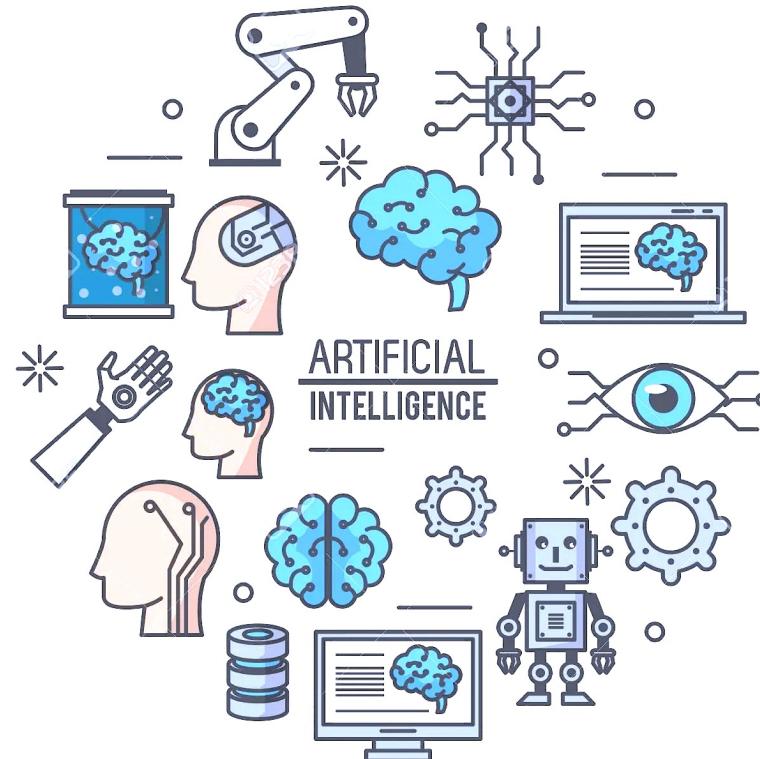




Machine Learning I

Inteligencia Artificial



La Inteligencia Artificial es una disciplina que intenta replicar la inteligencia humana en las máquinas.



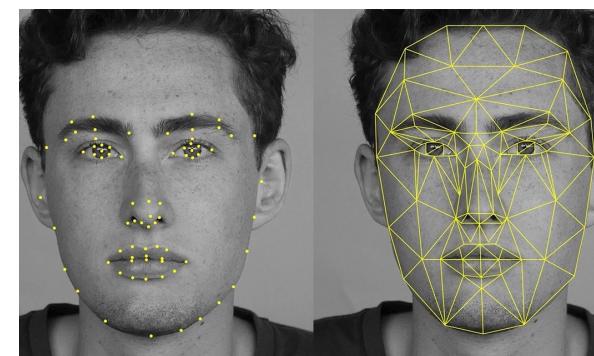
Busca crear un comportamiento inteligente en máquinas, dándoles la capacidad de percibir su entorno y realizar acciones que maximicen una recompensa.



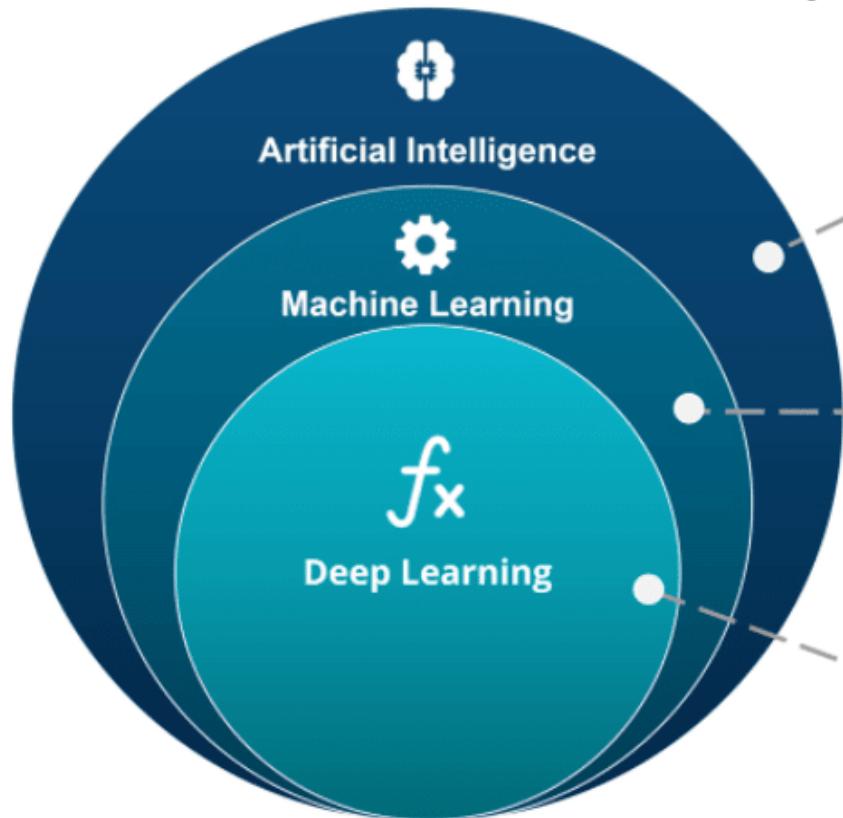
Uno de los objetivos más ambiciosos que se ha planteado la ciencia es lograr que una máquina tenga una inteligencia de tipo general similar a la humana,

Inteligencia Artificial

Automatización en Marketing	Reconocimiento de voz	Agentes virtuales
Reconocimiento de emociones	Creación de contenido	Aplicaciones Biométricas
Automatización de procesos	Defensa cibernética	Reconocimiento de señas
Generación de lenguaje natural	Plataformas machine learning	Predicción de valores



Inteligencia Artificial



ARTIFICIAL INTELLIGENCE

A technique which enables machines to mimic human behaviour

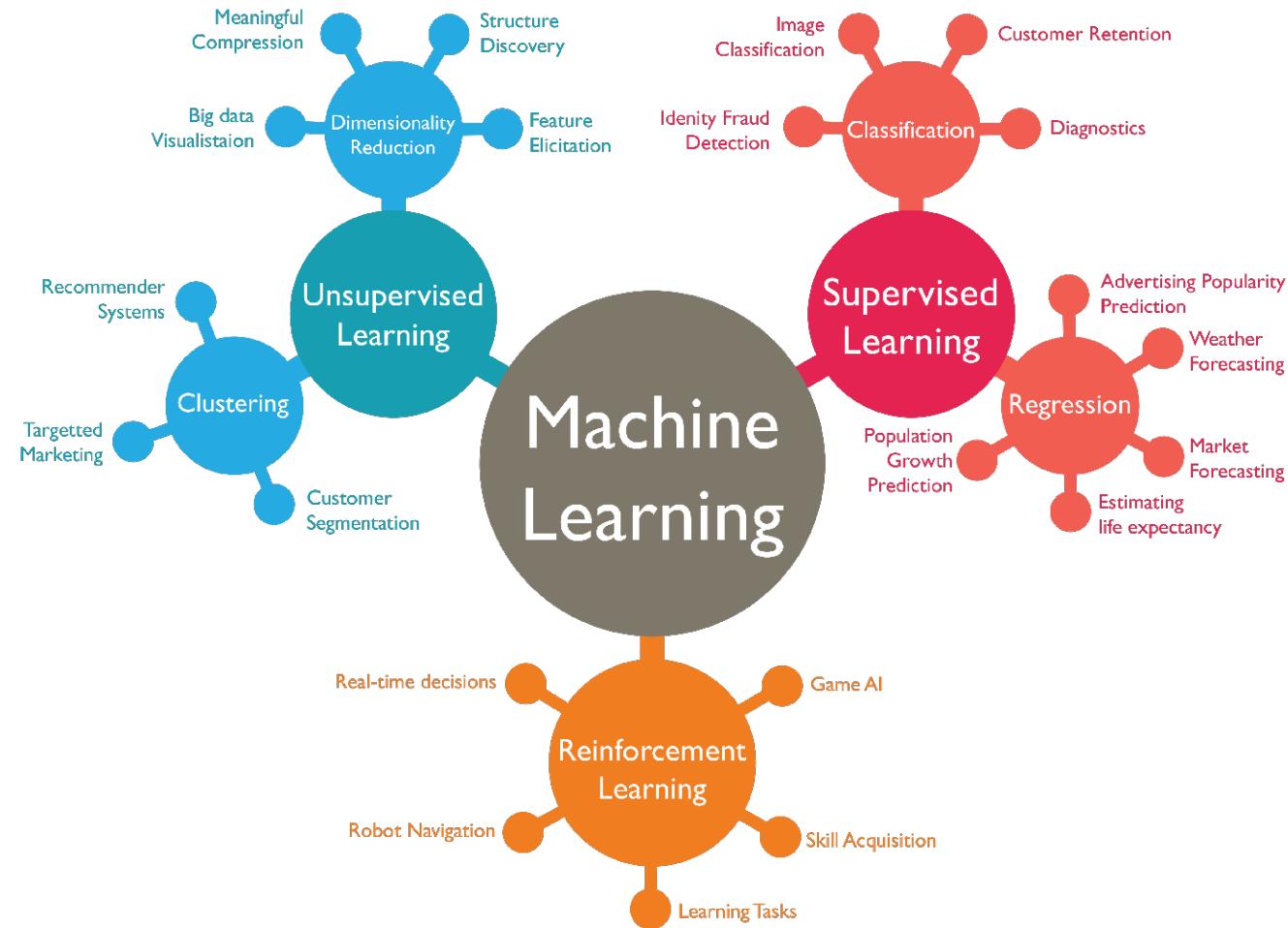
MACHINE LEARNING

Subset of AI technique which use statistical methods to enable machines to improve with experience

