

Regresión Lineal

La regresión lineal es un modelo matemático que describe la relación entre varias variables. Los modelos de regresión lineal son un procedimiento estadístico que ayuda a predecir el futuro. Se utiliza en los campos científicos y en los negocios, y en las últimas décadas se ha utilizado en el aprendizaje automático.

La tarea de la regresión en el aprendizaje automático consiste en predecir un parámetro (Y) a partir de un parámetro conocido X.

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.

La regresión lineal no sólo se utiliza con fines de predicción: también ha demostrado su eficacia para describir sistemas. Si quieres modelar los valores de una variable numérica, tendrás una lista relativamente corta de variables independientes y, como esperas que el modelo sea comprensible, es probable que elijas la regresión lineal como herramienta de modelización.

Regresión lineal simple

En una regresión lineal, se trata de establecer una relación entre una variable independiente y su correspondiente variable dependiente. Esta relación se expresa como una línea recta.

No es posible trazar una línea recta que pase por todos los puntos de un gráfico si estos se encuentran ordenados de manera caótica. Por lo tanto, sólo se determina la ubicación óptima de esta línea mediante una regresión lineal.

Algunos puntos seguirán distanciados de la recta, pero esta distancia debe ser mínima.

El cálculo de la distancia mínima de la recta a cada punto se denomina función de pérdida.

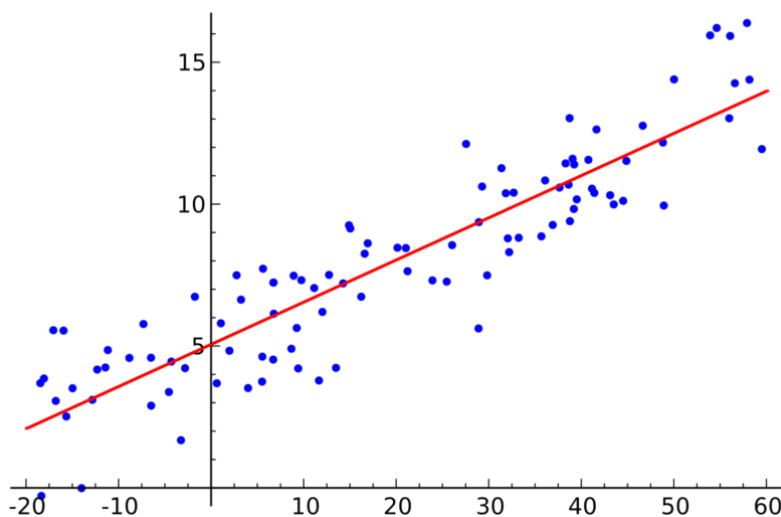
La ecuación de una línea recta tiene la siguiente forma:

$$Y = \beta_0 + \beta_1 X + \varepsilon,$$

donde:

1. Y es la variable independiente.
2. β_0 y β_1 son dos constantes desconocidas que representan el punto de intersección y la pendiente respectivamente.
3. ε (epsilon) es la función de pérdida.

A continuación se muestra un ejemplo gráfico de un modelo de una regresión lineal simple:



Aplicación de la regresión lineal simple:

1. Para predecir la cosecha en función de la precipitación, con la precipitación como variable independiente y la cosecha como variable dependiente.
2. Para saber qué calificación obtendrán los alumnos en función del número de horas que estudien: aquí la cantidad de horas de estudio representa la variable independiente y las calificaciones, la dependiente.
3. Para prever el salario basado en la experiencia: la experiencia se convierte en la variable independiente y el salario en la variable dependiente.

Limitaciones de la regresión lineal simple:

La regresión lineal simple establece que existe una relación entre las variables, pero no revela una relación causal: Y depende de X, pero no implica que genere a Y.

Si necesitas establecer algo más que la existencia de una relación, tendrás que hacer análisis adicionales.

Regresión lineal Múltiple

La regresión lineal múltiple encuentra la relación entre dos o más variables independientes y su correspondiente variable dependiente.

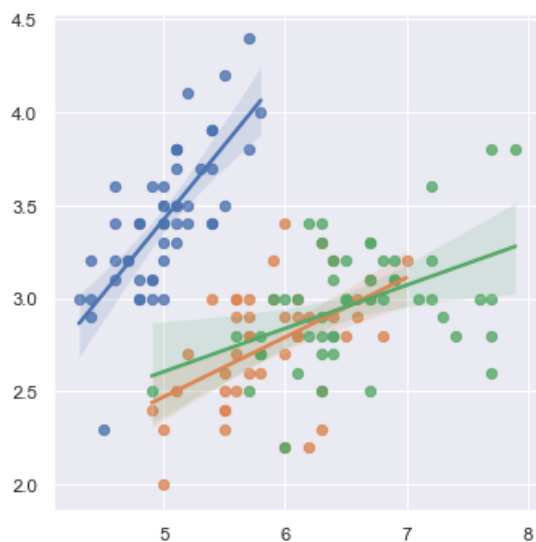
La ecuación de regresión lineal múltiple tiene la siguiente forma:

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_a X_a + \varepsilon$$

Donde:

1. Y es la variable dependiente.
2. X es una variable independiente.
3. β son coeficientes.
4. ε (epsilon) es la función de pérdida.

A continuación se muestra un ejemplo de gráfico de un modelo de regresión lineal múltiple:



Aplicaciones de la regresión lineal múltiple:

Este tipo de regresión permite predecir tendencias y valores futuros.

El análisis de regresión lineal múltiple ayuda a determinar el grado de influencia de las variables independientes sobre la variable dependiente, es decir, cuánto cambiará la variable dependiente cuando cambiemos las variables independientes.