



XUNTA
DE GALICIA

CENTRO DE
FORMACIÓN E
RECURSOS DE FERROL

Inteligencia Artificial para la Sociedad



Alma Mallo (UDC)

Francisco Bellas (UDC)

Noviembre de 2024



Organización del curso

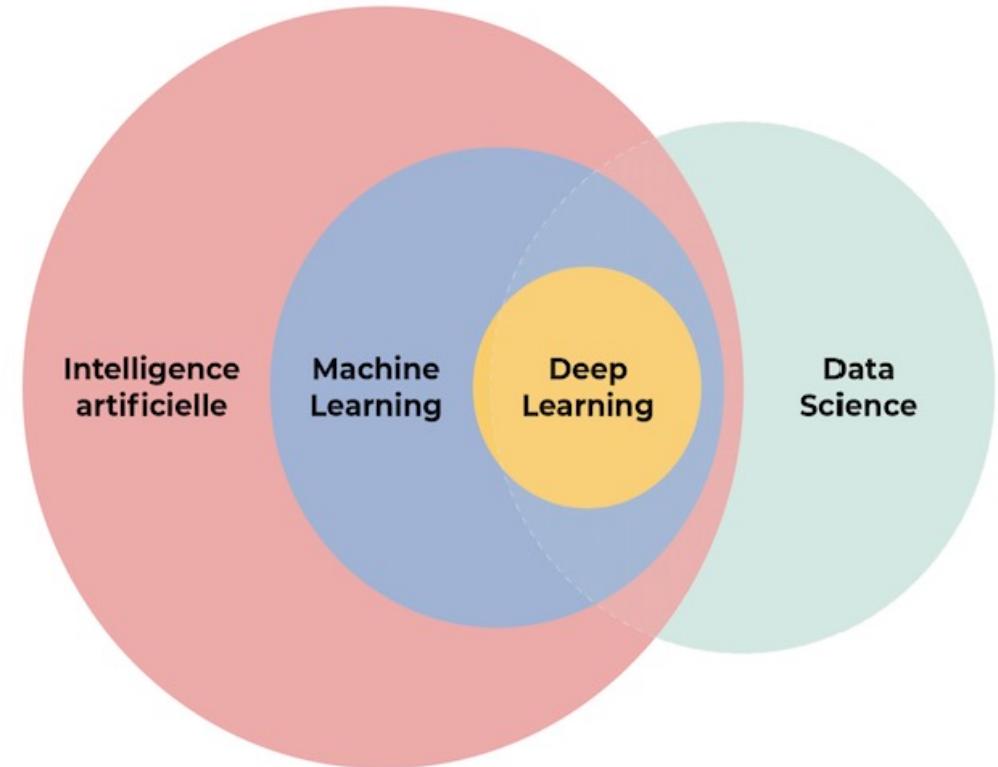
1. Sesión 1: introducción a la IA y aspectos ético-legales
2. Sesión 2: percepción y actuación
3. Sesión 3: representación y razonamiento
- 4. Sesión 4: aprendizaje automático 1**
5. Sesión 5: aprendizaje automático 2
6. Sesión 6: herramientas de IA para la educación (IA generativa)



Sesión 4: aprendizaje automático

Aprendizaje automático

- El **aprendizaje automático** es un campo de la inteligencia artificial en el cual se desarrollan algoritmos que permiten a los sistemas resolver tareas sin ser explícitamente programados para eso.
- En lugar de escribir el código de un programa, **se alimenta con datos al algoritmo**, y este construye la lógica (el **modelo**) en base a dichos datos.



Fonte: <https://medium.com/geekculture/machine-learning-data-science-and-artificial-intelligence-a45a2ffe9639>

Ciencia de datos

- La ciencia de datos es un campo interdisciplinario que combina la estadística, informática y el conocimiento específico de un dominio para extraer información útil a partir de datos.

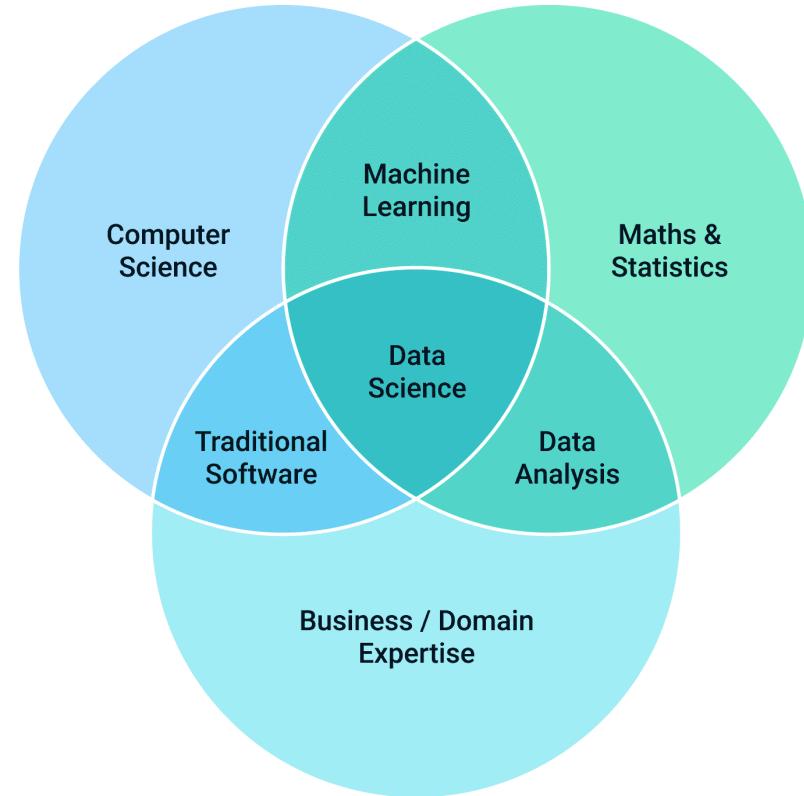
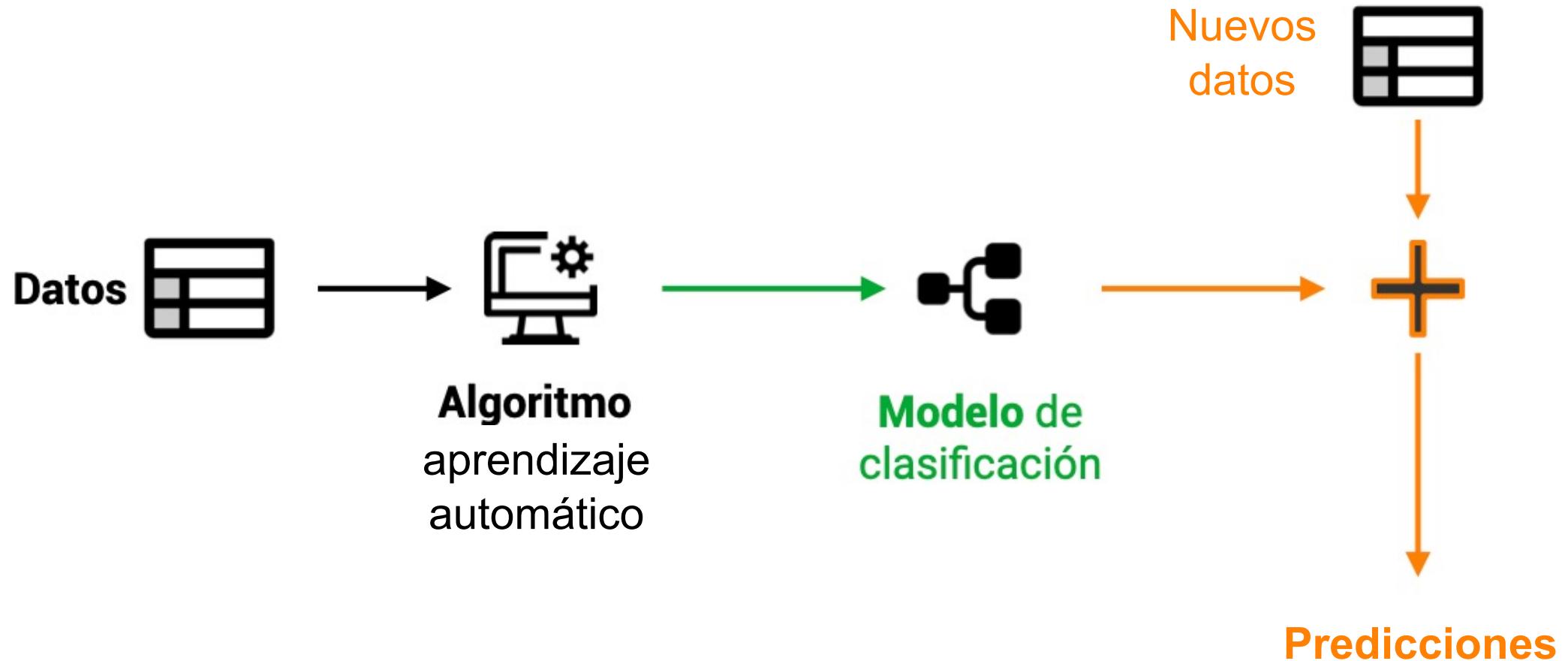


Imagen: <https://codilime.com/blog/data-science-vs-data-analytics/>

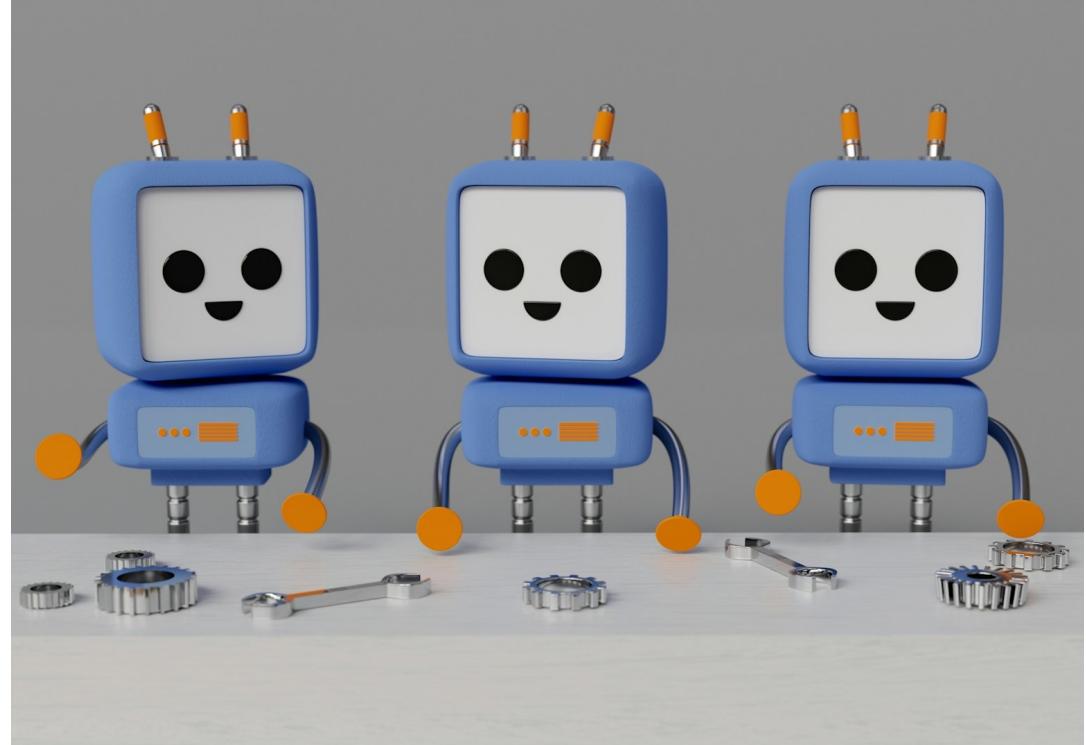
Aprendizaje automático



Tipos de aprendizaje automático

Existen tres tipos básicos:

- Aprendizaje supervisado
- Aprendizaje no supervisado
- Aprendizaje por refuerzo



Aprendizaje supervisado

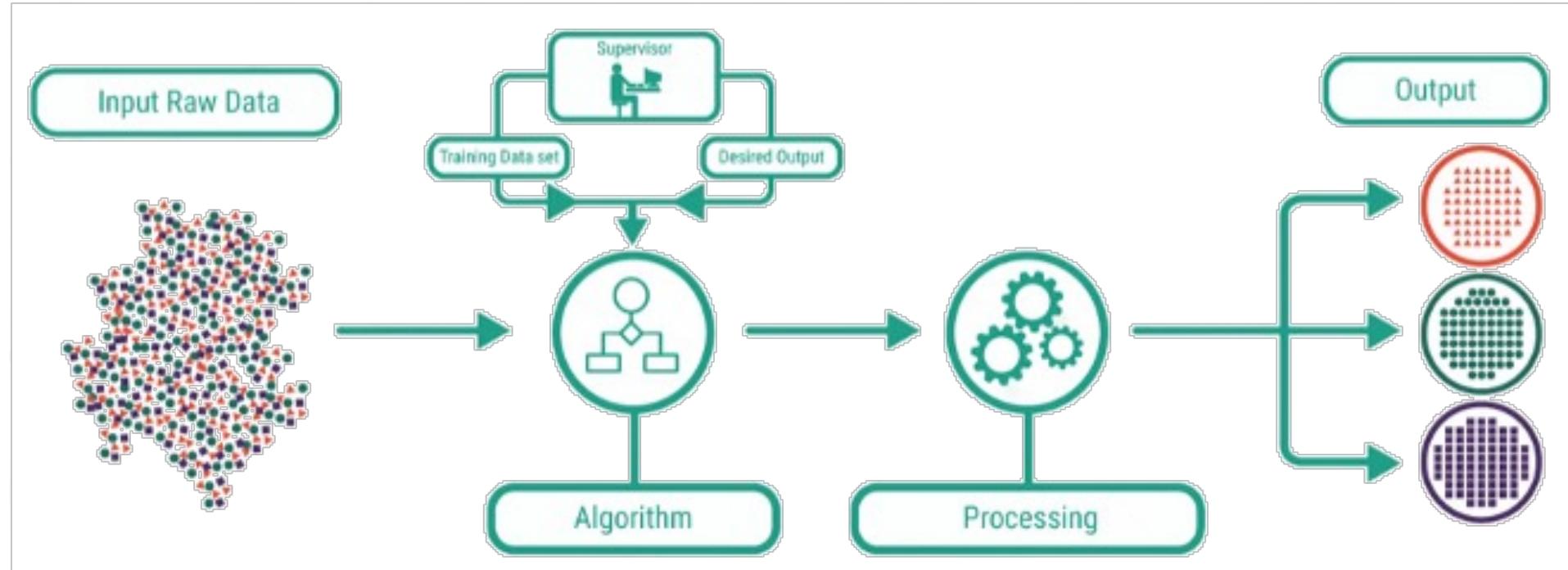


Imagen: <https://www.digitalvidya.com/blog/supervised-learning/>



Aprendizaje supervisado

Age	Sex	BP	Cholesterol	Na	K	Drug
23	F	HIGH	HIGH	0.793	0.031	drugY
47	M	LOW	HIGH	0.739	0.056	drugC
47	M	LOW	HIGH	0.697	0.069	drugC
28	F	NORMAL	HIGH	0.564	0.072	drugX
61	F	LOW	HIGH	0.559	0.031	drugY
22	F	NORMAL	HIGH	0.677	0.079	drugX
49	F	NORMAL	HIGH	0.79	0.049	drugY
41	M	LOW	HIGH	0.767	0.069	drugC
60	M	NORMAL	HIGH	0.777	0.051	drugY
43	M	LOW	NORMAL	0.526	0.027	drugY



Muestras clasificadas

Age	Sex	BP	Cholesterol	Na	K	Drug
36	F	LOW	HIGH	0.697	0.069	

Nueva muestra
Valor a predecir

Aprendizaje no supervisado

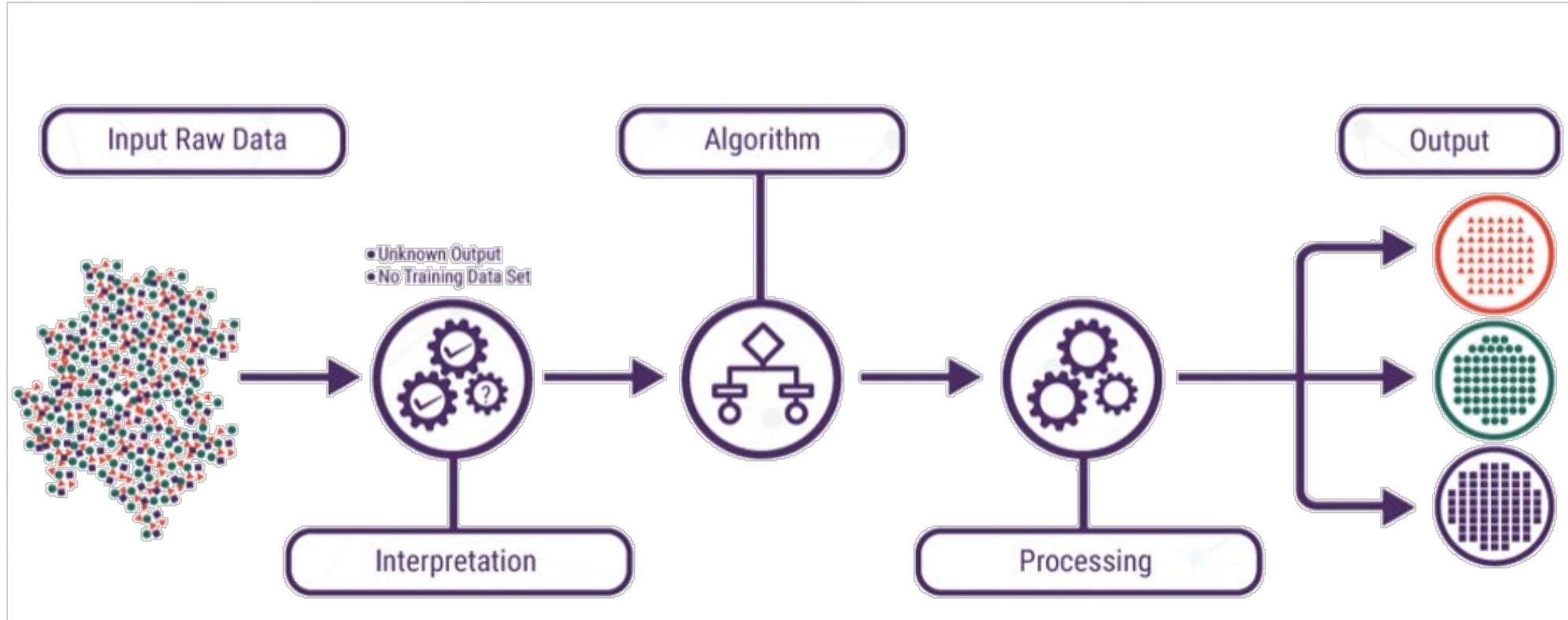
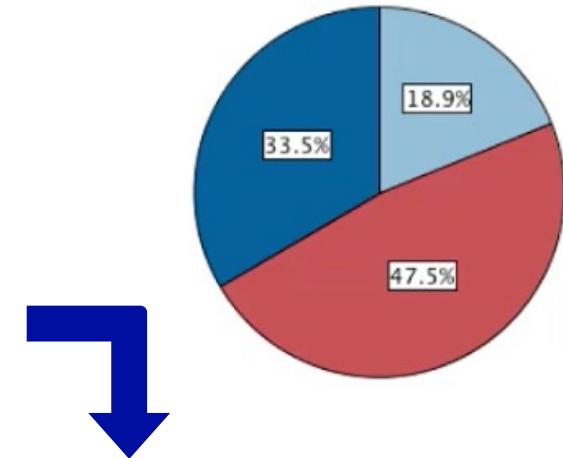


Imagen: <https://aditi22prerna.medium.com/unsupervised-learning-a24caf362e79>

Aprendizaje no supervisado

Customer Id	Age	Edu	Years Employed	Income	Card Debt	Other Debt	Address	DebtIncomeRatio
1	41	2	6	19	0.124	1.073	NBA001	6.3
2	47	1	26	100	4.582	8.218	NBA021	12.8
3	33	2	10	57	6.111	5.802	NBA013	20.9
4	29	2	4	19	0.681	0.516	NBA009	6.3
5	47	1	31	253	9.308	8.908	NBA008	7.2
6	40	1	23	81	0.998	7.831	NBA016	10.9
7	38	2	4	56	0.442	0.454	NBA013	1.6
8	42	3	0	64	0.279	3.945	NBA009	6.6
9	26	1	5	18	0.575	2.215	NBA006	15.5



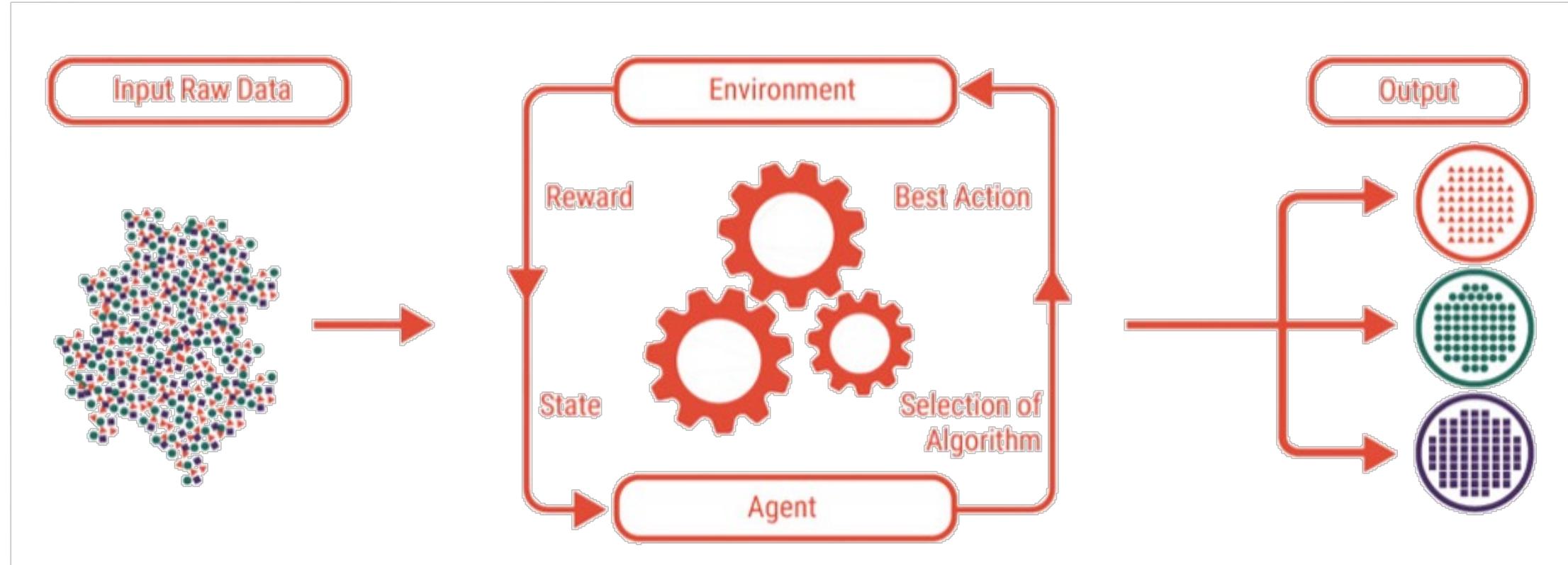
Cluster	Segment Name
cluster-1	AFFLUENT AND MIDDLE AGED
cluster-2	YOUNG EDUCATED AND MIDDLE INCOME
cluster-3	YOUNG AND LOW INCOME



