

Taller de aprendizaje profundo aplicado en biomedicina

Raúl Fernández Díaz



Organización

Sesión 1: Fundamentos
de aprendizaje profundo

Sesión 2: Aplicación de
arquitecturas avanzadas
en biomedicina

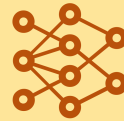
Sesión 3: Clasificación de
cánceres de piel



Sesión 1: Contenidos



¿Qué es el aprendizaje automático?



¿Cómo funcionan las redes neurales?



¿Qué problemas podemos encontrar?

Primera Parte

¿Qué es el aprendizaje automático?

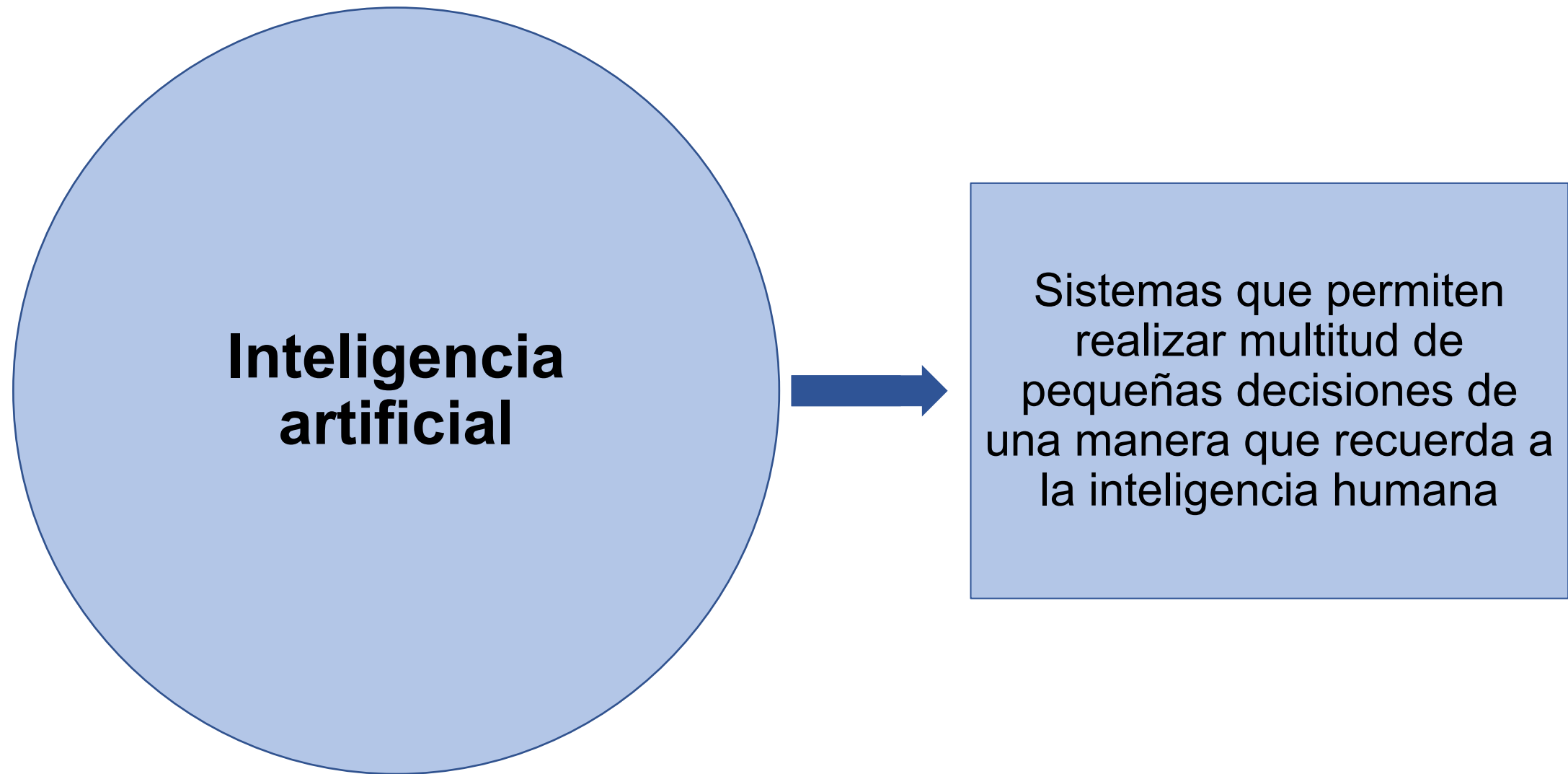
Inteligencia artificial: Máquinas que piensan



**Inteligencia
artificial**

Aprendizaje profundo

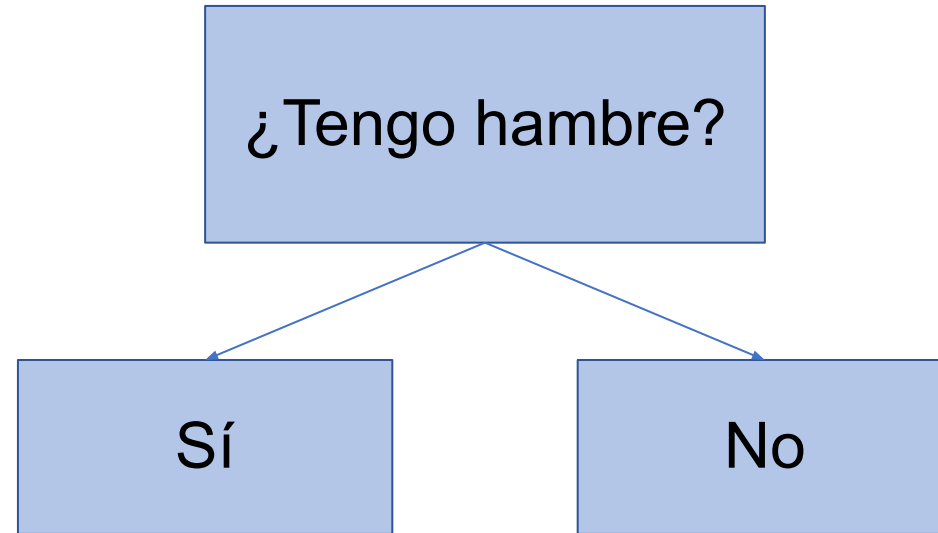
Inteligencia artificial: Máquinas que piensan



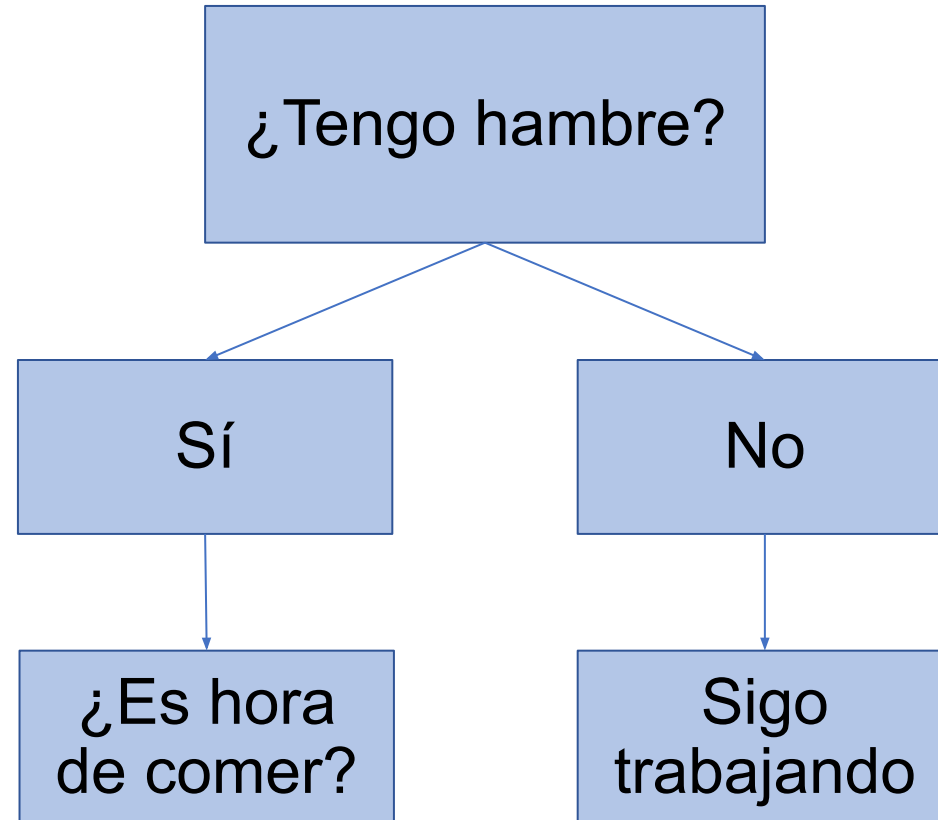
Inteligencia artificial: Decisiones

¿Tengo hambre?

