



Proyecto final

Aprendizaje Automático

Grupo 12A

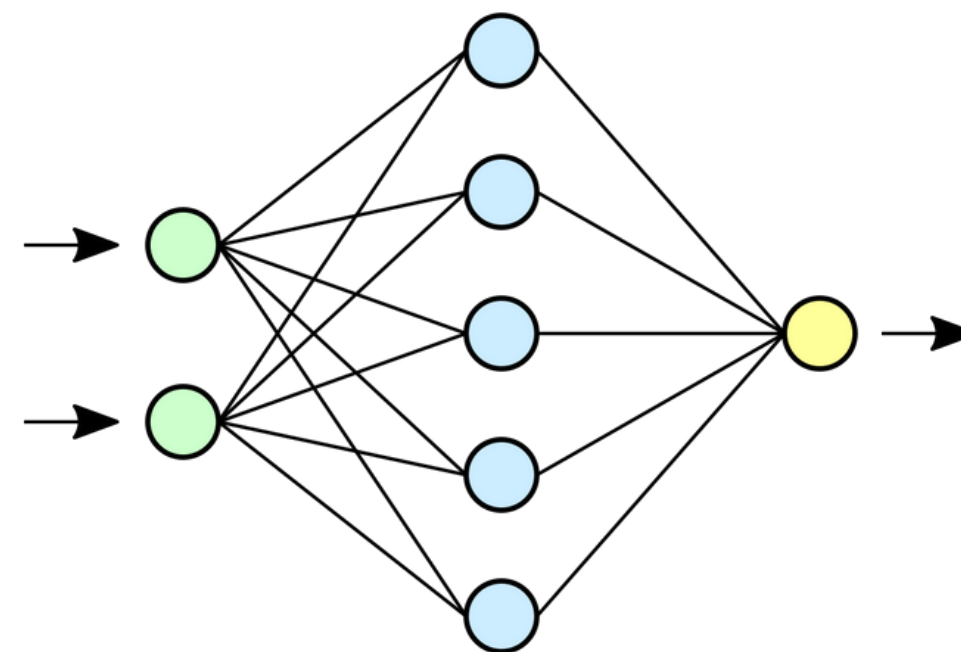
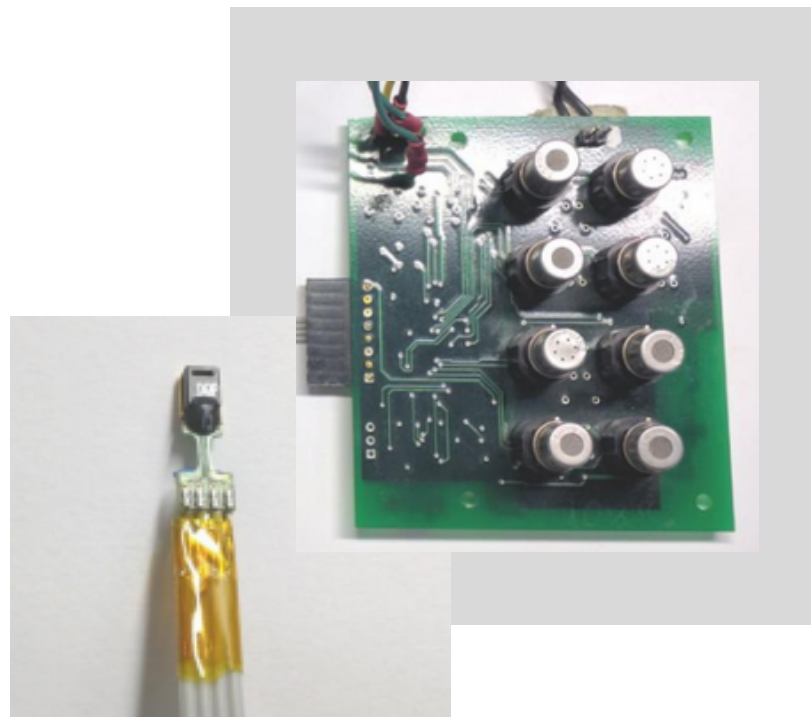
Alejandro Martín

Daniel Mohedano

Silvia Sopeña

Objetivo del problema

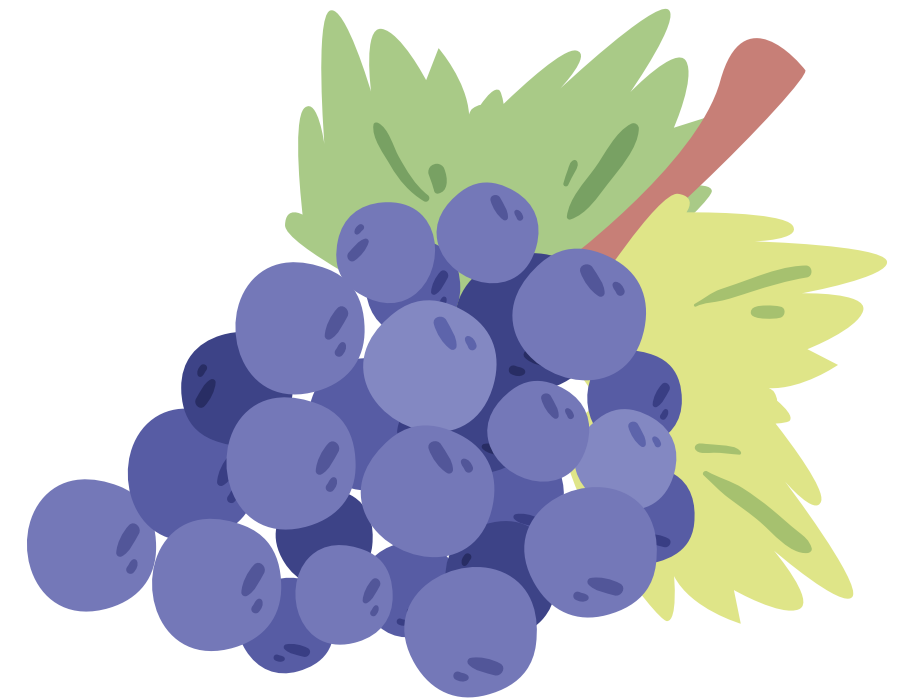
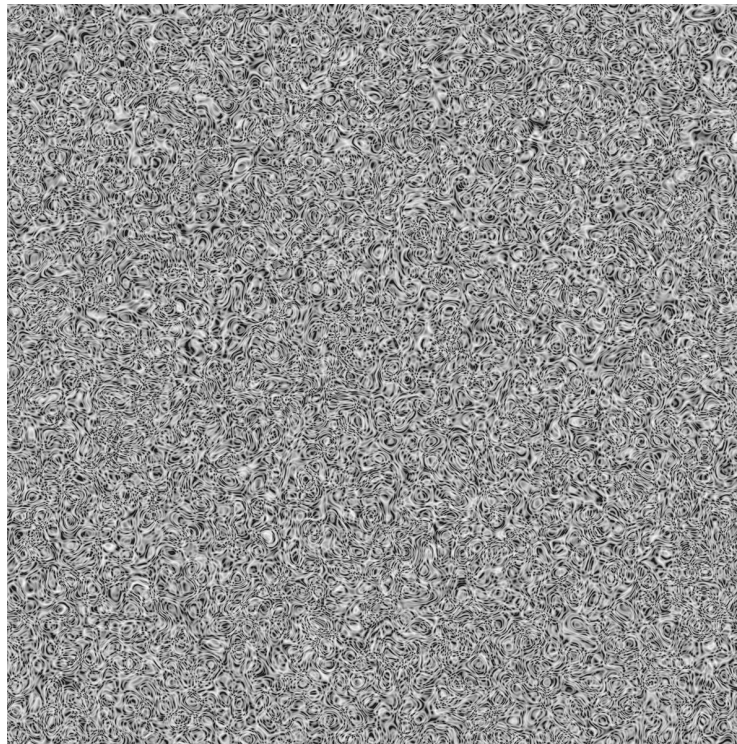
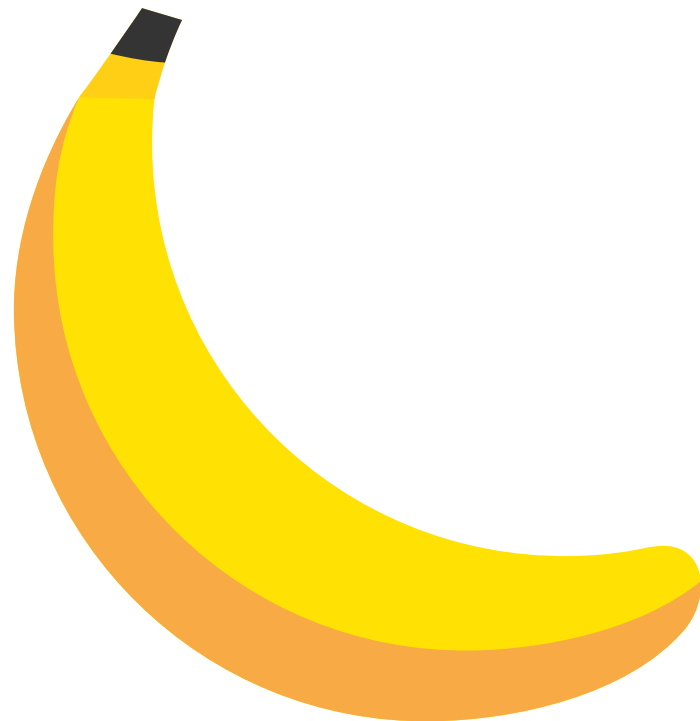
Clasificador elegido



