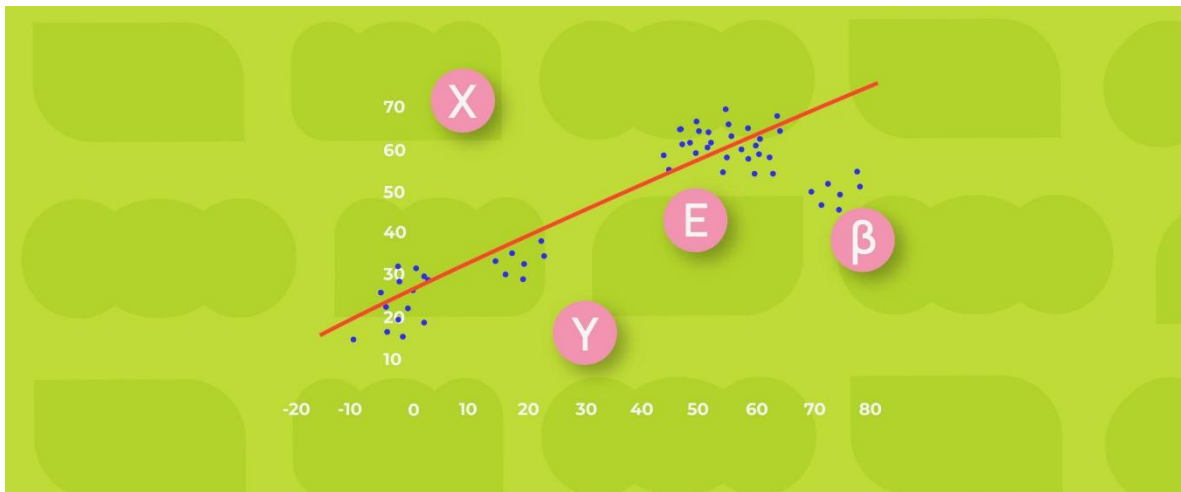


Regresión Lineal

Definición:

es intentar representar los puntos mediante una recta lineal (de ahí el nombre). Este algoritmo forma parte de los modelos supervisados, puesto que necesitamos datos etiquetados (la variable Y tiene que tener valores numéricos) para que el modelo aprenda a predecir.



¿Por qué es importante la regresión lineal?

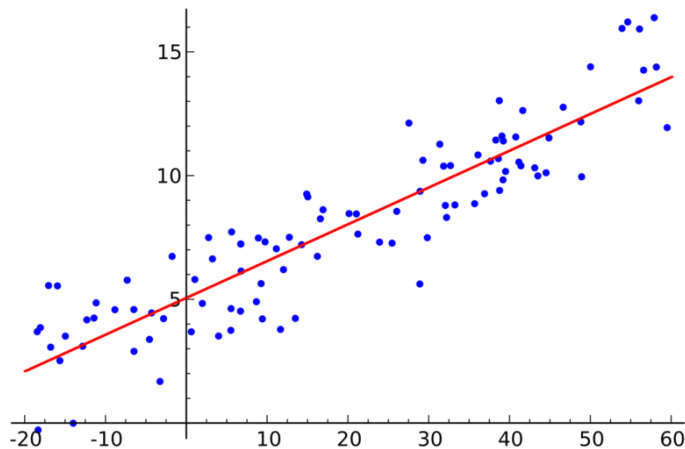
Los modelos de regresión lineal son muy populares en diversos campos de investigación gracias a su rapidez y facilidad de interpretación.

Debido a su capacidad para transformar datos, pueden utilizarse para simular una amplia gama de relaciones, y debido a su forma, que es más simple que la de las redes neuronales, sus parámetros estadísticos se analizan y comparan con facilidad, lo que permite que se les extraiga información valiosa.

Tipos de Regresiones lineales:

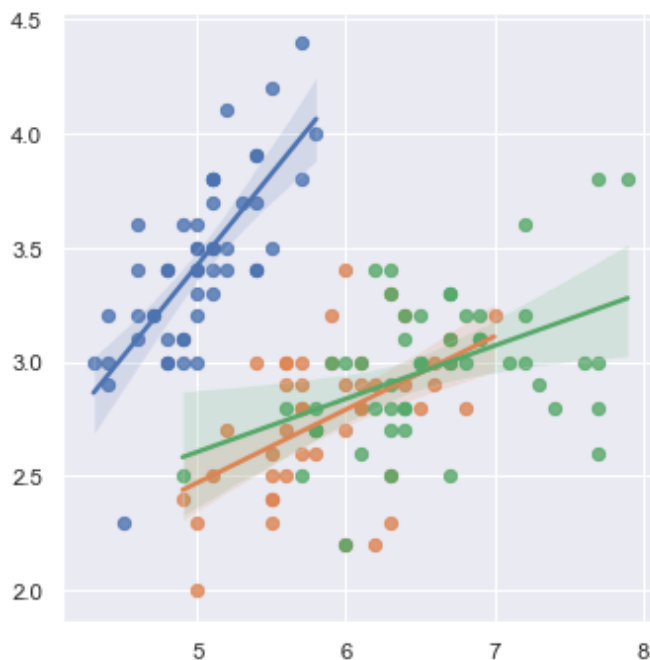
Regresión lineal simple (o univariante): en este tipo de regresión utilizamos sólo una variable más el intercepto para predecir algo. Por ejemplo, predecir el peso de una persona a partir de su altura:

$$\text{Peso} = \beta_0 + \beta_1 * \text{Altura}$$



Regresión lineal múltiple: utiliza más de una variable para hacer el ajuste de la recta. Por ejemplo, predecir el peso de una persona a partir de su altura, su edad, su sexo:

$$\text{Peso} = \beta_0 + \beta_1 * \text{Altura} + \beta_2 * \text{Edad} + \beta_3 * \text{Sexo}$$



¿En qué casos no deberías usar la regresión lineal?

No te aconsejo usarla si:

Los datos no tienen una relación lineal clara (por ejemplo, si la relación es curva o exponencial).

Existen valores atípicos extremos que pueden distorsionar el modelo.

Las variables no son independientes, lo que podría llevar a problemas de colinealidad.

El modelo no explica bien la variabilidad de los datos (bajo valor de R^2).