

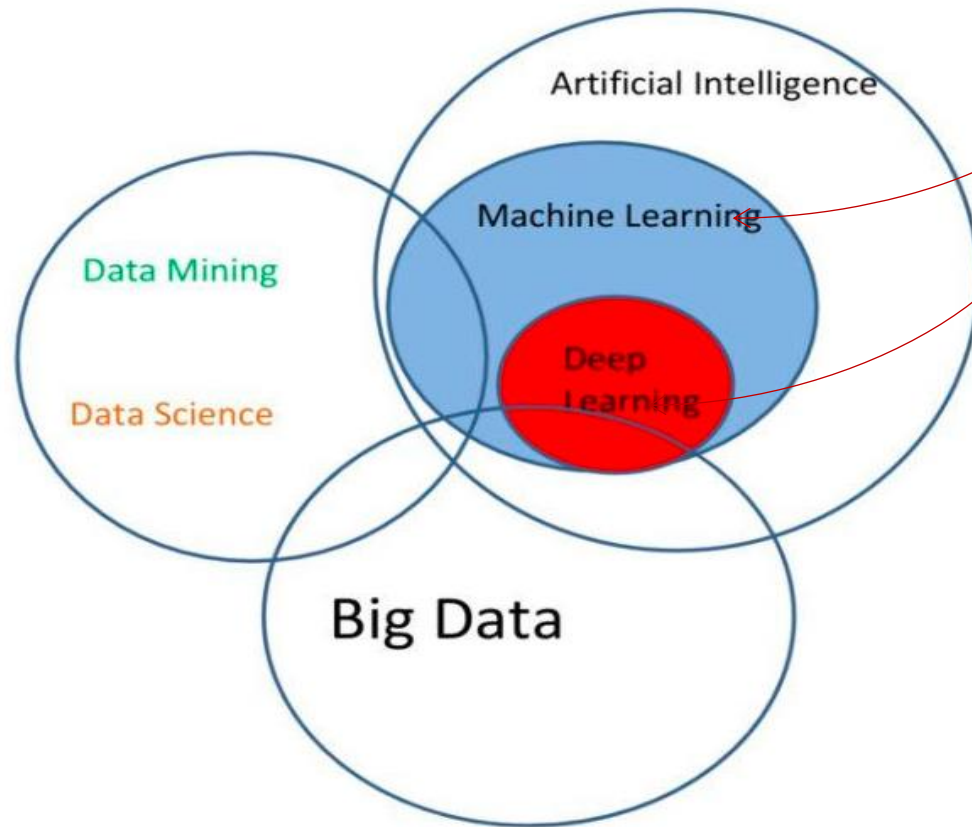
INTRODUCCIÓN AL APRENDIZAJE AUTOMÁTICO

TIPOS DE APRENDIZAJE:
SUPERVISADO - NO SUPERVISADO – POR REFUERZO

LAURA DIAZ DÁVILA – FRANCISCO TAMARIT

EL “QUIÉNES SOMOS” DE LA INTELIGENCIA ARTIFICIAL

HOY



—¿Sabes sumar?—le preguntó la Reina Blanca.— ¿Cuánto es uno más uno
más uno más uno más uno más uno más uno más uno más uno más uno?

—No lo sé—dijo Alicia—. Perdí la cuenta.

—No sabe hacer una adición—le interrumpió la Reina Roja.'

De Lewis Carrol, en Alicia a través del espejo. (1871)

MACHINE LEARNING: LOS SISTEMAS CONEXIONISTAS

REPRESENTACIONES SUBSIMBÓLICAS

- REDES NEURONALES
- ALGORITMOS GENÉTICOS
- REDES BAYESIANAS
- AUTÓMATAS CELULARES



LOS SISTEMAS CONEXIONISTAS

TIPOS DE APRENDIZAJE AUTOMÁTICO



" —¿Qué sabes de este asunto?— preguntó el Rey a Alicia.

—Nada— dijo Alicia.

—¿Absolutamente nada?— insistió el Rey.

—Absolutamente nada— dijo Alicia.

—Esto es importante— dijo el Rey, volviéndose hacia los jurados."

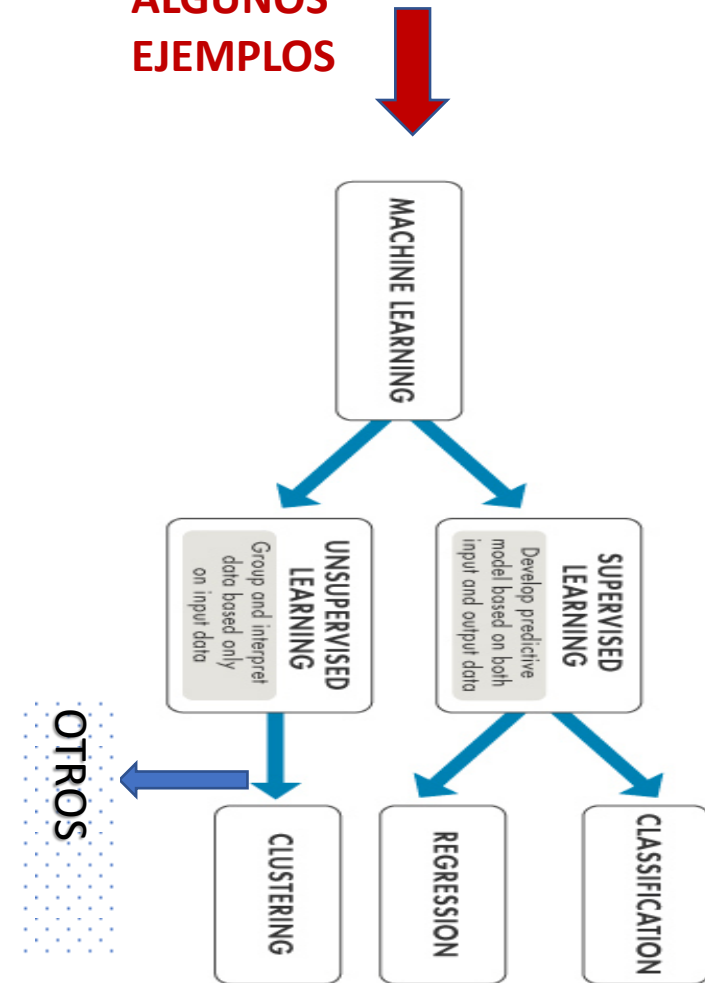
Lewis Carroll, Alicia en el país de la maravillas, en capítulo XII, La Declaración de Alicia (1865)

