



8. IDENTIFICACIÓN DE PERFILES CLÍNICOS EN PACIENTES CON ENFERMEDADES CARDIOVASCULARES MEDIANTE ALGORITMOS DE CLUSTERING

Jorge Vélez García¹, Nicolás Rosillo Ramírez¹, Lourdes Vicent Alaminos¹, Guillermo Moreno Muñoz¹, José Luis Bernal Sobrino² y Héctor Bueno Zamora¹

¹Servicio de cardiología. Hospital Universitario 12 de Octubre, Madrid, España y ²Departamento de Análisis de Información y Control de Gestión. Hospital Universitario 12 de Octubre, Madrid, España.

Resumen

Introducción y objetivos: El análisis de perfiles de pacientes hospitalizados por enfermedades cardiovasculares es un reto (debido a factores como las comorbilidades o la heterogeneidad en perfiles con un mismo diagnóstico). Además, la estadística descriptiva no siempre es suficiente para comprender la complejidad de las interacciones entre variables y su relación con el pronóstico del paciente. El objetivo de este trabajo es la determinación de diferentes fenotipos (*clusters*) de pacientes para los diagnósticos cardiovasculares más frecuentes mediante aprendizaje automático no supervisado.

Métodos: Se utilizó como fuente de información el conjunto mínimo básico de datos de hospitalización de los años 2016 a 2022, de donde se extrajeron los pacientes con ingresos cardiovasculares. Posteriormente, se aplicaron técnicas de *clustering* para agrupar a los pacientes en diferentes perfiles, utilizando algoritmos de agrupamiento como k-means y *clustering* jerárquico. Mediante el método *elbow*, se estableció el número de clústeres óptimos en 5. Finalmente, se evaluaron los perfiles resultantes utilizando tablas descriptivas de perfiles de comorbilidad y letalidad por *cluster*, y se examinó la distribución de los diagnósticos principales cardiovasculares más frecuentes por *cluster*. Se utilizó el lenguaje de programación R para realizar el procesamiento de la base de datos, el análisis por *clustering* y la creación de tablas y figuras.

Resultados: Los 5 clústeres agruparon 2.901, 4.242, 4.222, 3.857 y 1.904 episodios respectivamente, lo que permitió identificar perfiles de pacientes con características clínicas y pronósticos diferentes. Se identificaron grupos con diferencias significativas en comorbilidades y perfiles con mayor riesgo de mortalidad (4,64%). Sin embargo, se sugiere un análisis más complejo utilizando distintas variables y métodos de agrupamiento para identificar perfiles de riesgo que se beneficiarían de intervenciones de seguimiento estrecho al alta hospitalaria.



Diagrama de Sankey de distribución de diagnósticos principales por cluster.

Conclusiones: La utilización de técnicas de *clustering* permite identificar grupos clínicamente relevantes con diferentes características clínicas, diagnósticos y pronósticos. Esta metodología puede ser útil para el

desarrollo de estrategias prevención, atención y tratamiento personalizadas, basadas en las características específicas de cada perfil, además de ayudar a mejorar la comprensión de la complejidad de los pacientes con enfermedades cardiovasculares.