

# CALIDAD DEL VINO

**Autor:** Jorge Luis Cohen Torres

**Descripción:** Se tiene una base de datos con varias características químicas de algunos tipos de vinos rojos para predecir su calidad a través de regresiones lineales con machine learning.

**Datos:** Los datos fueron tomados del repositorio de machine learning UCI cuyos autores son :

- Paulo Cortez, Universidad de Minho, Guimarães (Portugal)
- A. Cerdeira, F. Almeida, T. Matos y J. Reis, Comisión de Viticultura de la Región del Vino Verde (CVRVV) ,Oporto (Portugal)

**Metodología:** se utilizaron las técnicas de machine learning con regresión lineal múltiple y con Random Forests.

**Resultados:** Para las técnicas con regresión lineal múltiple y Random Forest se obtuvieron los siguientes indicadores del score y RMS:

	SCORE	RMS2
LR	0,246	0,700
LR2	0,206	0,719
RF	0,931	0,211

Apreciamos que la mejor metodología usada fue la de Random Forest puesto que el score paso de 0,246 y 0,206 con las técnicas múltiples a 0,931 con RF, valor muy cercano al 1. Apreciamos también un error cuadrado medio que pasó de 0,70

y 0,719 con las técnicas múltiples a 0,211 con RF, valor muy cercano a 0. Los resultados anteriores y basado en la teoría estadística, la mejor técnica es la de Random Forest

**Trabajos futuros:** Este trabajo estadístico fue muy enriquecedor para predicciones a posteriori, por lo tanto el autor propone replicar dicho a estudio a otros tipos de bebidas (alcohólicas o no alcohólicas) para ver la incidencia en la calidad de los mismo y de gran ayuda para una mejor toma de decisiones

score\_SM is 0.2465278379986101  
score\_LR is 0.24652705028414046  
score\_LR2 is 0.2061784992879474  
score\_RF is 0.931198905259292

RMS2\_SM is 0.7007734463102371  
RMS2\_LR is 0.700773812620545  
RMS2\_LR2 is 0.7192923813258453  
RMS2\_RF is 0.21175891633920807

