



Fundamentos de la Ciencia de Datos. Prueba de Laboratorio 3 (PL3)

La Prueba de Laboratorio constará de dos partes:

1. La primera parte consiste en la realización de dos ejercicios en clase con ayuda del profesor:
 - 1.1. En el primero se va a realizar un análisis de clasificación de Datos con R aplicando todos los conceptos vistos en el tema. Se resolverá, utilizando el algoritmo construcción de árboles de decisión de Hunt, el mismo problema que el visto en la descripción teórica del tema. Es decir, para la misma muestra¹ que se utilizó para hacer de forma manual el primer ejercicio de clasificación para datos cualitativos, se deberá obtener, utilizando la medida de Ganancia de información, mediante la medida de impureza Gini, la función de clasificación.
 - 1.2. En el segundo para los datos de un fichero .txt generado a partir de los datos de la muestra² utilizada para hacer el segundo ejercicio de clasificación para datos cuantitativos, hacer un análisis de regresión lineal.
2. La segunda parte consiste en la realización de dos ejercicios:
 - 2.1. Desarrollo por parte de cada alumno del enunciado y la solución de un ejercicio en el que se realice un análisis con R de clasificación supervisada con árboles introduciendo modificaciones sobre el ejercicio hecho en clase.
 - 2.2. Desarrollo por parte de cada alumno del enunciado y la solución de un ejercicio en el que se realice un análisis con R de clasificación supervisada con regresión introduciendo modificaciones sobre el ejercicio hecho en clase

Se debe completar en el documento del grupo utilizando Sweave y Noweb, para realizar Literate-Programming y entregar los archivos .Rnw, .tex y .pdf generados por R con el resultado de las dos partes de la PL3 y los ficheros de datos utilizados. La PL3 se entregará como actividad en Blackboard. El nombre de los ficheros será: Grupo.NúmerodeGrupo-PL3

¹ Está formado por las siguientes 9 calificaciones de estudiantes: 1. {A, A, B, Ap}; 2. {A, B, D, Ss}; 3. {D, D, C, Ss}; 4. {D, D, A, Ss}; 5. {B, C, B, Ss}; 6. {C, B, B, Ap}; 7. {B, B, A, Ap}; 8. {C, D, C, Ss}; 9. {B, A, C, Ss}, donde las características de las calificaciones son: {Teoría, Laboratorio, Prácticas, Calificación Global}.

² Está formada por los siguientes 4 radios ecuatoriales y densidades de los planetas interiores : {Mercurio, 2.4, 5.4; Venus, 6.1, 5.2; Tierra, 6.4, 5.5; Marte, 3.4, 3.9}