DEEP LEARNING

Subset of ML which make the computation of multi-layer neural network feasible

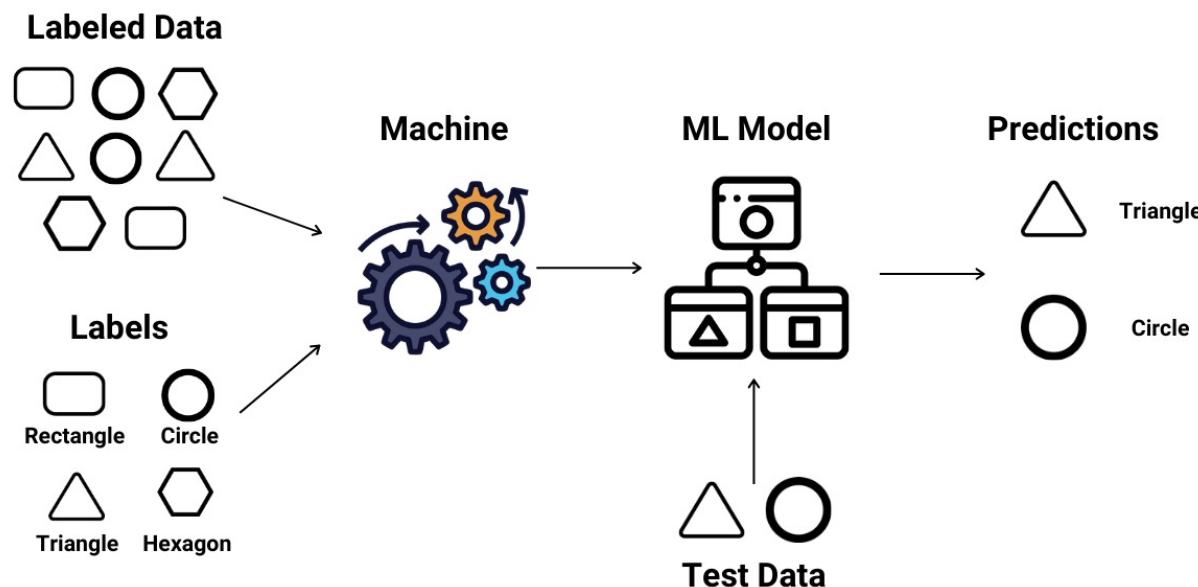
Inteligencia Artificial



Aprendizaje Supervisado



El aprendizaje supervisado es una subcategoría dentro de Machine Learning que se define por su uso de conjuntos de datos etiquetados para entrenar algoritmos que clasifican datos o predicen resultados con precisión.



Aprendizaje Supervisado

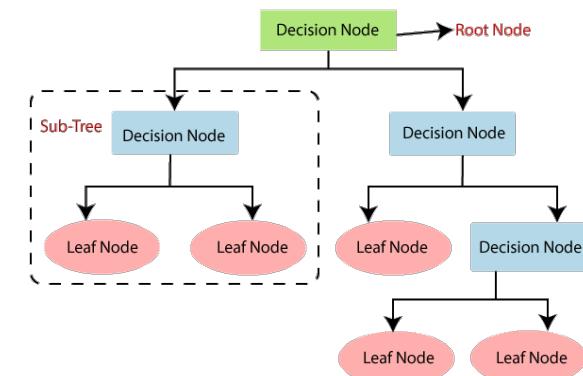
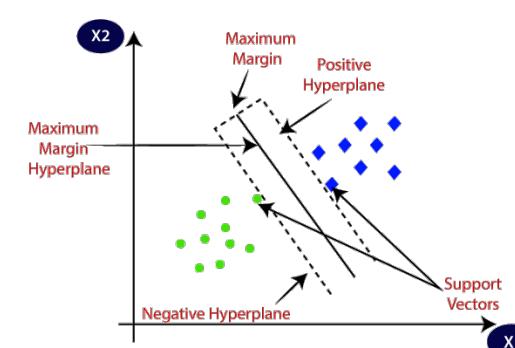
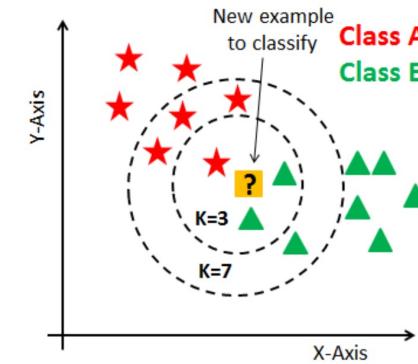
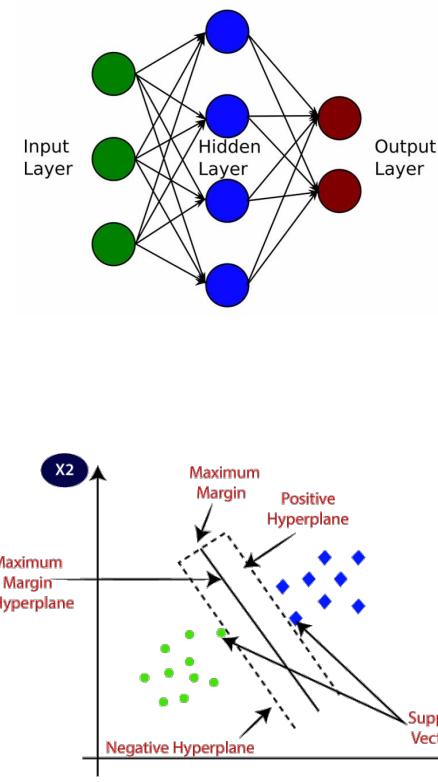


Los algoritmos de aprendizaje supervisado generan principalmente dos tipos de resultados: clasificación y regresión.

Input (X)	Output (Y)	Application
email	spam? (0/1)	spam filtering
audio	text transcripts	speech recognition
English	Spanish	machine translation
ad, user info	click? (0/1)	online advertising
image, radar info	position of other cars	self-driving car
image of phone	defect? (0/1)	visual inspection

Aprendizaje Supervisado

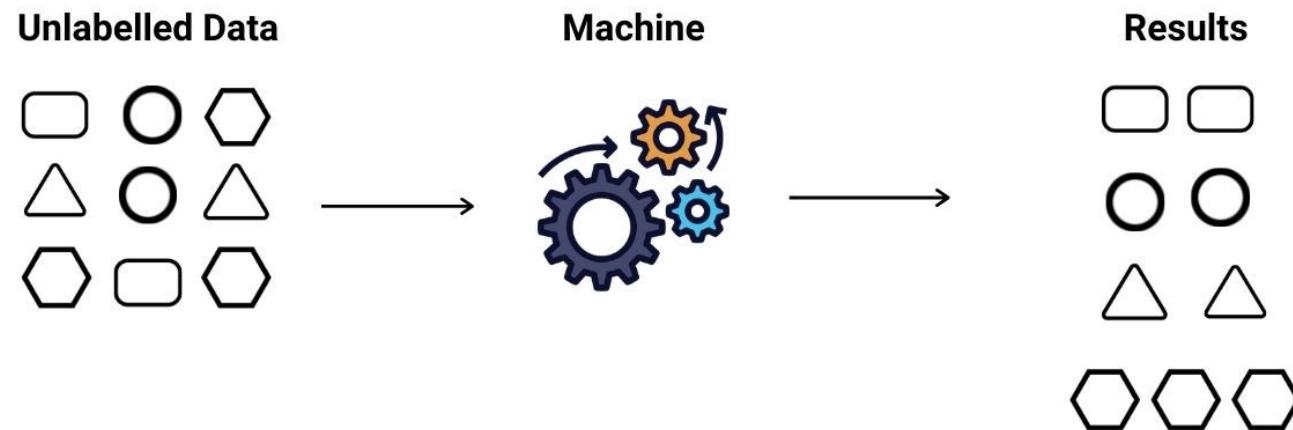
- Regresion Logistica
- Redes Neuronales
- Decision Trees
- Support Vector Machine (SVM)
- K-Nearest Neighboors (KNN)
- Random Forest
- Naive Bayes
- XGBoost
- Convolutional Neural Networks
- Recurrent Neural Networks
- y muchos mas



Aprendizaje No Supervisado



El aprendizaje no supervisado es una de las formas en que Machine Learning (ML) "aprende" los datos. El aprendizaje no supervisado tiene datos sin etiquetar que el algoritmo tiene que intentar entender por sí mismo.

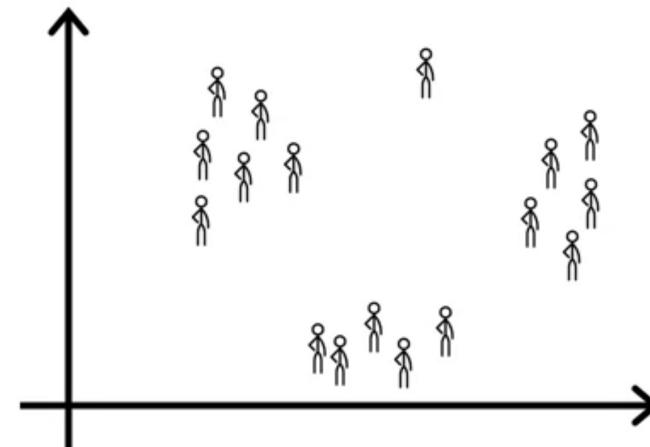


Aprendizaje No Supervisado



Estos algoritmos descubren patrones ocultos o agrupaciones de datos sin necesidad de intervención humana.

Clustering: Grouping customers

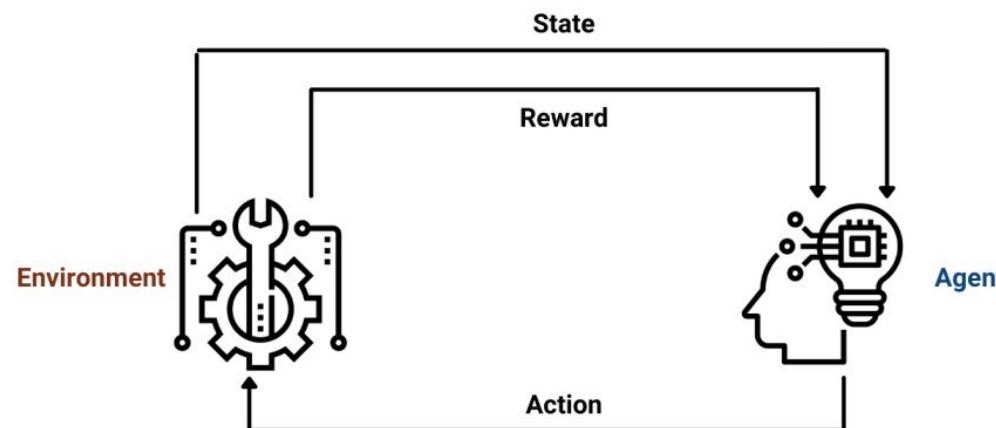


Aprendizaje por Refuerzo



El aprendizaje por refuerzo es un método de entrenamiento de aprendizaje automático basado en recompensar los comportamientos deseados y/o castigar los no deseados.