Aprendizaje por refuerzo

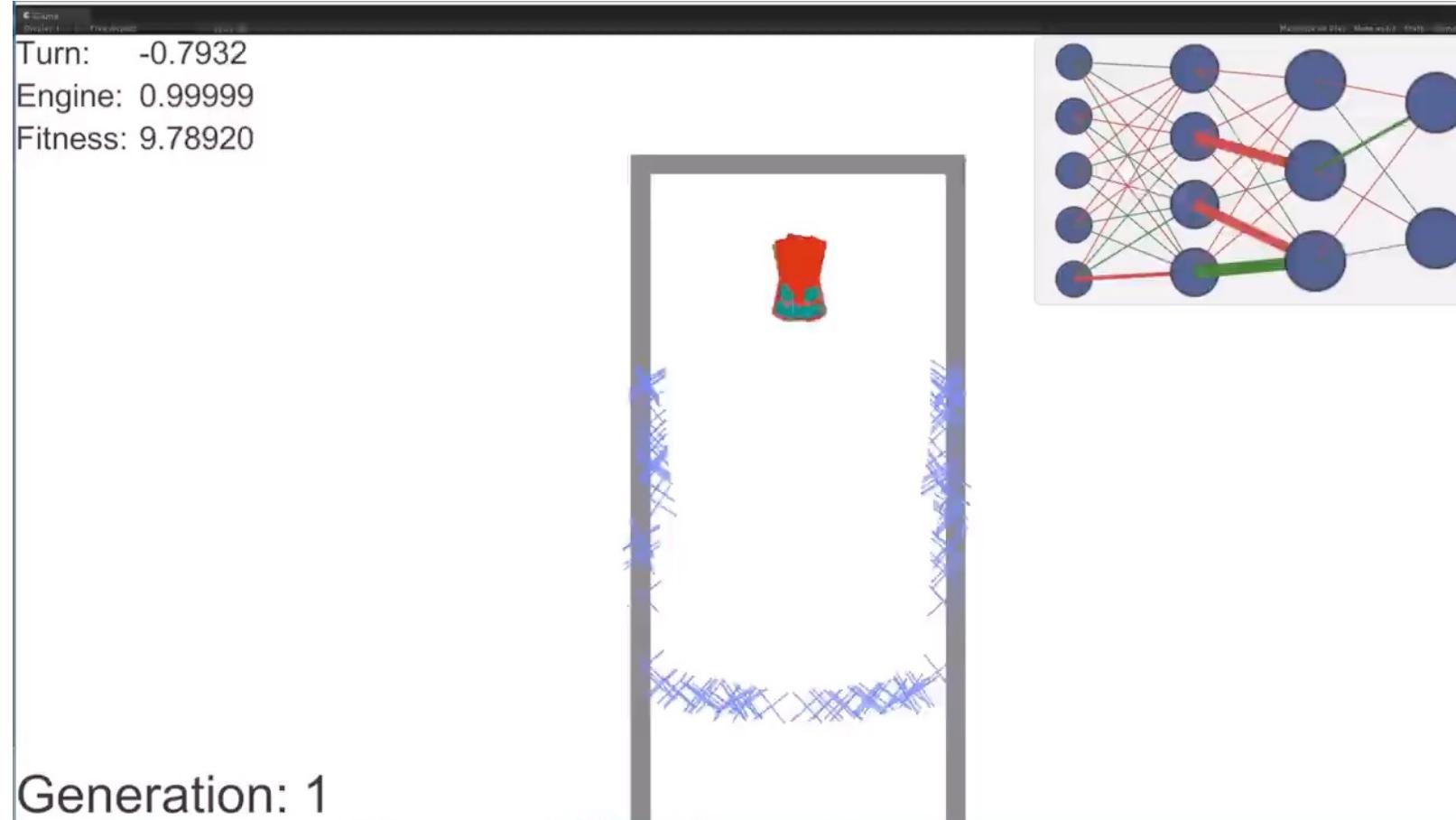
- El agente aprende mediante la **interacción con el entorno** a través de **refuerzos que serán positivos o negativos** en función de las acciones.
- Aprende por prueba y error, ajustando la estrategia a medida que avanza para mejorar su rendimiento.
- Ejemplos de uso:
 - Robótica: Control de robot para realizar tareas complejas, por ejemplo, navegación autónoma.
 - Videojuegos: crear agentes que aprendan a jugar o mejoren con la experiencia.

Aprendizaje por refuerzo





Aprendizaje por refuerzo



Deep Learning Cars- Deep Reinforcement Learning
2016. YouTube. <https://youtu.be/Aut32pR5PQA>



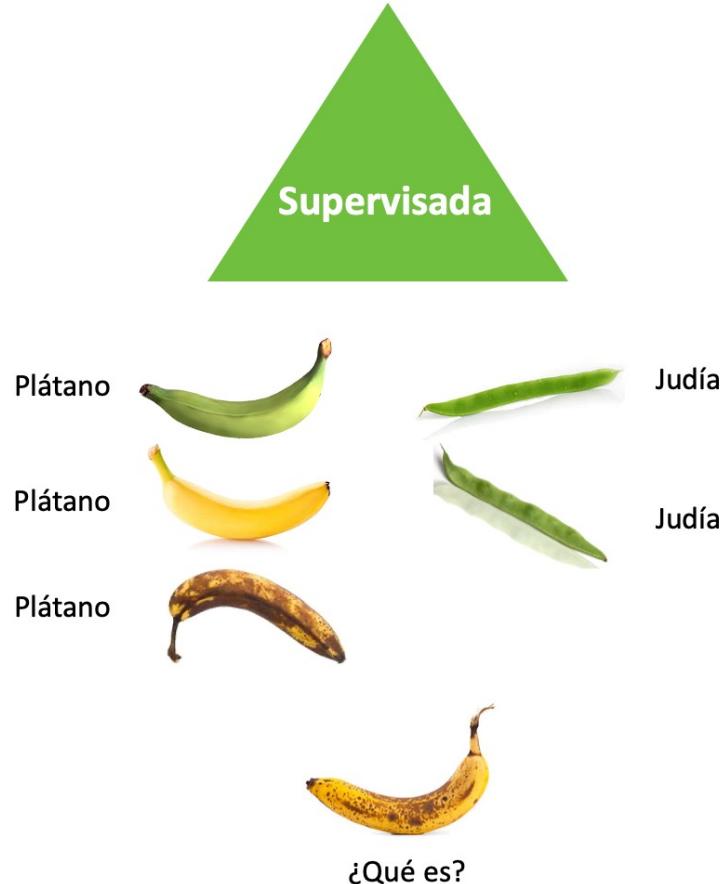
Técnicas de aprendizaje automático

Dos técnicas principales:

- **Clasificación**: consiste en tomar un vector de entradas etiquetadas. Para una nueva entrada, decide qué etiqueta le corresponde.
- **Regresión** se define como la creación de un modelo matemático que permita estimar las relaciones entre una variable dependiente y una o más variables independientes.



Tipos de clasificación



¿Cuántos grupos formarías?

Si tuvieses que formar 2 grupos, ¿cuáles serían?



Clasificación supervisada

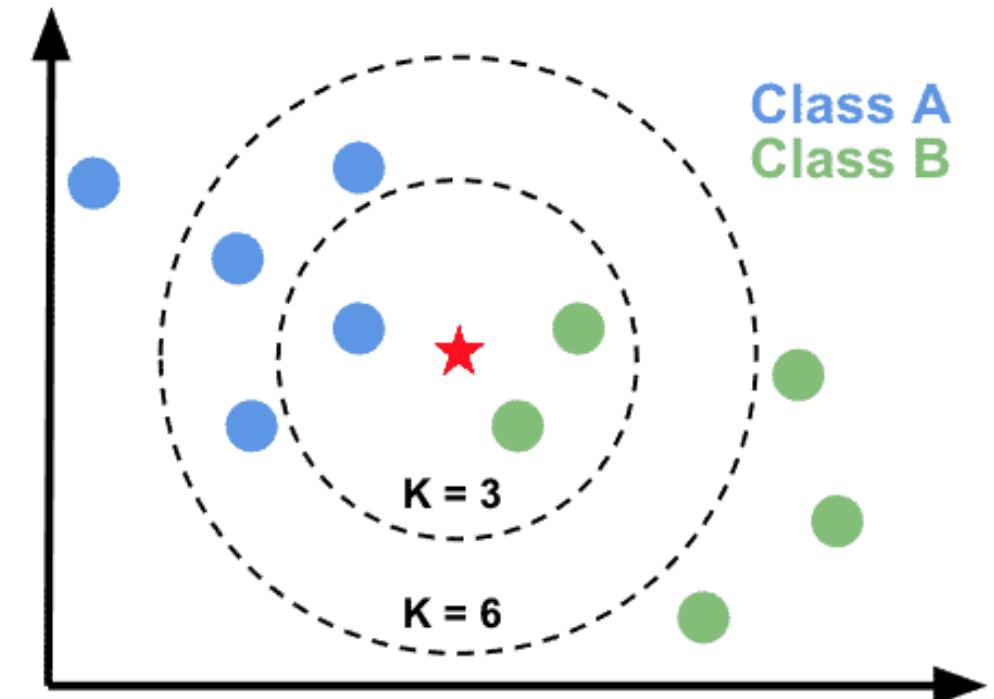
Se puede realizar con diferentes algoritmos:

- Vecinos más cercanos (kNN)
- Árboles de decisión
- Máquinas de vectores de soporte(SVM)
- Redes de neuronas
- ...

Clasificación supervisada

Vecinos más cercanos - kNN

- Basado en vecindad: Los métodos de este tipo emplean una métrica (medida de similitud o distancia) entre los distintos datos.
- Los datos pertenecientes a una misma clase, probablemente, serán más próximos en el espacio de representación.



<https://www.jcchouinard.com/k-nearest-neighbors/>

Clasificación supervisada

Árboles de decisión

- Utiliza una estructura de árbol para tomar decisiones a partir de los datos de entrada.
- Cada nodo interno, representa una pregunta sobre un atributo y cada nodo hoja una predicción.

