

DIPLOMADO EN CIENCIA DE DATOS

MODELACIÓN SUPERVISADA SEMANA 5

Facultad de Estudios Superiores Acatlán

OUTLINE

Módulo 2 - Semana 5

- 1.- Bosque Aleatorio**
- 2.- Redes Neuronales**

QUIZ 4 (ÚLTIMO DEL MÓDULO!!)



<https://b.socrative.com/login/student/>
Room name: IRENE2290

1.- Bosque Aleatorio

BOSQUE ALEATORIOS

CONCEPTOS BÁSICOS



Este algoritmo es un ensamble de árboles de decisión individuales



Definición

- Target numérica/catógorica → Regresión/Clasificación
- Un set de datos está compuesto por el par (x_n, y_n)
- Es un ensamble de árboles de decisión, generalmente entrenado con el método Bagging



Objetivo: Obtener un modelo de árboles muy diverso

Características

- Introduce más aleatoriedad al crecer cada árbol que lo compone
- Lo anterior resulta en un modelo de árboles de decisión muy diverso
- En vez de partir un nodo con base en la mejor variable, buscar seleccionar la mejor variable en un subconjunto aleatorio de variables
- Este algoritmo busca tener un mayor sesgo, a cambio de una menor varianza
- Normalmente un ensamble de k árboles de decisión logra buenas métricas para el problema global
- Debido a la forma de entrenamiento del modelo, esta técnica permite conocer la importancia que cada variable aporta al mismo

BOSQUE ALEATORIOS

INTUICIÓN



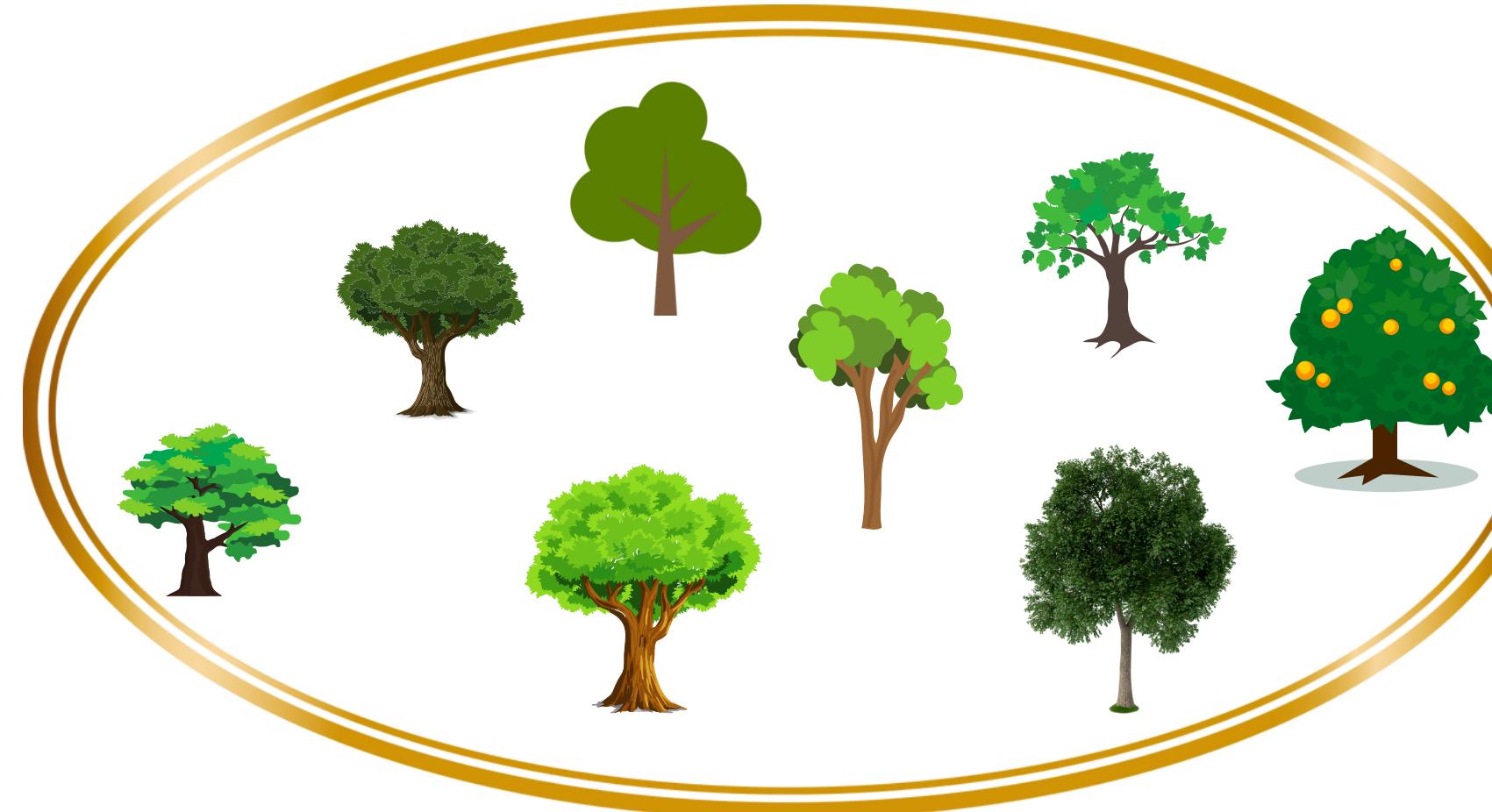
Este algoritmo es un ensamble de árboles de decisión individuales, generalmente utiliza el método de Bagging