MACHINE LEARNING: ALGORITMOS INTELIGENTES

Machine learning o Aprendizaje automático es el concepto asociado a las tecnologías inteligentes que aprenden a partir de un entrenamiento basado en datos. Los algoritmos adquieren la capacidad de responder asertivamente frente a una situación nueva. A partir de estas tecnologías es posible descubrir patrones de comportamiento, relacionar, clasificar, construyendo modelos predictivos basados en Inteligencia Artificial



ALGUNOS
EJEMPLOS





ALGORITMOS INTELIGENTES

The Machine Learning Process

- **EJEMPLOS (DATOS)**

- **MODELO**

- **ARQUITECTURA**

- **TIPO DE APRENDIZAJE**

- **REGULARIZACIÓN,
MÉTRICAS**

- **HIPERPÁRAMETROS
PARÁMETROS**



DATA SCIENCE Y MACHINE LEARNING



Datasets

Explore, analyze, and share quality data. [Learn more](#) about data types, creating, and collaborating.

+ New Dataset




Search datasets

Filters

All datasets Computer Science Education Classification Computer Vision NLP Data Visualization Pre-Trained Model

Trending Datasets

See All




National Universities Rankings

The Devastator · Updated 2 days ago

Usability **9.4** · 72 kB

1 File (CSV)

11




Fifa World Cup 2022: Complete Dataset

Iron486 · Updated a day ago

Usability **9.4** · 6 kB

1 File (CSV)

19




Stock Prediction Dataset - ADP

Elon Mask · Updated 11 days ago

Usability **10.0** · 338 kB

1 File (CSV)

15



Data Breaches

The Devastator · Updated 25 days ago

Usability **10.0** · 9 kB

1 File (CSV)

20

FUENTE: <https://www.kaggle.com/datasets>

IMAGENET

14,197,122 images, 21841 synsets indexed

[Home](#) [Download](#) [Challenges](#) [About](#)

Logged in as lauradiazdavila. [My Account](#) | [Logout](#)

An Update to the ImageNet Website and Dataset

March 11, 2021

We are proud to see ImageNet's wide adoption going beyond what was originally envisioned. However, the decade-old website was burdened by growing download requests. To serve the community better, we have redesigned the [website](#) and upgraded its hardware. The new website is simpler; we removed tangential or outdated functions to focus on the core use case—enabling users to [download the data](#), including the full ImageNet dataset and the [ImageNet Large Scale Visual Recognition Challenge \(ILSVRC\)](#).

Meanwhile, the computer vision community has progressed, and so has ImageNet. The dataset was created to benchmark object recognition—at a time when it barely worked. The problem then was how to collect labeled images at a sufficiently large scale to be able to train complex models in laboratories. Today, computer vision is in real-world systems impacting people's Internet experience and daily lives. An emerging problem now is how to make sure computer vision is fair and preserves people's privacy. We are continually evolving ImageNet to address these emerging needs.

In a [FAT* 2020 paper](#), we filtered 2,702 synsets in the "person" subtree that may cause problematic behaviors of the model. We have updated the full ImageNet data on the website to remove these synsets. The update does not affect the 1,000 categories in ILSVRC.

In a [more recent paper](#), we investigate privacy issues in ILSVRC. 997 out of 1000 categories in ILSVRC are not people categories; nevertheless, many incidental people are in the images, whose privacy is a concern. We first annotated faces in the images and then constructed a face-blurred version of ILSVRC. Experiments show that one can use the face-blurred version for benchmarking object recognition and for transfer learning with only marginal loss of accuracy. We release our [face annotations](#) to facilitate further research on privacy-aware visual recognition.

Team members working on these new improvements: [Kaiyu Yang](#) (Princeton), [Jacqueline Yau](#) (Stanford), [Li Fei-Fei](#) (Stanford), [Jia Deng](#) (Princeton), [Olga Russakovsky](#) (Princeton).