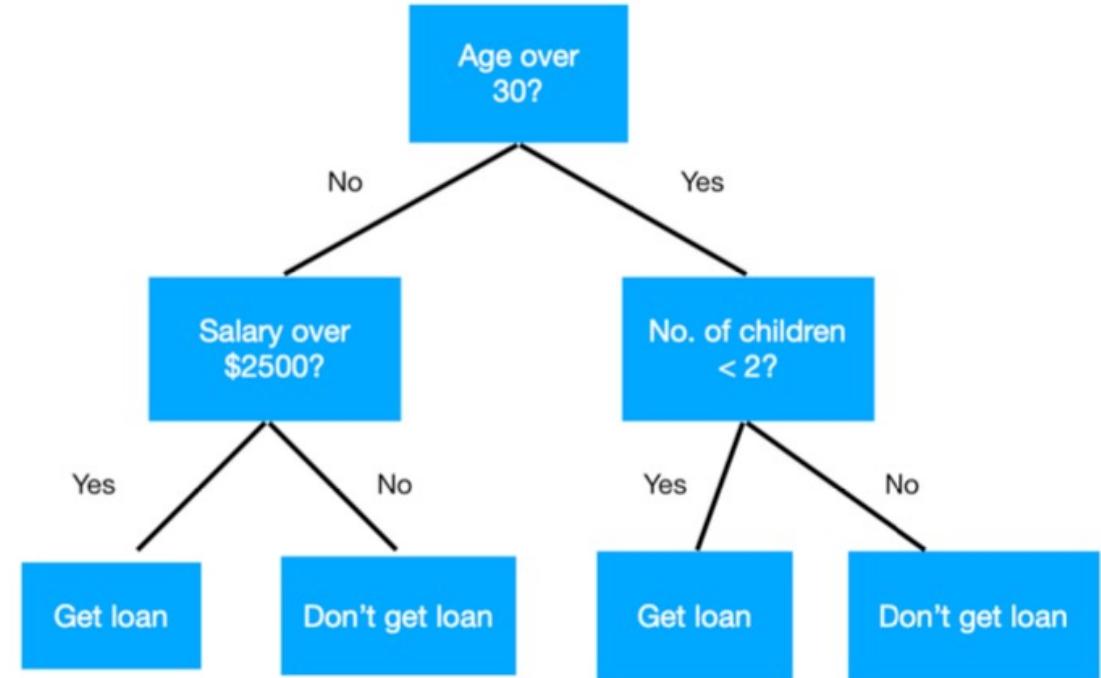


Imagen: <https://eloquentarduino.github.io/2020/10/decision-tree-random-forest-and-xgboost-on-arduino>

Clasificación supervisada

SVM

- Método matemático que consiste en buscar una frontera o hiperplano que permita separar los datos en las distintas clases.
- Puede trabajar con datos que no son linealmente separables utilizando un truco llamado núcleo (kernel).

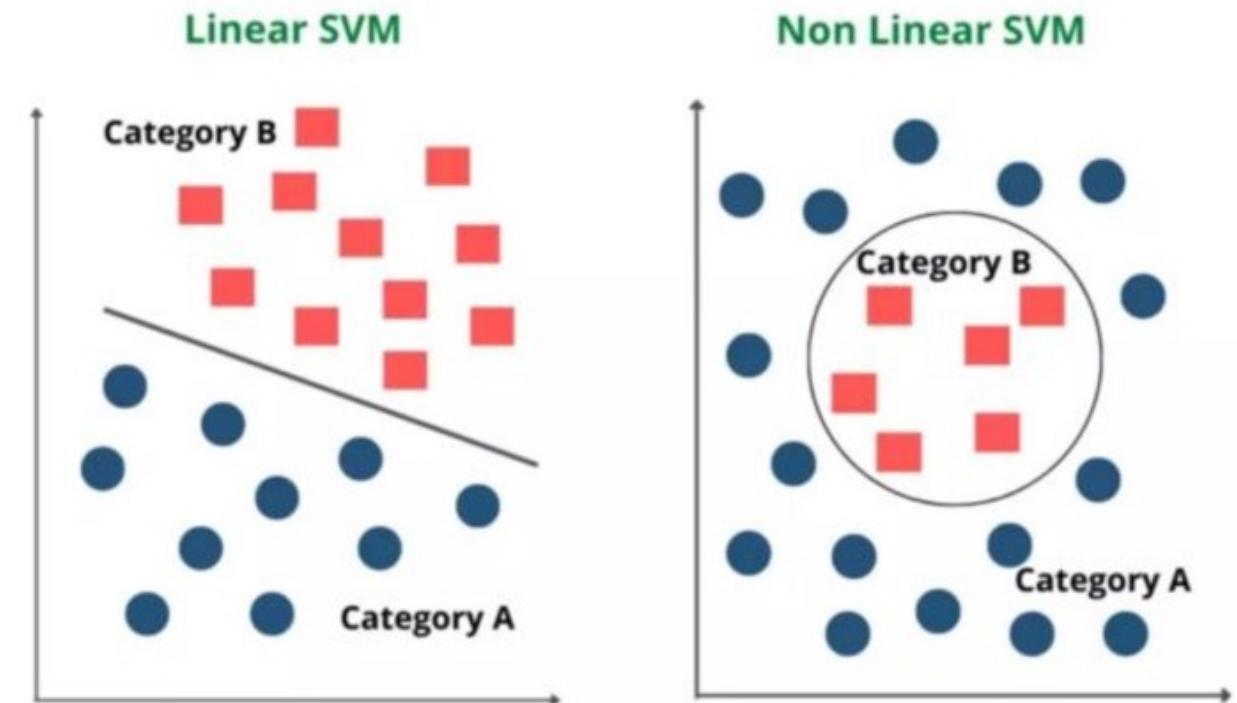
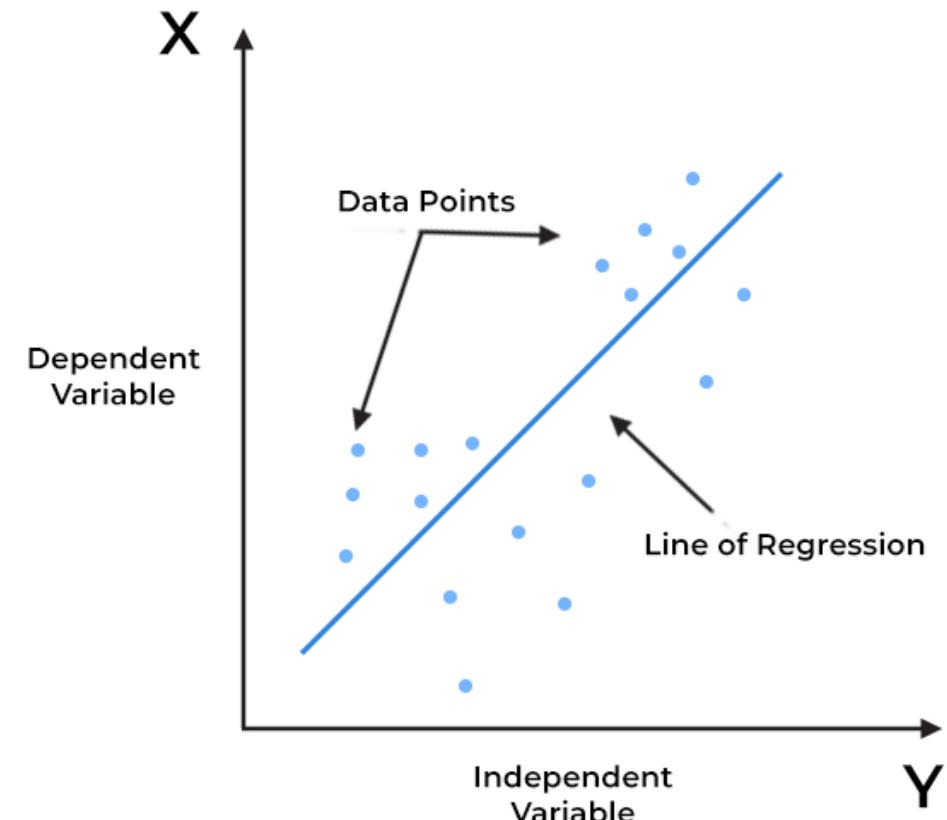


Imagen:
<https://copyassignment.com/support-vector-machine-svm-in-machine-learning>

Regresión

- El problema de regresión se define como la creación de un modelo matemático que permita estimar las relaciones entre una variable dependiente y una o más variables independientes.
- Se resuelve principalmente con:
 - Modelos lineales
 - Redes de neuronas



<https://www.spiceworks.com/tech/artificial-intelligence/articles/what-is-linear-regression>



Clasificación vs Regresión

Clasificación

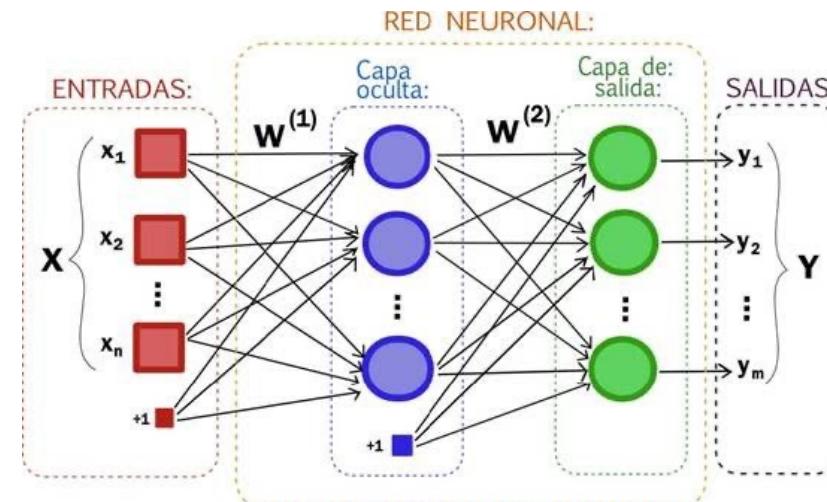
- Variables discretas / categóricas
- El objetivo es predecir la categoría de nuevos datos.

Regresión

- Variables continuas.
- El objetivo es modelar un fenómeno real, o predecir respuesta futura.

Redes de neuronas artificiales

- Las redes de neuronas artificiales (RNA) son un paradigma de aprendizaje y procesamiento automático inspirado en la forma en que funciona el sistema nervioso de los animales.
 - Interconexión de neuronas en una red que produce un estímulo de salida.
 - Generan modelos matemáticos de alta dimensionalidad y capaces de generalizar patrones no lineales



Neurona biológica

- **Soma** o cuerpo celular
- **Dendritas**: conectadas a otras neuronas.
- **Axón**: conduce impulso nervioso.

