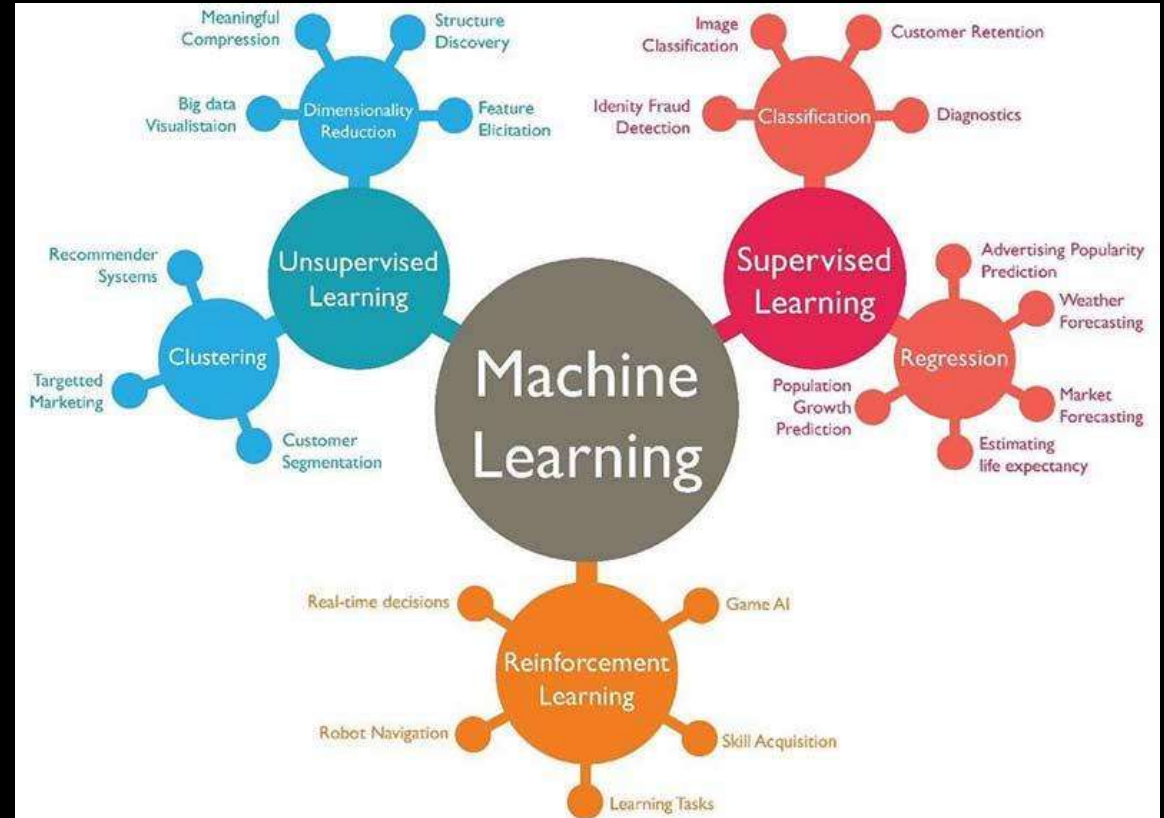


Machine Learning – Regresión Logística

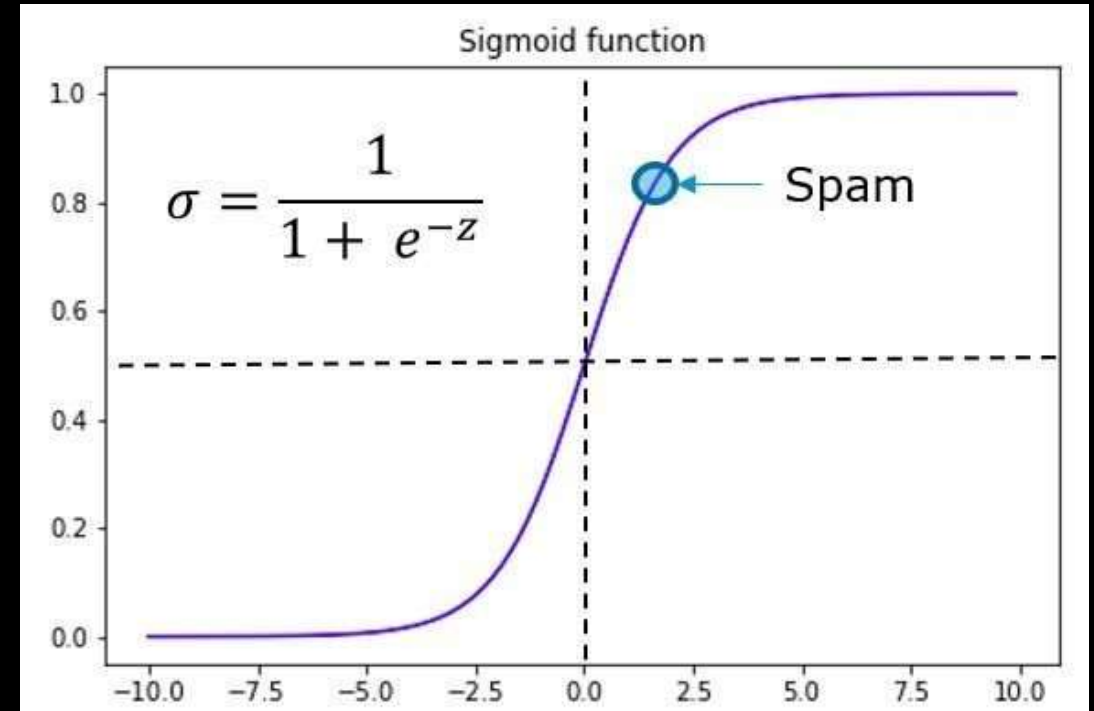
Algoritmo de clasificación

- Aprendizaje supervisado:
 - Regresión
 - Clasificación
- Aprendizaje no supervisado:
 - Clusterización
 - Reducción de dimensionalidad
- Aprendizaje por refuerzo



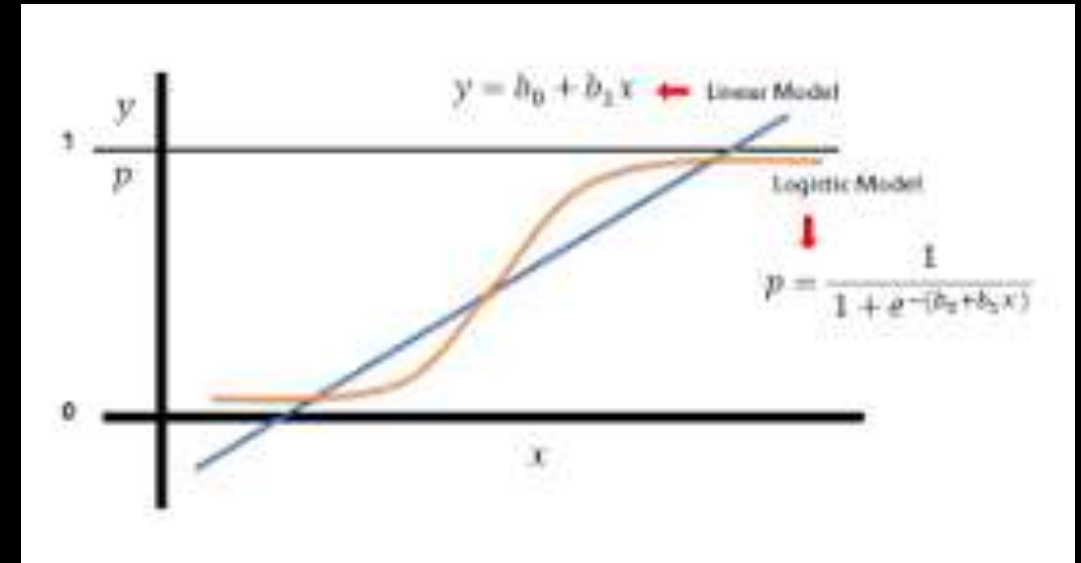
Regresión logística

- Don't get confused by its name! Es clasificación.
- Es un tipo de análisis de regresión utilizado para predecir el resultado de una variable categórica (y) en función de las variables independientes (x).
- Util para modelar la probabilidad de un evento ocurriendo como función de otros factores.



Regresión logística

- Su fórmula se obtiene a partir de la función sigmoide. Esta función tiene un valor de entrada y la operación que realiza ofrece un valor comprendido entre $[0,1]$.
- En la resolución de problemas de clasificación, este valor de entrada es toda la ecuación de la regresión lineal.
- Como hacíamos con la regresión lineal, la regresión logística trata de ajustar los valores para “a” y “b”



