

UNIVERSIDAD AUTÓNOMA DE
YUCATÁN

MAESTRÍA EN CIENCIAS DE LA
COMPUTACIÓN

Aprendizaje Automático

Reporte Ejercicio 2

WILLIAM FELIPE CETINA PECH

11 de Febrero del 2026

En este reporte se realizarán experimentos para buscar los parámetros correctos para nuestra regresión logística, cuya meta es una clasificación binaria, inicialmente tendremos 2 parámetros los cuales son:

A = Alfa (Tasa de aprendizaje)

E = Épocas a realizar

La base de datos a utilizar contiene una matriz de 99x3

Las primeras 50 filas son ejemplos positivos y las siguientes 49 son negativas

Cada fila tiene 3 parámetros (x_1 , x_2 y su etiqueta binaria (0, 1))

La ecuación a utilizar para el modelo será:

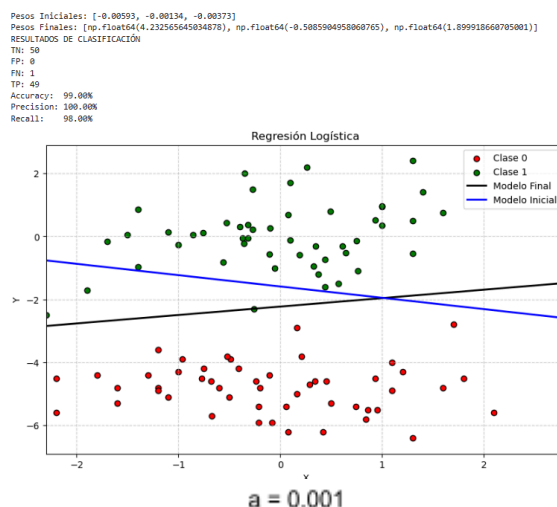
$$z = \theta_0 + \theta_1 x_1 + \theta_2 x_2$$

Si esta función se aplicará solamente estaríamos hablando de una regresión lineal; por lo que es necesaria una función de activación que centre los valores en 0 y 1 para la clasificación binaria, por lo que se usará la función sigmoide

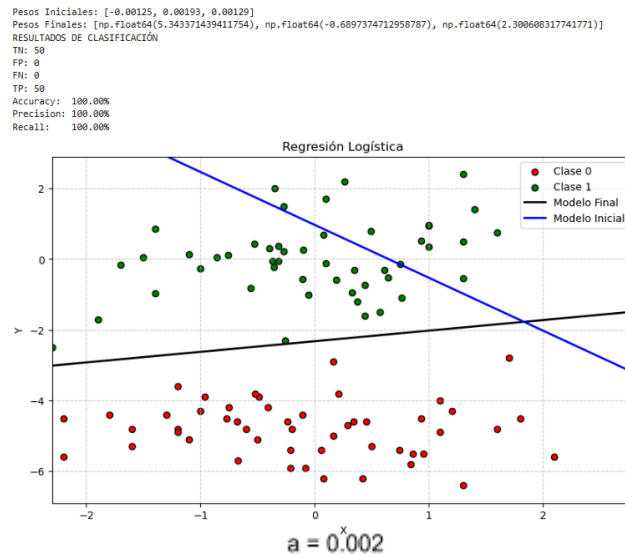
$$g(x) = \frac{1}{1+e^{-z}}$$

Donde z es la ecuación del modelo previamente definida, en estos experimentos la época a utilizar serán 1000. Para validar el modelo, al terminar las épocas se obtendrá su calificación en base a las métricas de Accuracy, Precision y Recall, estas requieren de las etiquetas de true-positive, true-negative, false-positive y false-negative que están también serán visualizadas. Finalmente se presentará una gráfica donde se podrá visualizar, los datos obtenidos, así como el modelo inicial (Parámetros aleatorios iniciales) y el modelo final (Parámetros finales)

Primeramente iniciamos con una tasa de aprendizaje de 0.001



Este modelo casi logra clasificar bien el 100 de los datos, es posible ver como todos los datos rojos están bien clasificados pero hay uno verde (positivo) que está debajo del modelo actual, por muy poco pero lo está por lo que lo categoriza como negativo que es erróneo (De ahí el FN encontrado), el modelo casi logra una categorización perfecta, posiblemente un reajuste mínimo en la tasa de aprendizaje debería ser suficiente, por lo que se aumentará la tasa de aprendizaje a 0.002



Justo cómo se hipótizo, el uso de una taza de aprendizaje de 0.002 parece ser el indicado, por lo que la ecuación final sería:

$$h(\theta) = 5.343371439411754 - 0.6897374712958787x_1 - 2.300608317741771x_2$$

Esto teniendo los siguientes parámetros:

A = 0.002

E = 1000