

ARTIFICIAL INTELLIGENCE ML

JOHNATAN GARZÓN M.

BEFORE ALL

- I'm not an expert



- Be free to ask



- Secure space



ARTIFICIAL INTELLIGENCE

- AI Bases:
- 250 a.C Ctesibio de Alejandria
 - First self-controlled machine
- 1843 Ada Lovelace
 - First Algorithm
- 1936 Alan Turing
 - First Calculator
- 1959 Marvin Minsky, J. McCarthy
 - They founded A.I Laboratory in MIT



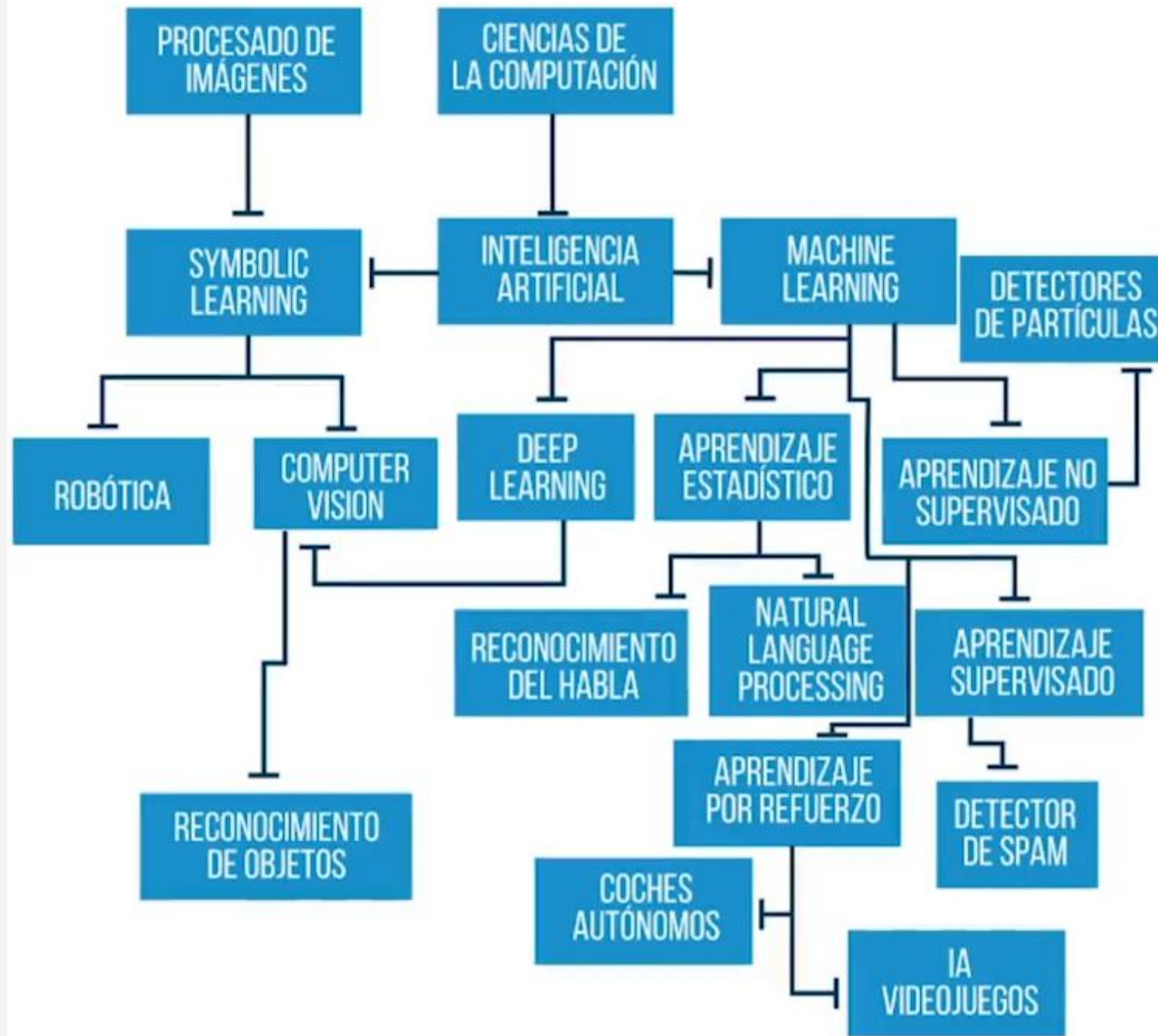


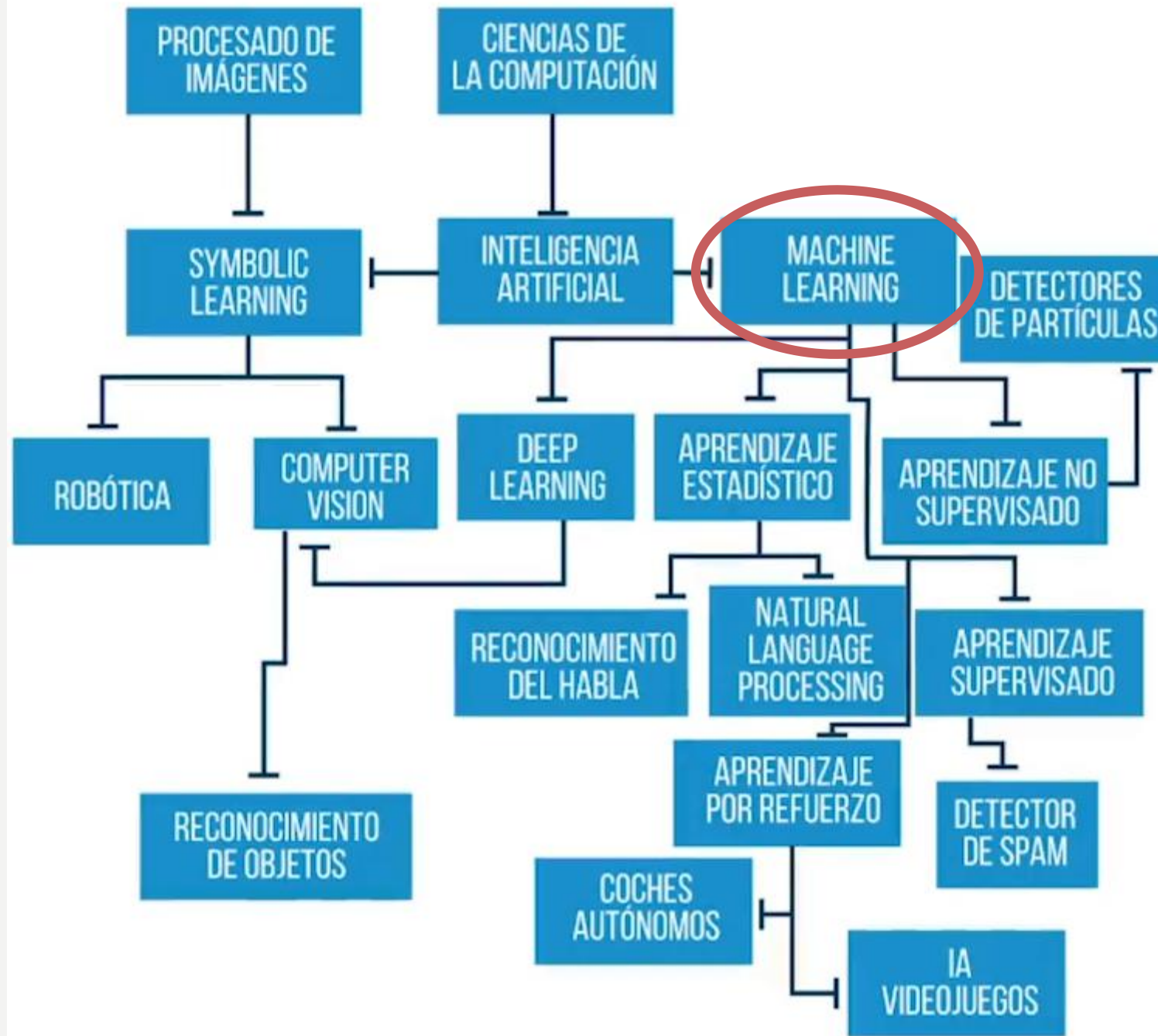
HUMAN BRAIN VS. AI

MAIN IDEA: DEVELOPMENT OF METHODS AND ALGORITHMS
THAT ALLOW COMPUTERS TO BEHAVE INTELLIGENTLY

WHICH COMPANIES USE AI?





