

Práctica 3

```
# N : numero de habitantes
# R : radio de la poblacion
# S : funcion desconocida del número de habitantes
# T : funcion desconocida del radio de la poblacion
# U : datos aleatorios con el mismo rango que la solución
# G : datos aleatorios con la misma media y desviación típica que la solución
# L : longitud de cable en la población
```

En el archivo 'cables.csv' se encuentran datos de un problema real de modelado de la longitud de línea eléctrica en poblaciones que se explicará en la sesión presencial. Los datos pueden leerse con las siguientes instrucciones:

```
import pandas as pd
cables_filename = 'cables.csv'
cables = pd.read_csv(cables_filename, sep=',', decimal='.')
```

El objeto de esta práctica es:

- Limpiar el dataset, quitando los valores perdidos
- Escalar o normalizar las variables
- Detectar las variables irrelevantes o redundantes
- Construir un modelo lineal y otro con random forest
- Realizar la validación cruzada de ambos modelos y decidir cuál es la precisión del modelo conseguido