Instructions



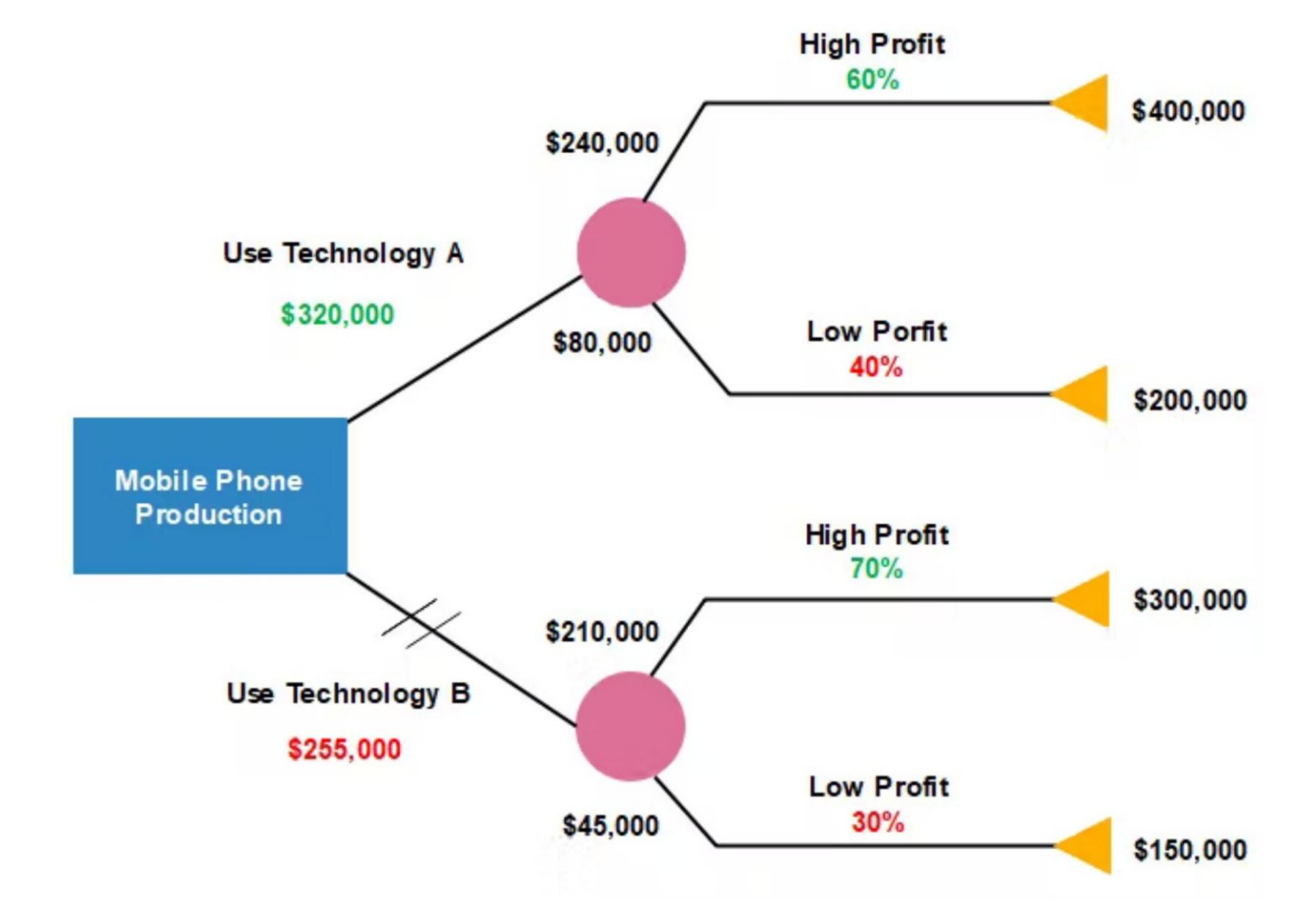
Arboles de Decisión