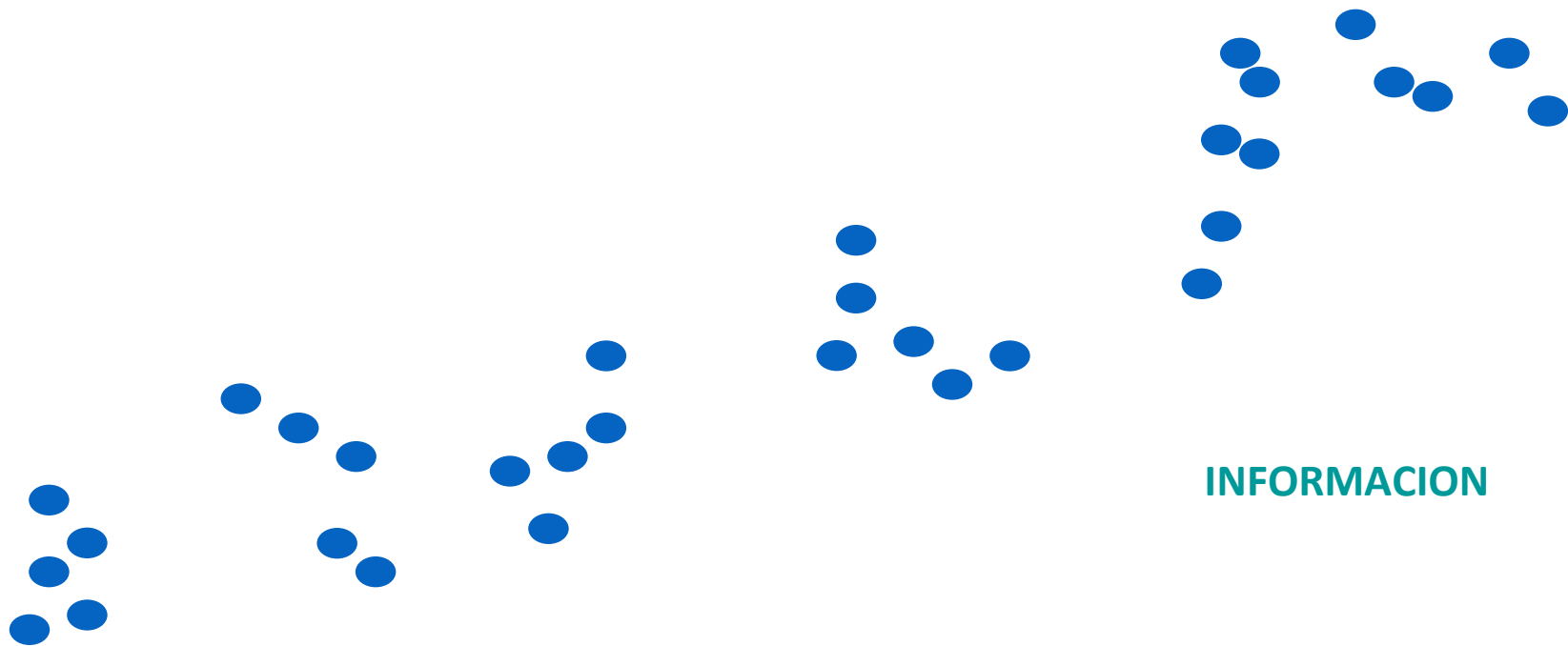
Fuente: <https://www.image-net.org/update-mar-11-2>

¿Qué es un Patrón de Conocimiento?

EL APRENDIZAJE DE LOS SISTEMAS INTELIGENTES

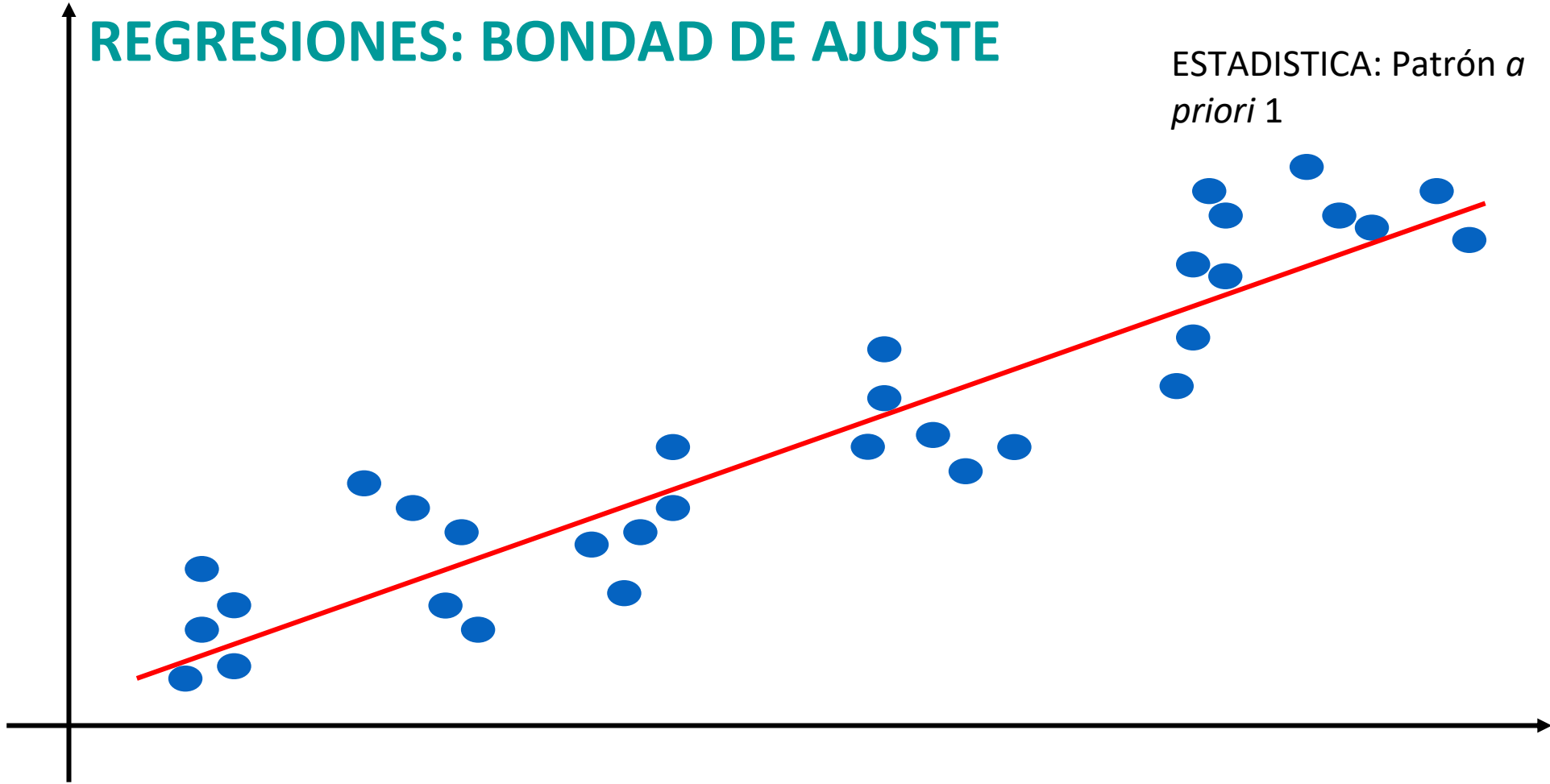
BASADO EN DATOS:

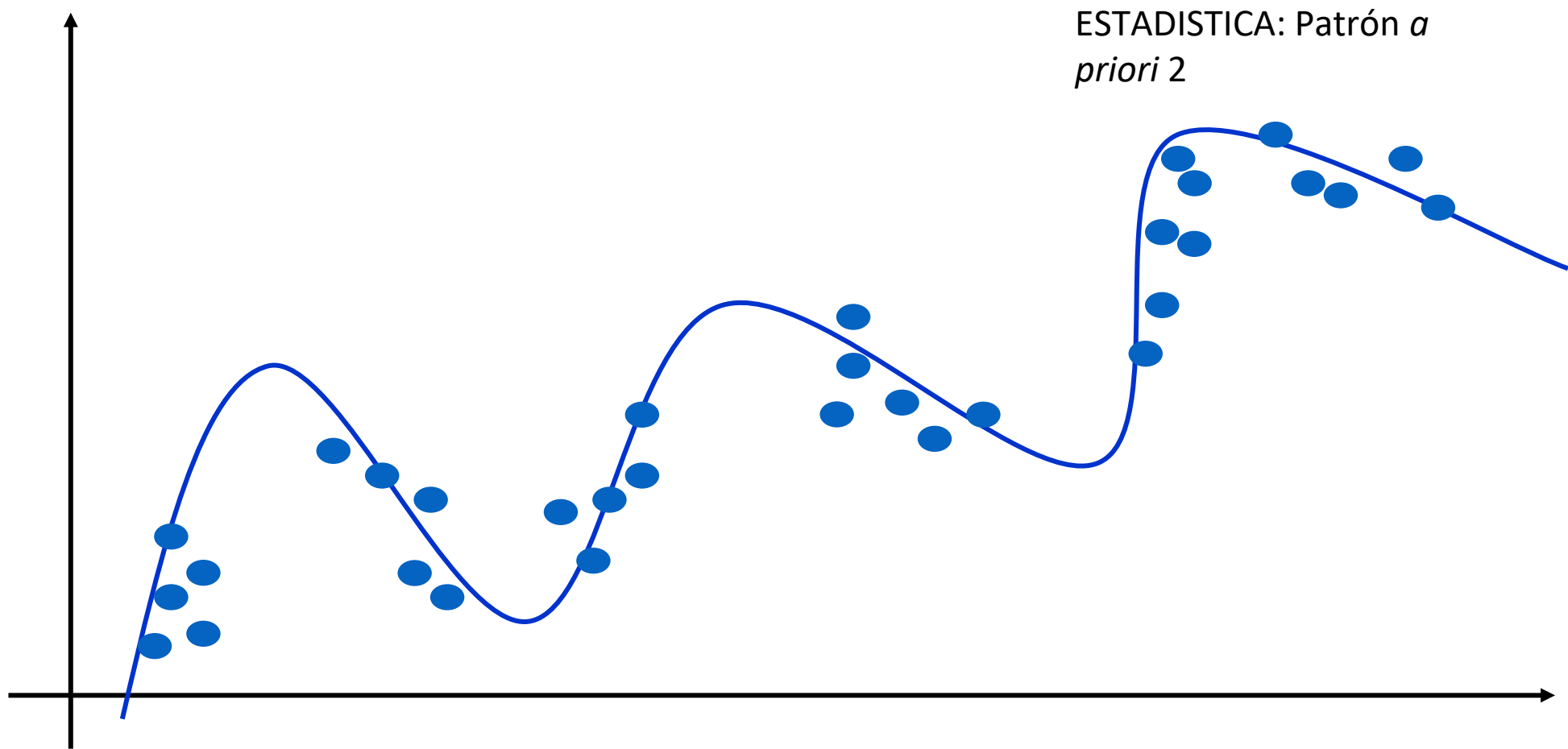
MACHINE LEARNING



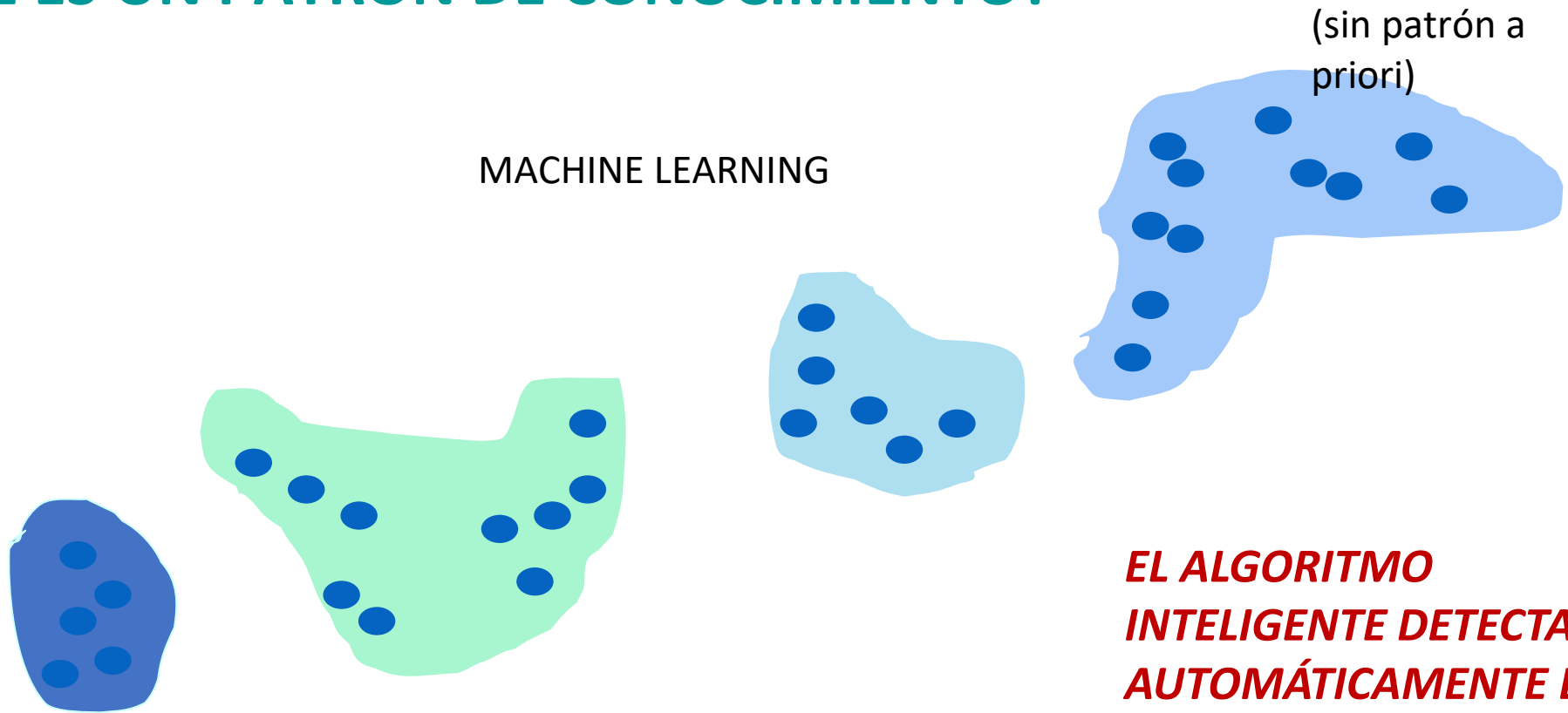
REGRESIONES: BONDAD DE AJUSTE

ESTADISTICA: Patrón α
priori 1





¿QUÉ ES UN PATRON DE CONOCIMIENTO?