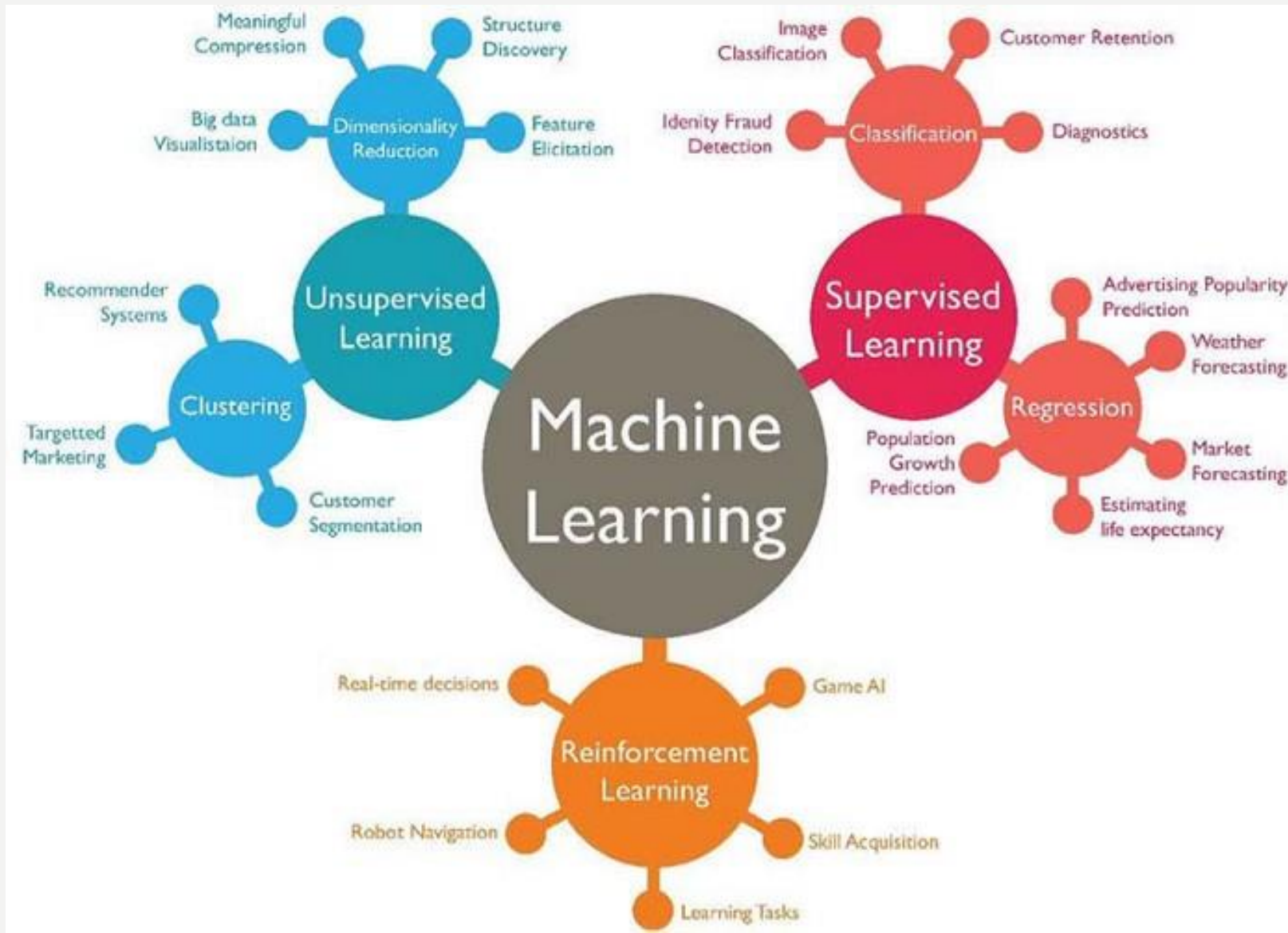


WHAT IS MACHINE LEARNING



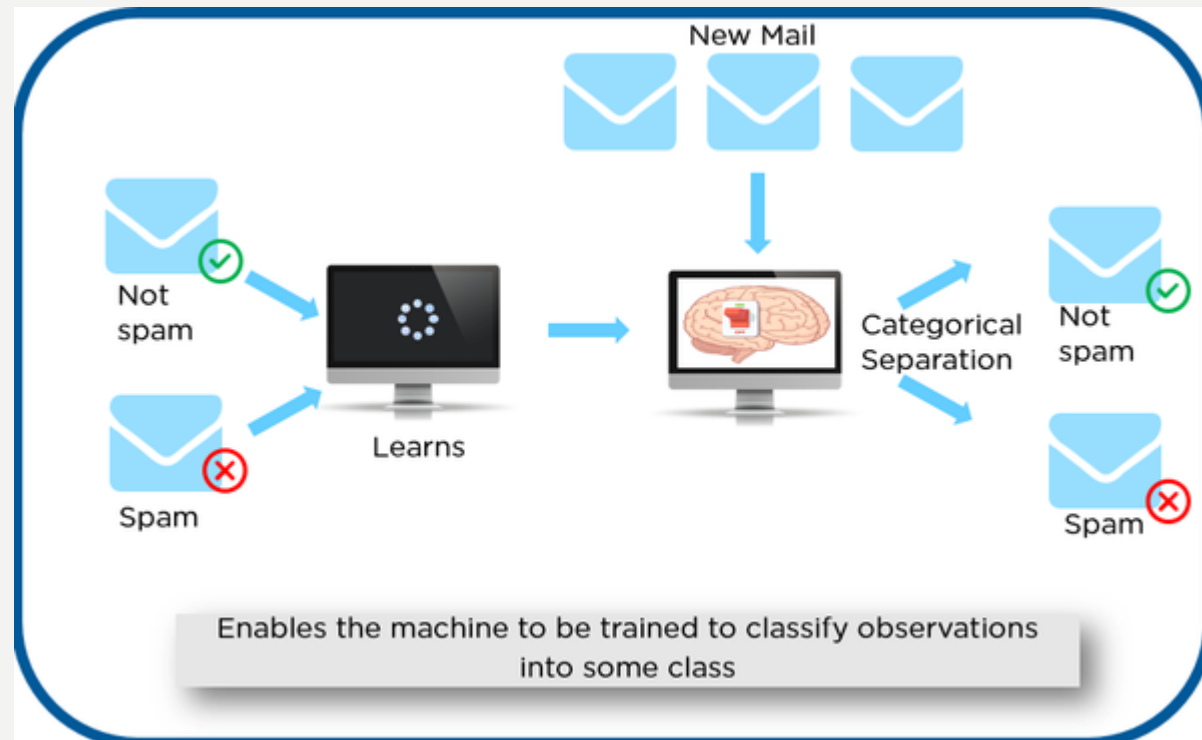
- According to Arthur Samuel, Machine Learning algorithms enable the computers to learn from data, and even improve themselves, without being explicitly programmed.

