



Aprendizaje Automático I

Ejercicio: Métricas de evaluación

DSLAb

noviembre, 2023

El objetivo de este ejercicio es construir un **árbol de decisión** para el conjunto de datos `adult` de la librería `liver`.

1. Como primera parte del ejercicio, explora las características básicas de las variables en la base de datos.
2. Divide la base de datos en entrenamiento y validación. ¿Qué porcentaje eliges para cada partición? Justifícalo.
3. Empleando la función `trainControl` de la librería `caret` fija el número de particiones que consideres oportunas para entrenar el modelo en la de entrenamiento.
4. Entrena un modelo `knn` empleando como variables explicativas la edad (`age`) y el número de horas por semana `hours.per.week`. Evalua el rendimiento del modelo.
5. A continuación entrena un **árbol de decisión** empleando las variables que consideres oportunas. Justifica tu respuesta.
6. Obten una visualización del mejor árbol entrenado.
7. ¿Cómo funciona sobre la muestra de validación?