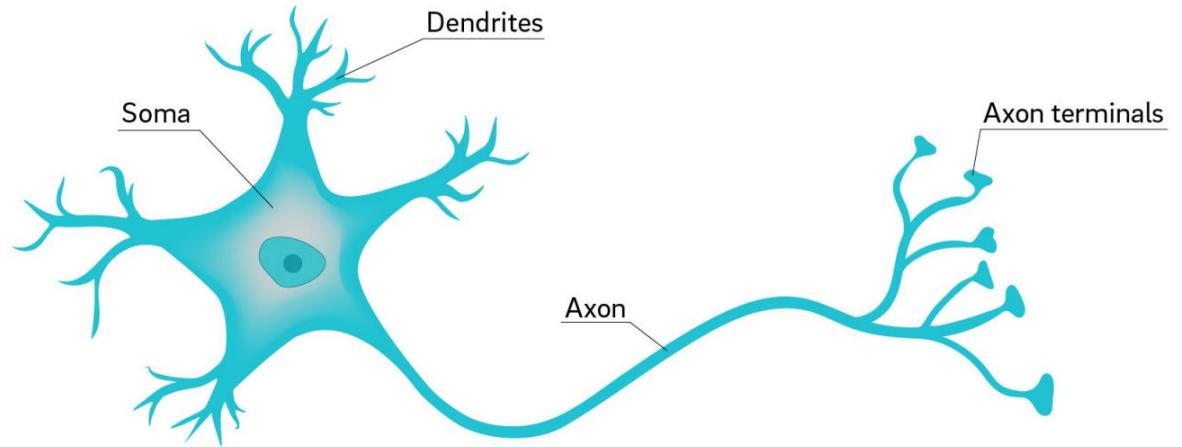


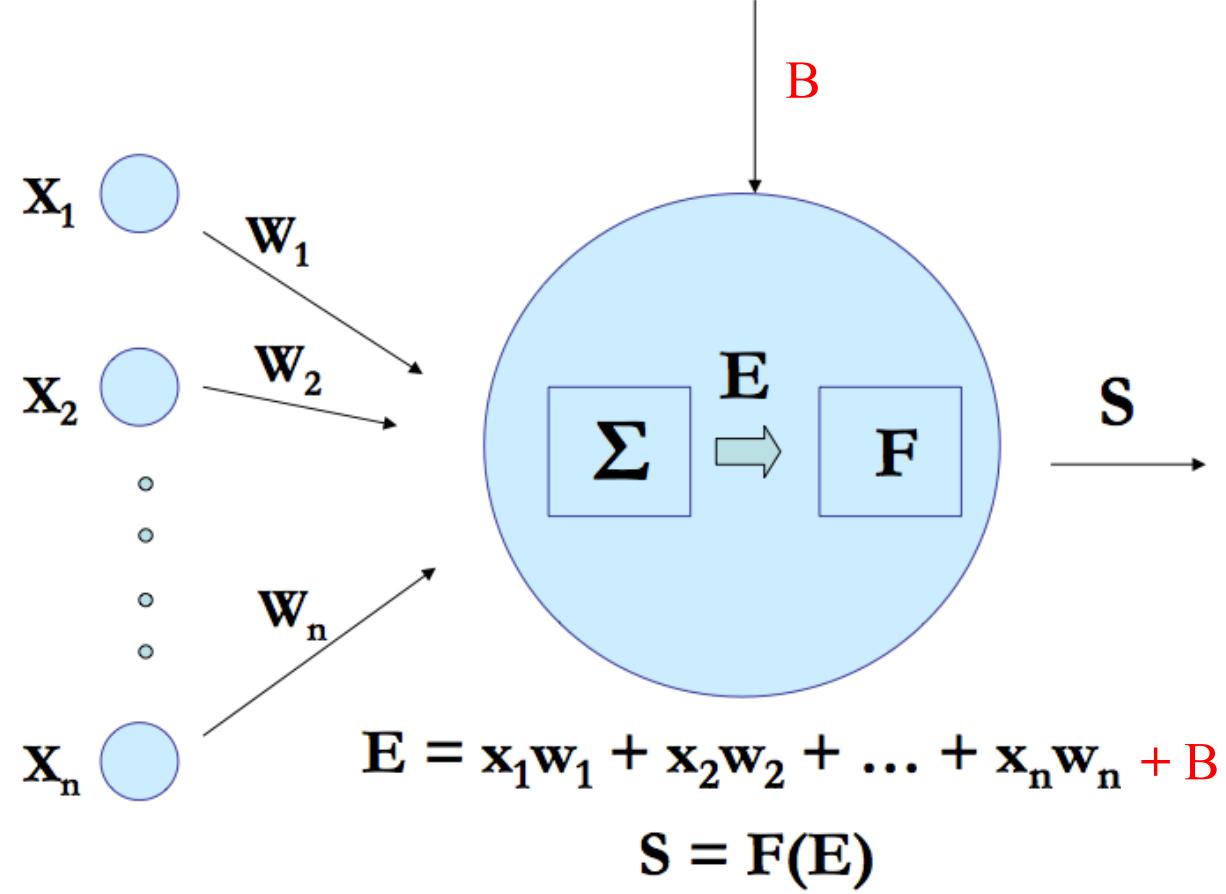
Imagen:

<https://www.upgrad.com/blog/biological-neural-network/>

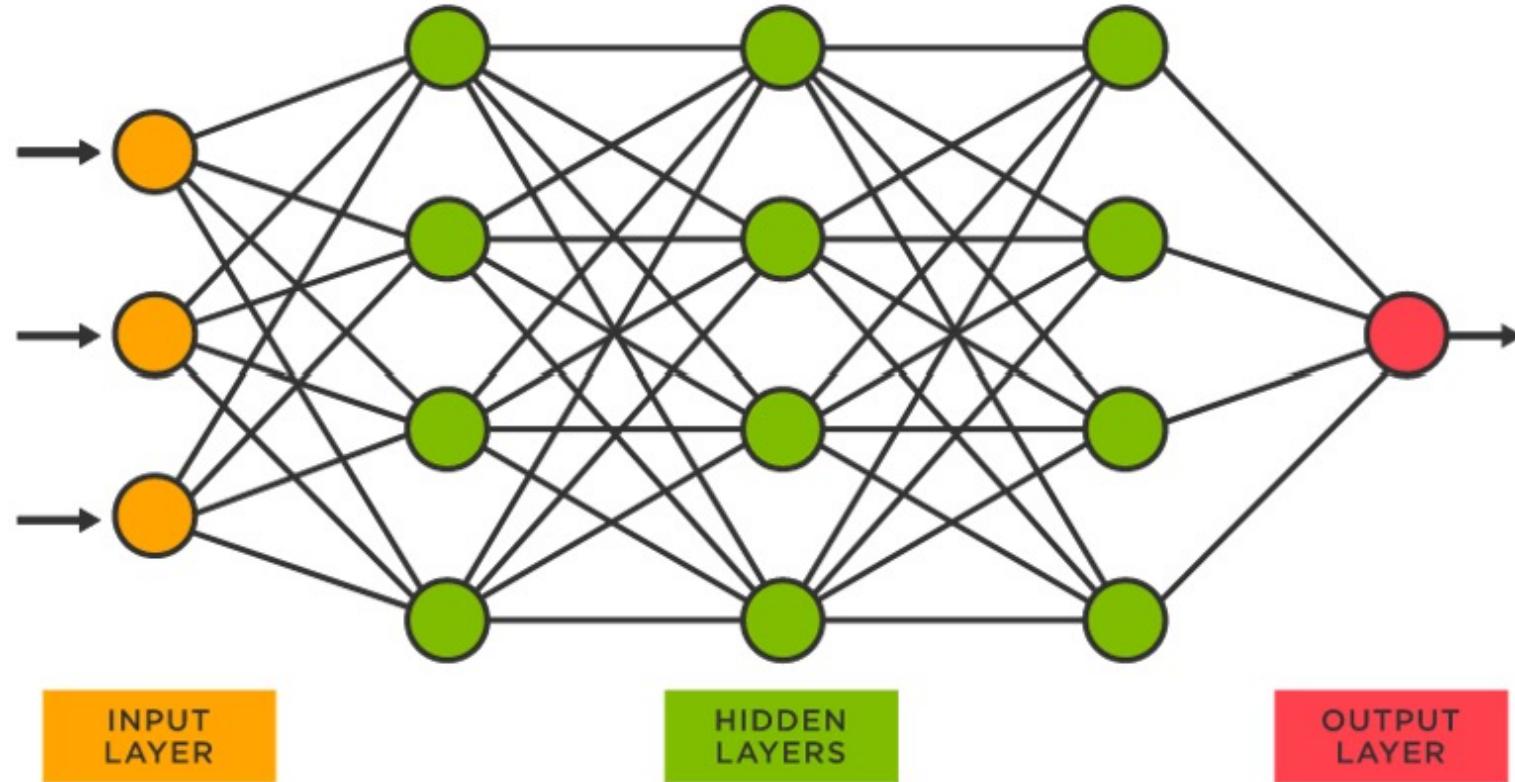
Redes de neuronas artificiales

Neurona computacional:

- Entradas
- Sinapsis (pesos, w)
- Bias
- Función de activación
- Salida



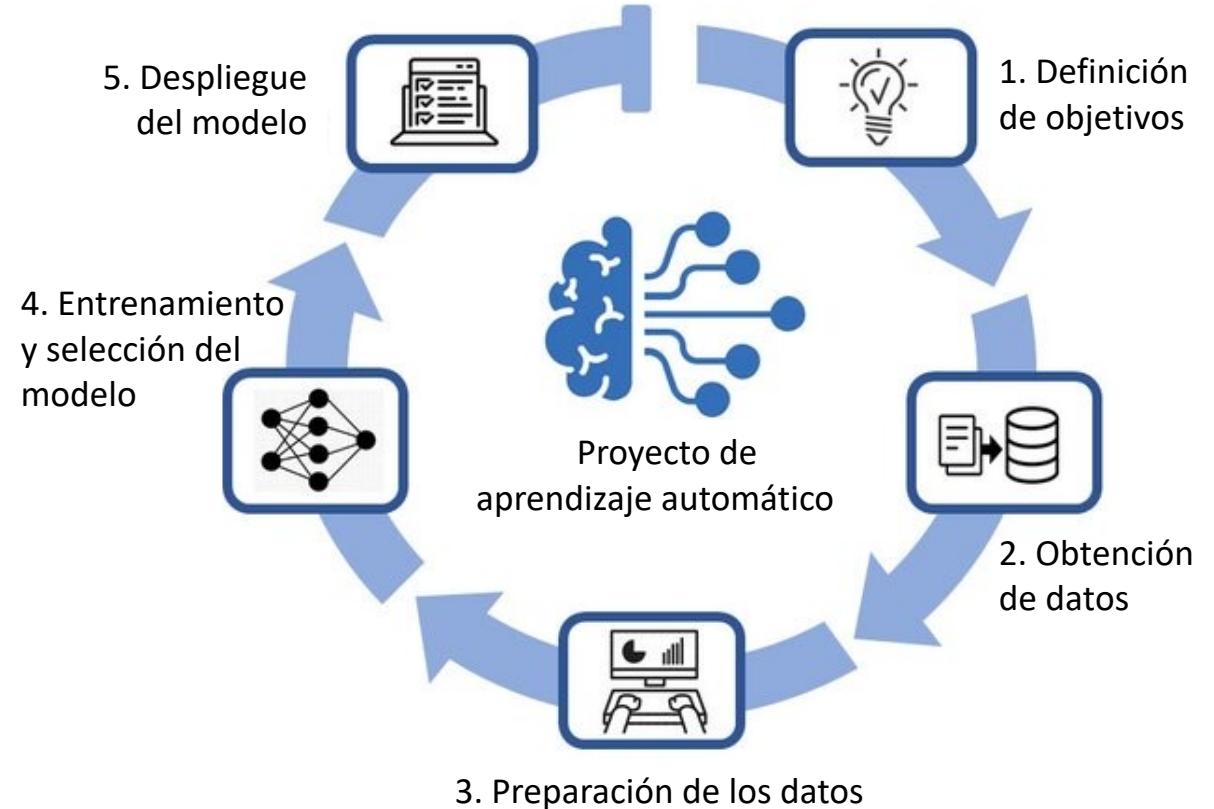
Redes neuronales



Fuente: <https://medium.com/@srempe/an-introduction-to-neural-networks-8c61b880f358>

Proyecto de aprendizaje automático

1. Definición de objetivos.
2. Obtención de los datos.
3. Preparación de los datos.
4. Entrenamiento de los modelos.
5. Despliegue del modelo.



https://www.researchgate.net/publication/364284062_Using_Machine_Learning_for_Dynamic_Authentication_in_Telehealth_A_Tutorial

