

# PREDICCIÓN DE ATAQUE CARDÍACO CON UN MODELO DE CLASIFICACIÓN LOGÍSTICA



## TIPO DE PROYECTO:

ABP (APRENDIZAJE BASADO EN PROYECTOS) CON ENFOQUE DE CIENCIA DE DATOS APLICADO A SALUD.

FUENTE: [\[README\]](#)

## ESPACIO CURRICULAR/MÓDULO:

CIENCIA DE DATOS · ESTADÍSTICA Y EXPLORACIÓN DE DATOS II

## EJES | UNIDADES CONCEPTUALES:

PRIMERO EXPLORAMOS Y ORDENAMOS LOS DATOS; DESPUÉS REVISAMOS SUPUESTOS Y RELACIONES ENTRE VARIABLES (SPEARMAN Y ANOVA CON INTERACCIÓN); Y, POR ÚLTIMO, ENTRENAMOS UNA REGRESIÓN LOGÍSTICA QUE MEDIMOS CON LA MATRIZ DE CONFUSIÓN, LA CURVA ROC Y LOS ODDS RATIOS.

## PROBLEMÁTICA | NECESIDAD | CASO:

CLASIFICAMOS EL DIAGNÓSTICO DE ATAQUE CARDÍACO (POSITIVO/NEGATIVO) A PARTIR DE MEDICIONES HOSPITALARIAS DE RUTINA —EDAD, SEXO, PRESIÓN ARTERIAL, PULSO, GLUCOSA Y BIOMARCADORES CARDÍACOS (CK-MB Y TROPONINA)— USANDO UN CONJUNTO ETIQUETADO DE 1.319 PACIENTES. EL TRABAJO ES REPRODUCIBLE Y ORIENTADO A APOYAR LA TOMA DE DECISIONES CLÍNICAS.

FUENTE: [\[MEDICALDATASET.CSV\]](#)

## OBJETIVO GENERAL:

DESARROLLAR Y EVALUAR UN MODELO DE CLASIFICACIÓN QUE ANTICIPE EL DIAGNÓSTICO POSITIVO O NEGATIVO DE ATAQUE CARDÍACO UTILIZANDO VARIABLES CLÍNICAS DISPONIBLES EN LA ADMISIÓN.

FUENTE: [\[ABP\\_FINAL\\_PREDICCIÓN.IPYNB\]](#)

## OBJETIVOS ESPECÍFICOS:

DESCRIBIR EL CONJUNTO DE DATOS, ESTUDIAR RELACIONES Y DIFERENCIAS POR DIAGNÓSTICO Y SEXO, ENTRENAR UNA REGRESIÓN LOGÍSTICA Y VALORAR SU DESEMPEÑO CON MÉTRICAS ESTÁNDAR, INTERPRETANDO SUS EFECTOS MEDIANTE ODDS RATIO E INTERVALOS DE CONFIANZA.

FUENTE: [\[CORRELACIONES\\_LINEALES.IPYNB\]](#)



## FUNDAMENTACIÓN | HIPÓTESIS:

LA FUNDAMENTACIÓN PRIORIZA UN MODELO INTERPRETABLE.

REGRESIÓN LOGÍSTICA PARA ESTIMAR UNA PROBABILIDAD CLÍNICAMENTE ACCIONABLE Y COMUNICAR EFECTOS MEDIANTE ODDS RATIOS. EL EDA Y EL ANOVA DE DOS VÍAS GUIARON LA SELECCIÓN CON CRITERIO DE PARSIMONIA Y RELEVANCIA CLÍNICA, BUSCANDO ESTABILIDAD Y EVITANDO COLINEALIDAD; OPERATIVAMENTE, EL FOCO ES REDUCIR FALSOS NEGATIVOS EN LA PRIORIZACIÓN DE PACIENTES.

PLANTEAMOS COMO HIPÓTESIS QUE TROPONINA Y CK-MB INCREMENTAN DE FORMA SIGNIFICATIVA LA PROBABILIDAD DE DIAGNÓSTICO POSITIVO; QUE LA EDAD SE ASOCIA POSITIVAMENTE CON EL RIESGO; QUE EL SEXO MASCULINO APORTA UN EFECTO MODERADO; QUE LA GLUCEMIA NO EXHIBE UN EFECTO PRINCIPAL CONSISTENTE PERO PODRÍA INTERACTUAR CON EL SEXO; Y QUE PRESIÓN ARTERIAL Y FRECUENCIA CARDÍACA NO AÑADEN MEJORA ESTABLE UNA VEZ CONSIDERADOS BIOMARCADORES Y EDAD. LA VERIFICACIÓN SE REALIZA CON LOS COEFICIENTES Y ODDS RATIOS (IC95%) DEL MODELO LOGÍSTICO Y CONTROLES DE ESTABILIDAD (VIF BAJO), MANTENIENDO UN MODELO COMPACTO Y CLÍNICAMENTE LEGIBLE.

FUENTE: [\[EV2\\_ANOVA.IPYNB\]](#)

FUENTE: [\[EV3\\_HEARTATTACK\\_.IPYNB\]](#)

## ACCIONES | RECURSOS | TIEMPO:

**ACCIONES:** EDA Y CONTROL DE SUPUESTOS - CORRELACIONES (SPEARMAN) - ANOVA 2 VÍAS - REGRESIÓN LOGÍSTICA - EVALUACIÓN (ROC, MATRIZ DE CONFUSIÓN) - SÍNTESIS.

**RECURSOS:** PYTHON (PANDAS, NUMPY, MATPLOTLIB, SEABORN, STATSMODELS, SCIKIT-LEARN); DATASET DE HOSPITAL ZHEEN (MENDELEY DATA). SE CORRIGEN OUTLIERS DE “HEART RATE” DETECTADOS.

**TIEMPO:** SECUENCIA POR ETAPAS (ANÁLISIS - MODELADO - REPORTE) SEGÚN CRONOGRAMA DEL MÓDULO.

## PRODUCTO FINAL | CONCLUSIONES | RESULTADOS ESPERADOS:

EL MODELO FINAL CON EDAD, SEXO, CK-MB Y TROPONINA ALCANZA ACCURACY 0,785, PRECISIÓN 0,819, RECALL 0,835, F1 0,827 Y ROC-AUC 0,871. TROPONINA Y CK-MB SON LOS PREDICTORES DOMINANTES; LA EDAD CONTRIBUYE EN MENOR MAGNITUD Y EL SEXO MUESTRA UN EFECTO MODERADO. PRESIÓN Y PULSO NO EVIDENCIAN DIFERENCIAS RELEVANTES; LA GLUCEMIA SUGIERE INTERACCIÓN CON EL SEXO. SE PROPONE COMO HERRAMIENTA DE APOYO A LA PRIORIZACIÓN, SIN REEMPLAZAR EL CRITERIO MÉDICO.

## INTEGRANTES:

● BARBERO, MACIEL ● MARINI, IAN DENIS ● MOLINA, JONATHAN ARIEL ● MOLINA, MAURICIO LEONEL ● ROBLES, EMILCE LUCIA NICOLE ● SOSA, SEBASTIAN CRISTHIAN ● VIRINNI, MARCO

## DOCENTE:

● NAHUEL PRATTA ● MARCOS UGARTE

## TECNICATURA:

CIENCIA DE DATOS II

## COHORTE:

20214