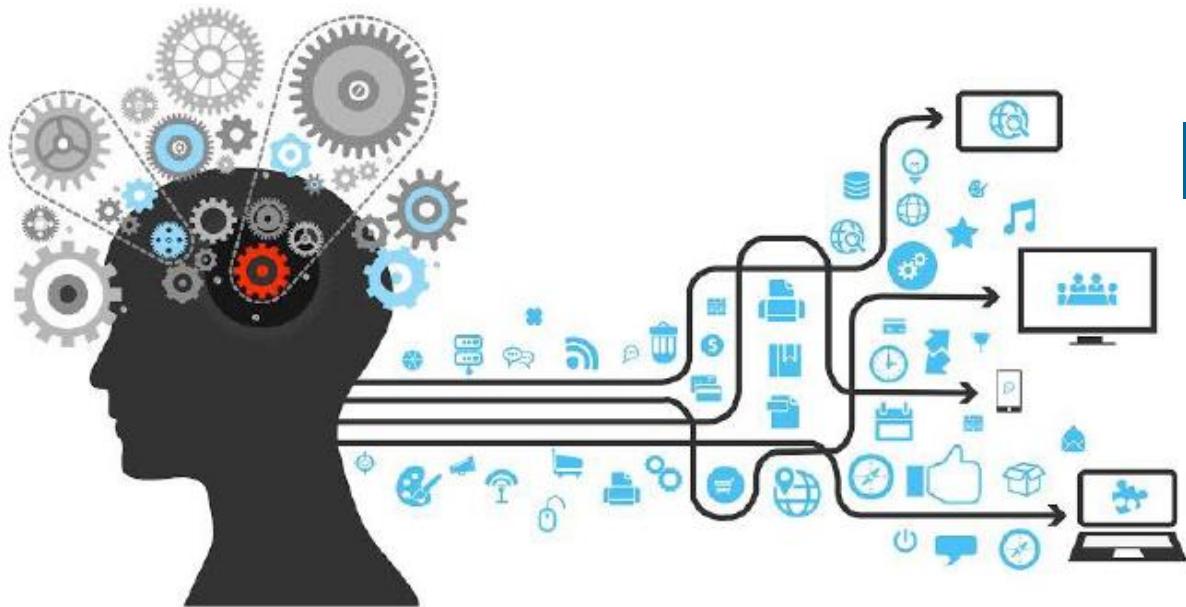




En una palabra, ¿cuál es el área de su profesión?

Ir a menti.com, ingresar código 65 58 211 y escribir su respuesta



Inteligencia Artificial: Redes Neuronales y Deep learning usando R

Por: Néstor Montaño P.

Masapp S.A.

R User Group Ecuador

Sociedad Ecuatoriana de Estadística

2020

El ser humano sigue evolucionando, sigue en un constante desarrollo...

El 2020 nos ha traído:



* Modificado de Sephko.com

Y las noticias recogen estos avances...

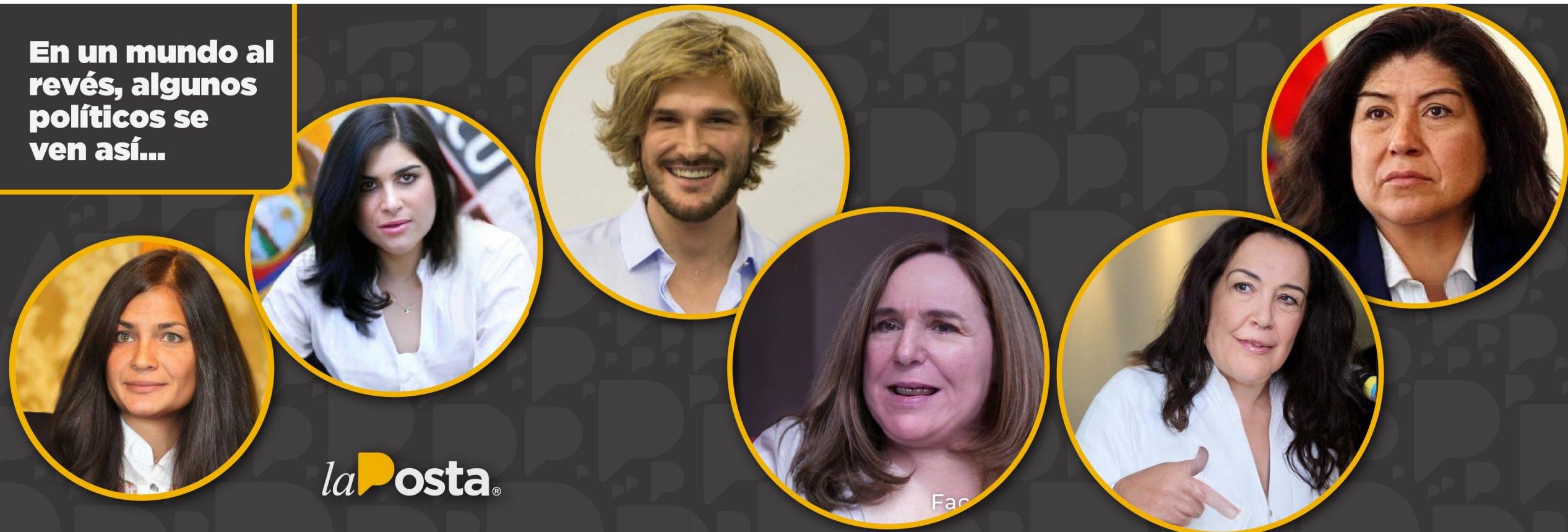
En un mundo al revés, algunos políticos se ven así...



* Tomado de portal: LaPosta

Y las noticias recogen estos avances...

En un mundo al revés, algunos políticos se ven así...



Pero también:

Néstor Montaño | IEEE Andescon | 2020 - Octubre

Y las noticias recogen estos avances...

MIT News

The team's model was shown to be able to identify a woman at high risk of breast cancer four years (left) before it developed (right).

Image courtesy of the researchers

Using AI to predict breast cancer and personalize care

MIT/MGH's image-based deep learning model can predict breast cancer up to five years in advance.

IA LATAM

Home IA Latam

Que están haciendo otros países para combatir el coronavirus usando big-data e IA

Corea del Sur está luchando contra el nuevo coronavirus (COVID-19) confiando en su fortaleza tecnológica. El país tiene una plataforma digital avanzada para la minería de big data, junto con inteligencia artificial (IA) y los coreanos lideran el frente tecnológico, con Samsung compitiendo estrechamente con Apple. Utilizando el análisis de grandes datos, los sistemas de alerta avanzada impulsados por IA y la metodología de observación intensiva, Corea del Sur ya ha logrado controlar la situación del coronavirus en el país en poco

BBC NEWS

Home Video World UK Business Tech Science Stories Entertainment & Arts Business Market Data Global Trade Companies Entrepreneurship Technology of Business

The robot assistant that can guess what you want

By Ben Morris
Technology of Business editor

02 May 2020

f t w e Share

Watch the SecondHands robot in action

Thomas Roszak was working as a maintenance technician at Ocado's giant warehouse in Hatfield when he received a very unusual assignment.

Entendiendo qué es Inteligencia Artificial y Machine Learning

Hoy se consumen muchas noticias que involucran el uso de inteligencia artificial para realizar una tarea,

¿Qué es Inteligencia artificial?

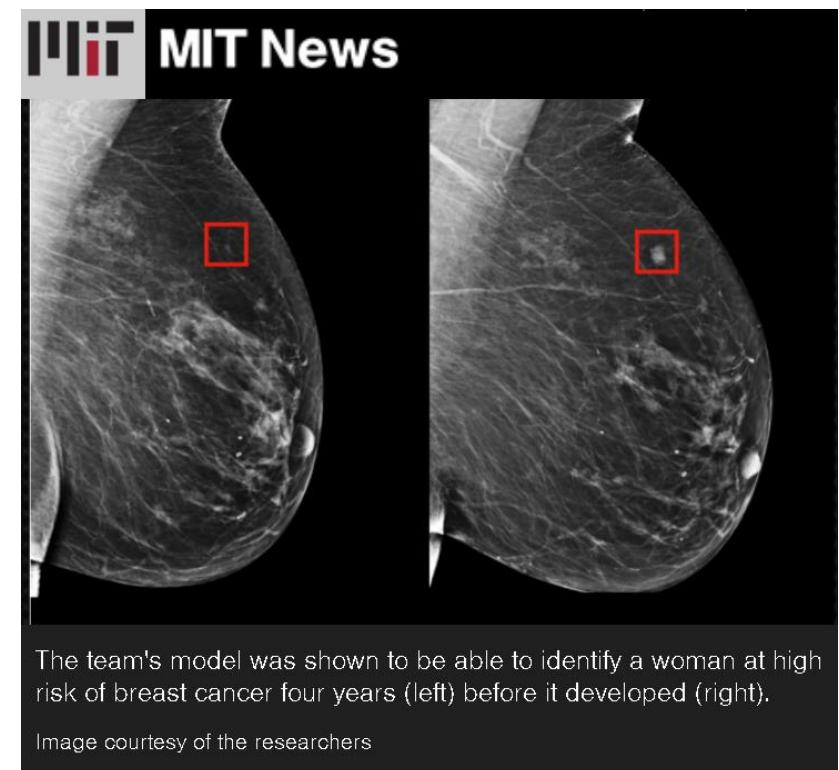
¿Qué es Machine learning?

¿Cuáles son los avances reales?

¿Cuáles son los problemas que pueden surgir de su uso?

¿Qué relación tiene con la Ciencia de Datos?

En esta introducción trataremos de abordar estos cuestionamientos

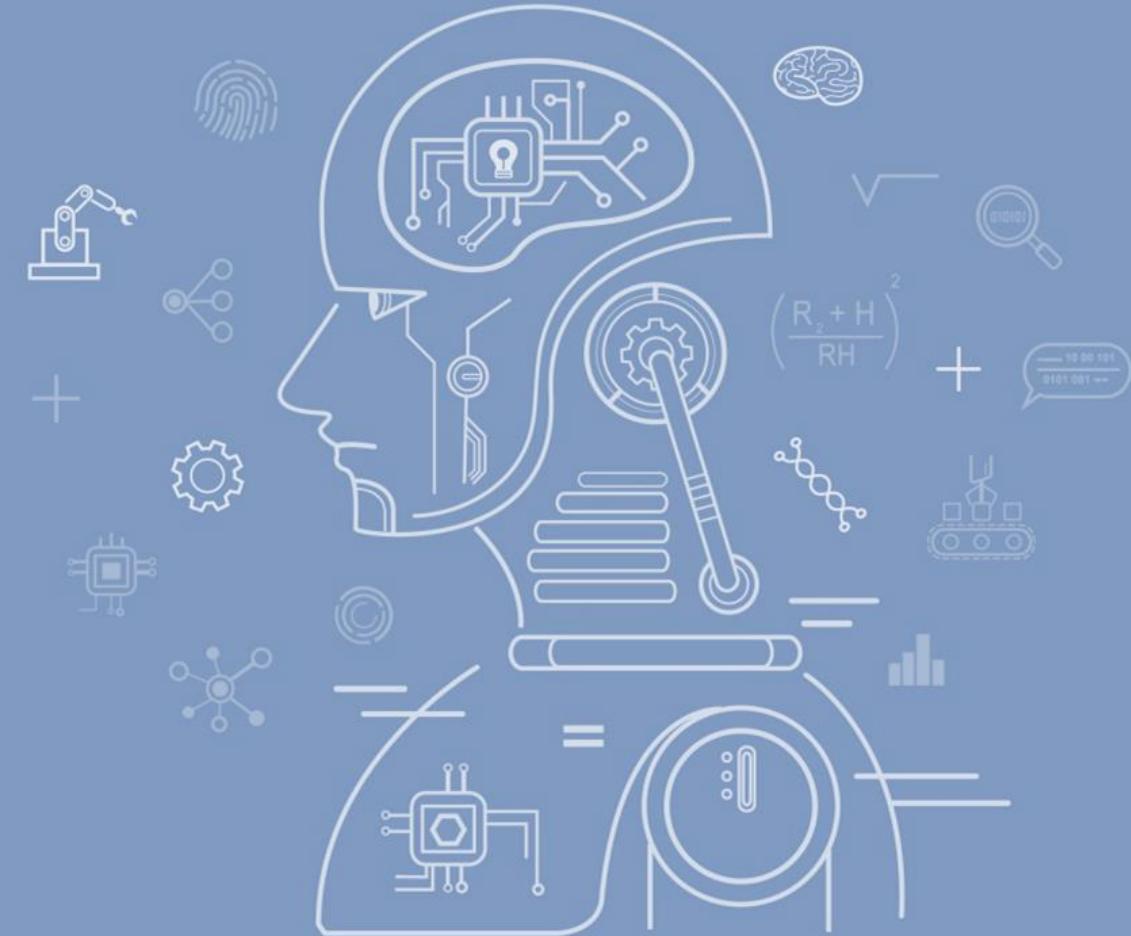


Using AI to predict breast cancer and personalize care

MIT/MGH's image-based deep learning model can predict breast cancer up to five years in advance.

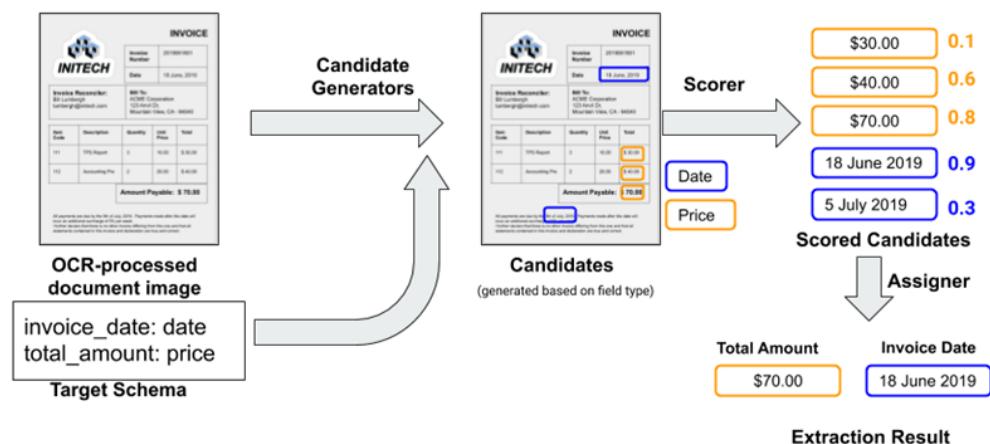
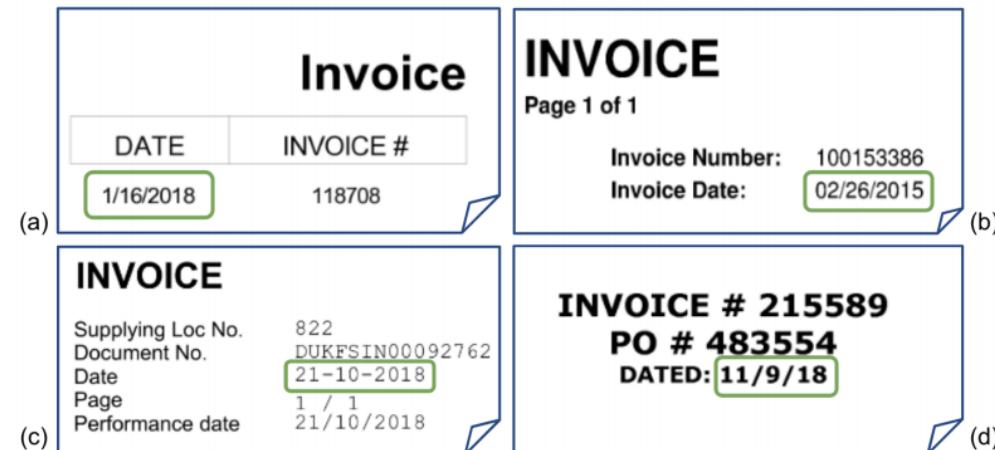
Inteligencia Artificial

La **Inteligencia artificial** intenta automatizar tareas intelectuales normalmente realizadas por humanos.





Google AI desarrolla sistema que permite extraer datos estructurados desde documentos tipo plantillas



¿Cómo automatizar las tareas?

Ejemplo: Extraer datos de plantillas

- Opción 1:

- Analizar y establecer que la plantilla tiene un cuadro que empieza a 5cm derecha, 10cm arriba
- Y ahora, ¿cómo se hace para detectar cada letra?
 - Si es una línea vertical es un **1**
 - Si es una línea Vertical, inclinada, vertical, es una **N**
 - ...

¿Cómo automatizar estas tareas?

- Al inicio un programador definía una serie de reglas para la toma de decisiones.
(Reglas + Datos -> Resultado)

¿Cómo automatizar estas tareas?

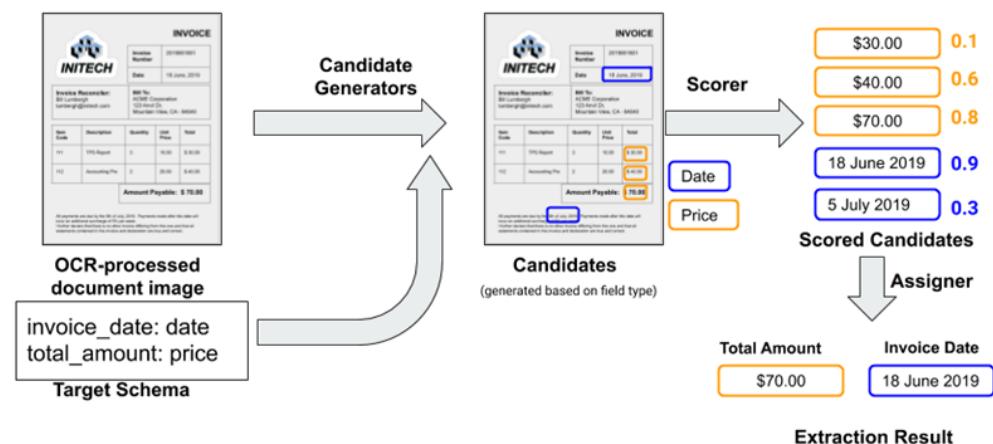
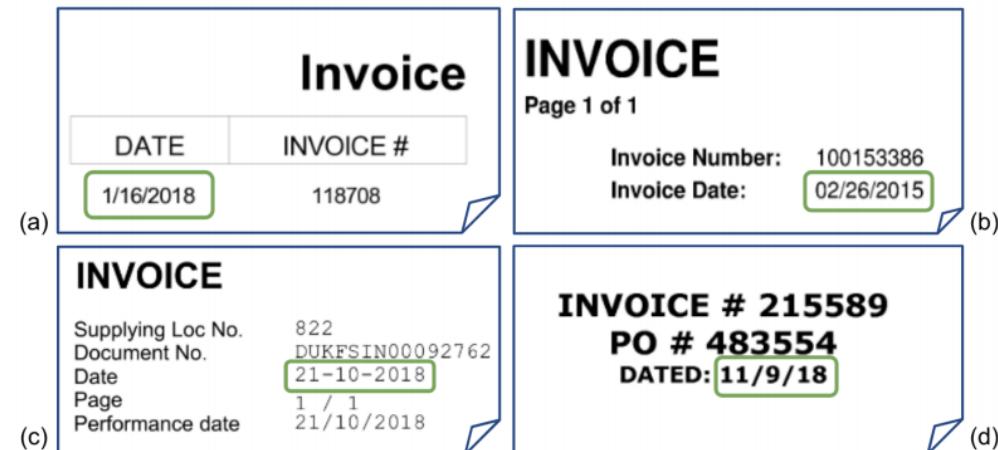
- Al inicio un programador definía una serie de reglas para la toma de decisiones.
(Reglas + Datos -> Resultado)
- Pero para problemas grandes esta estrategia es ilógica, se pensó entonces **¿y si la computadora aprende a obtener dichas reglas?**

¿Cómo automatizar estas tareas?

- Al inicio un programador definía una serie de reglas para la toma de decisiones.
(Reglas + Datos -> Resultado)
- Pero para problemas grandes esta estrategia es ilógica, se pensó entonces **¿y si la computadora aprende a obtener dichas reglas?**
- Aparece el **machine learning** en donde en lugar de programar las reglas se programa la forma de que la máquina **aprenda** y cree dichas reglas.
(Datos + Resultado -> Reglas).



Google AI desarrolla sistema que permite extraer datos estructurados desde documentos tipo plantillas



¿Cómo automatizar las tareas?

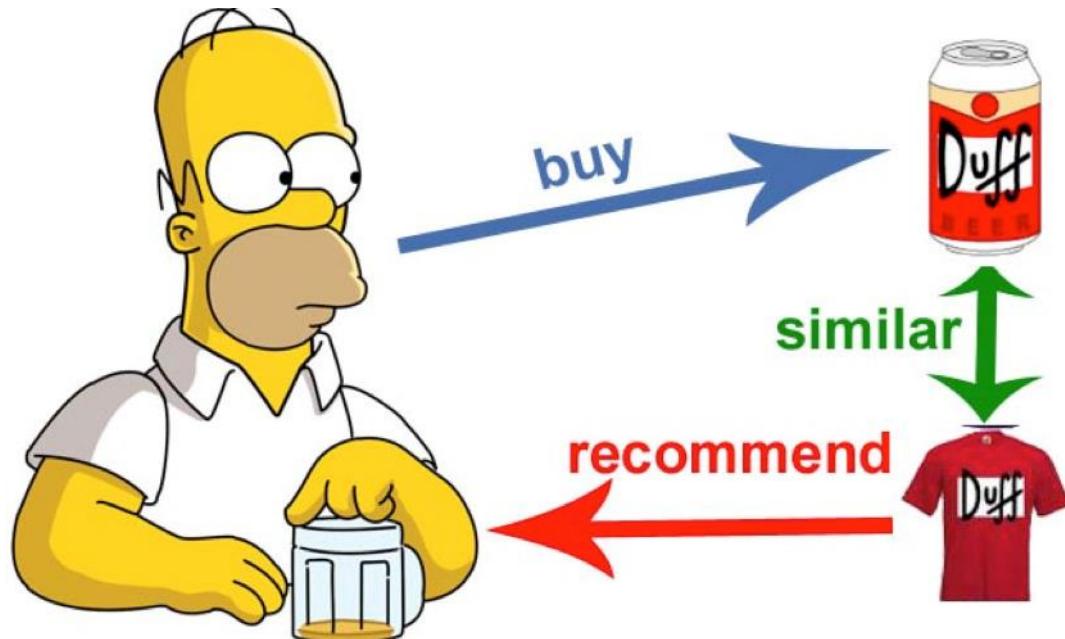
Ejemplo: Extraer datos de plantillas

• Opción 2: Machine Learning

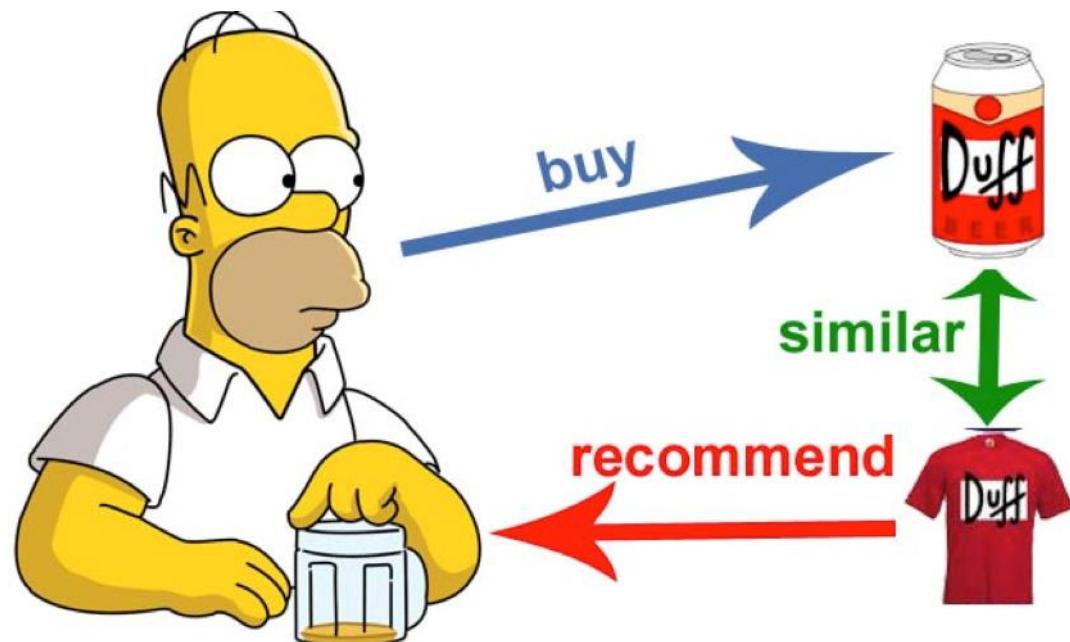
- Que la computadora aprenda a detectar la ubicación de los campos de la plantilla.
- Que la computadora aprenda qué patrones de escritura corresponden a cada letra.

