

# MODELOS DE APRENDIZAJE AUTOMÁTICO EXPLICABLES PARA PREDECIR MORTALIDAD HOSPITALARIA EN UCI: COMPARACIÓN CON SOFA Y APACHE II EN UNA COHORTE DE 1.200 PACIENTES

Magdalena Vera-Alarcón<sup>1</sup>, Alvaro Passi-Solar<sup>1</sup>, María Luz Riquelme<sup>2</sup>, Rosana Lagos<sup>2</sup>, José Antonio Leiva<sup>1</sup>, Daniela Maldonado<sup>1</sup>, Omar Azar<sup>1</sup>, Milovan Gamboa<sup>1</sup>, Catalina Sauterel<sup>1</sup>, Florencia Sauterel<sup>1</sup>, Paz Uribe<sup>1</sup>

1. Pontificia Universidad Católica de Chile

2. Red de Salud UC-CHRISTUS

## INTRODUCCIÓN:

La predicción de mortalidad en UCI es clave para la toma de decisiones clínicas y la asignación eficiente de recursos. Las escalas tradicionales, como SOFA y APACHE II, tienen limitaciones en precisión individual y aplicabilidad a contextos específicos. Este estudio propone modelos de aprendizaje automático explicables para anticipar mortalidad hospitalaria en pacientes críticos, integrando rendimiento predictivo, interpretabilidad clínica y comparación con scores clásicos.

## MÉTODOS:

Estudio retrospectivo en una cohorte de 1.200 pacientes adultos ingresados a la UCI del Hospital Clínico de la Red de Salud UC-CHRISTUS, entre el 1 de agosto de 2023 y el 31 de diciembre de 2024. Se utilizaron variables clínicas, de laboratorio y demográficas recolectadas durante las primeras 24 horas. Se entrenaron cinco algoritmos de clasificación (KNN, Naive Bayes, Regresión Logística, Random Forest, XGBoost y LightGBM), comparando tres configuraciones: (1) variables seleccionadas mediante SHAP, (2) solo puntaje SOFA y (3) solo puntaje APACHE II. El conjunto de datos se dividió en 80% entrenamiento y 20% prueba. Se evaluó desempeño mediante AUC, F1-score, sensibilidad, especificidad, VPP y VPN. Se utilizó SHAP (SHapley Additive exPlanations) para interpretación global e individual de los modelos. Estudio aprobado por comité de ética institucional (ID 220907003).

## RESULTADOS:

La cohorte fue equilibrada por sexo (50.1% mujeres), con predominio de pacientes entre 60–80 años (42.2%). La mortalidad hospitalaria fue 15.0%. El modelo LightGBM con variables SHAP obtuvo el mejor desempeño global (AUC: 0.934; sensibilidad: 94.6%; VPN: 0.99), seguido de Random Forest con SHAP (AUC: 0.901; F1-score: 0.63). En todos los algoritmos, las versiones basadas en SHAP superaron a las que usaron solo SOFA (AUC promedio 0.81) o APACHE II (AUC promedio 0.84). En términos clínicos, las variables más influyentes fueron ventilación mecánica, tipo de diagnóstico (quirúrgico o infeccioso), creatinina, índice de comorbilidad de Charlson, BUN, lactato, edad y presión arterial media. Los modelos con solo scores tradicionales, si bien alcanzaron buena sensibilidad (hasta 100% en APACHE-LightGBM), mostraron menor precisión global (accuracy 62–76%) y F1-scores más bajos (<0.48). La visualización SHAP permitió validar la coherencia clínica del modelo y facilitó la interpretación para el equipo asistencial.

## CONCLUSIÓN:

Los modelos de machine learning predicen mortalidad hospitalaria con alta precisión, superando de forma consistente a los scores tradicionales. La combinación de rendimiento técnico y transparencia explicativa convierte a SHAP en una herramienta potente para fortalecer la confianza clínica en modelos predictivos. Este enfoque promueve una medicina intensiva más proactiva, personalizada y basada en datos. Futuras etapas incluirán la ampliación progresiva de la cohorte para robustecer el modelo y explorar su generalización a distintos subgrupos clínicos dentro de la UCI.