TYPES OF MACHINE LEARNING



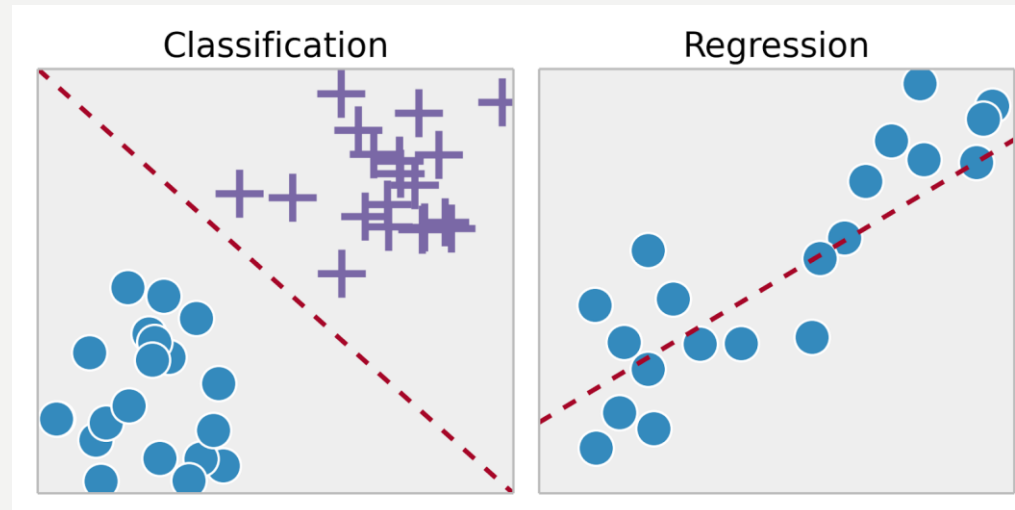
OVERVIEW OF SUPERVISED LEARNING ALGORITHM

- In Supervised learning, an AI system is presented with data which is labeled, which means that each data tagged with the correct label.



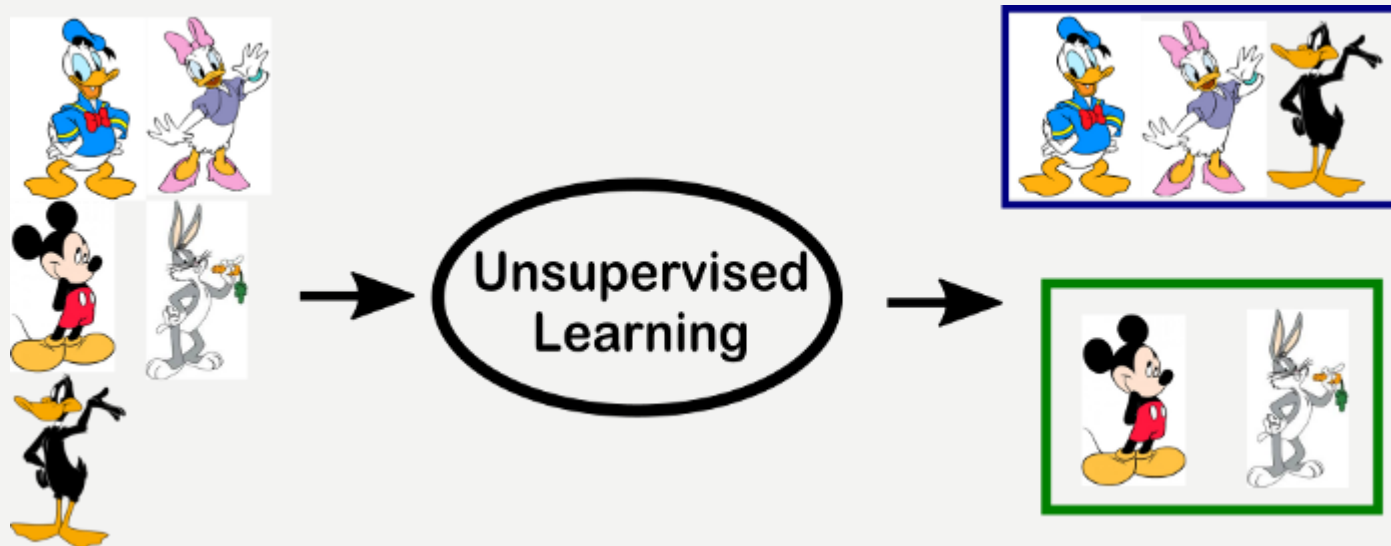
TYPES OF SUPERVISED LEARNING

- **Classification:** A classification problem is when the output variable is a category, such as “red” or “blue” or “disease” and “no disease”.
- **Regression:** A regression problem is when the output variable is a real value, such as “dollars” or “weight”.



OVERVIEW OF UNSUPERVISED LEARNING ALGORITHM

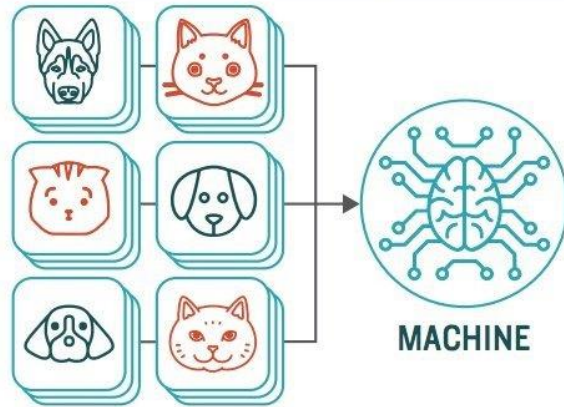
- In unsupervised learning, an AI system is presented with unlabeled, uncategorized data and the system's algorithms act on the data without prior training. The output is dependent upon the coded algorithms.



