CURSO ESPECIALIZACIÓN INTELIGENCIA ARTIFICIAL Y BIG DATA

Tema 4: Regresión Logística

Sebastián Rubio

Mayo 2025

Sistema aprendizaje automático

1 Algoritmo de regresión logística

https://codificandobits.com/blog/regresion-logistica-y-neurona-artificial/ La regresión logística es un algoritmo de clasificación binaria que estima la probabilidad de que una observación pertenezca a una de dos clases, como vota o no votar en función de un conjunto de datos determinado de variables Hay tres tipos de modelos de regresión logística, que se definen en función de la respuesta categórica.

- Regresión logística binaria
- Regresión logística multinomial
- Regresión logística ordinal

La regresión logística pertenece a la familia de modelos de aprendizaje automático supervisado. También se considera un modelo discriminativo, lo que significa que intenta distinguir entre clases (o categorías). En la regresión logística, la variable dependiente es dicotómica (0 ó 1) y se estima la probabilidad de que se produzca la enfermedad si la persona considerada tiene una edad, un sexo y un hábito relativo al tabaco. Para explicarlo vamos a partir de un problema. Supongamos que se tiene un set de 100 datos. Cada dato se representa con dos características(dos cantidades numéricas) y además pertenece a una de dos posibles categorías: 0 o 1 El objetivo de la regresión logística es clasificar de forma automática estos datos en las dos categorías existentes. Esto equivale a encontrar una **frontera** que permita separar los datos en dos agrupaciones diferentes. La frontera de decisión se puede obtener automáticamente con el algoritmo de regresión logística Para facilitar el proceso de clasificación es necesario realizar una transformación de los datos de entrada. Esto se logra usando la siguiente ecuación:

$$z_i = \vec{w}.\vec{x_i} + b \tag{1}$$

donde $\vec{x_i}$ es un vector que contiene dos valores para cada uno de los datos de entrada:

$$\vec{x_i} = (x_{1i}, x_{2i}) \tag{2}$$

el indice i se corresponde a cada uno de los 100 datos que se están usando en este caso Por su parte, los parámetros $\vec{w}$ y b corresponden a los coeficientes o parámetros del modelo que se obtiene durante el proceso de entrenamiento. En este algoritmo se siguen los mismos pasos que en el de regresión lineal. En este caso utilizaremos la **entropía cruzada**, en vez del error cuadrático medio. La entropía cruzada se define matemáticamente como:

$$error = -\frac{1}{N}\sum_{i=1}^{N}[y_i.ln\hat{y}_i + (1-y_i).ln(1-\hat{y}_i] \tag{3}$$

donde N es el número total de elementos usados en la regresión y ln es la función del logaritmo natural. A esta fórmula se le conoce función de costo, y cuanto menor sea mejor será nuestro modelo. Por lo que en sucesivas iteraciones debemos hacerlo lo más pequeño posible.

$$J(\theta) = -\frac{1}{N}\sum_{i=1}^{N}[y_i.ln\hat{y}_i + (1-y_i).ln(1-\hat{y}_i] \tag{4}$$

$$J(\theta) = -\frac{1}{m}\sum_{i=1}^{m}\left[y^{(i)}\log(h_\theta(x^{(i)})) + (1-y^{(i)})\log(1-h_\theta(x^{(i)}))\right] \tag{5}$$

La hipótesis de regresión logística es

$$h_\theta(x) = \frac{1}{1+e^{-z}} \tag{6}$$

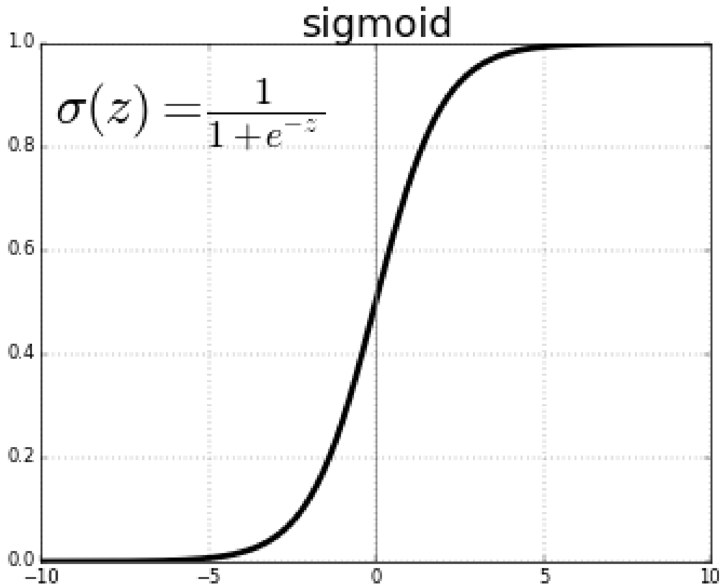

Figure 1: Enter Caption

- x son las características
- θ son los parámetros

- $h_\theta(x)$ es la probabilidad de que y=1
- $z = \theta^T x$

Donde z es una combinación lineal de las características de entrada y los pesos

$$z = \theta_0 + \theta_1 x_1 + \theta_2 x_2 + ... + \theta_n x_n \tag{7}$$

EJEMPLO:Regresión Logística. El objetivo es predecir si un estudiante aprueba (1) o no aprueba (0) un examen, en función de una sola característica: el número de horas de estudio de la asignatura. Usaremos el dataset de la siguiente tabla.