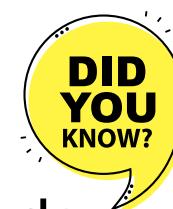
Bosque aleatorio

- Este modelo hace el ensamble de k árboles de decisión para realizar una predicción

Utiliza la sabiduría
de la multitud
para tomar una
decisión



Este modelo es la
versión optimizada de
un ensamble de árboles
de decisión con método
bagging

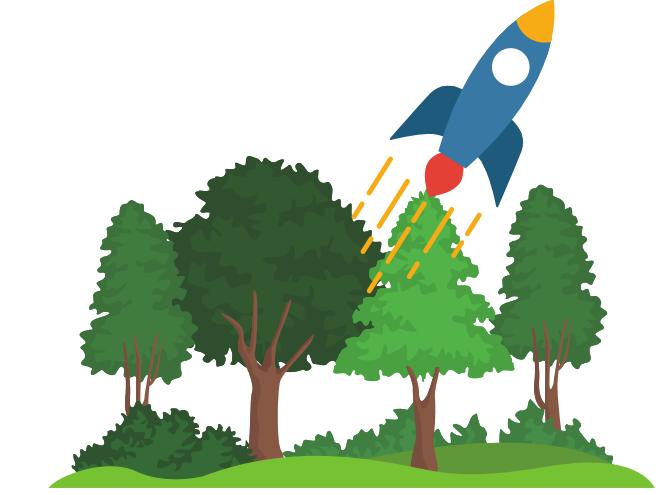


BOSQUE ALEATORIOS

EXTRA TREES



Es una versión de los bosques aleatorios que introduce aún más aleatoriedad



Árboles extras

- También llamado **Extremely Randomized Trees**
- Es una versión que introduce más componentes aleatorios al modelo de bosque aleatorio
- Cuando se entrena un modelo de bosque aleatorio se selecciona un subconjunto de variables al azar para particionar un nodo
- En un modelo de árboles extras, se hace que los árboles de decisión sean más aleatorios, al usar thresholds aleatorios para cada variable que partitiona un nodo
- Esta técnica hace que el modelo tenga una métrica de sesgo mayor, pero una métrica de varianza menor, haciendo que el modelo generalice mejor
- Tienen un menor tiempo de ejecución que un modelo de bosque aleatorio, ya que encontrar el threshold óptimo que particione un nodo es de las tareas que le toman más tiempo al modelo de bosque aleatorio
- Es difícil decir qué modelo tendrá mejores métricas de decisión antes de ser entrenados. La única forma de saberlo es al entrenar ambos modelos utilizando validación cruzada.

BOSQUE ALEATORIOS

IMPORTANCIA DE VARIABLES



Debido a la manera en la que se lleva a cabo el entrenamiento de este tipo de modelos, es posible determinar la importancia que algunas variables aportan al mismo

Importancia de variables

- Es sencillo medir la importancia relativa de cada variable
- Dicha importancia se mide al ver cuánto los nodos de cada árbol de decisión utilizan una determinada variable para reducir la impureza en promedio.
- Sklearn calcula la métrica de importancia de manera automática para cada variable en el set de entrenamiento
- La suma de la importancia de cada variable es igual a 1
- Puedes acceder a dichas métricas utilizando la variable `feature_importances_` del objeto `RandomForestClassifier`