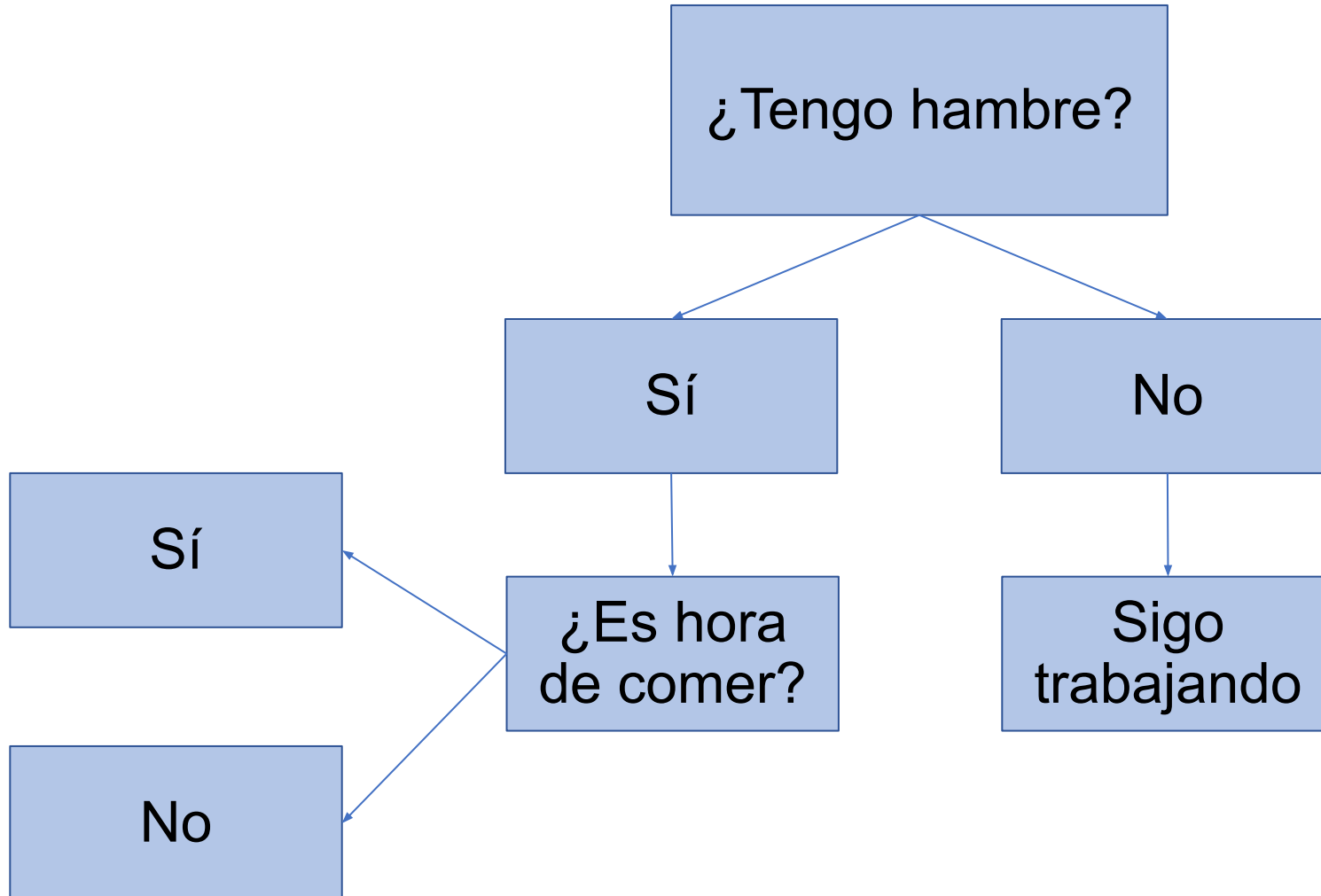
Inteligencia artificial: Decisiones



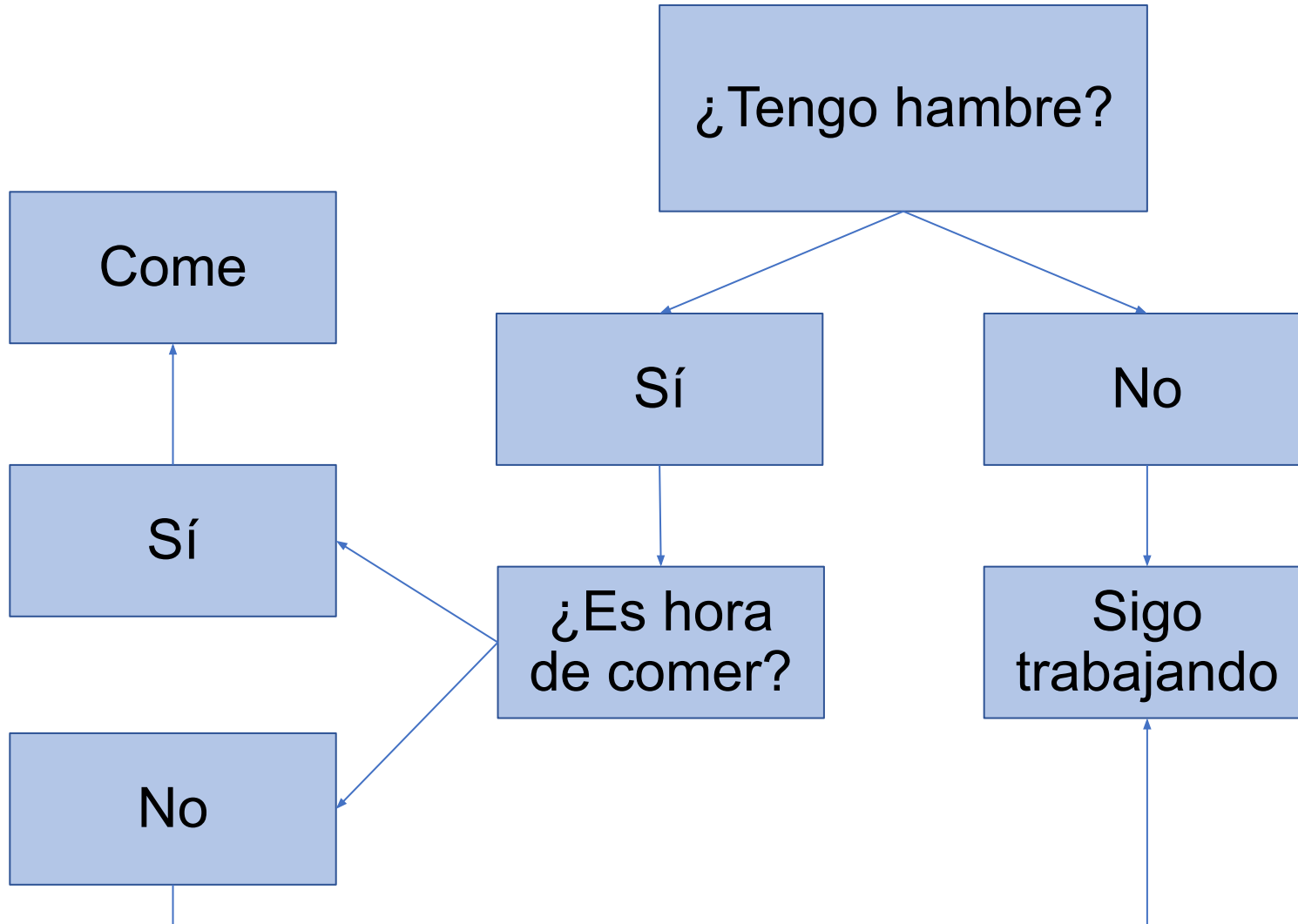
Inteligencia artificial: Decisiones



Inteligencia artificial: Decisiones



Inteligencia artificial: Decisiones



Inteligencia artificial: Decisiones



Aprendizaje profundo

Inteligencia artificial: Decisiones

Si Mario a la izqda:
Mueve a la izqda
Si no:
Mueve a la dcha



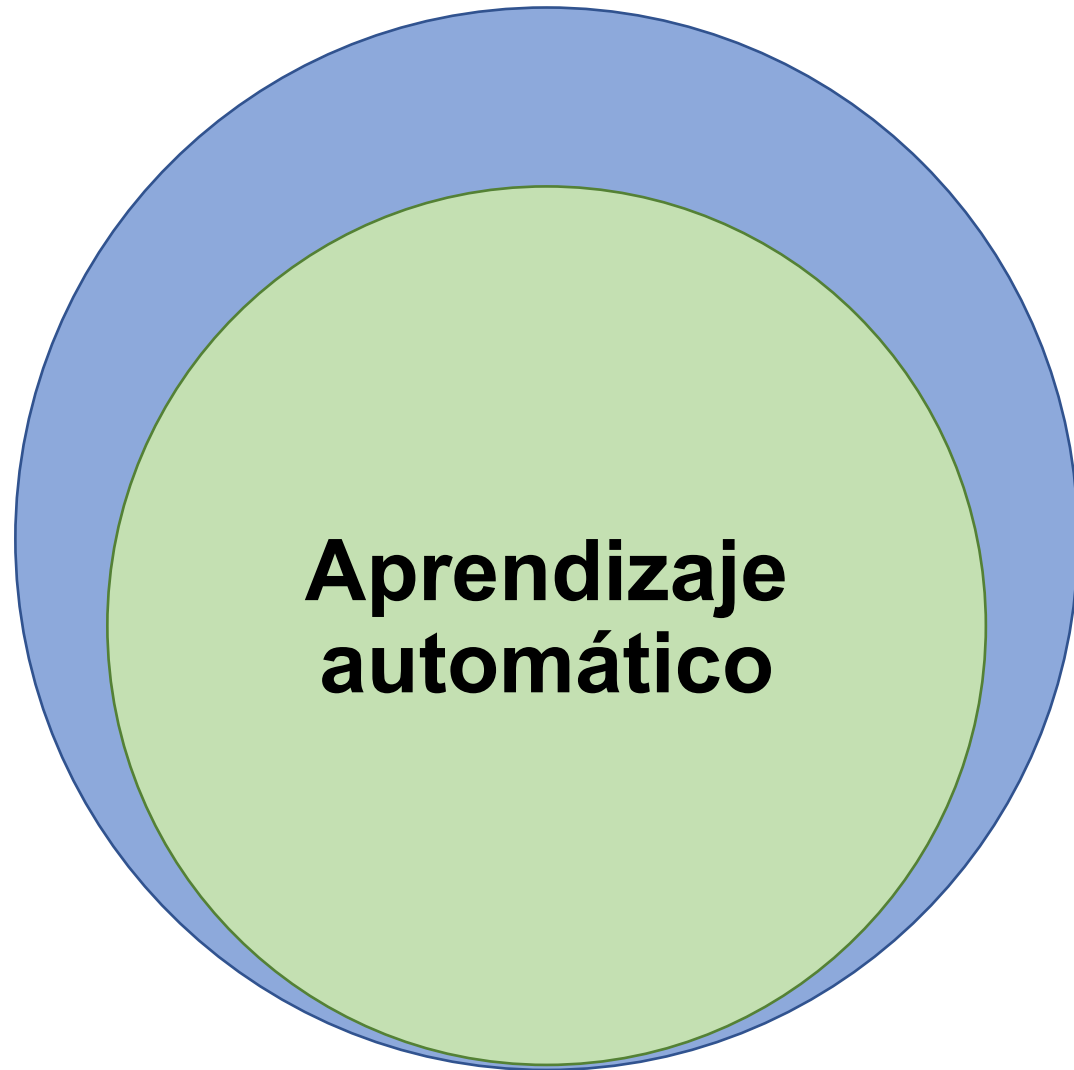
Aprendizaje profundo

Inteligencia artificial: Decisiones



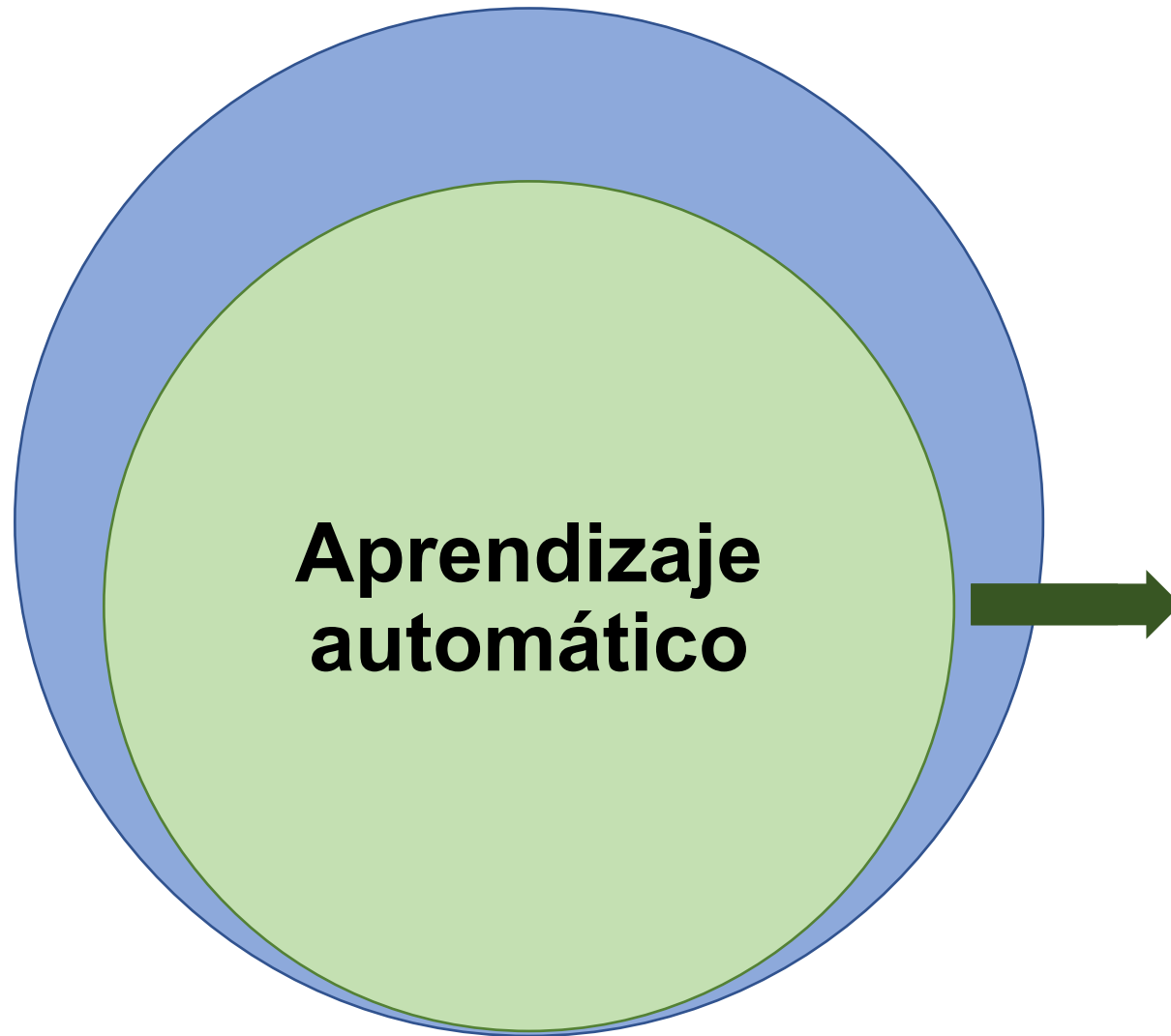
Aprendizaje profundo

Aprendizaje automático: Máquinas que aprenden



Aprendizaje profundo

Aprendizaje automático: Máquinas que aprenden



**Aprendizaje
automático**

El ordenador desarrolla sus propios criterios y estrategias de toma de decisiones que mejoran con la experiencia o el uso de datos

Aprendizaje profundo

Aprendizaje automático: Tipos de aprendizaje

Supervisado

No supervisado

Reforzado

Aprendizaje profundo

Aprendizaje automático: Tipos de aprendizaje



Aprendizaje profundo

Aprendizaje automático: Tipos de aprendizaje

Supervisado

Lista de
ejemplos con
sus soluciones



Aprendizaje profundo

Aprendizaje automático: Tipos de aprendizaje

Supervisado

Lista de
ejemplos con
sus soluciones

Perro



Perro



Pez



Pez



Aprendizaje profundo

Aprendizaje automático: Tipos de aprendizaje

No supervisado

Separa por su
cuenta

Aprendizaje profundo

Aprendizaje automático: Tipos de aprendizaje



No supervisado

Separa por su
cuenta



Modelo 1



Aprendizaje profundo

Aprendizaje automático: Tipos de aprendizaje



No supervisado

Separa por su
cuenta



Modelo 2



Aprendizaje profundo

Aprendizaje automático: Tipos de aprendizaje



No supervisado



Separa por su
cuenta

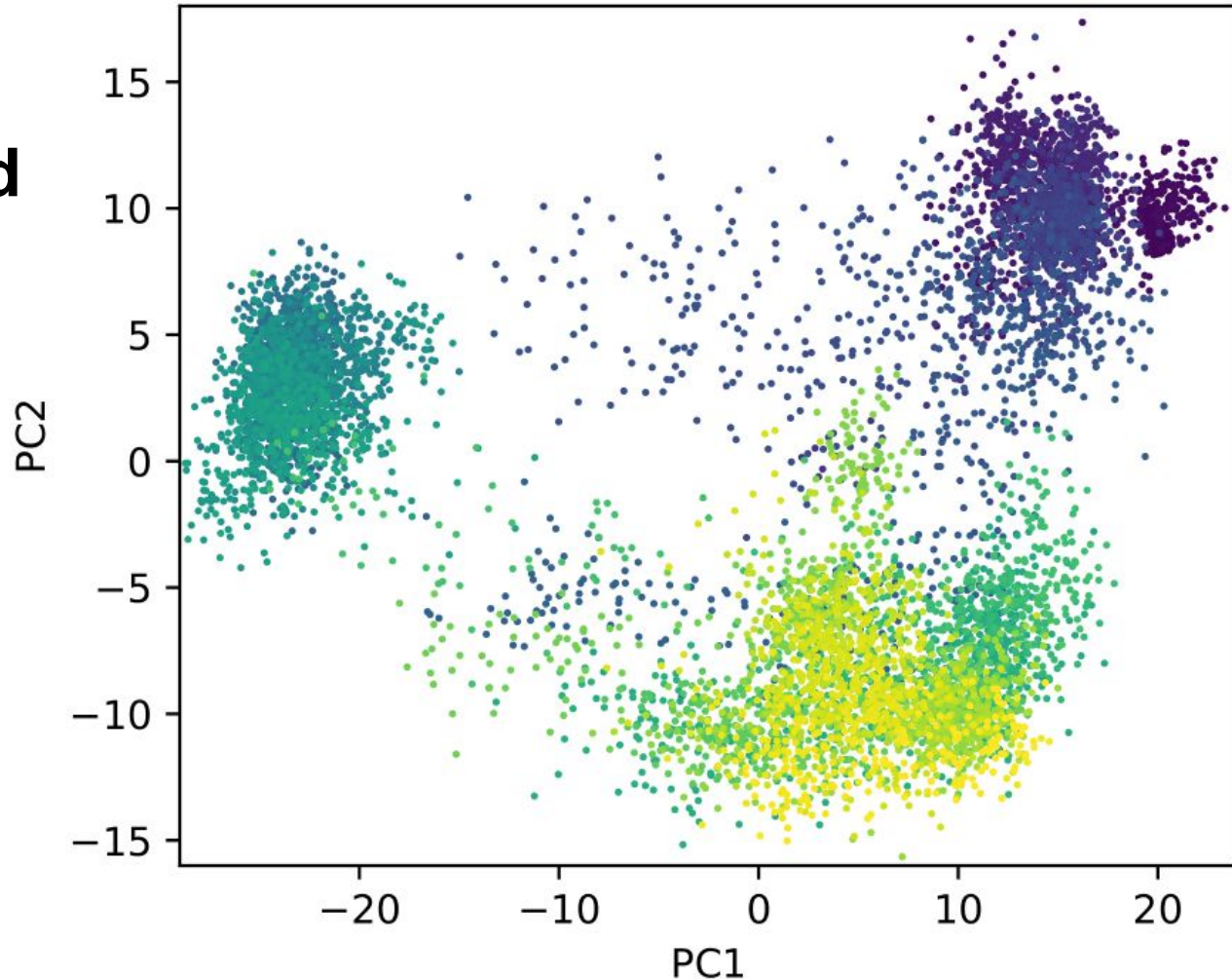
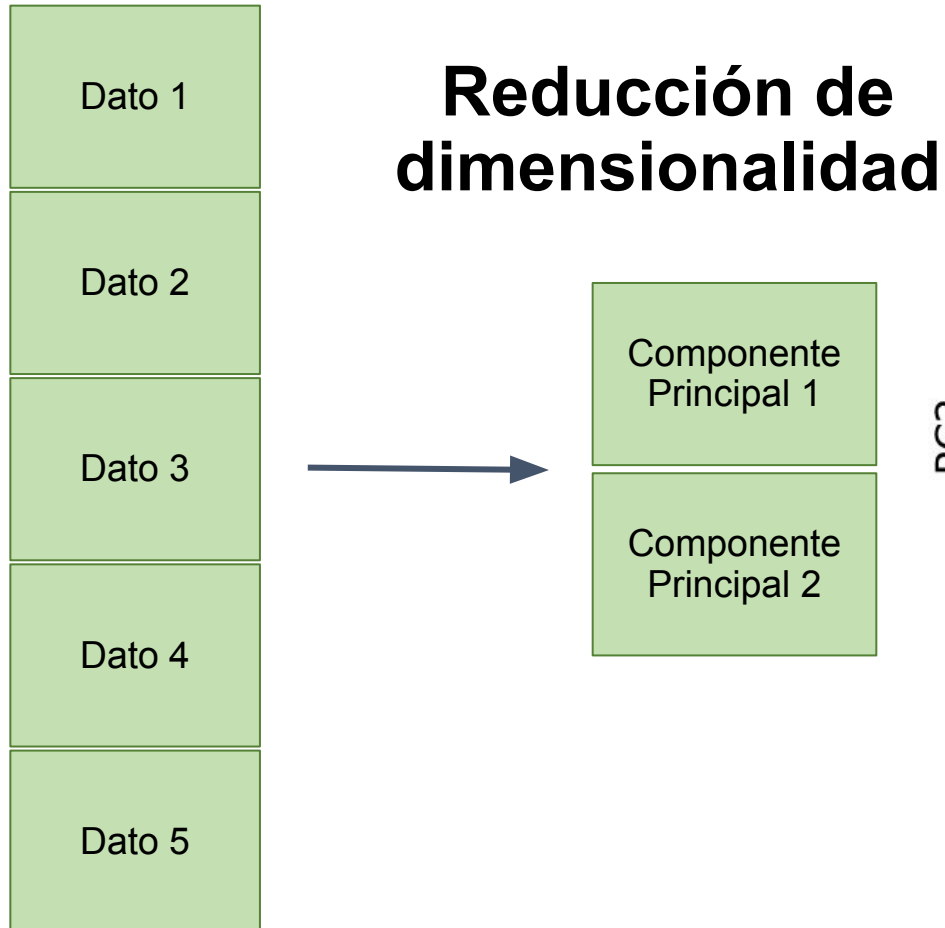


Modelo 3



Aprendizaje profundo

Aprendizaje automático: Tipos de aprendizaje



Aprendizaje profundo

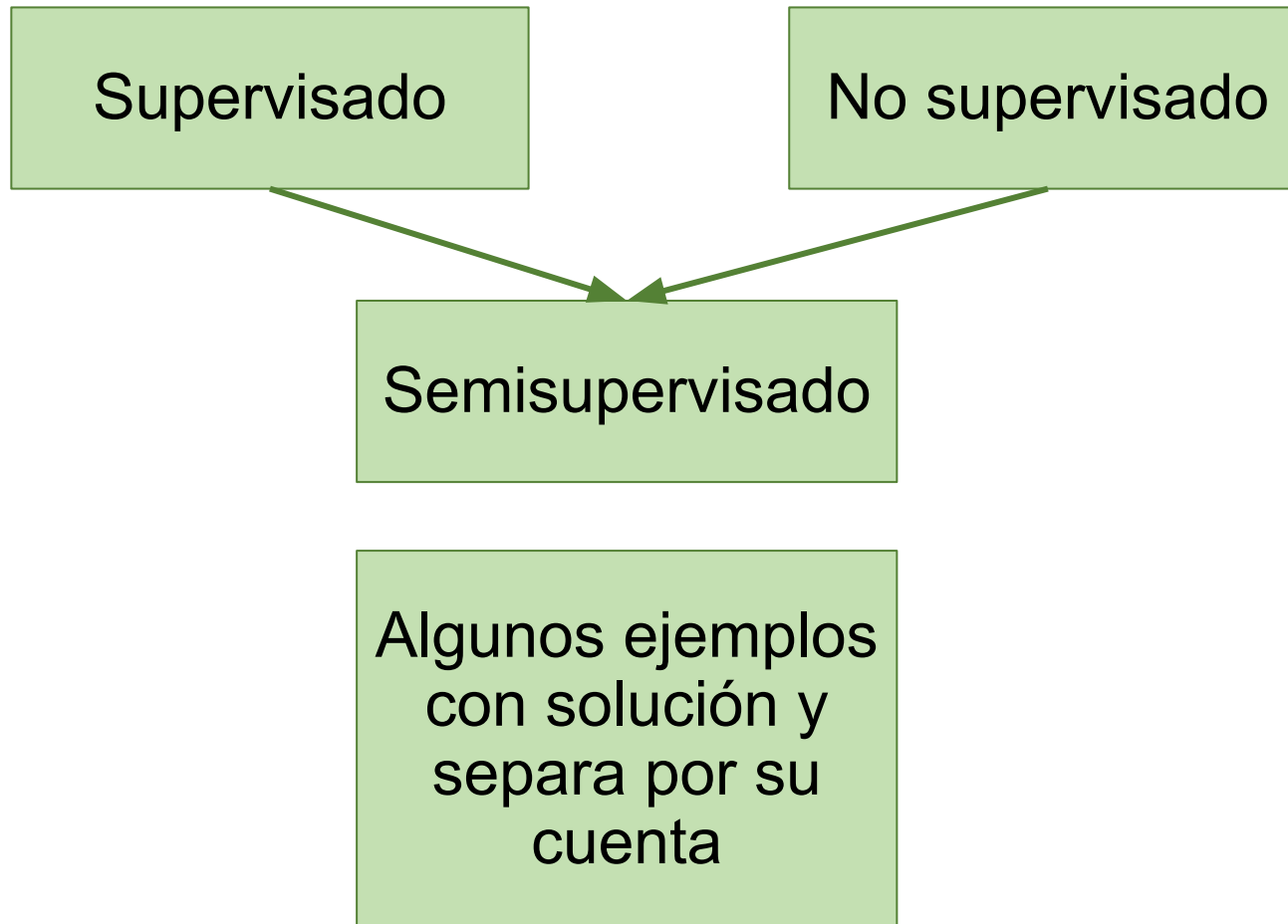
Aprendizaje automático: Tipos de aprendizaje

