



XLVIII Coloquio Argentino de Estadística

VI JORNADA DE EDUCACIÓN ESTADÍSTICA "MARTHA DE ALIAGA"

27 al 30 oct 2020

Poster:

Predicción del reintento de suicidio: Random survival forest versus regresión de Cox

Luciana C. Chiapella, Leandro Grendas, Federico M. Daray



Esta obra está bajo una Licencia
Creative Commons Atribución-
NoComercial 4.0 Internacional



FACULTAD
DE CIENCIAS
ECONÓMICAS



Universidad
Nacional
de Córdoba



PREDICCIÓN DEL REINTENTO DE SUICIDIO: RANDOM SURVIVAL FOREST VERSUS REGRESIÓN DE COX

LUCIANA C. CHIAPELLA^{1,2}, LEANDRO GRENDAS³, FEDERICO M. DARAY^{2,3}

¹Facultad de Ciencias Bioquímicas y Farmacéuticas, Universidad Nacional de Rosario

²Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET) ³Instituto de Farmacología, Facultad de Medicina, Universidad Buenos Aires - lucianachiapella@conicet.gov.ar



Introducción: En Psiquiatría, resulta de interés poder predecir un reintento de suicidio en pacientes asistidos por conducta suicida, de acuerdo a sus características sociodemográficas y clínicas.

Habitualmente, se utilizan modelos de regresión de Cox a fin de seleccionar los mejores predictores y, en base a ellos, estimar la probabilidad del evento de interés, pero resulta de interés evaluar métodos alternativos.

Los datos: Cohorte de 308 pacientes atendidos por intento de suicidio en 3 hospitales de la ciudad de Buenos Aires, con seguimiento entre 3 meses y dos años. Se registraron 26 características sociodemográfica y clínicas como posibles predictores del reintento de suicidio y el tiempo hasta la ocurrencia del evento o censura.

Objetivo: Comparar la capacidad discriminante y predictiva de cuatro estrategias para el ajuste de los datos: regresión de Cox con todas las covariables (*CoxComp*), regresión de Cox mediante eliminación hacia atrás (*CoxBack*), Random Survival Forest (RSF) con todas las covariables (*RSFComp*) y RSF con eliminación de las variables con importancia negativa en *RSFComp* (*RSFelim*).

Métodos: Los datos se dividieron en una base de entrenamiento y otra de prueba. Sobre la primera, se ajustaron los modelos y, con ellos, se estimó la probabilidad de reintento a dos años para los datos de prueba. Se midió la sensibilidad, especificidad y precisión para predecir el evento y se calculó el área bajo la curva ROC (AUC) y el score integrado de Brier (IBS) a fin de evaluar el rendimiento de predicción general. El proceso se repitió 500 veces.

Resultados: *RSFelim* presentó los mejores resultados respecto a AUC, especificidad y precisión, siendo *CoxComp* y *CoxBack* quienes presentaron mejores resultados para sensibilidad. *RSFComp* mostró los resultados más desfavorables en las medidas calculadas, salvo para especificidad. Respecto al IBS, *RSFelim* mostró los mejores resultados y *CoxComp*, los más desfavorables (datos no mostrados). Si bien *RSFComp* y *CoxBack* presentaron valores similares de la mediana de IBS, *CoxBack* mostró mayor variabilidad. De acuerdo a estos resultados, *RSFelim* fue considerado el método con mejor capacidad predictiva para este conjunto de datos.