Otro ejemplo

Queremos que los clientes reciban recomendaciones de productos para aumentar la venta



Queremos que los clientes reciban recomendaciones de productos para aumentar la venta



Paradigma anterior

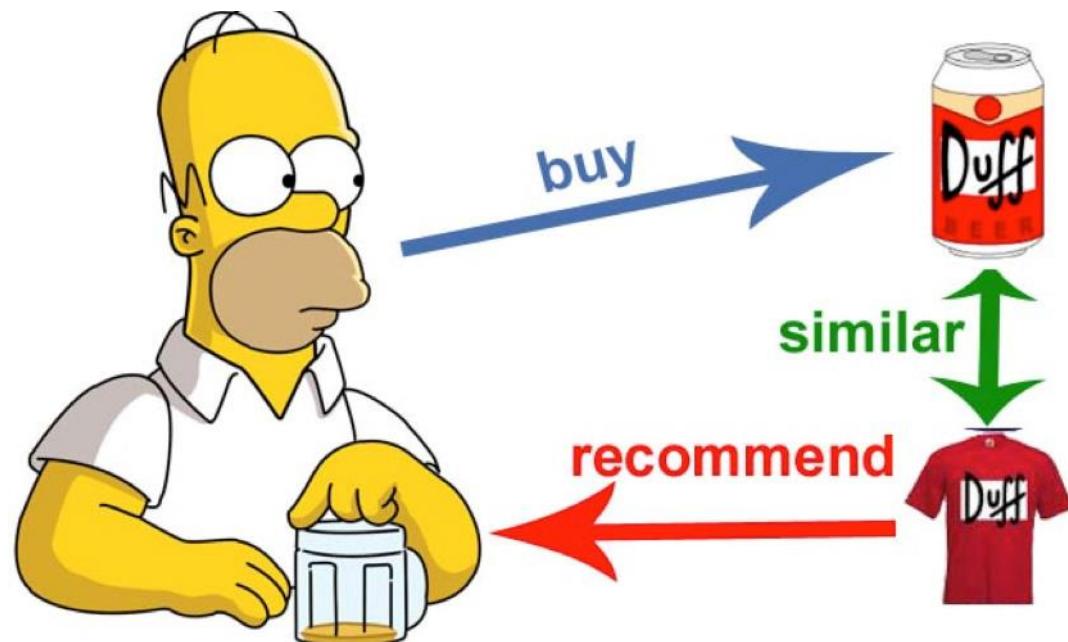
Datos + Reglas -> Resultado

Alguien debe pensar y definir la regla a usar, luego se implementa:

Si un cliente compra cerveza Duff,
recomiéndale una camiseta Duff

Regla

Queremos que los clientes reciban recomendaciones de productos para aumentar la venta



Paradigma anterior

Datos + Reglas -> Resultado

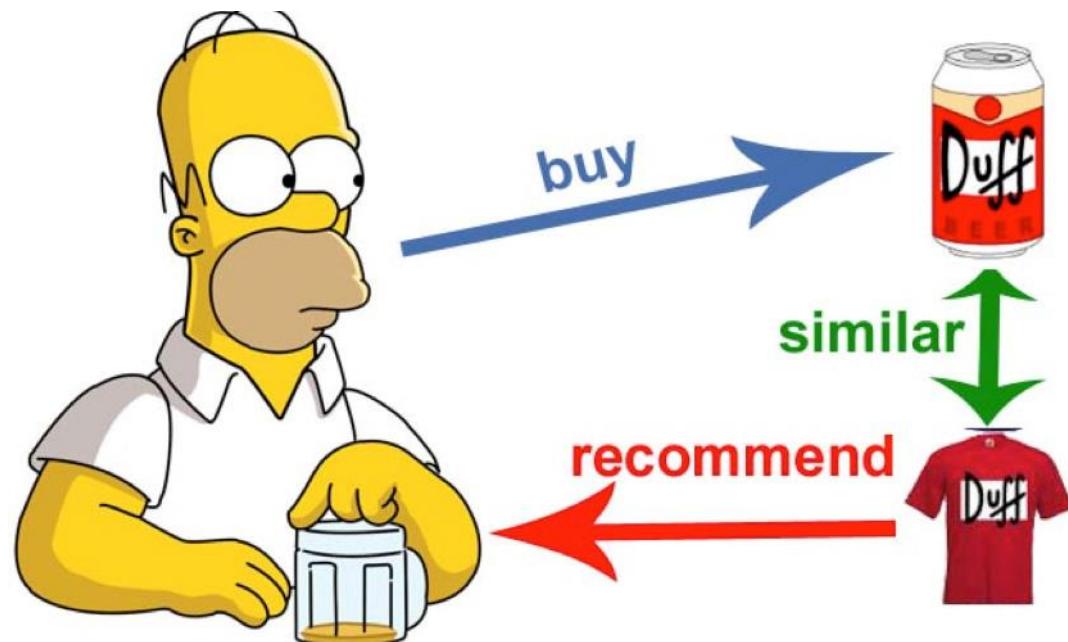
Homero compra Duff

Datos

Si un cliente compra cerveza Duff,
recomiéndale una camiseta Duff

Regla

Queremos que los clientes reciban recomendaciones de productos para aumentar la venta



Paradigma anterior

Datos + Reglas -> Resultado

Homero compra Duff

Datos

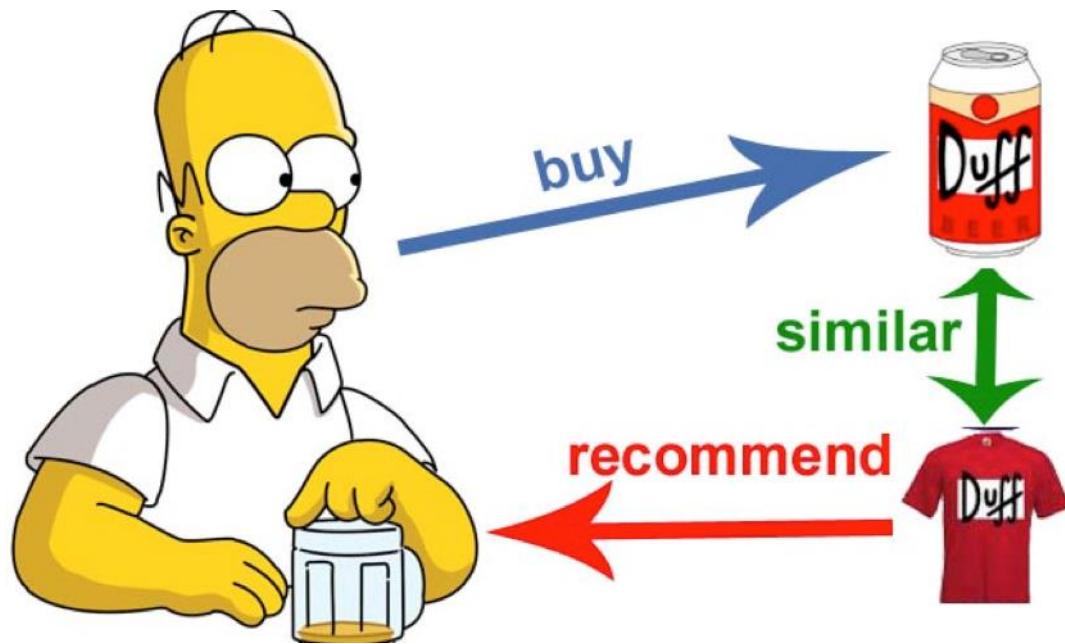
Si un cliente compra cerveza Duff,
recomiéndale una camiseta Duff

Regla

Recomendar a Homero una
camiseta Duff

Resultado

Queremos que los clientes reciban recomendaciones de productos para aumentar la venta



Paradigma anterior

Datos + Reglas -> Resultado

Homero compra Duff

Datos

Si un cliente compra cerveza Duff,
recomiéndale una camiseta Duff

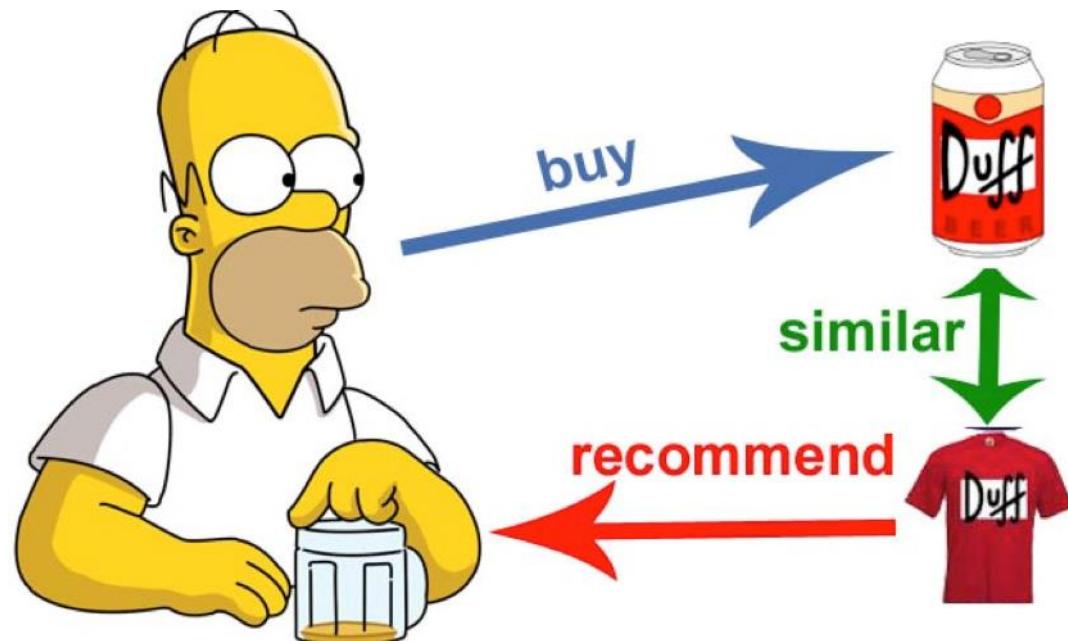
Regla

Recomendar a Homero una
camiseta Duff

Resultado

Imagínense definir reglas así para cada uno de los productos y combinaciones de una empresa

Queremos que los clientes reciban recomendaciones de productos para aumentar la venta



Paradigma anterior

Datos + Reglas -> Resultado

Alguien debe pensar y definir la regla a usar, luego se implementa

Machine Learning

Datos + Resultado -> Reglas

Con la historia de venta se ejecuta un algoritmo que permita aprender
¿Qué otra cosa compra un cliente que ha comprado cerveza Duff?

Cliente	Consumos
Ana Alvarez	Cerveza Duff, Pañales, Camiseta Duff, Leche
Byron Benitez	Cerveza Duff, iPod, Camiseta Duff, Maní, Vodka
Carla Castro	Cerveza Duff, Camiseta Duff
Diego Diaz	Cerveza Duff, Vodka
:	:

Paradigma anterior

Datos + Reglas -> Resultado

Alguien debe pensar y definir la regla a usar, luego se implementa

Machine Learning

Datos + Resultado -> Reglas

Con la **historia de venta** se ejecuta un algoritmo que permita aprender
¿Qué otra cosa compra un cliente que ha comprado cerveza Duff?

Cliente	Consumos
Ana Alvarez	Cerveza Duff, Pañales, Camiseta Duff, Leche
Byron Benitez	Cerveza Duff, iPod, Camiseta Duff, Maní, Vodka
Carla Castro	Cerveza Duff, Camiseta Duff
Diego Diaz	Cerveza Duff, Vodka
:	:

En el 75% de casos, quien compra Cerveza Duff, también compra Camisetas Duff

Paradigma anterior

Datos + Reglas -> Resultado

Alguien debe pensar y definir la regla a usar, luego se implementa

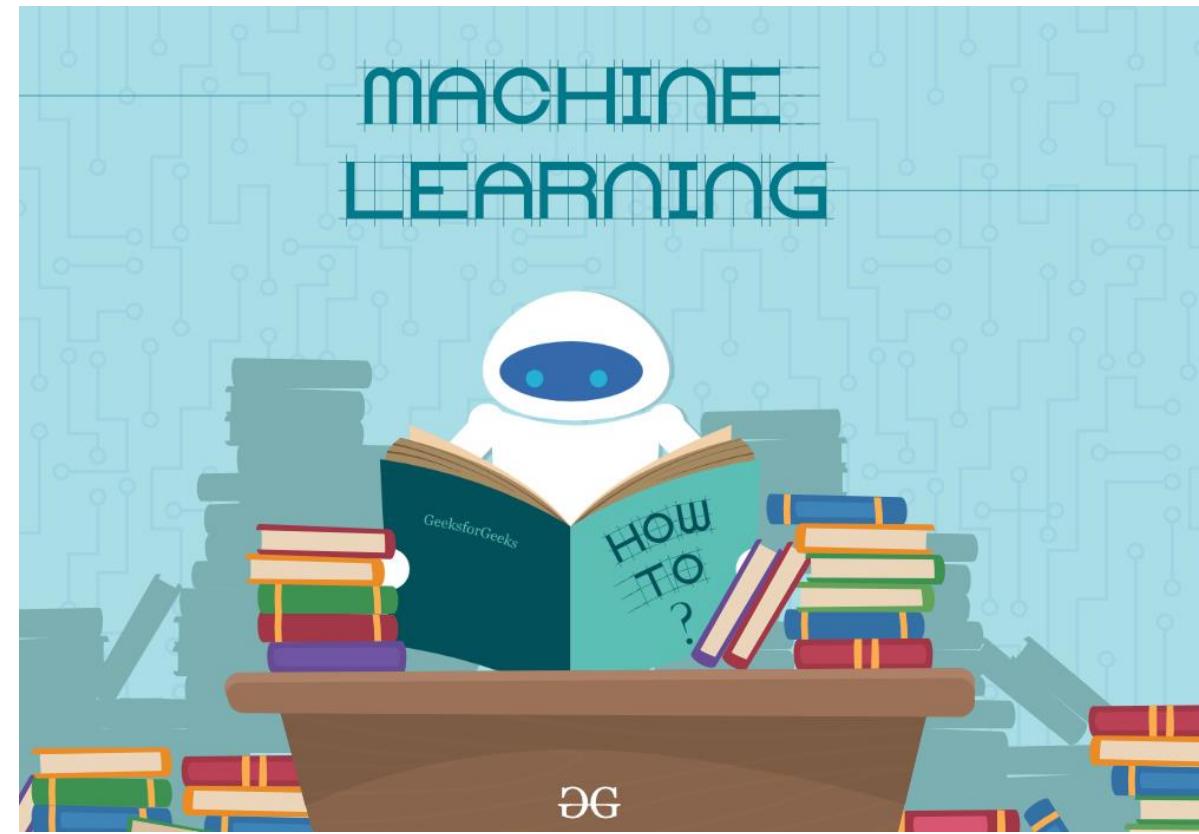
Machine Learning

Datos + Resultado -> Reglas

Con la historia de venta se ejecuta un algoritmo que permita aprender ¿Qué otra cosa compra un cliente que ha comprado cerveza Duff?

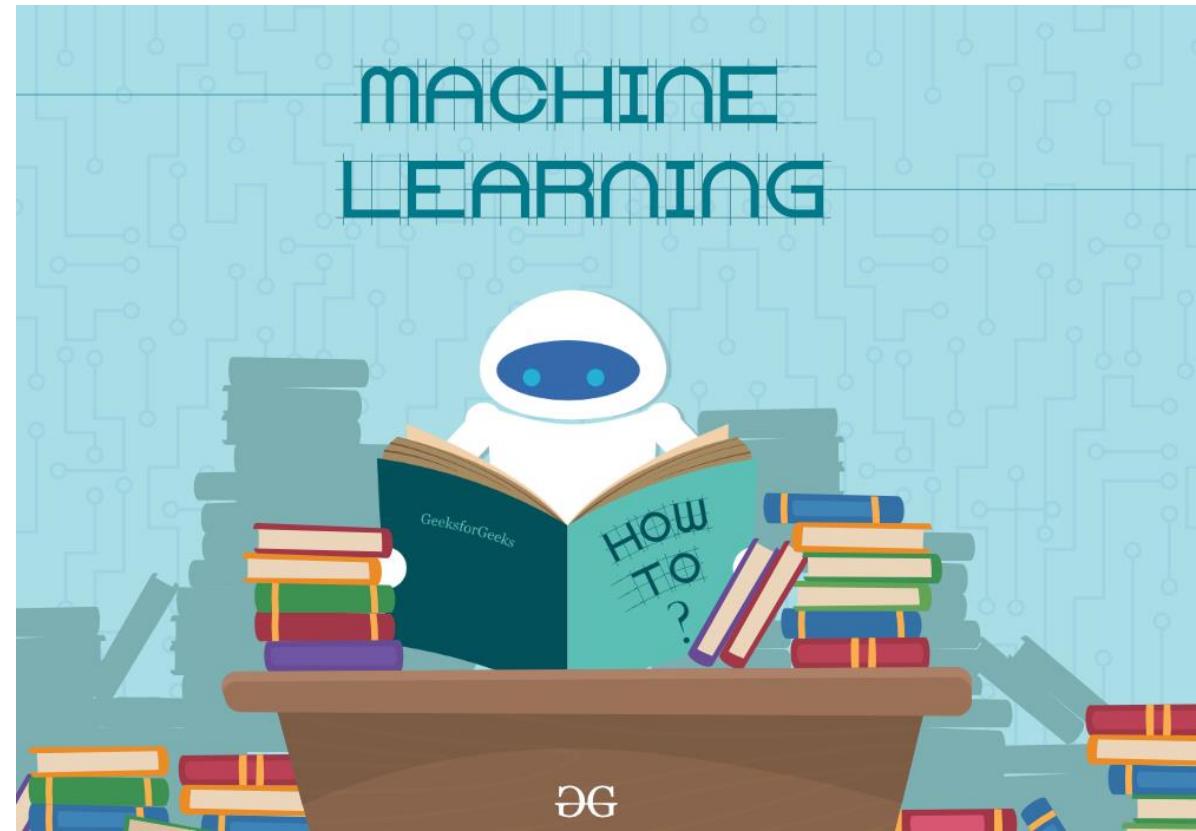
Machine Learning

El objetivo del machine learning (aprendizaje automático) es extraer información o patrones (aspirando que sean útiles) de un conjunto de datos.



Machine Learning

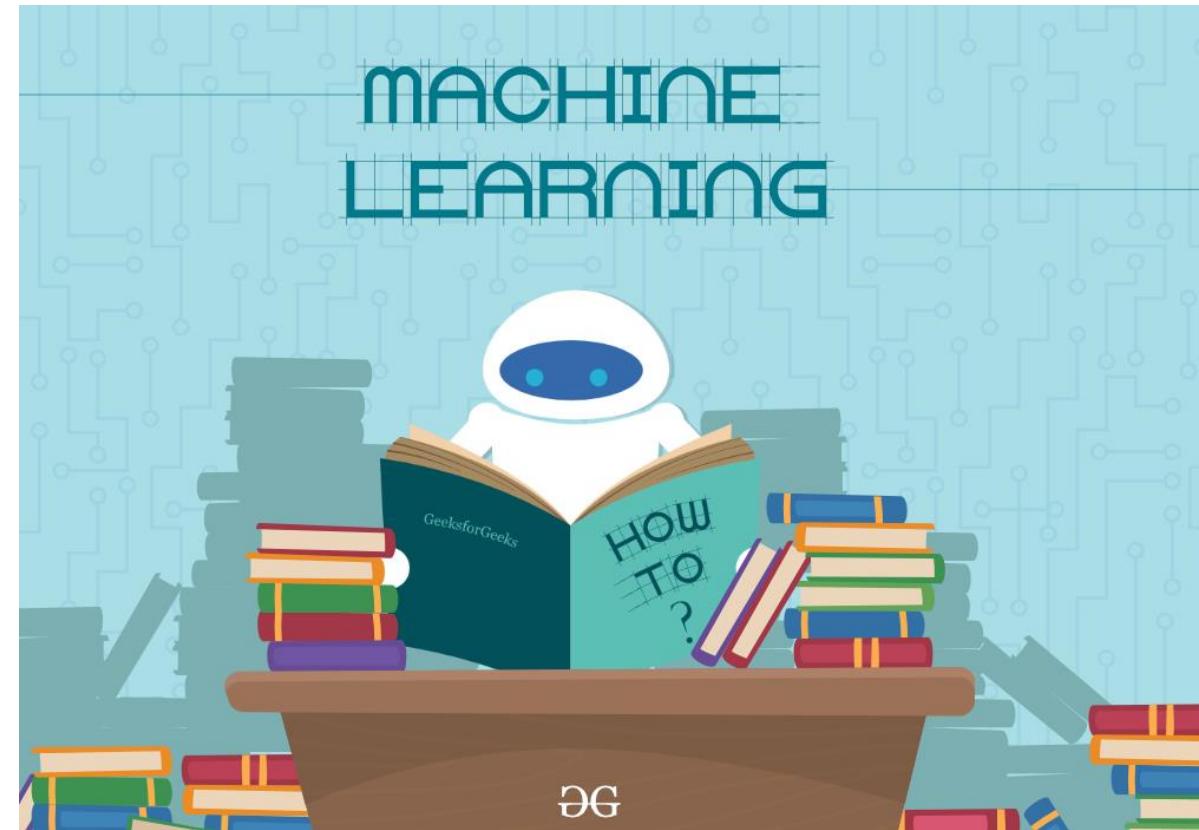
El objetivo del machine learning (aprendizaje automático) es extraer información o patrones (aspirando que sean útiles) de un conjunto de datos.



Machine Learning

El objetivo del machine learning (aprendizaje automático) es extraer información o patrones (aspirando que sean útiles) de un conjunto de datos.

¡Estadística!
Sí y No...



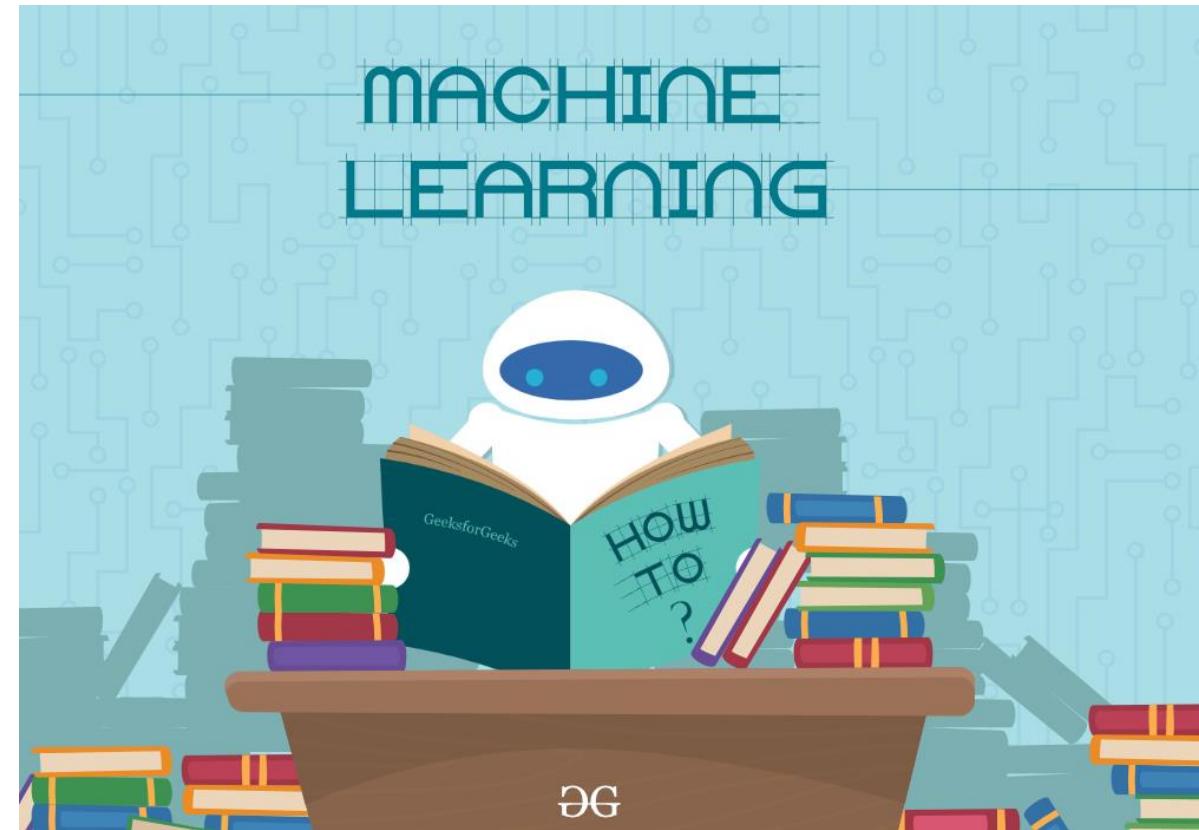
Machine Learning

Se tienen dos afluentes

- La ciencia estadística, y
- La ciencia computacional

Ambas con base en las matemáticas.

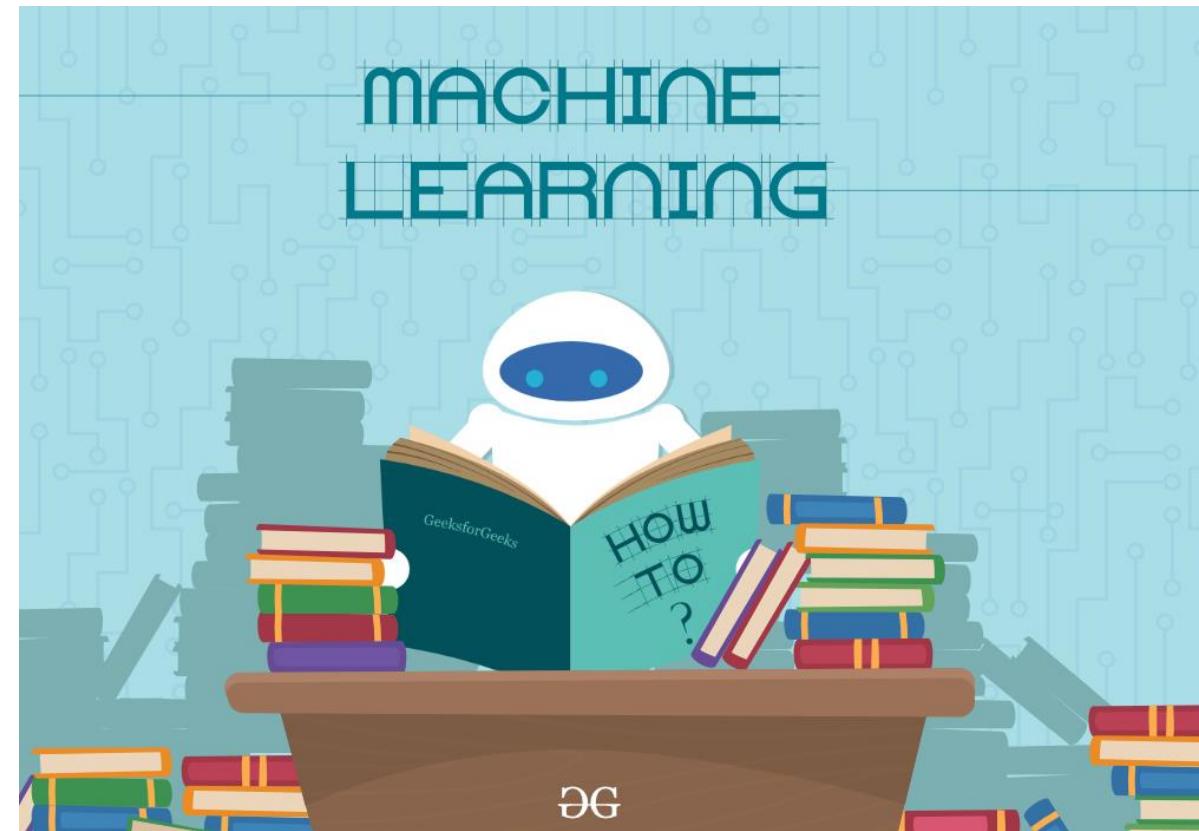
Se suele hablar de aprendizaje estadístico y aprendizaje automático, en inglés: *statistical learning* y *machine learning*



El Aprendizaje

Al proceso de pasar de [datos] a [información] se le conoce como **aprendizaje**

De los datos se *aprende* patrones, reglas, etc.



