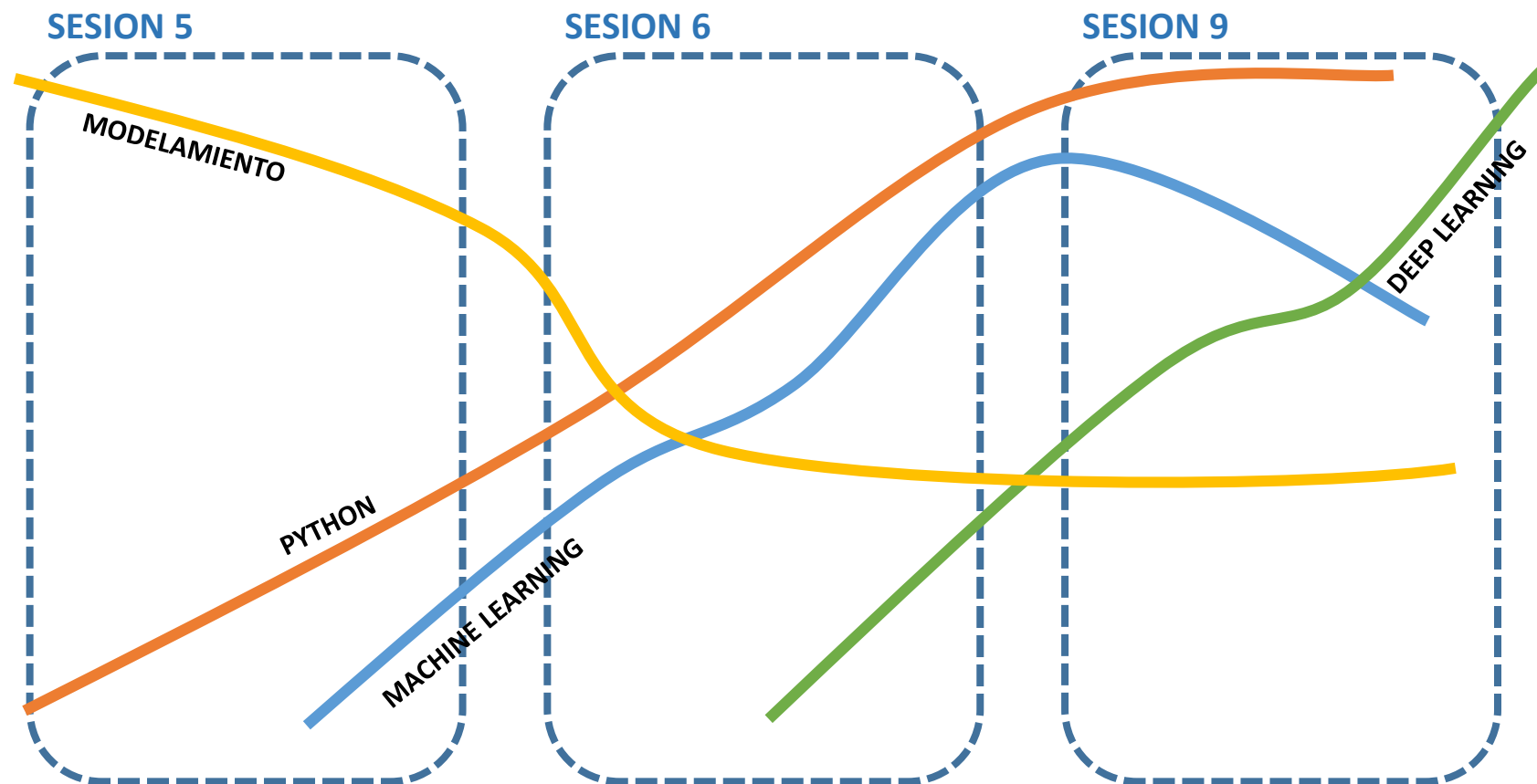


Ritmo de Temas en las Próximas Sesiones



Si era muy difícil entender “complejidad” en estas sesiones se va a poder entender muy bien.