En general, un agente de aprendizaje por refuerzo es capaz de percibir e interpretar su entorno, tomar acciones y aprender a través de prueba y error.



Pregunta

De los siguientes ejemplos ¿Cuales son aplicaciones de algoritmos de aprendizaje supervisado ?

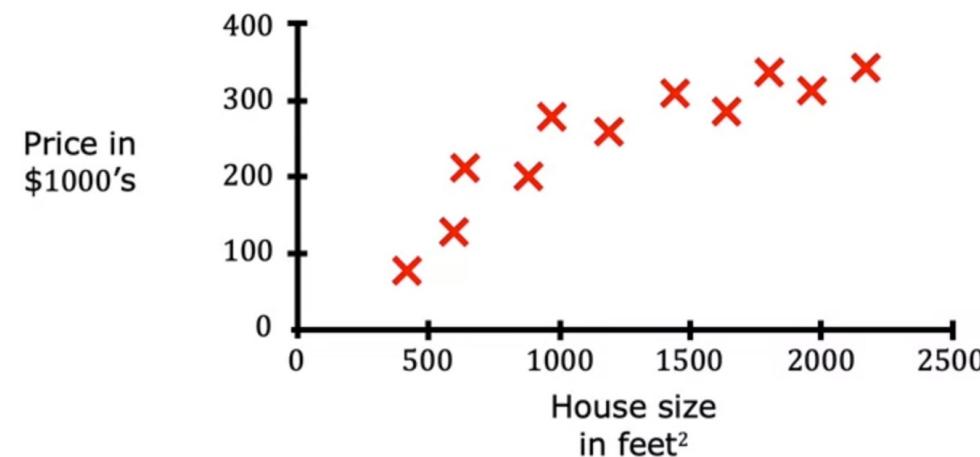
- A. Dado emails etiquetados como spam/no spam, aprender un filtro de spam
- B. Dado un conjunto de nuevos articulos encontrados en la web, separarlos en grupos de articulos acerca del mismo tema.
- C. Dado una base de datos de clientes, automaticamente descubrir segmentos de marketing.
- D. Dado una base de datos de pacientes diagnosticados como diabeticos y no diabeticos, aprender a clasificar nuevos pacientes como diabeticos o no diabeticos.

Regresión Lineal



La regresión establecerá un modelo para ajustar la relación de dependencia entre una característica específica independiente y el valor "resultado" correspondiente

Regression: Housing price prediction



Regresión Lineal



Data table

size in feet ²	price in \$1000's
2104	400
1416	232
1534	315
852	178
...	...
3210	870

X → Y

x = Variable de entrada (input)

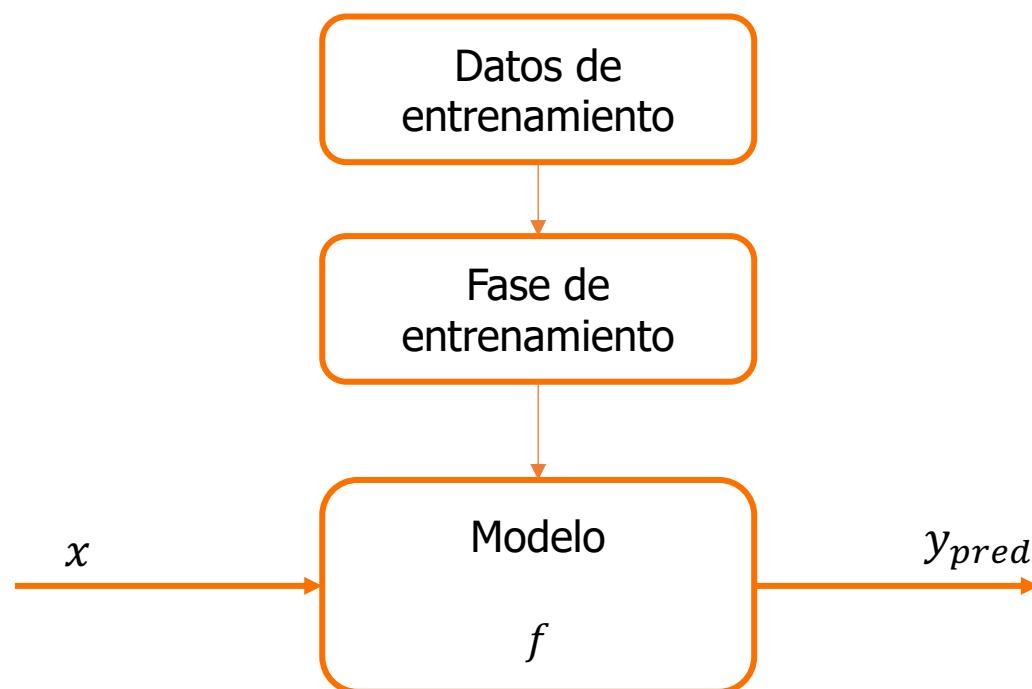
y = Variable de salida (output)

m = Numero de muestras de entrenamiento

$(x^{(i)}, y^{(i)})$ = Muestra de entrenamiento i^{th}



Regresión Lineal



Regresion Lineal

$$y_{pred} = f(x) = wx + b$$

w, b : parametros o pesos
(weights)



Regresión Lineal

Train – Test Split

A

Training

Test

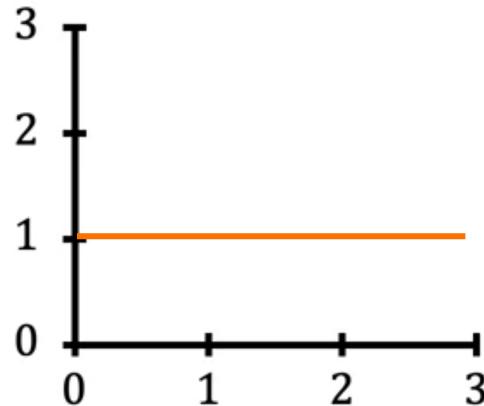
Single Dataset

Data table	
size in feet ²	price in \$1000's
2104	400
1416	232
1534	315
852	178
...	...
3210	870

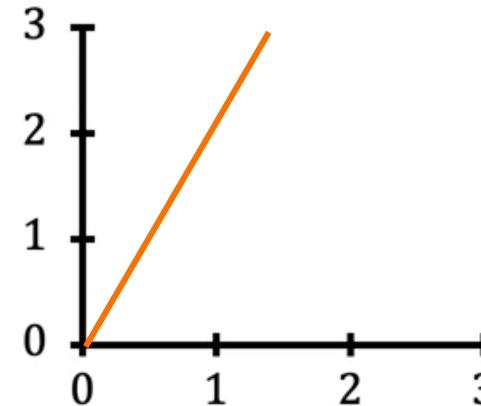
Regresión Lineal

$$f_{w,b}(x) = wx + b$$

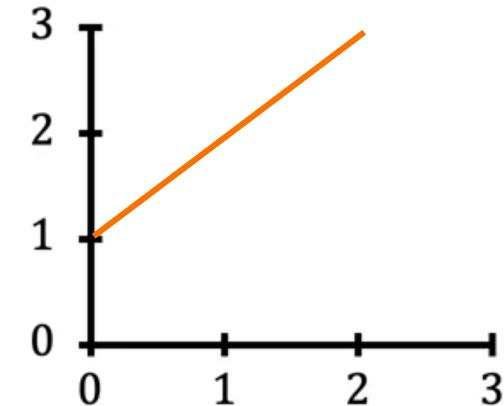
$f(x)$



$$\begin{aligned} w &= 0 \\ b &= 1 \end{aligned}$$

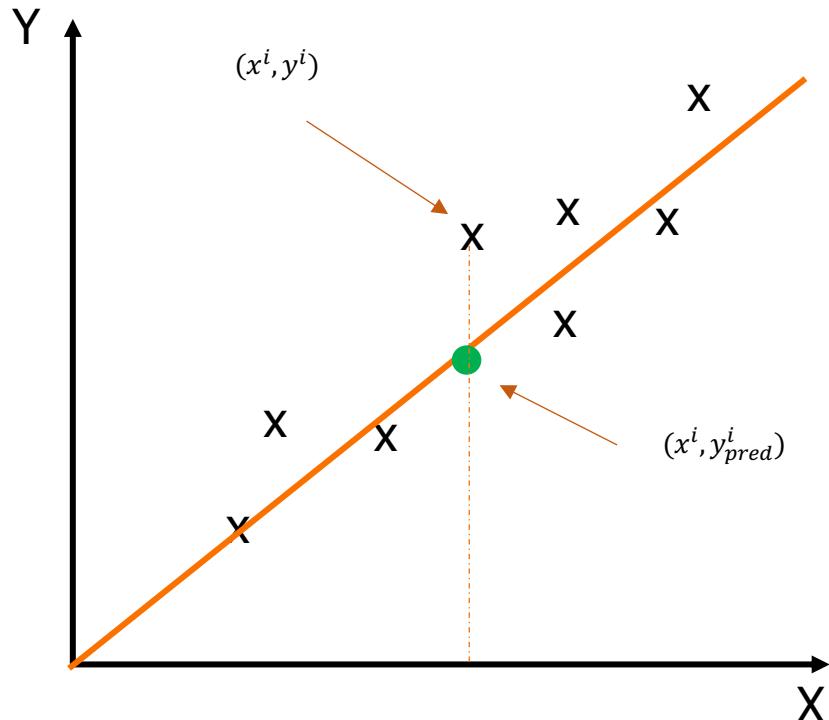


$$\begin{aligned} w &= 2 \\ b &= 0 \end{aligned}$$



$$\begin{aligned} w &= 1 \\ b &= 1 \end{aligned}$$

Regresión Lineal



Funcion costo : Funcion que determina el error entre el valor estimado y el valor real, con el fin de optimizar los parámetros de un modelo