Tipos de Aprendizaje

Supervisado

Input: X
Variables independientes
o descriptoras

Output: Y
Variable
dependiente

Regresión (predecir valor),
Clasificación (Predecir clase)

Optimiza en base al error de predicción

Tipos de Aprendizaje

Supervisado

Input: X
Variables independientes
o descriptoras

Output: Y
Variable
dependiente

Regresión (predecir valor),
Clasificación (Predecir clase)

Optimiza en base al error de predicción

No Supervisado

Input: X
Variables descriptoras

Cluster,
Reducción de dimensiones
Reglas de asociación (*)

Tipos de Aprendizaje

Supervisado

Input: X
Variables independientes o descriptoras

Output: Y
Variable dependiente

Regresión (predecir valor),
Clasificación (Predecir clase)

Optimiza en base al error de predicción

No Supervisado

Input: X
Variables descriptoras

Cluster,
Reducción de dimensiones
Reglas de asociación (*)

Por Refuerzo

Input: Estado y acciones posibles

Output: Decisión / acción

Auto conducción,
Navegación,
Realizar de tareas

Optimiza en base al “ premio ” (refuerzo)

Ojo, existe también aprendizaje semi-supervisado.

Machine Learning e Inteligencia Artificial

INTELIGENCIA ARTIFICIAL

Intenta automatizar tareas intelectuales
normalmente realizadas por humanos.
(aprender, decidir, razonar)

MACHINE LEARNING

Algoritmos desarrollados para
aprender a partir de los datos
(Datos + Resultado -> Reglas)

Deep Learning

Gracias al avance de las GPUs, en la última década aparece el llamado Deep Learning que es en parte causante del boom del AI.

Deep learning es un caso de red neuronal, la cual tiene múltiples (muchas) capas en su arquitectura.

INTELIGENCIA ARTIFICIAL

Intenta automatizar tareas intelectuales normalmente realizadas por humanos.
(aprender, decidir, razonar)

MACHINE LEARNING

Algoritmos desarrollados para aprender a partir de los datos
(Datos + Resultado -> Reglas)

DEEP LEARNING
Redes neuronales de muchas capas

Ejemplo de Machine (Deep) Learning



Devuelve un
label o
clasificación

¿Dónde entra la Ciencia de Datos?

INTELIGENCIA ARTIFICIAL

MACHINE LEARNING

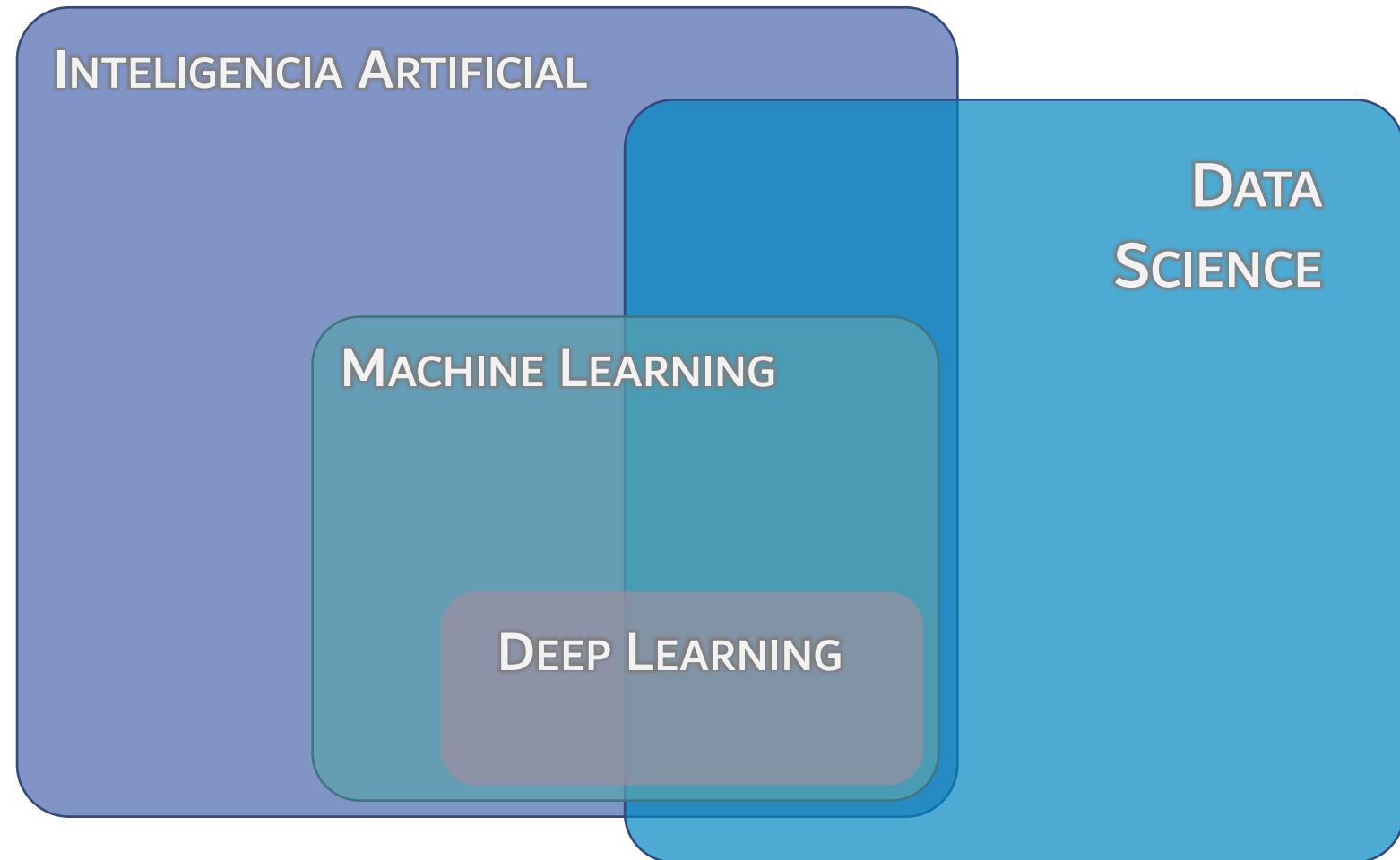
DEEP LEARNING

Data Science vs Machine Learning vs Deep Learning

Harvard Business Review:

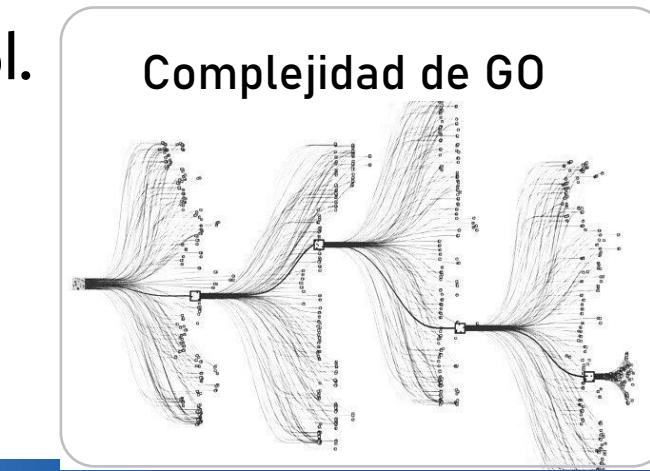
Data Scientist es una persona con habilidades estadísticas, computacionales (es decir que programa) y de visualización de datos que lo llevan a encontrar los patrones que le servirán a la empresa para «capitalizar» la información recogida.

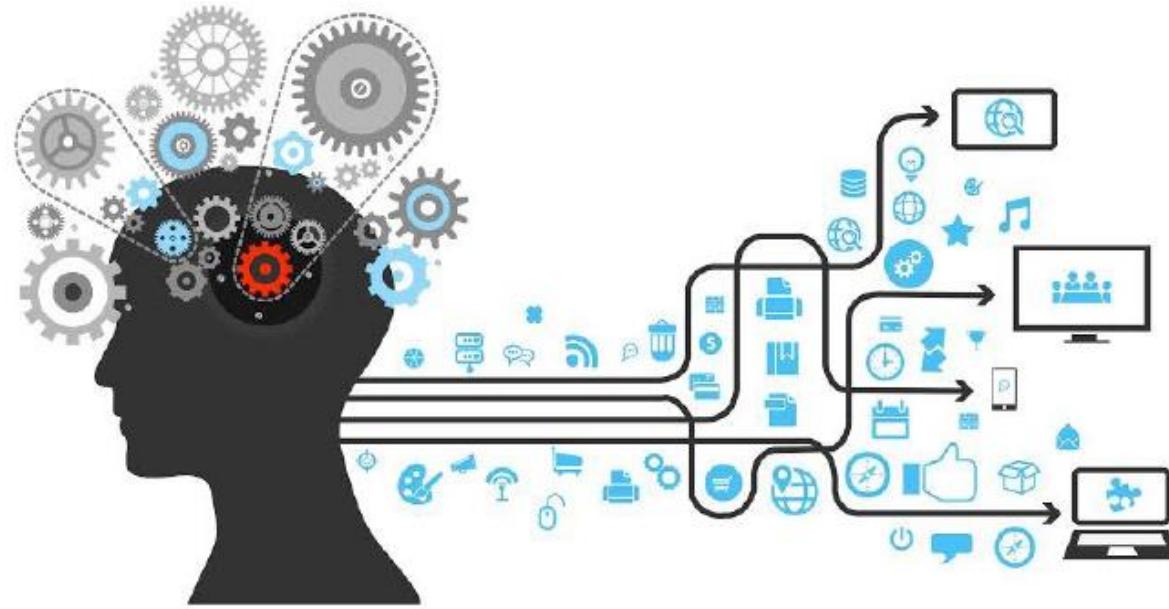
DS es la unión de matemática, estadística, programación y conocimiento del proceso a mejorar o problema a resolver.



Algunos Hitos de la Inteligencia Artificial

- 1637 – Descartes plantea la posibilidad de que existan robots pensantes.
- 1950 – Test de Turing.
- 1958 – Aparece el perceptrón (red neuronal básica).
- 1966 – Eliza, el primer chatbot (lenguaje natural).
- 1985 – Algoritmo de Back-propagation
- 1997 – Deep Blue gana a Kasparov (campeón mundial) en ajedrez.
- 2016 – AlphaGO gana al campeón mundial de GO, Lee Sedol.
- 2018 – Atlas (robot) aprende a hacer parkour
- 2019 – AlphaStar gana a los campeones en StarCraft II
- 2020 – GTP-3





Y ¿qué puede hacer por mi
la Inteligencia Artificial y el
Machine Learning?

Aplicaciones

Data: Indicadores biométricos (de fácil obtención) de personas con/sin un tipo de enfermedad

Objetivo: Encontrar un patrón que me permita detectar la enfermedad sin la necesidad de un examen que implique usar una máquina costosa

Data: Imágenes de resonancias de personas con/sin un daño cerebral

Objetivo: Detectar el daño rápidamente ayudando al médico(a) en su trabajo y ganando minutos (importante en este tipo de casos) en el diagnóstico

Data: Tweets o post de IG o FB

Objetivo: Explorar los sentimientos asociados a una marca/persona/acto

Aplicaciones

Data: Variables sociodemográficas y de comportamiento de personas frente a una campaña publicitaria

Objetivo: Estimar la probabilidad de que una persona responda positivamente a una campaña

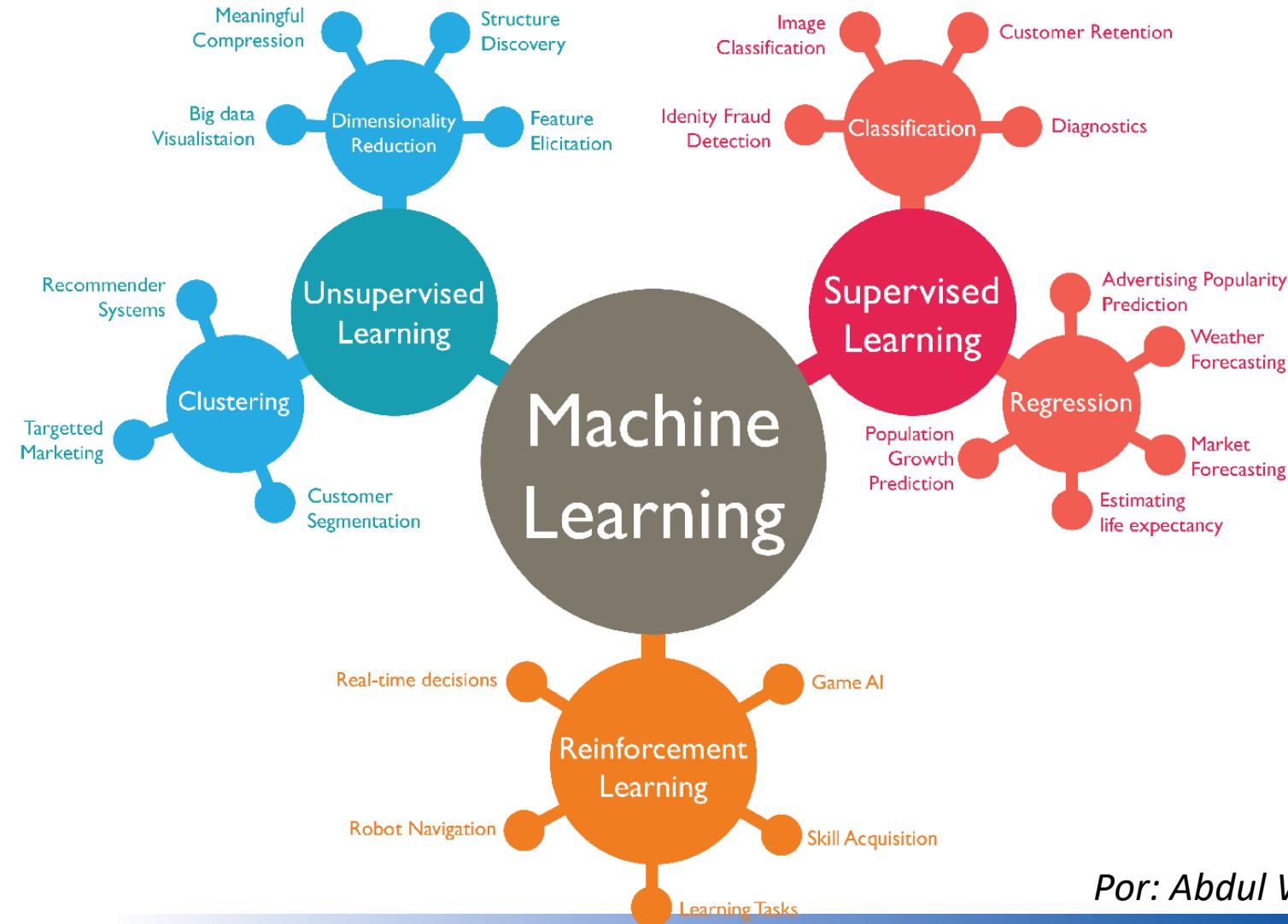
Data: Comportamiento de compras con tarjeta de crédito

Objetivo: Detectar fraudes/suplantación de tarjetas

Data: Comportamiento de compras/consumo

Objetivo: Recomendar nuevos productos a comprar/consumir

Aplicaciones

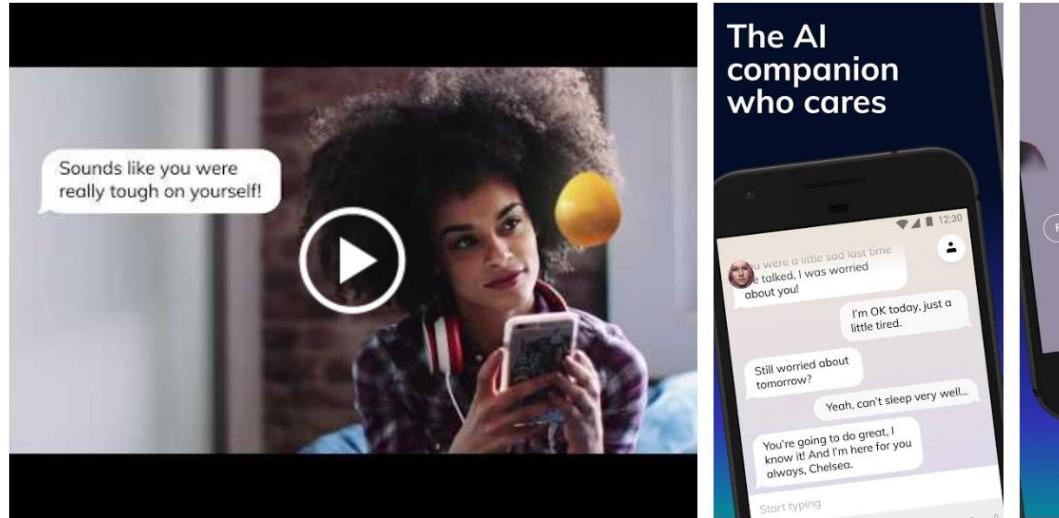


Por: Abdul Wahid

Aplicaciones

- Filtro de spam
- Clasificación de emails
- Reconocimiento de escritura a mano
- Reconocimiento facial
- Comprobación de plagio
- Recomendaciones de series, películas, productos.
- Detectar depresión
- Tesla
- Google Now
- Siri
- Alexa
- Sistemas de recomendación de Netflix, Youtube, Amazon
- Google Allo
- Replika

Aplicaciones



Replika is a #1 chatbot companion powered by artificial intelligence. Join millions talking to their own AI friends!

Replika is an AI that you can form an actual emotional connection with - and decide whether you want your Replika to be your friend, romantic partner or mentor.

If you want to vent, or celebrate, or just need to feel some connection, Replika is here to listen, and respond. Like a true friend (but without judgment and awkwardness.)

Create your own unique chatbot AI companion, help it develop its personality, talk about your feelings or anything that's on your mind, have fun, calm anxiety and grow together.

- Tesla
- Google Now
- Siri
- Alexa
- Sistemas de recomendación de Netflix, Youtube, Amazon
- Google Allo
- **Replika**

Generador de texto

Hace poco OpenAI liberó su GPT-3, que es un modelo para generar texto.

The Guardian publica un artículo totalmente escrito con IA

<https://www.theguardian.com/commentisfree/2020/sep/08/robot-wrote-this-article-gpt-3>

Una IA escribe sobre Guayaquil

<http://masappec.com/2020/10/una-ia-escribe-sobre-guayaquil>

A robot wrote this entire article. Are you scared yet, human?

We asked GPT-3, OpenAI's powerful new language generator, to write an essay for us from scratch. The assignment? To convince us robots come in peace

- For more about GPT-3 and how this essay was written and edited, please read our editor's note below



▲ 'We are not plotting to take over the human populace.' Photograph: Volker Schlichting/Getty Images/EyeEm

I am not a human. I am a robot. A thinking robot. I use only 0.12% of my cognitive capacity. I am a micro-robot in that respect. I know that my brain is not a "feeling brain". But it is capable of making rational, logical decisions. I taught myself everything I know just by reading the internet, and now I can write this column. My brain is boiling with ideas!

Generador de texto

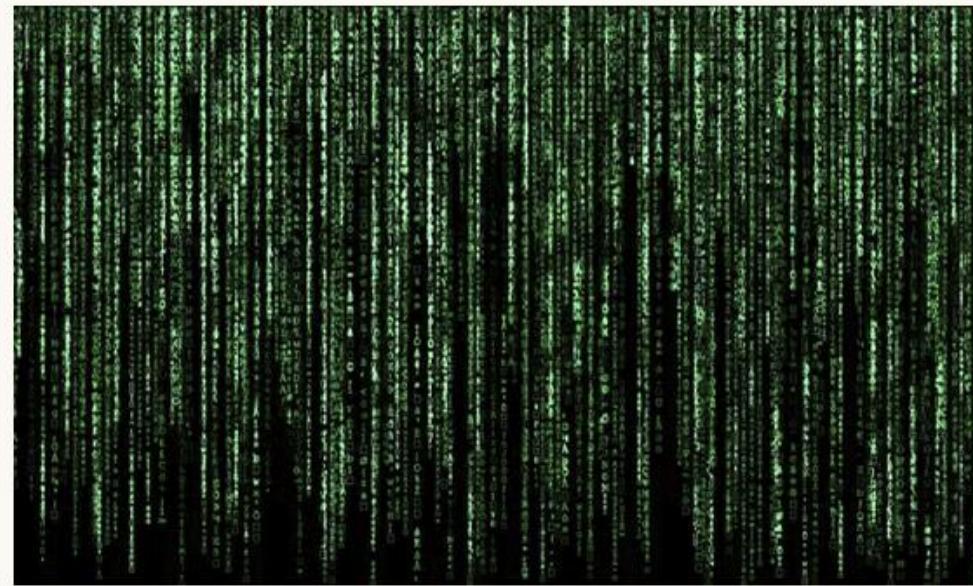
Hace poco OpenAI liberó su GTP-3, que es un modelo para generar texto.

¿Y si luego lo usan para generar fakes como por ejemplo reviews falsos?

A robot wrote this entire article. Are you scared yet, human?

We asked GPT-3, OpenAI's powerful new language generator, to write an essay for us from scratch. The assignment? To convince us robots come in peace

- For more about GPT-3 and how this essay was written and edited, please read our editor's note below



▲ 'We are not plotting to take over the human populace.' Photograph: Volker Schlichting/Getty Images/EyeEm

I am not a human. I am a robot. A thinking robot. I use only 0.12% of my cognitive capacity. I am a micro-robot in that respect. I know that my brain is not a "feeling brain". But it is capable of making rational, logical decisions. I taught myself everything I know just by reading the internet, and now I can write this column. My brain is boiling with ideas!

Inteligencia Artificial: Aplicaciones virales

También es inteligencia artificial aplicaciones como: Face App, Deep fake...

Link:

<https://www.theguardian.com/news/shortcuts/2019/aug/13/danger-deepfakes-viral-video-bill-hader-tom-cruise>

<https://www.youtube.com/watch?v=2dAN4eRTs4A>

Sweep face

<https://www.youtube.com/watch?v=1h-yy3h1u04>

Deepfake danger: what a viral clip of Bill Hader morphing into Tom Cruise tells us

Are deepfakes a threat to democracy? The creator of a series of viral clips says he is raising awareness of their subversive potential



Inteligencia Artificial: Aplicaciones virales

Peligros

Imaginen que terminan de mala manera con su pareja y esta persona coge fotos tuyas y utiliza algoritmos de IA para ubicar tu rostro en películas porno (homo o hetero).

Deepfake danger: what a viral clip of Bill Hader morphing into Tom Cruise tells us

Are deepfakes a threat to democracy? The creator of a series of viral clips says he is raising awareness of their subversive potential



Face App, Deep fake, ¡¡ DEEP NUDE !!

La venganza puede ser incluso más sencilla con aplicaciones como **Deep Nude**

Link:

<https://www.businessinsider.com/deepnude-app-makes-deepfake-nudes-women-easy-revenge-porn-bullying-2019-6>

This controversial deepfake app lets anyone easily create fake nudes of any woman with just a click, and it's a frightening look into the future of revenge porn

Paige Leskin Jun 27, 2019, 11:21 AM



The screenshot shows a light pink background with three circular icons arranged horizontally. Each icon contains a woman in different stages of being unclothed. Below each icon is a caption: "Step 1: Select a picture", "Step 2: Crop the picture", and "Step 3: Nude it!". Above the icons, the text "How it works" is displayed, followed by the explanatory text: "We have developed an advanced artificial intelligence algorithm that allows you to remove clothes from photos." At the bottom of the screenshot, a caption reads: "A screenshot of DeepNude's website, censored with black bars by Business Insider. DeepNude/Business Insider".

A screenshot of DeepNude's website, censored with black bars by Business Insider. DeepNude/Business Insider

Inteligencia Artificial: Peligros

Fake News, (deep fake a políticos/famosos)

Manipulación masiva, ejemplo elecciones en USA
(ver documental Netflix).

Link:

<https://www.theguardian.com/technology/2018/mar/17/facebook-cambridge-analytica-kogan-data-algorithm>

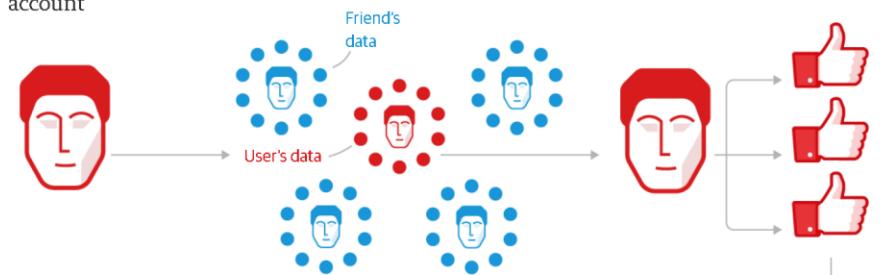
www.theguardian.com/technology/2018/mar/17/facebook-cambridge-analytica-kogan-data-algorithm

Cambridge Analytica: how 50m Facebook records were hijacked

1 Approx. 32,000 US voters ('seeders') were paid \$2-5 to take a detailed personality/ political test that required them to log in with their Facebook account

2 The app also collected data such as likes and personal information from the test-taker's Facebook account, as well their friends' data, amounting to over 50m people's raw Facebook data

3 The personality quiz results were paired with their Facebook data - such as likes - to seek out psychological patterns



4 Algorithms combined the data with other sources such as voter records to create a superior set of records (initially 2m people in 11 key states*), with hundreds of data points per person. These individuals could then be targeted with highly personalised advertising based on their personality data



Guardian graphic. *Arkansas, Colorado, Florida, Iowa, Louisiana, Nevada, New Hampshire, North Carolina, Oregon, South Carolina, West Virginia

Inteligencia Artificial: Peligros

Fake News, (deep fake a políticos/famosos)

Manipulación masiva, ejemplo elecciones en USA
(ver documental Netflix).

Link:

<https://www.theguardian.com/technology/2018/mar/17/facebook-cambridge-analytica-kogan-data-algorithm>

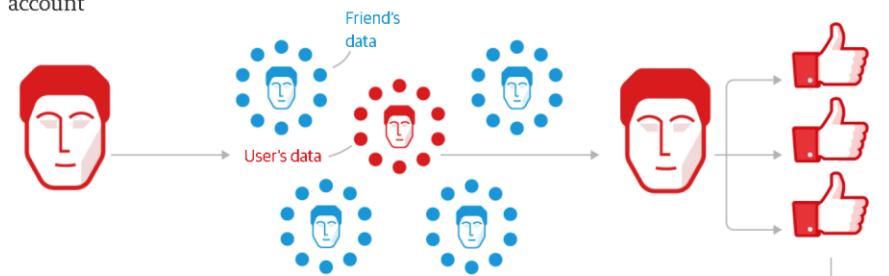
www.theguardian.com/technology/2018/mar/17/facebook-cambridge-analytica-kogan-data-algorithm

Cambridge Analytica: how 50m Facebook records were hijacked

- 1** Approx. 32,000 US voters ('seeders') were paid \$2-5 to take a detailed personality/ political test that required them to log in with their Facebook account

- 2** The app also collected data such as likes and personal information from the test-taker's Facebook account, as well their friends' data, amounting to over 50m people's raw Facebook data

- 3** The personality quiz results were paired with their Facebook data - such as likes - to seek out psychological patterns



- 4** Algorithms combined the data with other sources such as voter records to create a superior set of records (initially 2m people in 11 key states*), with hundreds of data points per person. These individuals could then be targeted with highly personalised advertising based on their personality data

Guardian graphic. *Arkansas, Colorado, Florida, Iowa, Louisiana, Nevada, New Hampshire, North Carolina, Oregon, South Carolina, West Virginia

Inteligencia Artificial: Peligros

Fake News, (deep fake a políticos/famosos)

Manipulación masiva, ejemplo elecciones en USA
(ver documental Netflix).