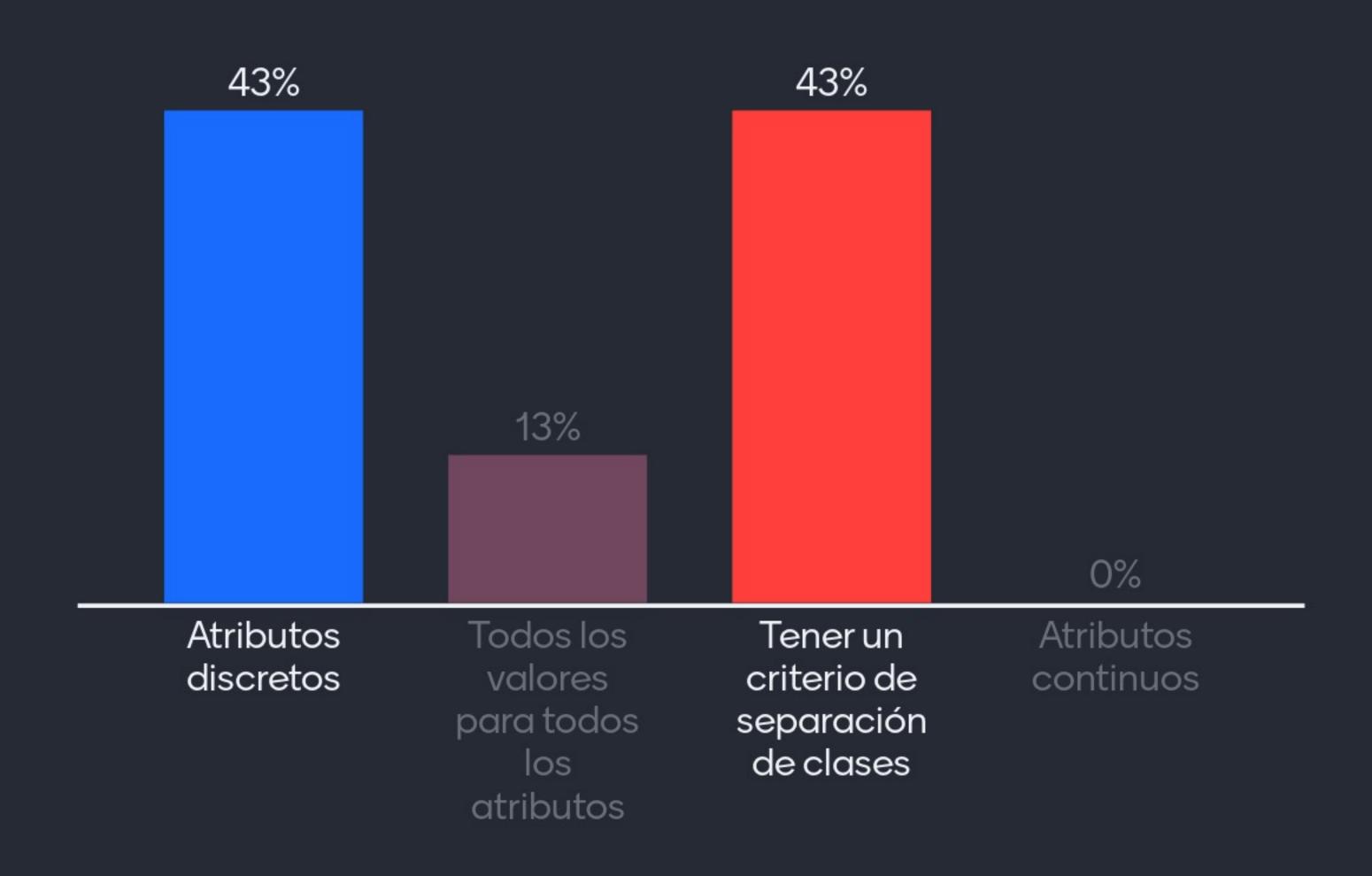






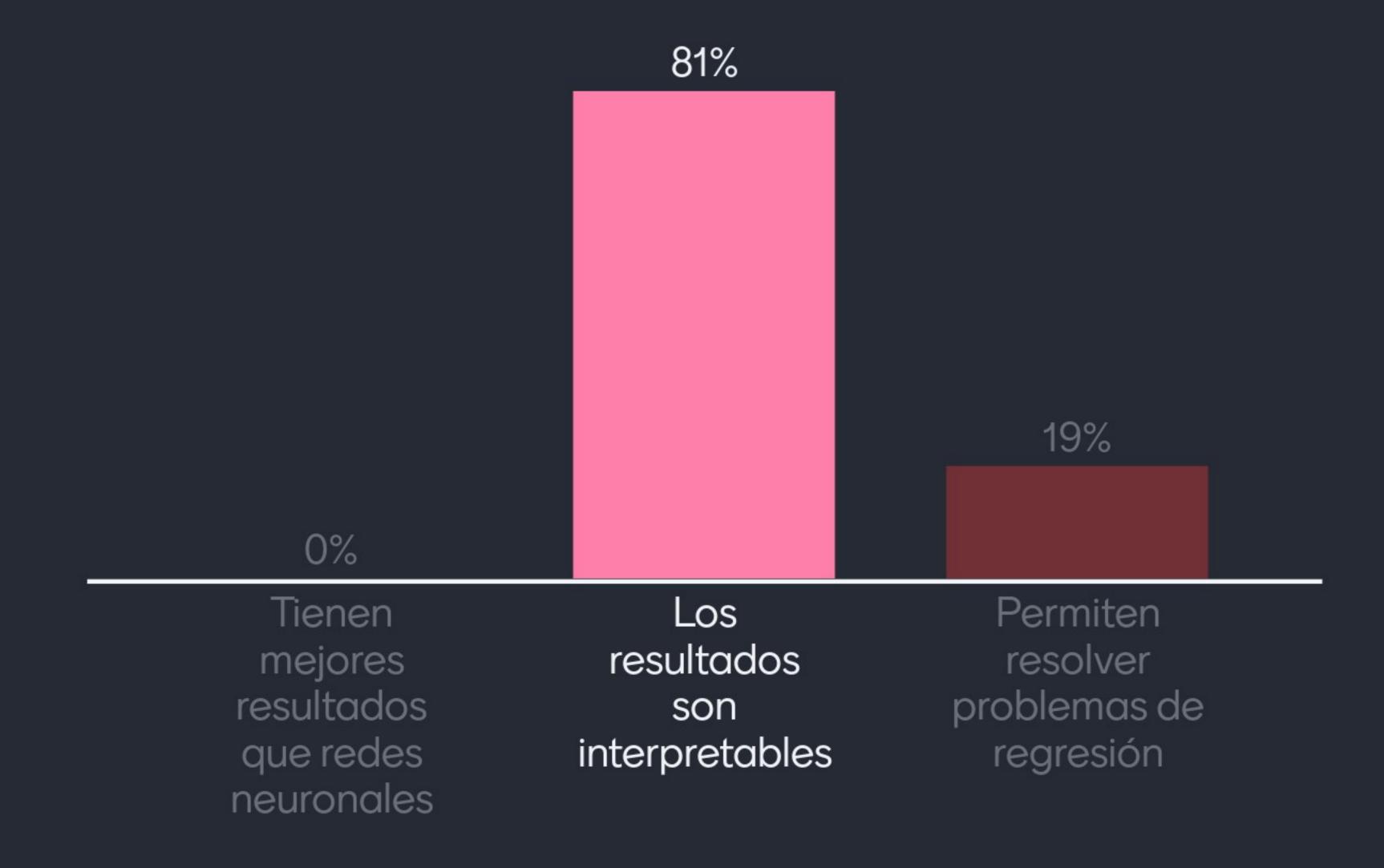


¿Cuáles son características principales de un árbol de decisión?