2.- Redes Neuronales

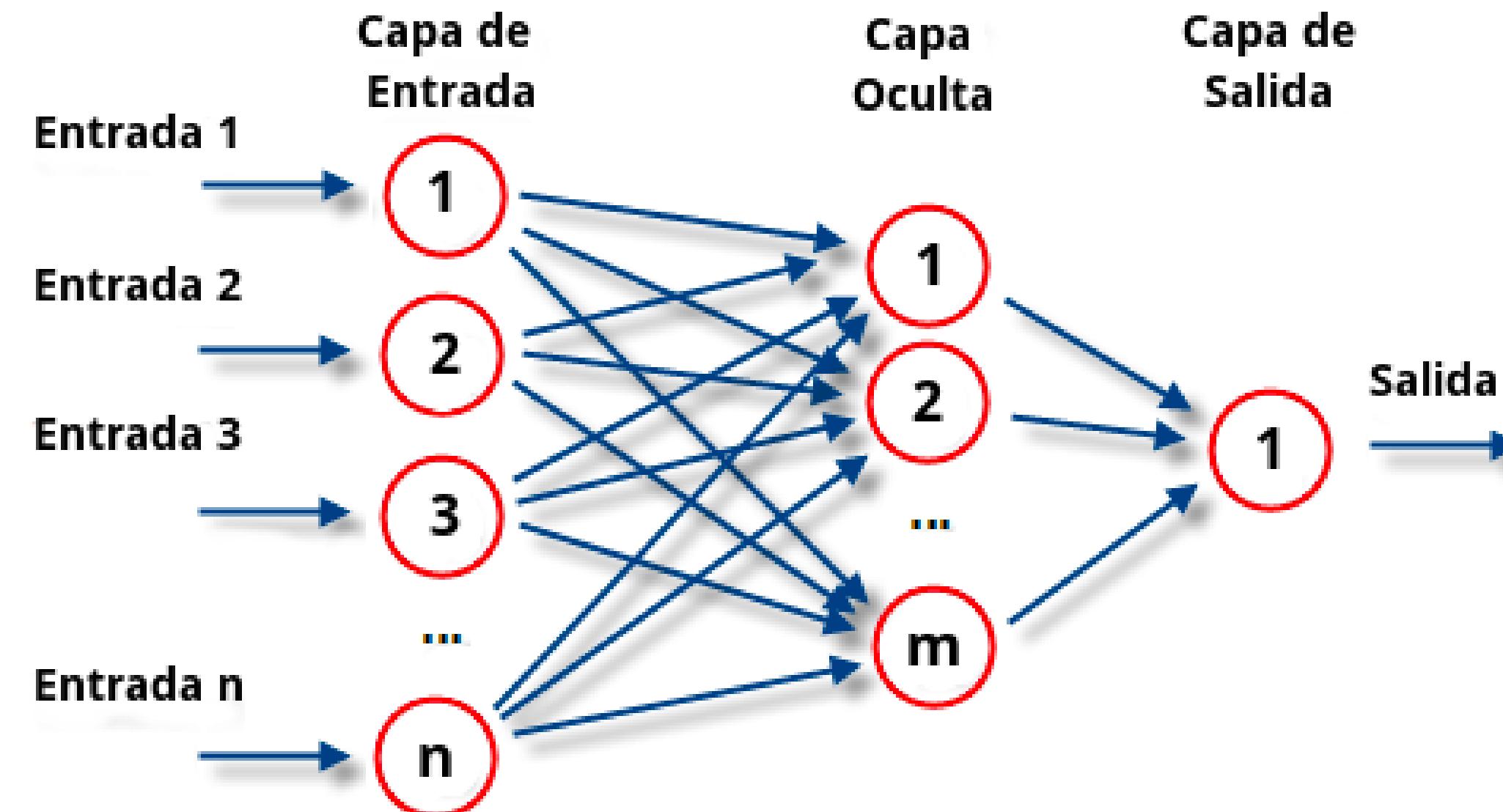
REDES NEURONALES

COMPONENTES



Una red neuronal con pocas capas ocultas se llaman redes poco profundas o "shallow"

Red neuronal poco profunda (shallow)