Banana

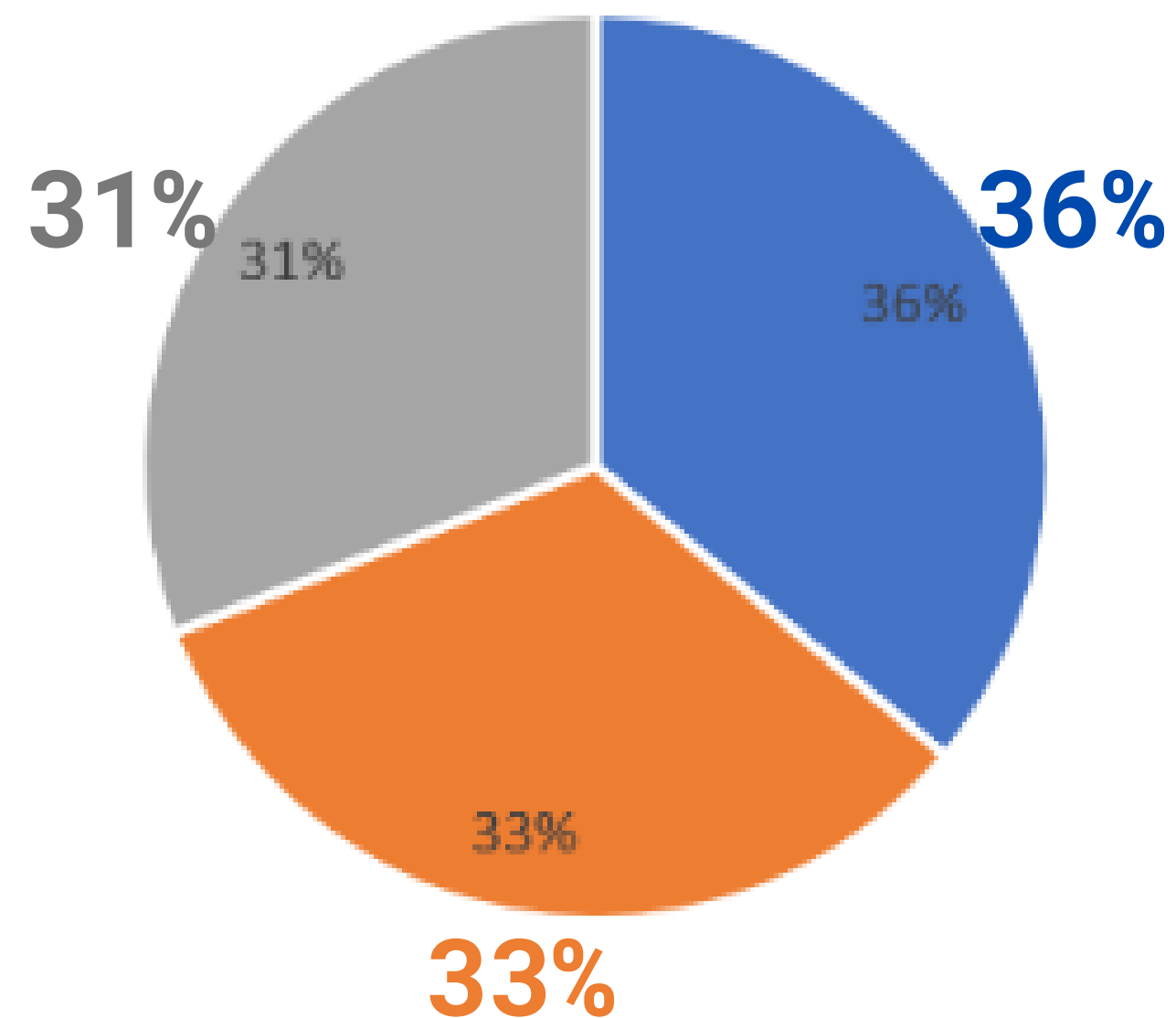
Análisis exploratorio de los datos

100 series de tiempo de experimentos divididas entre las 3 clases:



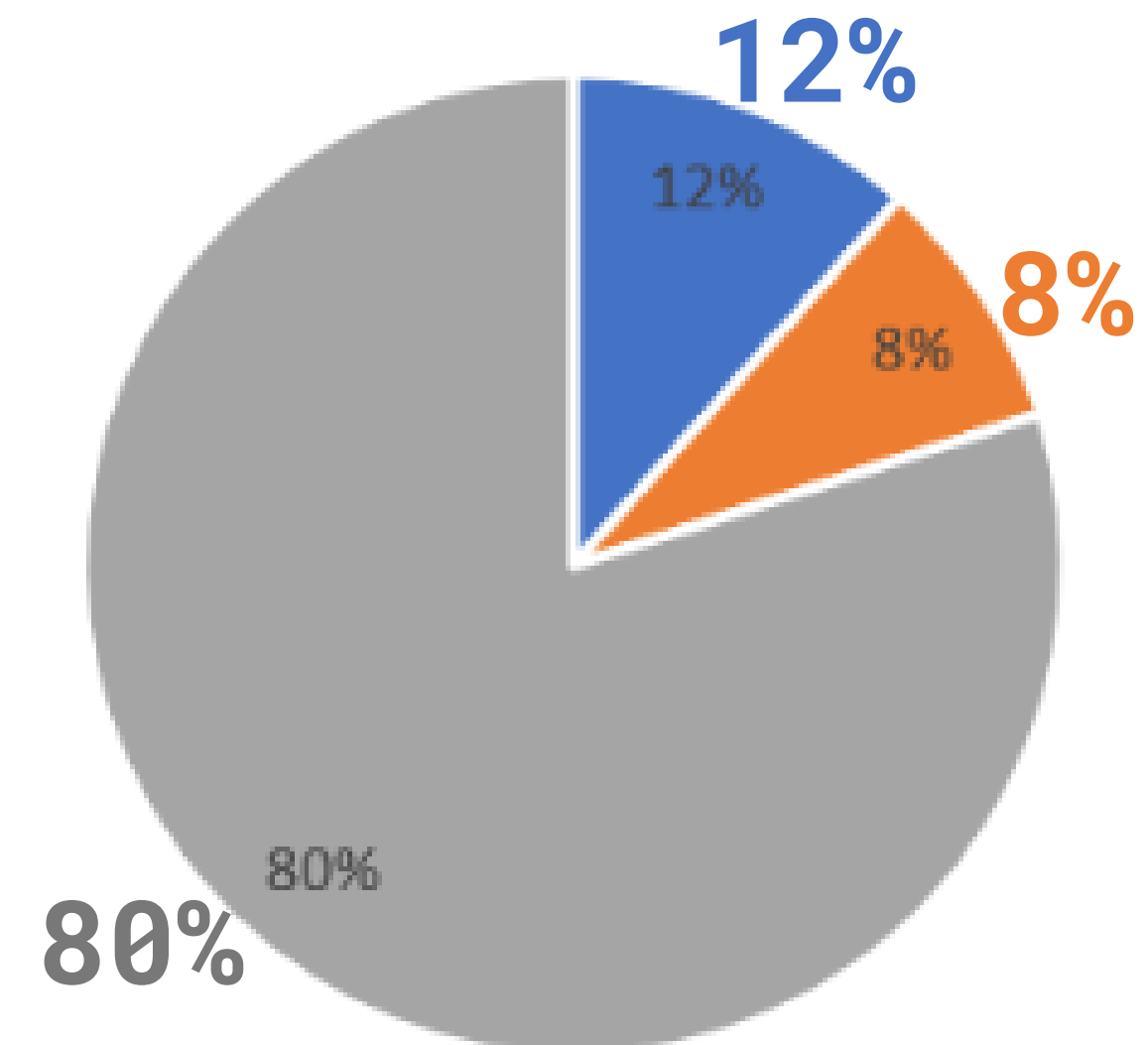
Distribución de experimentos

■ wine ■ banana ■ background



Distribución de datos

■ wine ■ banana ■ background

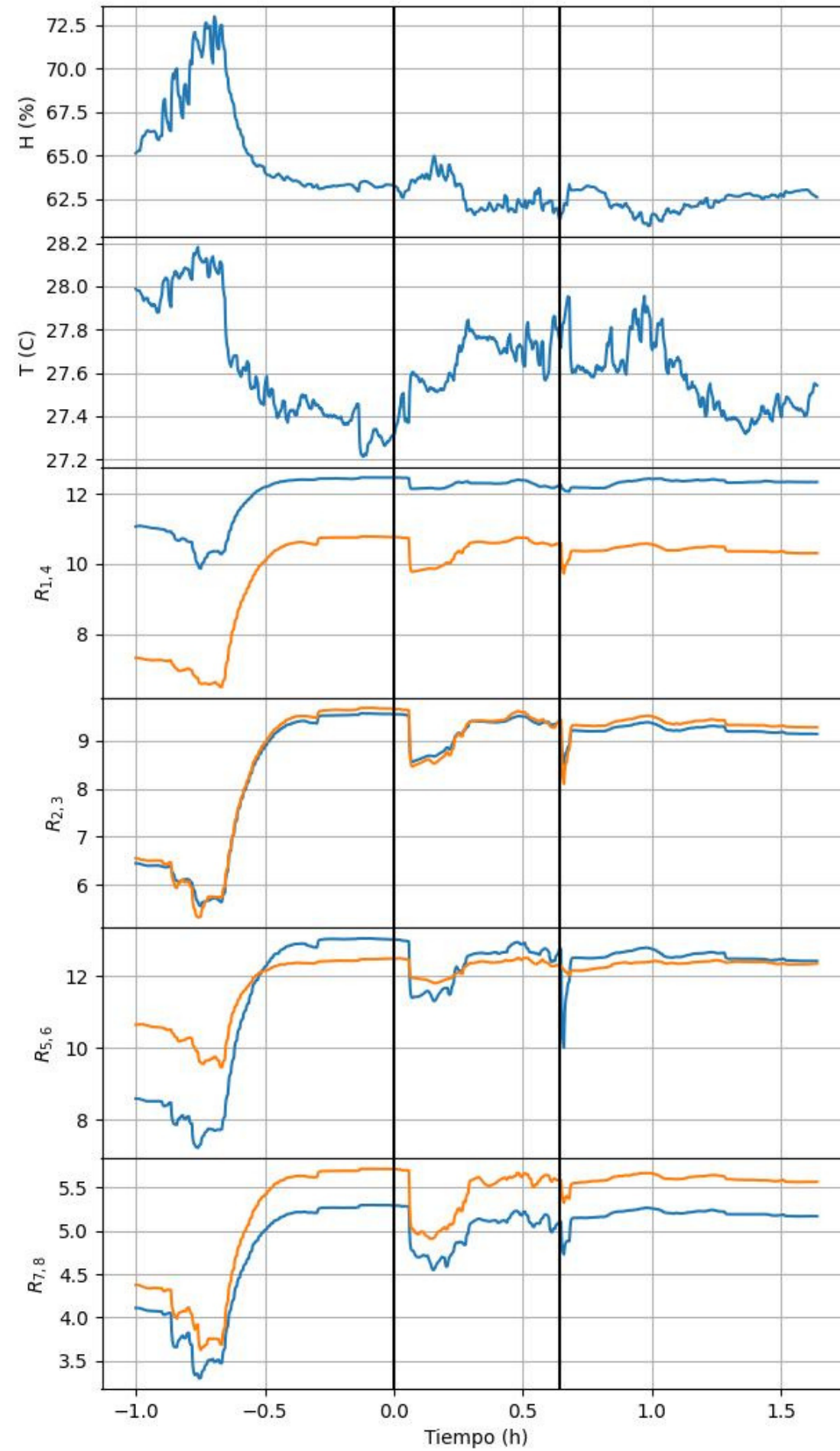


Descripción de los datos

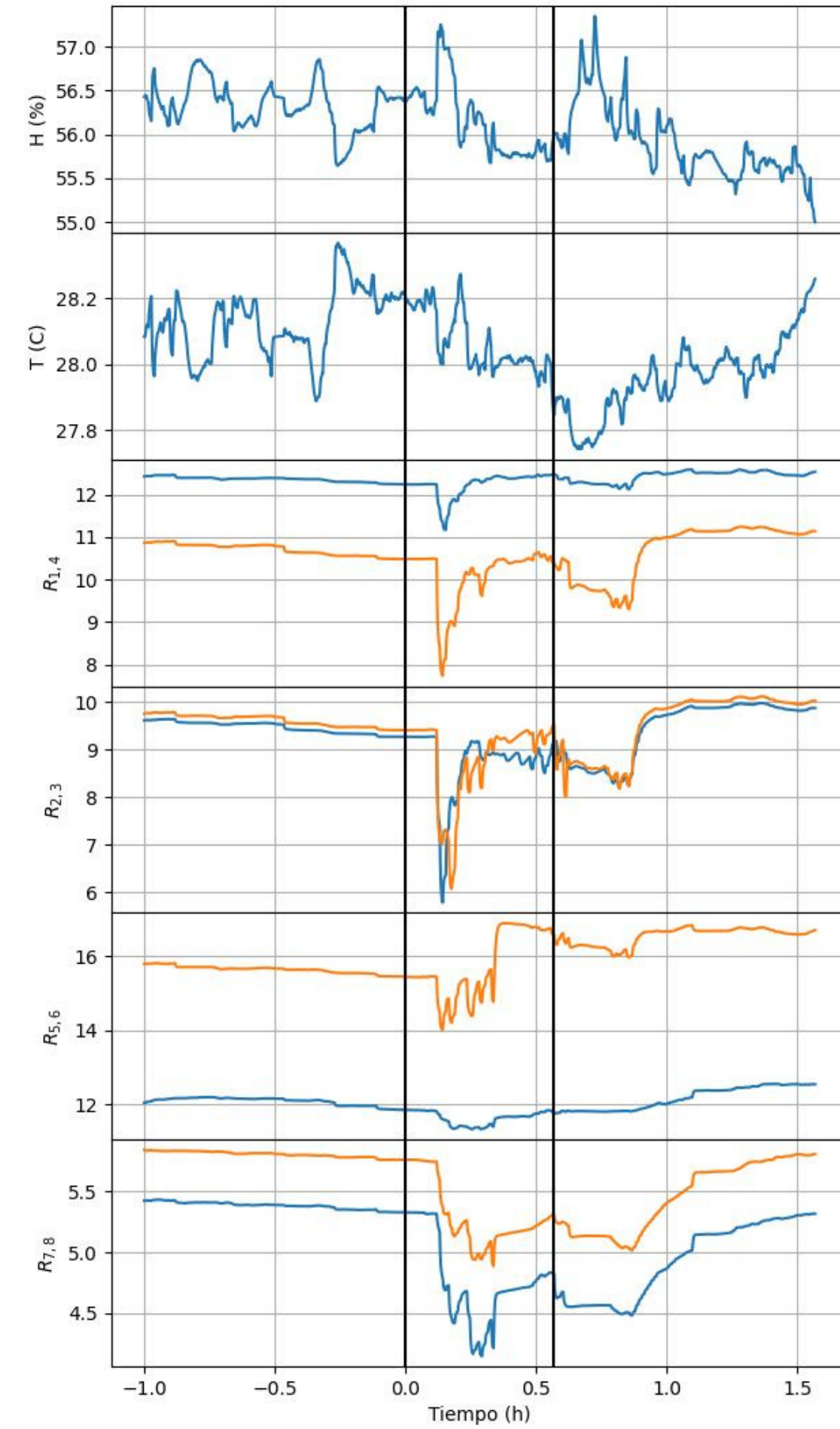
	count	mean	std
R1	928991	12.185	0.868
R2	928991	8.958	1.558
R3	928991	8.945	1.748
R4	928991	10.130	1.711
R5	928991	15.154	18.391
R6	928991	16.052	3.303
R7	928991	5.390	2.889
R8	928991	5.912	3.304
Temp.	928991	27.283	0.904
Humidity	928991	57.568	4.821

Table 1: Descripción de los datos.

