

Propuesta de Trabajos Fin de Grado, curso académico 2021-22

PROFESOR: Amparo Baíllo Moreno

Número máximo de TFG que solicita dirigir: 2

1.- TEMA: Procesos puntuales espaciales

Válido para 1 alumno.

Resumen/contenido: Un proceso puntual espacial es un conjunto de datos formado por las localizaciones (en el plano o en el espacio) de objetos o eventos. Este tipo de datos aparecen en múltiples disciplinas: localizaciones de árboles o animales en un terreno, depósitos de un mineral en una prospección geológica, epicentros de terremotos, localizaciones de móviles, casos de una enfermedad infrecuente, ... El objetivo de este trabajo es estudiar los métodos estadísticos utilizados para analizar la distribución de un proceso puntual y aplicarlos en el análisis de conjuntos de datos reales (mediante el programa R).

Requisitos: No es necesario saber programar en R (basta tener interés por aprender y haber utilizado algún programa tipo Matlab o Python).

Asignaturas de cuarto relacionadas/compatibles: Estadística II

Bibliografía/referencias:

Baddeley, A., Rubak, E. y Turner, R. (2016). *Spatial Point Patterns. Methodology and Applications with R*. Chapman & Hall/CRC.

2.- TEMA: Árboles y *random forests*

Válido para 1 alumno

Resumen/contenido: En el problema de regresión el objetivo es describir los valores de una variable respuesta cuantitativa en términos de otras variables cuantitativas llamadas regresores o covariadas. La función de regresión de la respuesta sobre los regresores se estima en base a la información proporcionada por una muestra de todas las variables en n individuos. En el problema de clasificación supervisada, también llamado “reconocimiento de patrones”, la variable “respuesta” es cualitativa e indica el grupo al que pertenece el individuo. El objetivo en este caso es clasificar correctamente a un individuo del que se conocen sus valores de unas variables auxiliares llamadas características (*features*), basándose en la información de una muestra de n individuos correctamente clasificados, en los que se han observado los valores de las características de interés. El objetivo del trabajo será, bien en el contexto de la regresión o en el de la clasificación, estudiar la técnica de árboles y su versión agregada, los *ranfom forests*, que es un procedimiento de referencia en la actualidad. Se practicará lo aprendido analizando un conjunto de datos reales con R.

Requisitos: No es necesario saber programar en R (basta tener interés por aprender y haber utilizado algún programa tipo Matlab o Python).

Asignaturas de cuarto relacionadas/compatibles: Estadística II

Bibliografía/referencias:

Devroye, L., Györfi, L. and Lugosi, G. (1996). *A Probabilistic Theory of Pattern Recognition*. Springer.

Hastie, T., Tibshirani, R. and Friedman, J. (2009). *The Elements of Statistical Learning: Data Mining, Inference and Prediction*. Springer.

<https://web.stanford.edu/~hastie/Papers/ESLII.pdf>