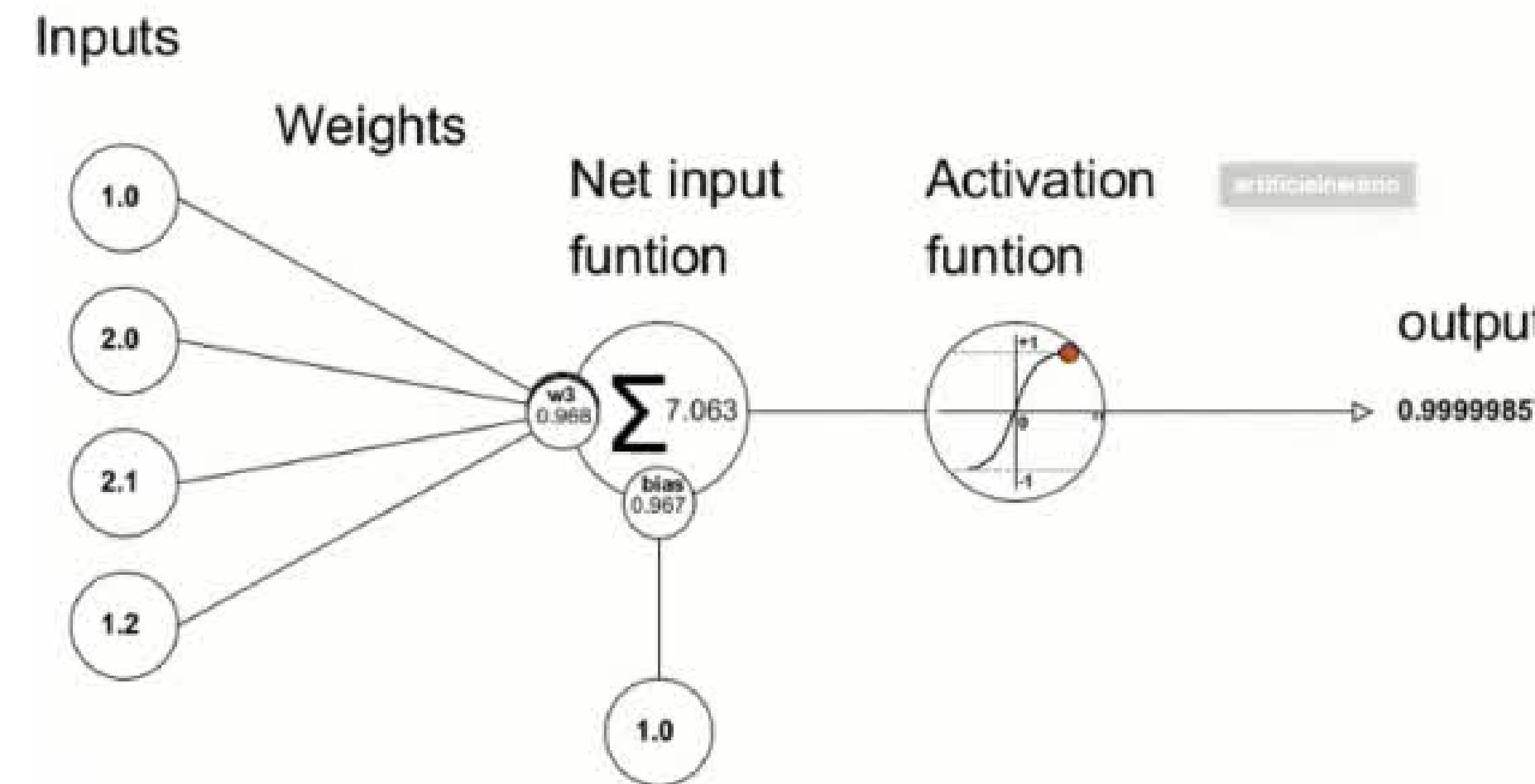
REDES NEURONALES

INTUICIÓN



Una red neuronal con pocas capas ocultas se llaman redes poco profundas o "shallow"

Red neuronal poco profunda (shallow)