Supervisado

No supervisado

Aprendizaje profundo

Aprendizaje automático: Tipos de aprendizaje



Aprendizaje automático: Tipos de aprendizaje



Aprendizaje automático: Tipos de aprendizaje



Aprendizaje automático: Tipos de aprendizaje



Aprendizaje automático: Tipos de aprendizaje

Supervisado

No supervisado

Reforzado

Aprende los
pasos óptimos
para realizar una
tarea

Aprendizaje profundo

Aprendizaje automático: Tipos de aprendizaje

Intenta andar:

**Si caes:
Movimiento -10**

**Si no:
Movimiento +10**



Reforzado

**Aprende los
pasos óptimos
para realizar una
tarea**

Aprendizaje profundo

Aprendizaje automático: Tipos de aprendizaje

Supervisado

Lista de
ejemplos con
sus soluciones

No supervisado

Separa por su
cuenta

Reforzado

Aprende los
pasos óptimos
para realizar una
tarea

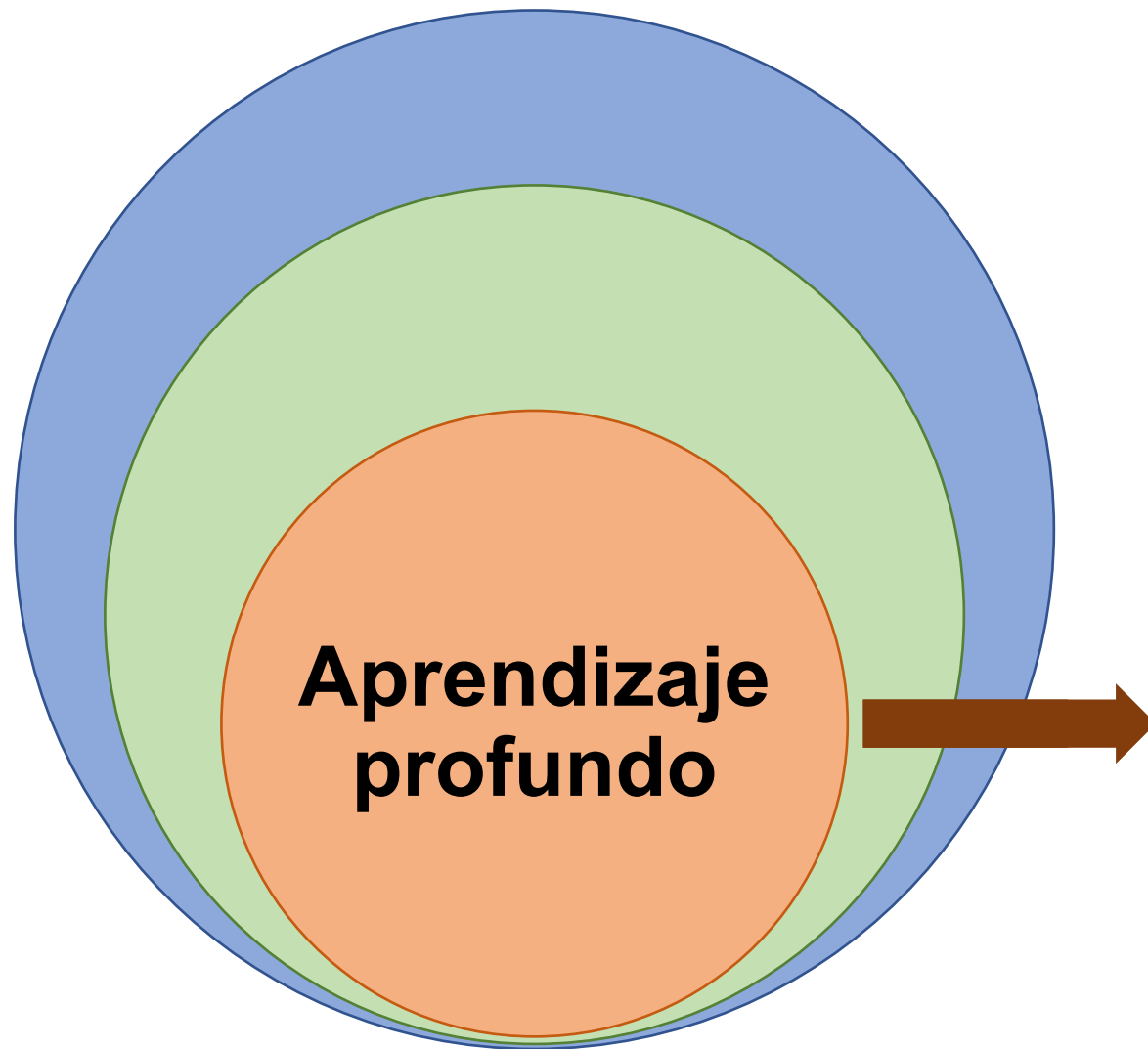
Aprendizaje profundo

Aprendizaje profundo: Cerebros de silicio



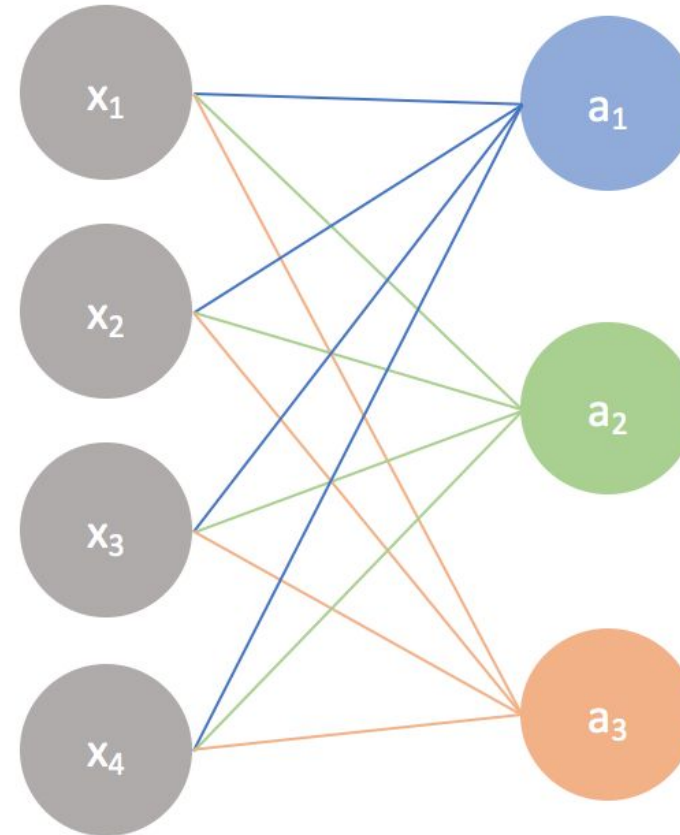
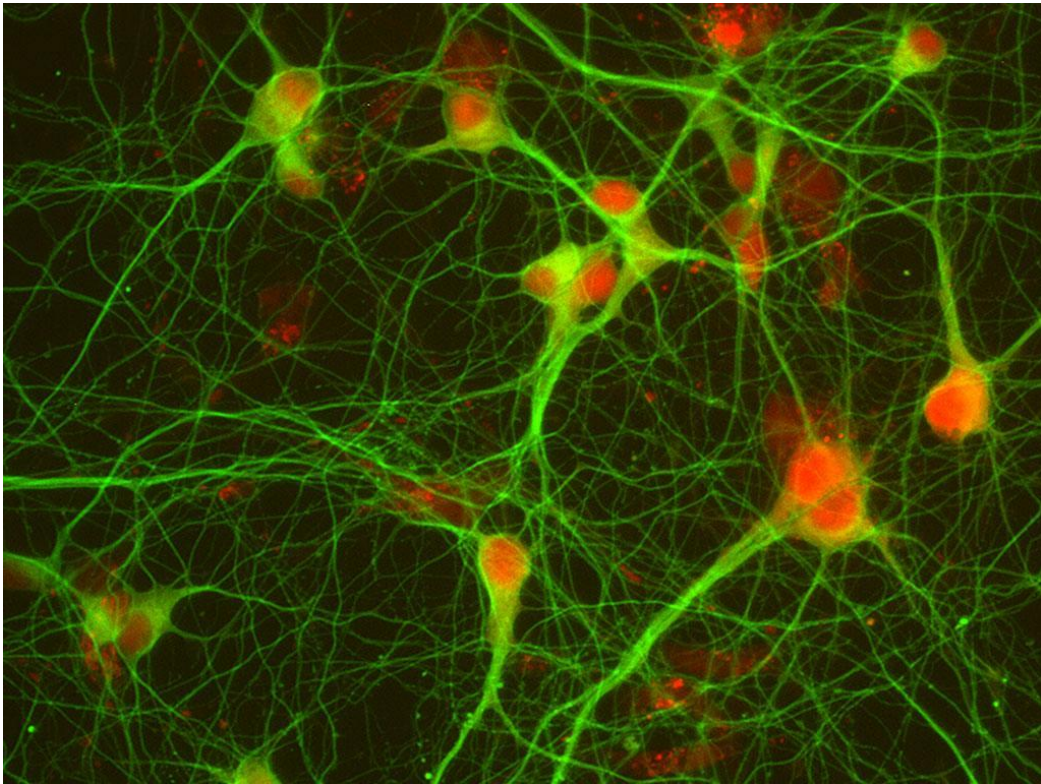
Aprendizaje profundo

Aprendizaje profundo: Cerebros de silicio



Técnica de aprendizaje automático inspirada en el funcionamiento de nuestro cerebro (redes neurales)

Aprendizaje profundo: Cerebros de silicio



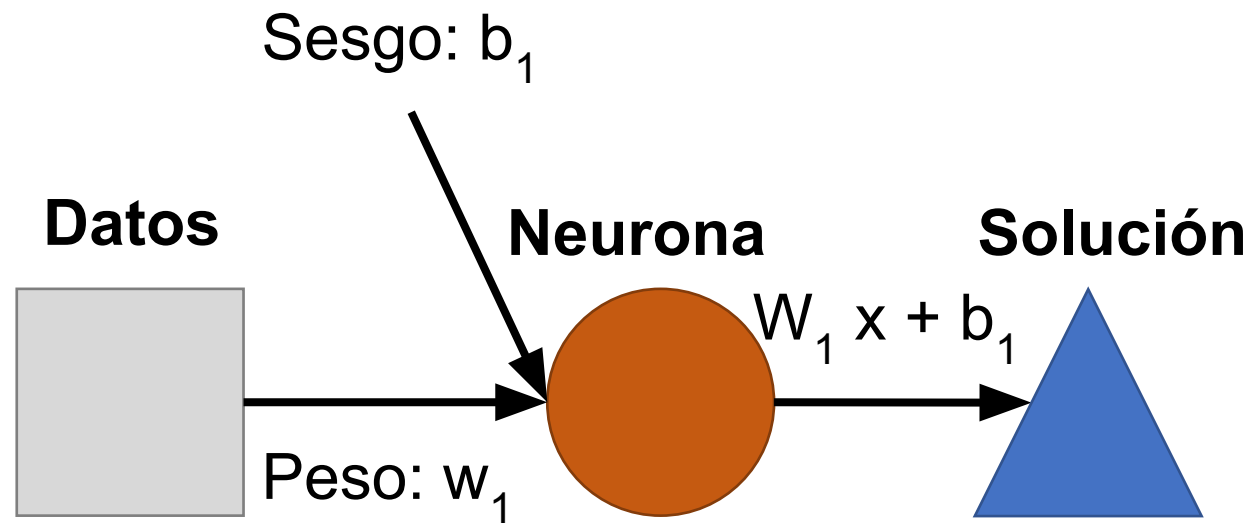
Aprendizaje profundo

Primera parada: ¿Preguntas hasta aquí?

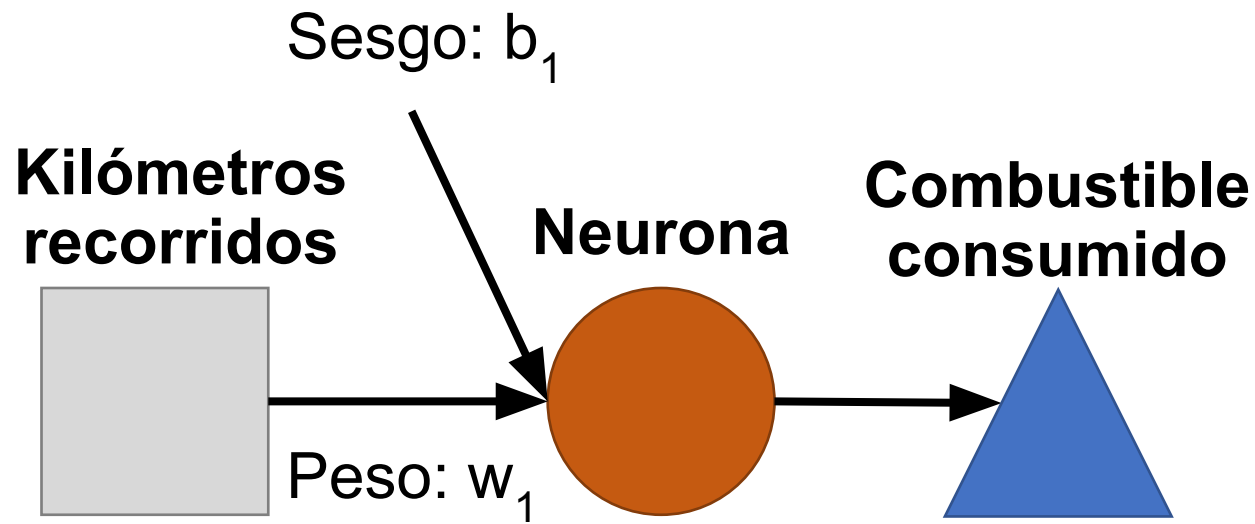
Segunda Parte

¿Cómo funcionan las redes neurales?