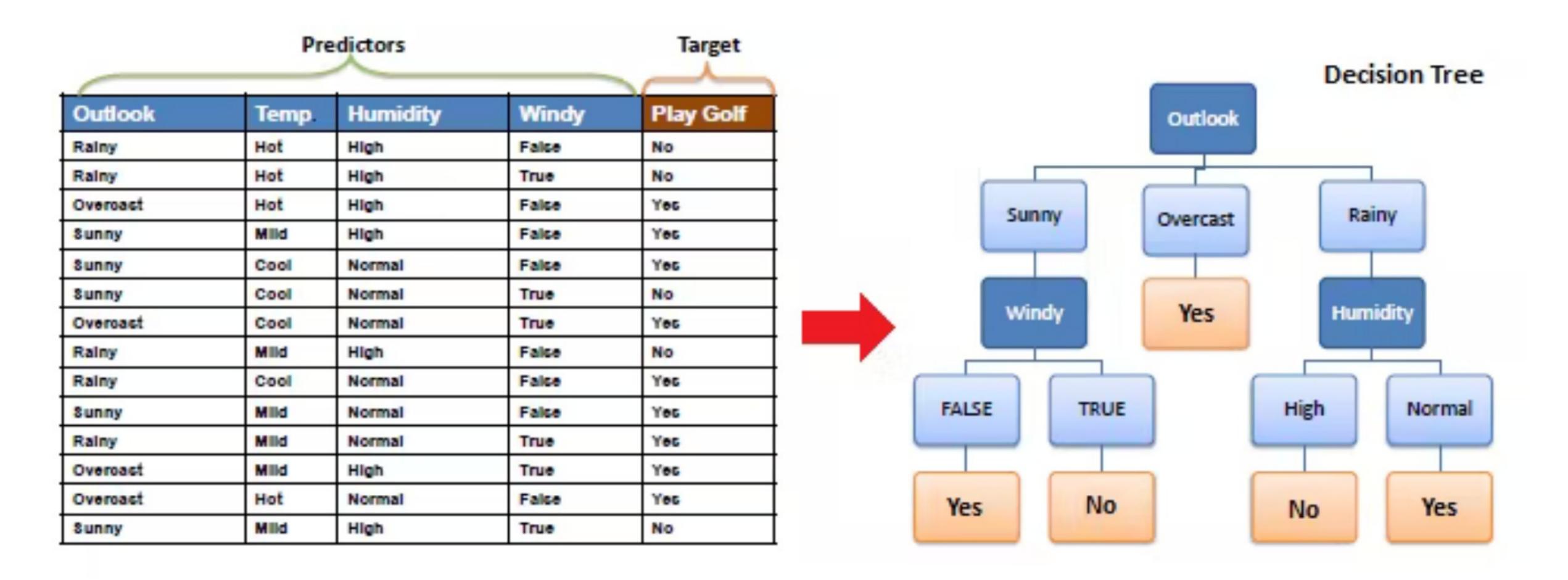




¿Cuáles son ventajas de un árbol de decisión?





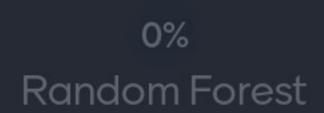


Sobre la construcción de un árbol





Cual es la mejor forma de determinar el atributo más discriminativo?