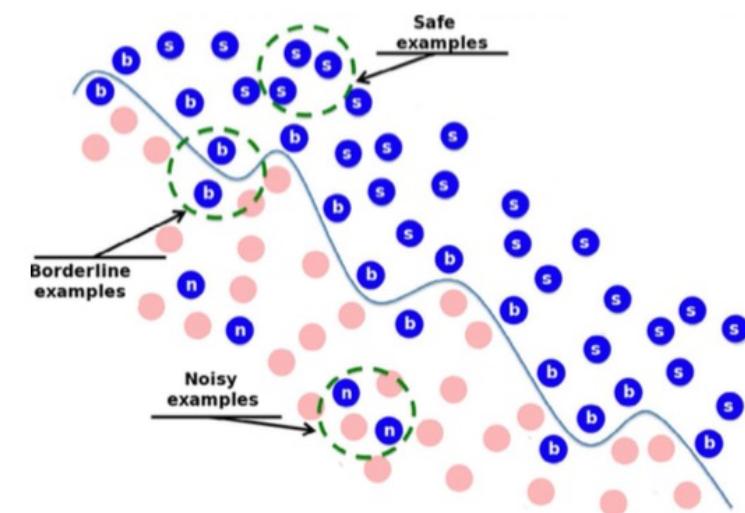
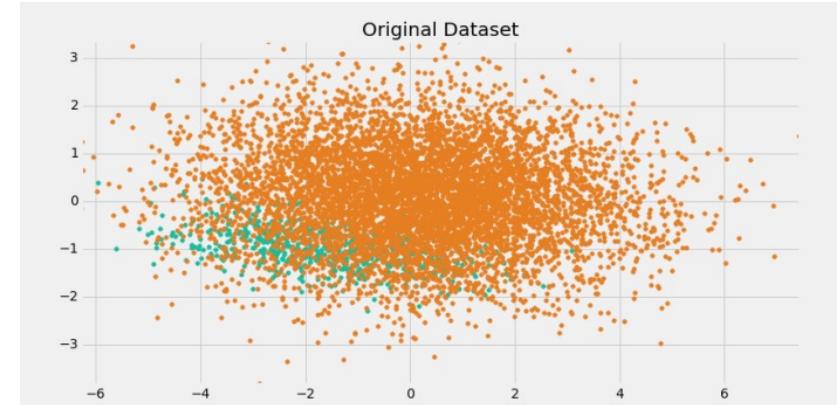


Obtención de los datos

- Los datos para entrenar y evaluar del modelo, en un caso real:
 - Datos históricos de la empresa, organización ...
 - Generados para el propio proyecto → difícil ya que la cantidad de datos necesaria para obtener un modelo válido es muy alta.
- Existen repositorios con conjuntos de datos:
 - Contienen variedad de conjunto de datos para proyectos de aprendizaje automático de con distintos objetivos: clasificación, regresión, reconocimiento de objetos, ...
- Ejemplos de repositorios:
 - Kaggle: <https://www.kaggle.com/datasets>
 - UCI repository: <http://archive.ics.uci.edu/datasets>

Preparación de los datos

- Antes de aplicar cualquier algoritmo de aprendizaje se deben analizar los datos para:
 - Tratar o eliminar casos con **información incompleta**
 - Tratar o eliminar, si es posible, **casos atípicos y ruido.**
 - **Balancear** los conjuntos de datos (número de muestras similares de cada caso).
 - **Reescalar** los datos para que tengan rangos similares.



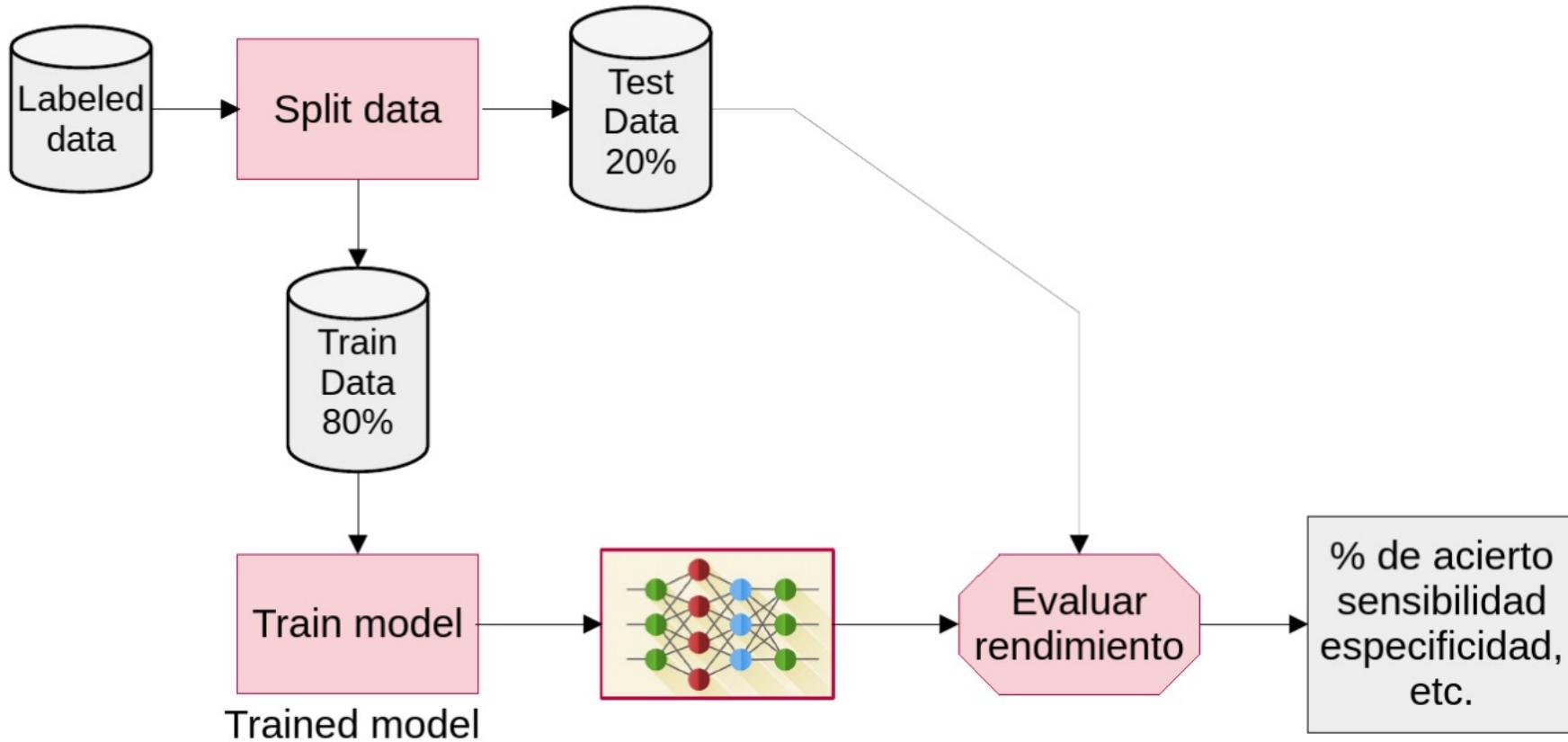


División del conjunto de datos

Cuando los datos están preparados, se dividen en dos o tres conjuntos:

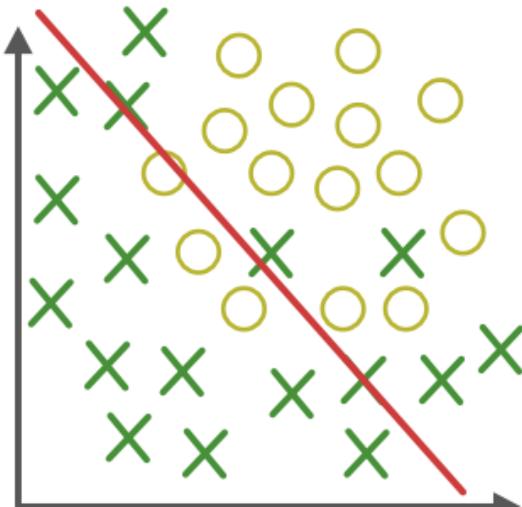
- **Conjunto de entrenamiento:** Para que el modelo "aprenda" los patrones a partir de los datos.
- **Conjunto de validación (opcional):** para ajustar los hiperparámetros del modelo y mejorarlo.
- **Conjunto de test:** Para evaluar la capacidad de generalización del modelo con datos nuevos que nunca vio antes.

Entrenamiento

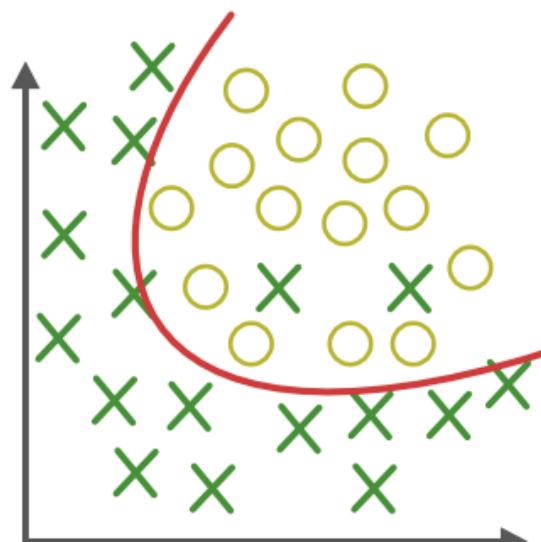


Entrenamiento

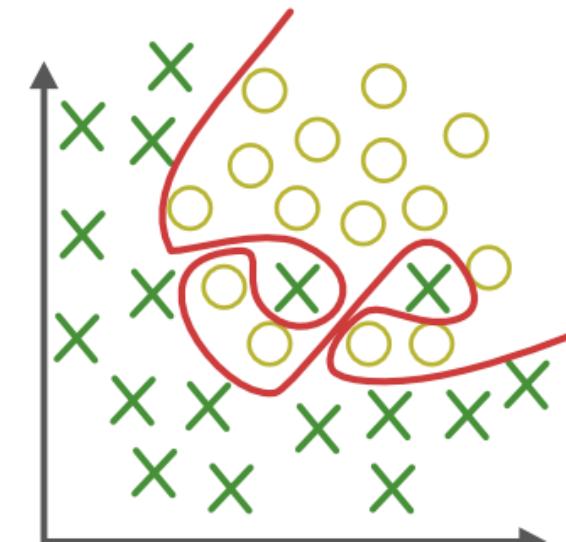
Tipos de ajuste del modelo tras el entrenamiento:



Sub-ajuste



Ajuste adecuado



Sobre-ajuste

Imagen: <https://www.geeksforgeeks.org/underfitting-and-overfitting-in-machine-learning/>



