



Línea de Trabajo fin de Máster 2025-2026

Máster Universitario en Estadística Aplicada.	
Título	Métodos de remuestreo para el tratamiento de datos desbalanceados en problemas de clasificación
Tipo	INVESTIGACIÓN ☒ ORIENTACIÓN PRÁCTICA ☐
Número de alumnos admitidos	1
Profesor(es)/ email	José Antonio Sáez Muñoz (joseasaezm@ugr.es)
Descripción	En numerosos contextos reales, los conjuntos de datos presentan distribuciones desbalanceadas entre clases, lo que dificulta la construcción de modelos de clasificación robustos y precisos. Este trabajo propone que los estudiantes estudien en profundidad las principales técnicas de remuestreo, como el submuestreo, el sobremuestreo y métodos avanzados como SMOTE y sus variantes. Asimismo, se abordarán estrategias combinadas con algoritmos de clasificación y métricas específicas para evaluar el desempeño en escenarios desbalanceados. Los estudiantes aplicarán estas metodologías a conjuntos de datos reales y simulados, compararán resultados y discutirán las ventajas y limitaciones de cada enfoque. El objetivo es que adquieran tanto una base teórica sólida como habilidades prácticas para enfrentar este tipo de problemas en la investigación y en aplicaciones profesionales.
Objetivos particulares	• Desarrollar tareas de investigación en un entorno multidisciplinar. • Comprender los principales retos relacionados con el remuestreo para el tratamiento de datos desbalanceados. • Aplicar y evaluar técnicas de remuestreo/clasificación en distintos conjuntos de datos.
Prerrequisitos y recomendaciones	• Conocimientos avanzados en el lenguaje R. • Conocimientos sobre técnicas de clasificación y su aplicación.
Plan de trabajo	• Revisión bibliográfica. • Análisis comparativo de las diferentes metodologías y paquetes en R. • Implementación de funciones en R para el remuestreo en datos no balanceados. • Aplicación de las técnicas a conjuntos de datos reales. • Interpretación y discusión de los resultados.
Competencias generales y específicas	CB: 6, 7, 8, 9, 10 CG: 1, 2, 3, 6, 9 CE: 10, 13, 15, 16, 17, 18, 19, 20, 22, 23, 26, 28, 29
Bibliografía	• Chawla, N. V., Bowyer, K. W., Hall, L. O., & Kegelmeyer, W. P. (2002). SMOTE: Synthetic minority over-sampling technique. <i>Journal of Artificial Intelligence Research</i>, 16, 321–357. • Carvalho, M., Pinho, A. J., & Brás, S. (2025). Resampling approaches to handle class imbalance: A review from a data perspective. <i>Journal of Big Data</i>, 12(1), 71. • Kumar, P., Bhatnagar, R., Gaur, K., & Bhatnagar, A. (2021). Classification of imbalanced data: Review of methods and applications. <i>IOP Conference Series: Materials Science and Engineering</i>, 1099(1), 012077.