

1687 - GENERACIÓN MEDIANTE *MACHINE LEARNING* DE UN MODELO DE CLASIFICACIÓN PARA EL DIAGNÓSTICO DE SARCOPENIA

Jose Carlos Arévalo Lorio¹, Juana Carretero Gómez¹, Nadia Mayoral Testón² y Daniel Fernández-Berges Gurrea³

¹Medicina Interna, Hospital Universitario de Badajoz/INUBE, Badajoz, España. ²Unidad de Investigación de Villanueva de la Serena, GRIMEX/FUNDESALUD, Villanueva de la Serena, España. ³Unidad de investigación de Villanueva de la Serena, GRIMEX/FUNDESALUD, Villanueva de la Serena, España.

Resumen

Objetivos: Según el grupo europeo para el estudio de la sarcopenia en personas mayores (EWGSOP2)¹, el diagnóstico de sarcopenia está basado en una disminución de la fuerza muscular generalmente valorado mediante dinamometría, y en una disminución de la masa muscular valorado por composición corporal mediante DXA (absorciometría dual de rayos X) o BIAS (bioimpedancia). El problema es que estas técnicas de composición corporal no están fácilmente disponibles. Nuestro objetivo es obtener un algoritmo que prediga la disminución de la masa muscular basada en los criterios de EWGSOP2, sin necesidad de utilizar DXA o BIAS.

Métodos: Mediante los registros obtenidos en NHANES (National Health and Nutrition Examination survey) en los años 2003-2005 y 2011-2018, se configura una base de datos matriz que contiene entre otros las variables de composición corporal medida por DXA. De esta matriz se separan por sexo dos submatrices calculándose en cada una de ellas el ASMI (índice de masa muscular esquelética apendicular) y se crea una variable (ASMI) que clasifique a los individuos según los puntos de corte sugeridos por EWGSOP (7 kg/m² en varones, 5,7 kg/m² en mujeres). Ambas matrices se subdividen en un set de entrenamiento (75% de los datos) y un set de test (restante 25%). Se aplican tres algoritmos de clasificación: dos SVM (*support vector machine*, radial y sigmoide) y un modelo XGBoost (Extreme Gradient Boosting) que se generan con el set de entrenamiento y posteriormente se validan con el set de test. Se obtienen las curvas ROC de cada uno de los modelos, junto a la precisión y los valores predictivos. Se utiliza el entorno R (The R Project for Statistical Computing), versión 4.3.0.

Resultados: De unas matrices iniciales de 10.927 hombres, edad media 33,7 (15,6) y de 10.362 mujeres, edad media 34,7 (17,7), se obtienen dos sets de entrenamiento (8.196 hombres y 7.772 mujeres) y dos de test (2.731 hombres y 2.590 mujeres). Las principales características se muestran en la tabla 1. Para los modelos se realiza una selección de variables basada en criterios antropométricos y bioquímicos (tabla 1). En el caso de los hombres (tabla 2, fig.), el mejor modelo obtenido es el realizado mediante XGBoost, y en el caso de las mujeres elegimos el modelo radial de SVM en base a los parámetros obtenidos.

Tabla 1

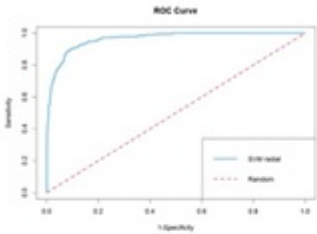
Variable	Mujeres	Hombres	p
N	10.362	10.927	
Edad (años)	34,7 (17,7)	33,7 (17,6)	0,000
IMC (Kg/m ²)	27,6 (7,1)	26,7 (5,9)	0,000
Perímetro abdominal (cm)	91,4 (16,7)	93,4 (16,9)	0,00
Circunferencia del brazo (cm)	31 (5,5)	32,6 (5)	0,00
Raza			
Latino-México	2.138	2.254	0,07
Latino-Sur	694	829	
Blanco	3.864	3.996	
Negro (no latino)	2.573	2.652	
Otras	1.093	1.196	
Albúmina (g/dl)	4,2 (0,3)	4,4 (0,3)	0,00
GOT (U/L)	22,9 (14,6)	27,7 (23,8)	0,00
GPT (U/L)	20,4 (15,1)	28,2 (27,7)	0,00
Creatinina (mg/dl)	0,7 (0,3)	0,9 (0,4)	0,00
BUN (mg/dl)	10,9 (4,3)	12,4 (4,6)	0,00
Colesterol total (mg/dl)	187 (41,3)	183,4 (42,4)	0,00
Glucosa (mg/dl)	94,5 (31,6)	96,8 (31,7)	0,00