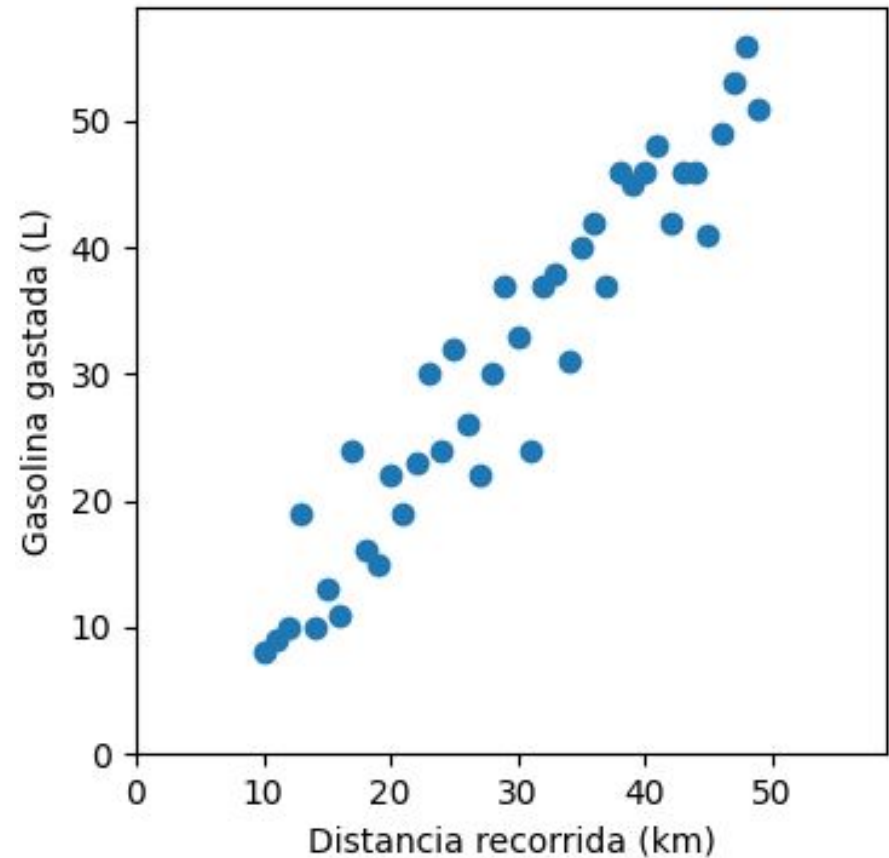
Redes neurales: caso más simple



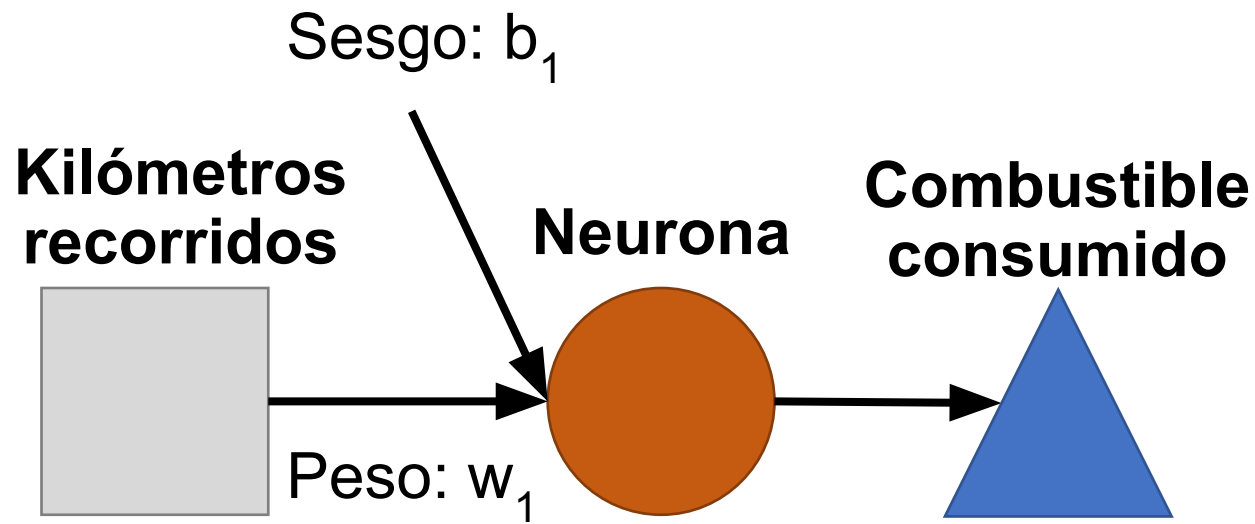
Redes neurales: caso más simple



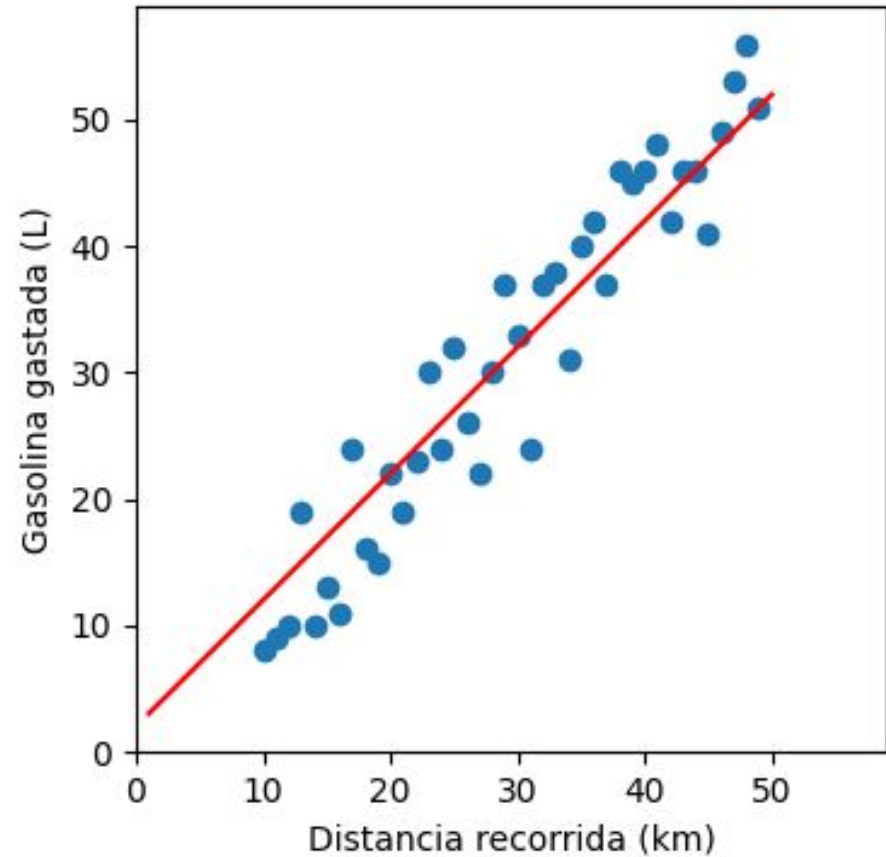
$$y = w_1 x + b_1$$



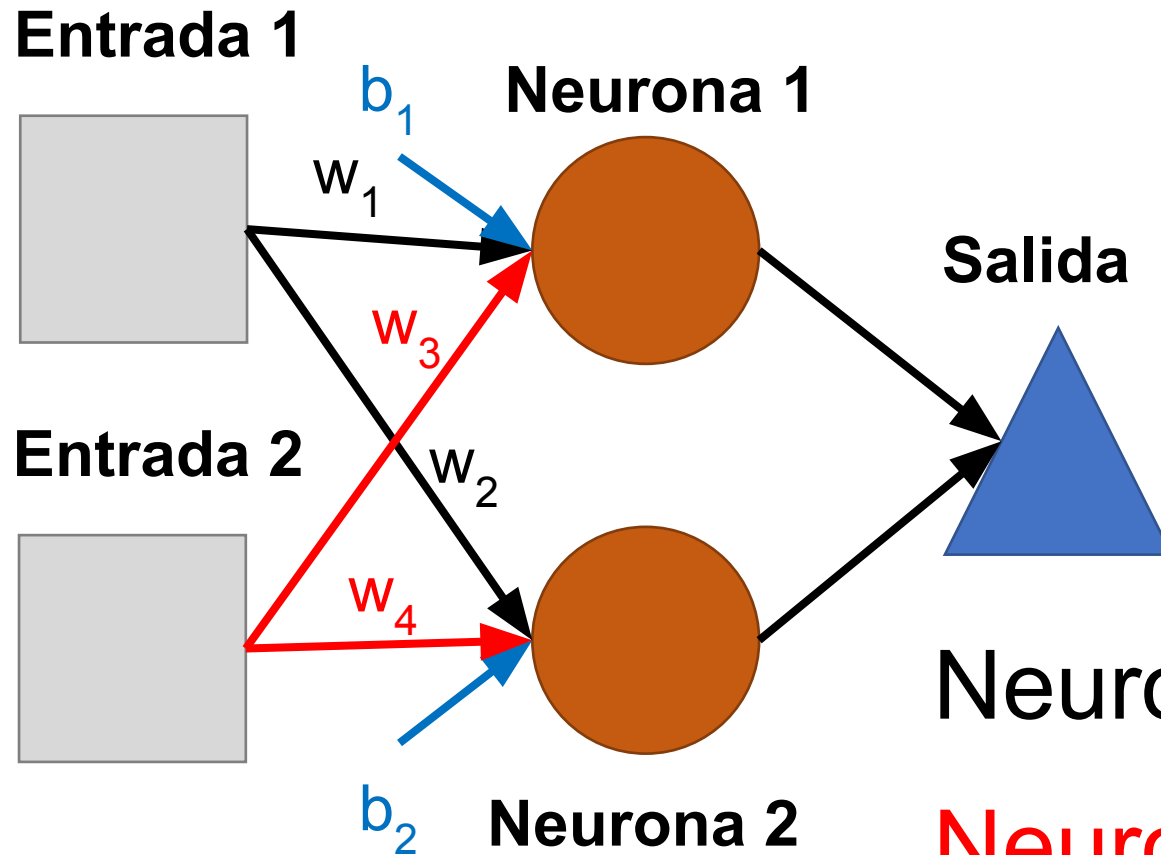
Redes neurales: caso más simple



$$y = x + 2$$



Redes neurales: aumentando la complejidad



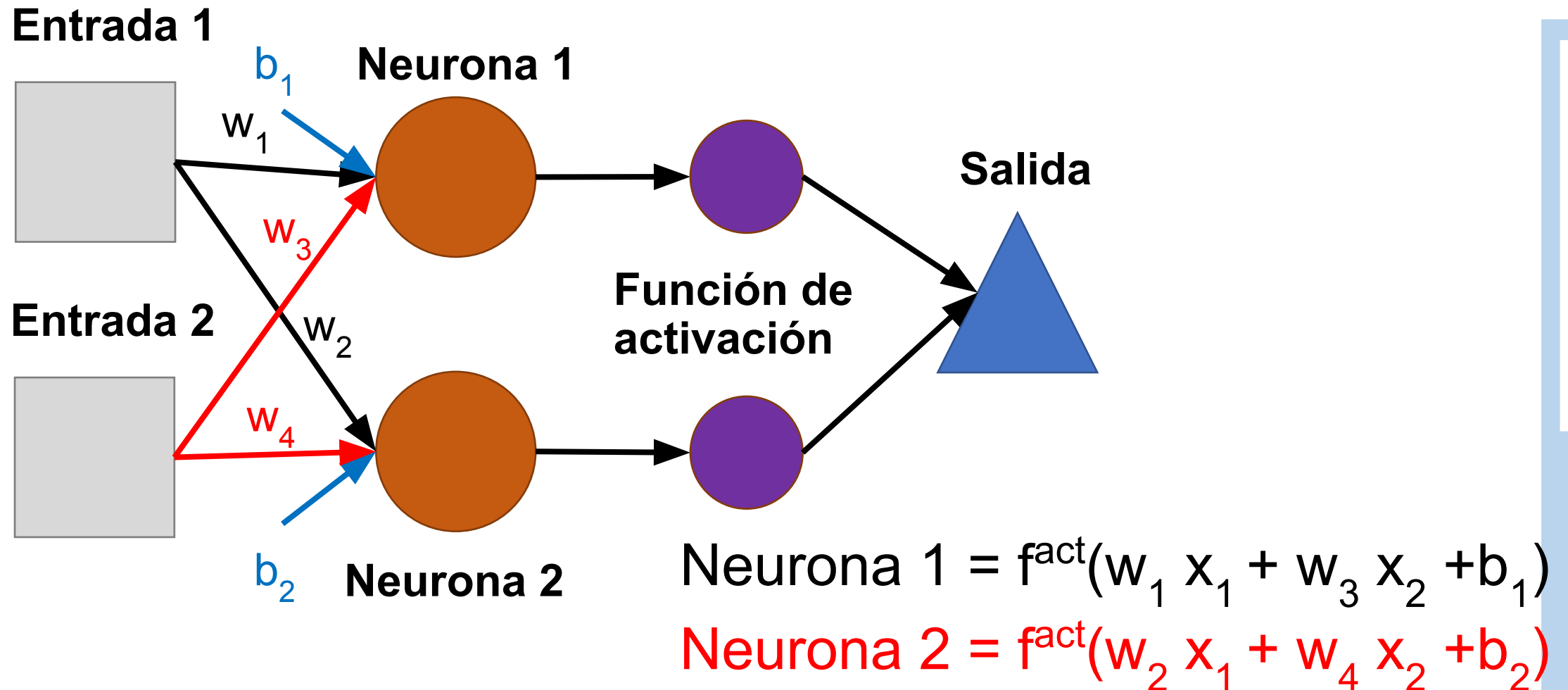
$$\text{Neurona 1} = w_1 x_1 + w_3 x_2 + b_1$$

$$\text{Neurona 2} = w_2 x_1 + w_4 x_2 + b_2$$

Tensorflow Playground

<https://playground.tensorflow.org/>

Redes neurales: funciones de activación



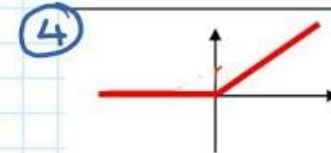
Funciones de activación

3. Linear function: $f(z) = z$



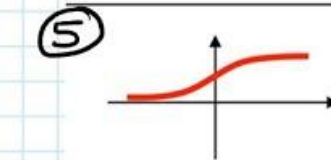
$(-\infty, \infty)$

4. ReLU function: $f(z) = \begin{cases} 0 & z < 0 \\ z & z \geq 0 \end{cases}$



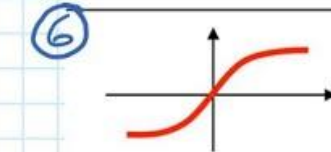
$(0, \infty)$

5. Sigmoid function: $f(z) = \frac{e^z}{1 + e^z}$



$(0, 1)$

6. Hyperbolic tan: $\tanh(z) = \frac{e^z - e^{-z}}{e^z + e^{-z}}$



$(-1, 1)$

by Dr. Pankaj Kumar Porwal (BTech - IIT Mumbai, PhD - Cornell University) : Principal, Techno India NJR Institute of Technology, Udaipur

Tensorflow Playground

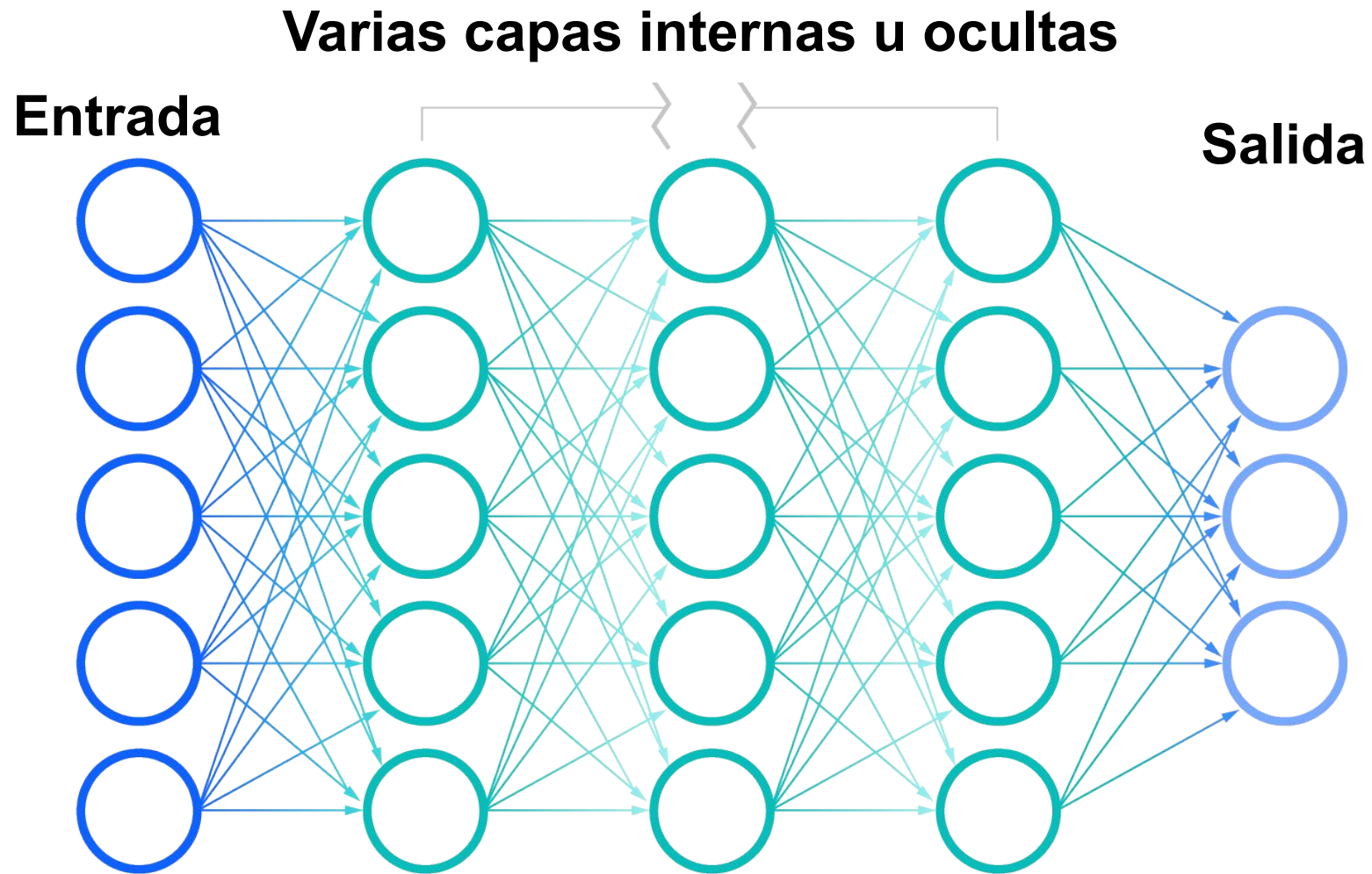
<https://playground.tensorflow.org/>

Segunda Parada: ¿Preguntas hasta aquí?

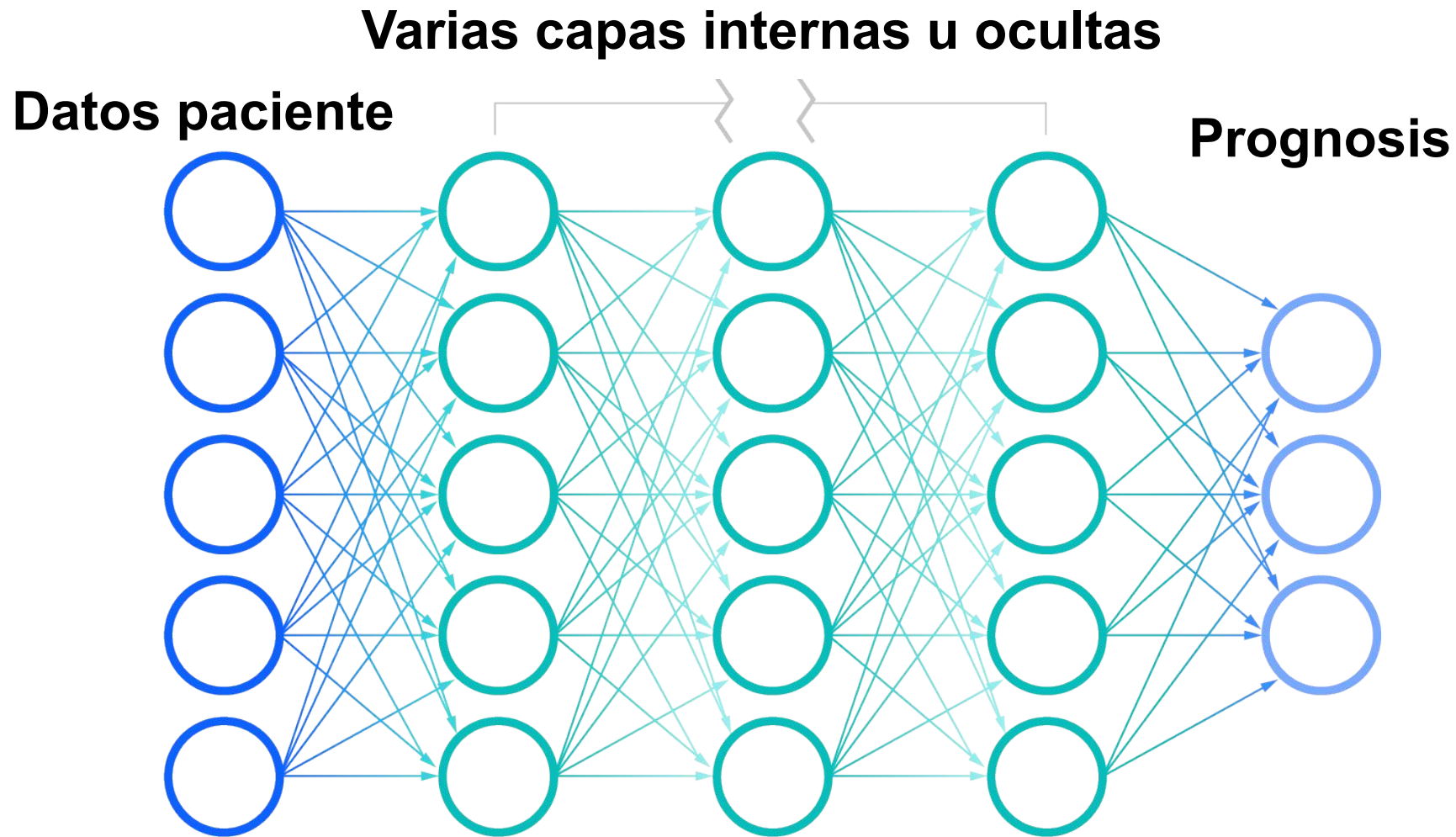
Tercera Parte

¿Qué problemas podemos encontrar al usar redes neurales profundas?