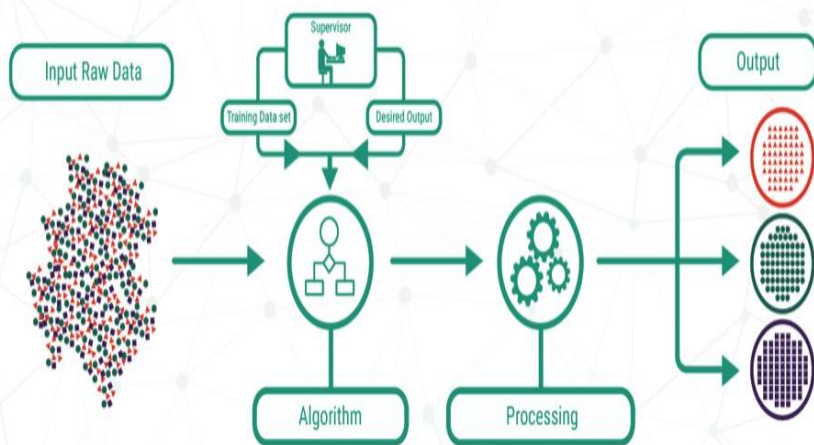


***EL ALGORITMO
INTELIGENTE DETECTA
AUTOMÁTICAMENTE LAS
SIMILITUDES ENTRE LOS
DATOS***

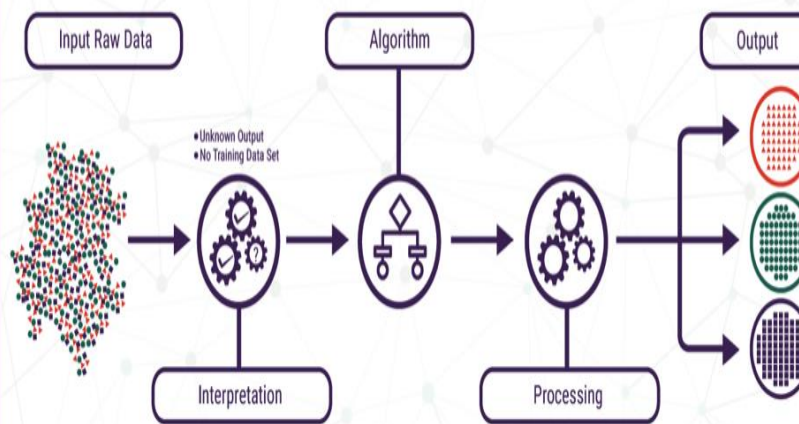


Tipos de aprendizaje automatizado

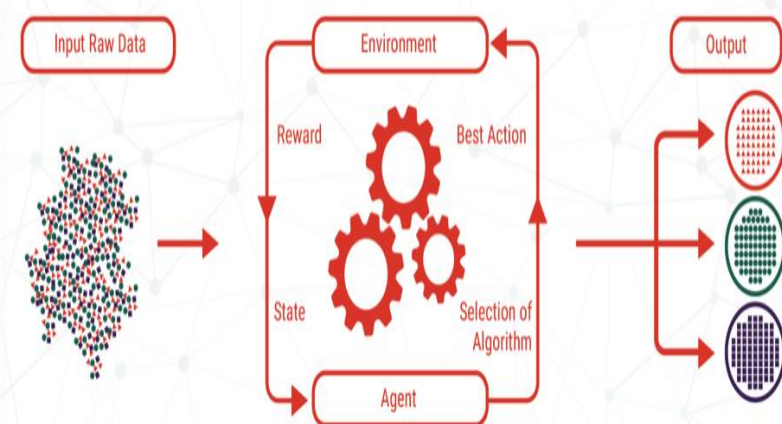
