

UNIVERSIDAD MAYOR REAL Y PONTIFICIA DE  
SAN FRANCISCO XAVIER DE CHUQUISACA



**LABORATORIO 2 – REGRESIÓN LOGÍSTICA**

**CARRERA:** Ing. Ciencias de la Computación

**MATERIA:** Inteligencia Artificial I

**NOMBRE:** Sandra Villca Señoranis

## Informe Laboratorio 2: Regresión Logística

1. Se ha trabajado con el dataset "WineQuality" para hacer la predicción de la calidad del vino. El preprocesamiento del dataset fue realizado con Pandas.

El dataset tiene 32485 ejemplos y 12 características.

	fixed acidity	volatile acidity	citric acid	residual sugar	\
0	7.4	0.170	0.29	1.4	
1	5.3	0.310	0.38	10.5	
2	4.7	0.145	0.29	1.0	
3	6.9	0.260	0.29	4.2	
4	6.4	0.450	0.07	1.1	
...	...	...	...	...	
32480	5.0	0.255	0.22	2.7	
32481	6.6	0.360	0.52	11.3	
32482	6.3	0.200	0.24	1.7	
32483	6.2	0.200	0.33	5.4	
32484	8.1	0.280	0.46	15.4	

  

	chlorides	free sulfur dioxide	total sulfur dioxide	density	pH	\
0	0.047	23.0	107.0	0.99390	3.52	
1	0.031	53.0	140.0	0.99321	3.34	
2	0.042	35.0	90.0	0.99080	3.76	
3	0.043	33.0	114.0	0.99020	3.16	
4	0.030	10.0	131.0	0.99050	2.97	
...	...	...	...	...	...	
32480	0.043	46.0	153.0	0.99238	3.75	
32481	0.046	8.0	110.0	0.99660	3.07	
32482	0.052	36.0	135.0	0.99374	3.80	
32483	0.028	21.0	75.0	0.99012	3.36	
32484	0.059	32.0	177.0	1.00040	3.27	

  

	sulphates	alcohol	Type	quality
0	0.65	10.4	White Wine	6
1	0.46	11.7	White Wine	6
2	0.49	11.3	White Wine	6
3	0.31	12.5	White Wine	6
4	0.28	10.8	White Wine	5
...	...	...	...	...
32480	0.76	11.3	White Wine	6
32481	0.46	9.4	White Wine	5
32482	0.66	10.8	White Wine	6

## **2. Valores de calidad de vino (y)**

Originalmente oscilan entre 0 y 8. Se han adecuado esos valores según la condición:

0 – 4 → mala calidad

5 – 8 → buena calidad

## **3. Regresión Logística**

- Los datos se ha dividido en dos partes: Entrenamiento (80% de los ejemplos) y Prueba (20% de los ejemplos).
- Se han normalizado los valores de las características y se ha agregado el término de intersección  $X_0$  o bias.
- En el algoritmo se ha definido las funciones: sigmoidea, de costo y gradiente.
- Se han realizado 40000 iteraciones para calcular el Costo J cuyo valor ha descendido hasta cerca de 0.1.
- Los ejemplos de prueba también han sido procesados. Se ha hecho la predicción de la calidad para un vino y para todo el conjunto de prueba.