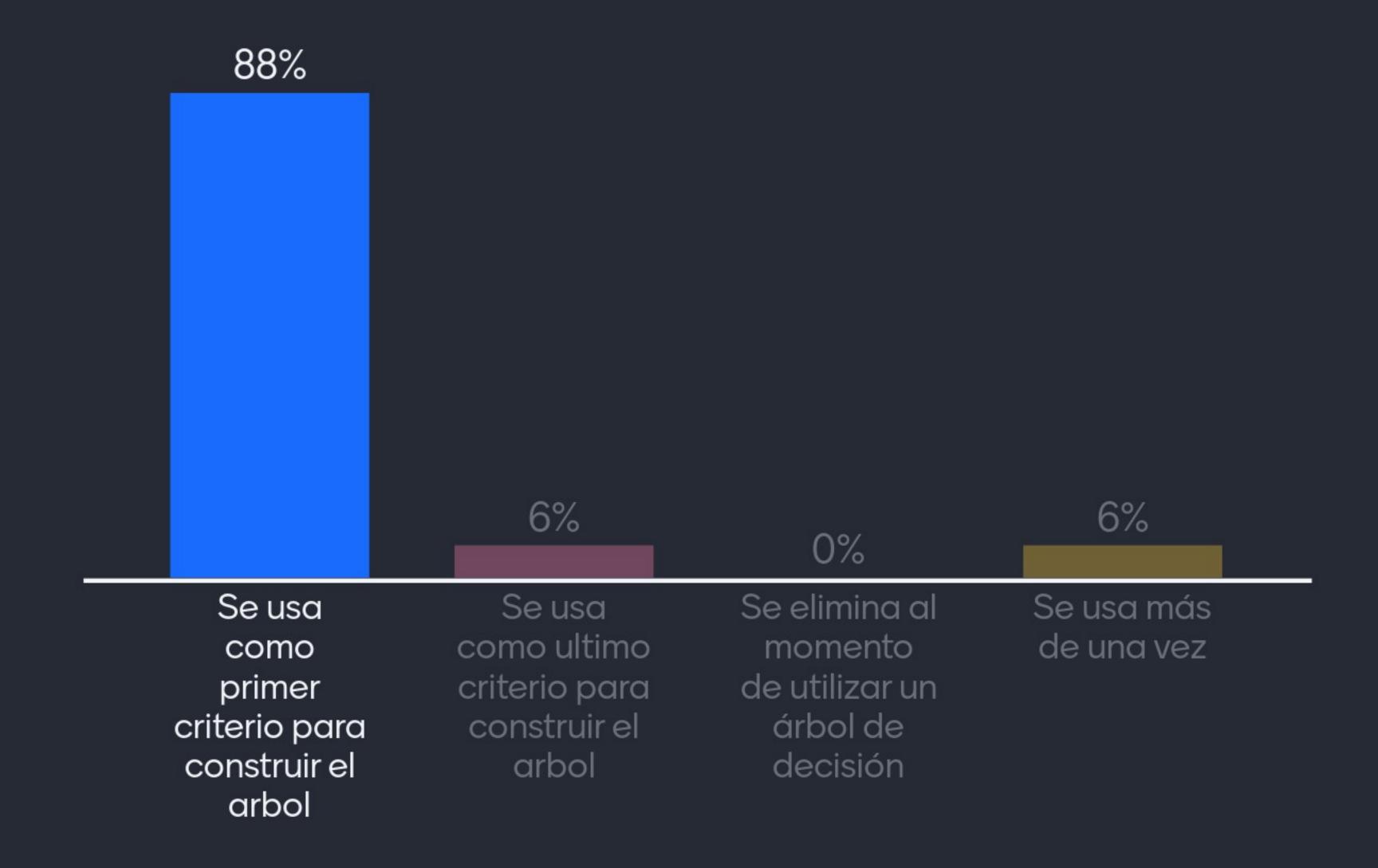


Ejemplo 1:

- 10 registros con clase A
- 20 registros con clase B
- 30 registros con clase C
- 40 registros con clase D
- Entropía= -[(.1 log .1) + (.2 log .2) + (.3 log .3) + (.4 log .4)]
- Entropía = 1.85



Que sucede con este atributo discriminativo?





¿Cuándo conviene utilizar GainRatio?