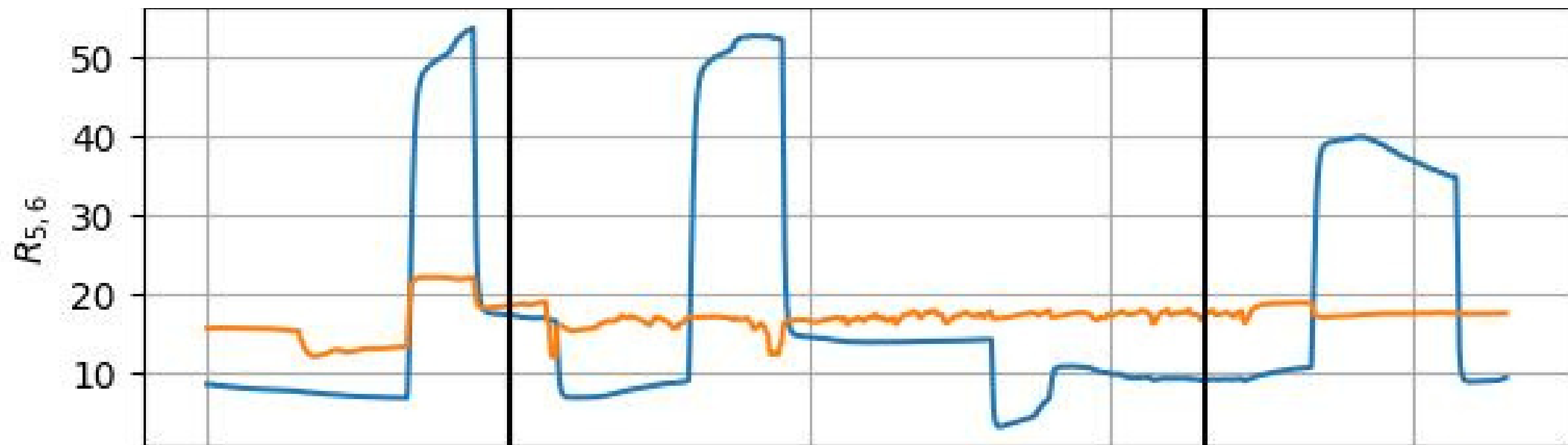
Experimento 17: banana



Experimento 42: wine

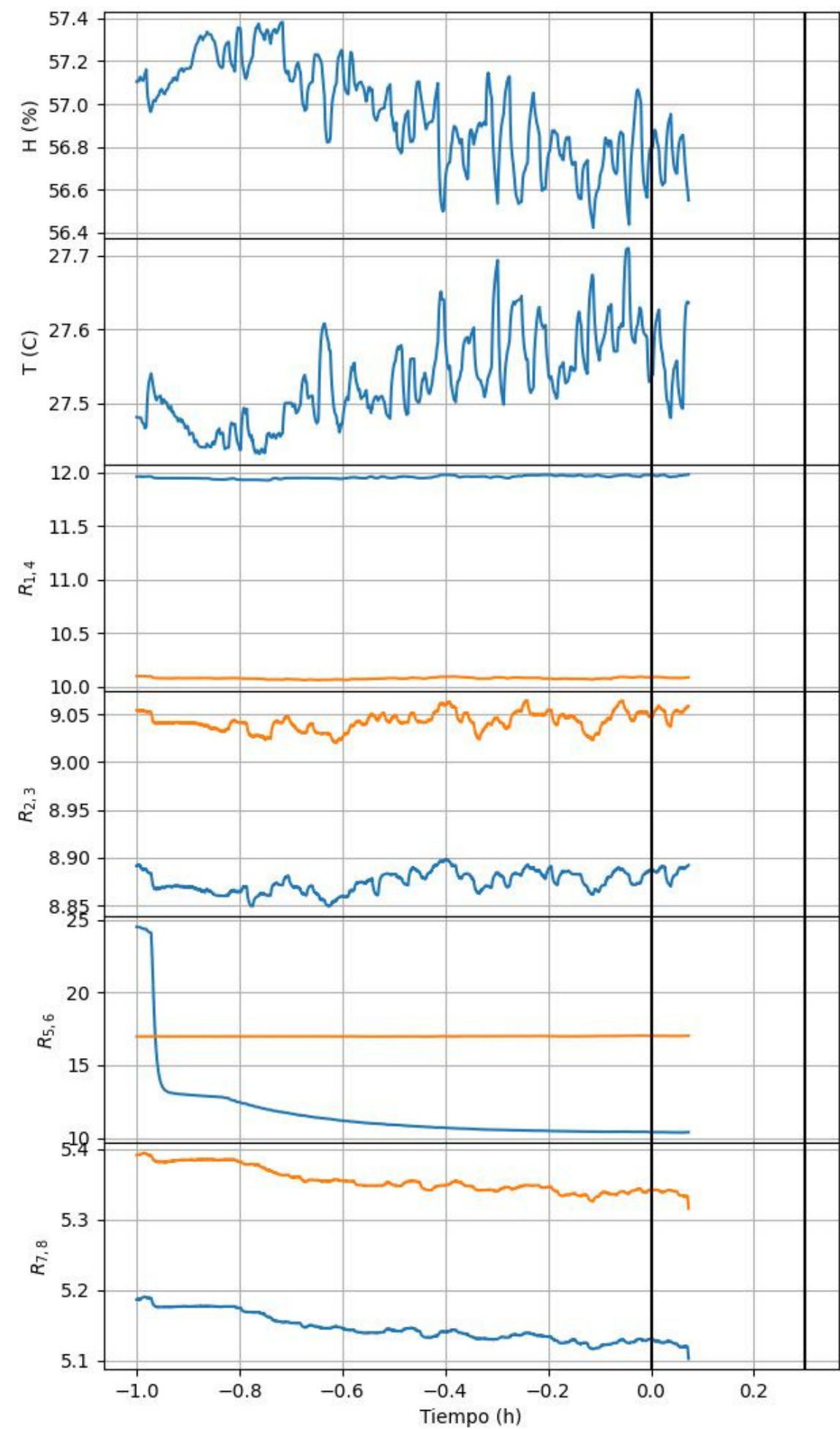


Imprecisión del R5

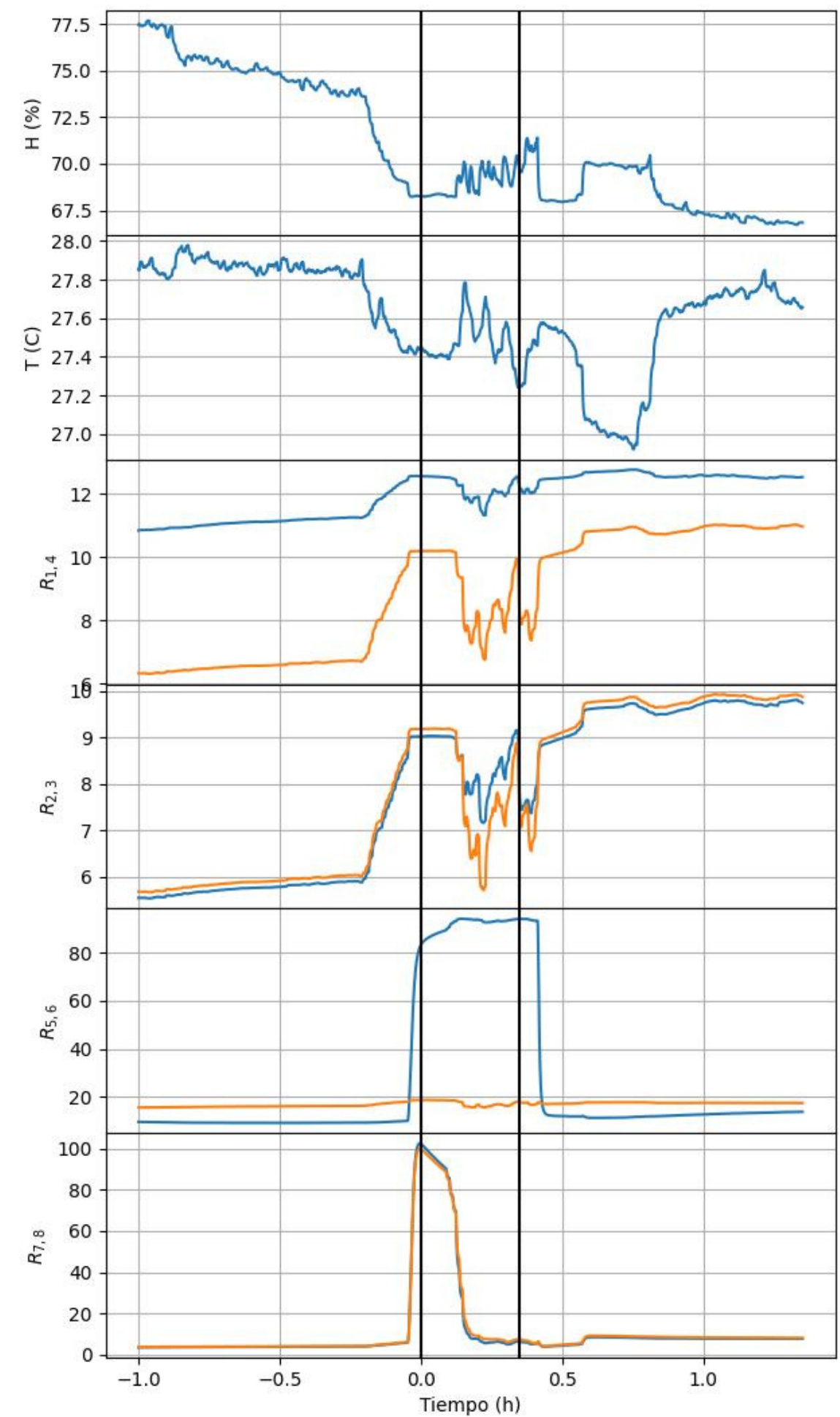



Experimento 38

Experimento 76: background



Experimento 47: banana





Descripción de los atributos propuestos y su obtención

1. Seleccionar sólo los datos entre los límites del estímulo.
2. Eliminar los datos del experimento 47.
3. Eliminar el atributo R5.
4. Debaticimos añadir atributos extra pero decidimos no hacerlo.



Clasificadores

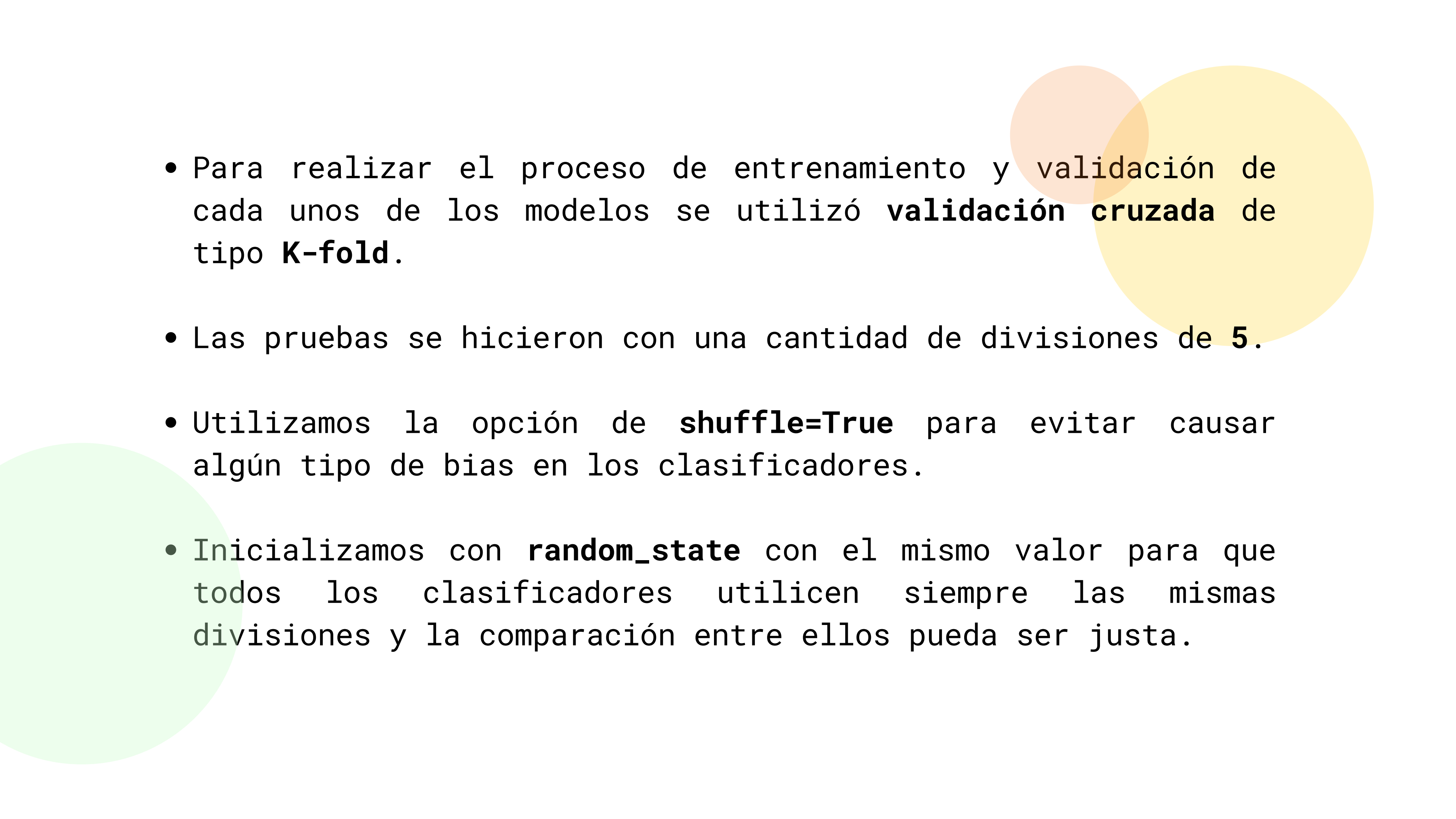
Naive-Bayes

KNN

Redes Neuronales

Regresión logística

Random Forest

- 
- Para realizar el proceso de entrenamiento y validación de cada uno de los modelos se utilizó **validación cruzada** de tipo **K-fold**.
 - Las pruebas se hicieron con una cantidad de divisiones de **5**.
