



Aprendizaje Automático I

Ejercicio: Árboles de Decisión

DSLlab

noviembre, 2023

El objetivo de este ejercicio es construir un **árbol de decisión** para el conjunto de datos `Carseats` de la librería `ISLR2`.

1. Como primera parte del ejercicio, explora las características básicas de las variables en la base de datos. Además, construye una nueva variable respuesta `High` de tal forma que:

$$High = \begin{cases} "No", & \text{si } Sales \leq 8 \\ "Yes", & \text{otro caso} \end{cases}$$

2. Divide la base de datos en entrenamiento y validación. ¿Qué porcentaje eliges para cada partición? Justifícalo.
3. A continuación entrena un **árbol de decisión** empleando las variables que consideres oportunas. Justifica tu respuesta.
4. Obten una visualización del árbol entrenado.
5. ¿Qué variable es la más importante para conseguir ventas elevadas? ¿Qué valores de dicha variable están asociados a un mayor número de ventas? Obten una visualización que refleje ese hecho.
6. ¿Cuál es el error de clasificación? ¿Qué valor es mayor, la **precisión** o la **recuperación**? ¿Qué significado tiene?
7. Si existe sobreajuste en el modelo, ¿qué podrías hacer para corregirlo?