SUPERVISED LEARNING



UNSUPERVISED LEARNING

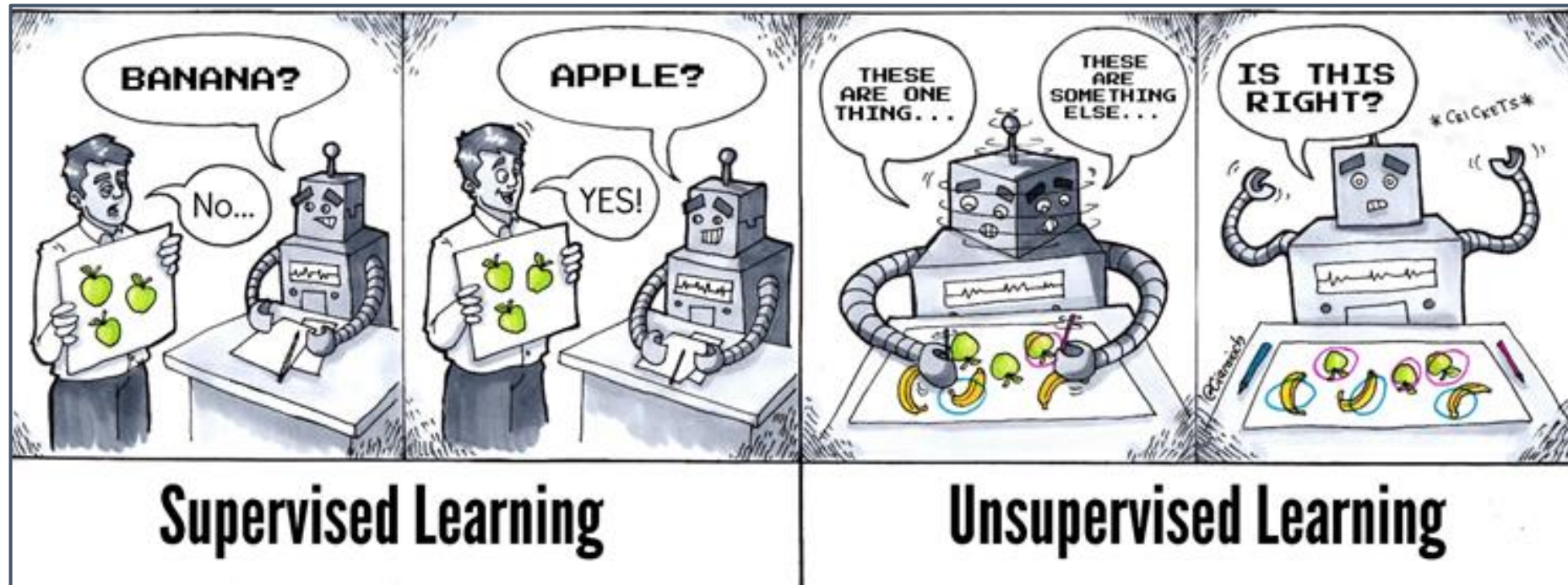


REINFORCEMENT LEARNING



Fuente: bigdata-madesimple.com

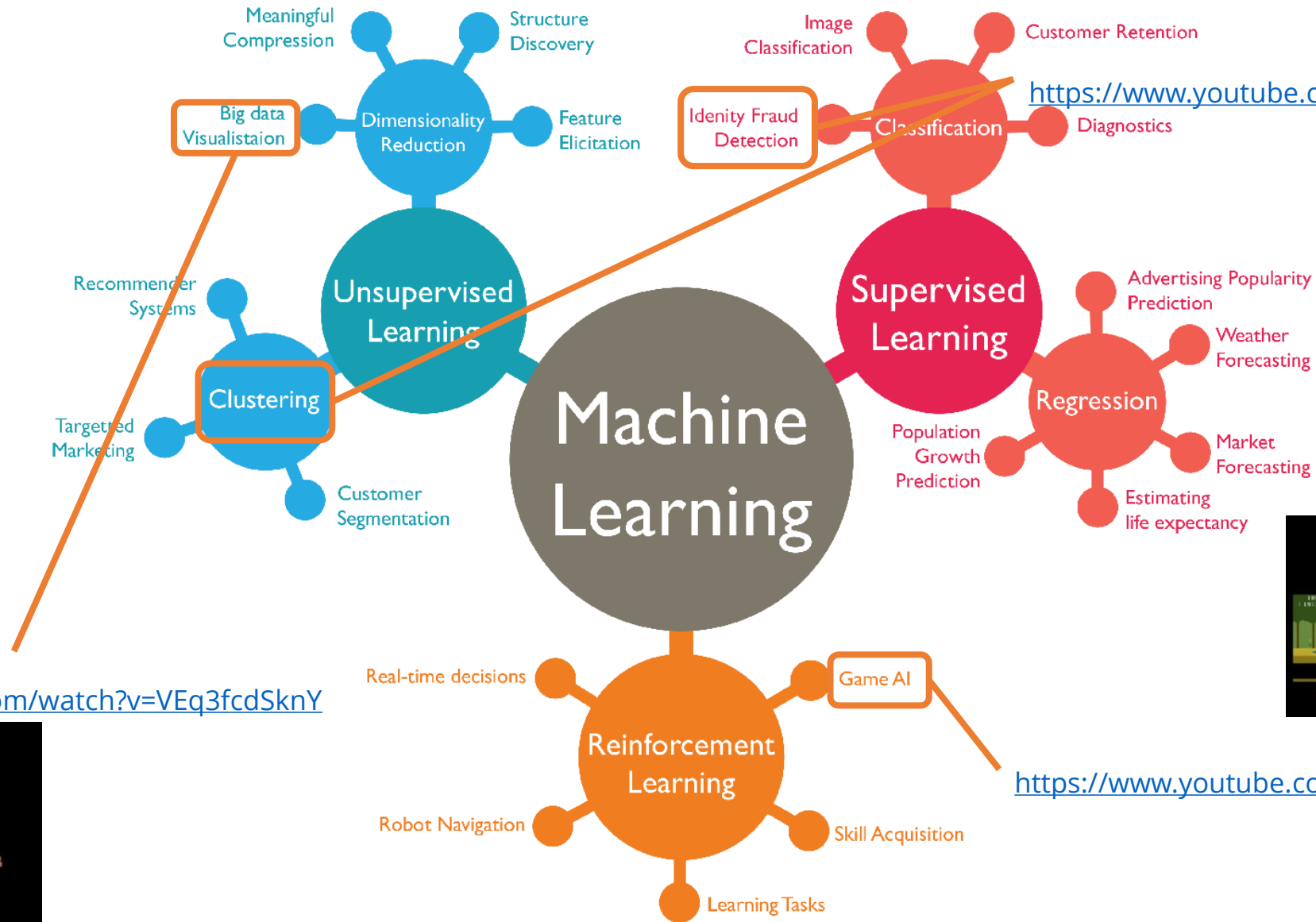
Aprendizaje supervisado vs. no supervisado