0%

Cuando se usa Random Forest

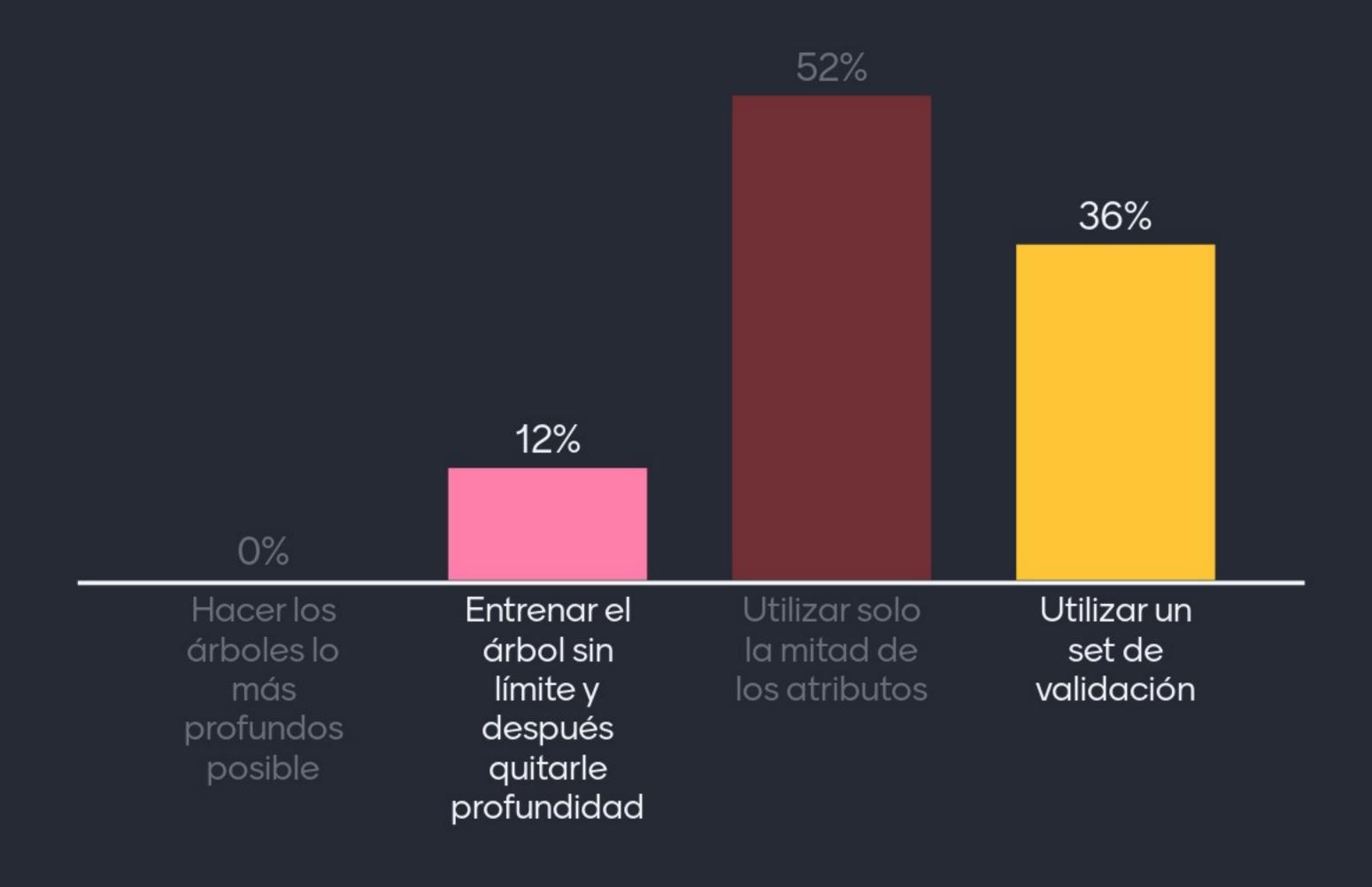
0%

Cuando se tienen pocos atributos



