

## Propuesta de Trabajos Fin de Grado, curso académico 2018-19

PROFESOR: Amparo Baíllo Moreno

### 1.- TÍTULO: Aprendizaje estadístico "con muchos ceros" (*sparsity*)

Resumen/contenido:

Los conjuntos de datos estadísticos actuales tienen en general un "exceso" de información; por ejemplo, en los problemas de regresión puede haber un número ingente de regresores en comparación con el tamaño de la muestra. En este contexto, es conveniente utilizar modelos estadísticos sencillos. Una manera de simplificar consiste en utilizar modelos *sparse* (podríamos traducirlo como "con muchos ceros"), que son aquéllos en los que sólo un número relativamente pequeño de parámetros son distintos de cero. La idea de este trabajo es leer y trabajar algunos capítulos del libro de Hastie et al. (2015) para entender cómo se plantean y ajustan este tipo de modelos en problemas estadísticos habituales (como regresión lineal, modelos lineales generalizados, clasificación,...). Todos los procedimientos estudiados se probarían sobre conjuntos de datos reales con librerías del programa R. Es importante que el estudiante haya cursado la asignatura Estadística I y al menos esté matriculado en Estadística II. No es necesario saber manejar R, si se sabe manejar un mínimo de Matlab u otro lenguaje de programación.

Bibliografía/referencias:

Hastie, T., Tibshirani, R., Wainwright, M. (2015). [\*Statistical Learning with Sparsity. The Lasso and Generalizations\*](#). Chapman and Hall/CRC