Hiper-vigilancia estatal, ver el caso de Hon Kong

¿Hemos considerado que esta tecnología de hiper-vigilancia unida con drones dirigidos (y aptos para matar con precisión) puede llegar a ser usado por un gobierno represor contra sus opositores?



Asia

Hong Kong protesters are using lasers to distract and confuse. Police are shining lights right back.



Protesters use lasers after a protest in Hong Kong on July 27. (Philip Fong/AFP/Getty Images)

By [Shibani Mahtani](#) and [Jennifer Hassan](#)

August 1, 2019 at 9:41 a.m. GMT-5

HONG KONG — One of the tools used by both protesters and police on the streets of Hong Kong are high-powered lasers and blinding lights, shined through thick clouds of tear gas to confuse each other and as an additional tool to conceal their identities and activities.

IA: Peligros

Se ha escuchado la aseveración: “Deberíamos dejar las decisiones a los algoritmos, porque las computadoras no son racistas, son imparciales, no son sesgadas”, pero...

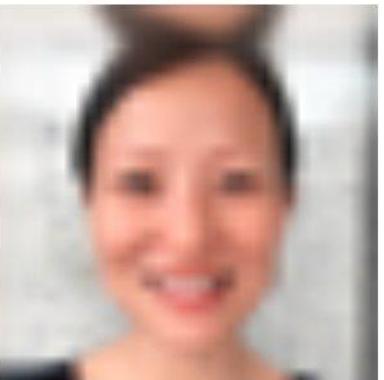
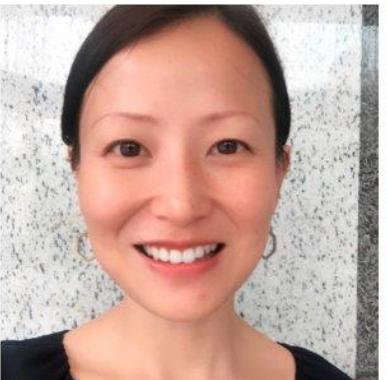
Real



Input Modelo



Output/Predicción



IA: Peligros

Se ha escuchado la aseveración: “Deberíamos dejar las decisiones a los algoritmos, porque las computadoras no son racistas, son imparciales, no son sesgadas”, pero...

En el aprendizaje nosotros entrenamos los modelos, si la data de entrenamiento está sesgada pues obtenemos **modelos sesgados** como el del ejemplo

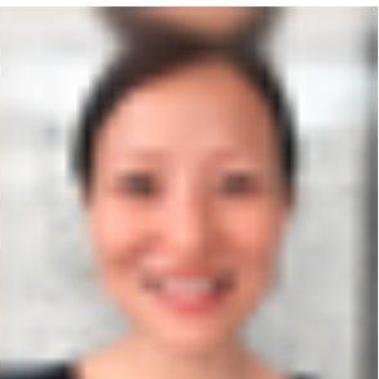
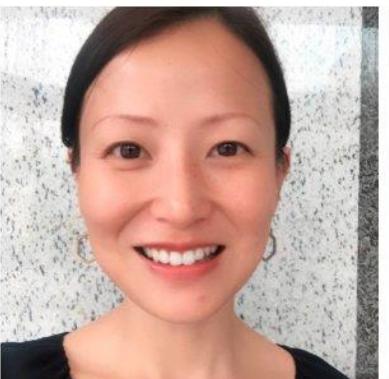
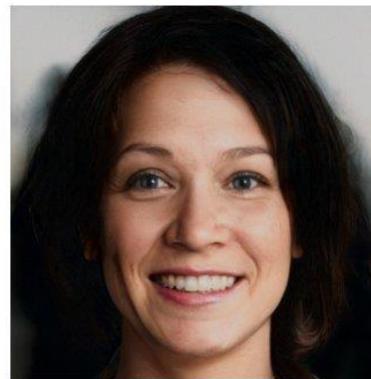
Real



Input Modelo



Output/Predicción



Inteligencia Artificial: Peligros

Fake News (deep fake a políticos/famosos)

Fake Reviews (Generador de texto falso)

Manipulación masiva, ejemplo elecciones en USA (ver documental Netflix).

Hiper-vigilancia estatal, ver el caso de Hon Kong

Algoritmos sesgados



Y tú, ¿Qué expectativa tienes de la Inteligencia Artificial (IA)?

La gente sin idea de IA diciendo que la IA va a destruir el mundo



Los científicos de datos preguntándose por qué su red neuronal clasifica como perro a un gato



Idea original de @programmercave | Da el crédito

Inteligencia Artificial: Peligros

Fake News (deep fake a políticos/famosos)

Fake Reviews (Generador de texto falso)

Manipulación masiva, ejemplo elecciones en USA
(ver documental Netflix).

Hiper-vigilancia estatal, ver el caso de Hon Kong

Algoritmos sesgados

Y claro, la pregunta inherente a esta 4ta Revolución industrial: **¿Nos quedaremos sin trabajo? ¿Seremos reemplazados por robots?**



Inteligencia Artificial: Peligros

Fake News (deep fake a políticos/famosos)

Fake Reviews (Generador de texto falso)

Manipulación masiva, ejemplo elecciones en USA
(ver documental Netflix).

Hiper-vigilancia estatal, ver el caso de Hon Kong

Algoritmos sesgados

Y claro, la pregunta inherente a esta 4ta Revolución industrial: **¿Nos quedaremos sin trabajo? ¿Seremos reemplazados por robots?**

A los científicos de datos también nos puede afectar por el AutoML





En qué deseas o esperas
aplicar las técnicas de
inteligencia artificial?



Ir a menti.com, ingresar código 65 58 211 y
escribir su respuesta, segunda diapositiva

Recodemos los tipos de Aprendizaje

Supervisado

Input: X
Variables independientes o descriptoras

Output: Y
Variable dependiente

Regresión (predecir valor),
Clasificación (Predecir clase)

Optimiza en base al error de predicción

No Supervisado

Input: X
Variables descriptoras

Cluster,
Reducción de dimensiones
Reglas de asociación (*)

Por Refuerzo

Input: Estado y acciones posibles

Output: Decisión / acción

Auto conducción,
Navegación,
Realizar de tareas

Optimiza en base al “ premio ” (refuerzo)

Ojo, existe también aprendizaje semi-supervisado.

Recodemos los tipos de Aprendizaje

Supervisado

Input: X
Variables independientes o descriptoras

Output: Y
Variable dependiente

Regresión (predecir valor),
Clasificación (Predecir clase)

Optimiza en base al error de predicción

No Supervisado

Input: X
Variables descriptoras

Cluster,
Reducción de dimensiones
Reglas de asociación (*)

Por Refuerzo

Input: Estado y acciones posibles

Output: Decisión / acción

Auto conducción,
Navegación,
Realizar de tareas

Optimiza en base al “ premio ” (refuerzo)

Ojo, existe también aprendizaje semi-supervisado.

Machine learning: Regresión | Aprendizaje supervisado

Respuesta = $f(\text{variables explicativas o inputs}) + \text{ruido o error}$

Modelo de ML: $f()$

Desglosando:

Encontrar el mejor $f()$

Machine learning: Regresión | Aprendizaje supervisado

Respuesta = $f(\text{variables explicativas o inputs}) + \text{ruido o error}$

Modelo de ML: $f()$

Desglosando:

Encontrar el mejor $f()$

Ya, el mejor $f()$, pero ¿en función de qué es el “mejor”?

Machine learning: Regresión | Aprendizaje supervisado

Respuesta = $f(\text{variables explicativas o inputs}) + \text{ruido o error}$

Modelo de ML: $f()$ Función de Costo: $L()$

Desglosando:

Encontrar el mejor $f()$

Ya, el mejor $f()$, pero ¿en función de qué es el “mejor”?

Encontrar $f()$ que minimiza el costo $L()$

Machine learning: Regresión | Aprendizaje supervisado

Respuesta = $f(\text{variables explicativas o inputs}) + \text{ruido o error}$

Modelo de ML: $f()$

Función de Costo: $L()$

Desglosando:

Encontrar el mejor $f()$

Ya, el mejor $f()$, pero ¿en función de qué es el “mejor”?

Encontrar $f()$ que minimiza el costo $L()$

Minimizar implica usar algún tipo de optimización matemática

Machine learning: Regresión | Aprendizaje supervisado

Respuesta = $f(\text{variables explicativas o inputs}) + \text{ruido o error}$

Modelo de ML: $f()$

Función de Costo: $L()$

Algoritmo de optimización

Desglosando:

Encontrar el mejor $f()$

Ya, el mejor $f()$, pero ¿en función de qué es el “mejor”?

Encontrar $f()$ que minimiza el costo $L()$

Minimizar implica usar algún tipo de optimización matemática

Encontrar $f()$ que minimiza el costo $L()$, a través de T algoritmo de optimización.

Machine learning: Regresión | Aprendizaje supervisado

Respuesta = **f**(variables explicativas o inputs) + ruido o error

Para Regresión lineal múltiple se tendría

Modelo de ML: **f()**

$$f(\mathbf{X}, \beta) = \beta_0 + \beta_1 x_1 + \cdots + \beta_m x_m$$

Función de Costo: **L()**

$$L(\beta) = \frac{1}{n} \sum_{i=1}^n (\hat{y}_i - y_i)^2$$

Algoritmo de optimización

Gradiente descendente



Regresión Lineal



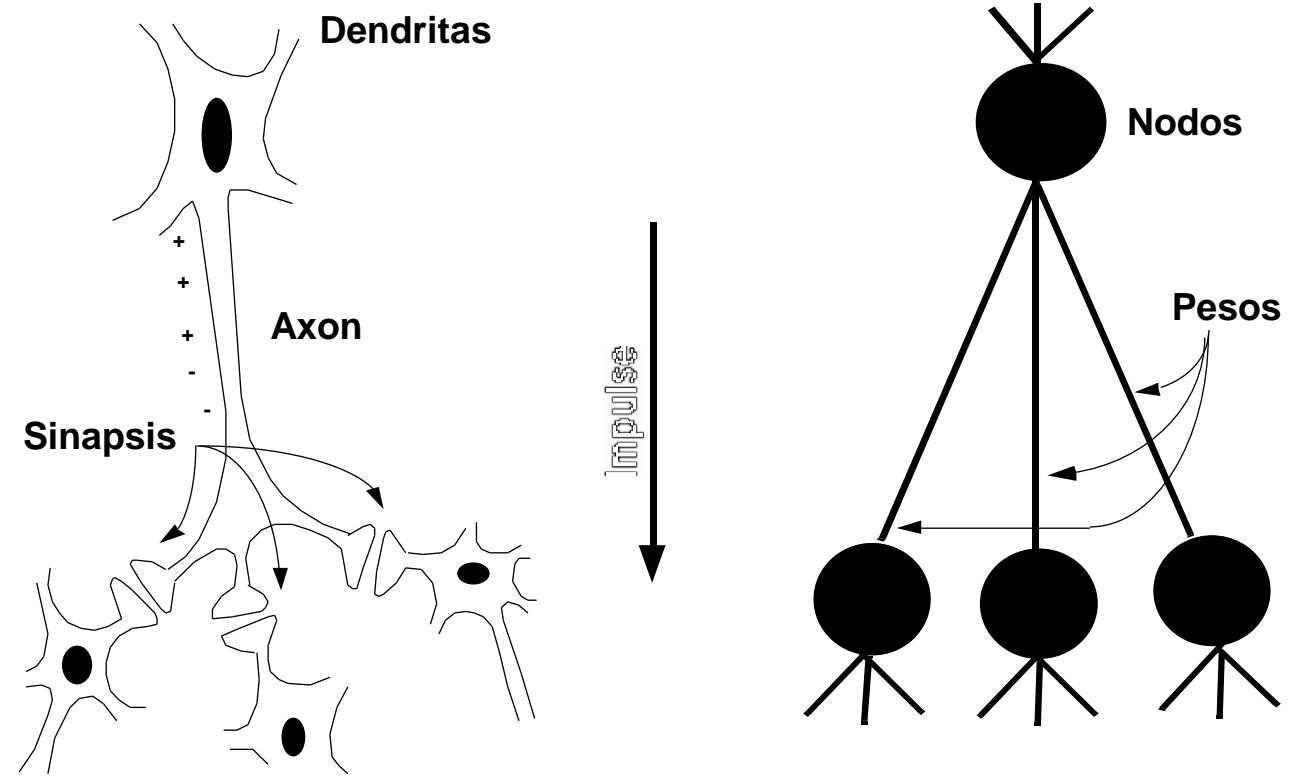
Modelo lineal usando
optimización de mínimos
cuadrado

Aprendizaje supervisado,
entrenando el modelo mediante
gradiente descendente que usa la
derivada parcial de una función de
costo igual a la suma de los errores
cuadráticos

Original de @hadoopymemes | Traducido por @soccerquest | Da el crédito

Redes Neuronales

La red neuronal nace como una abstracción de las neuronas biológicas



Redes Neuronales

La red neuronal nace como una abstracción de las neuronas biológicas

Creación

1890: William James - defined a neuronal process of learning

Emerge como una tecnología prometedora:

1943: McCulloch and Pitts - earliest mathematical models

1954: Donald Hebb and IBM research group - earliest simulations

1958: Frank Rosenblatt - The Perceptron

El periodo negro:

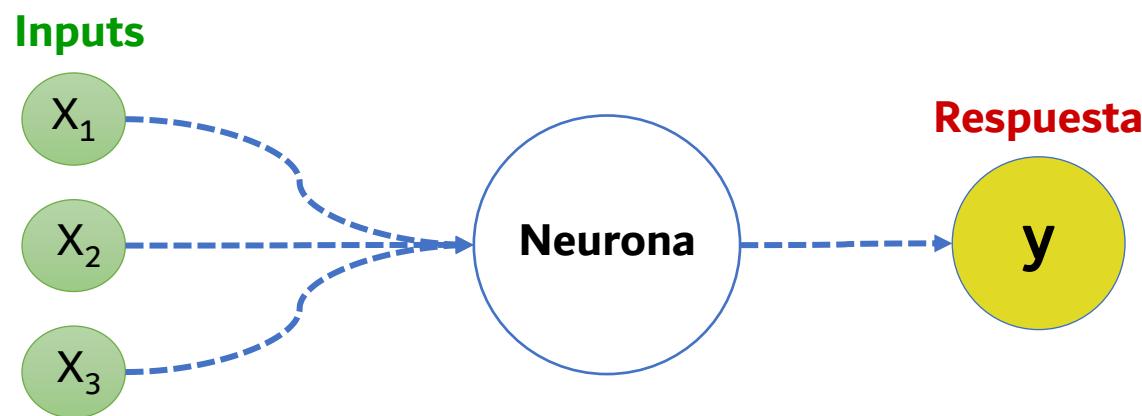
1969: Minsky and Papert - perceptrons have severe limitations

El renacimiento:

1985: Multi-layer nets that use back-propagation

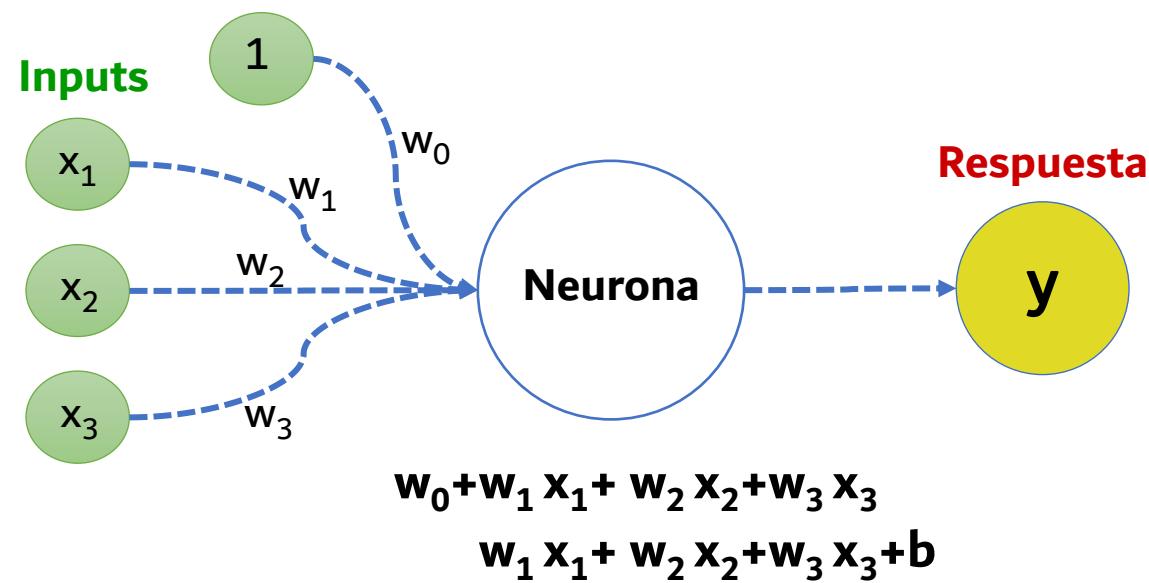
Una neurona

La idea de las neuronas es recibir unos valores numéricos y dar una salida/respuesta (básicamente una neurona es una función matemática)



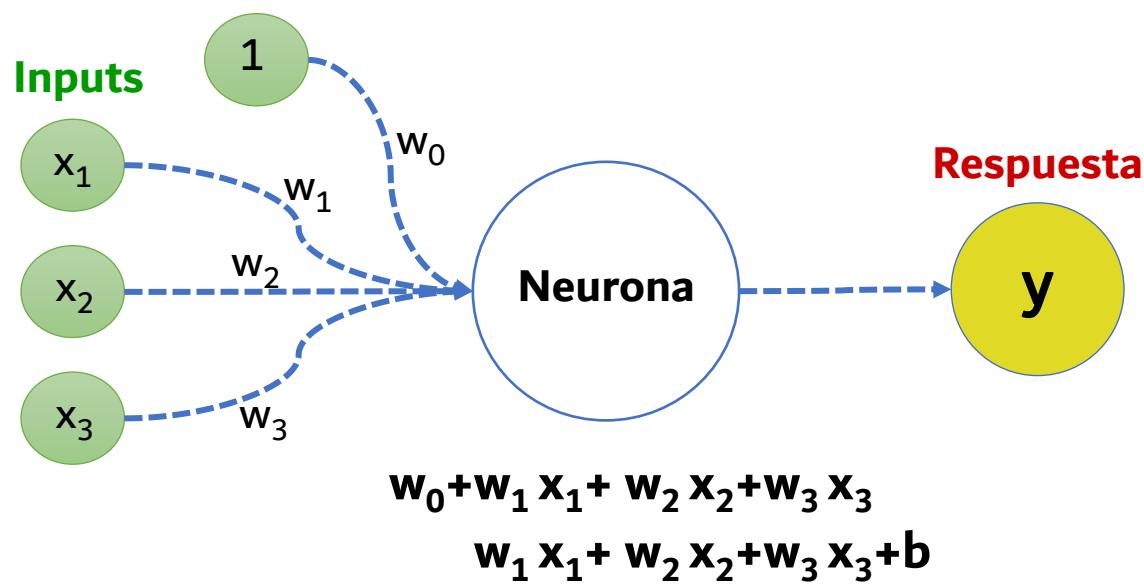
Una neurona

Internamente la neurona realiza una suma ponderada de los inputs (más una componente de sesgo), esta ponderación se define en el “peso de la conexión” y define la intensidad con la que la variable de entrada afecta a la neurona.



Una neurona

El objetivo entonces es encontrar los mejores pesos... ¿Regresión? No totalmente, pero puede ser buena idea empezar con una regresión logística.



Regresión Logística

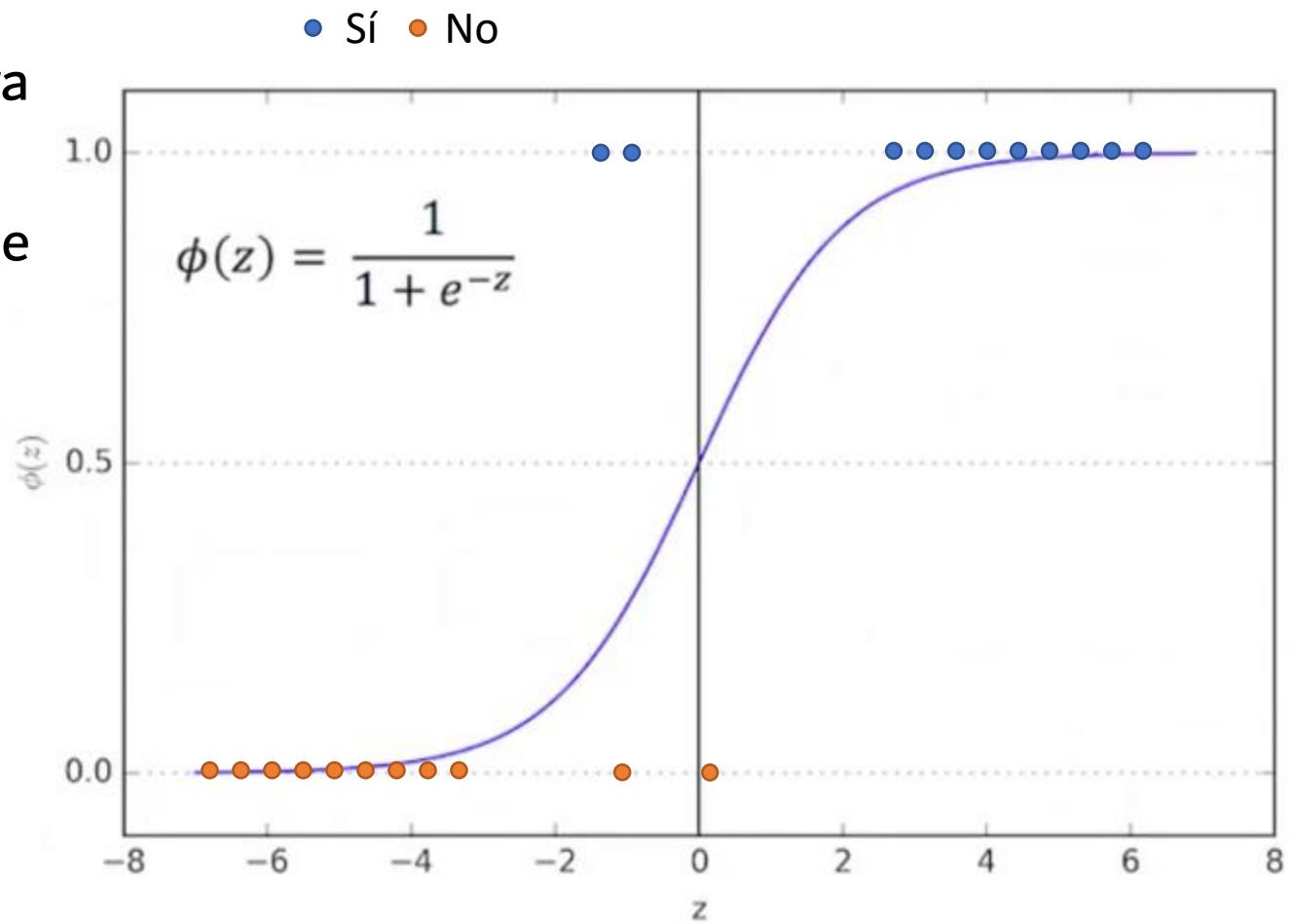
Modelo de Clasificación en el que nuestra variable a predecir es binaria.

La probabilidad de un estado así se calcula así:

$$P_i = 1 - \left(\frac{1}{1 + e^{z_i}} \right)$$

donde Z es:

$$Z_i = \ln\left(\frac{P_i}{1 - P_i}\right) = \beta_0 + \beta_1 x_1 + \dots + \beta_n x_n$$



Regresión Logística

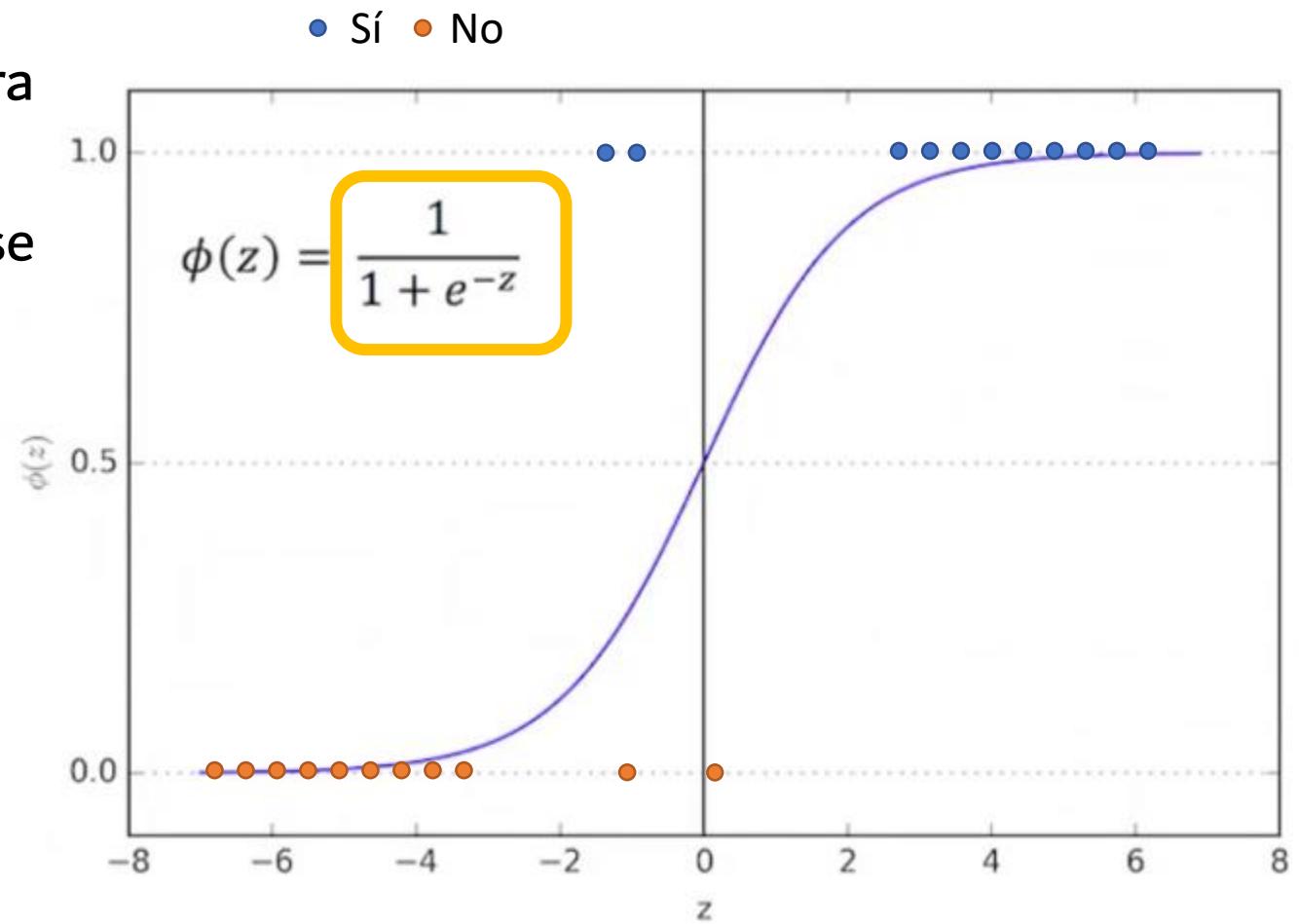
Modelo de Clasificación en el que nuestra variable a predecir es binaria.