$$J(w, b) = \frac{1}{2m} \sum_{i=1}^m (y_{pred}^i - y^i)^2$$

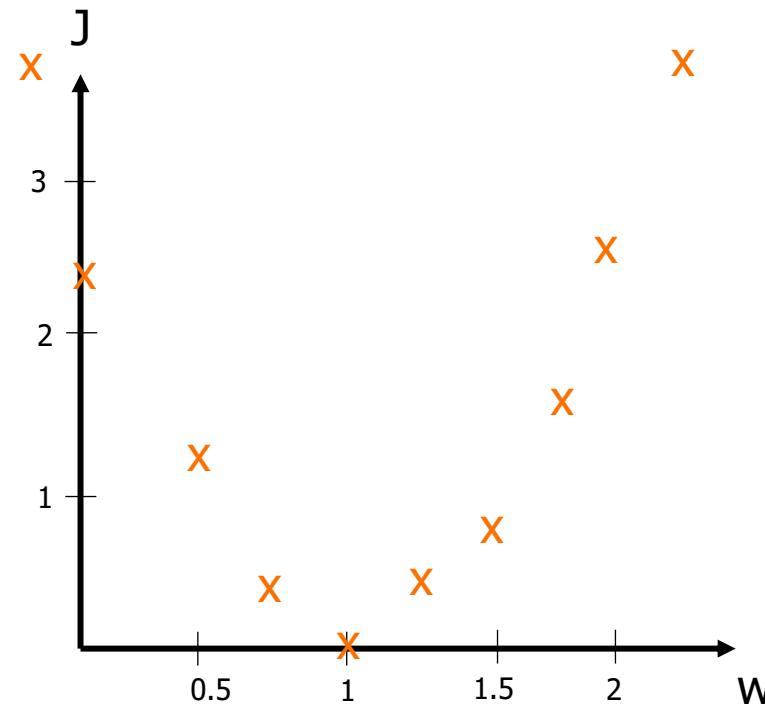
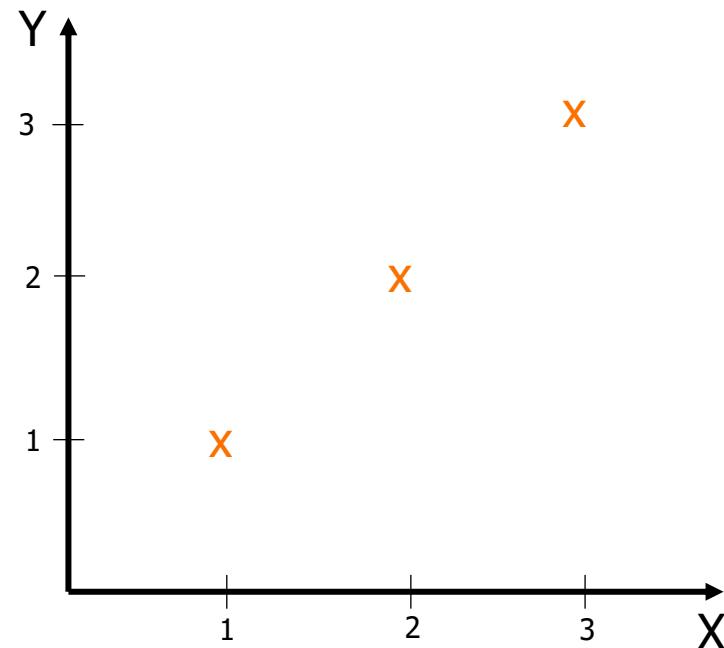
$$J(w, b) = \frac{1}{2m} \sum_{i=1}^m (f(x^i) - y^i)^2$$

El objetivo es encontrar los parametros que minimicen el valor de J

Regresión Lineal

$$y_{pred} = f(x) = wx + b$$

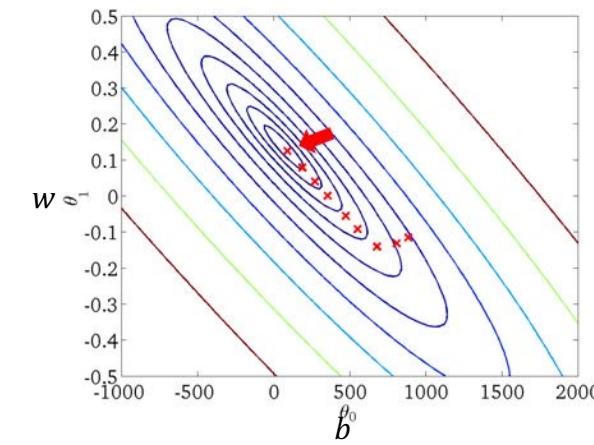
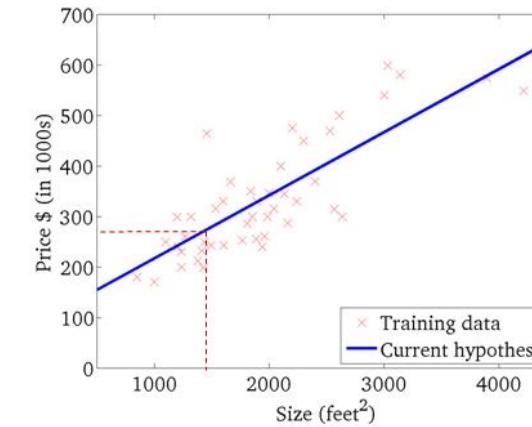
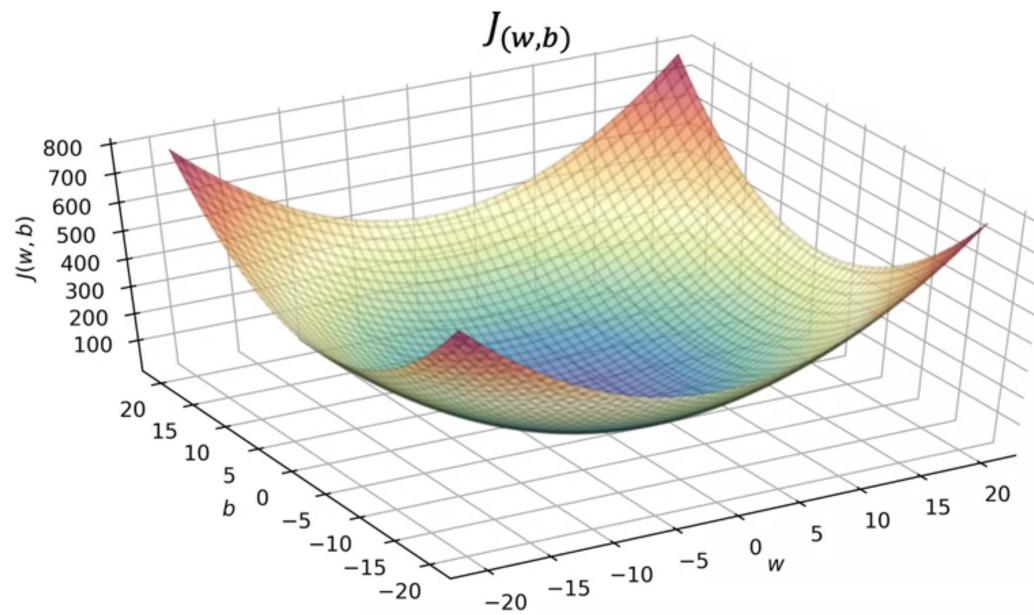
$$J(w, b) = \frac{1}{2m} \sum_{i=1}^m (y_{pred}^i - y^i)^2$$



