MODELO PRONÓSTICO DE MORTALIDAD EN ICTUS ISQUÉMICO NO REPERFUNDIDO Y MÉTRICAS A LA LUZ DE LA CIENCIA DE DATOS

García-Torrecillas JM[1], Iglesias-Espinosa M[2], Amaya-Pascasio L[2], Galván-Espinosa J[3], Martínez-Sánchez P[2], Aparicio-Mota A[4], Lea-Pereira MC[5]

1. UGC Urgencias. Unidad de Investigación Biomédica. HU Torrecárdenas. 2. Servicio de Neurología. HU Torrecárdenas. 3. Fundación FIBAO. 4. Unidad de Investigación Biomédica. HU Torrecárdenas 5.Servicio de Medicina Interna. HU Poniente. El Ejido (Almería)..

IV CONGRESO INTERNACIONAL DE INNOVACIÓN E INVESTIGACIÓN EN EL ÁMBITO DE LA SALUD

OBJETIVOS

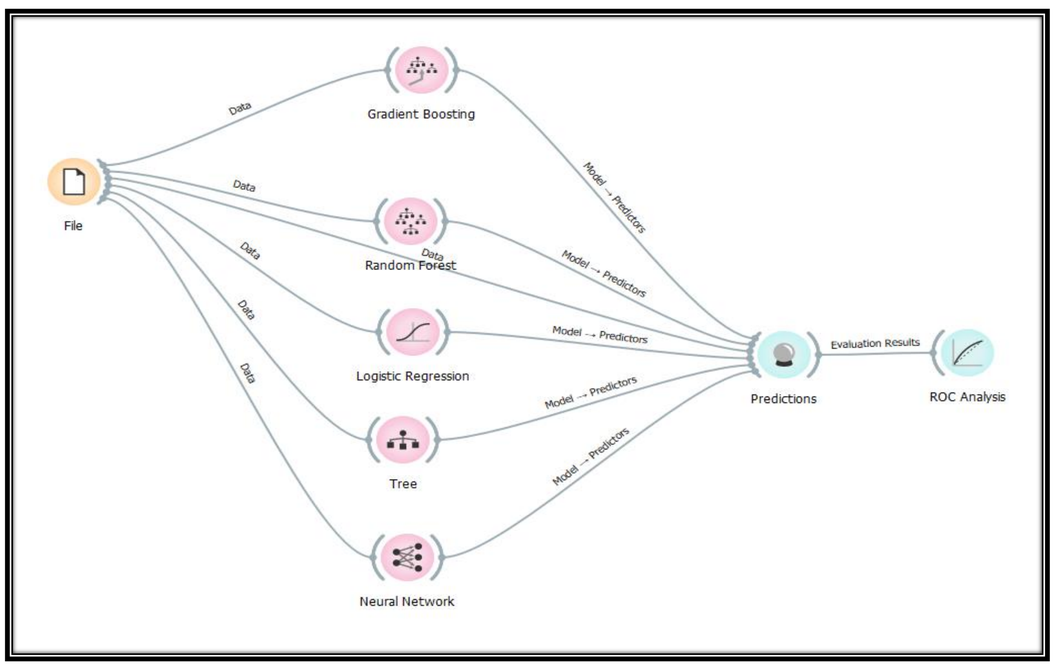
- Los modelos predictivos para mortalidad en ictus proporcionan información clave para el manejo estratificado según gravedad. La ciencia de datos proporciona herramientas útiles y complementarias a la estadística clásica para su elaboración.

- Objetivos: Estimar y evaluar mediante cinco metodologías de la ciencia de datos (regresión logística, árboles de clasificación, random forest, g-boosting y redes neuronales) las diferencias en la capacidad discriminativa (ROC) y la exactitud de las predicciones (Accuracy) de los modelos pronósticos obtenidos según cada método específico.