La probabilidad de un estado así se calcula así:

$$P_i = 1 - \left(\frac{1}{1 + e^{z_i}} \right)$$

donde Z es:

$$Z_i = \ln\left(\frac{P_i}{1 - P_i}\right) = \beta_0 + \beta_1 x_1 + \dots + \beta_n x_n$$

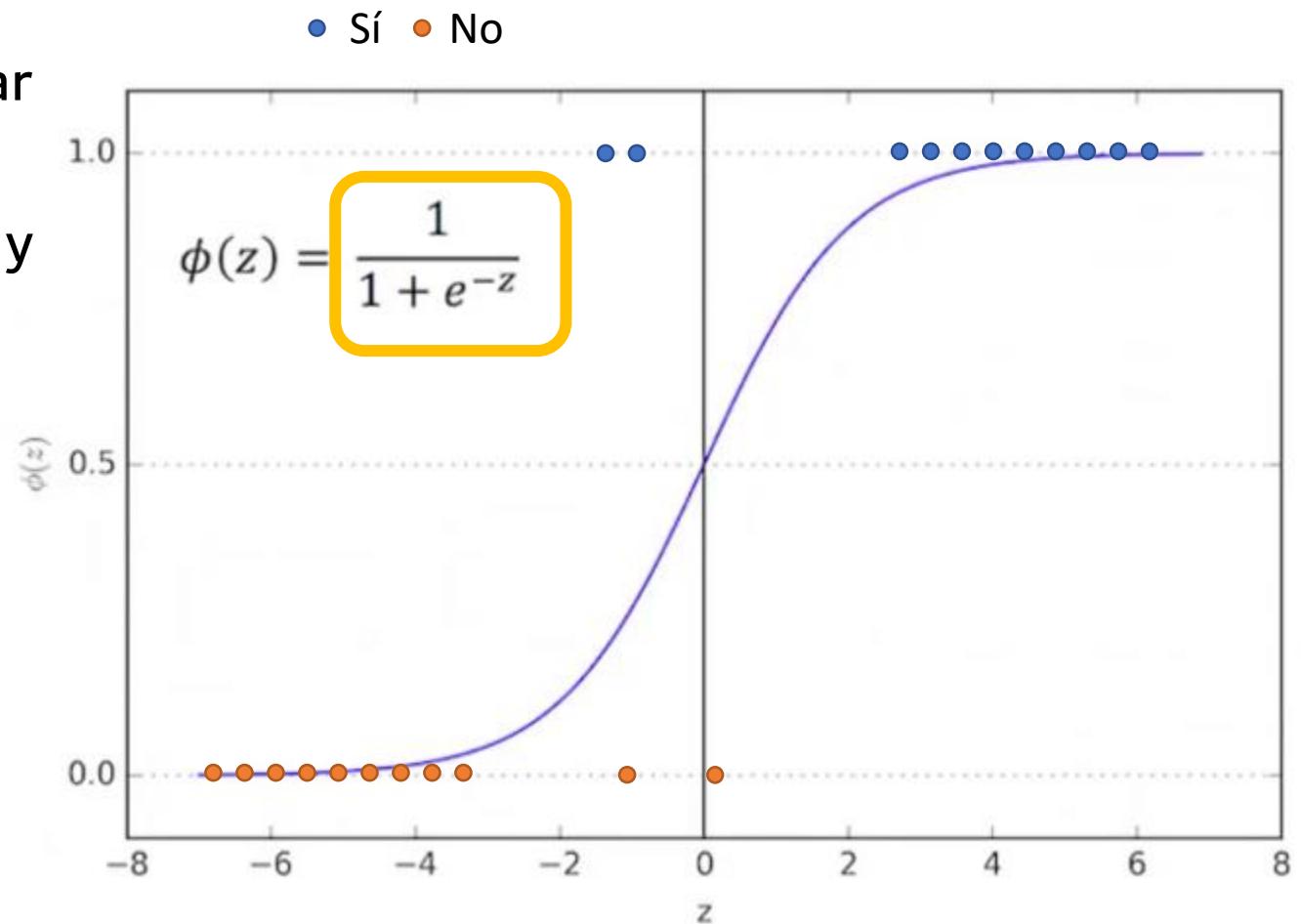


Regresión Logística

La función sigmoide para transformar nuestra data y adaptarla al problema.

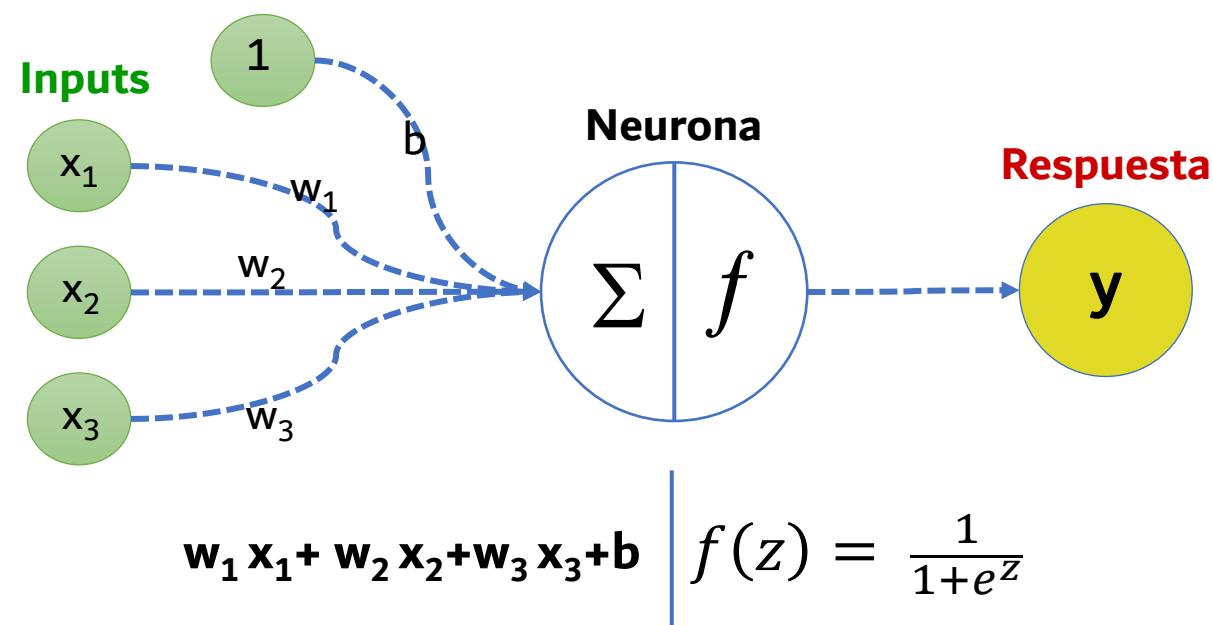
¿Qué sucede con valores de z grandes? y con valores muy negativos?

$$P_i = 1 - \left(\frac{1}{1 + e^{z_i}} \right)$$



Una neurona - función de activación

Esto nos abre paso a la siguiente parte de la neurona, la función de activación que si definimos como una sigmoide, resulta que una regresión logística puede ser vista como una red neuronal simple de una neurona.



Predecir si se va a tener una buena noche

Tener una buena noche = +

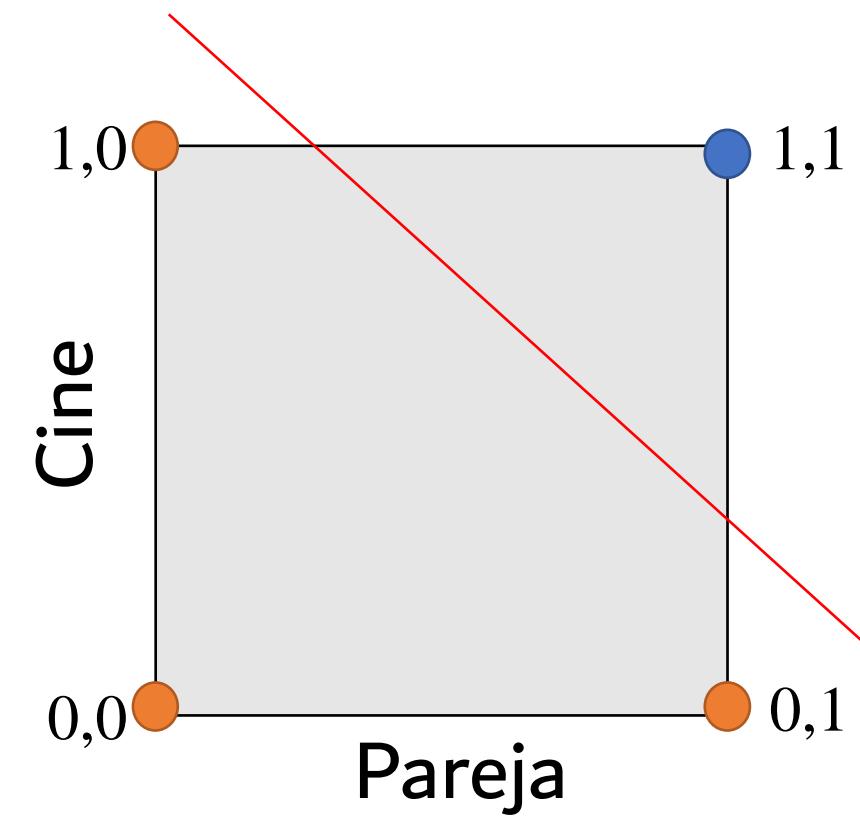
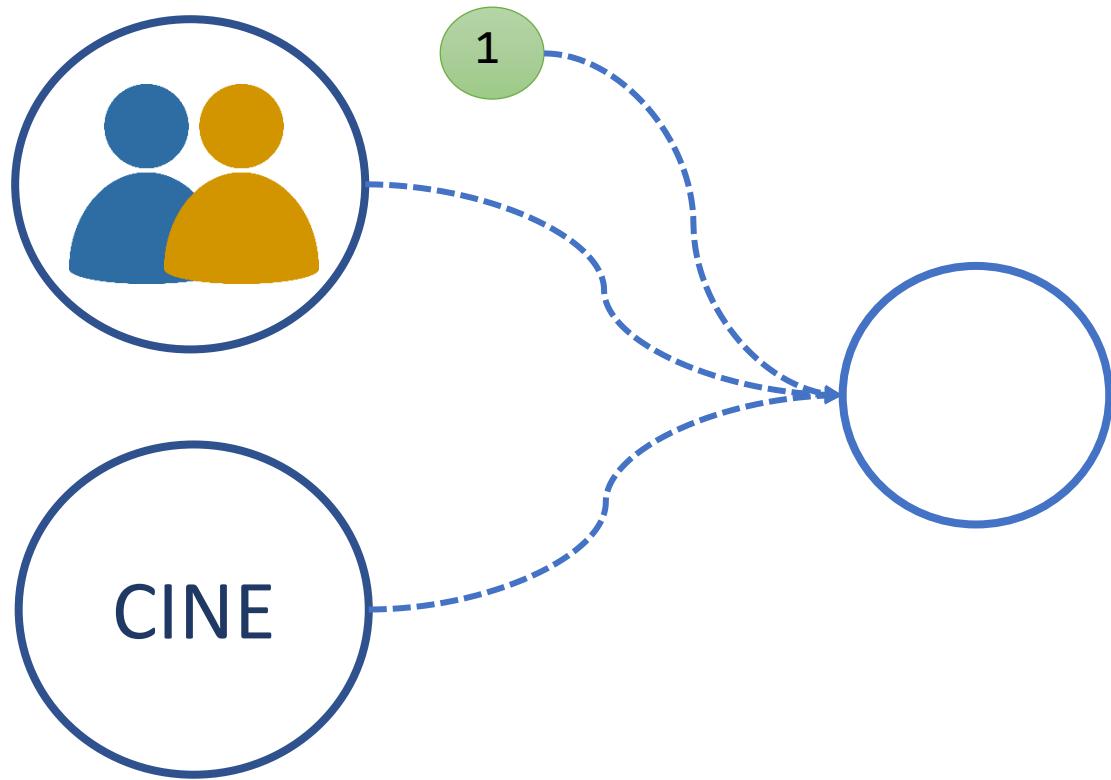
Predecir si se va a tener una buena noche

Tener una buena noche = Estar con tu pareja +



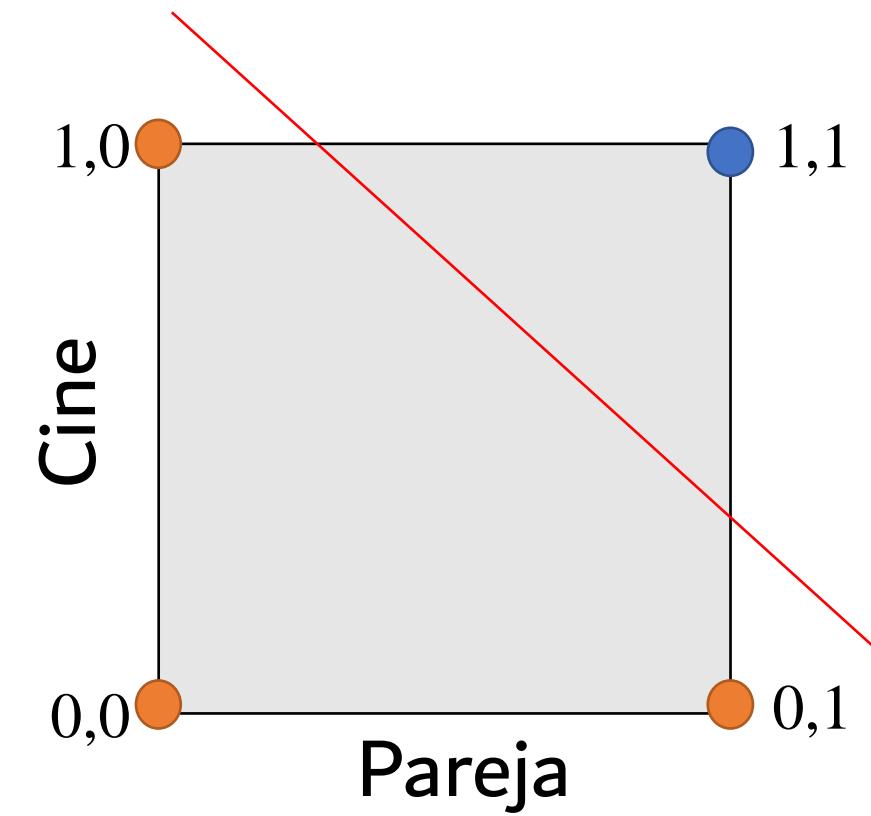
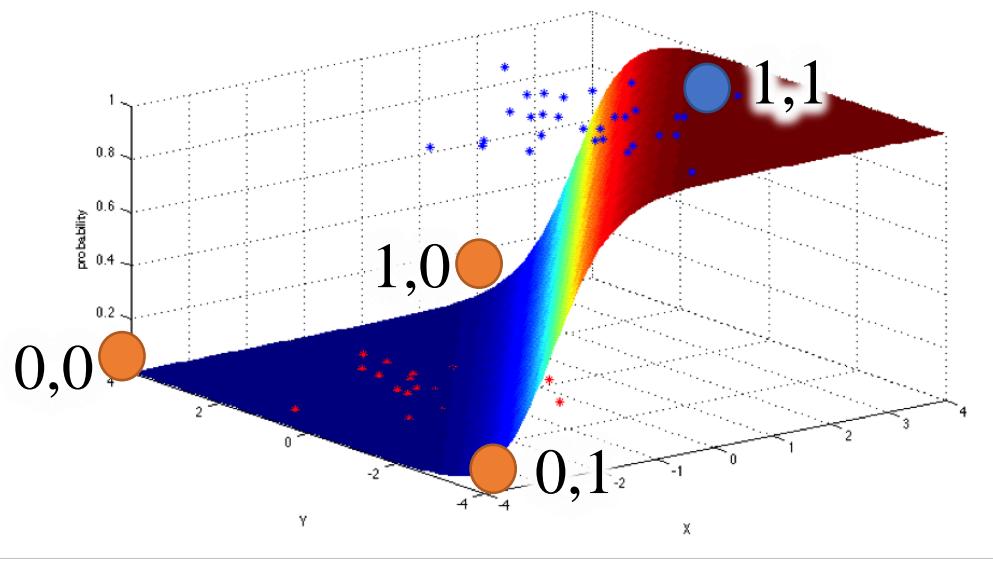
Predecir si se va a tener una buena noche

Tener una buena noche = Estar con tu pareja + Cine



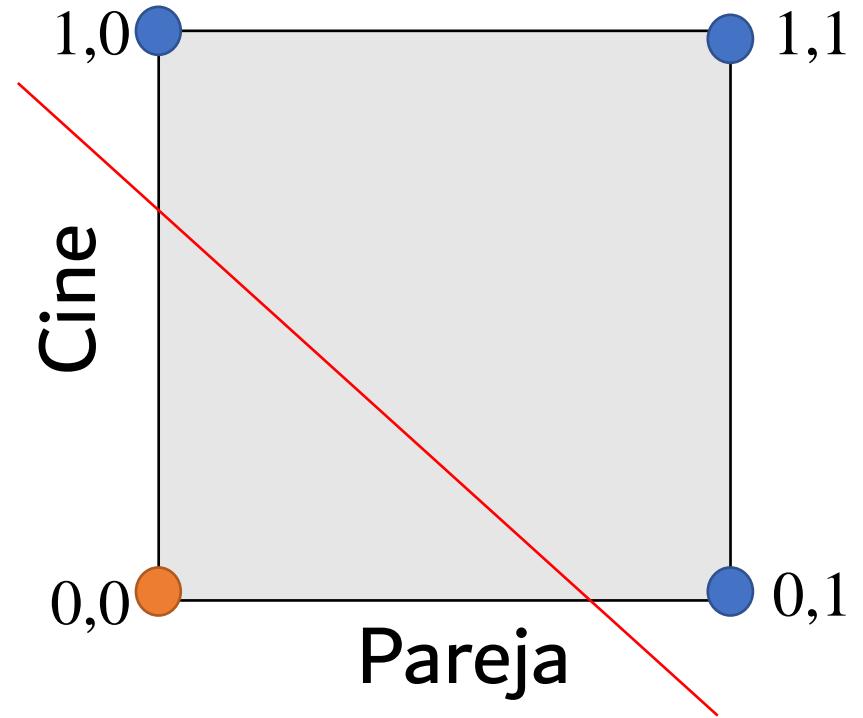
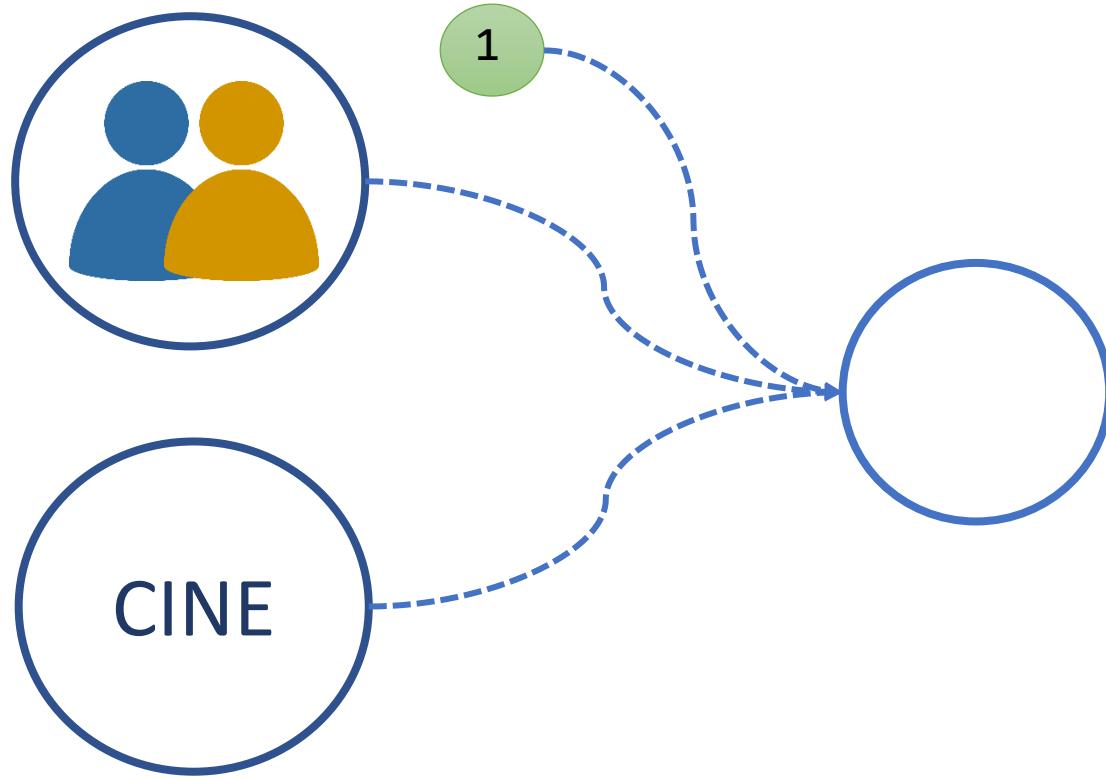
Predecir si se va a tener una buena noche

Notar que lo que estamos haciendo (viéndolo en 3D) es que con la función sigmoide encontramos un plano de esta forma:



Predecir si se va a tener una buena noche

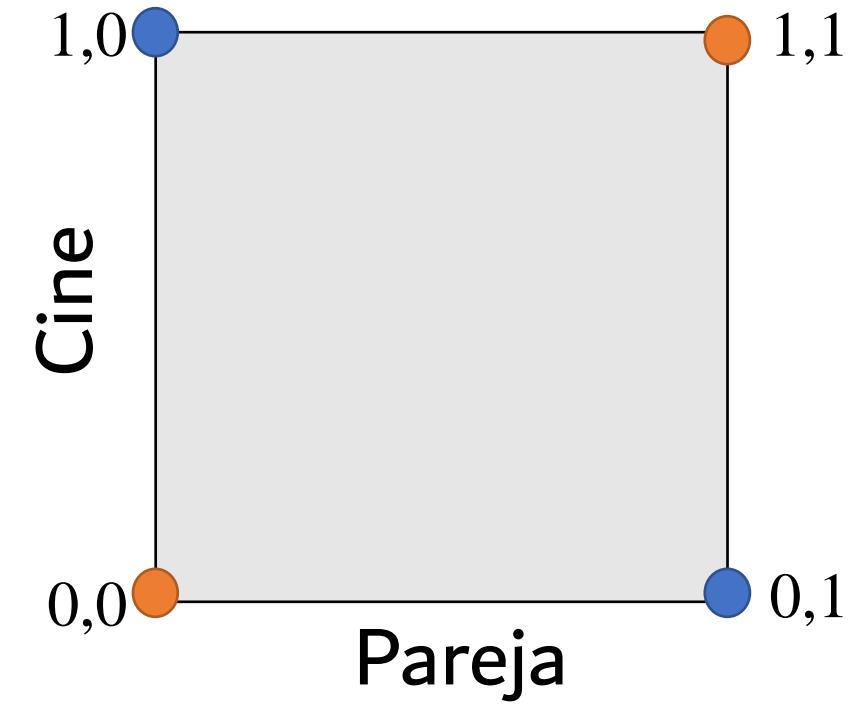
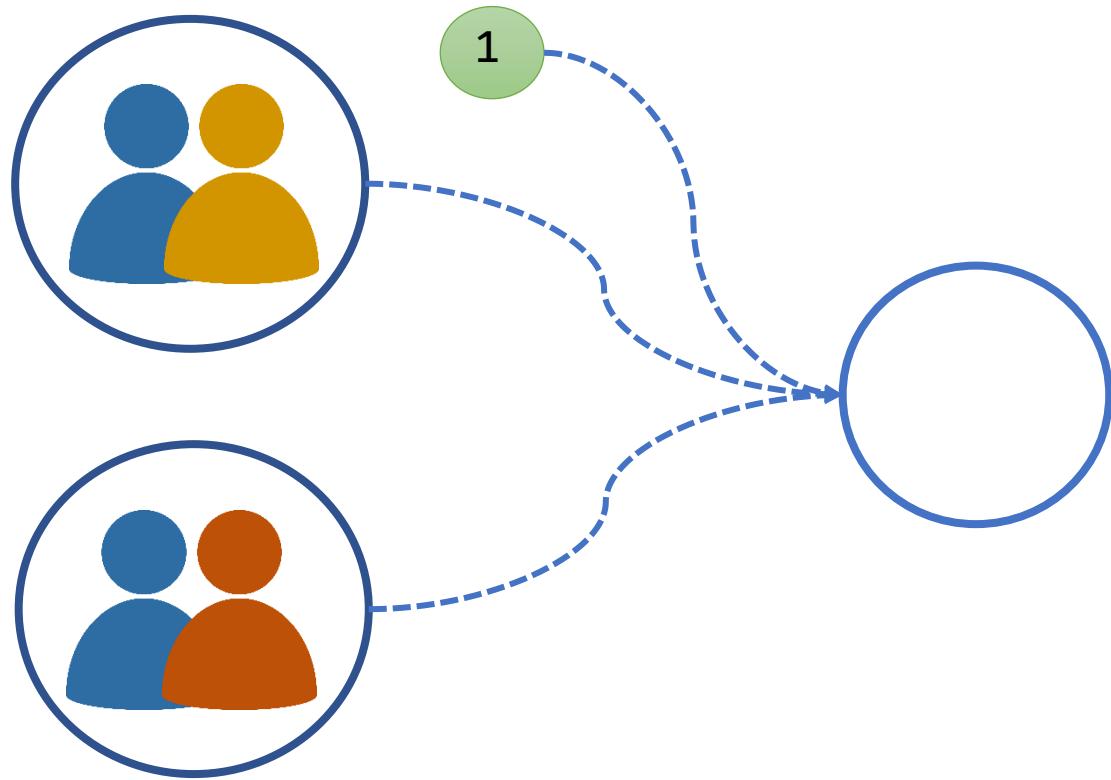
Tener una buena noche = Estar con tu pareja  Cine