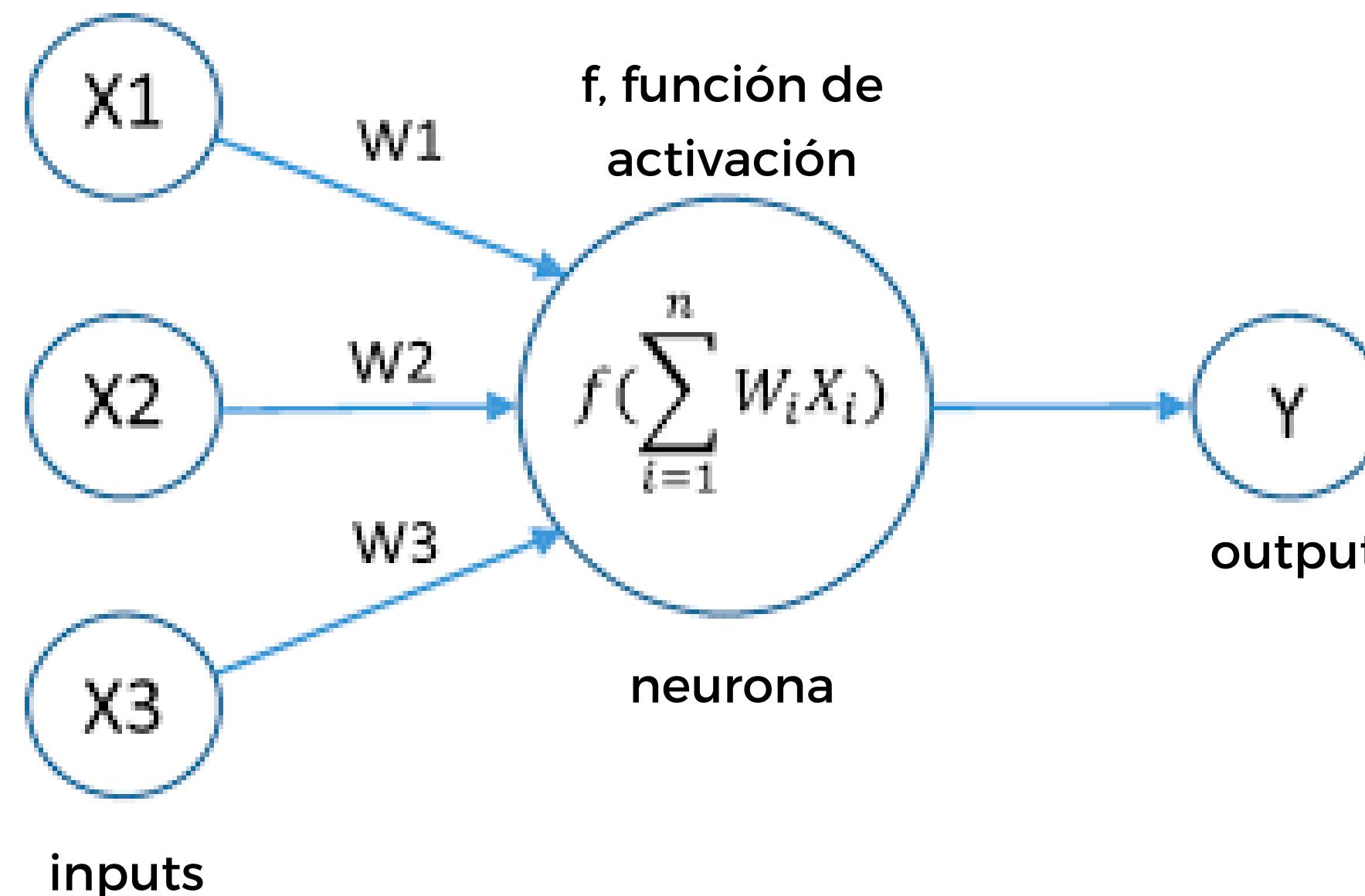


REDES NEURONALES

MODELO MATEMÁTICO

Las neuronas crean una combinación lineal con los inputs de la capa anterior y después aplican una función para convertir el resultado de lineal a no lineal.

Modelo en una neurona



REDES NEURONALES

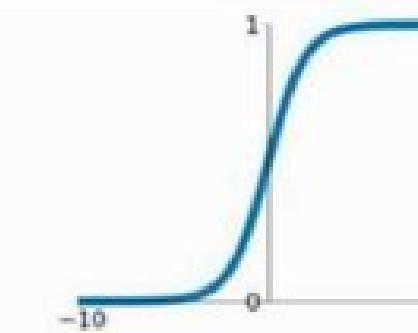
FUNCIONES DE ACTIVACIÓN

🎯 Las funciones de activación nos ayudan a transformar el output lineal de una neurona en un output no lineal. Lo anterior hace que en cada capa, la información sea procesada de mejor manera.

