

ANÁLISIS DE DATOS MULTIVARIANTES

Grado en Matemáticas. Doble Grado en Matemáticas y Estadística

Conjuntos de datos para los trabajos

Se trabajará con los ficheros “train.csv” y “test.csv”.

Objetivo: construir modelos de predicción de la variable mortality_rate a partir del resto de variables (excepto ID que es una identificación del registro).

Los modelos se construyen sobre el conjunto train.csv y se evalúan sobre el conjunto test.csv.

Entre los criterios para medir la calidad de las predicciones se incluirán al menos: R^2 y Raíz cuadrada del Error Cuadrático Medio.

También se construirán gráficos, al menos la nube de puntos comparando valores observados y predicciones.

Cómo leer en R estos datos:

```
train = read.csv("train.csv")  
test= read.csv("test.csv")
```