Para distintos w y manteniendo $b=0$

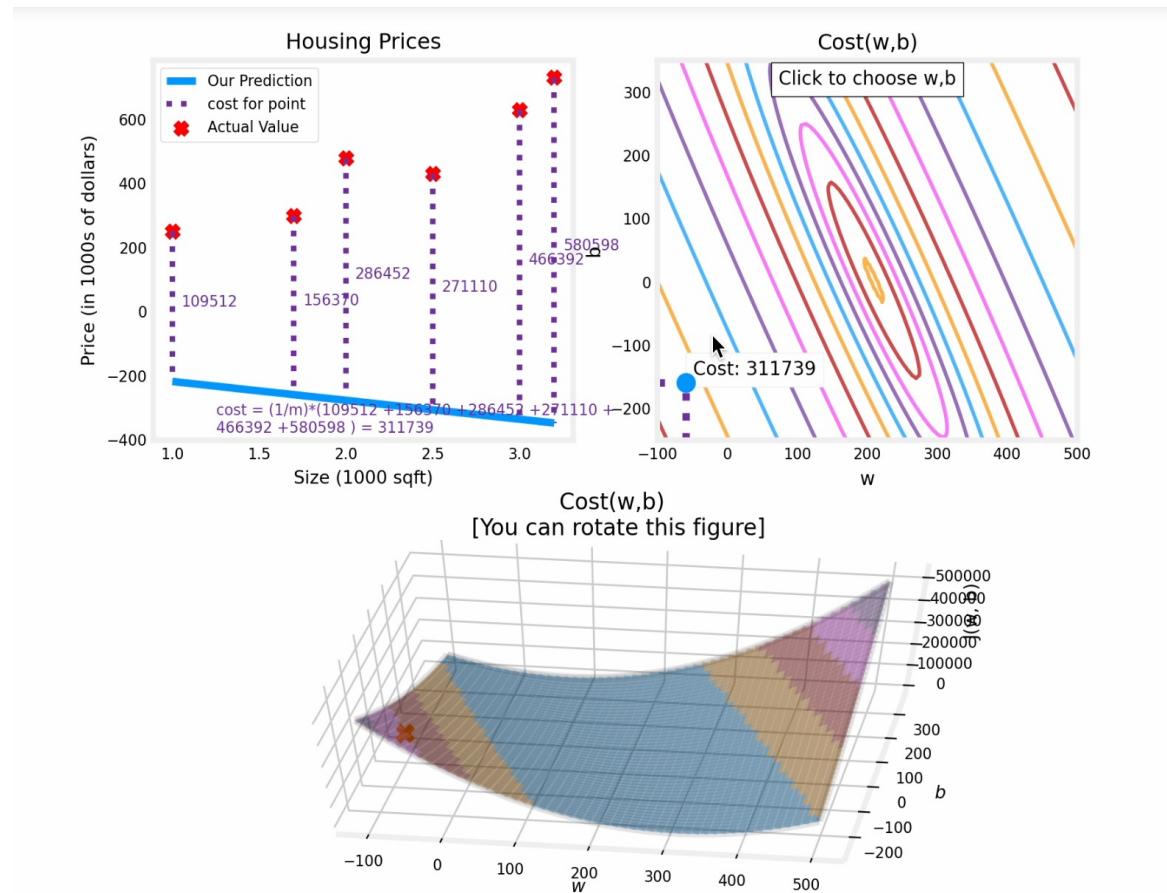
Regresión Lineal

Landscape – Función costo



Regresión Lineal

Landscape – Función costo



Regresión Lineal

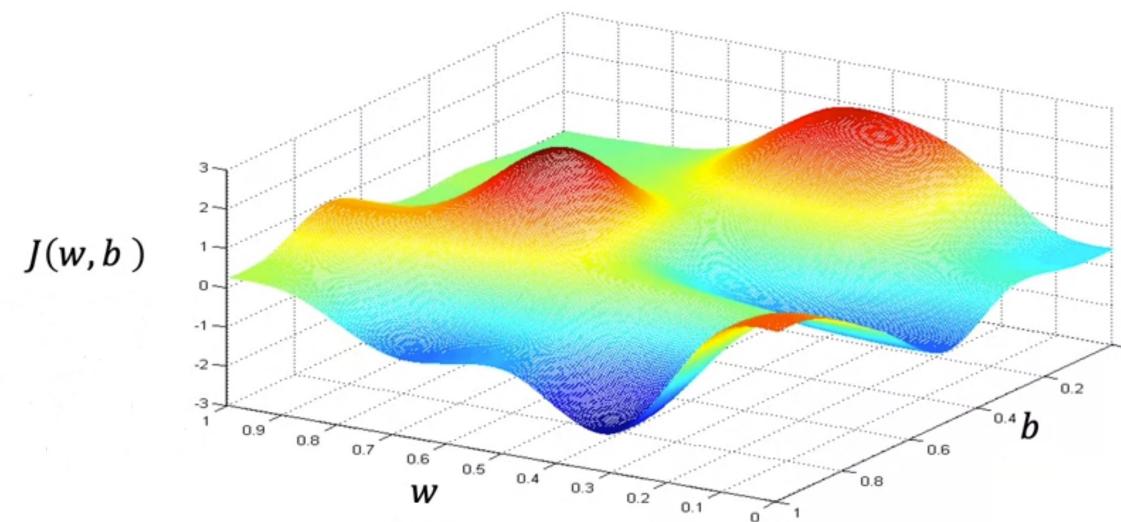
Landscape – Función costo

Objetivo

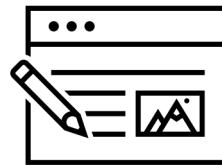
Encontrar los parametros que minimicen el valor de J

Como lograrlo

Gradient Descent



Pregunta



Para la regresión lineal, si se encuentra los parámetros w y b de modo que $J(w,b)$ esté muy cerca de cero, ¿qué puede concluir?

- A. Los valores seleccionados de los parámetros w y b hacen que el algoritmo se ajuste muy bien al conjunto de entrenamiento
- B. Esto nunca es posible; debe haber un error en el código.
- C. Los valores seleccionados de los parámetros w y b hacen que el algoritmo se ajuste muy mal al conjunto de entrenamiento.

Regresión Lineal

Gradient Descent

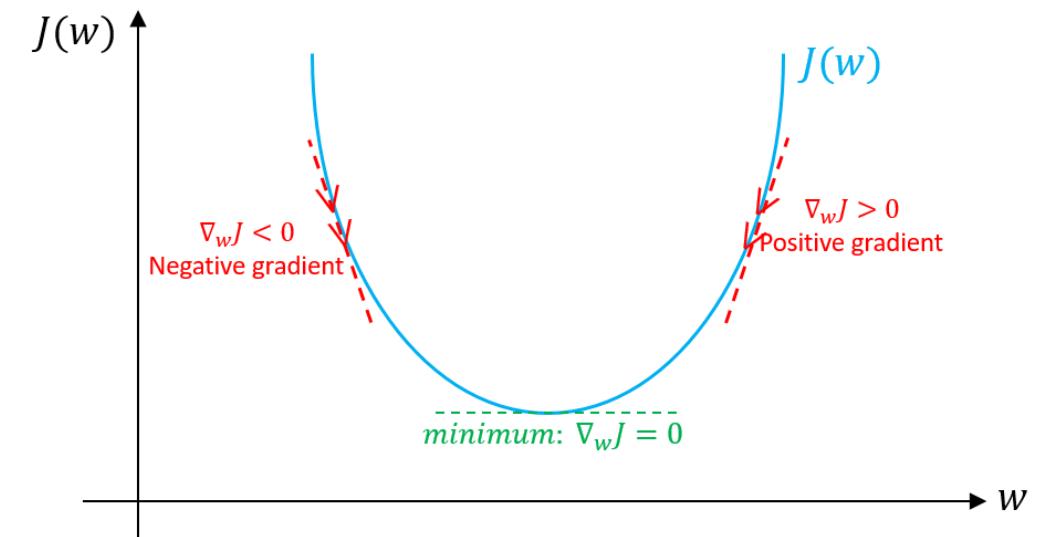


$$J(w, b) = \frac{1}{2m} \sum_{i=1}^m (f(x^i) - y^i)^2$$

Repetir

$$w := w - \alpha \frac{\partial}{\partial w} J(w, b)$$

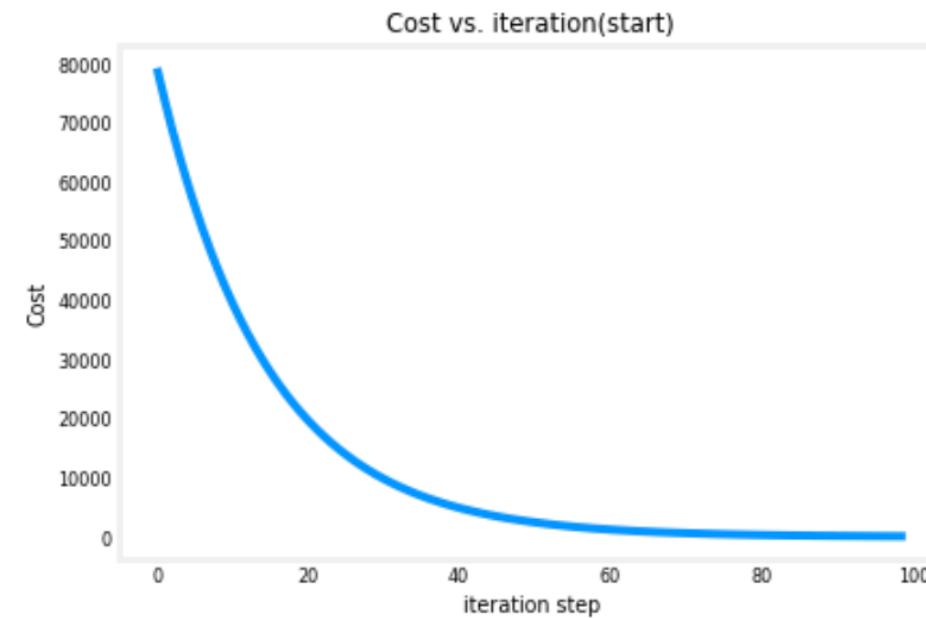
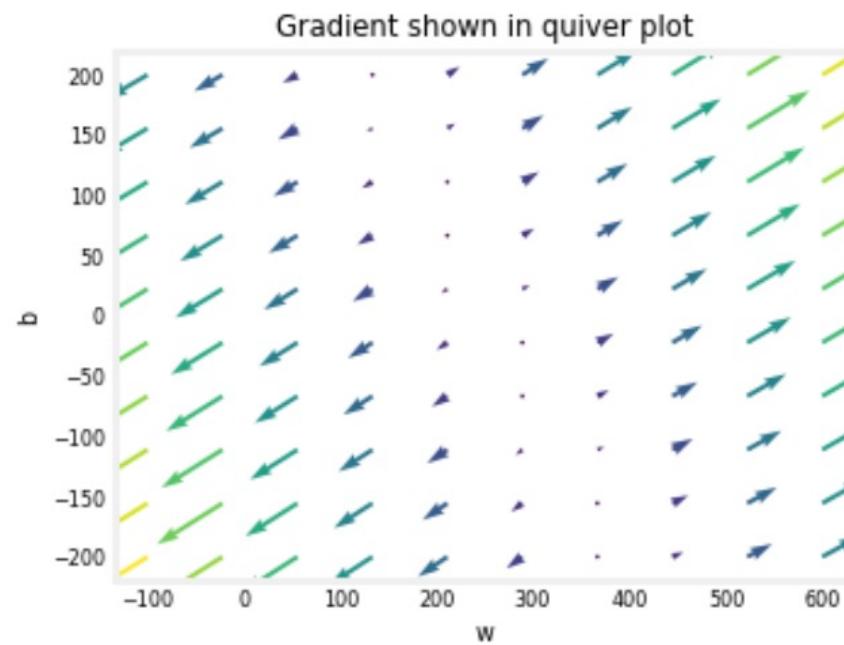
$$b := b - \alpha \frac{\partial}{\partial b} J(w, b)$$