Redes neurales profundas: muchas capas



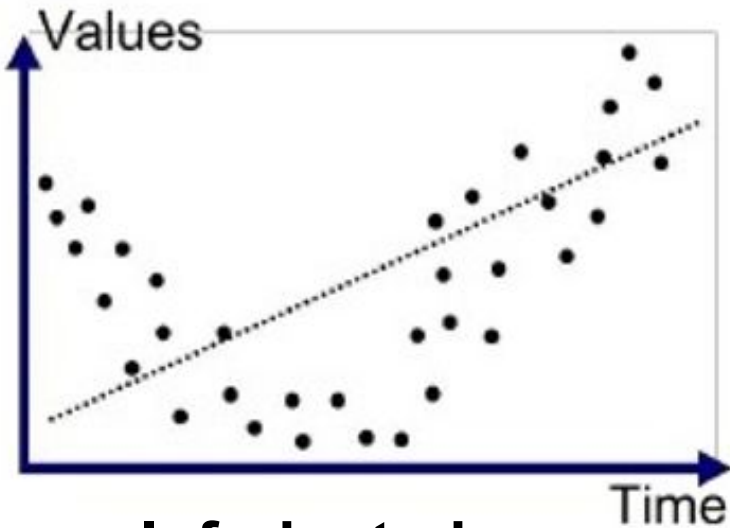
Redes neurales profundas: muchas capas



Problemas: ¿Qué están haciendo?

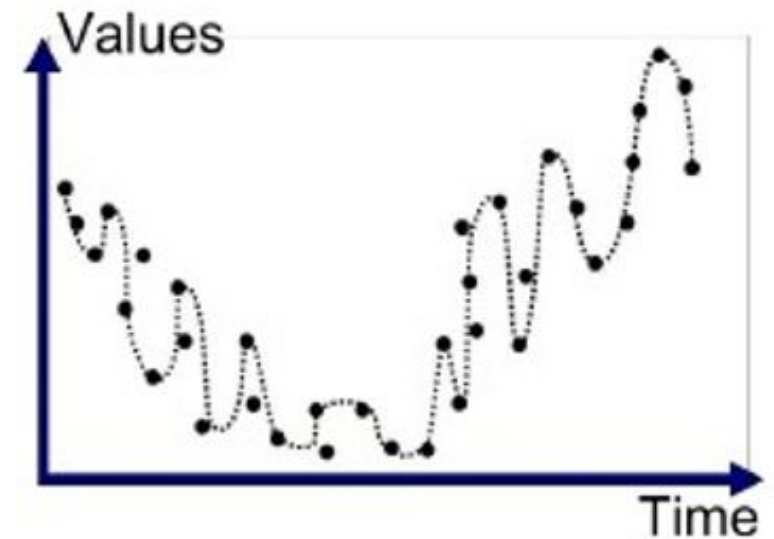


Problemas: Sobreajuste del modelo a los datos



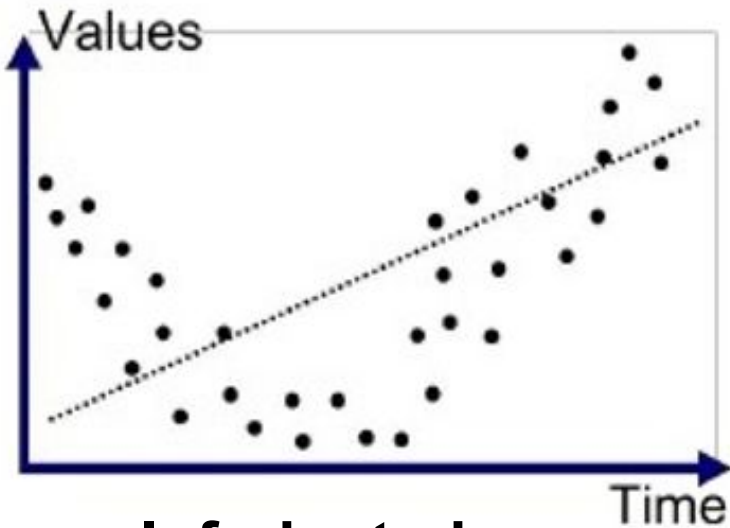
Infrajustado:
modelo demasiado
sencillo o no
suficientes
ejemplos para
aprender

Problemas: Sobreajuste del modelo a los datos

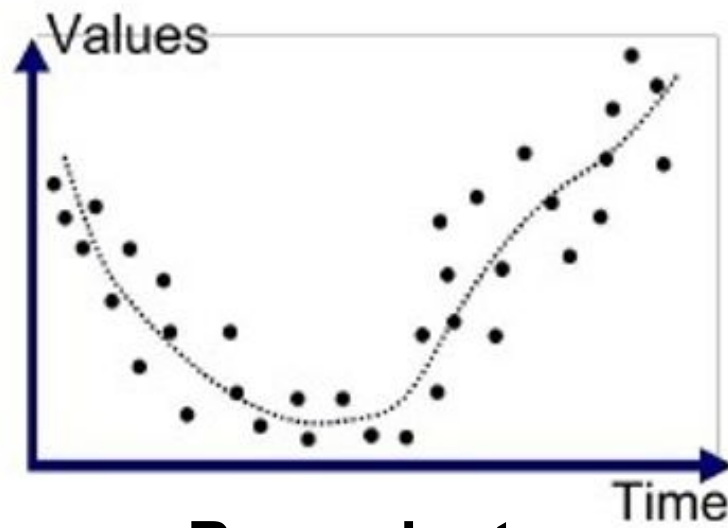


Sobreajustado:
el modelo aprende
de memoria los
ejemplos y no se
puede utilizar sobre
datos nuevos

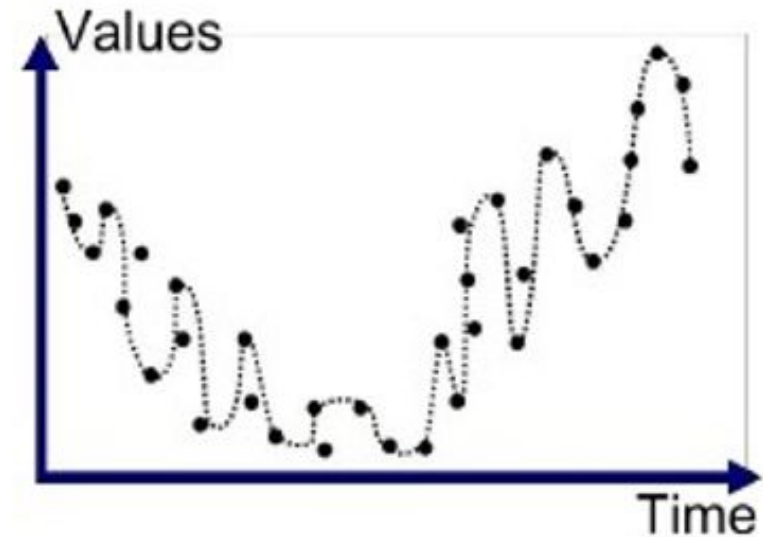
Sobreajuste del modelo a los datos



Infrajustado:
modelo demasiado
sencillo o no
suficientes
ejemplos para
aprender



Buen ajuste:
modelo ajustado a
la complejidad del
problema y
suficientes
ejemplos



Sobreajustado:
el modelo aprende
de memoria los
ejemplos y no se
puede utilizar sobre
datos nuevos

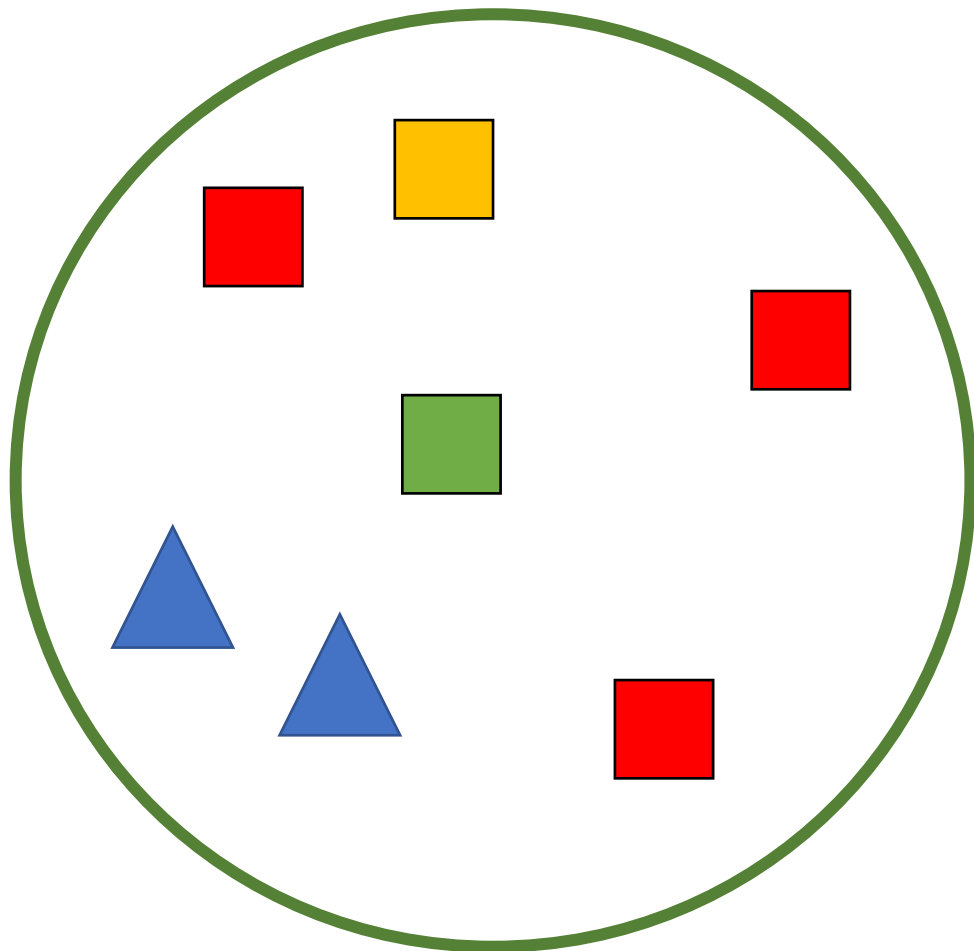
Tensorflow Playground

<https://playground.tensorflow.org/>

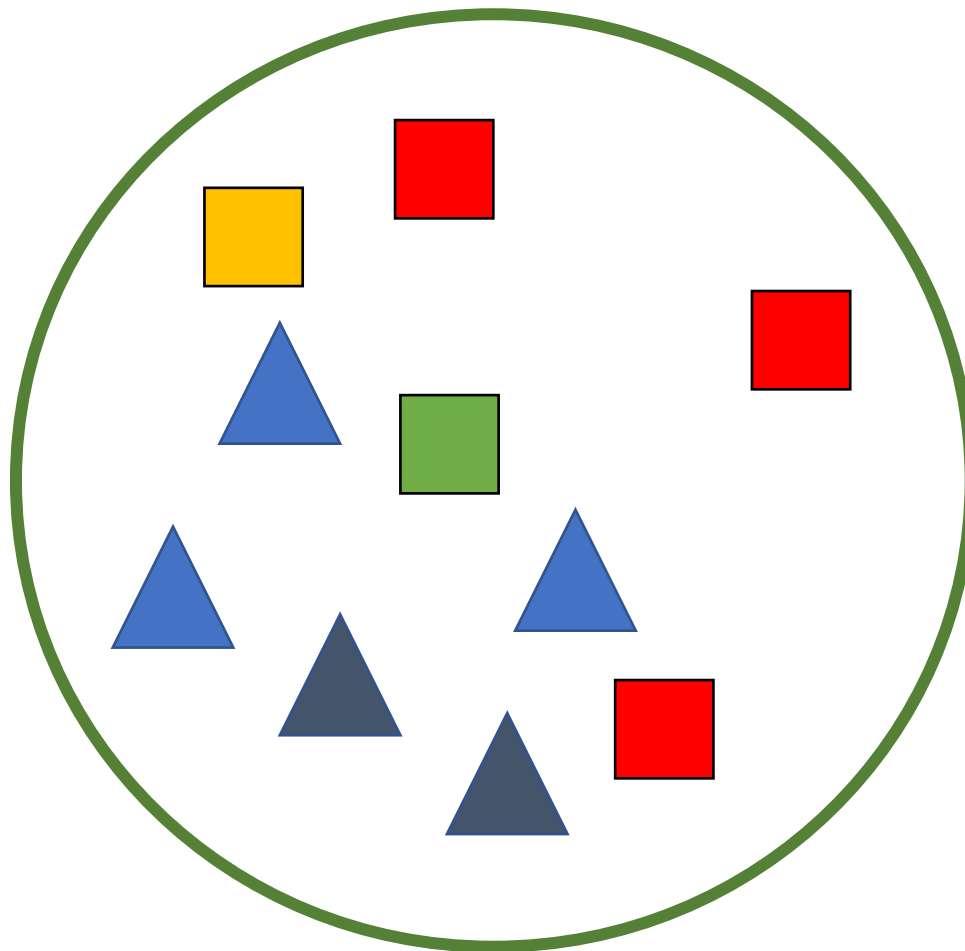
Solución al sobreajuste

1- Utilizar conjuntos de datos lo más amplios posibles evitando repeticiones y sesgos