 - Utilizamos la opción de **shuffle=True** para evitar causar algún tipo de bias en los clasificadores.
 - Inicializamos con **random_state** con el mismo valor para que todos los clasificadores utilicen siempre las mismas divisiones y la comparación entre ellos pueda ser justa.

Resultados obtenidos

Naive Bayes		
Tipo	Precisión	Desv.
Multinomial	0.649884	0.002006
Gaussian	0.778738	0.001009

Regresión Logística		
	Precisión	Desv.
	0.764829	0.003179
Regresión Logística SGD		
c	Precisión	Desv.
0.1	0.660019	0.075157
0.3	0.642637	0.094687
0.5	0.583213	0.065870
0.7	0.599605	0.116481
0.9	0.579930	0.118764

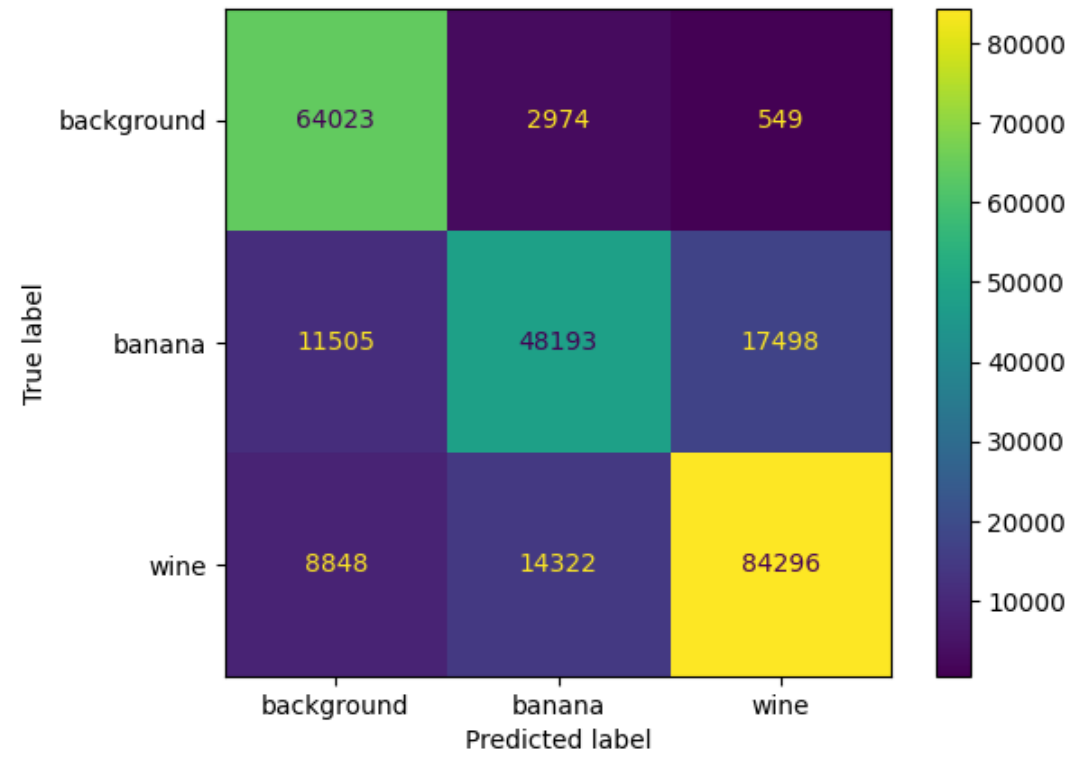
KNN		
K	Precisión	Desv.
5	0.999885	0.000063
15	0.999663	0.000074
25	0.999437	0.000113
35	0.999290	0.000092
45	0.999124	0.000122
55	0.998854	0.000069
65	0.998553	0.000159
75	0.998378	0.000125
85	0.998105	0.000216
95	0.997867	0.000164