El gradiente apunta en la dirección del ascenso más empinado

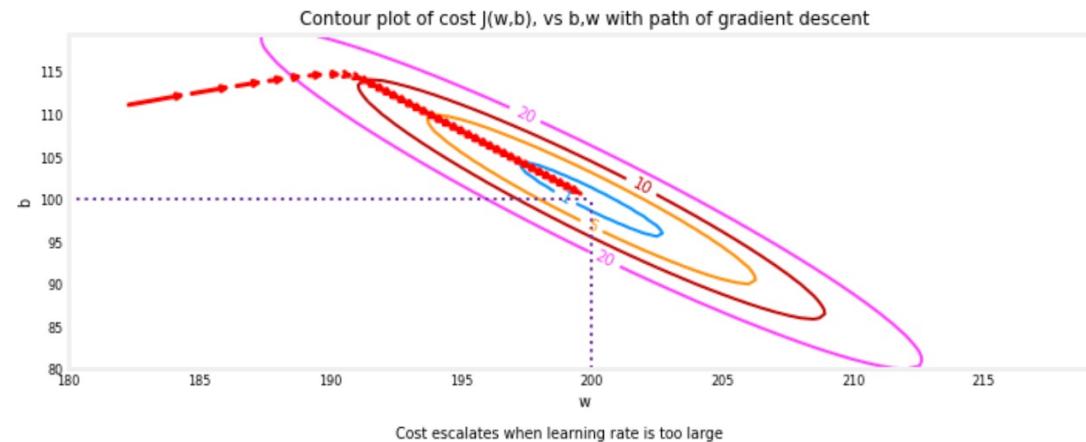
Regresión Lineal

Gradient Descent



Regresión Lineal

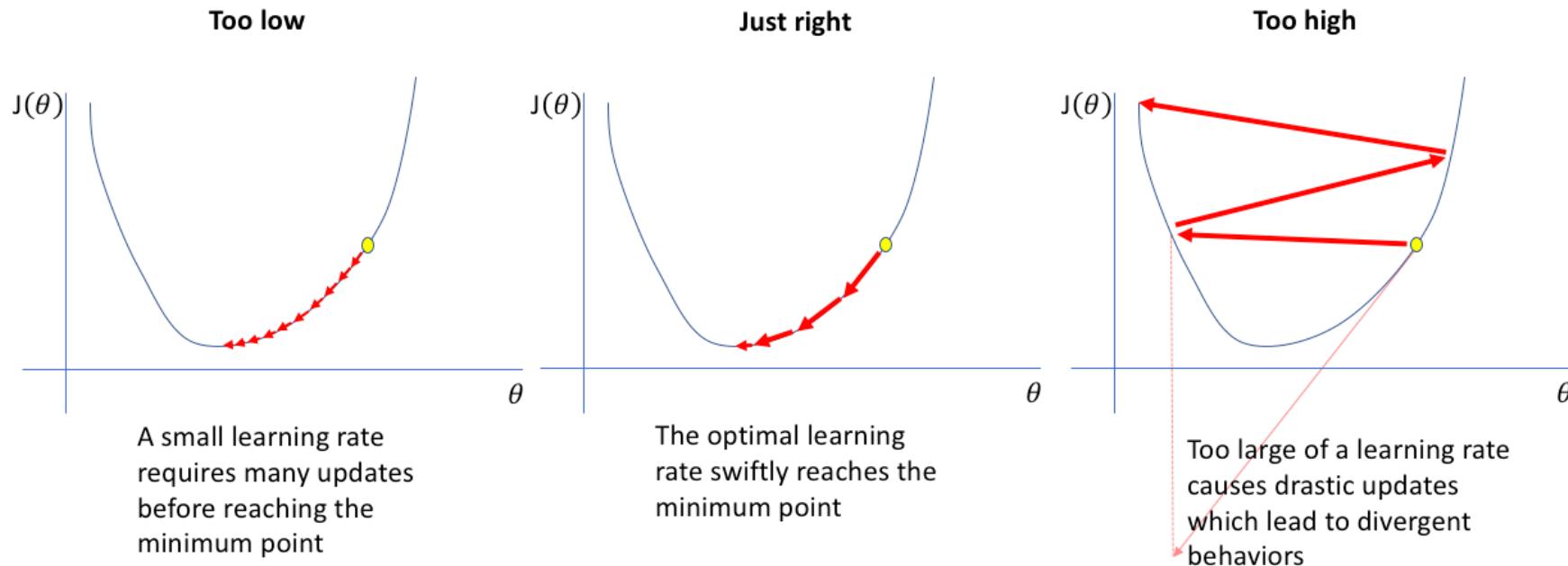
Gradient Descent



Regresión Lineal

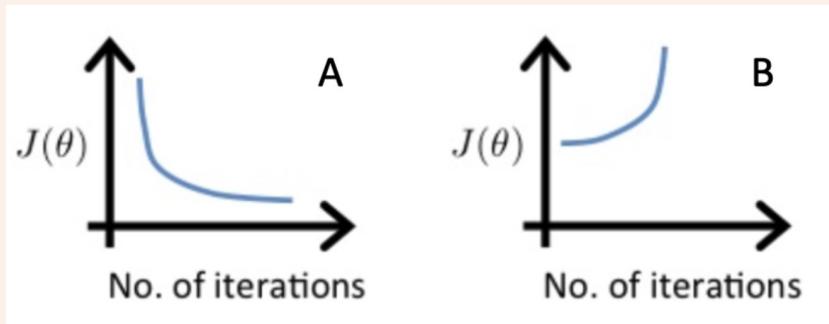
Learning rate

$$w := w - \alpha \frac{\partial}{\partial w} J(w, b)$$



Pregunta

¿Para qué caso, A o B, el valor de learning rate (α) probablemente fue demasiado alta?



- A. Solo caso B
- B. Ambos casos A y B
- C. Ni el caso A ni el B
- D. caso A solamente

Regresión Lineal

Feature Scaling

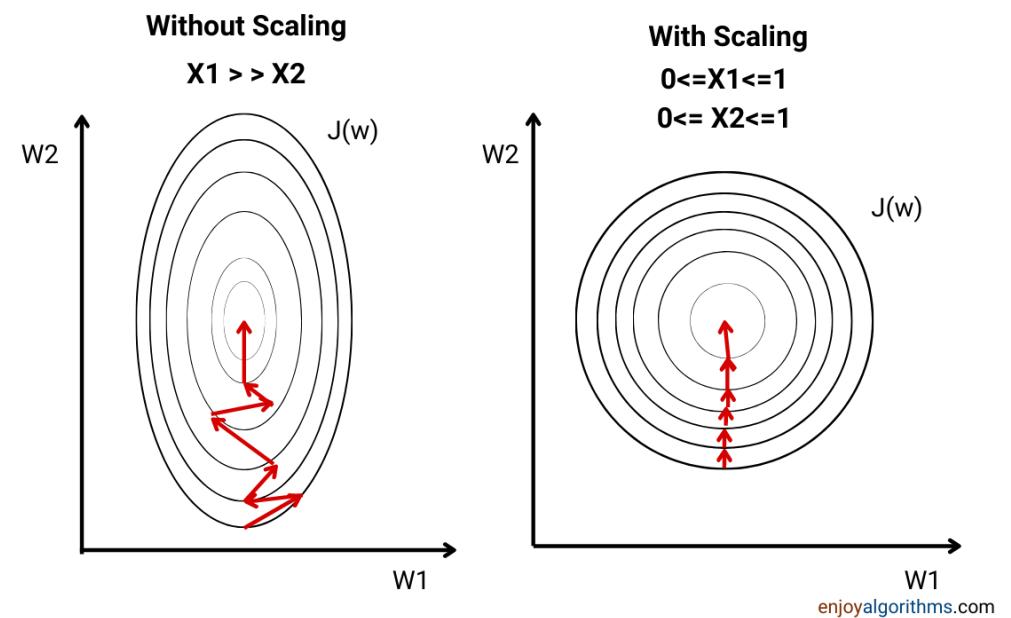


$$y_{pred} = f(x) = w_1x_1 + w_2x_2 + b$$

Ejemplo:

$$\begin{aligned} 300 < x_1 < 2000 \\ 0 < x_2 < 5 \end{aligned}$$

Como afecta a los valores de w_1 y w_2 ?



Regresión Lineal

Feature Engineering



En base a los datos encontrados, seleccionar los atributos (features) adecuados , modificando, transformando o combinando los atributos originales

$$f(x) = w_1x_1 + w_2x_2 + w_3x_3 + b$$

Ejemplos:

$$f(x) = w_1x_1 + w_2x_2 + w_3x_1x_2 + b \quad \longrightarrow \quad x_1 = x_1, x_2 = x_2, x_3 = x_1x_2$$

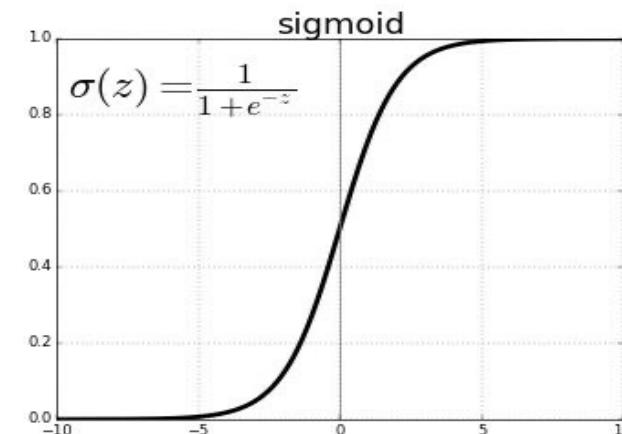
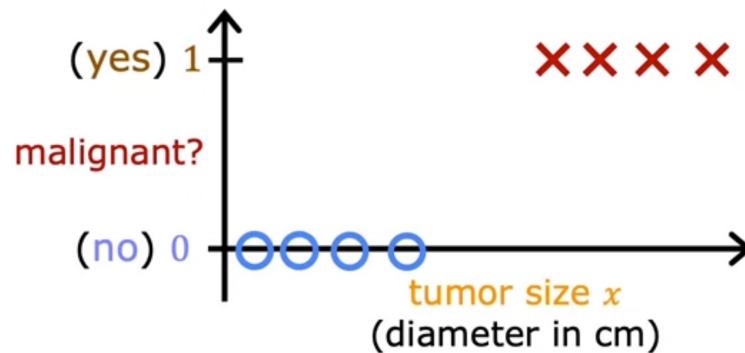
$$f(x) = w_1x_1 + w_2x_1^2 + w_3x_1^3 + b \quad \longrightarrow \quad x_1 = x_1, x_2 = x_1^2, x_3 = x_1^3$$

$$f(x) = w_1x_1 + w_2\sqrt{x_1} + w_3x_3 + b \quad \longrightarrow \quad x_1 = x_1, x_2 = \sqrt{x_1}, x_3 = x_3$$

Regresión Logística



La regresión logística estima la probabilidad de que ocurra un evento, como votar o no votar, en función de un conjunto de datos determinado de variables independientes.