SINCRONA

Machine Learning para el Modelamiento y Gestión de Sistemas Complejos

Sesión 5:

Modelamiento, Simulación

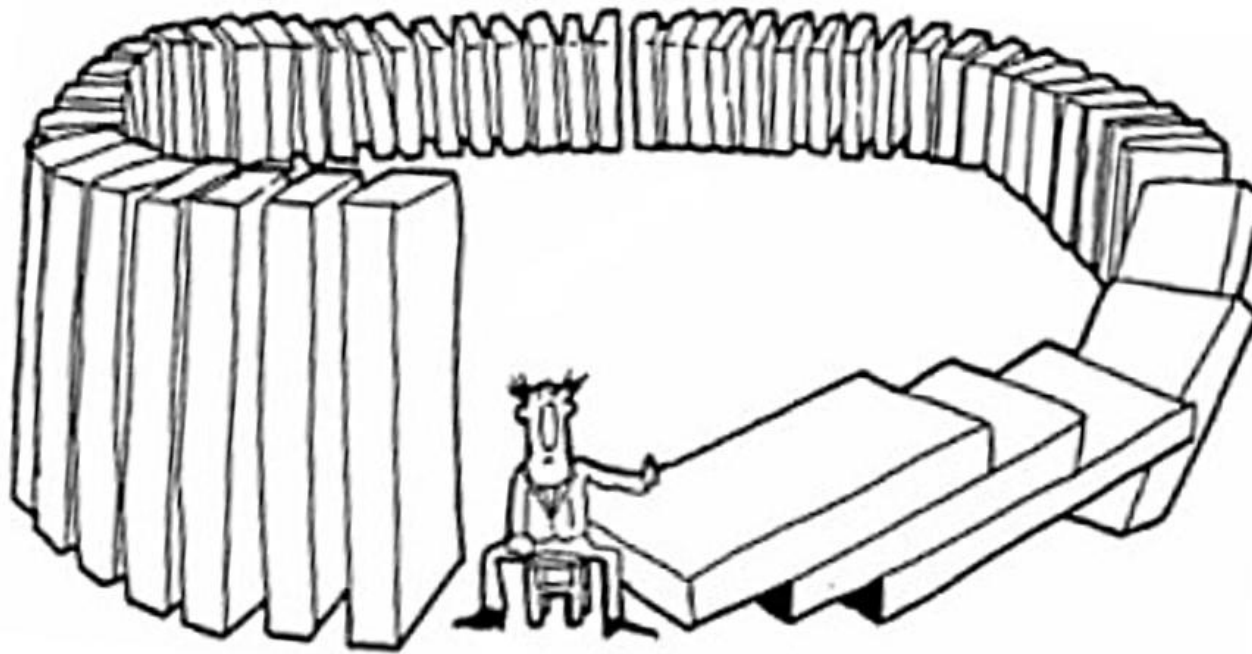
José Carlos Machicao



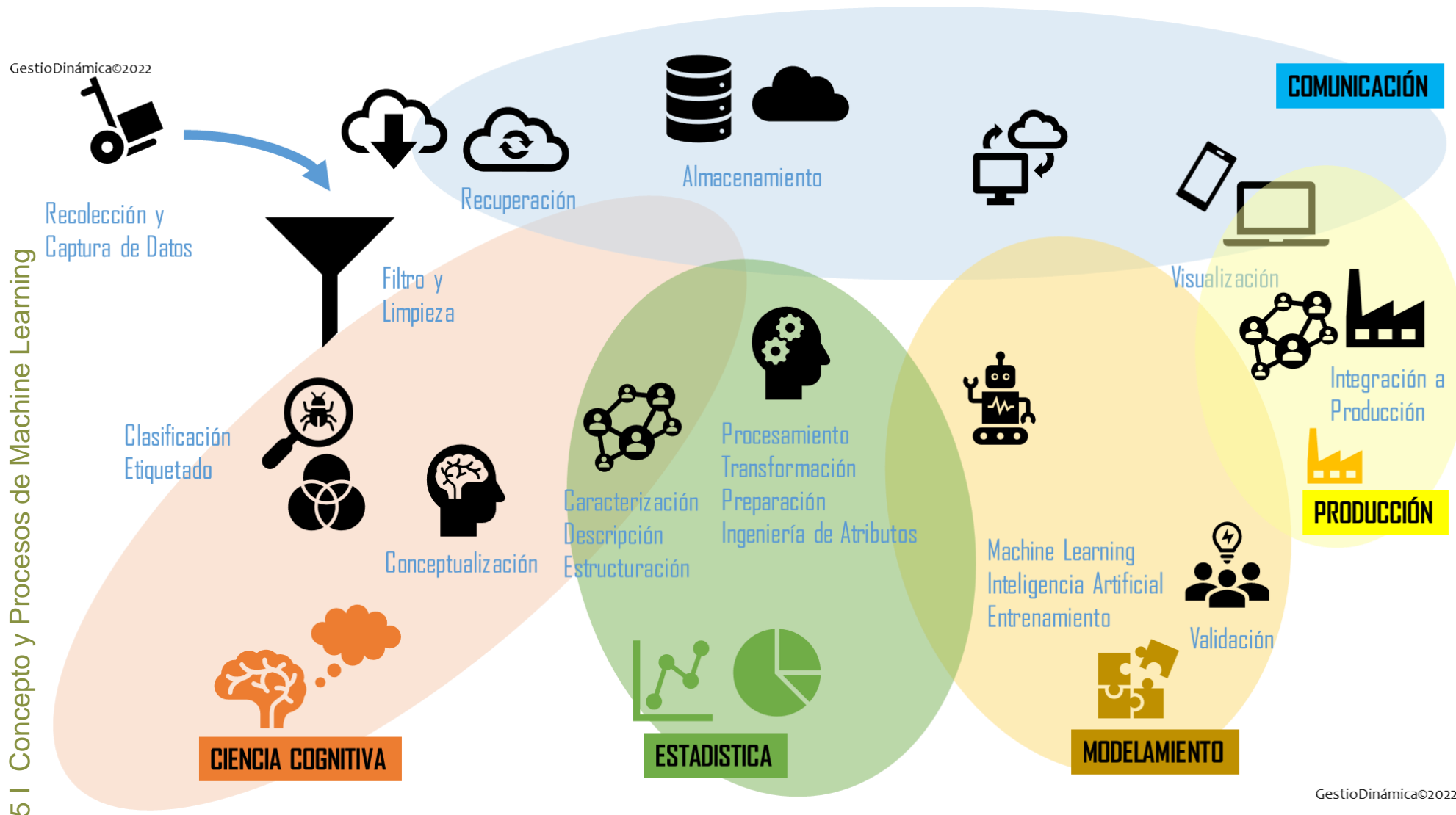
S5

Modelamiento: Concepto







Contra-intuición en el funcionamiento de los sistemas



Ciclo de vida de los datos como contexto del modelamiento



TIPOS DE ANÁLISIS EN SISTEMAS

						
	Análisis Exploratorio	Análisis Decriptivo	Análisis Diagnóstico	Análisis Inferencial	Análisis Prescriptivo	Análisis Predictivo
	Se aplica cuando hay un conocimiento bajo o intermedio pero se quiere ampliar sin mucha idea de los principios de búsqueda.	Se aplica cuando hay mucho desconocimiento o cuando hay muchas interpretaciones acerca del sistema.	Se quiere comprender fundamentalmente el comportamiento y su grado de idoneidad comparado con algún modelo de cómo debería comportarse.	Se tiene un conocimiento elevado sobre una parte de los factores. Se pretende deducir algunos de los factores a partir de los ya conocidos.	