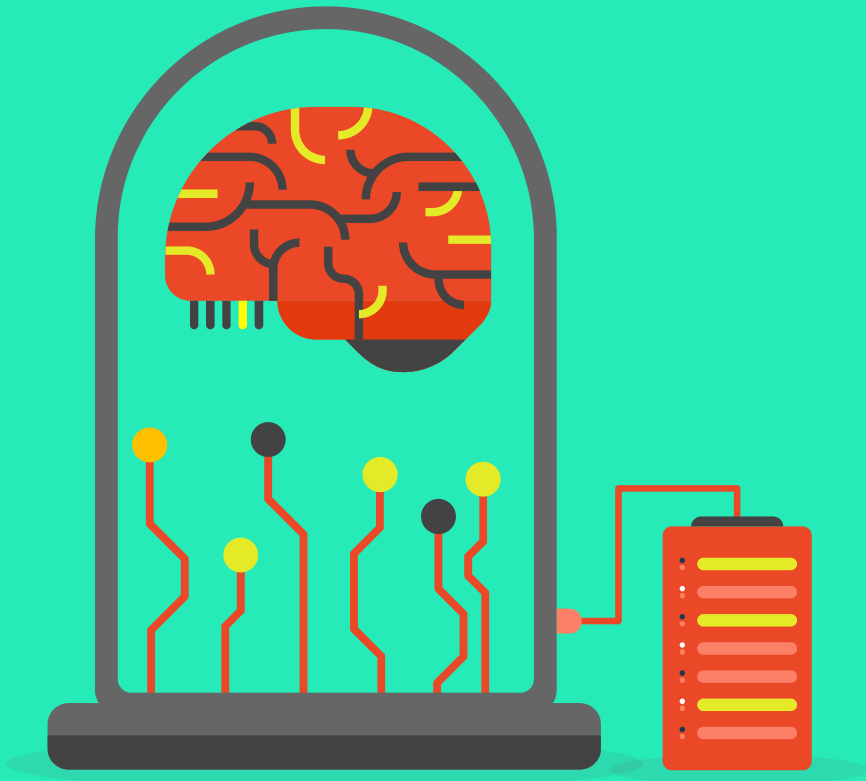


## 14. Clasificación con Bosques Aleatorios



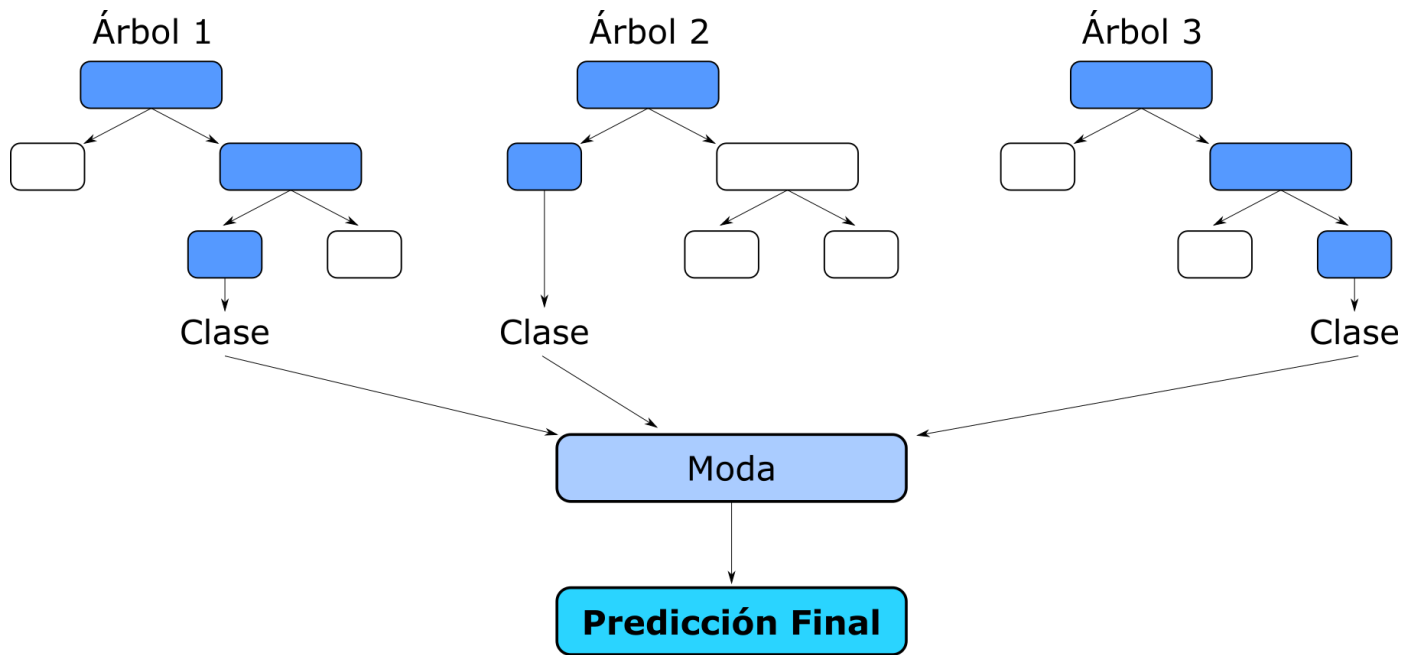
# Métodos de Conjuntos (Ensemble Learning)

Como ya mencionamos anteriormente, los **Bosques Aleatorios** son un tipo de modelo de Conjuntos o *Ensemble*, en donde se toman un grupo de árboles de decisión entrenados de manera independiente y con características diferentes, y luego se construye con ellos un modelo de predicción en conjunto.

En el caso de la regresión, la predicción final se obtenía al promediar la salida de cada uno de los árboles. Cuando se trata de clasificación, la salida final estará dada por el voto de la mayoría (o la moda) entre las clases predichas por los distintos árboles que conforman el bosque.

# Bosques Aleatorios

De la misma forma, el bosque como conjunto es capaz de ofrecer mejores resultados de clasificación que cada uno de los árboles independientes:



# Bosques Aleatorios

Los principios vistos para el entrenamiento de los Bosques Aleatorios (bagging, pasting, etc) se cumplen de la misma forma para la clasificación, así como la selección e influencia de los hiperparámetros de este tipo de modelos de Machine Learning. Y en general, es modelo es aplicable tanto para casos de clasificación binaria como de múltiples clases.

Los modelos de clasificación basados en Bosques Aleatorios generalmente ofrecen resultados sumamente buenos, incluso llegando a sobre entrenar con facilidad en muchos casos, por lo que es importante la selección de los parámetros correctos de los mismos.

Especialmente en el caso de datos estructurados (tabulares), está junto con el XGBoost (otro tipo de algoritmo de ensemble) entre los algoritmos más empleados para la clasificación.

# **Décimo Quinto Notebook Práctico**

Implementemos la clasificación a partir de bosques aleatorios, aplicado sobre el mismo conjunto de datos que venimos trabajando y que ya conocemos, el UCI Wine Dataset.

## **Décimo Quinto Notebook Práctico**