

Módulo 5 – Aprendizaje de Máquina Supervisado

# Actividad: Boosting Caso Wines

Especialización en Ciencia de Datos

## Actividad: Boosting Caso Wines





# Caso Wines

El siguiente dataset, corresponde a mediciones de propiedades químicas de vinos procedentes de 3 cultivos distintos de un área específica de Italia. El dataset, contiene los resultados de 178 mediciones de 13 variables químicas medidas para cada muestra.



# Caso Wines

A continuación, el set de datos:

Alcohol	Malic Acid	Ash	Alcalinity of ash	Magnesium	Total phenols	Flavanoids	Nonflavanoid phenols	Proanthocyanins	Color intensity	Hue	D280/OD315 of diluted wines	Proline
14.23	1.71	2.43	15.6	127	2.80	3.06	0.28	2.29	5.64	1.040	3.92	1065
13.20	1.78	2.14	11.2	100	2.65	2.76	0.26	1.28	4.38	1.050	3.40	1050
13.16	2.36	2.67	18.6	101	2.80	3.24	0.30	2.81	5.68	1.030	3.17	1185
14.37	1.95	2.50	16.8	113	3.85	3.49	0.24	2.18	7.8	0.860	3.45	1480
13.24	2.59	2.87	21.0	118	2.80	2.69	0.39	1.82	4.32	1.040	2.93	735
14.20	1.76	2.45	15.2	112	3.27	3.39	0.34	1.97	6.75	1.050	2.85	1450

Alcohol

Alcohol

Malic

Malic acid

Ash

Ash

Alcalinity

Alcalinity of ash

Magnesium

Magnesium

Phenols

Total phenols

Flavanoids

Flavanoids

Nonflavanoids

Nonflavanoid phenols

Proanthocyanins

Proanthocyanins

Color

Color intensity.

Hue

Hue

Dilution

D280/OD315 of diluted wines.

Proline

Proline

Asuncion, A. & Newman, D.J. (2007). *UCI Machine Learning Repository* [<http://www.ics.uci.edu/~mllearn/MLRepository.html>]. Irvine, CA: University of California, School of Information and Computer Science.



# Caso Wines

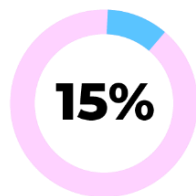
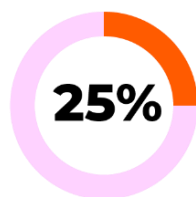
El objetivo, es **elaborar un modelo predictivo** que permita identificar el cultivo a partir de las características químicas. Para esto, utilizaremos un **algoritmo de boosting**.



# Instrucciones

**A partir de un notebook, ordenado y reproducible, realice lo siguiente:**

- Haga un análisis exploratorio de los datos.
- Realice entrenamiento con distintos modelos de boosting.
- Realice optimización de hiperparámetros.
- Evalúe el performance de los modelos entrenados.



¡Éxito!