

Benemérita Universidad Autónoma de Puebla  
Facultad de Ciencias de la Computación

# Máquinas de aprendizaje

## Reporte: Ejemplo árboles de regresión



# BUAP

**Docente: Abraham Sánchez López**

**Alumno**

Taisen Romero Bañuelos

**Matrícula**

202055209

## Árbol modelo (Cubist)

Los enfoques híbridos suelen tener ventajas importantes que los enfoques puros no tienen. En este caso, noto que el chiste de hacer primero un árbol y después regresión es que simplifiquemos los datos al hacerlos homogéneos y gracias a ello el modelo de regresión lineal se puede ajustar mejor para detectar las relaciones suaves entre características. Si no dividieramos los datos en grupos homogéneos el modelo de regresión lineal no funcionará tan bien porque los datos todavía son complejos.

Aún así, una ventaja que prevalece para los árboles de clasificación es que estos manejan bien datos categóricos (como determinar si un vino es bueno o no). En la práctica se observó que pese a que el trabajo exclusivamente categórico no es malo, se puede mejorar, y es por eso que optamos por agregar un enfoque híbrido. Si bien no pudo superar los resultados esperados para una máquina de soporte vectorial, creo que vale la pena considerarlo por su sencillez y fácil interpretabilidad.

Hay un detalle que me causa ruido en la cabeza, ¿La clasificación seguirá igual si tratamos los datos para volverlos categóricos?, es decir, en una parte de la práctica se nota que los datos medios los analiza bien (tanto el cubista como el de clasificación), pero si usamos un umbral para categorizar la calidad en bueno, malo y medio quizá el árbol capture mejor los valores extremos. Por ejemplo, en una práctica reciente se hizo evidente que los modelos de clasificación son más radicales al dividir los datos en grupos. Quizá este enfoque sea más adecuado cuando se quiera predecir específicamente una clase, por ejemplo, si un vino es bueno, y no tanto si es malo o medio.