Horas de estudios (x)	**Aprueba (y)**
1	0
2	0
3	0
4	1
5	1
6	1

Table 1: Notas matemáticas y lengua

PASO 1: Definición de variables Añadimos una columna de unos para el sesgo(bias)

$$X = \begin{bmatrix}1, x\end{bmatrix} \tag{8}$$

$$X = \begin{bmatrix} 1 & 1 \\ 1 & 2 \\ 1 & 3 \\ 1 & 4 \\ 1 & 5 \\ 1 & 6 \end{bmatrix}, \quad y = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \\ 1 \\ 1 \end{bmatrix} \tag{9}$$

PASO 2: Inicialización

- Parámetros θ=[θ_0,θ_1]
- Inicializamos θ=[0,0]

- Tasa de aprendizaje α=0.1

PASO 3: Definimos el modelo Nuestra función para el modelo la siguiente sigmoide

$$h_\theta(x) = \frac{1}{1+e^{-z}} \tag{10}$$

- x son las características
- θ son los parámetros (pesos)
- $h_0(x)$ es la probabilidad de que y=1

PASO 4: Primer paso del descenso de gradiente

1. Calcular predicciones:
   Con θ=[0,0]:

$$z = X \cdot \theta = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix} \implies h = \sigma(z) = \begin{bmatrix} 0.5 \\ 0.5 \\ 0.5 \\ 0.5 \\ 0.5 \end{bmatrix} \tag{11}$$

2. Calcular el gradiente:

$$\frac{1}{m} X^T (h - y) \tag{12}$$

$$h - y = \begin{bmatrix} 0.5 \\ 0.5 \\ 0.5 \\ 0.5 \\ 0.5 \\ 0.5 \end{bmatrix} - \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 0.5 \\ 0.5 \\ 0.5 \\ -0.5 \\ -0.5 \\ -0.5 \end{bmatrix} \tag{13}$$

$$X^T (h - y) = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 2 & 3 & 4 & 5 & 6 \end{bmatrix} \cdot \begin{bmatrix} 0.5 \\ 0.5 \\ 0.5 \\ -0.5 \\ -0.5 \\ -0.5 \end{bmatrix} = \begin{bmatrix} 0 \\ -6 \end{bmatrix} \tag{14}$$

$$\text{Gradiente} = \frac{1}{6} \begin{bmatrix} 0 \\ -6 \end{bmatrix} = \begin{bmatrix} 0 \\ -1 \end{bmatrix} \tag{15}$$

3. Actualizar parámetros:

$$\theta = \theta - \alpha \cdot \text{Gradiente} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} - 0.1 \cdot \begin{bmatrix} 0 \\ -1 \end{bmatrix} = \begin{bmatrix} 0 \\ 0.1 \end{bmatrix} \tag{16}$$

PASO 5: Segundo paso del descenso de gradiente

Con θ=[0,0.1]

$$z = X.\theta = \begin{bmatrix} 1.0 + 1.01 = 0.1 \\ 1.0 + 2.0.1 = 0.2 \\ 1.0 + 3.0.1 = 0.3 \\ 1.0 + 4.0.1 = 0.4 \\ 1.0 + 5.0.1 = 0.5 \\ 1.0 + 6.0.1 = 0.6 \end{bmatrix} \rightarrow h = \sigma(z) \approx \begin{bmatrix} 0.525 \\ 0.550 \\ 0.574 \\ -0.402 \\ -0.378 \\ -0.355 \end{bmatrix} \tag{17}$$

$$h - y \approx \begin{bmatrix} 0.525 \\ 0.550 \\ 0.574 \\ -0.402 \\ -0.378 \\ -0.355 \end{bmatrix} \tag{18}$$

$$X^T(h - y) \approx \begin{bmatrix} 0.514 \\ -0.822 \end{bmatrix} \Rightarrow 1/6. \begin{bmatrix} 0.514 \\ -0.822 \end{bmatrix} \approx \begin{bmatrix} 0.0086 \\ -0.1137 \end{bmatrix} \tag{19}$$

$$\theta = \theta - 0.1. \begin{bmatrix} 0.086 \\ -0.137 \end{bmatrix} = \begin{bmatrix} -0.0086 \\ 0.1137 \end{bmatrix} \tag{20}$$

PASO 6: Repetimos el paso anterior
Cada paso reduce el error y ajusta θ para que la curva logística se adapte a los datos