Multilayer Perceptron		
Neuronas	Precisión	Desv.
25	0.924792	0.007318
50	0.967091	0.015553
75	0.984382	0.005693
100	0.989306	0.002850
(25, 25)	0.982792	0.007527
(25, 50)	0.993057	0.002102
(50, 25)	0.994655	0.003016
(50, 50)	0.998045	0.000934
(75, 25)	0.971718	0.026802
(75, 50)	0.996737	0.001266
(100, 25)	0.997066	0.001095
(100, 50)	0.994861	0.005239

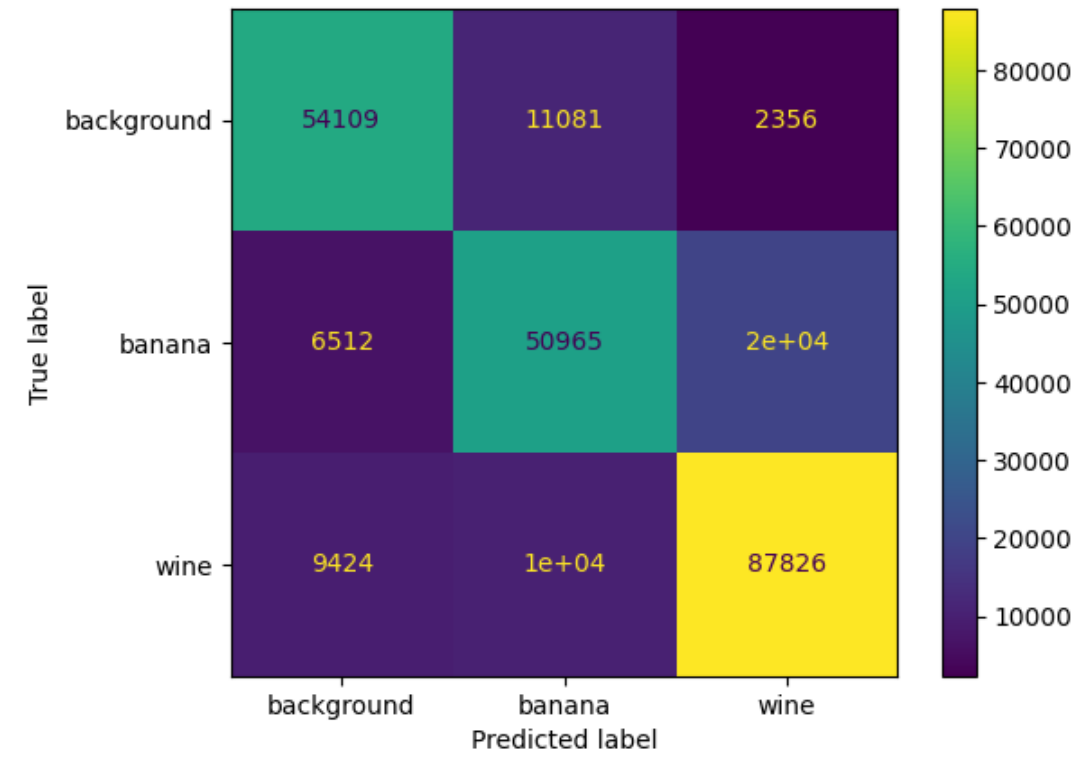
Random Forest		
N	Precisión	Desv.
10	0.999877	0.000038
35	0.999909	0.000063
60	0.999925	0.000044
85	0.999909	0.000046
110	0.999913	0.000048
135	0.999921	0.000050
160	0.999925	0.000044
185	0.999917	0.000049

