

# Actividad 6

## Clasificación con Árboles de Decisión

Desarrollar un modelo simple de aprobación de crédito usando el algoritmo C5.0. Los datos empleados, correspondientes al fichero 'german\_credit.csv', muestran los préstamos obtenidos de una agencia de crédito alemana. Como los datos del préstamo se obtuvieron en Alemania, la moneda se registra en marcos alemanes (DM - Deutsche Marks).

El *dataset* está formado por 1000 observaciones y 17 variables. La variable de incumplimiento '*default*' indica si el solicitante del préstamo pudo (*default* = 1) o no (*default* = 2) cumplir las condiciones de pago acordadas y entró en demora. Un 30 por ciento del total de los préstamos entraron en demora.

Una alta tasa de impago es indeseable para un banco porque significa que es poco probable que el banco recupere totalmente su inversión, por lo que nuestro modelo debería identificar con éxito a los solicitantes que es probable que incumplan el pago, para que este número pueda reducirse.

Imprimir el árbol final resultante e indicar tanto la tasa de error como la tasa media de acierto y el correspondiente coeficiente kappa.