How **Unsupervised** Machine Learning Works

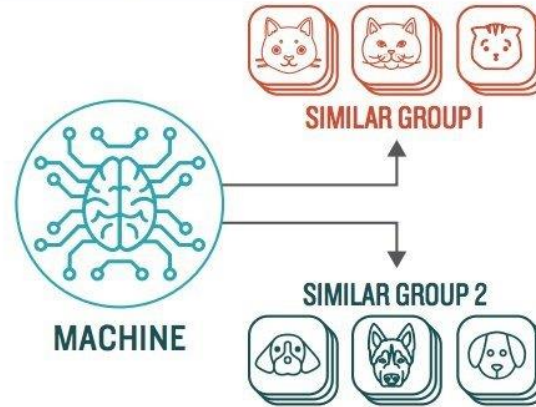
STEP 1

Provide the machine learning algorithm uncategorized, unlabeled input data to see what patterns it finds



STEP 2

Observe and learn from the patterns the machine identifies

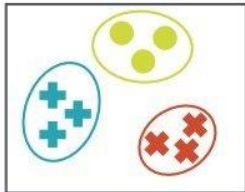


TYPES OF PROBLEMS TO WHICH IT'S SUITED

CLUSTERING

Identifying similarities in groups

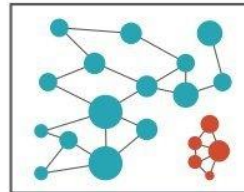
For Example: Are there patterns in the data to indicate certain patients will respond better to this treatment than others?



ANOMALY DETECTION

Identifying abnormalities in data

For Example: Is a hacker intruding in our network?