Discusión de resultados obtenidos

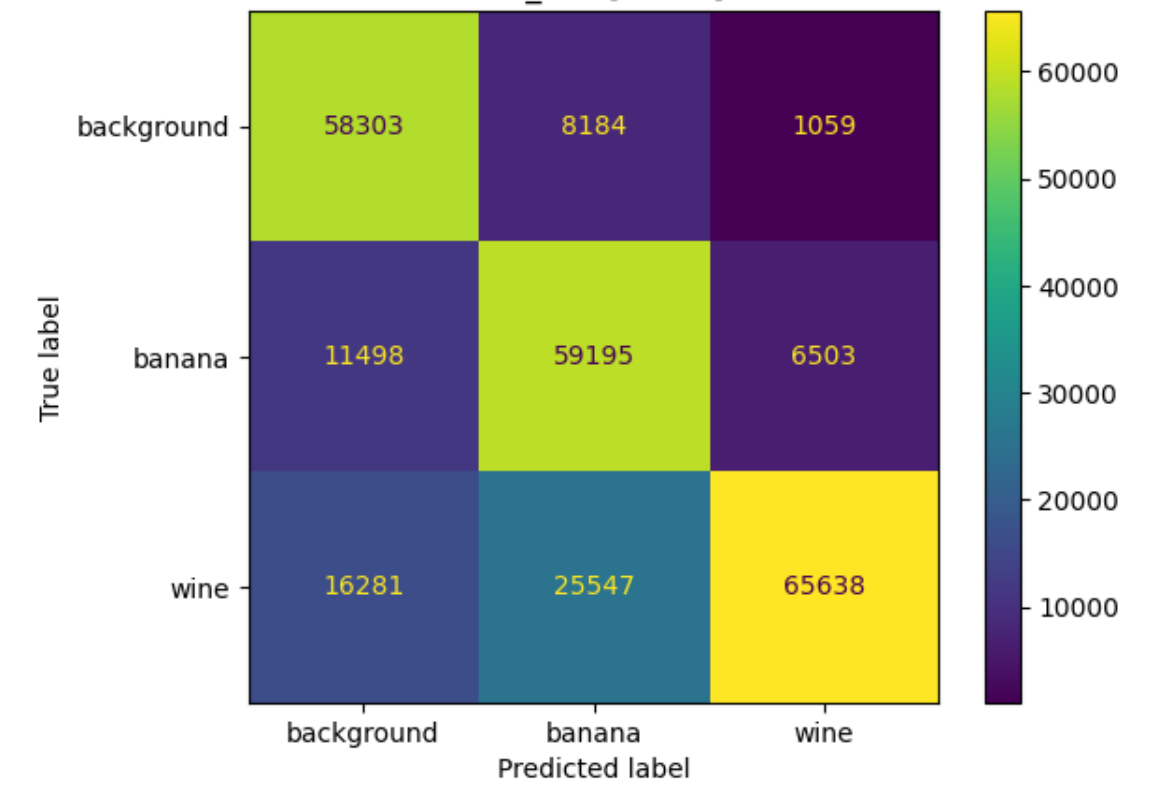
NBGaussian



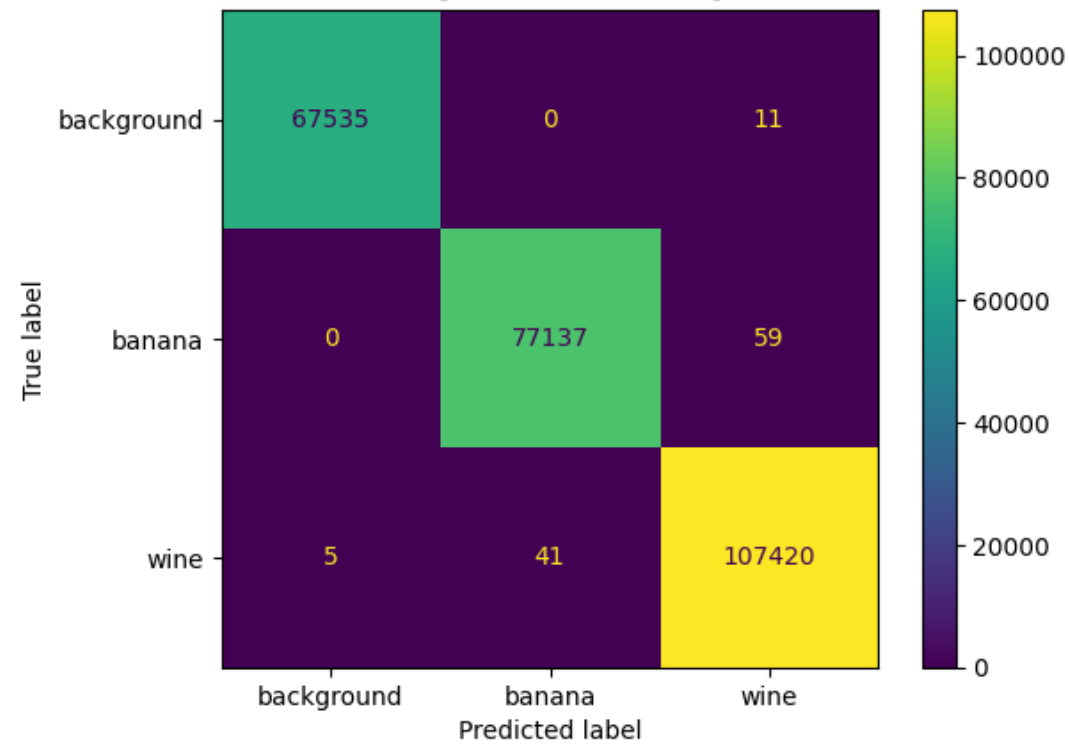
LR



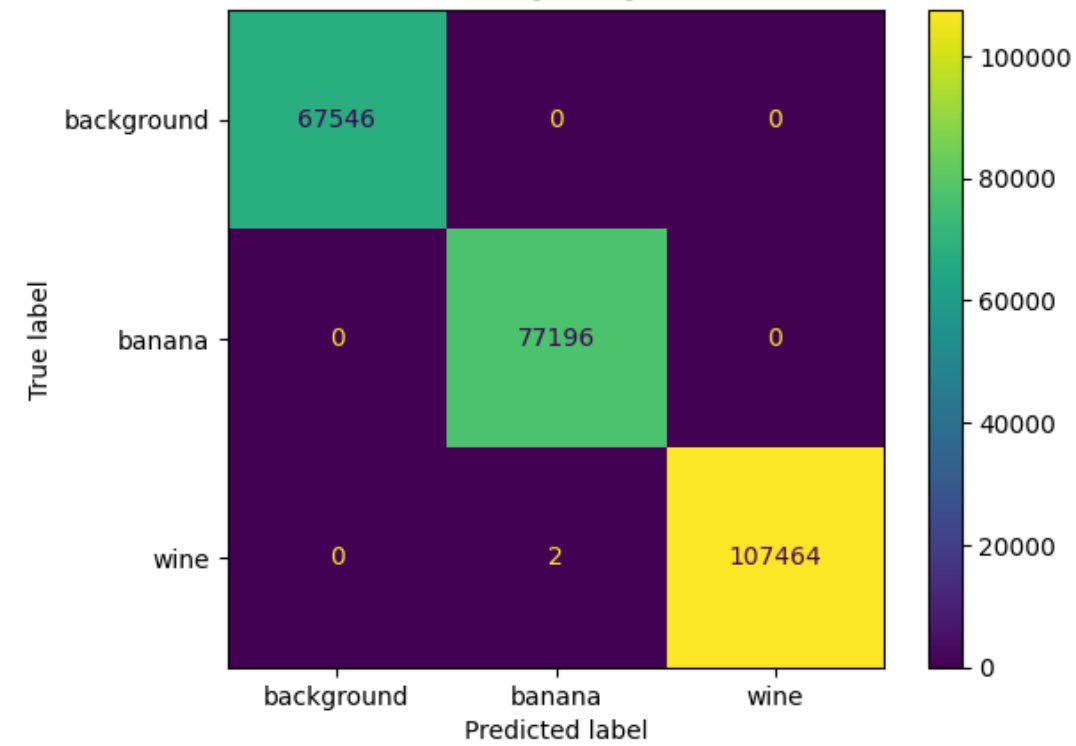
SGD_OVR[c=0.1]



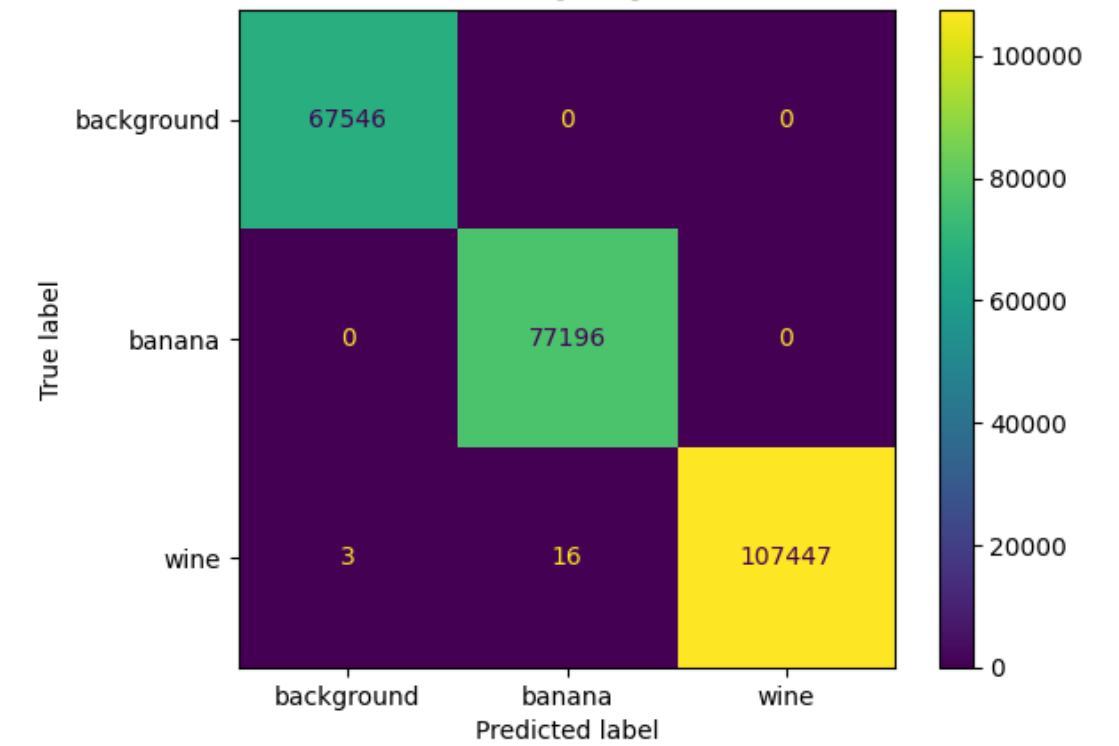
MLP[hidden=(50, 100)]



RFo[n=60]

