

Actividad 3

Clasificación con K-NN

Usar un clasificador K-NN con el conjunto de datos 'Caravan', que es parte del paquete 'ISLR'. Este conjunto de datos incluye 85 predictores que miden características demográficas de 5.822 personas. La variable de respuesta es *Purchase* (Compra), que indica si un individuo determinado compra o no una póliza de seguro de caravanas. En este conjunto de datos, sólo el 6.3% de las personas adquirieron un seguro de caravana, es decir, las clases no están balanceadas.

```
> dim(Caravan)
[1] 5822  86

> summary(Caravan$Purchase)
  No  Yes
5474 348
```

Comparar el resultado, dando las matrices de confusión y los ratios de acierto, para distintos valores de k y dibujar las correspondientes curvas ROC, al menos para la mejor y la peor predicción encontradas.

Probar además distintos tamaños para los conjuntos de entrenamiento y test y valorar cual es el mejor porcentaje para la partición del *dataset*.