

March 6<sup>th</sup>, 2019

## Soluciones consistentes para problemas de clasificación desequilibrada

**Aníbal R. Figueiras Vidal**

Universidad Carlos III de Madrid, Spain

### Abstract

Se denominan desequilibrados los problemas de clasificación en que las diferencias de tamaño de las poblaciones de las distintas clases o/y de los costes de los errores de clasificación son tan acusadas que las máquinas de clasificación convencionales proporcionan pobres prestaciones. Dichos problemas son tanto frecuentes como importantes: defensa ante intrusiones y otras alarmas de seguridad, diagnóstico médico, atención personalizada, concesión de créditos, lucha contra el fraude, detección de defectos y averías, mantenimiento preventivo y previsión de cambios sociales son ejemplos que lo acreditan



Desde hace más de dos decenios se han venido proponiendo procedimientos para salvar las dificultades que plantea el desequilibrio

y demostrando la utilidad de muchos de ellos. Sin embargo, la inmensa mayoría de estas técnicas carecen de fundamentos analíticos que permitan controlar su manejo: por ello es posible que no ofrezcan resultados de verdadera calidad, e incluso, lo que es aún más grave, que sus prestaciones ni siquiera lleguen a las que ofrece un diseño que emplee los datos desequilibrados.

En esta presentación se considerarán problemas binarios; para ellos, una reflexión sobre la teoría bayesiana permite concluir que puede establecerse un par de condiciones necesarias y suficientes para que el reequilibrado que se aplique permita estimar la característica de operación de Neyman-Pearson (óptima), y así garantizar que conduce a resultados potencialmente mejores que otras alternativas y que se evitan los riesgos de degradación: es decir, la consistencia de esta familia de procedimientos.

Una presentación de las direcciones abiertas para extender estas investigaciones concluirá la exposición.

## Short bio



Aníbal R. Figueiras Vidal obtuvo su título de Ingeniero de Telecomunicación para la UPM (1973) y el de Doctor de igual denominación por la UPC (1976). Ha prestado sus servicios en las dichas universidades y las de Santiago de Compostela y Carlos III de Madrid, donde actualmente se encuentra como Catedrático Numerario, nivel que alcanzó en 1978.

Los ámbitos de investigación a los que se dedica son el Tratamiento de Señales y el Aprendizaje Máquina, junto con sus aplicaciones. Es (co)autor de casi un centenar de artículos JCR, ha dirigido el mismo número de proyectos y contratos de investigación

(una cuarta parte de proyectos y contratos de investigación (una cuarta parte de ellos internacionales) y supervisado 33 tesis doctorales.

Entre otras distinciones, es Académico Numerario de la Real de Ingeniería, cuya presidencia desempeñó entre 2007 y 2011; “Life Fellow” del IEEE; Doctor “Honoris Causa” por las universidades de Vigo y San Pablo (Arequipa, Perú); y se le concedió el Premio “Miguel Catalán” a la trayectoria investigadora por la Comunidad de Madrid en 2018.