Proteínas totales (g/dl)	7,2 (0,4)	7,3 (0,4)	0,00
Urato (mg/dl)	4,6 (1,1)	5,9 (1,2)	0,00
ASMI (Kg/m ²)	6,8 (1,4)	8,4 (1,5)	0,00

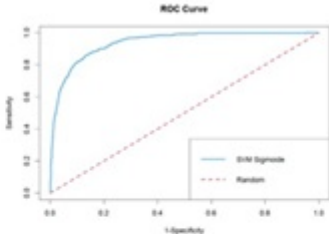
Tabla 2

	<i>Hombres</i>			<i>Mujeres</i>		
<i>Variable</i>	SVM-Radial	SVM-Sigmoide	XGBoost	SVM-Radial	SVM-Sigmoide	XGBoost
<i>Precisión</i>	0,932	0,917	1	0,878	0,866	0,827
<i>Sensibilidad</i>	0,686	0,609	1	0,661	0,632	0,172
<i>Especificidad</i>	0,977	0,614	1	0,936	0,929	0,997
<i>VPP</i>	0,837	0,967	1	0,734	0,704	0,939
<i>VPN</i>	0,948	0,761	1	0,911	0,904	0,823
<i>Balance</i>	0,832	0,935	1	0,798	0,780	0,584

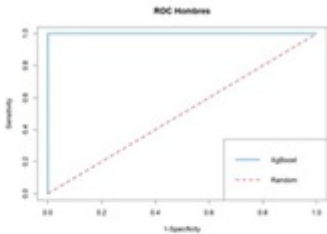
HOMBRES



SVM Radial en hombres

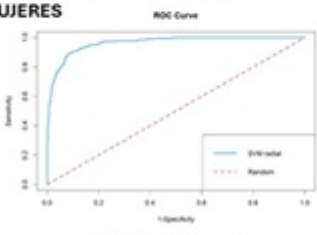


SVM sigmoide en hombres

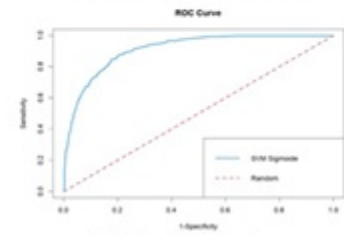


Extreme Gradient Boosting en hombres

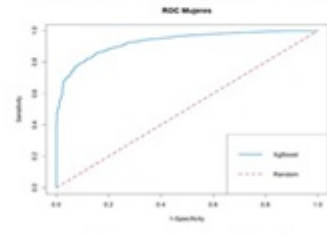
MUJERES



SVM Radial en mujeres



SVM sigmoide en mujeres



Extreme Gradient Boosting

Conclusiones: Nuestros resultados muestran una excesiva precisión al clasificar a los varones posiblemente relacionada con un sobreentrenamiento del modelo, y una aceptable precisión en general, que mejora otros modelos clásicos y que precisarán en cualquier caso de una validación externa posterior que confirme la utilidad de estos resultados.

Bibliografía

1. Cruz-Jentoft AJ, Bahat G, Bauer J, Boirie Y, et al; Writing Group for the European Working Group on Sarcopenia in Older People 2 (EWGSOP2), and the Extended Group for EWGSOP 2 Sarcopenia: revised European consensus on definition and diagnosis. *Age Ageing*. 2019;48:16-31. doi: 10.1093/ageing/afy169.