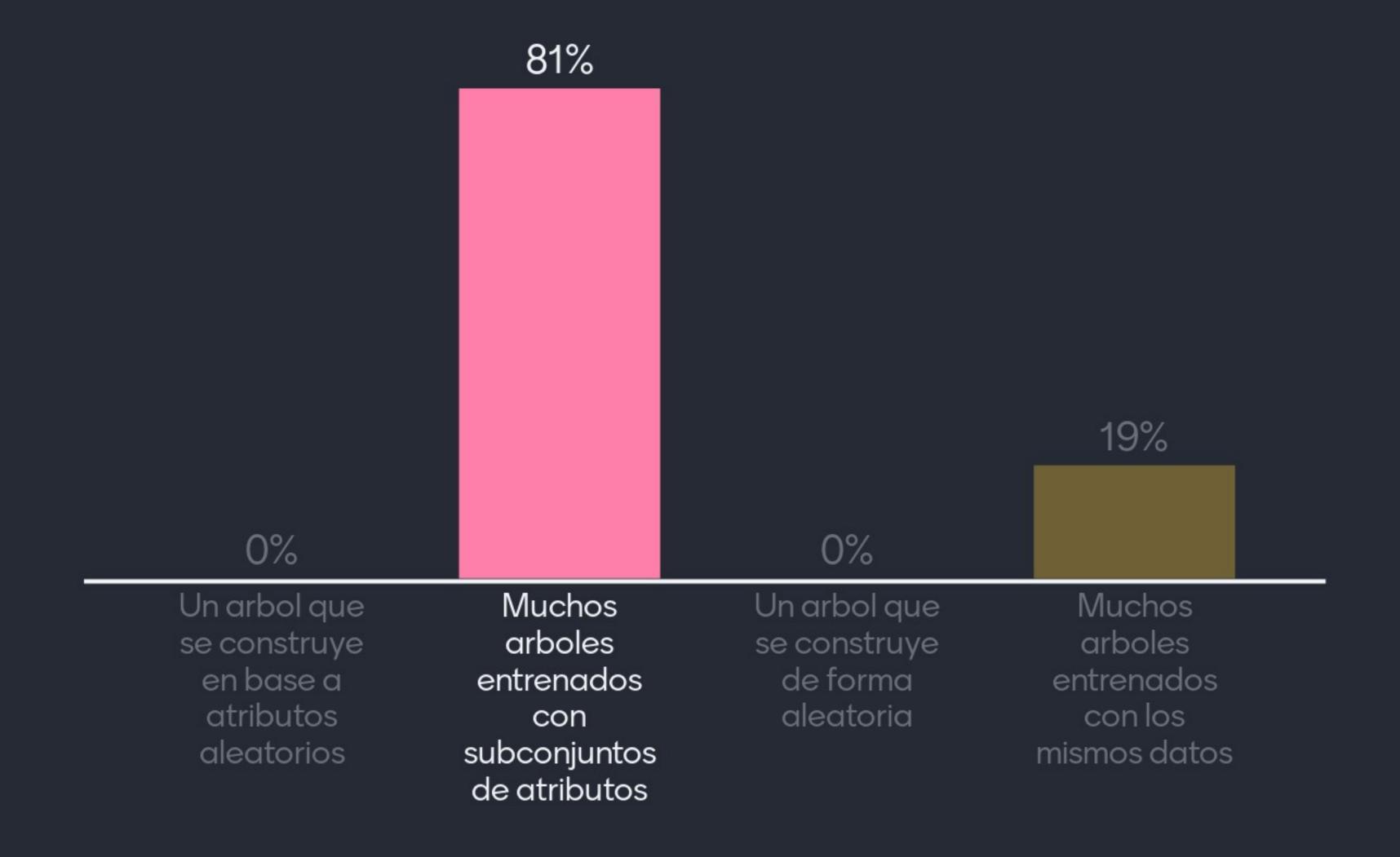


¿Cómo se puede evitar el "overfitting" en árboles de decisión?