Funciones de activación populares

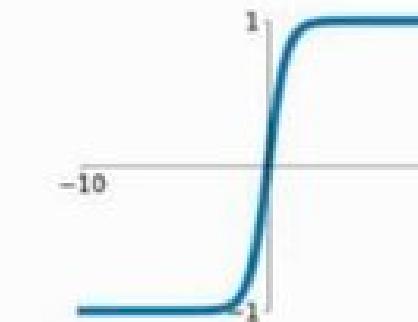
Sigmoid

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



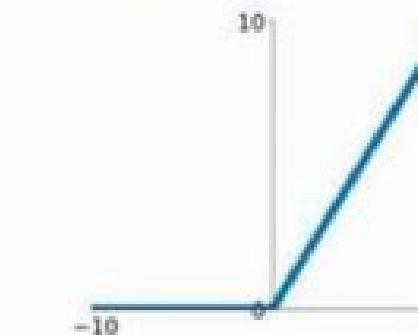
tanh

$$\tanh(x)$$



ReLU

$$\max(0, x)$$



Leaky ReLU

$$\max(0.1x, x)$$

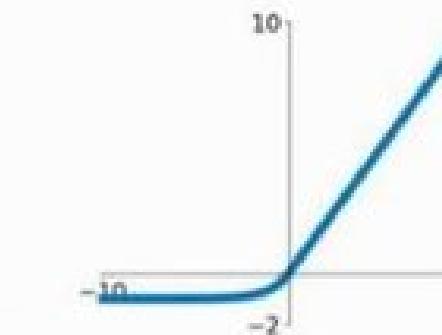


Maxout

$$\max(w_1^T x + b_1, w_2^T x + b_2)$$

ELU

$$\begin{cases} x & x \geq 0 \\ \alpha(e^x - 1) & x < 0 \end{cases}$$

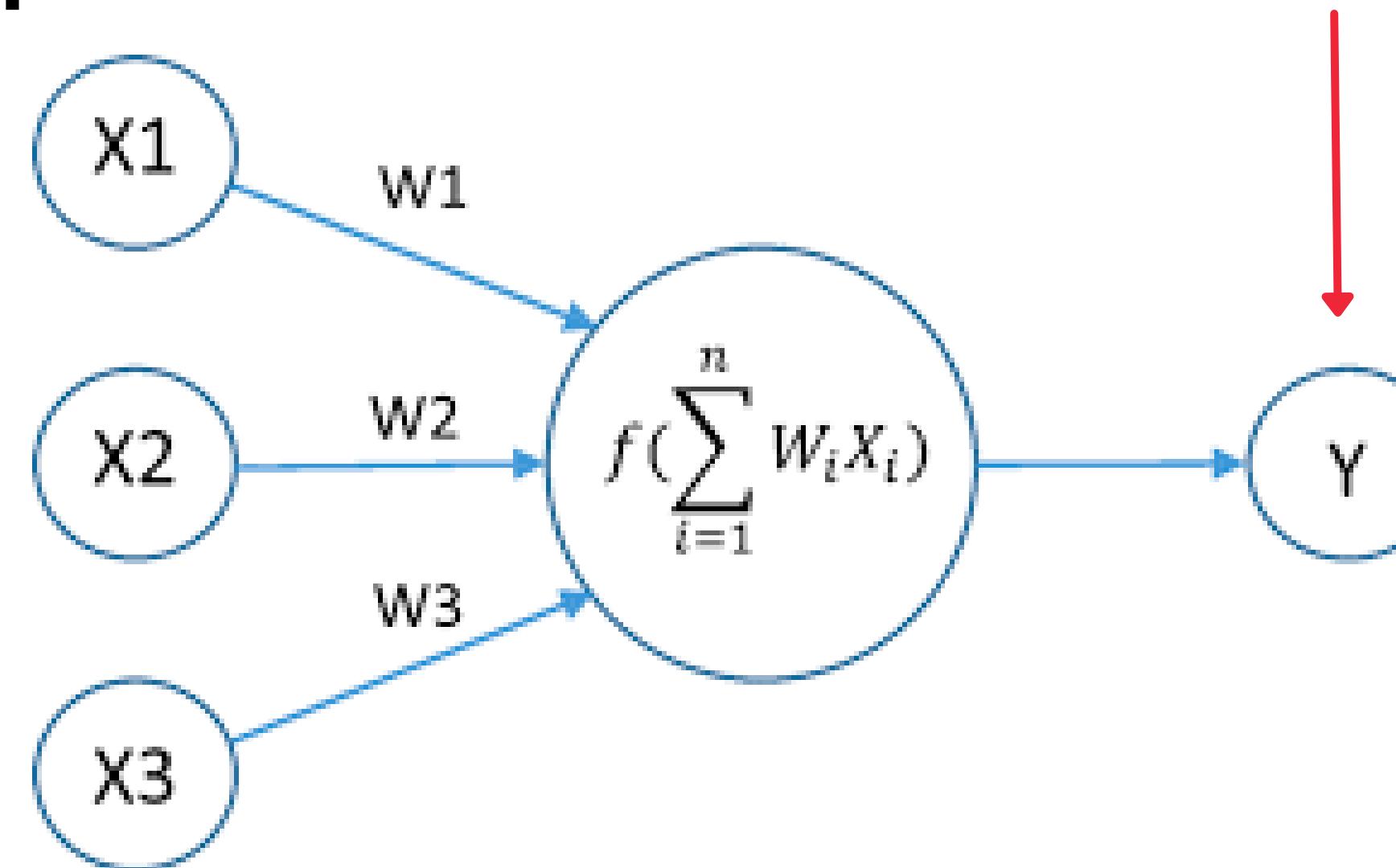


REDES NEURONALES

REGRESIÓN O CLASIFICACIÓN?

🎯 La función de activación en el nodo salida determina si el resultado de la red neuronal será numérico o categórico.

Si función lineal, entonces problema de regresión. Si sigmoide, tanh, etc, entonces problema de clasificación