Sesgo en los datos



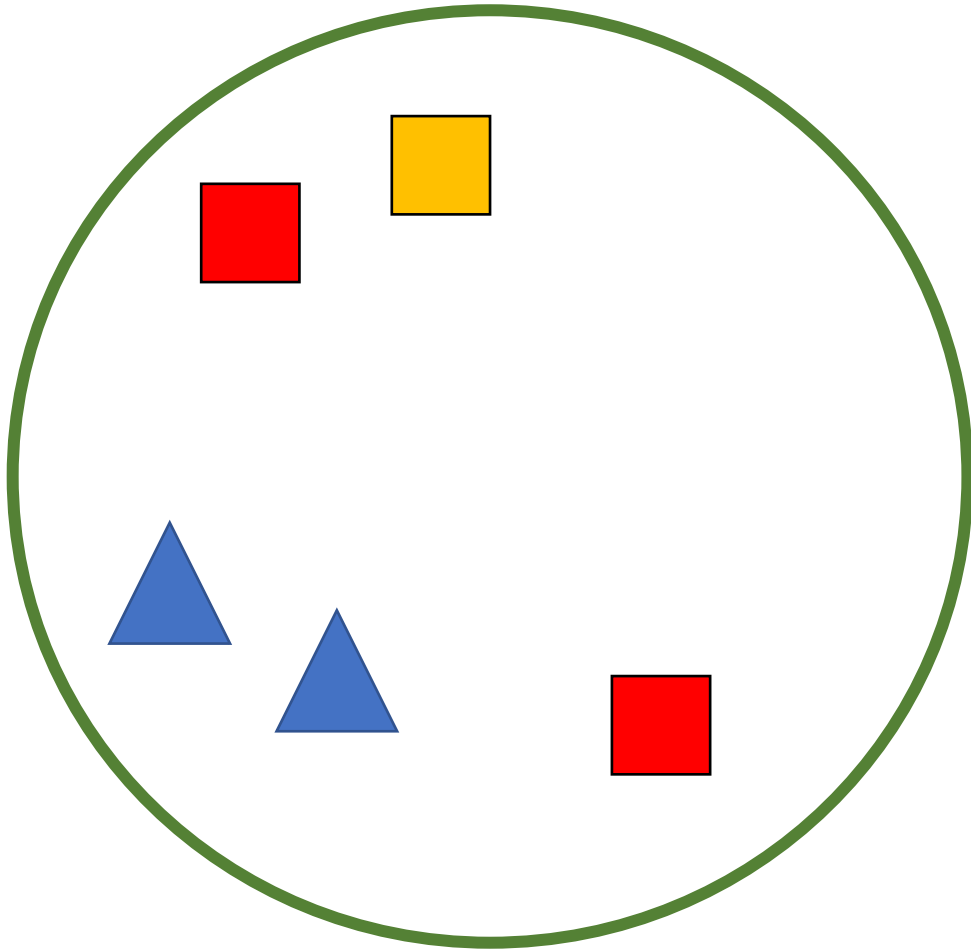
Datos



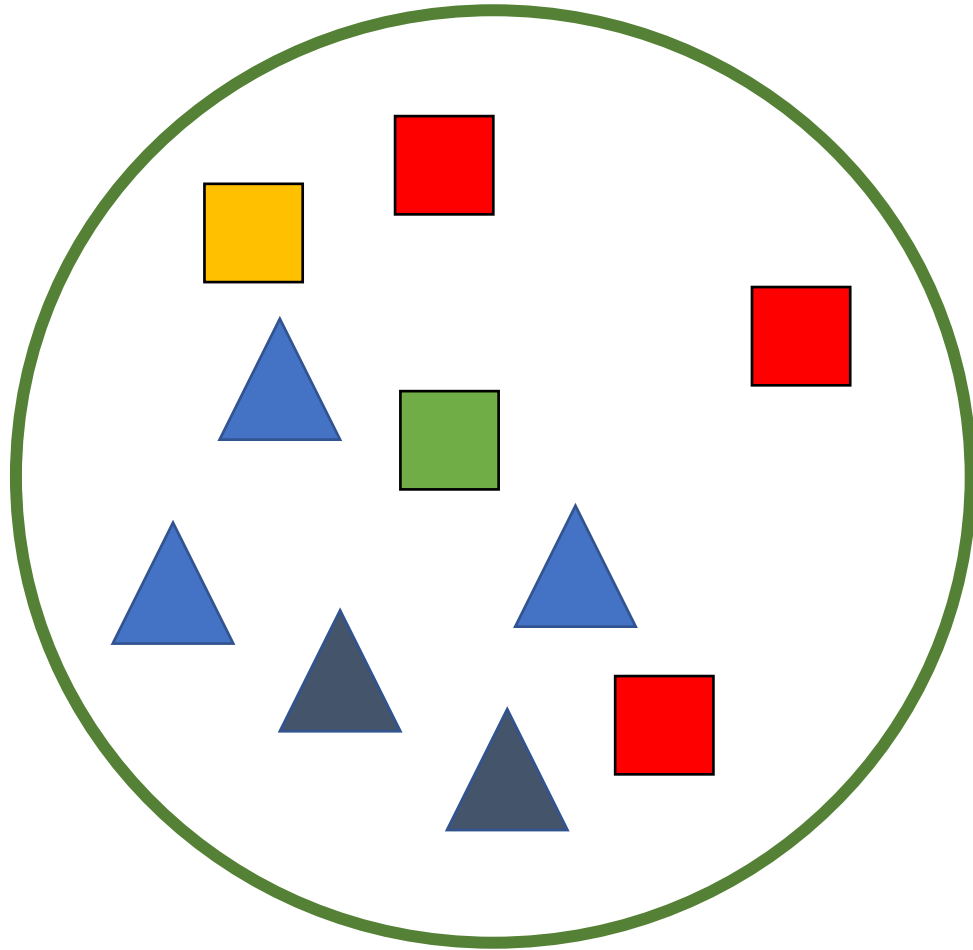
Mundo real

Problemas

Sesgo en los datos: Reducir datos

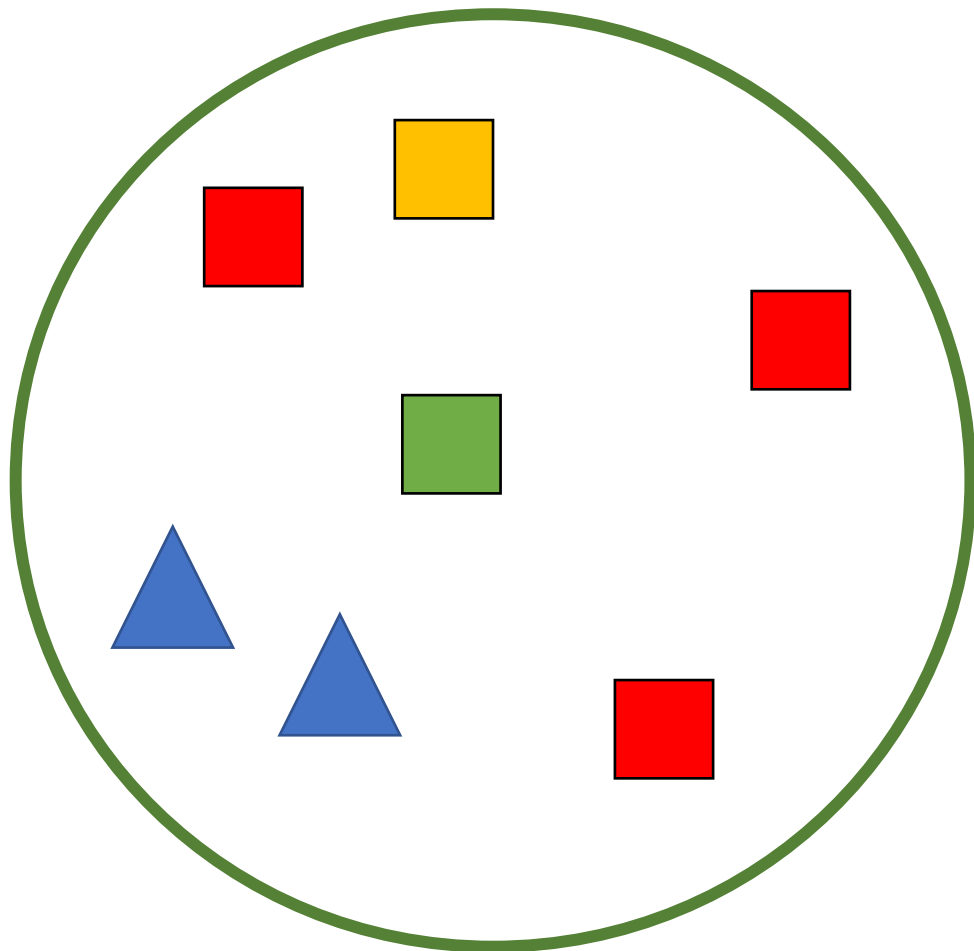


Datos

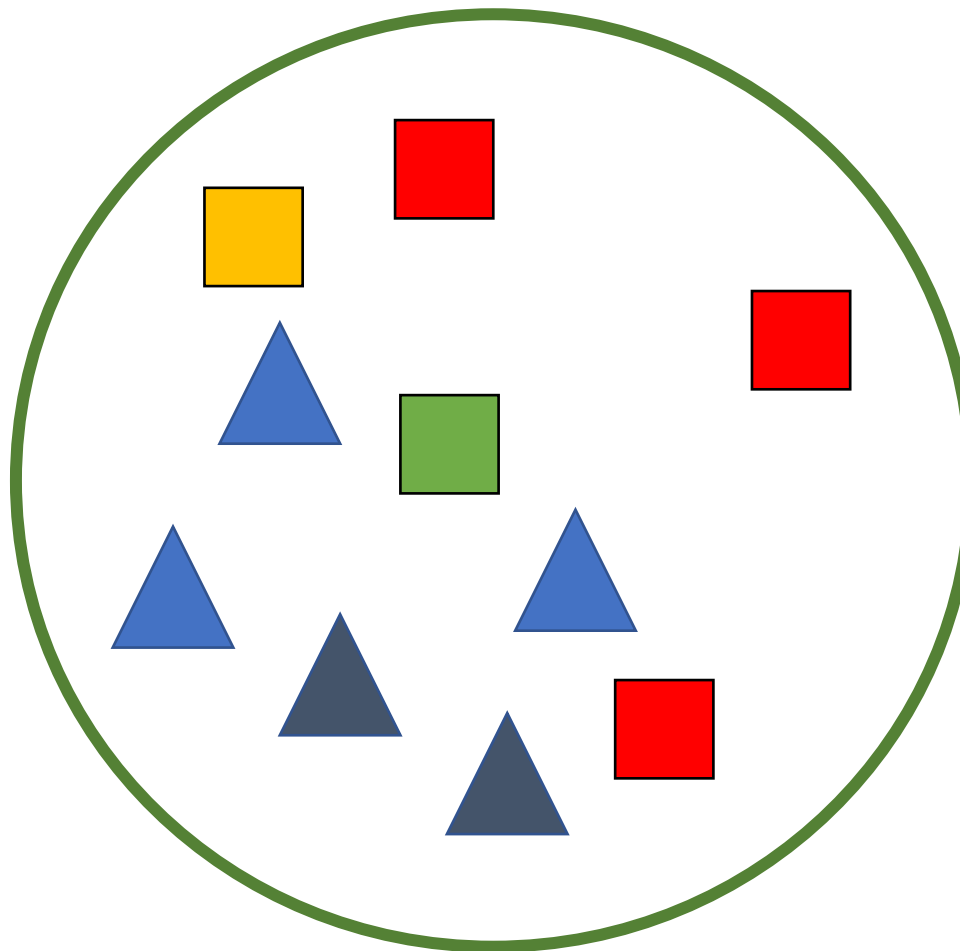


Mundo real

Sesgo en los datos



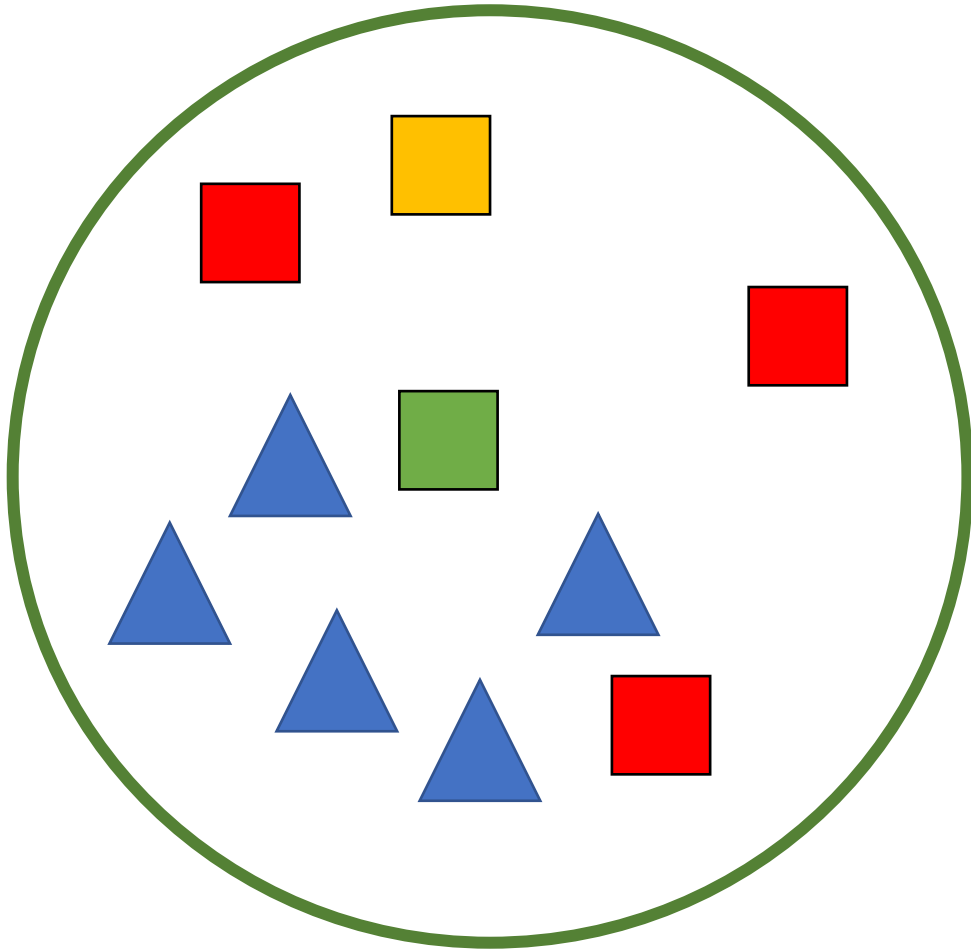
Datos



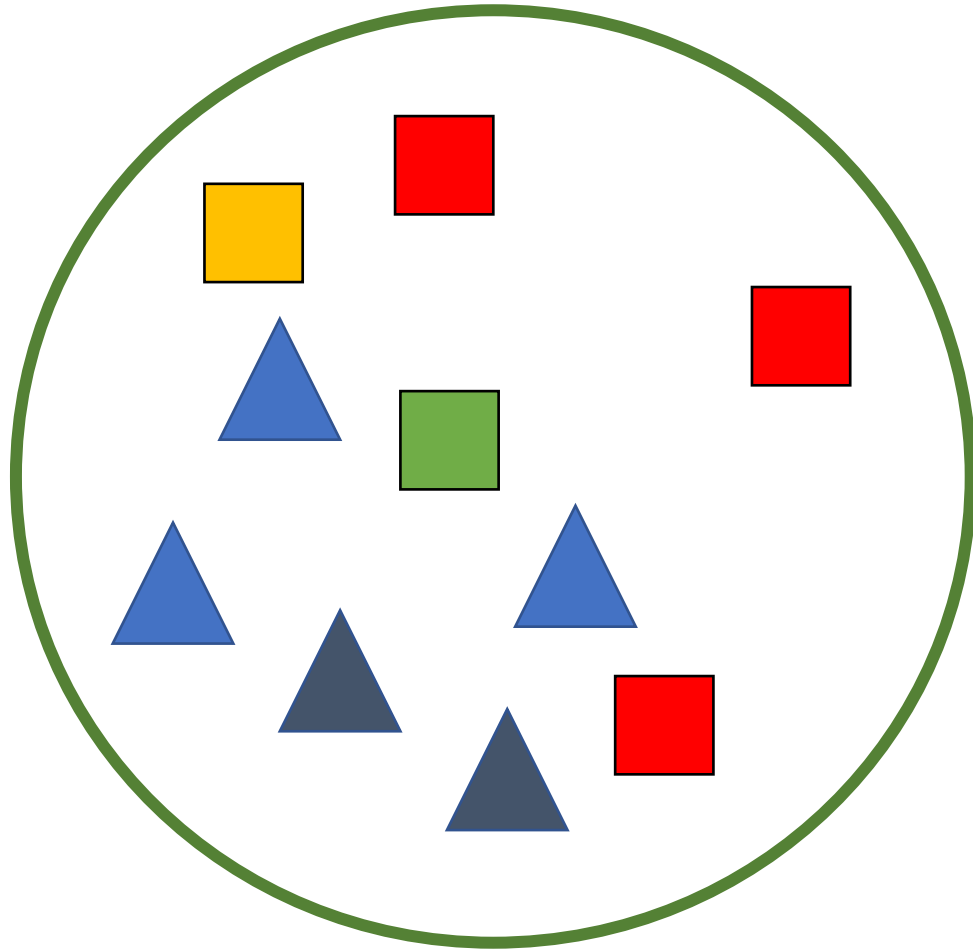
Mundo real

Problemas

Sesgo en los datos: Aumentar datos



Datos



Mundo real

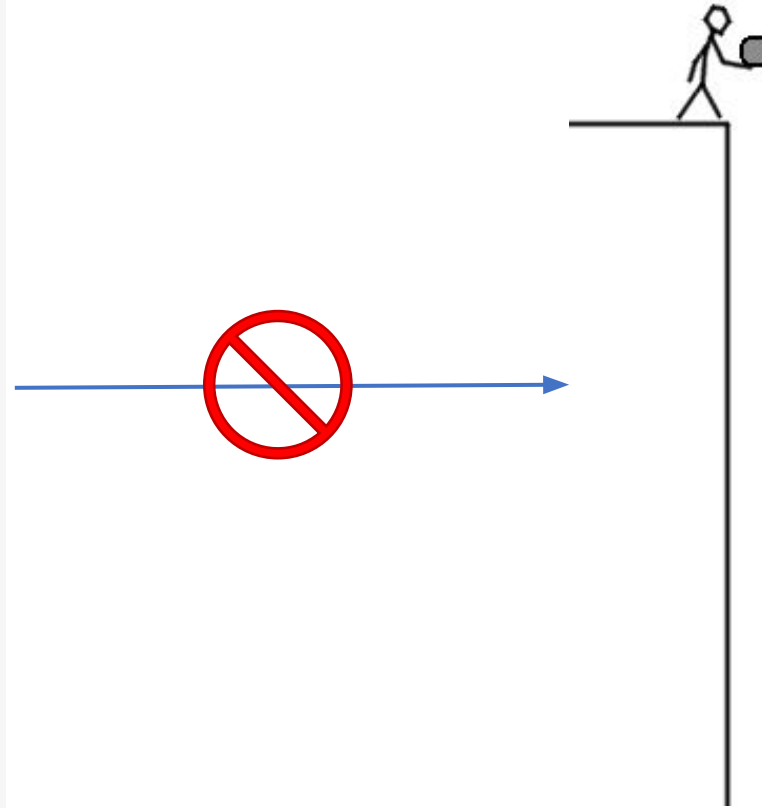
Solución al sobreajuste

1- Utilizar conjuntos de datos lo más amplios posibles evitando repeticiones y sesgos

2- Utilizar modelos cuya complejidad sea proporcional a la complejidad del problema