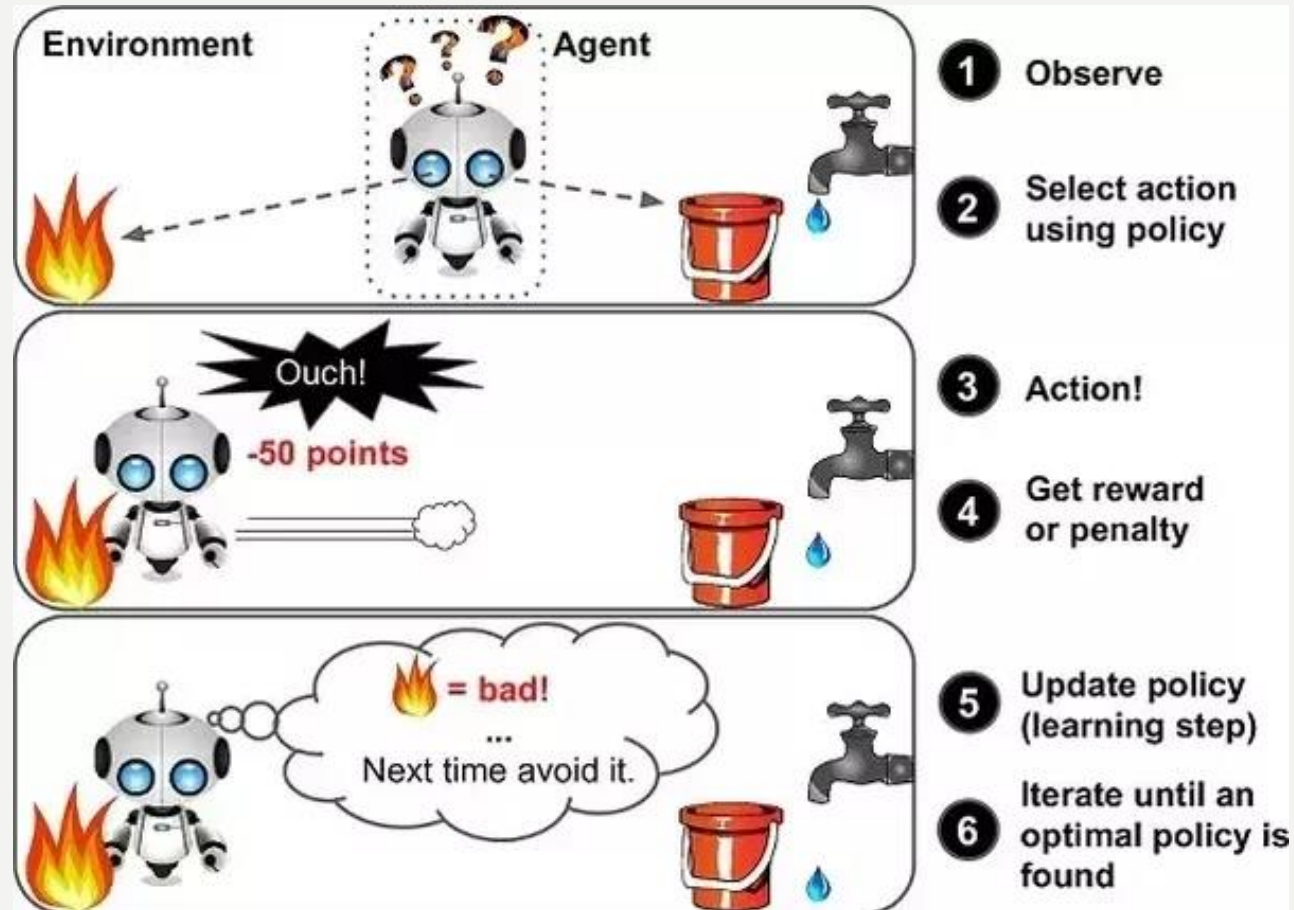


TYPES OF UNSUPERVISED LEARNING

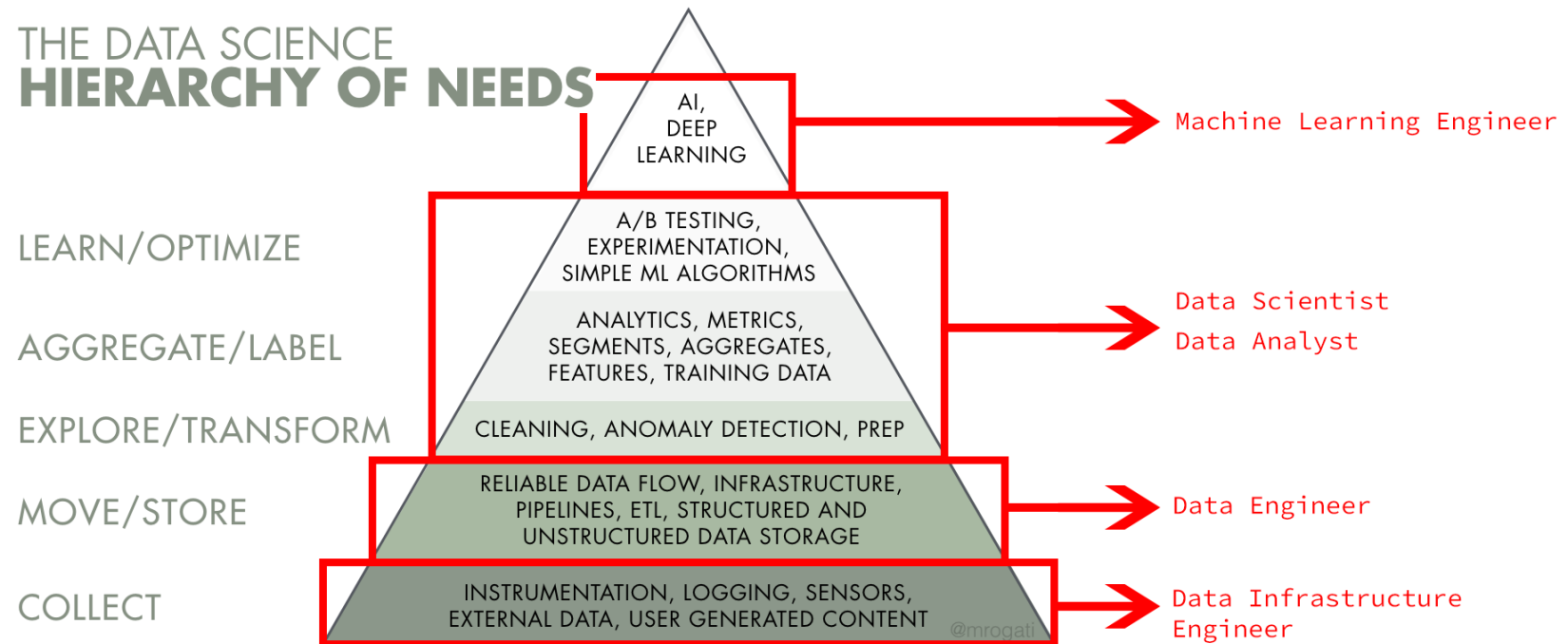
Clustering: A clustering problem is where you want to discover the inherent groupings in the data, such as grouping customers by purchasing behavior.