Fuente: Ciara Íoch

Aprendizaje por refuerzo

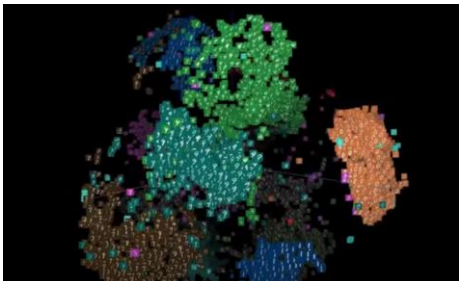




<https://www.youtube.com/watch?v=VEq3fcdSknY>



<https://www.youtube.com/watch?v=VEq3fcdSknY>



<https://www.youtube.com/watch?v=VEq3fcdSknY>

DISEÑANDO para INNOVAR:

- 1- SIMBÓLICO – SUBSIMBÓLICO
- 2- TIPOS DE APRENDIZAJE AUTOMÁTICO
- 3- CAJAS NEGRAS – CAJAS BLANCAS. EXPLICABILIDAD
- 4- DATOS: ENTRENAMIENTO VALIDACION
- 5- MÉTRICAS DE DESEMPEÑO
- 6- AJUSTE – GENERALIZACIÓN – DESPLIEGUE
- 7- HIPERPARÁMETROS – PARÁMETROS
- 8- SESGO Y DISCRIMINACIÓN
- 9- ÉTICA
- 10- IA FIABLE
- 11- IA PARA EL BIEN COMÚN



CÓMO INTERACTUAMOS CON “La Tecnología Inteligente” CUANDO:

- LA ENTRENAMOS
- VALIDAMOS SU DESEMPEÑO
- COMIENZA A GENERALIZAR, SE PREPARA PARA SU DESPLIEGUE

ASÍ ENTRENAMOS A LOS ALGORITMOS INTELIGENTES, ASÍ APRENDEN A DAR RESPUESTA EN SITUACIONES NUEVAS

VALIDACIÓN

HIPERPARÁMETROS

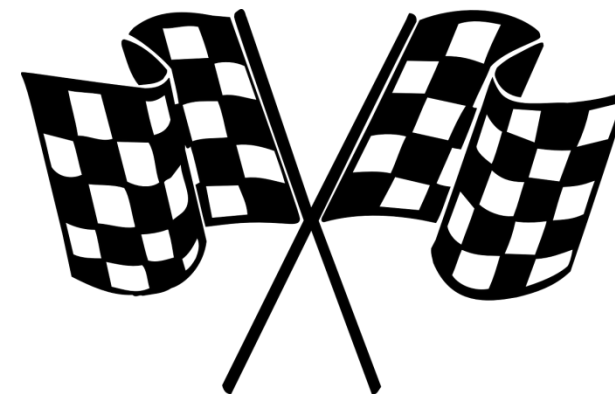
TÉCNICAS DE REGULARIZACIÓN



**EJEMPLOS NUEVOS
(DATOS QUE AÚN NO HA
VISTO)**



PARÁMETROS



MÉTRICAS

AHORA BIEN,

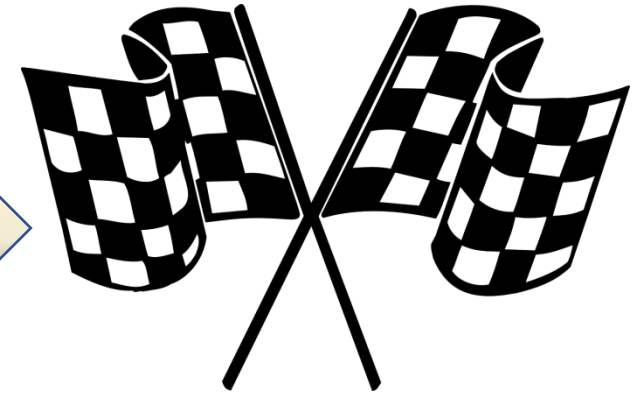
¿QUÉ LE DIMOS? Y ¿QUÉ NOS ENTREGÓ?

ENTRENAMIENTO

VALIDACIÓN

TÉCNICAS DE REGULARIZACIÓN
HIPERPARÁMETROS

- **EJEMPLOS (DATOS)**
- **EJEMPLOS NUEVOS (DATOS QUE AÚN NO HA VISTO)**



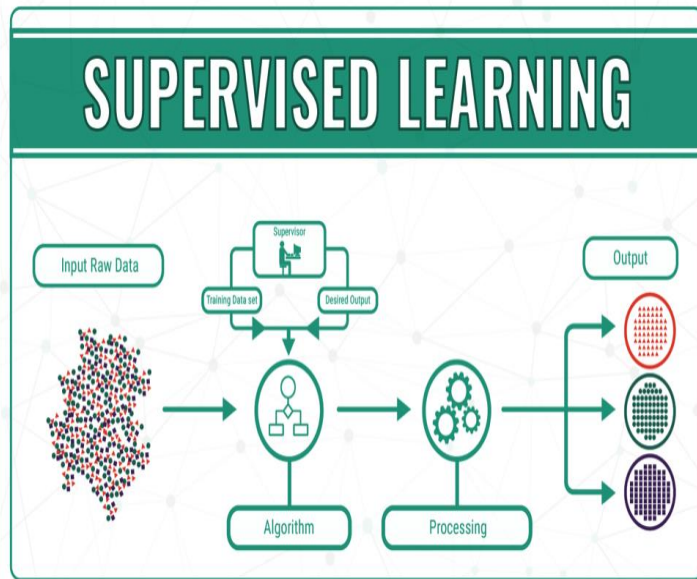
PARÁMETROS

MÉTRICAS

¡AHORA ES CAPAZ DE GENERALIZAR!



