

Predicción temprana de mortalidad en pacientes con sepsis usando técnicas de análisis de supervivencia y machine learning

Sebastián Mena Aliaga

Instituto de Estadística, Pontificia Universidad Católica de Valparaíso
División de Estadística, Banco Central de Chile

Resumen

La sepsis es una de las mayores causas de muertes al rededor del mundo. El año 2017 se registraron 48.9 millones de casos de sepsis en todo el mundo y se notificaron 11 millones de muertes relacionadas, lo que representó casi el 20% de todas las muertes mundiales. Este estudio realiza un análisis de supervivencia de las primeras 48 horas de registros clínicos de pacientes que han sido diagnosticados con sepsis en las Unidades de Cuidados Intensivos (UCI) del Beth Israel Deaconess Medical Center, base de datos MIMIC-IV, considerando un cohorte total de 4.509 pacientes. Se comienza realizando un análisis de las curvas de supervivencia por el estimador no paramétrico *Kaplan-Meier* y se constatan diferencias significativas entre las curvas estimadas de tres grupos etarios de pacientes a través de la prueba de bondad de ajuste de *log-rank*. Se propone un modelo de *Random Survival Forest*, extensión del famoso modelo de *Random Forest* con capacidad de trabajar con datos de supervivencia, para modelar una relación funcional entre los datos demográficos, mediciones de laboratorios y signos vitales de pacientes, y sus riesgos de supervivencia. El resultado del modelo predice la función de riesgo por estimación de *Nelson-Aalen* y la función de supervivencia para nuevos pacientes. El modelo propuesto logra resultados muy similares a los observados en el estado del arte, las métricas de desempeño especializadas para análisis de supervivencia obtenidas son: 0.69 de AUC Acumulativo/Dinámico, 0.68 de Índice de Concordancia y 0.72 de puntuación Brier.

Referencias Bibliográficas

1. Ishwaran, H., Kogalur, U.B., Blackstone, E.H., Lauer, M.S. (2008). Random survival forests. *The Annals of Applied Statistics* **2**, 841-860.
2. Wang, P., Li, Y., Reddy, C.K. (2019). Machine learning for survival analysis: A survey. *ACM Computing Surveys* **51** (6), 1-36.
3. Zhang, L., Huang, T., Xu, F., Li, S., Zheng, S., Lyu, J., Yin, H. (2022). Prediction of prognosis in elderly patients with sepsis based on machine learning (random survival forest). *BMC Emergency Medicine* **22**, 26.
4. Johnson, A., Bulgarelli, L., Pollard, T., Horng, S., Celi, L.A., Mark, R. (2022). MIMIC-IV (version 2.0). *PhysioNet*.