Explicabilidad

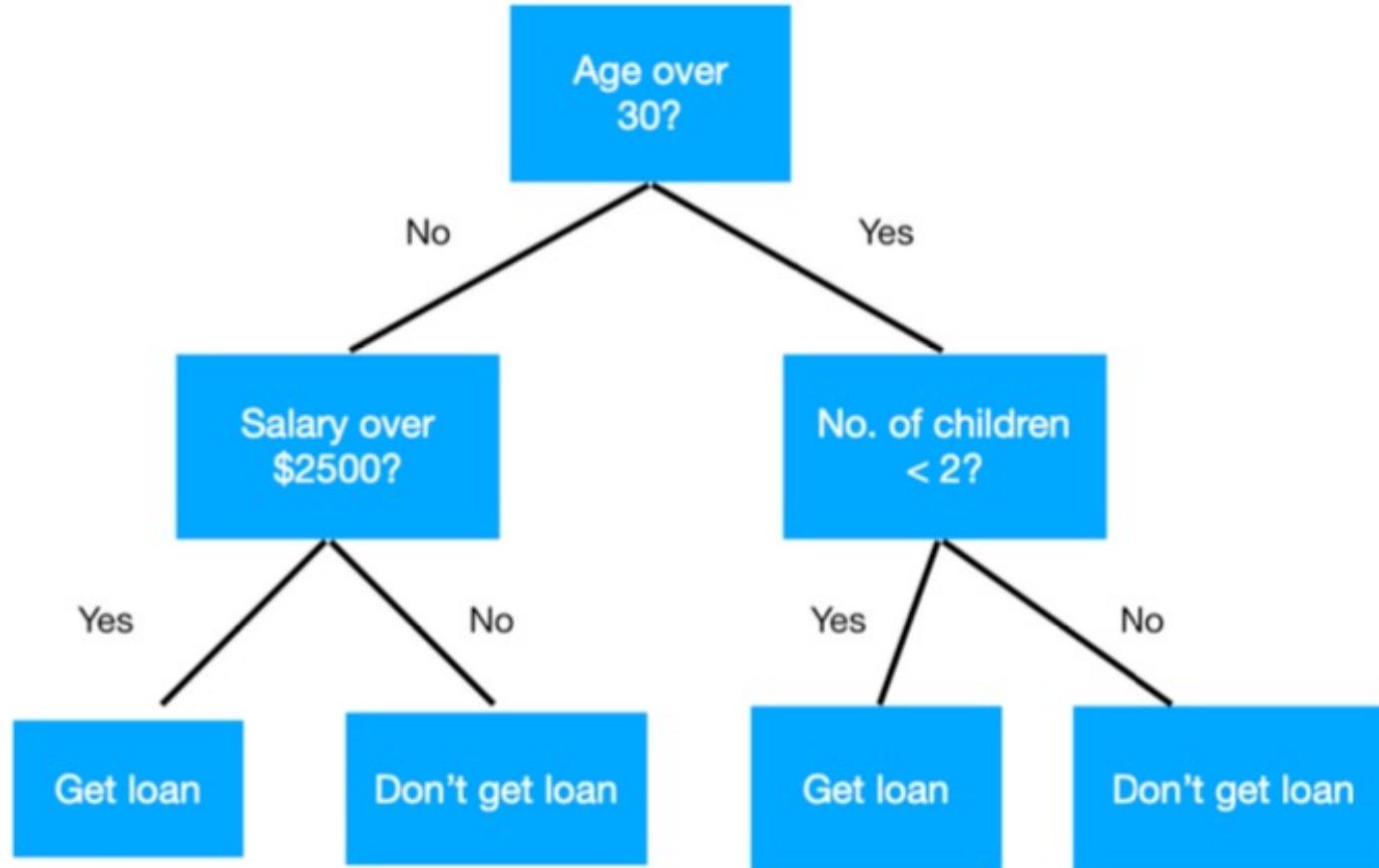
- IA explicable (interpretable) vs IA no explicable.
- **IA explicable:** sistemas de IA cuyas decisiones pueden ser fácilmente comprendidas y explicadas por los seres humanos.
 - **Objetivo:** proporcionar claridad y transparencia. Permiten entender como y por qué el modelo tomó una decisión particular.



Explicabilidad

- **IA no explicable:** Se refiere a sistemas que producen resultados muy precisos pero que sus decisiones son muy difíciles o imposibles de entender para los seres humanos.
 - Se llaman modelos de **caja negra** porque, aunque proporcionan una respuesta, el proceso interno que llevó a esa respuesta no es transparente ni fácilmente interpretable.
 - Es un problema para situaciones donde se requiere una explicación detallada.
- Ejemplo: redes de neuronas profundas, utilizadas por ejemplo para reconocimiento de imágenes.

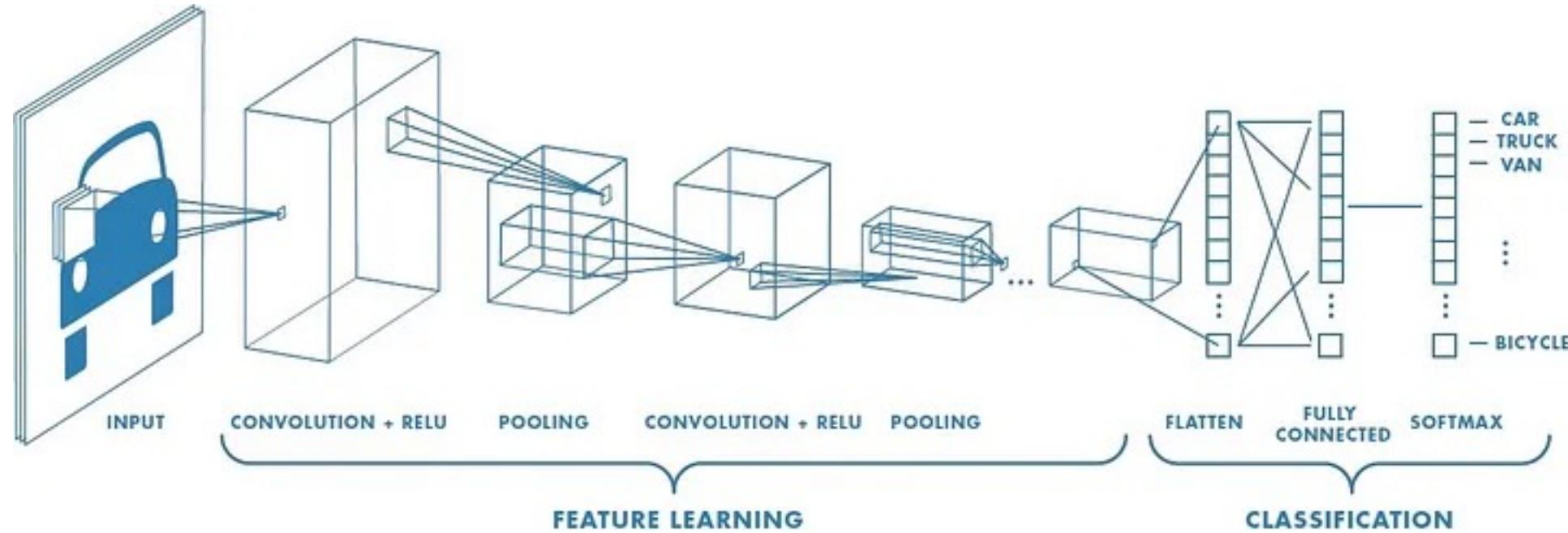
Explicabilidad: IA explicable



Fuente: <https://eloquentarduino.github.io/2020/10/decision-tree-random-forest-and-xgboost-on-arduino>

Explicabilidad: IA no explicable

- Redes convolucionales para clasificación de imágenes.



Fuente: <https://towardsdatascience.com/a-comprehensive-guide-to-convolutional-neural-networks-the-eli5-way-3bd2b1164a53>



Modelos pre-entrenados

- El entrenamiento de algoritmos para obtener un modelo es un proceso lento y costoso computacionalmente. Además, suele requerir gran cantidad de datos.
 - La situación empeora drásticamente cuando los datos son imágenes.
- Existen **modelos pre-entrenados** con conjuntos de imágenes de propósito general que pueden ser utilizados en nuestros programas.
- El uso de modelos pre-entrenados se conoce como **Transfer learning** (**transferencia de aprendizaje**).



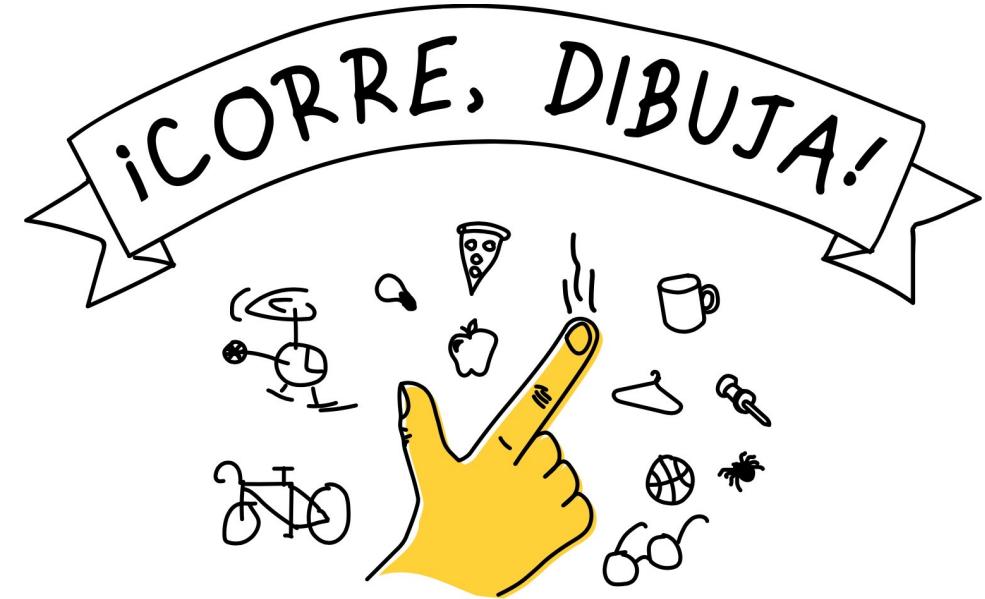
Actividades prácticas

Aprendizaje automático



Quickdraw

- <https://quickdraw.withgoogle.com>
- Red neuronal que reconoce dibujos
- Interesante con estudiantes jóvenes o como muestra de concepto
- Utiliza nuestros dibujos para continuar el entrenamiento



¿Puede una red neuronal reconocer tus dibujos?

Añade tus dibujos al [conjunto de datos de dibujos más grande del mundo](#), compartido públicamente, para ayudarnos con la investigación sobre el aprendizaje automático.

¡A dibujar!



IA para los océanos

- Aprende acerca de la inteligencia artificial (IA), el aprendizaje automático, los datos de entrenamiento y el sesgo, mientras exploras los problemas éticos y cómo la IA puede usarse para abordar los problemas mundiales

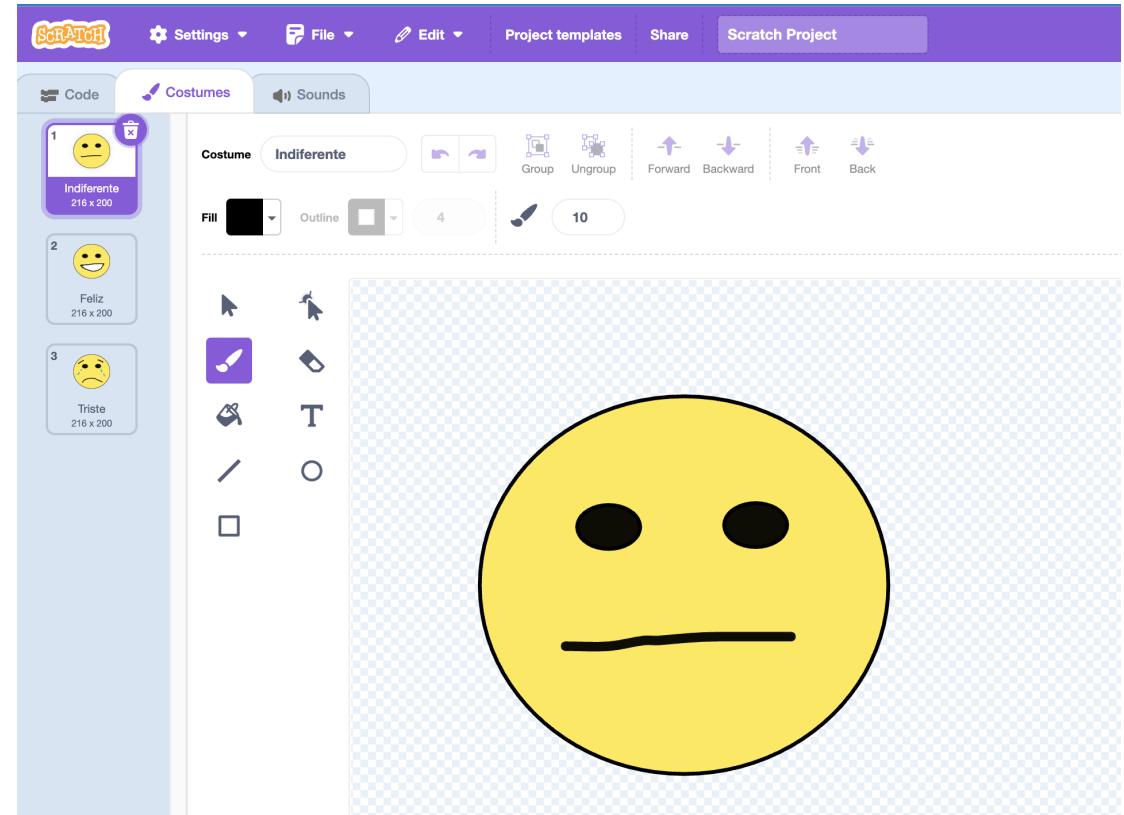


<https://studio.code.org/s/oceans/lessons/1/levels/1>



ML4K

- Make me Happy:
 - En este proyecto crearás un personaje que reaccionará a lo que le digas.
 - Si lo felicitas, parecerá contento.
 - Si lo insultas, parecerá triste.
 - <https://machinelearningforkids.co.uk/#!/worksheets>
 - Hoja de trabajo en español en el aula virtual:
 - [Actividad: Make me happy](#)