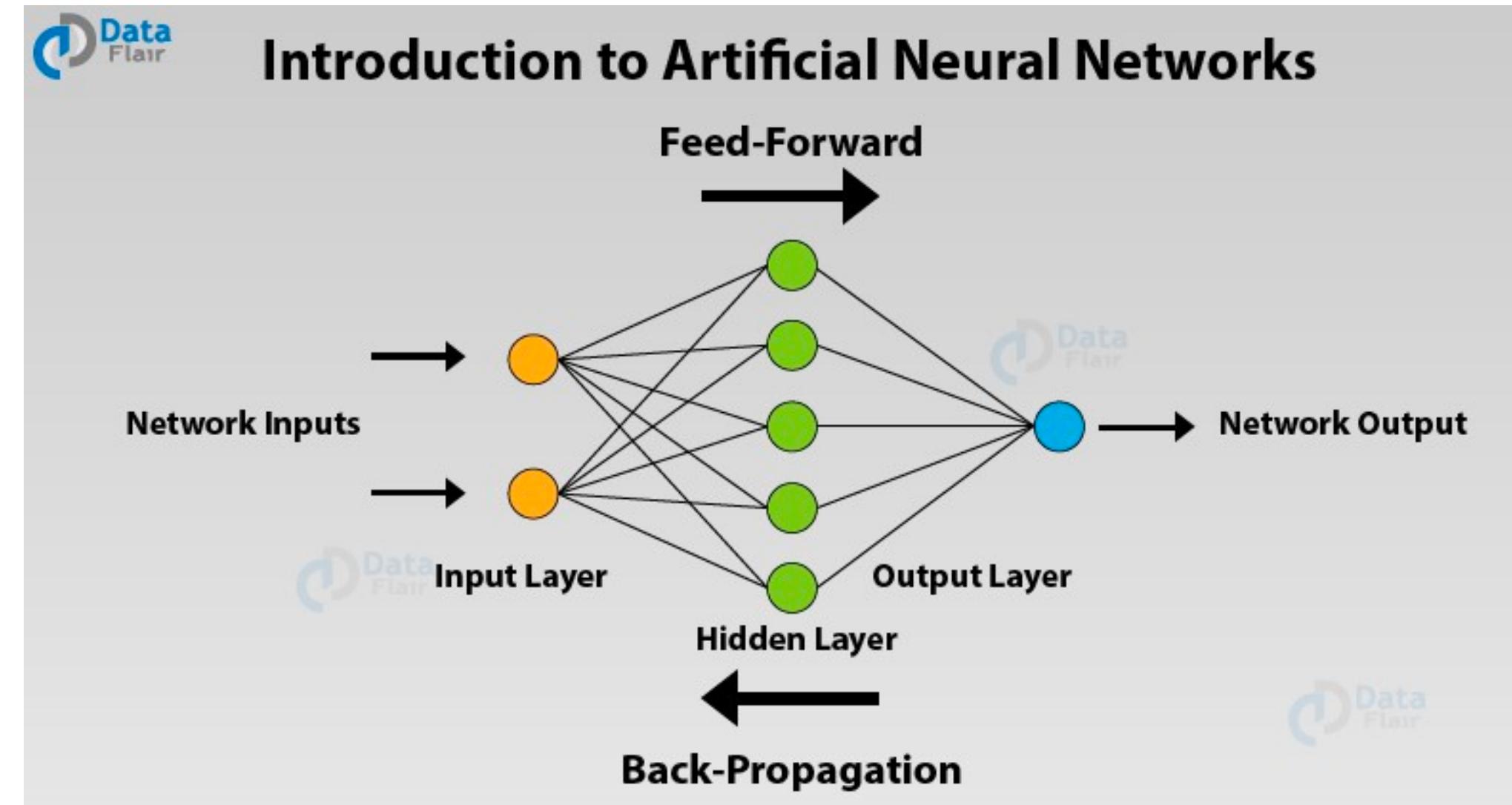
REDES NEURONALES

MÉTODO DE OPTIMIZACIÓN



La forma de optimizar parámetros de una red neuronal es utilizando el gradiente estocástico descendente y dos técnicas: Forward y Backpropagation.

Método consta de dos técnicas.



