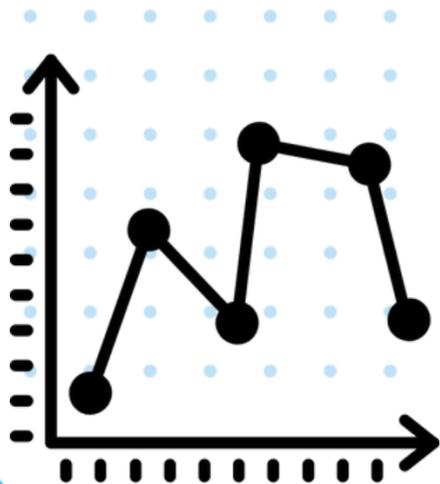


# Lección 1: Algoritmos de regresión



## × El problema de regresión

◁ La regresión es una técnica estadística para modelar la relación entre una variable dependiente y una o múltiples variables independientes. El resultado de la regresión es un modelo cuyas características describan el comportamiento de las variables analizadas.

▷ El análisis de regresión lineal se utiliza para estimar el valor de una variable en relación con el valor de otra variable. La variable que desea predecir se llama variable dependiente. La variable que utiliza para predecir el valor de la otra variable se llama variable independiente.

Los modelos de regresión lineal son relativamente simples y proporcionan una fórmula matemática fácil de interpretar que puede producir predicciones. La regresión lineal se puede aplicar a una variedad de campos de los estudios académicos y empresariales.

Debido a que la regresión lineal es un procedimiento estadístico bien establecido, es fácil entender las propiedades de un modelo y entrenar diferentes tipos de modelos según sea el caso.

# Regresión lineal simple



El modelo de regresión lineal simple es un método utilizado para estimar la variable dependiente con la ayuda de la variable independiente cuando existe una relación lineal entre la variable independiente y la variable dependiente. Es decir, los datos se pueden modelar como una línea recta.



En su forma canónica el modelo se puede escribir como:



$$y = A + Bx + e$$





En donde la pareja de datos  $(x, y)$  son los puntos de datos (proviene del conjunto de datos). Los coeficientes  $A$  y  $B$  modelan el intercepto de la recta con el cero y la pendiente de la recta respectivamente y el valor 'e' se agrega para modelar el error del modelo. Nota: Ya que se trabajan con datos medidos, siempre habrá un error asociado a los datos, generado en la adquisición o en el procesamiento, que debe modelarse.



# Regresión polinómica

La regresión polinómica se trata de encontrar un modelo que permita explicar datos no lineales mediante un polinomio. Para esto, se selecciona el grado del polinomio a ajustar y se entrenará el modelo, es decir, se encontrarán los coeficientes que mejor se ajusten a los datos. El modelo polinómico es de la forma:

$$y = A + Bx + Cx^2 + Dx^3 + \dots + W_n x^n + e$$

En donde los coeficientes  $A, B, C... W$  son el grado del polinomio y son los valores para ajustar de acuerdo con los datos. La variable  $x$  representa los valores independientes y la variable  $y$  los valores dependientes ( $x$  y  $y$  provienen de los datos).