



Línea de Trabajo fin de Máster 2024-2025

(Fecha última actualización:04/10/2024)

Máster Universitario en Estadística Aplicada.	
Título	Incrementando la robustez al ruido en clasificación con el uso de ensembles de clasificadores
Tipo	INVESTIGACIÓN <input checked="" type="checkbox"/> ORIENTACIÓN PRÁCTICA <input type="checkbox"/>
Número de alumnos admitidos	1
Profesor(es)/ email	José Antonio Sáez Muñoz (joseasaezm@ugr.es)
Descripción	La presencia de ruido en los datos es un factor crítico que puede afectar negativamente al rendimiento de los modelos de clasificación. Los algoritmos de ensemble han demostrado ser herramientas efectivas para incrementar la robustez y precisión de los modelos, particularmente en situaciones donde los datos contienen ruido. Ensembles como Bagging, Random Forest y XGBoost combinan múltiples clasificadores individuales para reducir la varianza y mejorar la estabilidad y precisión general del modelo, aprovechando la diversidad de los mismos. Este trabajo tiene como objetivo estudiar cómo los ensembles de clasificadores pueden mejorar la robustez frente al ruido, comparando su rendimiento con el de los componentes individuales en escenarios donde se introducen distintos niveles de ruido. Se explorarán diversas técnicas de ensemble y se analizará su comportamiento sobre diferentes conjuntos de datos, para evaluar su capacidad de mantener un alto rendimiento en la presencia de ruido.
Objetivos particulares	<ul style="list-style-type: none"> • Desarrollar tareas de investigación en un entorno multidisciplinar. • Comprender los principales retos relacionados con los datos erróneos en problemas de clasificación. • Aplicar y evaluar técnicas de ensembles de clasificadores en distintos conjuntos de datos.
Prerrequisitos y recomendaciones	<ul style="list-style-type: none"> • Conocimientos avanzados en el lenguaje R. • Conocimientos sobre técnicas de clasificación y su aplicación.
Plan de trabajo	<ul style="list-style-type: none"> • Revisión bibliográfica. • Análisis comparativo de las diferentes metodologías y paquetes en R. • Implementación de funciones en R para la creación de ensembles de clasificadores. • 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	<ul style="list-style-type: none"> • Oza, N. C., & Tumer, K. (2008). Classifier ensembles: Select real-world applications. <i>Information Fusion</i>, 9(1), 4-20. • Parmar, A., Katariya, R., & Patel, V. (2019). A review on random forest: An ensemble classifier. In <i>International Conference on Intelligent Data Communication Technologies and Internet of Things (ICICI) 2018</i> (pp. 758-763). Springer International Publishing. • Zhu, X., & Wu, X. (2004). Class noise vs. attribute noise: A quantitative study. <i>Artificial Intelligence Review</i>, 22, 177-210.