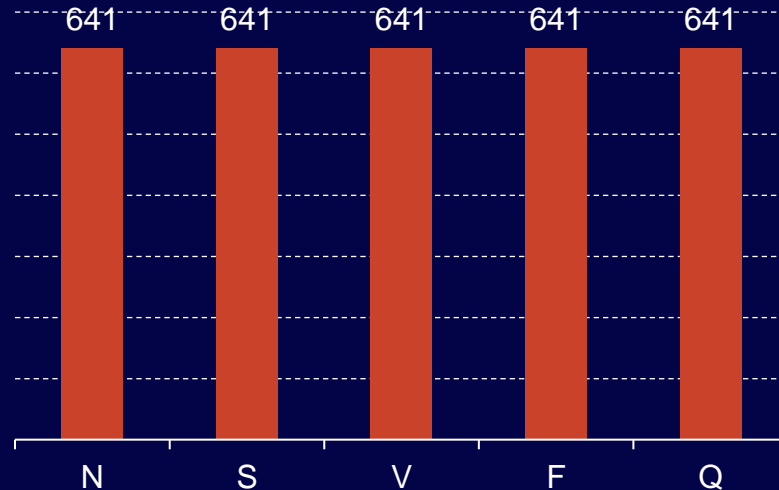


Incluyen las señales de ECG y la columna final corresponde a las etiquetas que indican el tipo de latido cardíaco.

# Balanceo de Datos



Conjunto de Datos Original



Conjunto de Datos Balanceado



Evitamos el sesgo hacia la clase "Normal"