Problemas: Complejidad del modelo

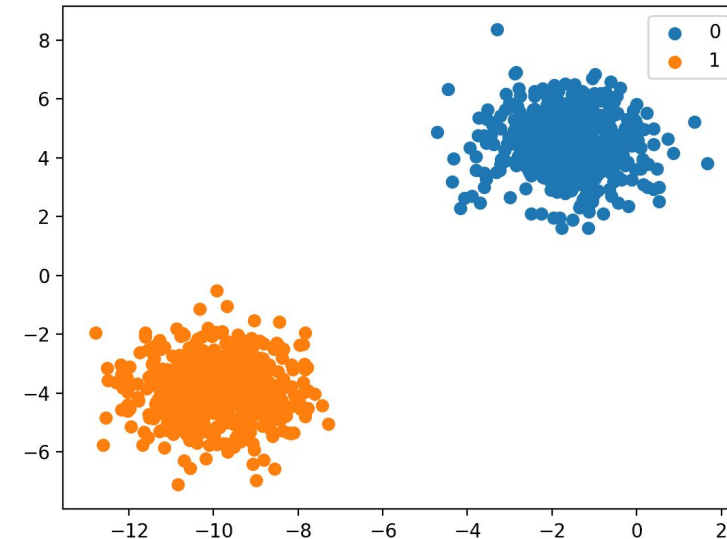
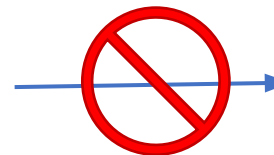
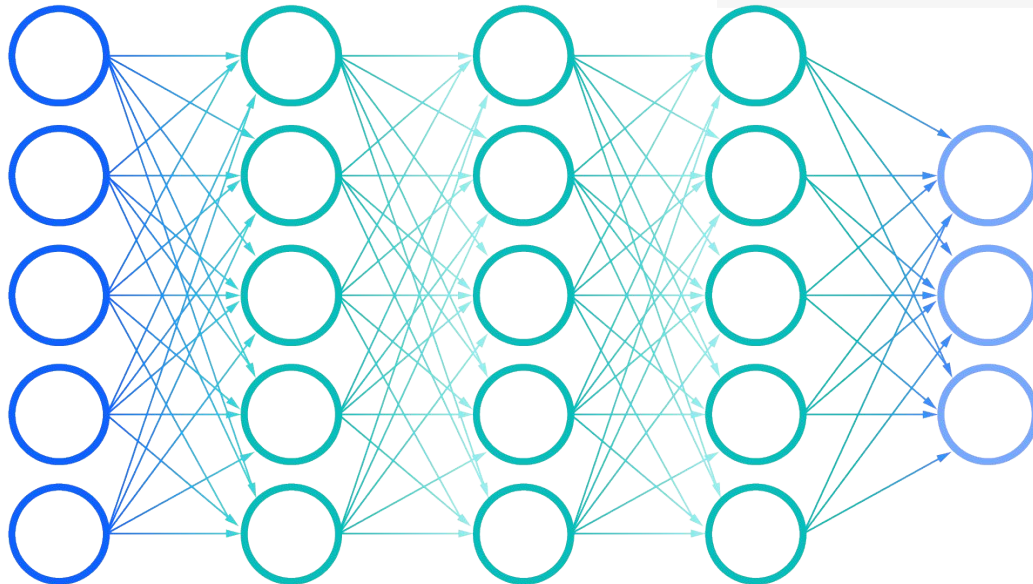
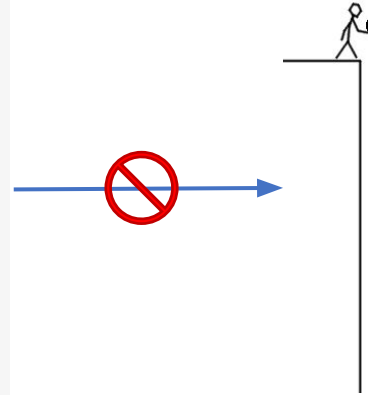
$$\begin{aligned}
 & W \left[\frac{\xi}{\alpha} \left(\frac{\partial f}{\partial t} - \beta^r \frac{\partial f}{\partial r} \right) + \frac{v}{\phi^2} \frac{\partial f}{\partial r} \right] - \frac{\varepsilon W^3}{r \alpha \phi^3} \frac{\partial f}{\partial \varepsilon} \\
 & \times \left\{ \beta^r \phi^3 \left(-\psi - r \mu \frac{\partial v_r}{\partial r} \right) + v_r^2 \phi \left[\beta^r \phi \left(2r \frac{\partial \phi}{\partial r} - \psi \phi \right) \right. \right. \\
 & \left. \left. + r \left(-\mu \frac{\partial \alpha}{\partial r} + \mu^2 \phi^2 \frac{\partial \beta^r}{\partial r} - \frac{\partial \phi^2}{\partial t} \right) \right] \right. \\
 & \left. + v_r^3 \left[r \mu \phi \left(-\mu \frac{\partial \alpha}{\partial r} + \frac{\partial \beta^r \phi^2}{\partial r} - \frac{\partial \phi^2}{\partial t} \right) - \psi \frac{\alpha}{\phi} \frac{\partial r \phi^2}{\partial r} \right] \right. \\
 & \left. + \phi \left[r \mu \left(\mu \alpha \frac{\partial v_r}{\partial r} + \frac{\partial \alpha}{\partial r} + \phi^2 \left(-\mu \frac{\partial \beta^r}{\partial r} + \frac{\partial v_r}{\partial t} \right) \right) \right. \right. \\
 & \left. \left. + r \frac{\partial \phi^2}{\partial t} - r \beta^r \frac{\partial \phi^2}{\partial r} \right] + v_r \alpha \left[\phi \left(\psi + r \mu \frac{\partial v_r}{\partial r} \right) \right. \right. \\
 & \left. \left. + 2r \psi \frac{\partial \phi}{\partial r} + \phi^2 \left(\mu \frac{\partial v_r}{\partial t} - \frac{\partial \beta^r}{\partial r} \right) + \frac{\partial \phi^2}{\partial t} \right] \right\} \\
 & + \frac{W^3 (1 - \mu^2)}{r \alpha \phi^3} \frac{\partial f}{\partial \mu} \left\{ \alpha \left[\phi \left(\frac{\xi}{W^2} - r v \frac{\partial v_r}{\partial r} \right) + 2r \frac{\xi}{W^2} \frac{\partial \phi}{\partial r} \right] \right. \\
 & \left. + \phi \left[\beta \phi^2 \left(r \xi \frac{\partial v_r}{\partial r} - \frac{v}{W^2} \right) - \frac{r}{W^2} \left(\xi \frac{\partial \alpha}{\partial r} - v \phi^2 \frac{\partial \beta^r}{\partial r} \right) \right. \right. \\
 & \left. \left. - r \xi \phi^2 \frac{\partial v_r}{\partial t} \right] \right\} = \mathfrak{C}[f], \tag{26}
 \end{aligned}$$



Problemas

Problemas: Complejidad del modelo

$$\begin{aligned}
 & W \left[\frac{\xi}{\alpha} \left(\frac{\partial f}{\partial t} - \beta' \frac{\partial f}{\partial r} \right) + \frac{v}{\phi^2} \frac{\partial f}{\partial r} \right] - \frac{\varepsilon W^3}{r \alpha \phi^3} \frac{\partial f}{\partial \varepsilon} \\
 & \times \left\{ \beta' \phi^3 \left(-\psi - r \mu \frac{\partial v_r}{\partial r} \right) + v_r^2 \phi \left[\beta' \phi \left(2r \frac{\partial \phi}{\partial r} - \psi \phi \right) \right. \right. \\
 & \left. \left. + r \left(-\mu \frac{\partial \alpha}{\partial r} + \mu^2 \phi^2 \frac{\partial \beta'}{\partial r} - \frac{\partial \phi^2}{\partial t} \right) \right] \right. \\
 & \left. + v_r^3 \left[r \mu \phi \left(-\mu \frac{\partial \alpha}{\partial r} + \frac{\partial \beta'}{\partial r} - \frac{\partial \phi^2}{\partial t} \right) - \psi \frac{\alpha}{\phi} \frac{\partial \phi^2}{\partial r} \right] \right. \\
 & \left. + \phi \left[r \mu \left(\mu \frac{\partial v_r}{\partial r} + \frac{\partial \alpha}{\partial r} + \phi^2 \left(-\mu \frac{\partial \beta'}{\partial r} + \frac{\partial v_r}{\partial t} \right) \right) \right. \right. \\
 & \left. \left. + r \frac{\partial \phi^2}{\partial t} - r \beta' \frac{\partial \phi^2}{\partial r} \right] + v_r \alpha \left[\phi \left(\psi + r \mu \frac{\partial v_r}{\partial r} \right) \right. \right. \\
 & \left. \left. + 2r \psi \frac{\partial \phi}{\partial r} + \phi^2 \left(\mu \frac{\partial v_r}{\partial t} - \frac{\partial \beta'}{\partial r} \right) + \frac{\partial \phi^2}{\partial t} \right] \right. \\
 & \left. + \frac{W^3 (1 - \mu^2)}{r \alpha \phi^3} \frac{\partial f}{\partial \mu} \left\{ \alpha \left[\phi \left(\frac{\xi}{W^2} - r v \frac{\partial v_r}{\partial r} \right) + 2r \frac{\xi}{W^2} \frac{\partial \phi}{\partial r} \right] \right. \right. \\
 & \left. \left. + \phi \left[\beta' \phi^2 \left(r \xi \frac{\partial v_r}{\partial r} - \frac{v}{W^2} \right) - \frac{r}{W^2} \left(\xi \frac{\partial \alpha}{\partial r} - v \phi^2 \frac{\partial \beta'}{\partial r} \right) \right. \right. \right. \\
 & \left. \left. \left. - r \xi \phi^2 \frac{\partial v_r}{\partial t} \right] \right\} \right\} = \mathbb{E}[f], \quad (26)
 \end{aligned}$$



Solución al sobreajuste

1- Utilizar conjuntos de datos lo más amplios posibles evitando repeticiones y sesgos

2- Utilizar modelos cuya complejidad sea proporcional a la complejidad del problema

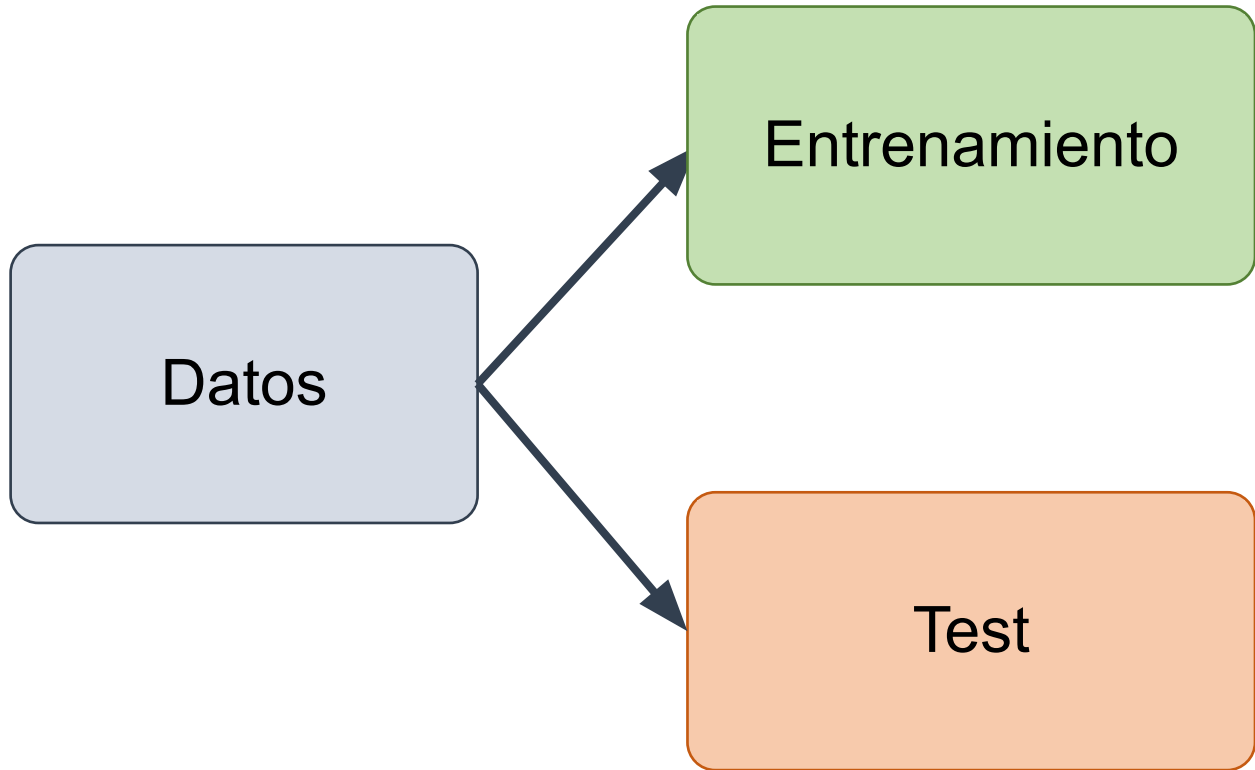
3- Diferenciar datos en grupos de entrenamiento, validación y test