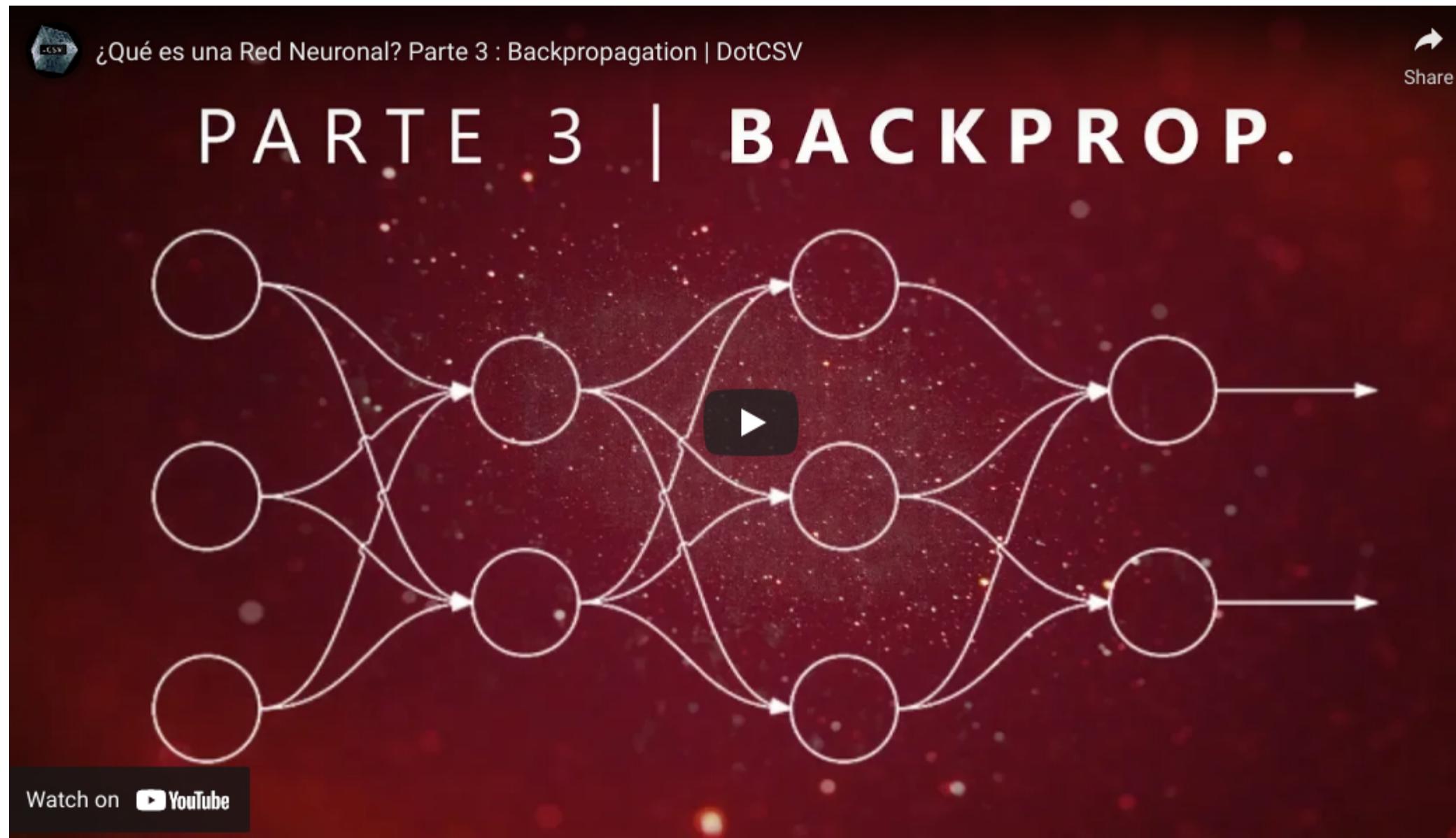
REDES NEURONALES

MÉTODO DE OPTIMIZACIÓN



Difícil eh... veamos un video para que quede más claro.

¿Qué es una Red Neuronal? Parte 3 : Backpropagation | DotCSV

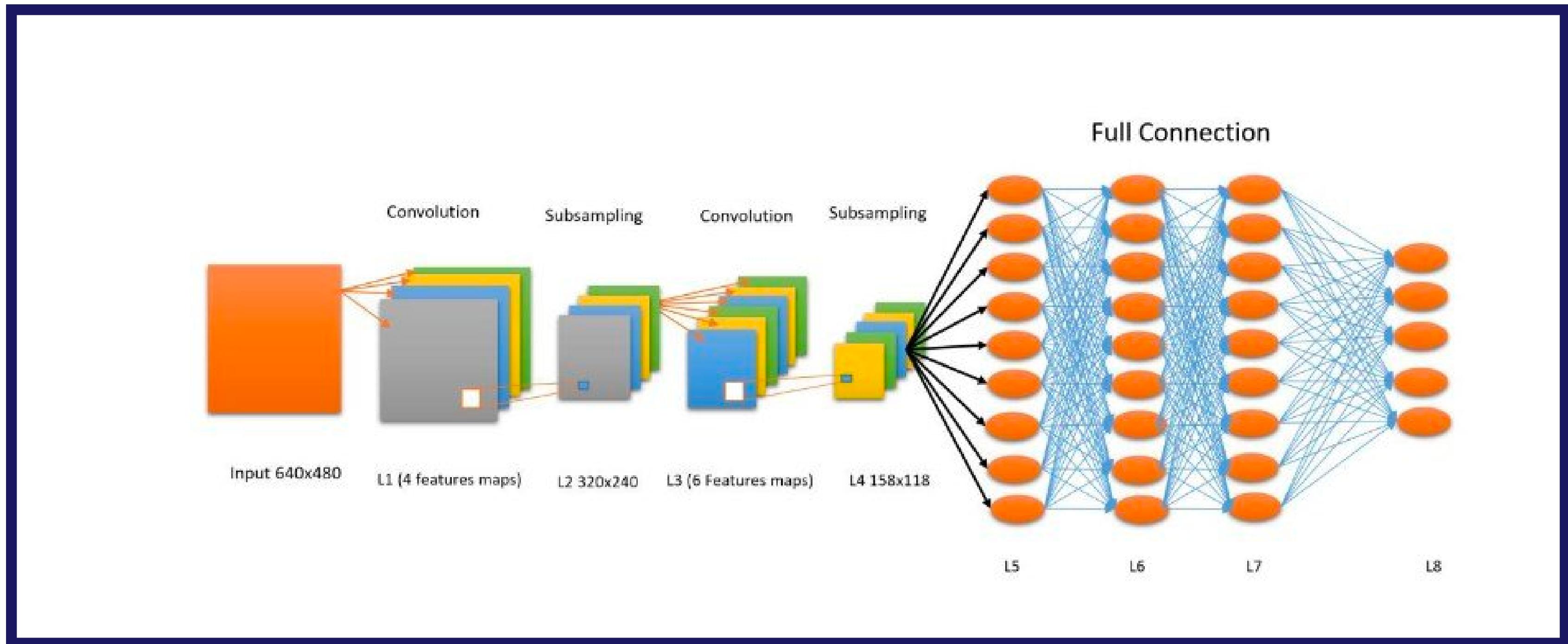


REDES NEURONALES CONVOLUCIONALES

INTUICIÓN



Las redes neuronales convolucionales crean filtros a los inputs. Posterior, son procesados por las capas de la red neuronal.



REDES NEURONALES

EJERCICIO



Ejercicio: Investiga casos de uso de empresas que usen las redes neuronales para resolver problemas de negocio.

Toma 30 minutos para investigar acerca de casos de uso que las empresas lleven a cabo con redes neuronales. Prepara un slide explicando un caso de uso y presentalo a tus compañeros.