



## 20671 - MODELO DE MACHINE LEARNING (RANDOM FOREST) PARA LA DETECCIÓN AUTOMÁTICA DE LA ESCLEROSIS MÚLTIPLE MEDIANTE EL ANÁLISIS ACÚSTICO DE LA VOZ

Delgado Hernández, J.<sup>1</sup>; Hernández Pérez, M.<sup>2</sup>; Abreu Rodríguez, R.<sup>2</sup>; Betancort Montesinos, M.<sup>3</sup>

<sup>1</sup>Facultad de Psicología y Logopedia. Universidad de La Laguna; <sup>2</sup>Servicio de Neurología. Complejo Hospital Universitario Nuestra Señora de Candelaria; <sup>3</sup>Departamento de Psicología Clínica, Psicobiología y Metodología. Universidad de La Laguna.

### Resumen

**Objetivos:** El objetivo de este trabajo es entrenar y probar un algoritmo de aprendizaje supervisado que permita detectar la EM a través del análisis de los parámetros acústicos de la voz.

**Material y métodos:** Participaron voluntariamente 110 personas con una media de edad de 50 años ( $DT = 10,3$ ) sin alteraciones orgánicas de la voz, 75 diagnosticadas de EM (GEM), con una evolución media de la enfermedad de 10,8 años, y 35 neurológicamente sanas (GC). Un 54,6% del GEM presenta un curso clínico RR, un 38,6% SP y un 6,6% PP. Se grabó a cada participante una vocal sostenida durante cuatro segundos con Praat.

**Resultados:** Se aplicó un modelo *random forest* con validación cruzada. Se estableció como variable criterio el grupo (GC y GEM) y como predictoras los parámetros acústicos: F0DS, Shimmer, HNR, CPPS y GNE. Se dividió la muestra aleatoriamente para la fase de entrenamiento del modelo (70%) y para la fase de test (30%). Los resultados muestran un  $mtry = 2$  con una precisión del 0,83, OBB = 16,83% y fiabilidad del 0,58 (kappa). El modelo fue probado en la fase test con una sensibilidad del 80% y una especificidad del 90% (AUC-ROC = 0,93). Se comparó la precisión en la clasificación de diferentes algoritmos de aprendizaje supervisado (Partial Least Squares, Random Forest, K-Nearest Neighbors algorithm y Decision Tree Classification) siendo el *random forest* utilizado en este estudio el que presenta una mayor precisión y fiabilidad.

**Conclusión:** El modelo de *machine learning* propuesto es capaz de clasificar automáticamente a personas con y sin EM con una alta sensibilidad y especificidad.