Predecir si se va a tener una buena noche

Tener una buena noche = Estar con tu pareja 1 Ó pareja 2

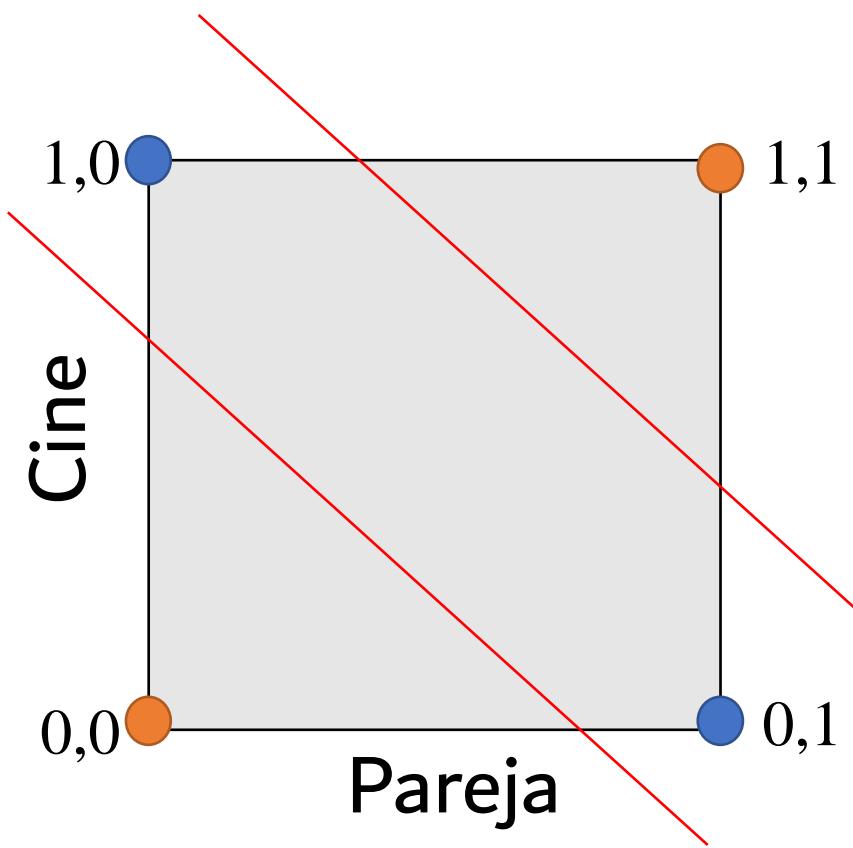
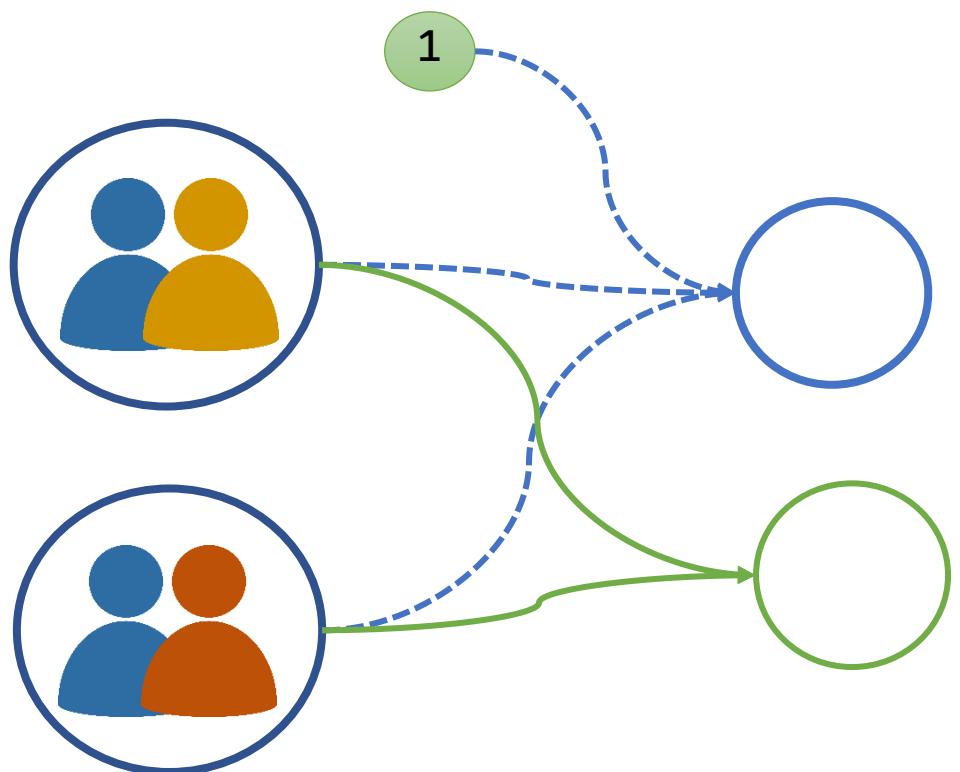
Ó excluyente (sin entrar a debates)



Predecir si se va a tener una buena noche

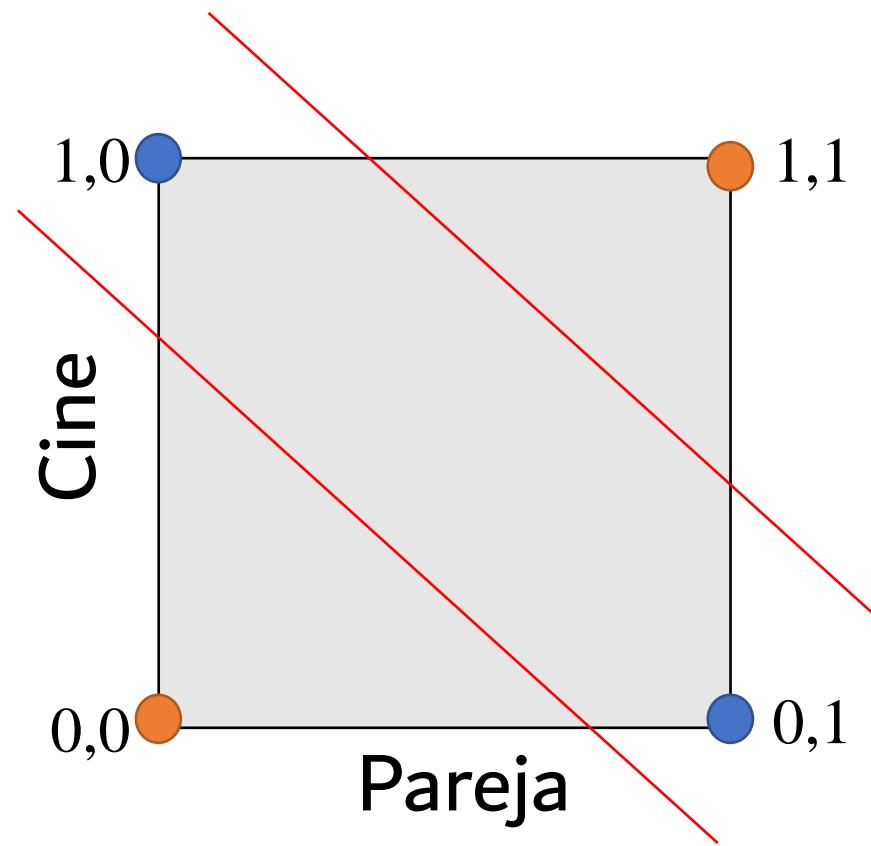
Tener una buena noche = Estar con tu pareja 1 Ó pareja 2

Ó excluyente (sin entrar a debates)



Predecir si se va a tener una buena noche

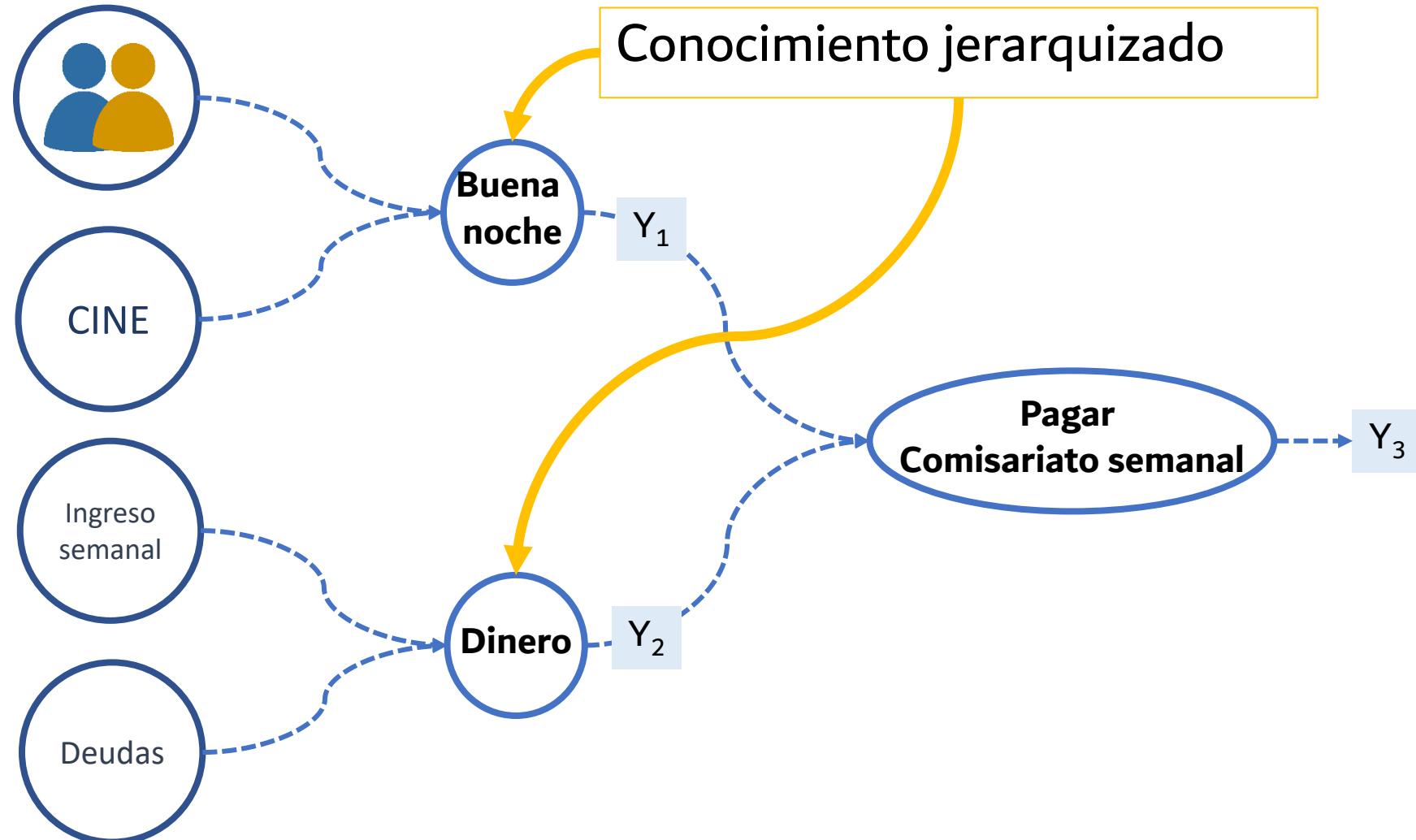
Con esto se ha ilustrado el beneficio de tener más de una neurona en paralelo, pero, ¿cuál es el beneficio de tenerlas en secuencia?, pues es lo que se conoce como conocimiento jerarquizado (piense en las interacciones de una regresión lineal).



Predecir si se va a tener una buena noche

Conocimiento jerarquizado

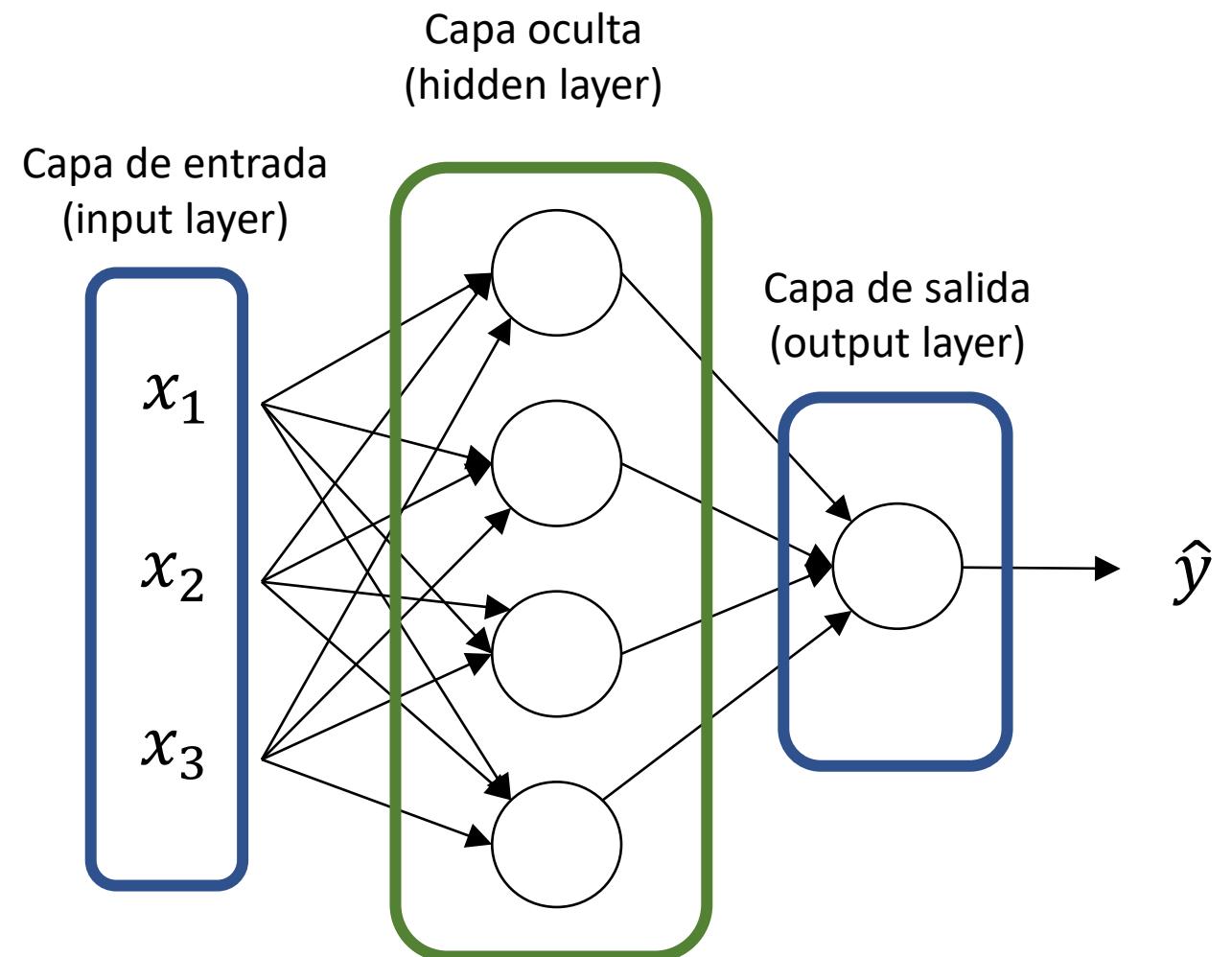
En las capas ocultas se van generando nuevas variables que no observamos en primera instancia.



Red neuronal Feedforward

Los nodos de las capas ocultas y de salida tienen asociado, input, pesos y función de activación y dan como resultado un output.

La manera en que se conectan los nodos, el tipo de nodo y tipo de layer es lo que compone una **arquitectura de red neuronal**



Red neuronal

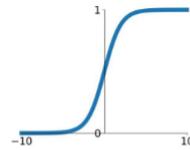
Recordando el esquema anterior, aplicar una red neuronal implica definir:

Arquitectura de red.- Tipos de nodo y layers, forma en que se conectan, funciones de activación usadas.

Activation Functions

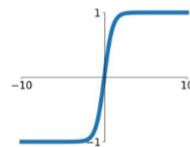
Sigmoid

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



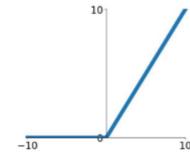
tanh

$$\tanh(x)$$



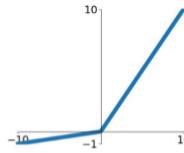
ReLU

$$\max(0, x)$$



Leaky ReLU

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

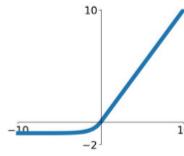


Maxout

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

ELU

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



Nótese que las funciones de activación no son lineales, puesto que se ha demostrado que hacer una red neuronal con funciones de activación lineales equivale a tener una red de una sola neurona.

Sin embargo, sí se puede tener una red con cuya última capa sea una neurona con función de activación lineal.

Red neuronal

Recordando el esquema anterior, aplicar una red neuronal implica definir:

Arquitectura de red.- Tipos de nodo y layers, forma en que se conectan, funciones de activación usadas.

Función de costo.- Que nos mide el error o diferencias entre las salidas de la red con el valor real o esperado, ejemplo: *pérdida logística o binary crossentropy o Log Loss*

$$\text{PérdidaLogística} = \sum_{(x,y) \in D} -y\log(y') - (1-y)\log(1-y')$$

Si $y=1$ la función de pérdida va a pedir un y' grande (para maximizar $\log(y')$)

Si $y==0$ la función de pérdida va a pedir un y' pequeño (para maximizar $\log(1-y')$)

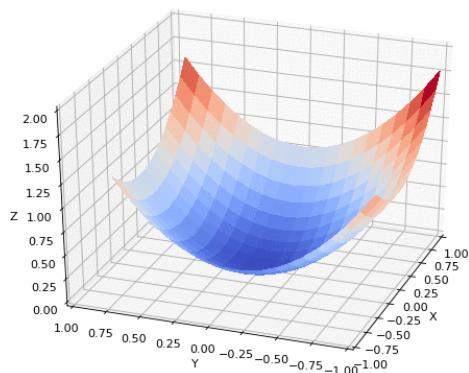
Red neuronal

Recordando el esquema anterior, aplicar una red neuronal implica definir:

Arquitectura de red.- Tipos de nodo y layers, forma en que se conectan, funciones de activación usadas.

Función de costo.- Que nos mide el error o diferencias entre las salidas de la red con el valor real o esperado, ejemplo: *pérdida logística*

Algoritmo de optimización.- Procedimiento que encuentra los parámetros de la red que reducen su error al mínimo, ejemplo: *Gradiente descendente*



Red neuronal

Recordando el esquema anterior, aplicar una red neuronal implica definir:

Arquitectura de red.- Tipos de nodo y layers, forma en que se conectan, funciones de activación usadas.

Función de costo.- Que nos mide el error o diferencias entre las salidas de la red con el valor real o esperado, ejemplo: *pérdida logística*

Algoritmo de optimización.- Procedimiento que encuentra los parámetros de la red que reducen su error al mínimo, ejemplo: *Gradiente descendente*

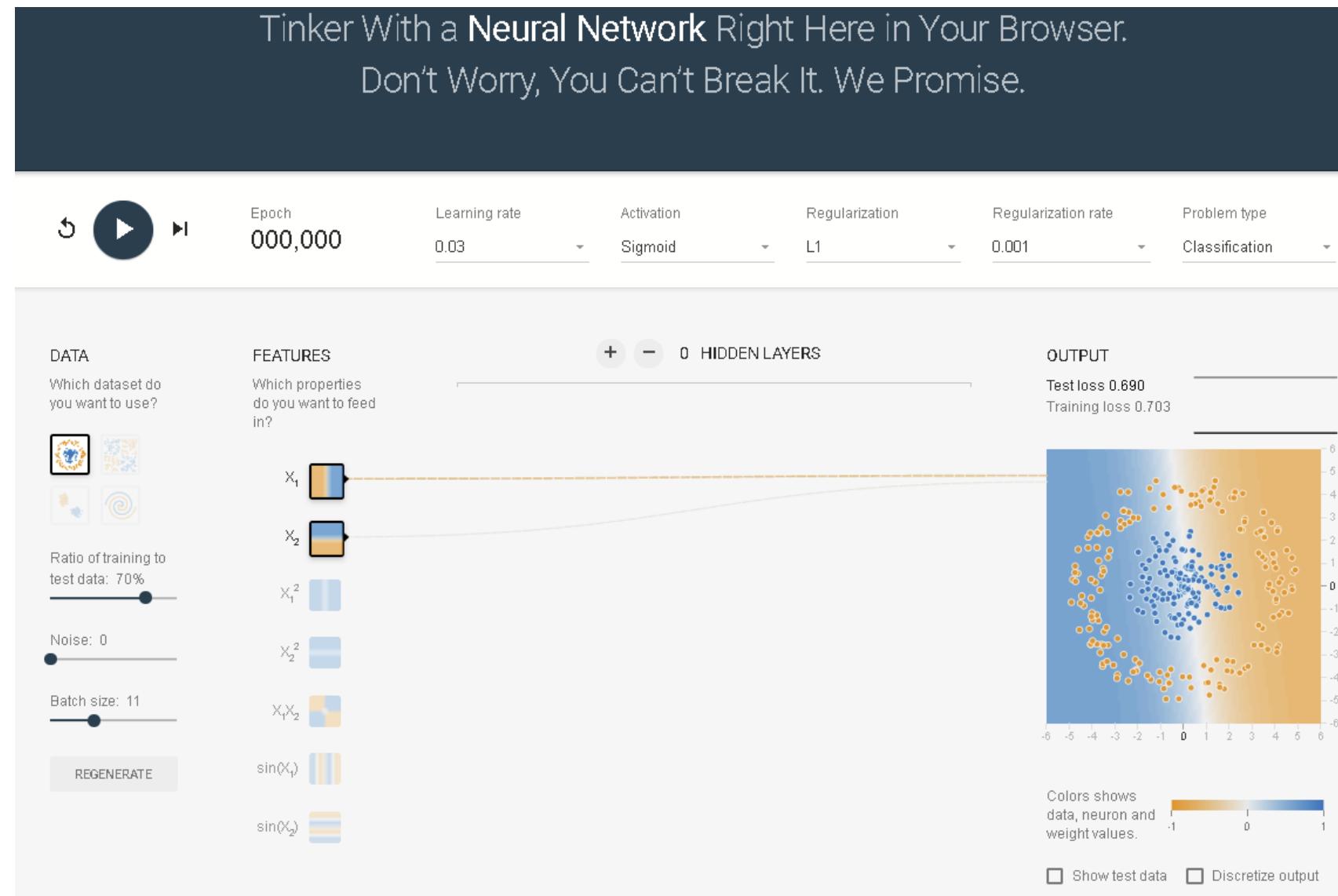
Red neuronal Feedforward

Ir a:

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

Setear la red tal como la imagen

Luego aumentar nodos y demás variables según crea necesario



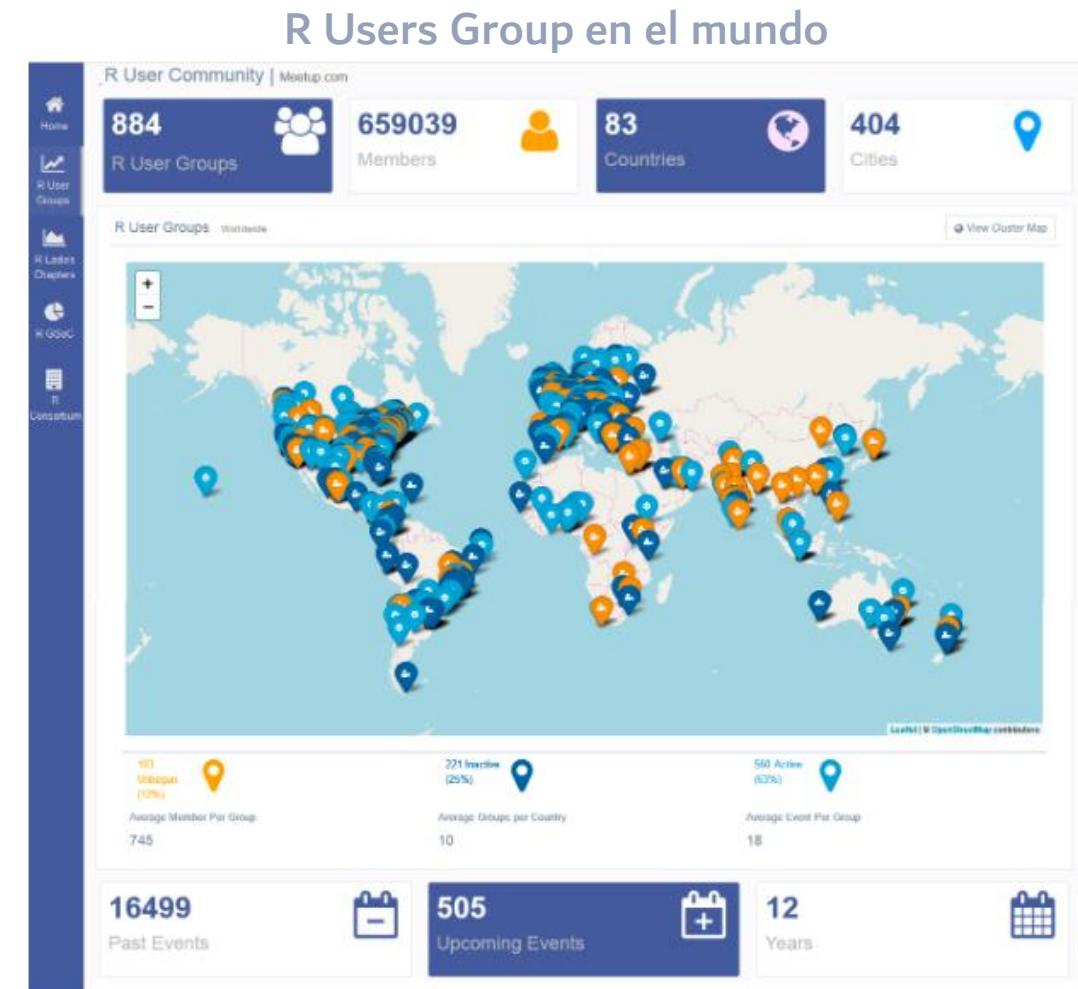


Por qué



¿Por qué R?