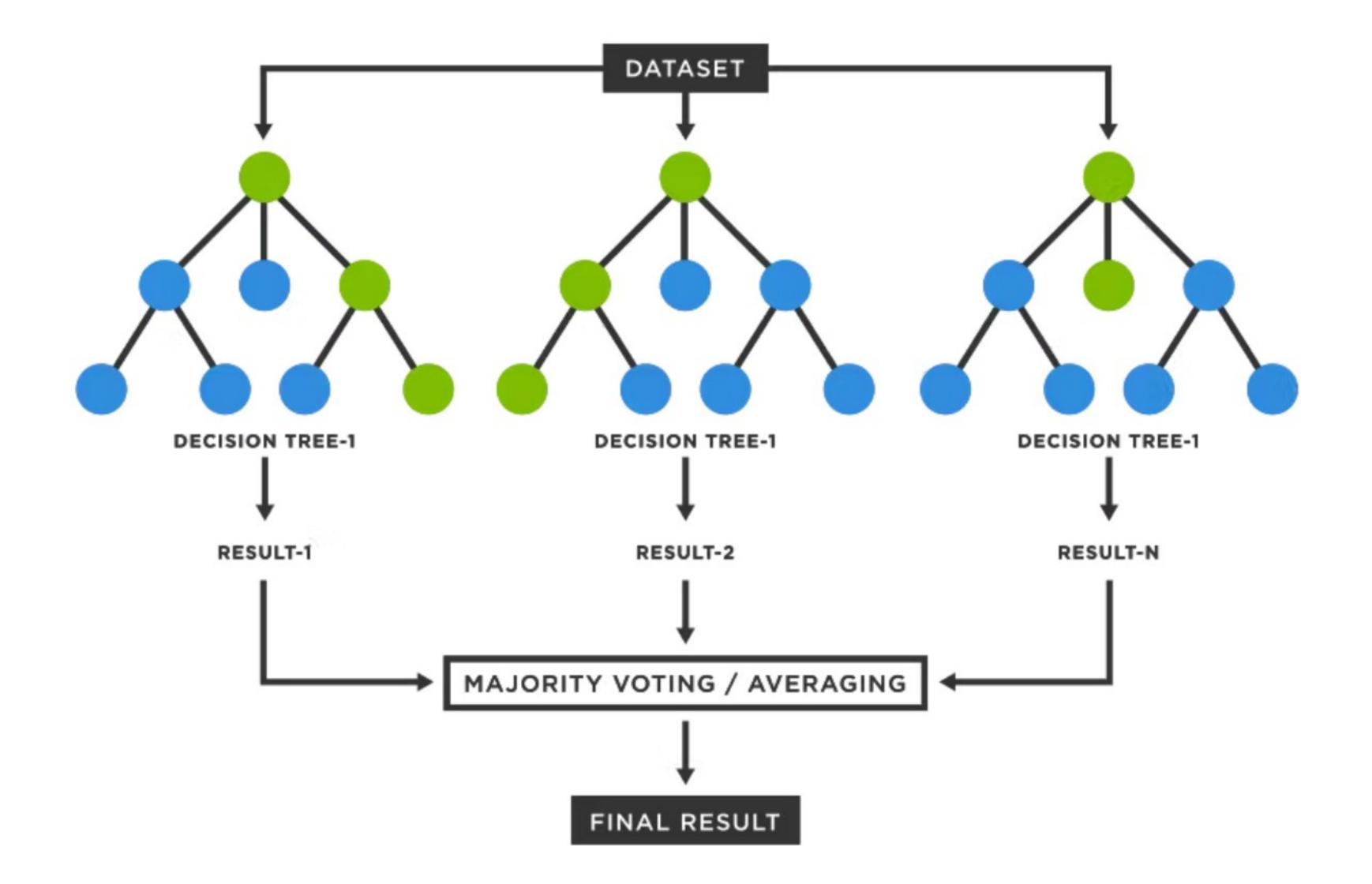




Que es un RandomForest









Sobre RandomForest

Si el # de atributos para cada árbol aumenta, entonces disminuye la diversidad pero aumenta la exactitud.

RandomForest soluciona el problema de overfitting?