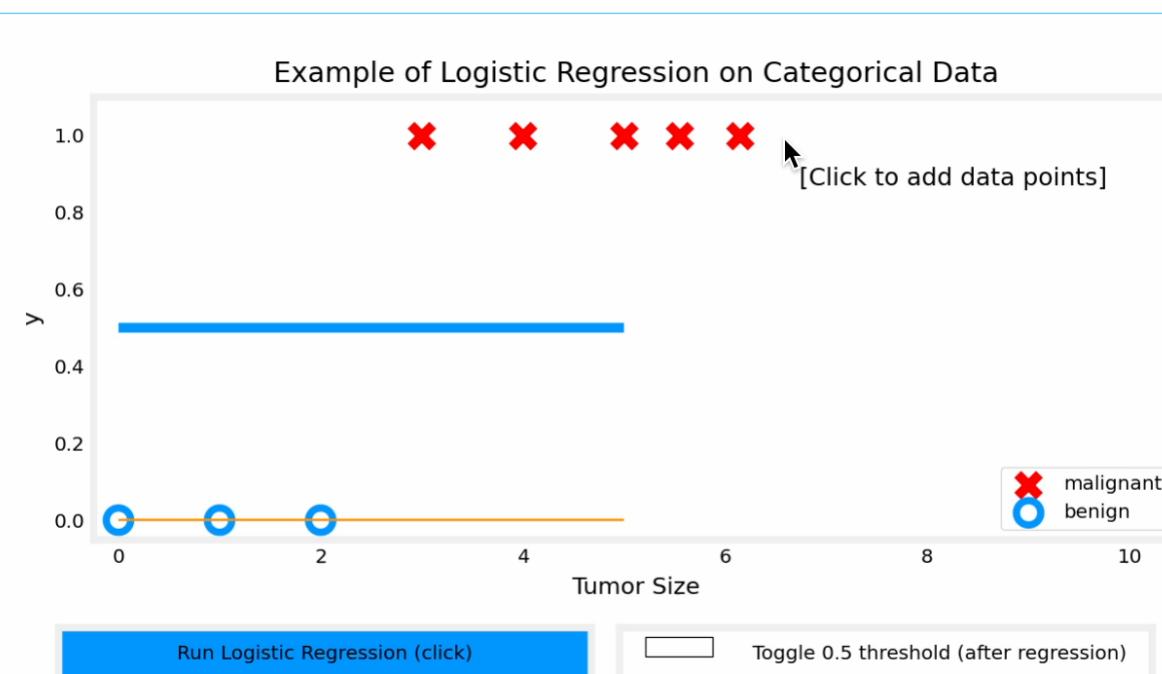


Funcion sigmoide

Regresión Logística



Se utiliza cuando la variable dependiente (objetivo) es categórica.



Regresión Logística



Para Regresion Lineal

$$y_{pred} = f(\vec{x}) = \vec{w}\vec{x} + b$$

Para Regresion Logistica

$$z = \vec{w}\vec{x} + b$$

$$\sigma(z) = \frac{1}{1 + e^{-z}}$$

$$f(\vec{x}) = \sigma(z) = \sigma(\vec{w}\vec{x} + b) = \frac{1}{1 + e^{-(\vec{w}\vec{x} + b)}}$$

Regresión Logística



Como predecir $y_{pred} = 0$ o $y_{pred} = 1$?

$$f(\vec{x}) = \sigma(z) = \sigma(\vec{w}\vec{x} + b) = \frac{1}{1 + e^{-(\vec{w}\vec{x}+b)}}$$

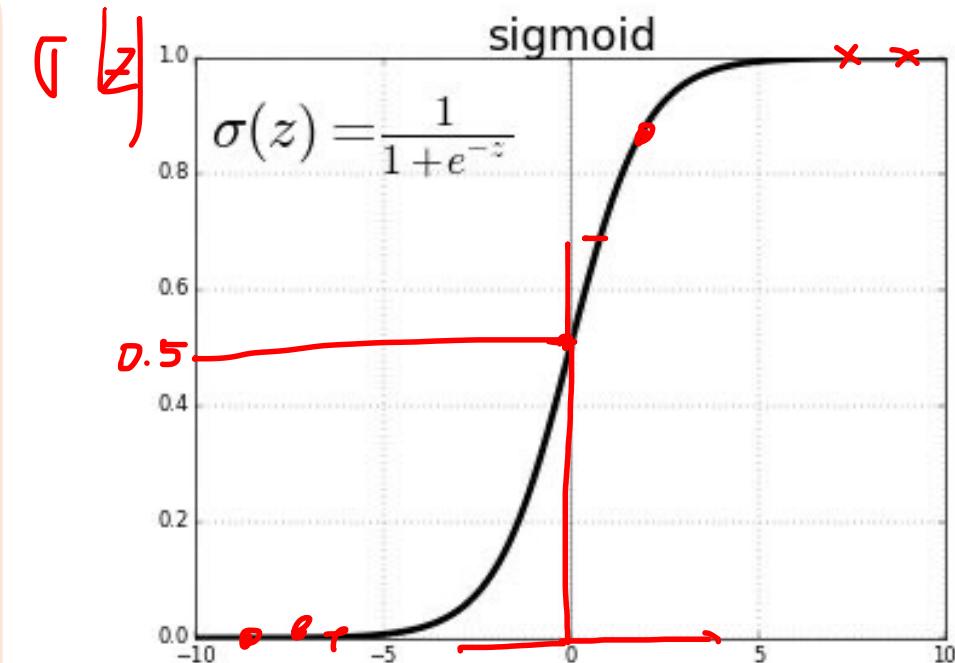
Si

$$y_{pred} = 1 ; f(\vec{x}) \geq 0.5 \text{ (threshold)}$$

$$\sigma(z) \geq 0.5$$

$$z \geq 0$$

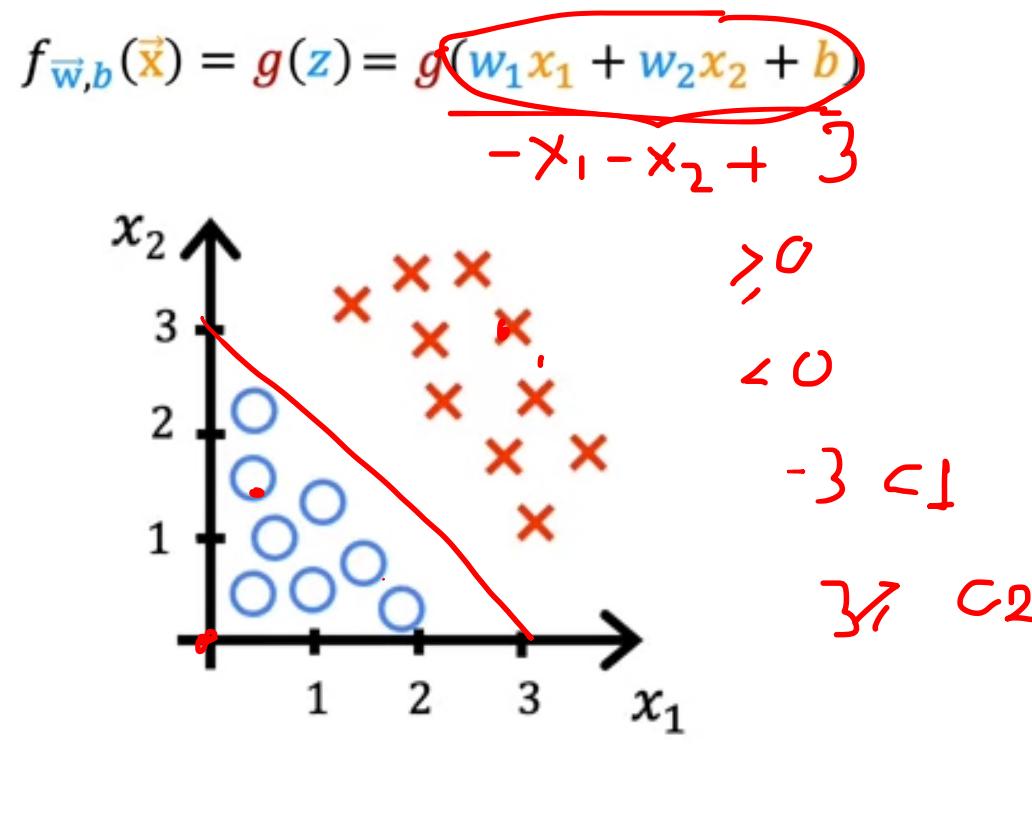
$$\begin{aligned} z &= \underbrace{w_x x + b}_{> 0} \\ &< 0 \end{aligned}$$



