



Guía de ejercicios de clasificación con árboles de decisión

Proponemos trabajar con un conjunto de datos de árboles de 4 especies: Jacarandá, Ceibo, Pindó y Eucaliptus. Cada árbol tiene 3 atributos (altura, ancho, inclinación) y la variable a predecir es la especie.



Primero deben descargar el archivo que está en el campus.

Ejercicios

1. Explorar los datos, graficar histogramas de los atributos.
2. Graficar scatterplot diámetro-altura con colores por especie.
3. Entrenar un árbol de decisión para clasificar árboles
 - a. probar con distintas profundidades
 - b. probar con criterio de información y de gini
 - c. probar no utilizar todos los atributos

En cada caso graficar el árbol de decisión y registrar la exactitud.

4. Elegir uno de los modelos entrenados
 - a. graficar el árbol de decisión
 - b. ¿cuál es el atributo utilizado en el primer corte?
5. Dado un árbol que mide 22 m, diámetro 56 e inclinación 8°:
¿cuál especie es?