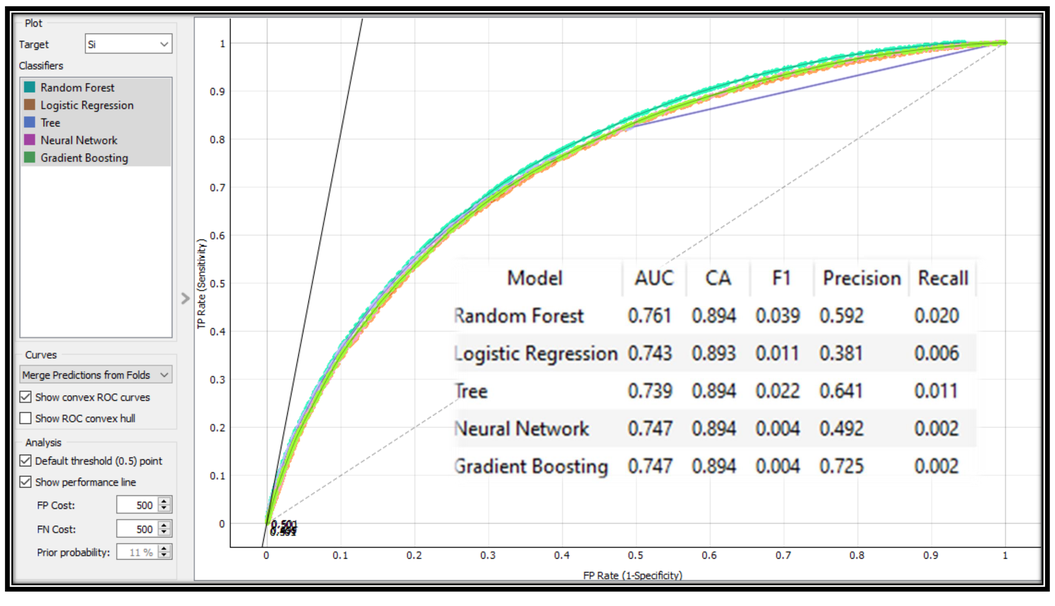
MÉTODO

Diseño: estudio empírico, del tipo observacional analítico, sobre una cohorte de inclusión prospectiva con metodología de Data Science. Información: 147.047 casos de ingresos de pacientes del conjunto mínimo básico de datos con el criterio de inclusión de estar adscritos al GRD45 (ictus isquémico) de la 10ª edición de la Clasificación Internacional de Enfermedades.
Ámbito: España, periodo temporal 2016-2018. Variable dependiente: mortalidad. Variables predictoras: edad, sexo, comorbilidades más frecuentes (hipertensión, cardiopatías, dislipemia, territorio vascular afectado, otras).
Procedimiento: aplicación de metodología de gradient boosting, árboles de clasificación, random forest, regresión logístisca y red neuronal para calcular las métricas ROC y Accuracy según cada procedimiento.
Software: Orange Data Mining v.3.34.0 basado en Python.

Algoritmo de Trabajo de Orange



MÉTODOS DE PREDICIÓN Y AUCs CORRESPONDIENTES



Model	AUC	CA	F1	Precision	Recall
Random Forest	0.761	0.894	0.039	0.592	0.020
Logistic Regression	0.743	0.893	0.011	0.381	0.006
Tree	0.739	0.894	0.022	0.641	0.011
Neural Network	0.747	0.894	0.004	0.492	0.002
Gradient Boosting	0.747	0.894	0.004	0.725	0.002

RESULTADOS

Todas las metodologías proporcionaron un AUC moderado-alto, destacando Random Forest (0.761), igual resultado de 0.747 en regresión logística, red neuronal y gradient boosting, así como el menor valor de ROC con árboles de clasificación (0.739).

La Accuracy fue de 0.894 en todos los métodos excepto en regresión logística (0.893).

CONCLUSIONES

La metodología basada en ciencia de datos para la elaboración de modelos pronósticos proporciona resultados óptimos, similares a la regresión clásica, y permite obtener métricas relativas a la capacidad discriminativa (ROC) y exactitud (Accuracy) equivalentes y que, en conjunto, nos facilitan comprender en qué grado y supuestos el modelo pronóstico evalúa la realidad.

FUNDING

Junta de Andalucía
Fundación para la Investigación Biosanitaria de Andalucía Oriental - Alejandro Otero
Consejería de Salud y Familias
Fundación Progreso y Salud
Consejería de Salud y Familias

Convocatoria FPS 2020. Expdte: AP-0013-2020-C1-F1