Distribución de los datos

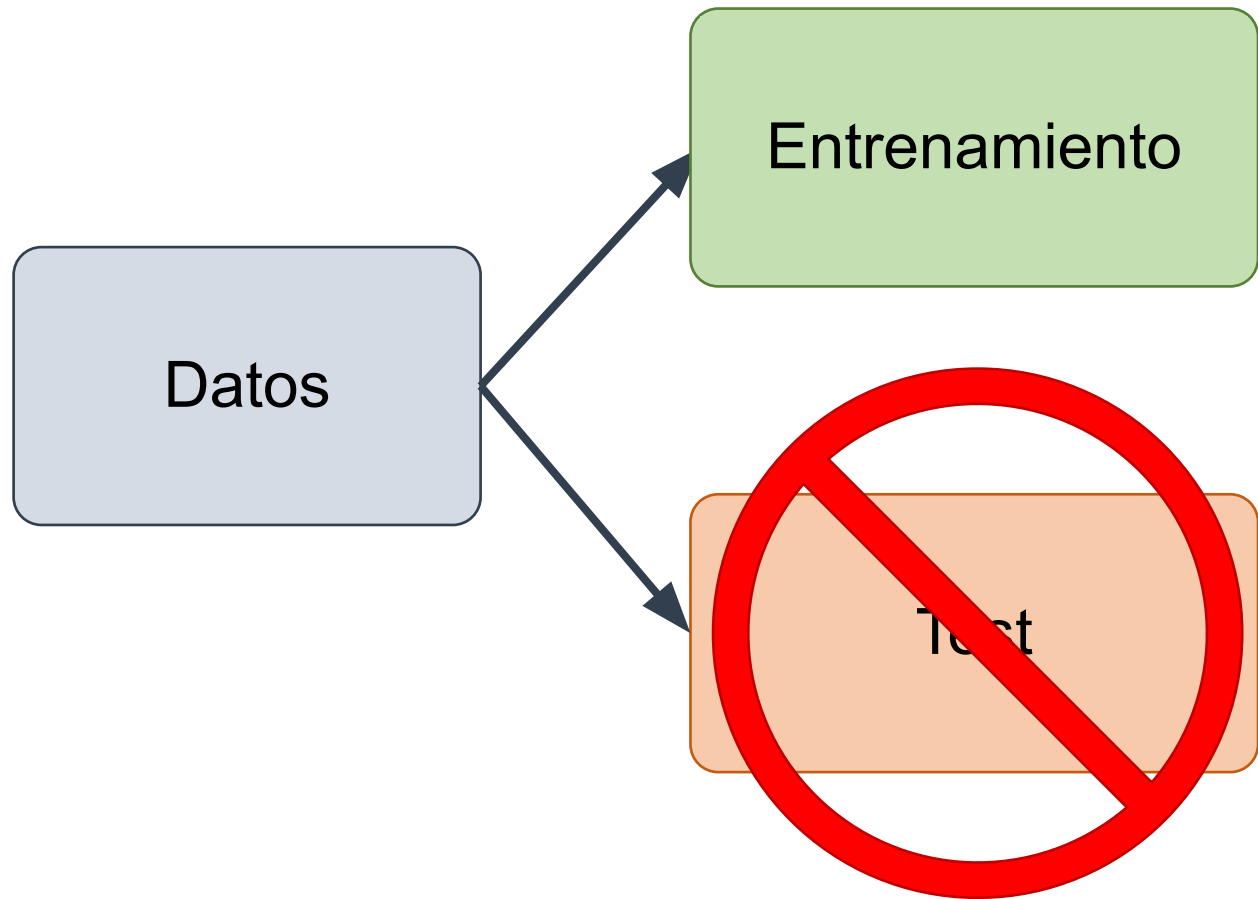
Datos

Problemas

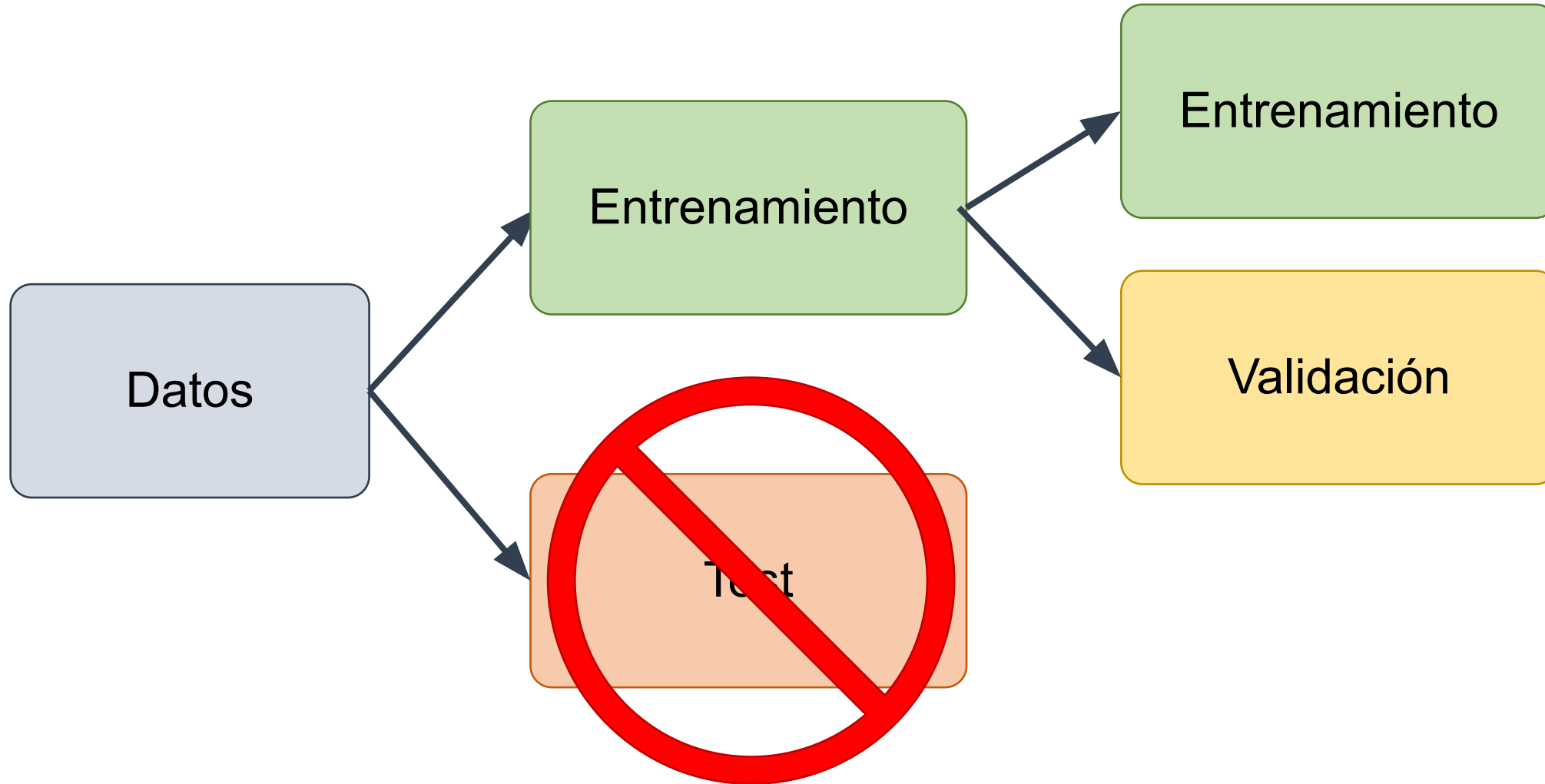
Distribución de los datos



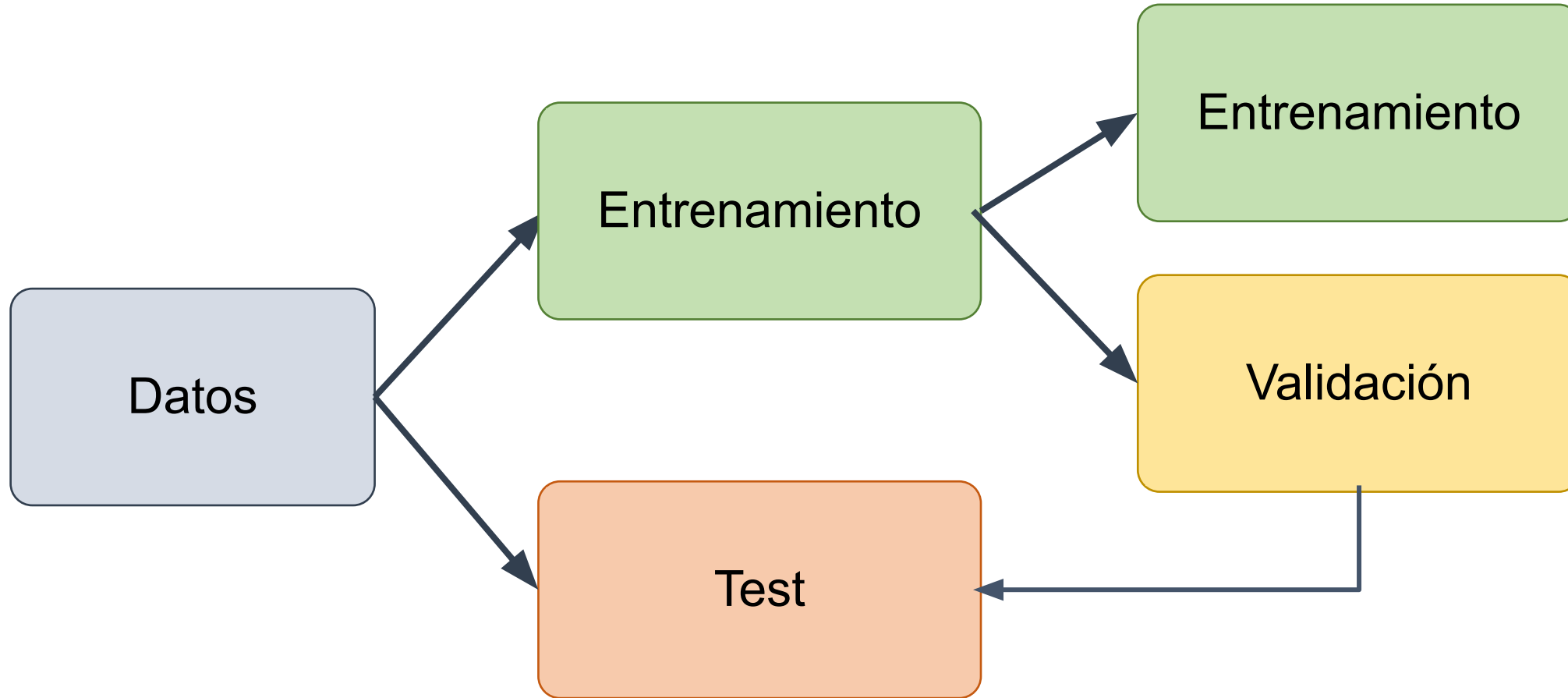
Distribución de los datos



Distribución de los datos



Distribución de los datos



Solución al sobreajuste

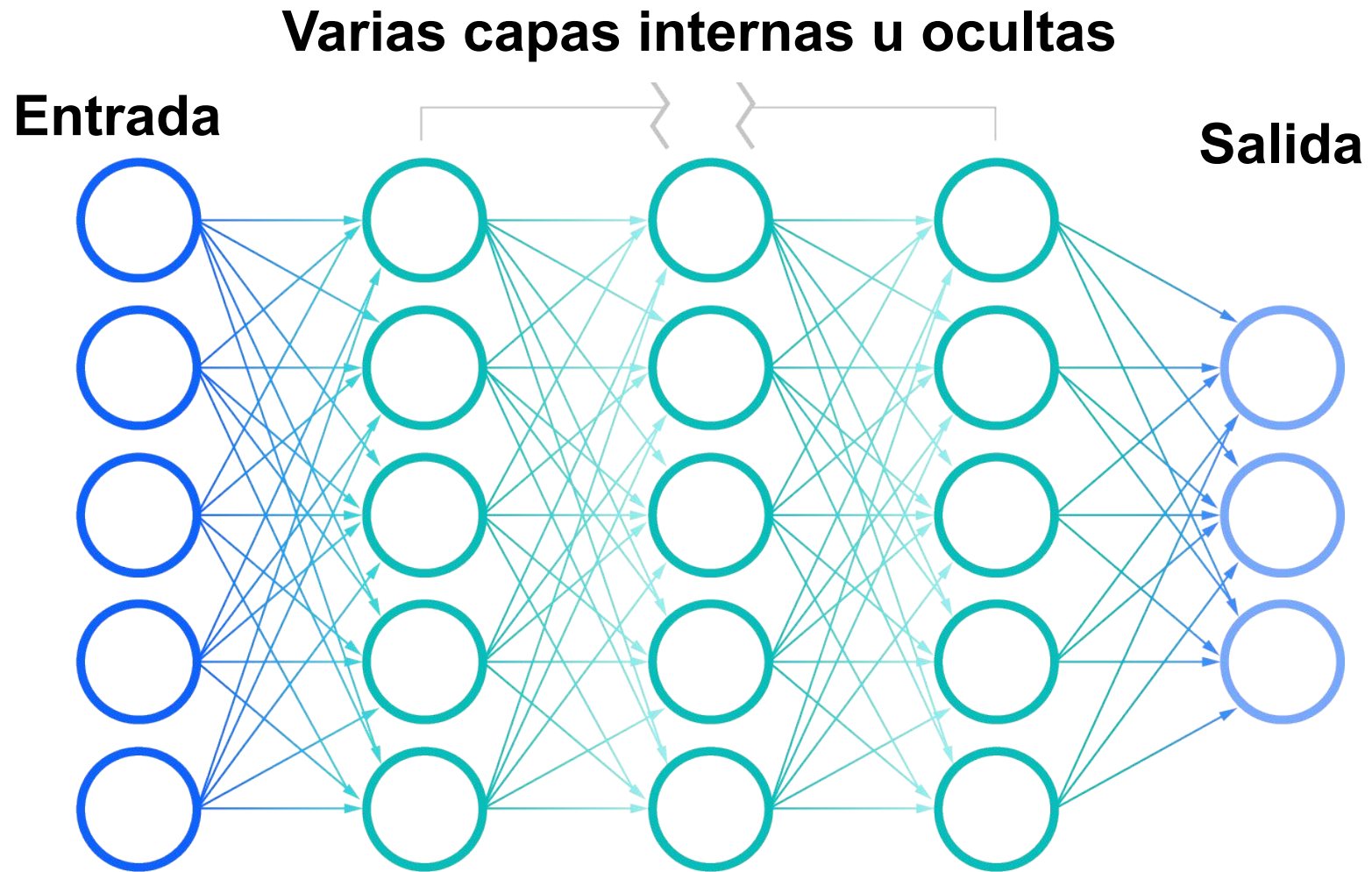
1- Utilizar conjuntos de datos lo más amplios posibles evitando repeticiones y sesgos

2- Utilizar modelos cuya complejidad sea proporcional a la complejidad del problema

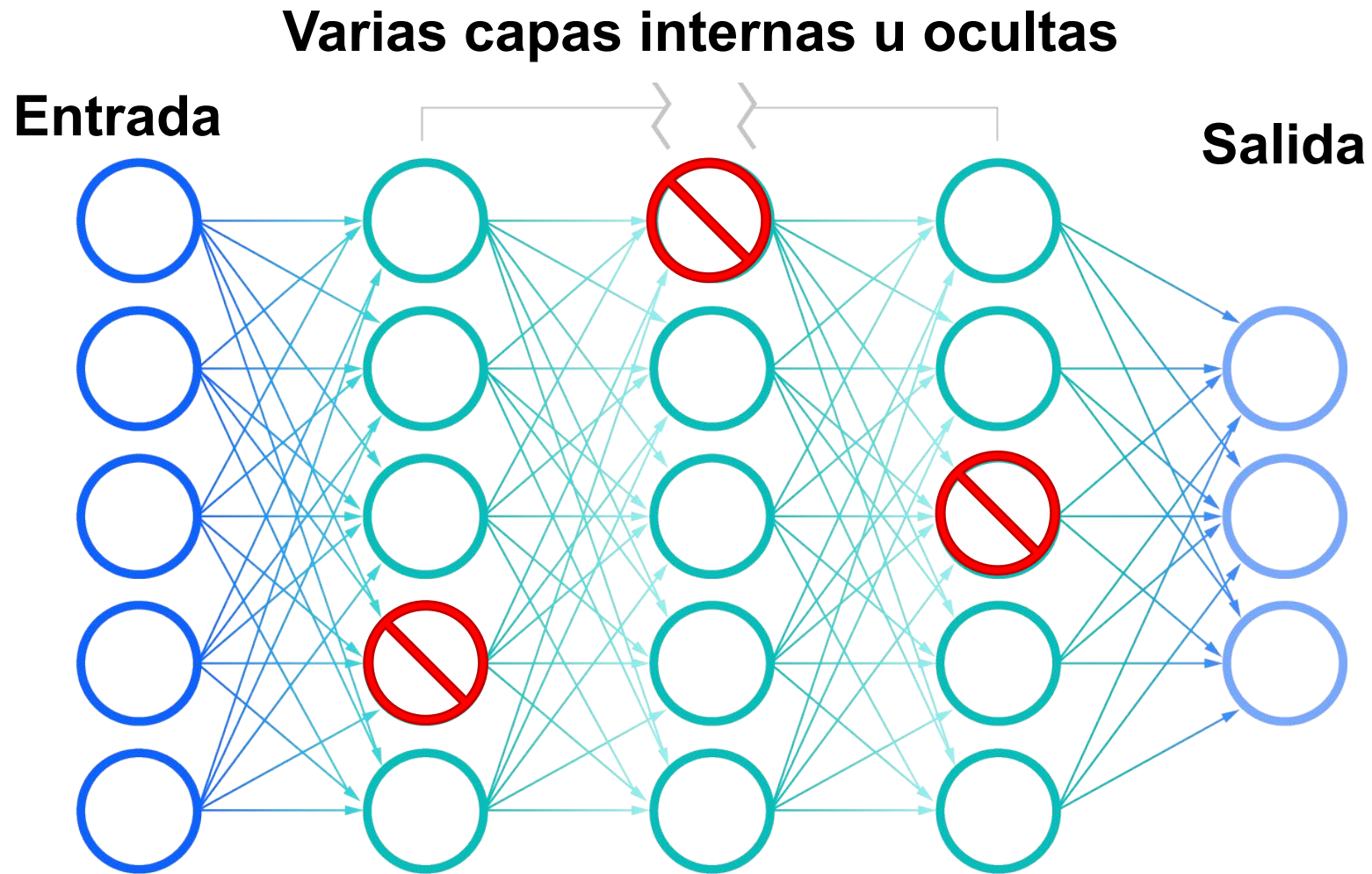
3- Diferenciar datos en grupos de entrenamiento, validación y test

4- Implementar *dropout*

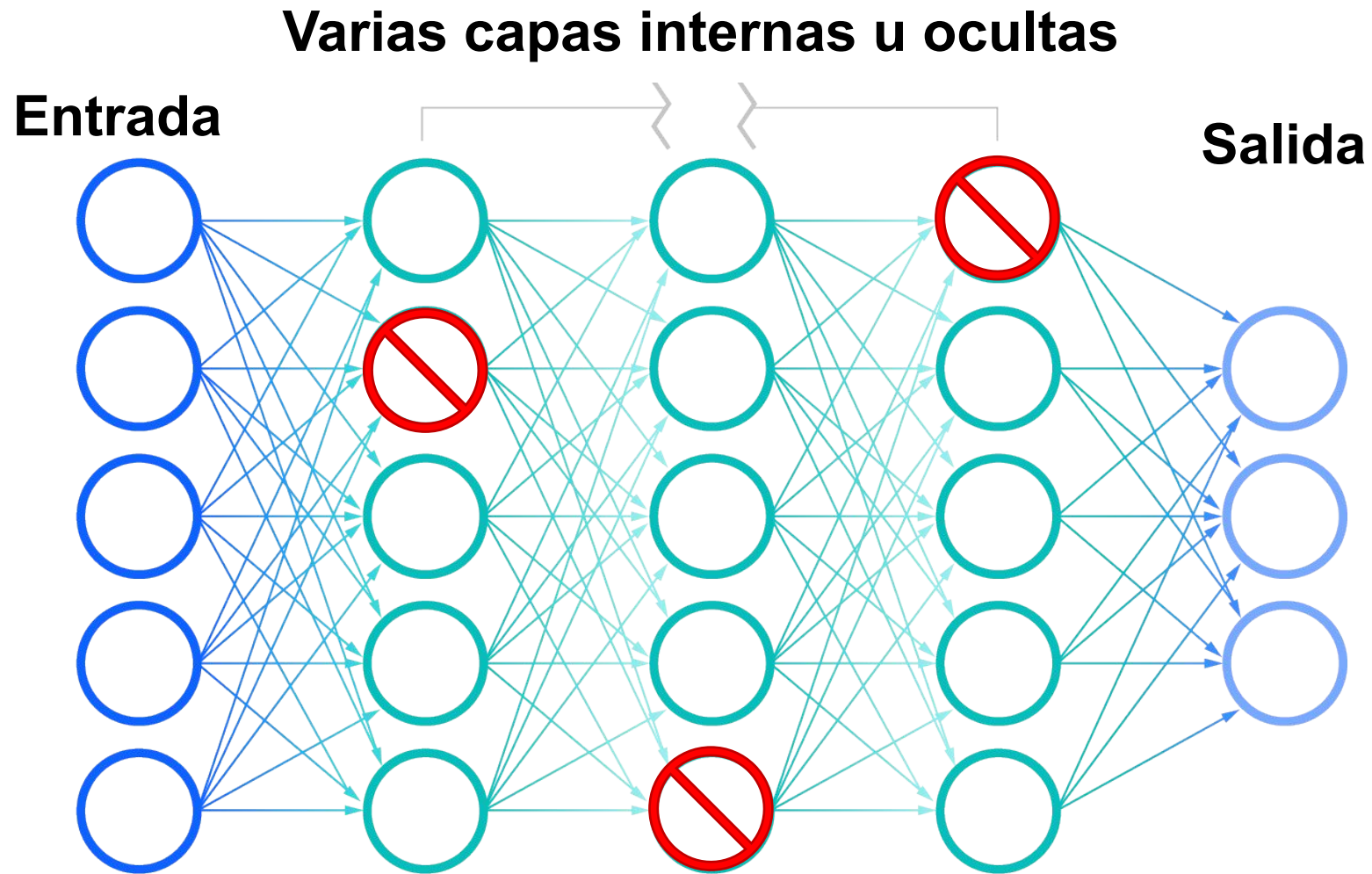
Dropout



Dropout



Dropout



Tercera Parada: ¿Preguntas hasta aquí?

¡Muchas gracias por su atención!

Taller de aprendizaje profundo aplicado en biomedicina

Raúl Fernández Díaz