z

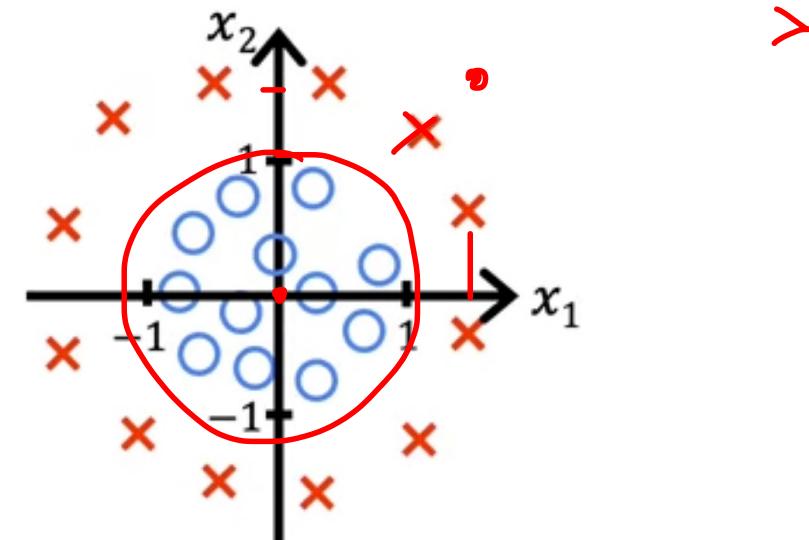
Regresión Logística

Decision Boundary



$$f_{\vec{w}, b}(\vec{x}) = g(z) = g(w_1 x_1^2 + w_2 x_2^2 + b)$$

$x_1^2 + x_2^2 + 1$



Regresión Logística

Función Costo



Para Regresion Lineal

$$J(w) = \frac{1}{2m} \sum_{i=1}^m (y_{pred}^i - y^i)^2$$



Para Regresion Logistica

$$J(w) = \frac{-1}{m} \sum_{i=1}^m (y^i \log(y_{pred}^i) + (1 - y^i) \log(1 - y_{pred}^i))$$

$y = 1$
 $y = 0$
 $y = \textcircled{O}$

Regresión Logística

Gradient Descent

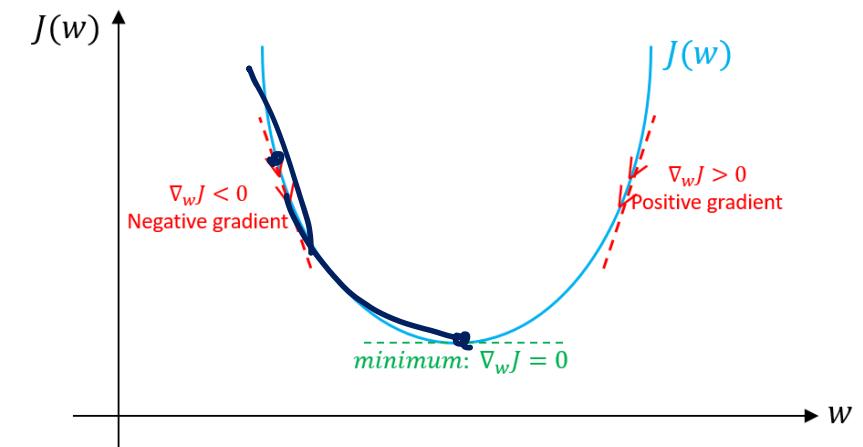


$$J(w) = \frac{-1}{m} \sum_{i=1}^m (y^i \log(y_{pred}^i) + (1 - y^i) \log(1 - y_{pred}^i))$$

Repetir

$$w := w - \alpha \frac{\partial}{\partial w} J(w, b)$$

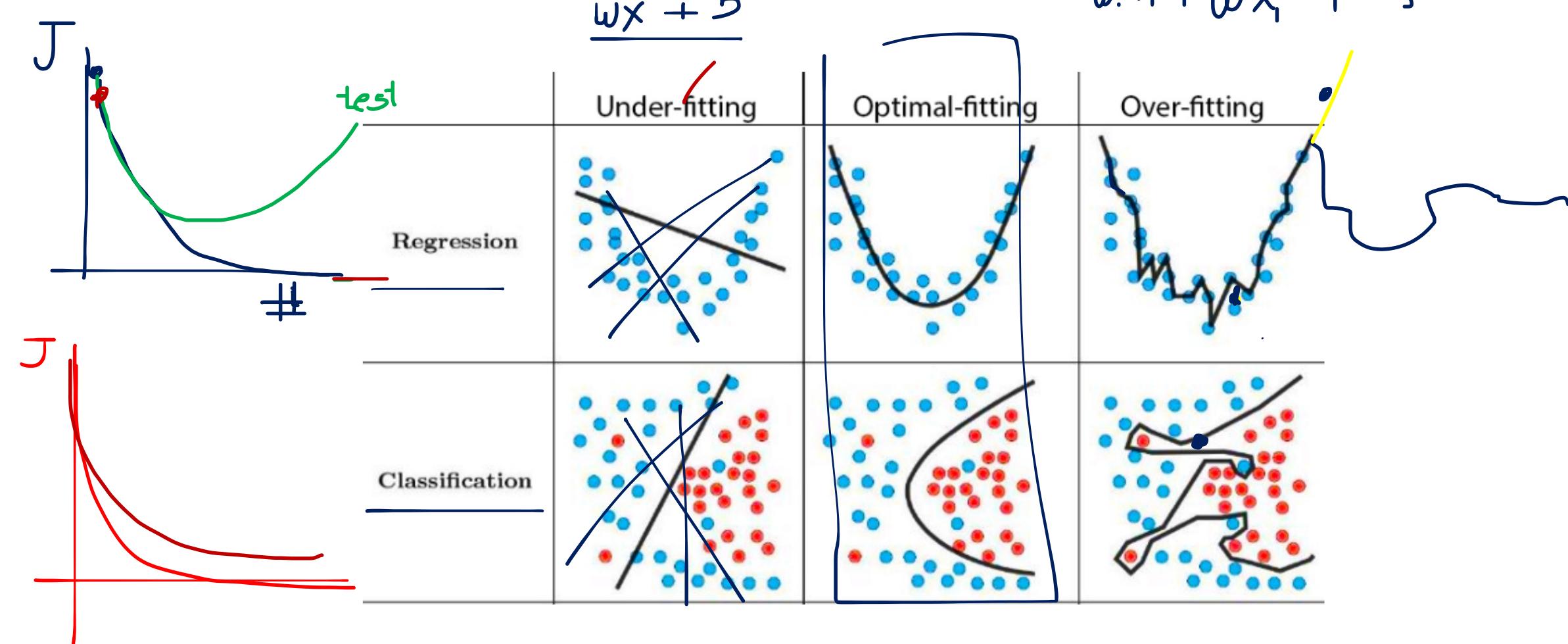
$$b := b - \alpha \frac{\partial}{\partial b} J(w, b)$$



El gradiente apunta en la dirección del ascenso más empinado

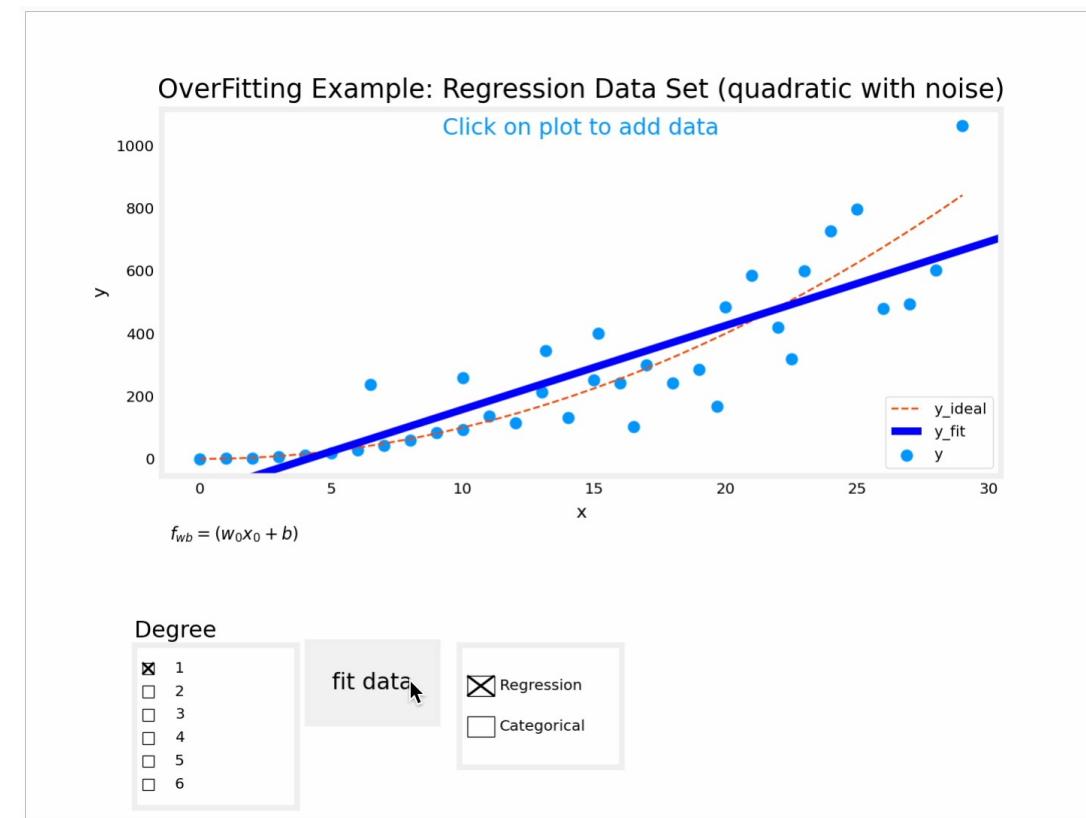
Regresión Logística

Overfitting y Underfitting

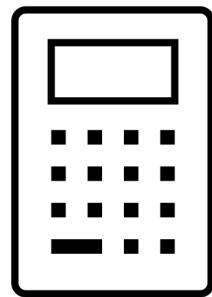


Regresión Logística

Overfitting y Underfitting



Formulario de la Sesión



<https://forms.gle/GXnX4WpwAvBgcegi9>

Link Google Colab de la Sesión



<https://colab.research.google.com/drive/1b0CNzWJ01Y2xX3LqRdbTpq5CCXRbKmKM?usp=sharing>



<https://colab.research.google.com/drive/1NyGnyHuDUBnryQA6ppLh-TbzL631sorN?usp=sharing>