- Face finder:
 - En este proyecto crearás un filtro facial basado en IA utilizando un modelo pre-entrenado.
 - <https://machinelearningforkids.co.uk/#!/worksheets>
 - Hoja de trabajo en español en el campus virtual
 - [Actividad: filtro facial](#)





LearningML

- Una forma rápida y sencilla de crear modelos de aprendizaje automático.
- Tiene un modo básico y otro avanzado.
 - Básico: <https://basic.learningml.org/editor/>
 - Avanzado: <https://advanced.learningml.org/editor/>
 - En el modo avanzado, se puede cambiar el algoritmo y la configuración de los hiperparámetros.



LearningML

LML



Archivo ▾

sin nombre

Aprende ▾



Modo avanzado activado

Acerca de

1. Entrenar

Primero necesito algunas imágenes de ejemplo

Añadir nueva clase de imágenes

2. Aprender

Llegó el momento de aprender a clasificar imágenes

Elije el algoritmo de ML para construir ...

KNN

Nº de vecinos 5

Porcentaje de ejemplos para validación

0

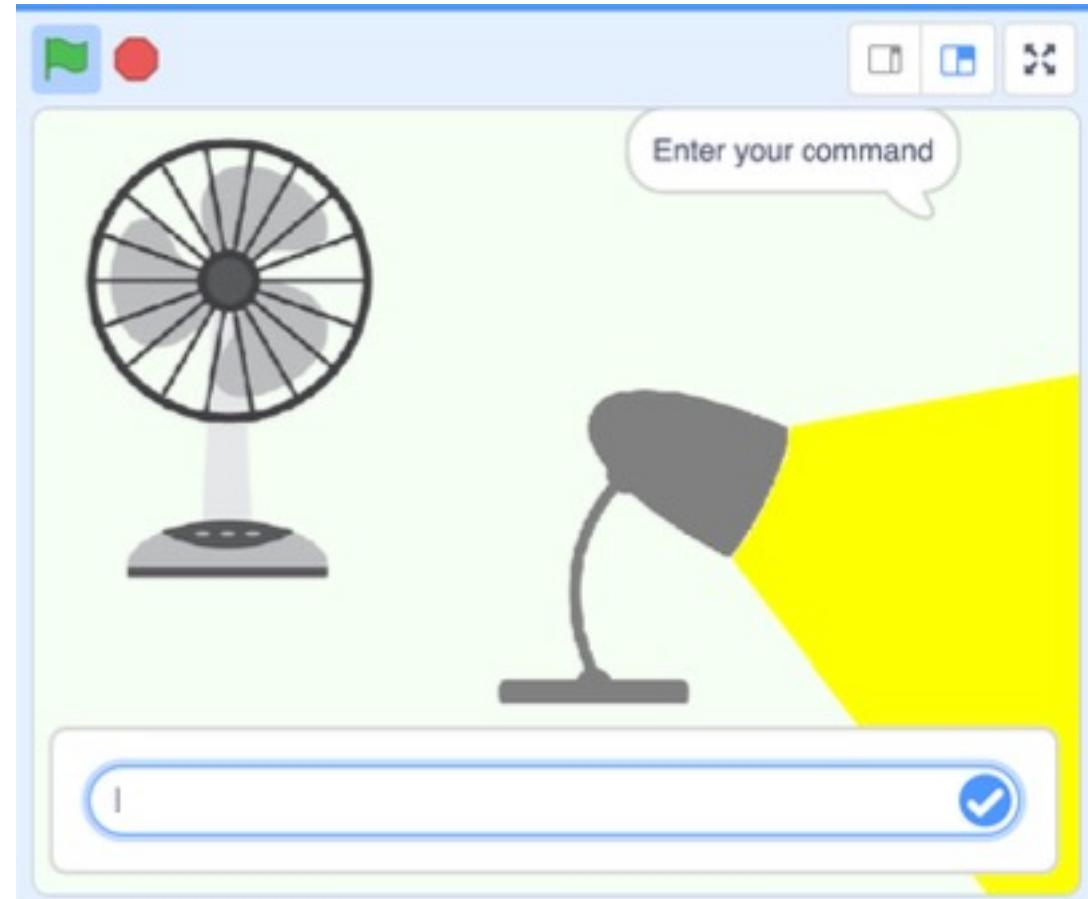
Aprender a reconocer imágenes

[LearningML](#) Copyright © 2022 Juan David Rodríguez García



ML4K y LearningML

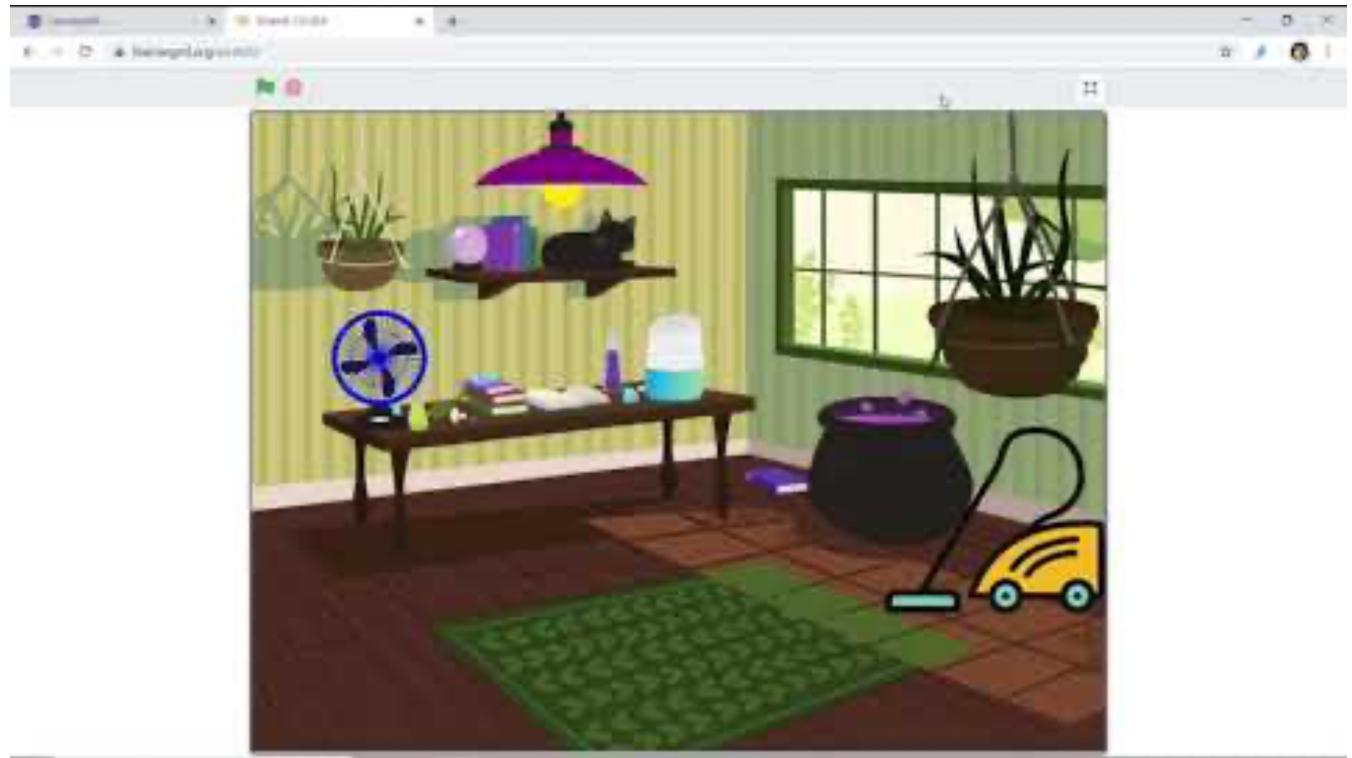
- Smart classroom: En este proyecto crearás un aula inteligente sencilla.
 - ML4K:
 - <https://machinelearningforkids.co.uk/#/worksheets>
 - Hoja de trabajo en español en el campus virtual:
 - Actividad: aula inteligente.
 - LearningML:
 - Plantilla y solución parcial con el ventilador en el aula virtual:
 - Actividad: aula inteligente.





LearningML

- Creación de un asistente virtual
 - Funciona en gallego.
 - <https://web.learningml.org/asistente-virtual/>



Asistente virtual. Youtube (<https://youtu.be/msSgOaQjBMw>)



Teachable Machine

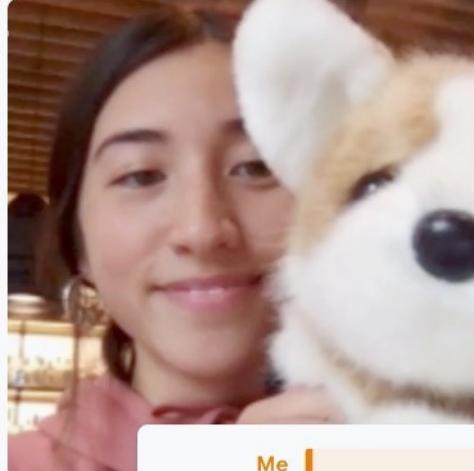
- Una forma rápida y sencilla de crear modelos de aprendizaje automático para tus sitios web, aplicaciones y mucho más, sin necesidad de conocimientos especializados ni de programar.

Teachable Machine

Prepara a un ordenador para que reconozca tus imágenes, sonidos y posturas.

Una forma rápida y sencilla de crear modelos de aprendizaje automático para tus sitios web, aplicaciones y mucho más, sin necesidad de conocimientos especializados ni de programar.

[Primeros pasos](#)



A photograph of a young woman with dark hair smiling, with a white dog partially visible behind her face. Below the photo is a small interface showing two options: "Me" and "Me + Dog <3". The "Me + Dog <3" option is highlighted with a red bar and shows a confidence level of 94%.

<https://teachablemachine.withgoogle.com/>



XUNTA
DE GALICIA

CENTRO DE
FORMACIÓN E
RECURSOS DE FERROL

Inteligencia Artificial para la Sociedad



Alma Mallo

alma.mallo@udc.es

Noviembre de 2024