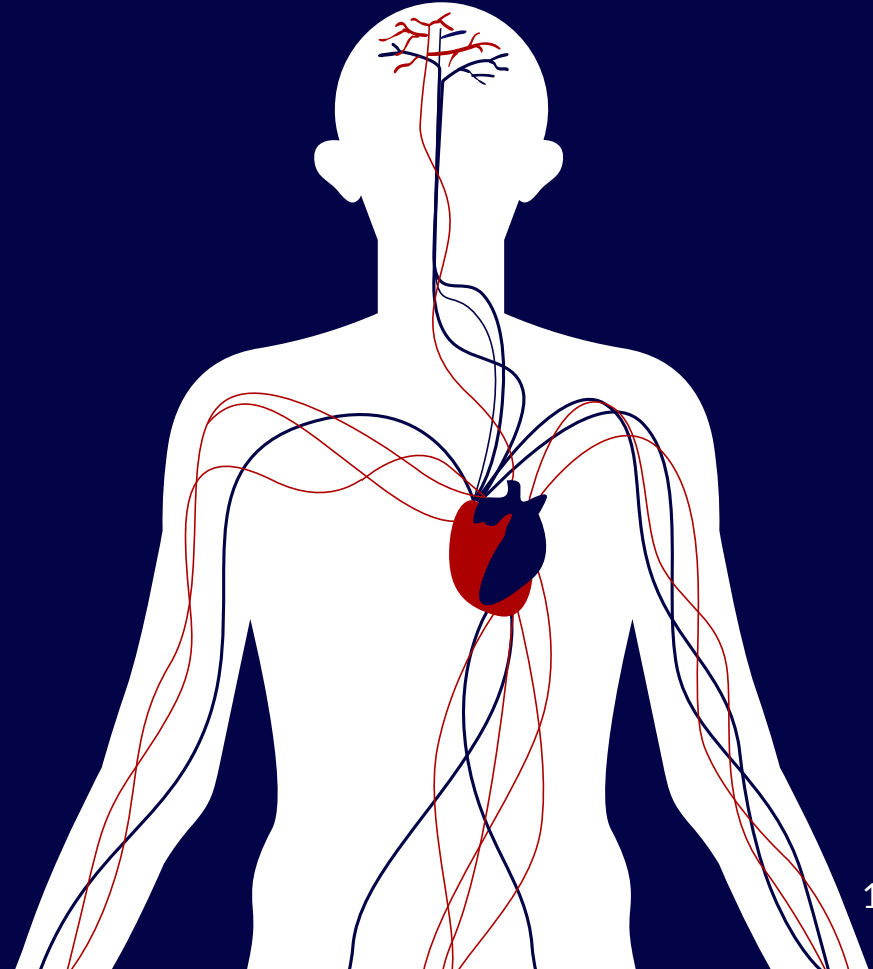


Capacidad de manejar grandes conjuntos de datos

04

# MODELOS

Modelación de las anomalías  
categorizadas del dataset





# MODELOS SOLO CON BALANCEO DE DATOS

Colores
Aceptable
Medio
Malo

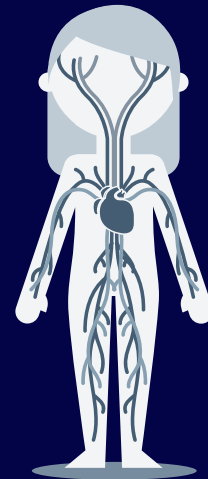
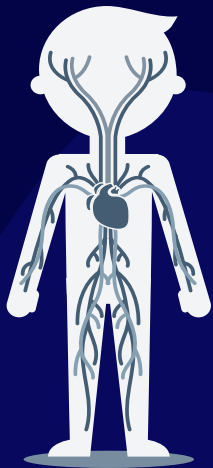
MODELO	ACCURACY	PRECISION	RECALL	F1-SCORE
Logistic Regression	74.85%	74.78%	74.85%	74.72%
KNN	83.65%	83.77%	83.65%	83.66%
Decision Tree	79.31%	79.28%	79.31%	79.23%
Random Forest	88.64%	89.14%	88.64%	88.76%
SVM	77.60%	78.58%	77.60%	77.84%
Naive Bayes	28.55%	39.54%	28.55%	21.28%
Gradient Boosting	85.34%	85.71%	85.33%	85.45%
MLP Classifier	86.15%	86.18%	86.15%	86.11%

# MODELOS CON SELECCIÓN DE HIPERPARÁMETROS

Colores
Aceptable
Medio
Malo

MODELO	ACCURACY	TIME FIT MODEL	TIME GRID SEARCH
KNN	0.97	1.5 min	173 min
Random Forest	0.91	1 min	67 min
Gradient Boosting	NA	NA	7337 min
MLP Classifier	0.98	20 seg	350 min