Association: An association rule learning problem is where you want to discover rules that describe large portions of your data, such as people that buy X also tend to buy Y.

OVERVIEW OF REINFORCEMENT LEARNING



THE DATA SCIENCE HIERARCHY OF NEEDS



IMPORTANCE OF DEFINING ML PROBLEM

- Not just programming!
- Common problems:
 - No problems to solve
 - There are easiest solutions
 - You don't know if someone else already resolved that problem
 - You need to define a real problem, in clean and precise way, to know if you need to apply ML or not

KEY QUESTIONS

- Para saber que tipo de aprendizaje necesitas:
 - Para quien piensas generar valor y como
 - Que tipo de funcionalidad queremos que cumpla esta solución de ML:
 - Predecir un valor
 - Predecir una etiqueta
 - Agrupar elementos similares
 - Optimizar un proceso con prueba y error

OUR FOCUS

- Nos enfocaremos en aprendizaje supervisado ya que es el que está teniendo más impacto en la industria (Andrew Ng).



KEY QUESTIONS

- Key questions for supervised learning:
 1. ¿De qué tipo es el valor que se quiere predecir?
 - a) Continuo
 - b) Discreto
 2. ¿Cuál es la definición de éxito en una predicción?
 3. ¿Con qué datos se contaría para hacer esa predicción?
 4. ¿La pregunta que se está tratando de resolver pertenece a alguna disciplina en particular?
 5. Considerando nuestra intuición en la disciplina ¿Los datos nos permiten predecir el objetivo?

WORKSHOP

Contexto: Somos un ente gubernamental que quiere definir sus políticas de financiamiento de producciones cinematográficas nacionales.

Reconoce el tipo de aprendizaje que necesitas

1. **Ayudar** a la producción de películas de calidad **que no logran ser autosustentables**.
2. Nos sería útil saber que películas tienen más dificultad para recuperar en sus presupuestos. Por consiguiente queremos **predecir una métrica**: el ingreso mundial generado por una película.

Aterriza tu problema de aprendizaje supervisado

1. Los ingresos de una película corresponden a **valores continuos**.
2. Mi éxito será "**qué tan cerca estoy del valor real de ingreso generado por la película**".
3. Me basaré en **bases de datos públicas de internet**
4. El dominio de trabajo es la **industria del cine**, en particular de la distribución de películas.
5. Sí, de forma general existen bastantes características que me pueden ayudar a saber que película será exitosa como: **calidad, actores, presupuesto, etc...**

- *Quiero predecir ingresos de películas, para tomar mejores decisiones de financiamiento, con base a una regresión sobre datos de películas extraídos de internet.*
- *Mi evaluación del éxito será la precisión de mis predicciones.*
- *Podré apoyarme en conocimientos específicos de la industria.*

PROGRAMMING TIME!

[HTTPS://GITHUB.COM
/GARZUZO/AI-TALK](https://github.com/GARZUZO/AI-TALK)

Google Colab

Google Colab es muy similar en su uso a Google Docs. Puedes ir directamente a su pagina <https://colab.research.google.com/> o abrir Google Drive, clicar en el botón “+ Nuevo” y desde ahí elegir la opción “Mas” y clicar en Colaboratory (como se ve en la imagen)