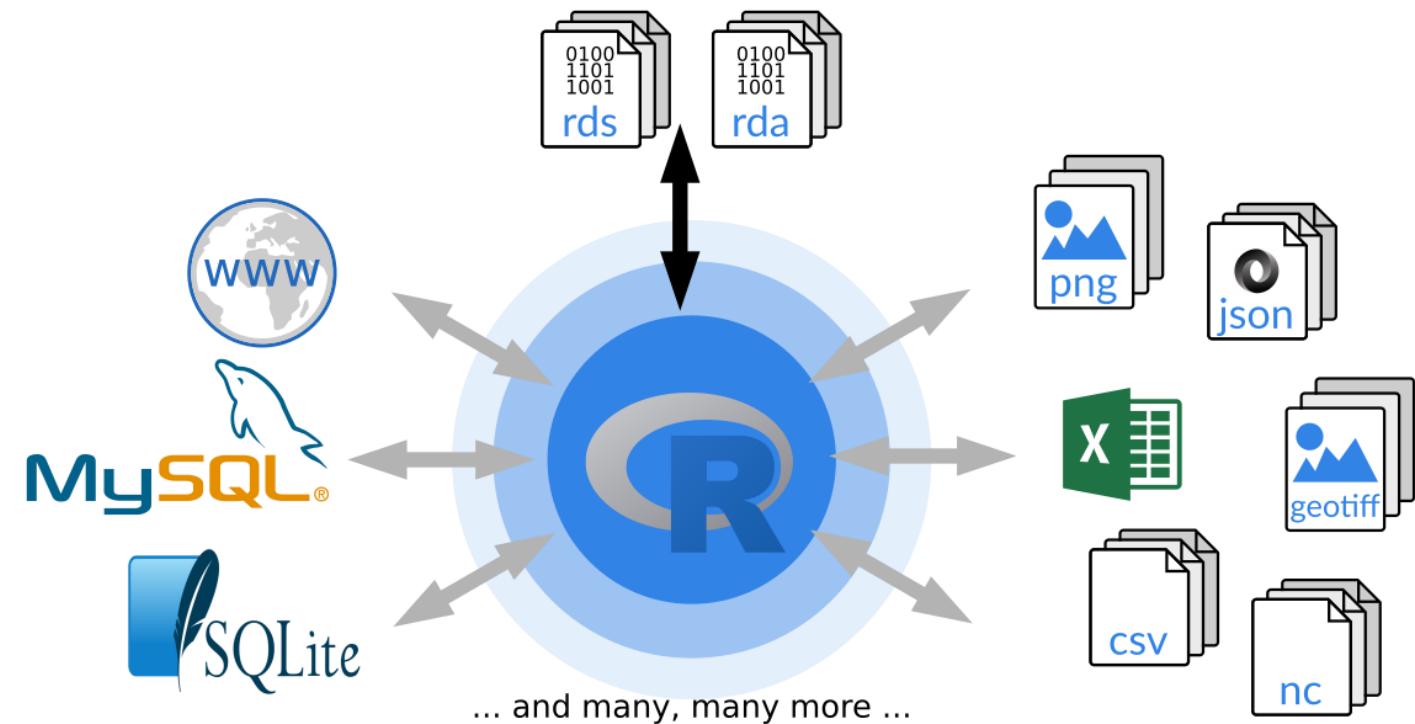
- Software libre, gratuito y de desarrollo independiente,
- Uno de los lenguajes más usados para Data Science,
- Gran y creciente comunidad de usuarios,
- Inmensa cantidad de material para aprender,
- Curva de aprendizaje simple (con el tidyverse),

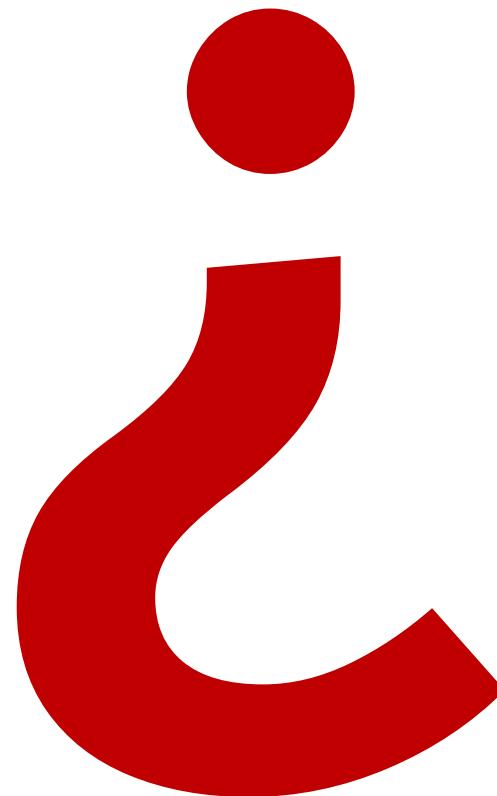


¿Por qué R?

- Rico ecosistema de paquetes, desarrollos derivados, integraciones y derivados,
- La mayoría de los nuevos métodos estadísticos se desarrollan primero en R,
- Varias IDEs y GUI disponibles, Rstudio es la IDE más usada,

Algunas de las integraciones de R





Por qué

K Keras

TensorFlow



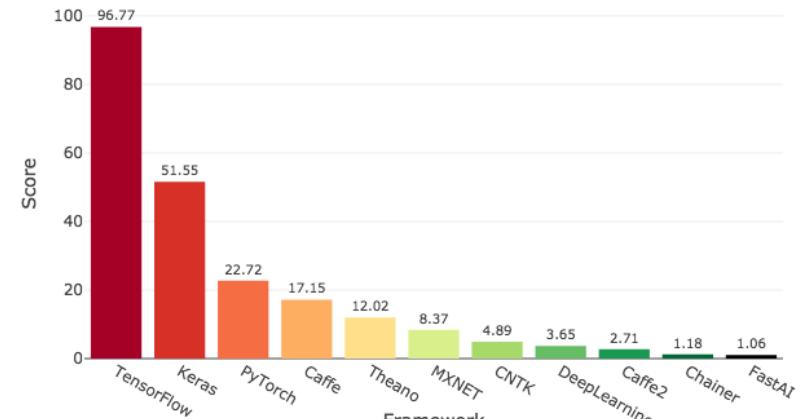
Frameworks para deep learning



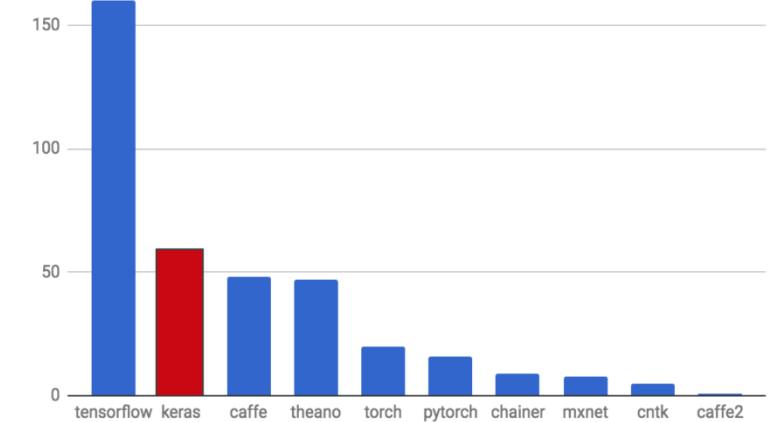
¿Por qué Tensorflow y Keras?

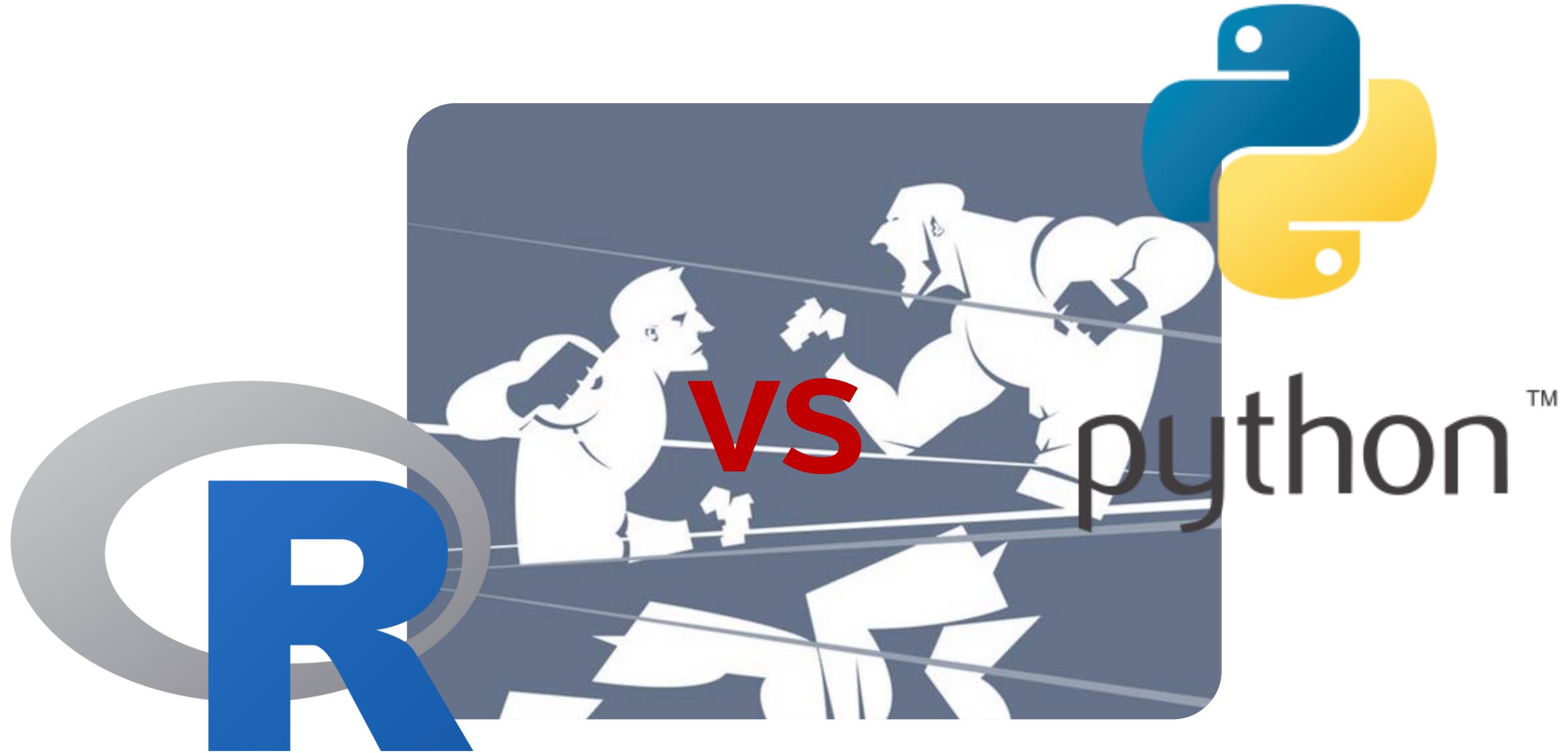
- Tensorflow creada y soportada por Google es el framework más usado, el que más actividad en github, búsquedas de Google, posts, etc.
- Keras es el segundo más usado y está construido bajo la idea de que sea fácil construir modelos de deep learning,
- Keras corre sobre tensorflow (y otros frameworks), ambos sobre **Python**.

Deep Learning Framework Power Scores 2018



arXiv mentions, October 2017











python™

- RStudio desarrolla reticulate para ejecutar Python desde R.
- RStudio connect soporta aplicaciones Flask y Dash
- Se puede ejecutar R desde Python con rpy2, pyRserve or pypeR

Deep Learning con keras y tensorflow desde R

- R se conecta con Python y ejecuta keras que a su vez ejecuta tensorflow.

Paquetes para que se...

- **reticulate**: ...conecte con Python
- **tensorflow**: ...conecte con Tensorflow
- **keras**: ...conecte con keras

Además

- tfruns, cloudml, y más.



Preeliminares

Python:

- Instalar Python (anaconda recomendado).
- Crear un environment en Python 3.6 y llamarlo r-reticulate
- Instalar tensorflow y keras
- Para GPU instalar drivers

R:

- Instalar reticulare, tensorflow y keras.

Si no quieres “topar” Python

R:

- Instalar reticulate, tensorflow y keras.
 - `install_tensorflow()`
 - `install_keras()`
- Para GPU (instalar requisitos)
- `install_tensorflow(version = "gpu")`
 - `install_keras(tensorflow = "gpu")`
- Más información en:
tensorflow.rstudio.com/installation/

Un Hello World: Titanic: Machine Learning from Disaster

Objetivo: Predecir cuáles pasajeros mueren

Link Kaggle: <https://www.kaggle.com/c/titanic/overview>

Kernel Kaggle: <https://www.kaggle.com/nestormontano/keras-r-titanic>

Google Colaboratory



Ir a <https://colab.research.google.com/drive/1WiKBVuGMKs9oe7ox1sxOa0gRg0UptjG3>
Copiar el colab en su Google Drive
Ojo: También pueden usar un kernel de kaggle

Red Neuronal en Keras

Se requieren 4 pasos:

- Especificar la arquitectura (cuántos layers, nodos, función de activacion en cada layer),
- Compilar el modelo, aquí se especifica la función de costo y cómo se la va a optimizar,
- Ajustar (fit) el modelo, que es el paso donde se optimiza haciendo backpropagation,
- Predecir.

Definir arquitectura en Keras

- `keras_model_sequential()` especifica que la red será de tipo secuencial, esto significa que cada layer va a tener pesos o conexiones sólo con el layer inmediato siguiente
- `layer_dense()` es un layer cuyos nodos están conectados con todos los nodos del layer inmediato anterior,
- En el layer se especifica la función de activación y el número de nodos ('relu', 'sigmoid', 'softmax'),
- El primer `layer_dense()` debe indicar que recibe los nodos correspondientes a las variables de la data,
- El último `layer_dense()` es el output, en regresión y clasificación binaria sería una neurona, para clasificación de varias clases se usan tantas neuronas como label posibles

Definir arquitectura en Keras

- Existen otros tipos de layers (además del dense), así como de modelos (además del secuencial).
- `layer_batch_normalization()` normaliza las activaciones del layer previo en cada batch,
- `layer_dropout()` indica al modelo que un porcentaje del input se definan en cero (aleatorio), ayuda a prevenir sobreajuste,
- `layer_gaussian_noise()` Agrega ruido gausiano para mitigar el sobreajuste,
- `layer_conv_2d()` permite realizar un layer convolucional (util para procesamiento de imágenes)

Configurar la compilación en Keras

- `compile()` define la forma cómo se va a compilar la red creada,
- Se debe definir el optimizador, el cual controla la taza de aprendizaje “learning rate”,
- El algoritmo más usado es `optimizer_adam()`, pero hay otras opciones que incluso podrían ser mejores para un problema en particular,
- Adam tiene la particularidad de ajustar la taza de aprendizaje a medida que ejecuta el gradiente descendente,
- Se debe establecer la función de costo, '`binary_crossentropy`', '`categorical_crossentropy`' y '`mean_squared_error`' son los más usados,
- Adicionalmente se puede definir unas métrica a ser evaluadas durante el train-test.

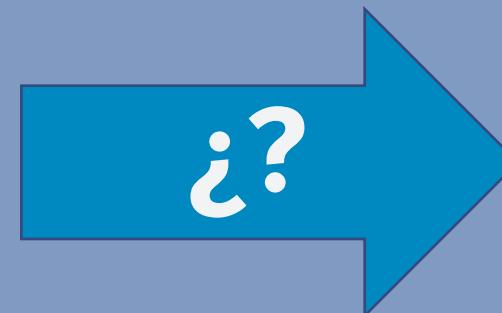
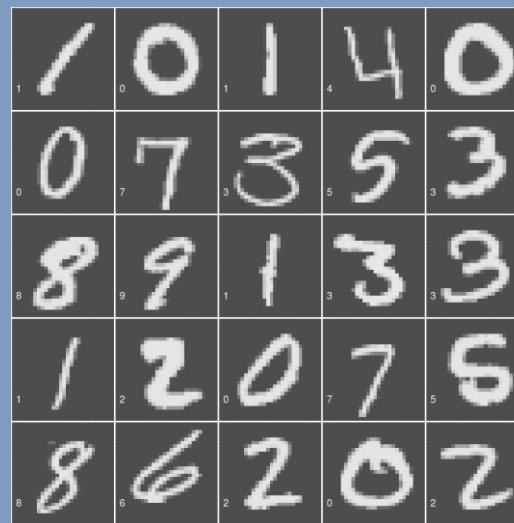
Ajustar el modelo en Keras

- `fit()` ajusta el modelo, es decir aplica backpropagation a la data para obtener los pesos óptimos,
- Se le debe definir input X y el output Y
- Además se le puede especificar: número de epochs, tamaño del batch, porcentaje para data de validación, entre otras opciones.
- Use `callback_early_stopping(patience = 100)` para indicar que pare si luego de 100 epoch no hay mejora,
- Para evaluar el fit, se usa `evaluate()`,
- `predict_classes()` y `predict_proba()` permiten hacer las predicciones.

Un Hello World: MNIST con Red Neuronal Convolucional

Objetivo: Detectar el número que está escrito a mano en la imagen.

Problema y Data: <https://www.kaggle.com/c/digit-recognizer>



1, 0, 1, 4, , 2, 0, 2

Redes Neuronales Convolucionales CNN

Se utilizan principalmente para el procesamiento de imágenes, pero también se pueden usar para otros tipos de entrada, como audio.

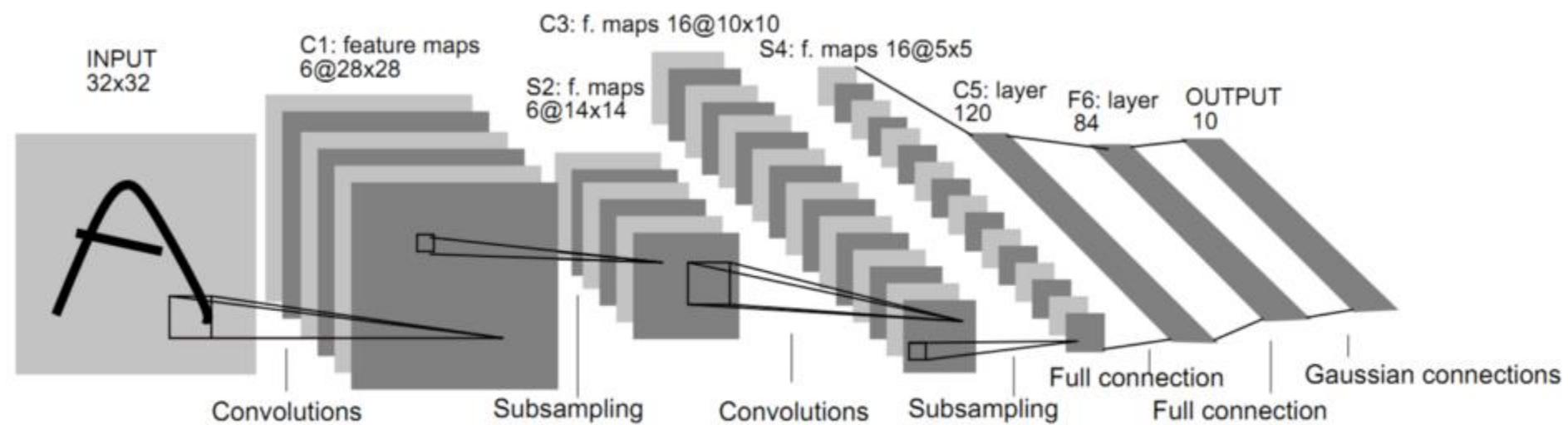
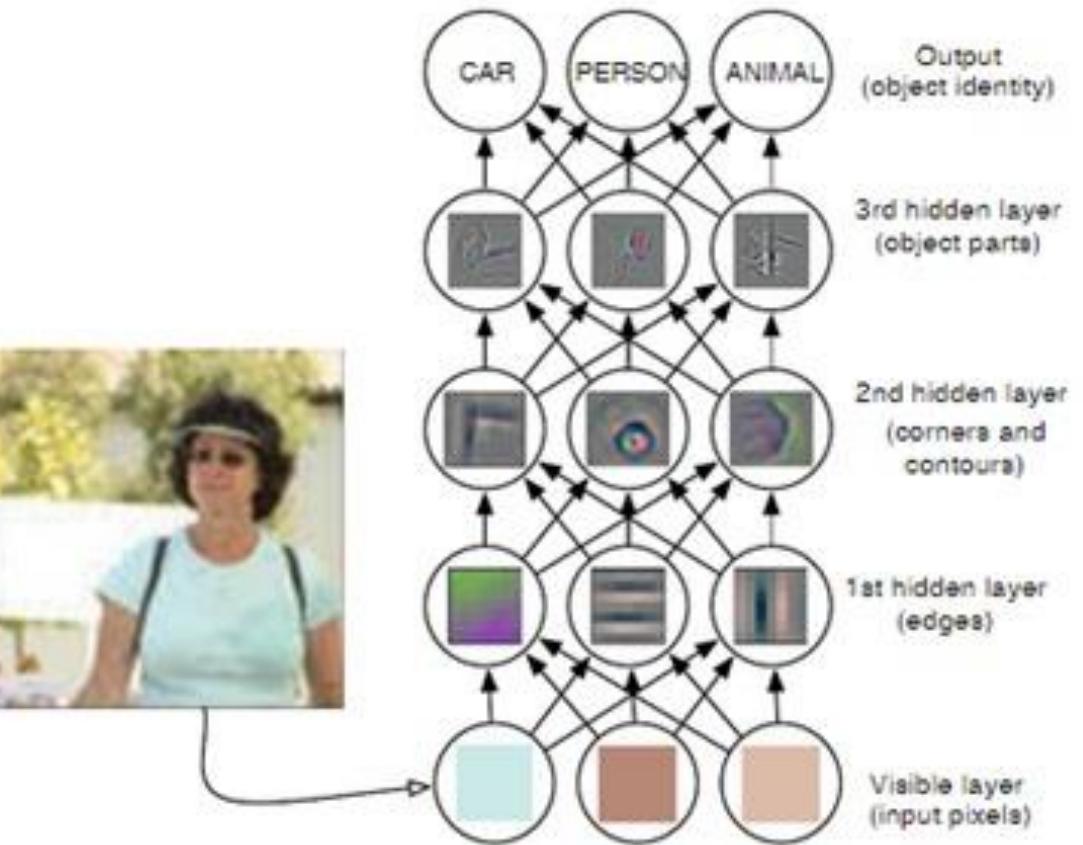


Fig. 2. Architecture of LeNet-5, a Convolutional Neural Network, here for digits recognition. Each plane is a feature map, i.e. a set of units whose weights are constrained to be identical.

Redes Neuronales Convolucionales CNN

Los layer convolucionales en el contexto de visión por computadora trabajarían así:



Redes Neuronales Convolucionales CNN

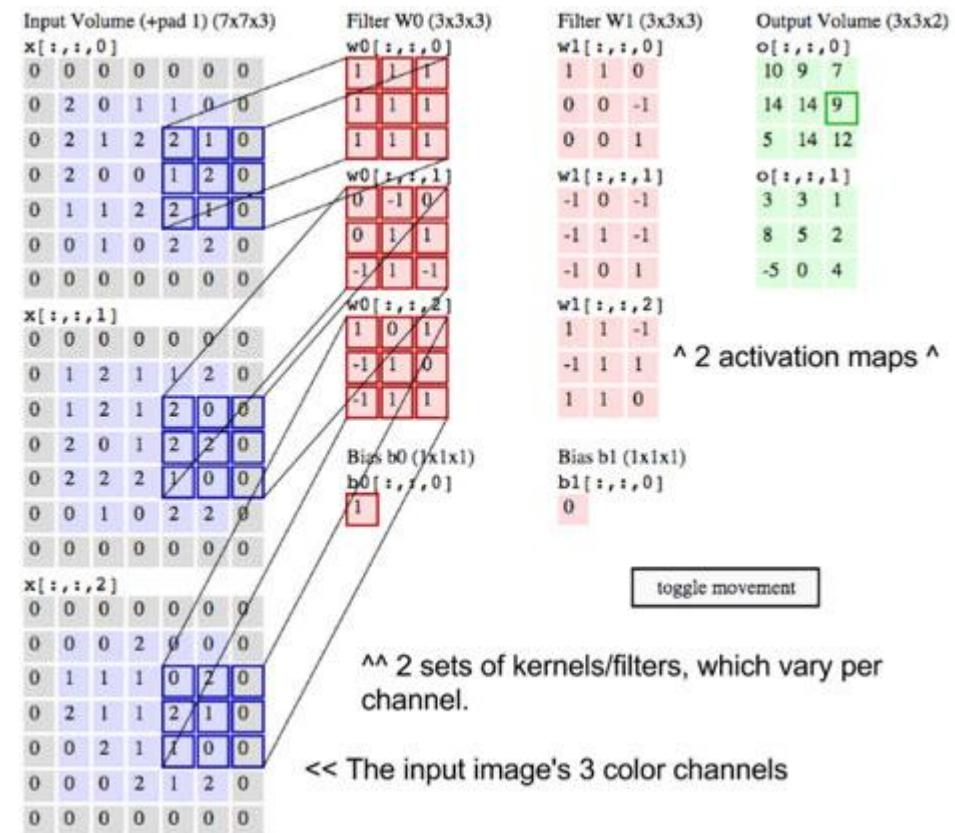
Convolución

1 x1	1 x0	1 x1	0	0
0 x0	1 x1	1 x0	1	0
0 x1	0 x0	1 x1	1	1
0	0	1	1	0
0	1	1	0	0

Image

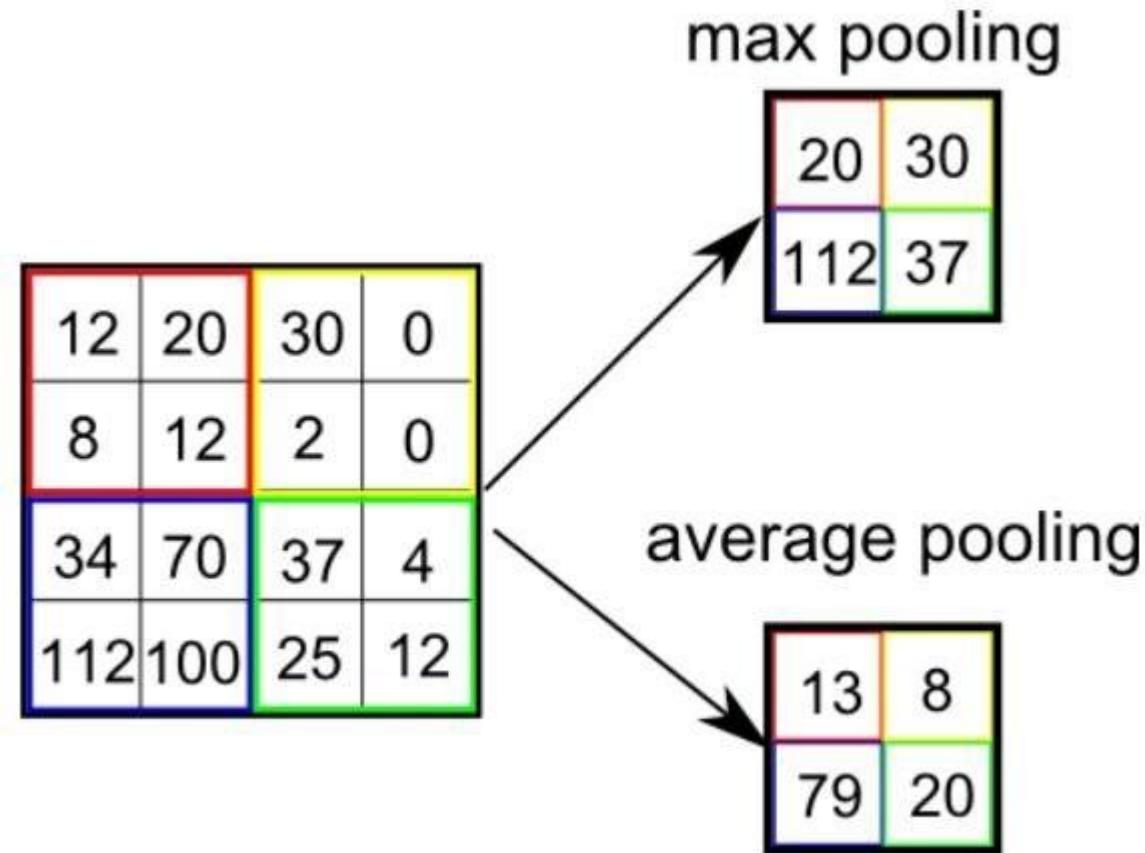
4		

Convolved Feature



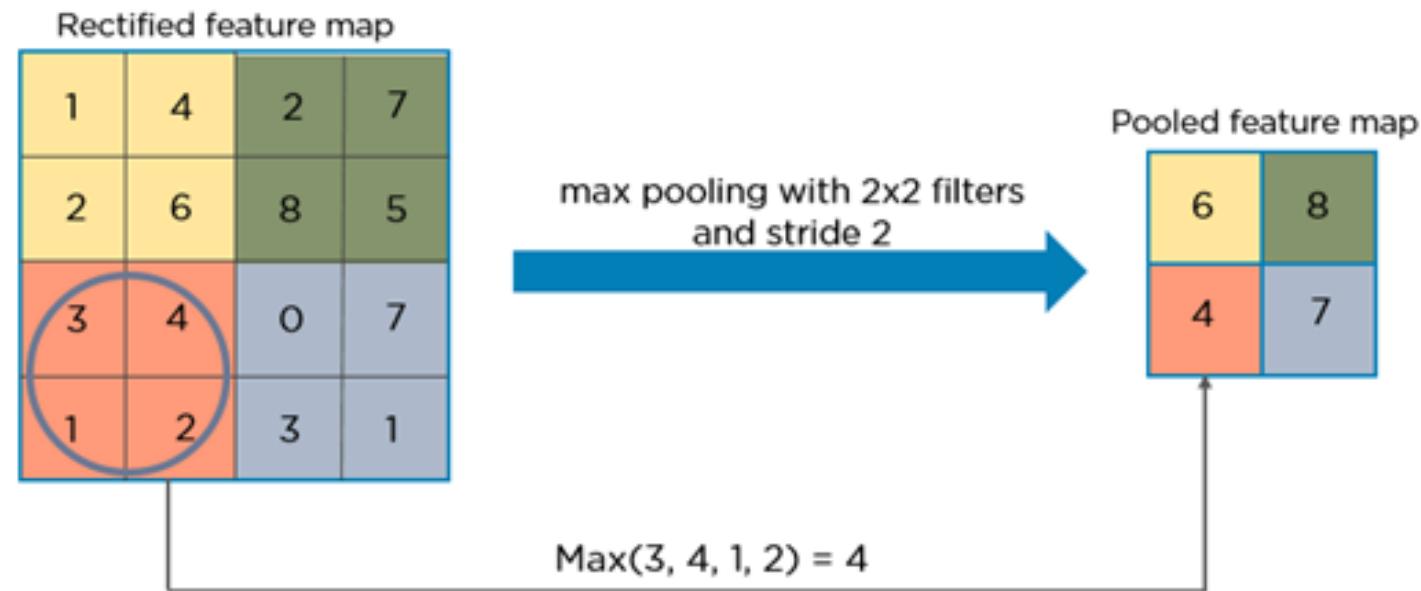
Redes Neuronales Convolucionales CNN

Pooling



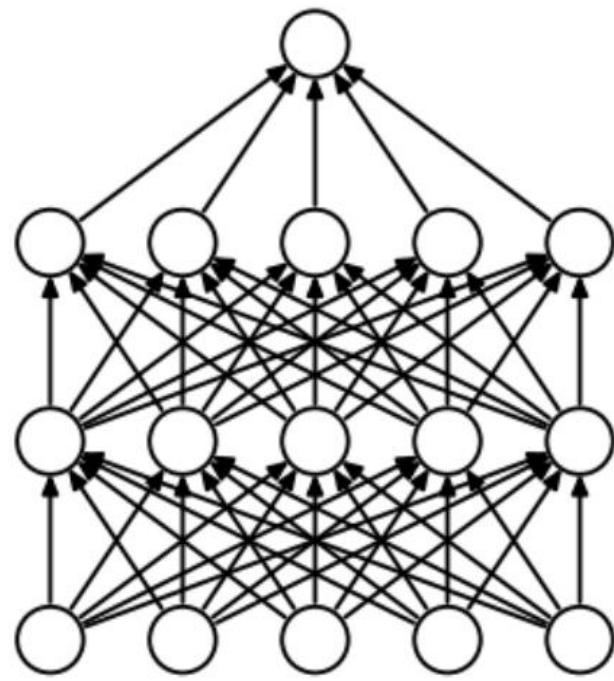
Redes Neuronales Convolucionales CNN

Stride es un salto que se da en la matriz para aplicar el filtro.

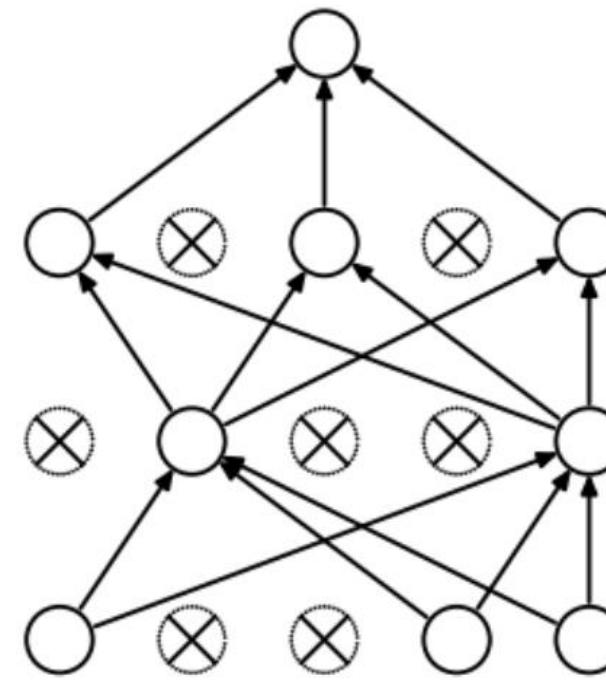


Redes Neuronales Convolucionales CNN

Drop out permite “apagar” algunas neuronas, ayuda a evitar overfitting



(a) Standard Neural Net



(b) After applying dropout.

Redes Neuronales

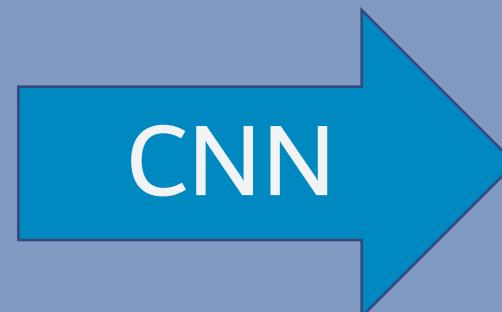
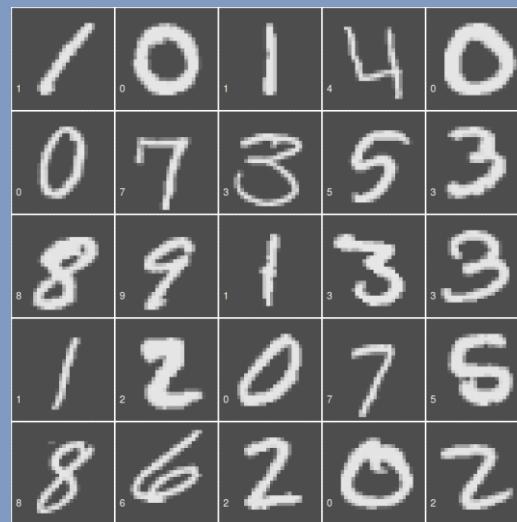
Esto es aún parte del arte de modelar, pero se puede tener como algoritmo básico lo siguiente:

- Empezar con una red simple,
- Obtener las métricas en la validacion,
- Aumentar los nodos en esta red hasta que se empiece a empeorar métricas,
- Aumentar una capa oculta con la misma cantidad de nodos,
- Verificar si aumenta la capacidad, si lo hace dejar el layer y probar quitando nodos.
- Probar transfer learning (investigar)

Un Hello World: MNIST con Red Neuronal Convolucional

Objetivo: Detectar el número que está escrito a mano en la imagen.

Idea: Construir un modelo de CNN (convolutional neural network) al que se le ingrese una imagen y este nos diga qué número está escrito en dicha imagen.



1, 0, 1, 4, , 2, 0, 2

Un Hello World: MNIST con Red Neuronal Convolucional

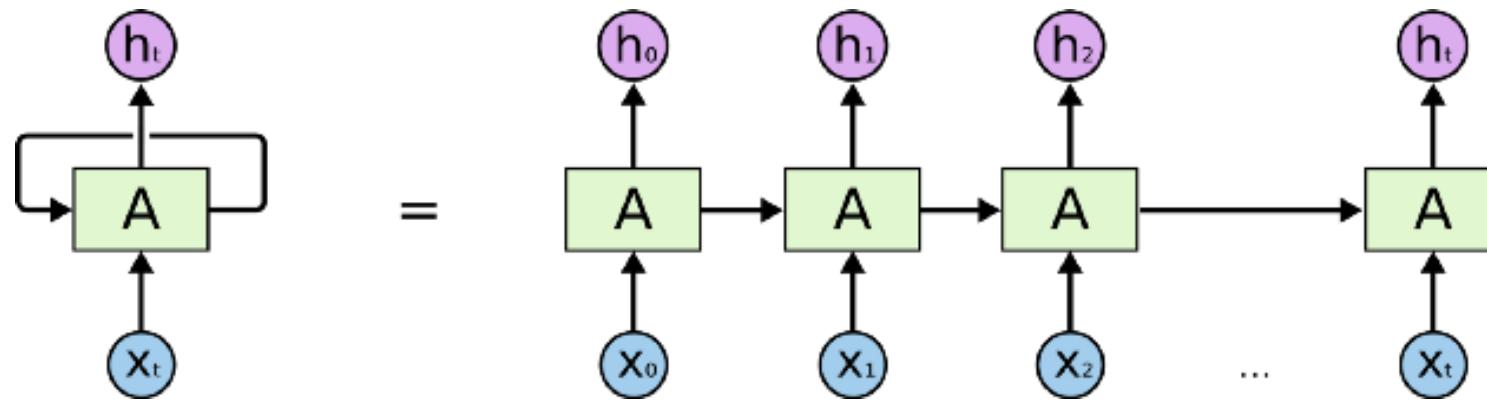
Objetivo: Detectar el número que está escrito a mano en la imagen.

Idea: Construir un modelo de CNN (convolutional neural network) al que se le ingrese una imagen y este nos diga qué número está escrito en dicha imagen.

Kernel Kaggle: <https://www.kaggle.com/nestormontano/flisol-2020-rkeras/>

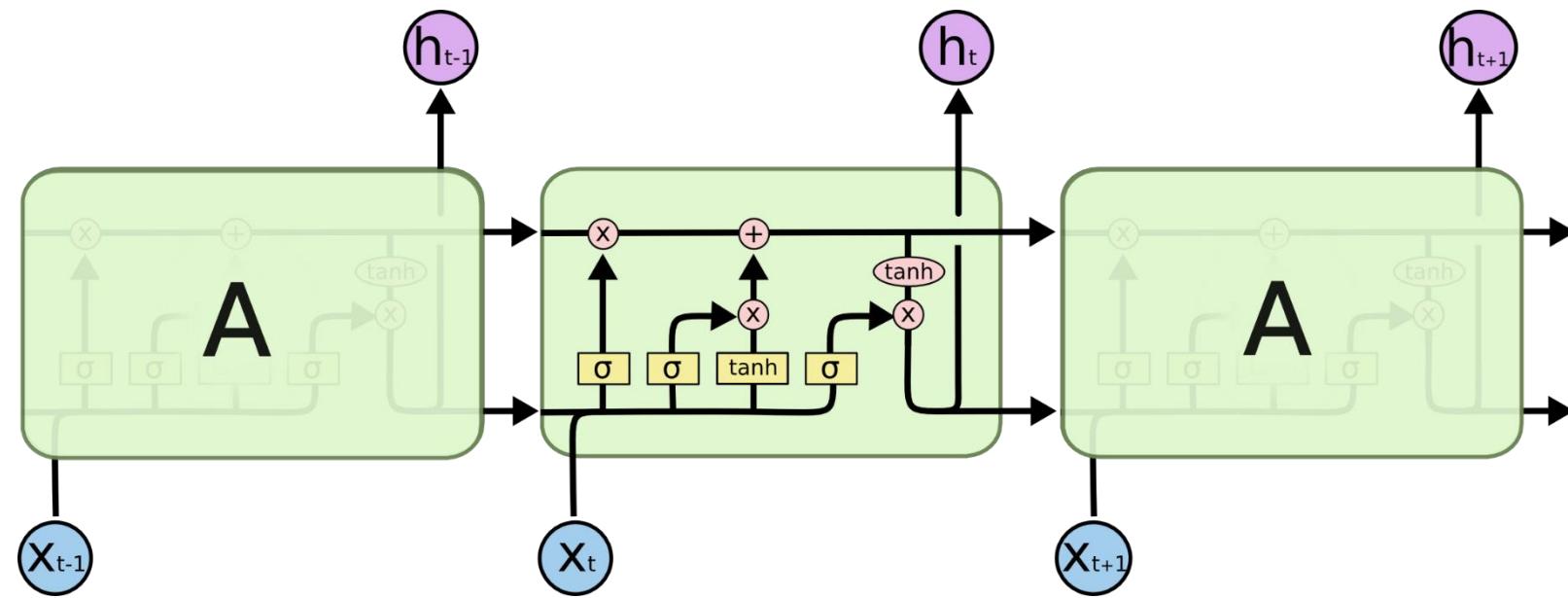
Redes Neuronales Recurrentes RNN

Especializados en recordar secuencias, la idea es que la RNN prediga la siguiente escena, posición, palabra, frase (chatbots), etc



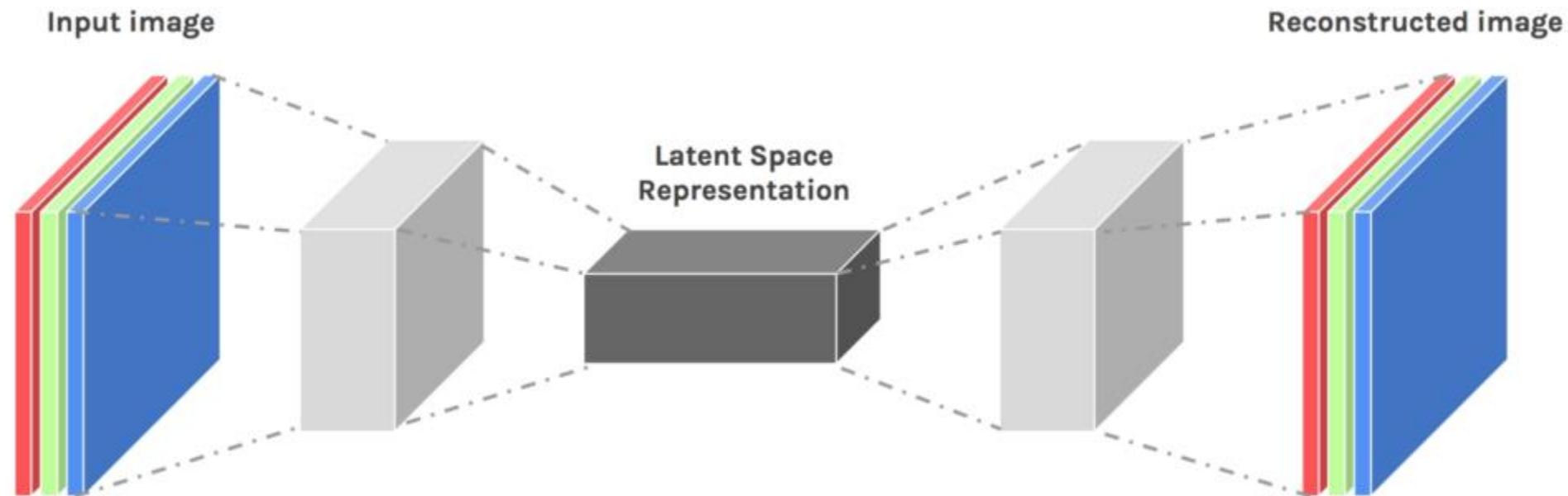
Long Short Term Memory Network LSTM

Es un tipo especial de RNN, capaz de aprender dependencias a largo plazo, fueron diseñada para guardar información durante largos períodos de tiempo



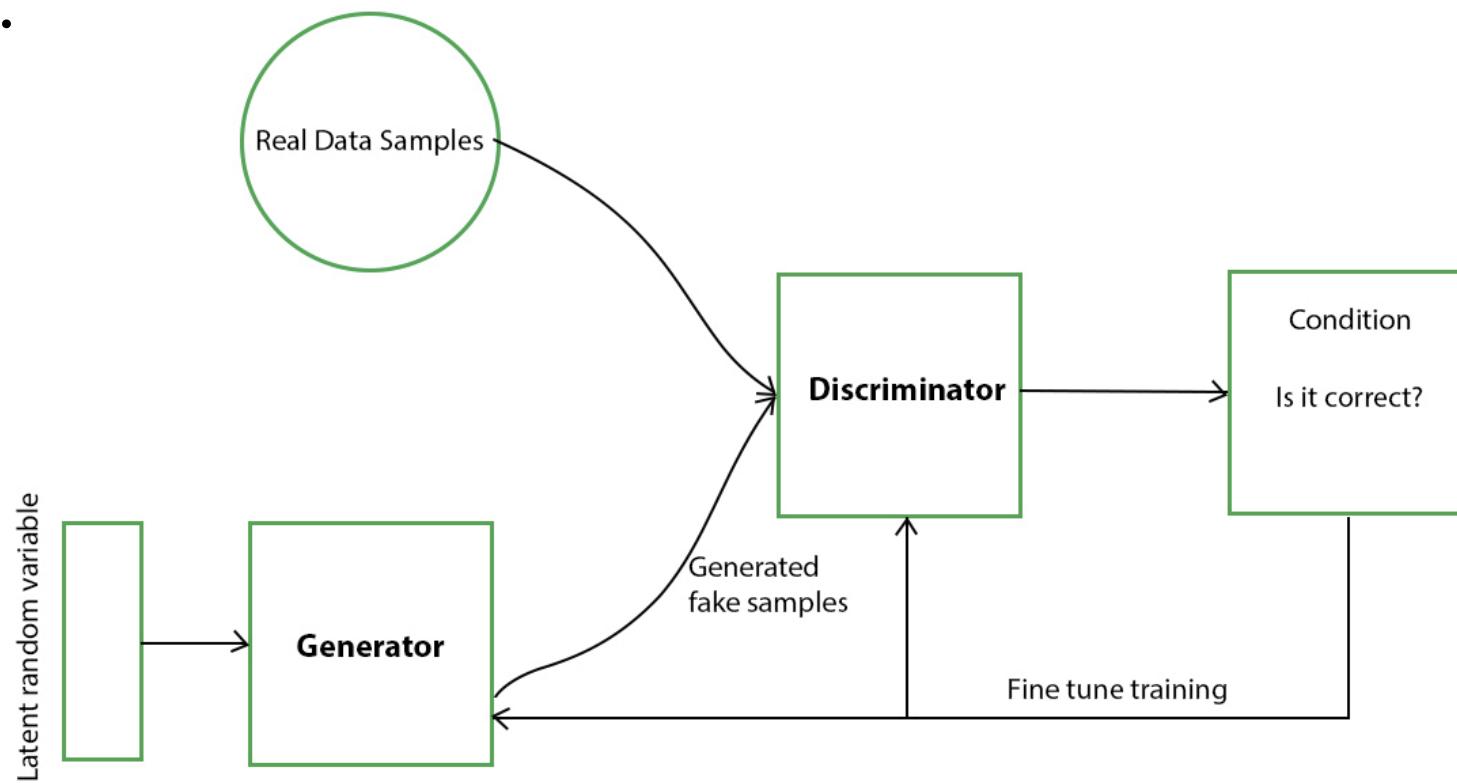
Autoencoders

Redes neuronales diseñadas para aprendizaje no supervisado, es un modelo de compresión de datos, se pueden usar para codificar una entrada determinada en una representación de dimensión más pequeña. Luego se puede usar un decodificador para reconstruir la entrada desde la versión codificada.



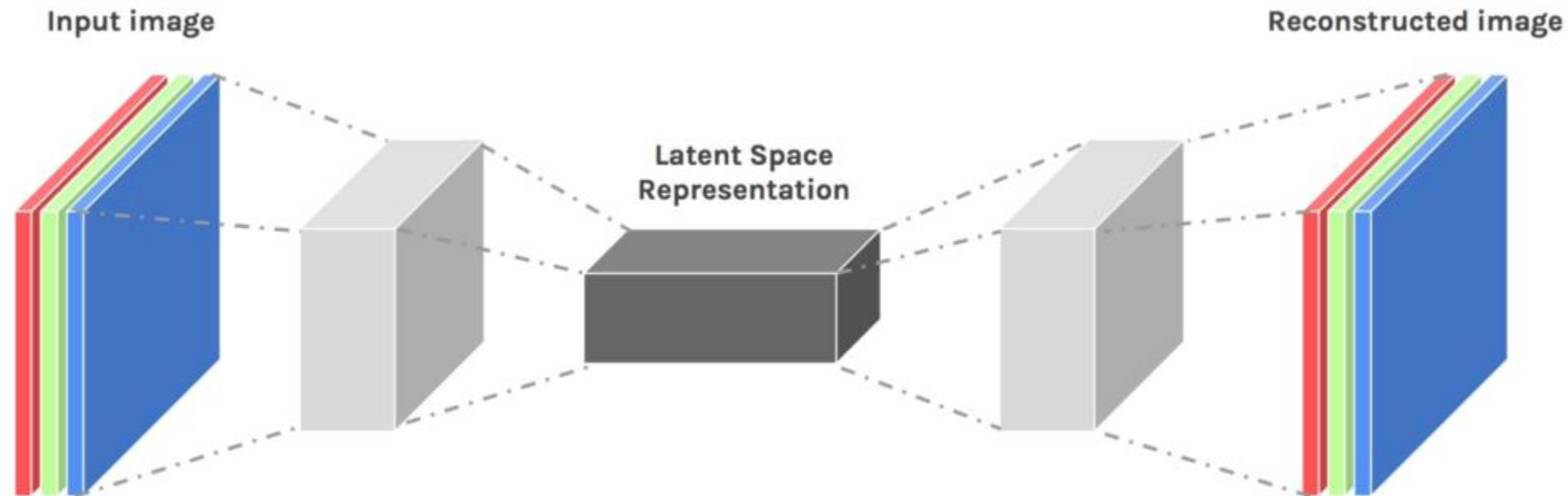
Generative Adversarial Network (GAN)

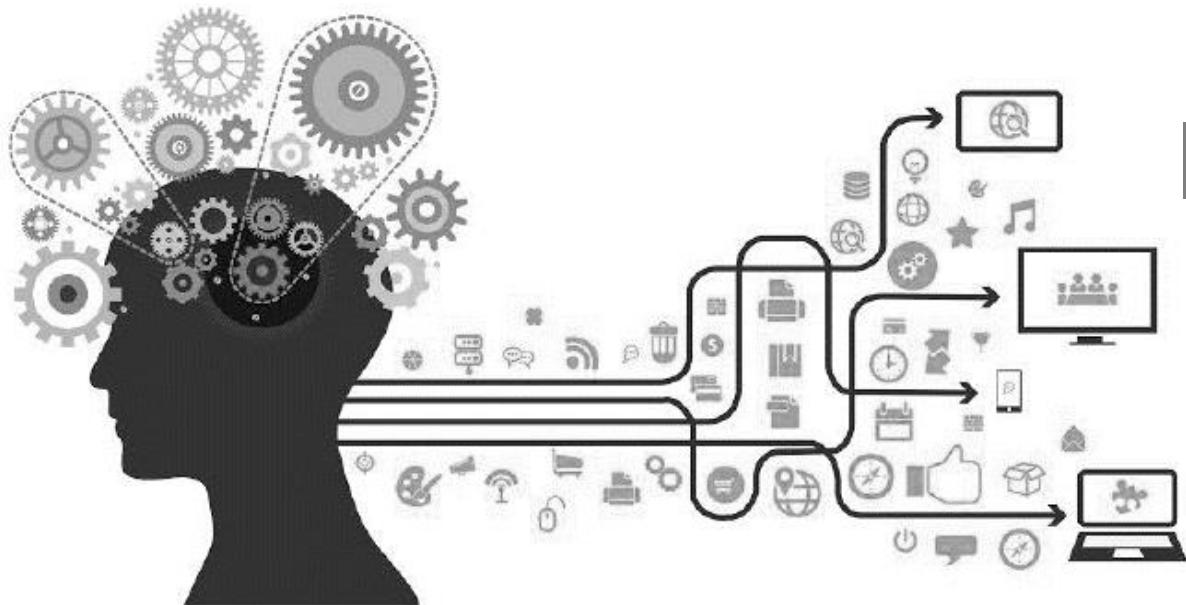
Se tiene un generador y un discriminador. El generador genera muestras falsas de datos (ya sea una imagen, audio, etc.) e intenta engañar al discriminador. El Discriminador, por otro lado, trata de distinguir entre las muestras reales y las falsas. El Generador y el Discriminador son redes neuronales que compiten entre sí en la fase de entrenamiento..



Autoencoders

Redes neuronales diseñadas para aprendizaje no supervisado, es un modelo de compresión de datos, se pueden usar para codificar una entrada determinada en una representación de dimensión más pequeña. Luego se puede usar un decodificador para reconstruir la entrada desde la versión codificada.





Inteligencia Artificial: Redes Neuronales y Deep learning usando R

Gracias

Por: **Néstor Montaño P.**

Masapp S.A.

R User Group Ecuador

Sociedad Ecuatoriana de Estadística