**Modelo seleccionado: Random Forest**



# Balanced Random Forest

MODELO	ACCURACY	RECALL	F-Score
Random Forest	0.91	0.96	0.73

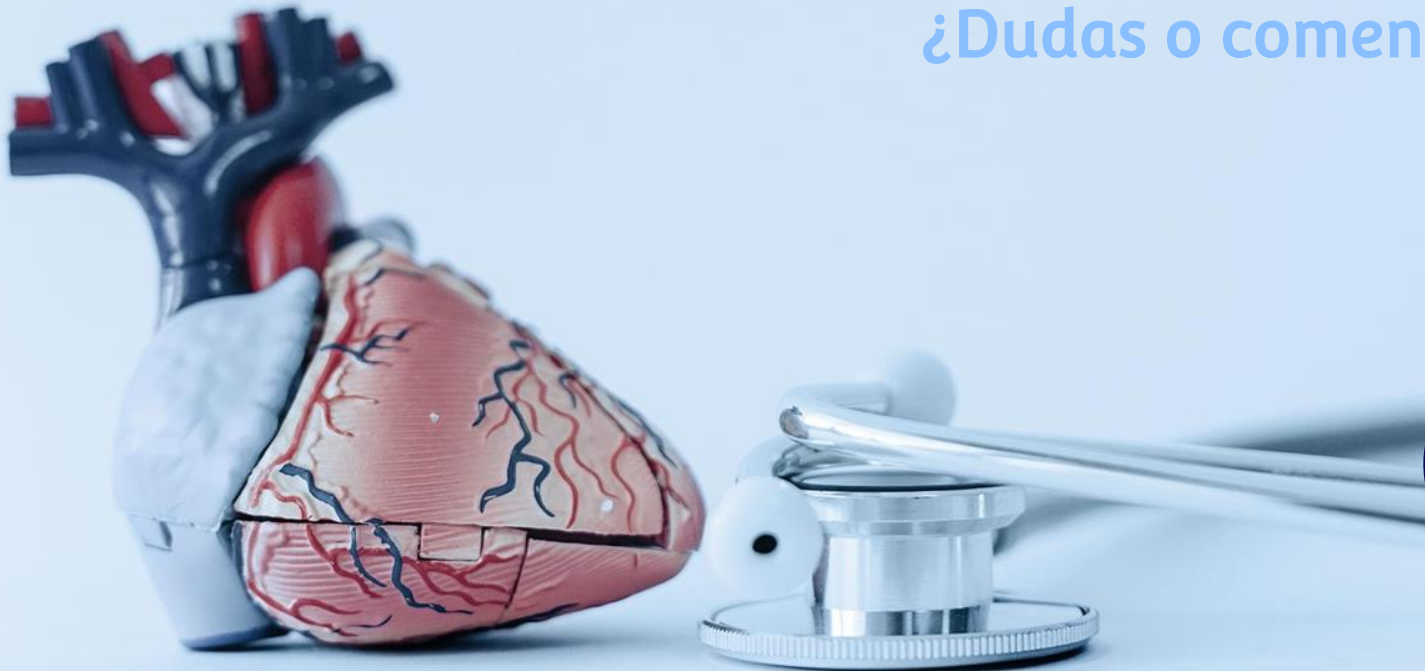
n_estimators	100	200	500
max_depth	10	20	None
min_samples_split	2	10	20
min_samples_leaf	1	5	10
bootstrap	True		False

Matriz de Confusión					
Etiqueta Real	0	1	2	3	4
	65001	2543	1284	3194	448
	30	2180	8	4	1
	137	25	5473	136	17
	0	0	0	641	0
4	93	12	29	8	6289
Etiqueta Predicha					

# Conclusión

# ¡Gracias!

¿Dudas o comentarios?



# REFERENCES

- Instituto de Seguridad y Servicios Sociales de los Trabajadores del Estado. (2016). La obesidad en México. Recuperado de <https://www.gob.mx/issste/articulos/la-obesidad-en-mexico>
- Lara, A., Rosas, M., Pastelín, G., Aguilar, C., & Attie, F. (2004). Hipercolesterolemia e hipertensión arterial en México: Consolidación urbana actual con obesidad, diabetes y tabaquismo. Archivos de cardiología de México, 74(3). Recuperado de [https://www.scielo.org.mx/scielo.php?script=sci\\_arttext&pid=S1405-99402004000300012](https://www.scielo.org.mx/scielo.php?script=sci_arttext&pid=S1405-99402004000300012)
- Medina-Mora, M. (2010). Tabaquismo en México ¿Cómo evitar 60,000 muertes prematuras cada año? Recuperado de <http://www.uade.inpsiquiatria.edu.mx/Articulos%20Jorge/2010/Tabaquismo%20en%20Mexico%20como%20evitar%2060%20000%20muertes%20prematurs%20cada%20ano.pdf>
- Sánchez-Arias, A. G., Bobadilla-Serrano, M. E., Dimas-Altamirano, B., Gómez-Ortega, M., & González-González, G. (2016). Enfermedad cardiovascular: primera causa de morbilidad en un hospital de tercer nivel. Revista Mexicana de Cardiología, 27(Supl. 3). Recuperado de <https://www.medigraphic.com/pdfs/cardio/h-2016/hs163a.pdf>
- Secretaría de Salud. (2022). 547. En México, 12.4 millones de personas viven con diabetes. Recuperado de <https://www.gob.mx/salud/prensa/547-en-mexico-12-4-millones-de-personas-viven-con-diabetes?idiom=es>
- UNAM Global. (2023). México: alarmantes cifras de estrés laboral. Recuperado de [https://unamglobal.unam.mx/global\\_revista/mexico-alarmantes-cifras-de-estres-laboral/](https://unamglobal.unam.mx/global_revista/mexico-alarmantes-cifras-de-estres-laboral/)