Regresión logística vs Regresión lineal

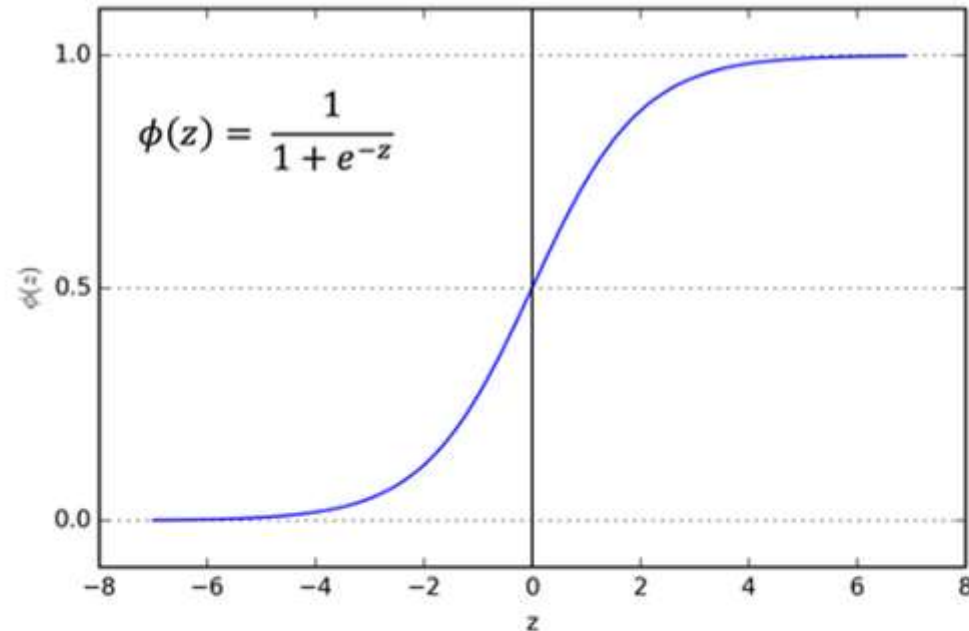
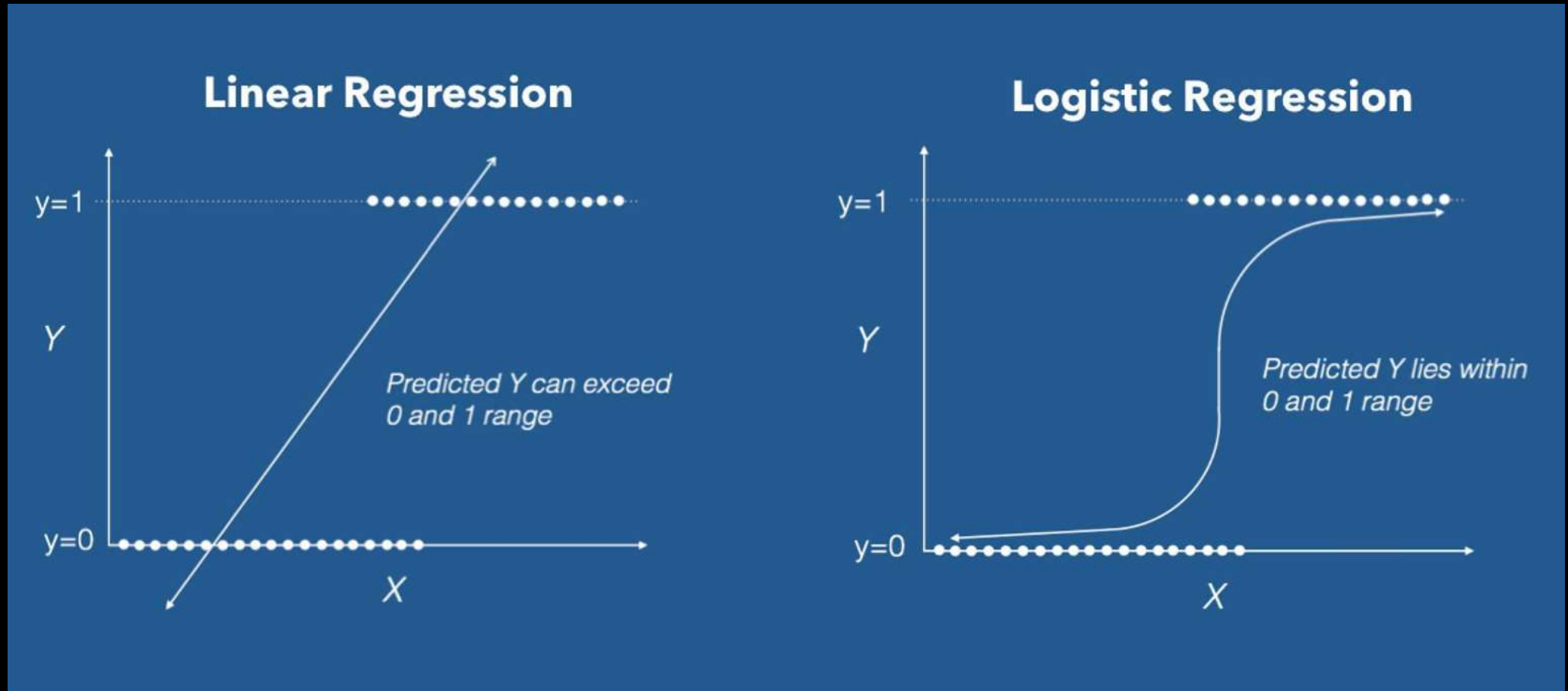


Fig: Sigmoid Function

1

If 'Z' goes to infinity, Y(predicted) will become 1 and if 'Z' goes to negative infinity, Y(predicted) will become 0.

Regresión logística vs Regresión lineal



Gradient Descent

Al igual que en Regresión Lineal, la Regresión Logística tiene su función de coste, que hay que minimizar para obtener los pesos (w) de la regresión que minimizan los errores.

¿Cómo solventamos esto? → De nuevo, **Gradient Descent**.

$$J(\theta) = \frac{1}{m} \sum_{i=1}^m \text{Cost}(h_{\theta}(x^{(i)}), y^{(i)})$$

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

m = number of samples

Ejemplo