

## Grupo 7: Myocardial Infarction Complications

### Miembros del Equipo

- Acostupa Del Carpio, Juan José (20244906)
- Muñoz Saavedra, Breno Valentino (20162955)
- Huerta Firma, Ronny Michael (20184255)

### Conjunto de datos a Utilizar

Se utilizará el conjunto de datos "Myocardial Infarction Complications" publicado en el Machine Learning Repository de la Universidad de California en Irvine. Dicho conjunto de datos contiene información de 1700 instancias de pacientes que pueden presentar o no, al menos una de las 12 posibles complicaciones relacionadas con el infarto de miocardio.

- [Clasificación] Myocardial infarction complications Data Set - Prediction of myocardial infarction complications. <https://archive.ics.uci.edu/ml/datasets/Myocardial+infarction+complications>

### Objetivo del proyecto

El objetivo de este proyecto es desarrollar modelos de machine learning utilizando el dataset mencionado para predecir las posibles complicaciones en pacientes que han sufrido infarto en el miocardio.

### Artículos científicos relevantes

1. Golovenkin, S. (2020). Trajectories, bifurcations, and pseudo-time in large clinical datasets: applications to myocardial infarction and diabetes data. doi: 10.1093/gigascience/giaa128
2. Benzakour, H. (2023). A State of Art of Cardiovascular Diseases Using Machine Learning Algorithms. doi: 10.1109/CiSt56084.2023.10409886
3. Mesinovic, M., & Yang, K. (2022). Multi-label Neural Model for Prediction of Myocardial Infarction Complications with Resampling and Explainability. doi: 10.1109/BHI56158.2022.9926915
4. Moore, A., & Bell, M. (2022). XGBoost, A Novel Explainable AI Technique, in the Prediction of Myocardial Infarction: A UK Biobank Cohort Study. doi: 10.1177/11795468221133611
5. Kononova Y. (2024) Machine learning prediction of in-hospital recurrent infarction and cardiac death in patients with myocardial infarction. doi: 10.1016/j.imu.2023.101443
6. Ghafari, R. (2023) Prediction of the Fatal Acute Complications of Myocardial Infarction via Machine Learning Algorithms. doi: 10.18502/jthc.v18i4.14827
7. Newaz, A. (2023) Predicting complications of myocardial infarction within several hours of hospitalization using data mining techniques. doi: 10.1016/j.imu.2023.101361




### Propuesta tentativa de modelos de clasificación a utilizar

El dataset contiene once variables objetivo de naturaleza binaria, además de una variable categórica que puede tomar 7 categorías diferentes. Por ello, se puede tratar el problema tanto como uno de clasificación con múltiples clases (multiclass) o múltiples etiquetas (multilabel). Algunos modelos que permiten trabajar con datos de esta naturaleza son:

- Regresión logística (Esquema One Vs. The Rest para el caso multiclass).
- Árboles de decisión.
- Redes neuronales.
- Modelos de gradient boosting (XGBoost, LightGBM).

### Declaración de Trabajo Grupal

Unidad académica: Maestría en Informática	Semestre: 2024 - I
Curso y horario: Aprendizaje Automático: Teoría y aplicaciones (INF648), lunes de 7-10pm.	Profesores: Dr. César Beltrán Castañón, Mag. César Olivares Poggi

Título del trabajo: Myocardial infarction complications.	
Diseño/planificación del trabajo grupal (definir cronograma de trabajo, etc.)  (Ver Anexo 1)	
Funciones (compromiso) de cada integrante	Nombre, firma y fecha:
Realizar un análisis estadístico descriptivo y visualización de los datos. Determinar la estrategia para el manejo de datos faltantes (si aplica). Seleccionar y justificar la medida de calidad a utilizar. Reproducir los resultados reportados en un artículo científico anterior (línea base). Revisar y editar el documento final.	Breno Valentino Muñoz Saavedra (20/05/2024) 
Investigar y sintetizar artículos científicos relevantes (con el dataset) Análisis descriptivo del conjunto de datos. Visualización de los datos. Identificar algoritmos a emplear. Establecer estrategia de validación. Implementar y evaluar el rendimiento del modelo propio.	Juan José Acostupa Del Carpio (20/05/24) 
Introducción y Estado del Arte Redactar la introducción (presentación del problema y objetivo del estudio). Investigar y sintetizar el aporte de artículos científicos relevantes (sin el conjunto de datos seleccionado). Identificar y resumir al menos 2 artículos científicos que hayan utilizado el conjunto de datos seleccionado. Código de procesamiento de datos	Ronny Michael Huerta Firma (20/5/24) 
Firma del profesor y fecha:	

## Anexo 1: Cronograma propuesto

			Cronograma										
Etapas	Sección	Tarea	14/05 - 20/05	21/05 - 27/05	28/05 - 03/06	04/06 - 10/06	11/06 - 17/06	18/06 - 24/06					
Primera Parte	1. Preparación	1.1 Definir roles y responsabilidades											
		1.2 Realizar una revisión preliminar de la literatura											
Segunda Parte	2. Introducción y estado del arte	2.1 Investigar y sintetizar artículos científicos relevantes (sin el dataset)											
		2.2 Investigar y sintetizar artículos científicos relevantes (con el dataset)											
		2.3 Redactar la introducción											
	3. Diseño del experimento	3.1 Análisis descriptivo del conjunto de datos											
		3.2 Visualización de los datos											
	4. Metodología	4.1 Determinar la estrategia de feature eng. y manejo de datos faltantes											
		4.2 Selección de variables											
		4.3 Elección de métrica objetivo											
		4.4 Identificar algoritmos a emplear y definir la estrategia para su ajuste											
		4.5 Establecer la estrategia de validación											
Tercera Parte	5. Experimentación y resultados	5.1 Reproducir resultados de un artículo científico anterior											
		5.2 Implementar y evaluar el rendimiento del método propio											
		5.3 Comparar resultados propios con línea base											
Cuarta Parte	6. Discusión y conclusiones	6.1 Interpretar y discutir los resultados obtenidos											
		6.2 Sugerir mejoras y posible trabajo futuro											
		6.3 Redactar las conclusiones											
	7. Revisión y entrega final	7.1 Revisar y editar el informe completo											
		7.2 Preparar la presentación final											
		7.3 Entrega del informe											