MÉTRICAS

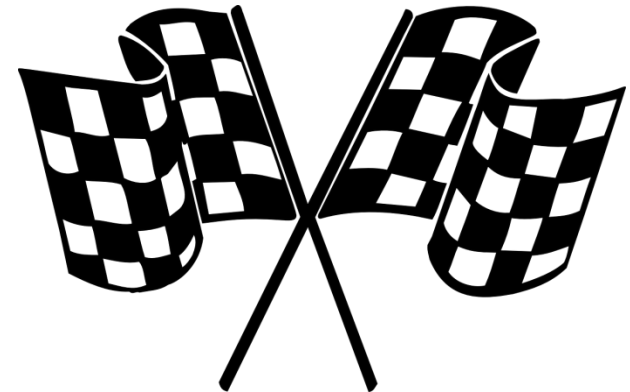


- **LOS DATOS QUE LES PROPORCIONAMOS TIENE UN OBJETIVO, TARGET, CATEGORÍA O ETIQUETA:**

- **FOTO DE PATO CON ETIQUETA “PATO”**
- **DATOS DE ESTUDIANTES CON CATEGORÍA “RENDIMIENTO ACADÉMICO”**



- **EJEMPLOS (DATOS)**



EL DESEMPEÑO DEL MODELO ENTRENADO

LAS MÉTRICAS Y LA ESTADÍSTICA

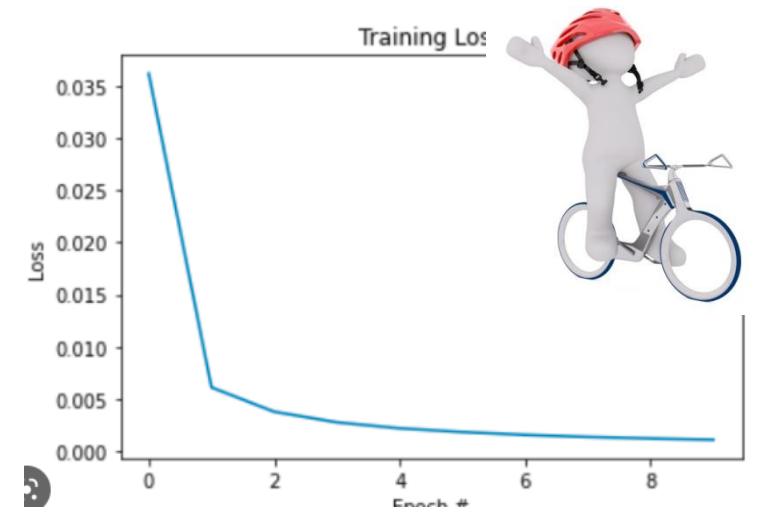
Confusion matrix			
	'republican'	'democrat'	Sum
'republican'	166	2	168
'democrat'	1	266	267
Sum	167	268	435



ACCURACY

		Actual Values	
		Positive (1)	Negative (0)
Predicted Values	Positive (1)	TP	FP
	Negative (0)	FN	TN

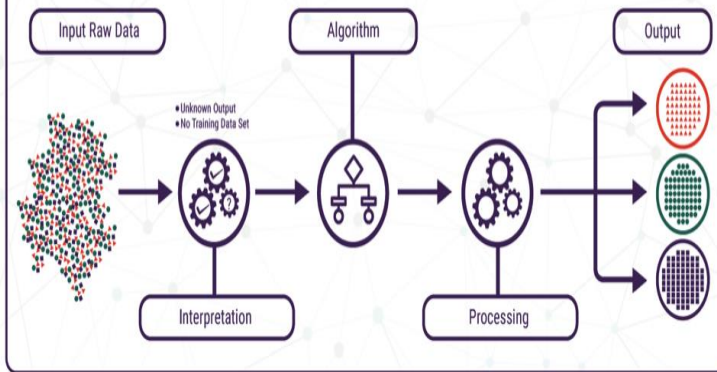
Confusion matrix					
	c_som_1_1	c_som_1_2	c_som_2_1	c_som_2_2	Sum
c_som_1_1	61	0	3	0	64
c_som_1_2	4	122	0	2	128
c_som_2_1	5	0	190	2	197
c_som_2_2	1	5	2	38	46
Sum	71	127	195	42	435



LOSS

MATRIZ DE
CONFUSIÓN

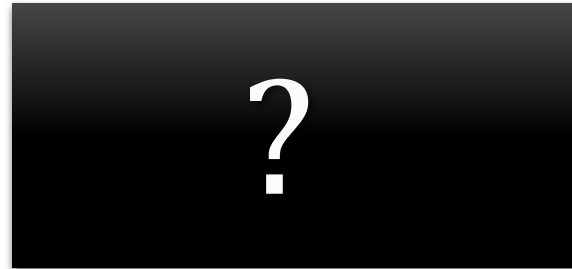
UNSUPERVISED LEARNING



- **LOS DATOS QUE LES PROPORCIONAMOS NO TIENEN UN OBJETIVO.**
- **SALEN A BUSCAR SIMILITUDES QUE ENCUENTRAN EN LOS DATOS!**
- **SON SIMILITUDES SUBSIMBÓLICAS!**
- **LOS AGRUPAN SEGÚN SUS HALLAZGOS**



- **EJEMPLOS (DATOS)**



**SE USAN OTRAS
MÉTRICAS**

¡Gracias!