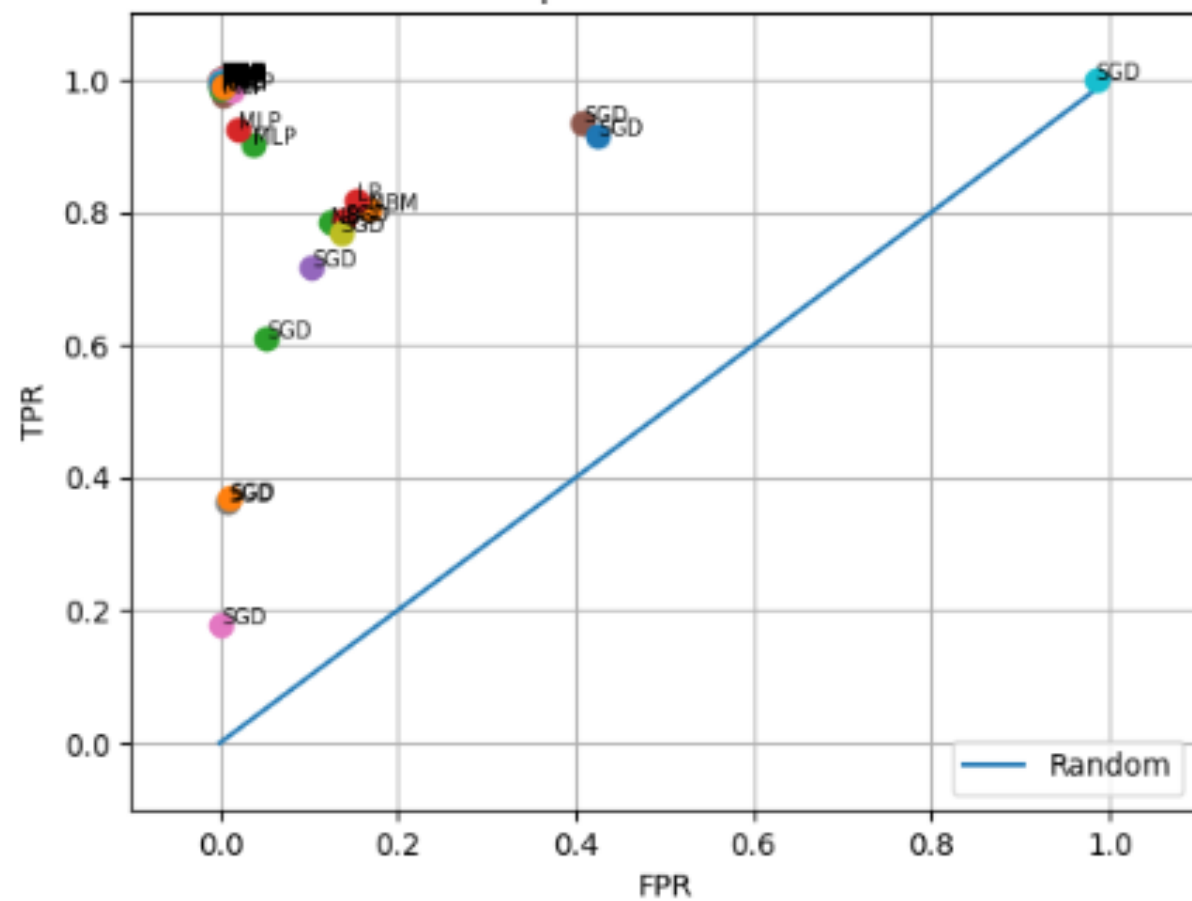


KNN[K=5]

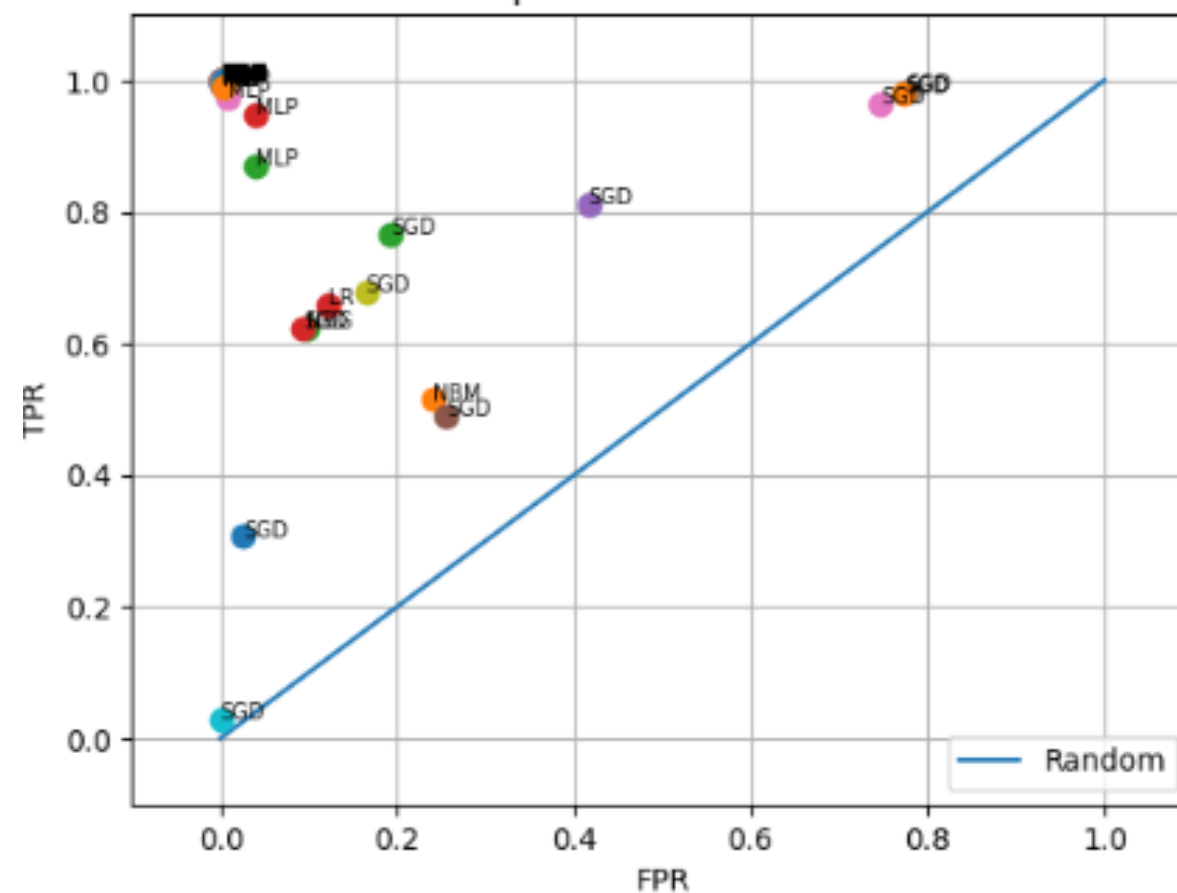


Discusión de resultados obtenidos

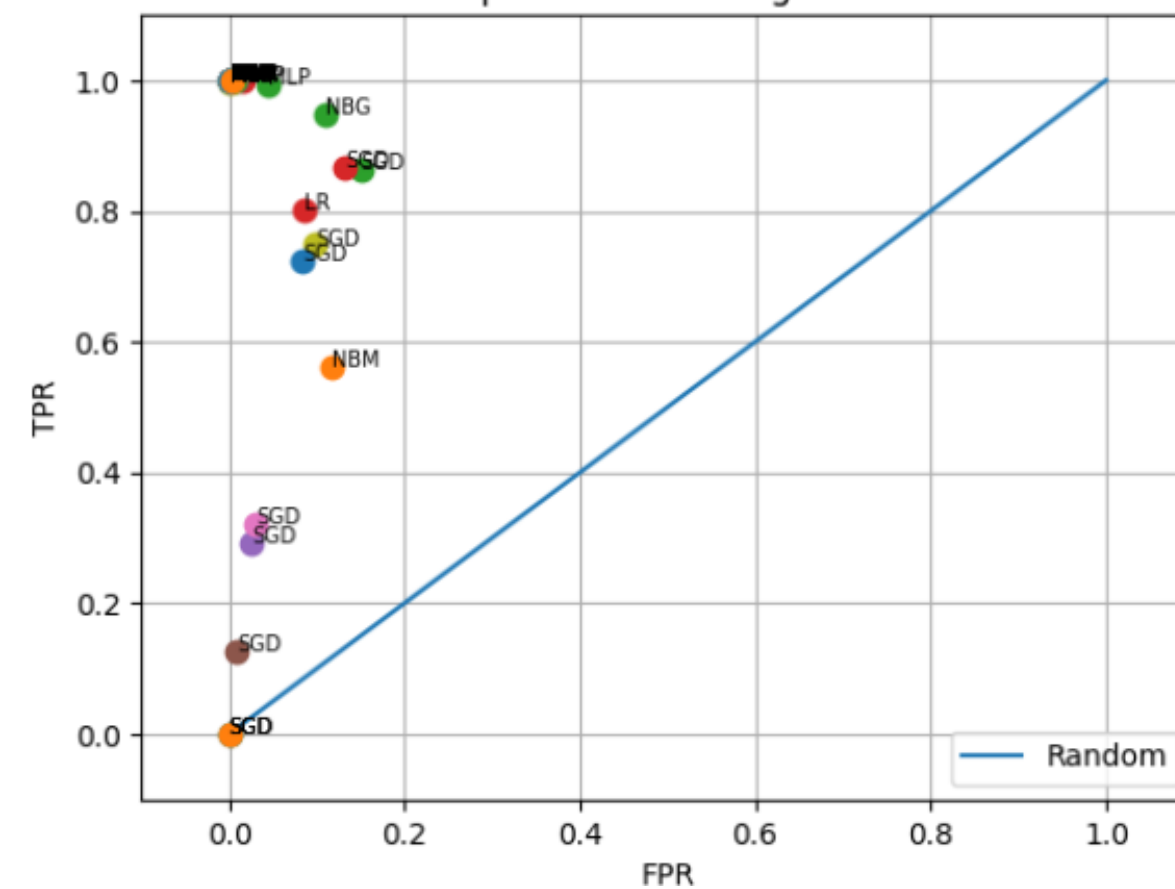
Espacio ROC: wine




Espacio ROC: banana



Espacio ROC: background





Gracias por tu atención