Se tiene un alto conocimiento de cómo se comporta el sistema y cómo debería comportarse el sistema. Se modela el salto que debe dar para pasar de uno a otro estado.	Se trata de tener los criterios para anticipar el futuro del sistema bajo condiciones diversas.
Identidad del sistema	No se conoce	Conocimiento parcial	Conocimiento parcial	Se conoce mejor	Se conoce mejor	Se conoce mejor
Horizonte de cambio del sistema	No se conoce	No se conoce	No se conoce	Conocimiento parcial	Se conoce mejor	Se conoce mejor
Horizonte de Mejora del Sistema	No se conoce	No se conoce	Conocimiento parcial	Conocimiento parcial	Se conoce mejor	Se conoce mejor
Comportamiento real del sistema	Conocimiento parcial	Conocimiento parcial	Conocimiento parcial	Conocimiento parcial	Conocimiento parcial	Conocimiento parcial
Comportamiento idóneo del sistema	Conocimiento parcial	No se conoce	Conocimiento parcial	Conocimiento parcial	Se conoce mejor	Se conoce mejor
Entorno del Sistema	No se conoce	Conocimiento parcial	Conocimiento parcial	Conocimiento parcial	Conocimiento parcial	Se conoce mejor
Elementos primarios del sistema	No se conoce	Conocimiento parcial	Conocimiento parcial	Conocimiento parcial	Conocimiento parcial	Conocimiento parcial
Elementos secundarios del sistema	No se conoce	Conocimiento parcial	No se conoce	No se conoce	Conocimiento parcial	Conocimiento parcial
Defectos del sistema	No se conoce	No se conoce	Conocimiento parcial	Conocimiento parcial	Se conoce mejor	Se conoce mejor
Caracterización de datos requeridos	No se puede tener mucha exigencia porque no hay datos. Recién se empieza a explorar.	Se requiere tener datos estructurados acerca de los elementos.	Se requiere tener datos del comportamiento del sistema. Los datos podrían ser dinámicos.	Se requiere datos y modelos definidos.	Se requiere datos abundantes acerca del sistema idóneo, elementos y comportamiento.	Se requiere datos sobre el comportamiento real o actual del sistema. Se requieren modelos o algoritmos generadores de modelos. Se requiere datos de prueba y validación.

GestioDinámica © 2021

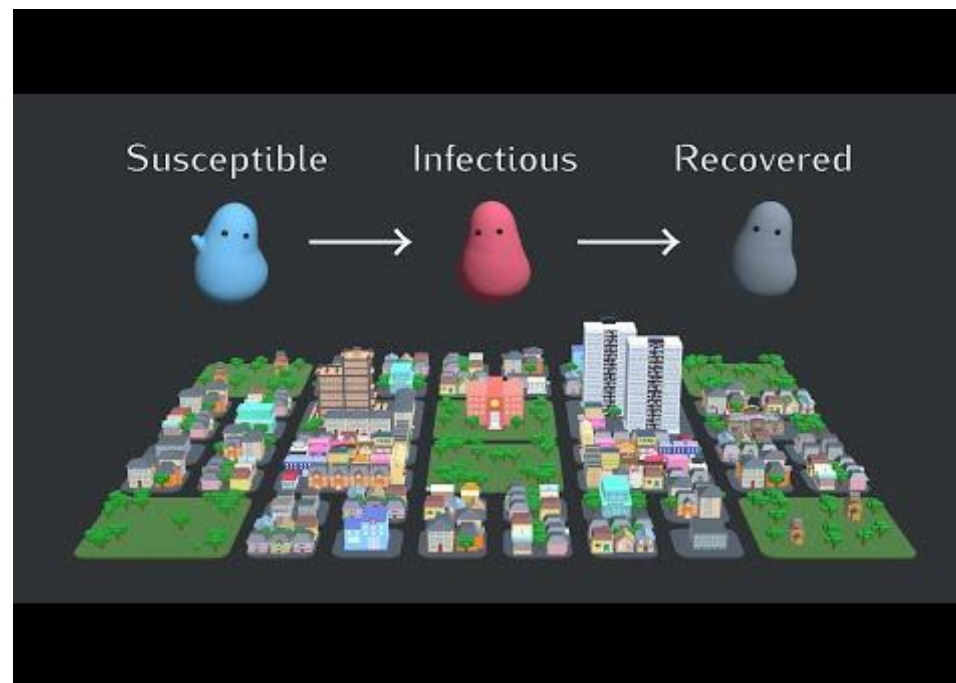
Ejemplo de Modelo: Epidemia, Endemia, Erradicación: Simulación

MODELO SIR:

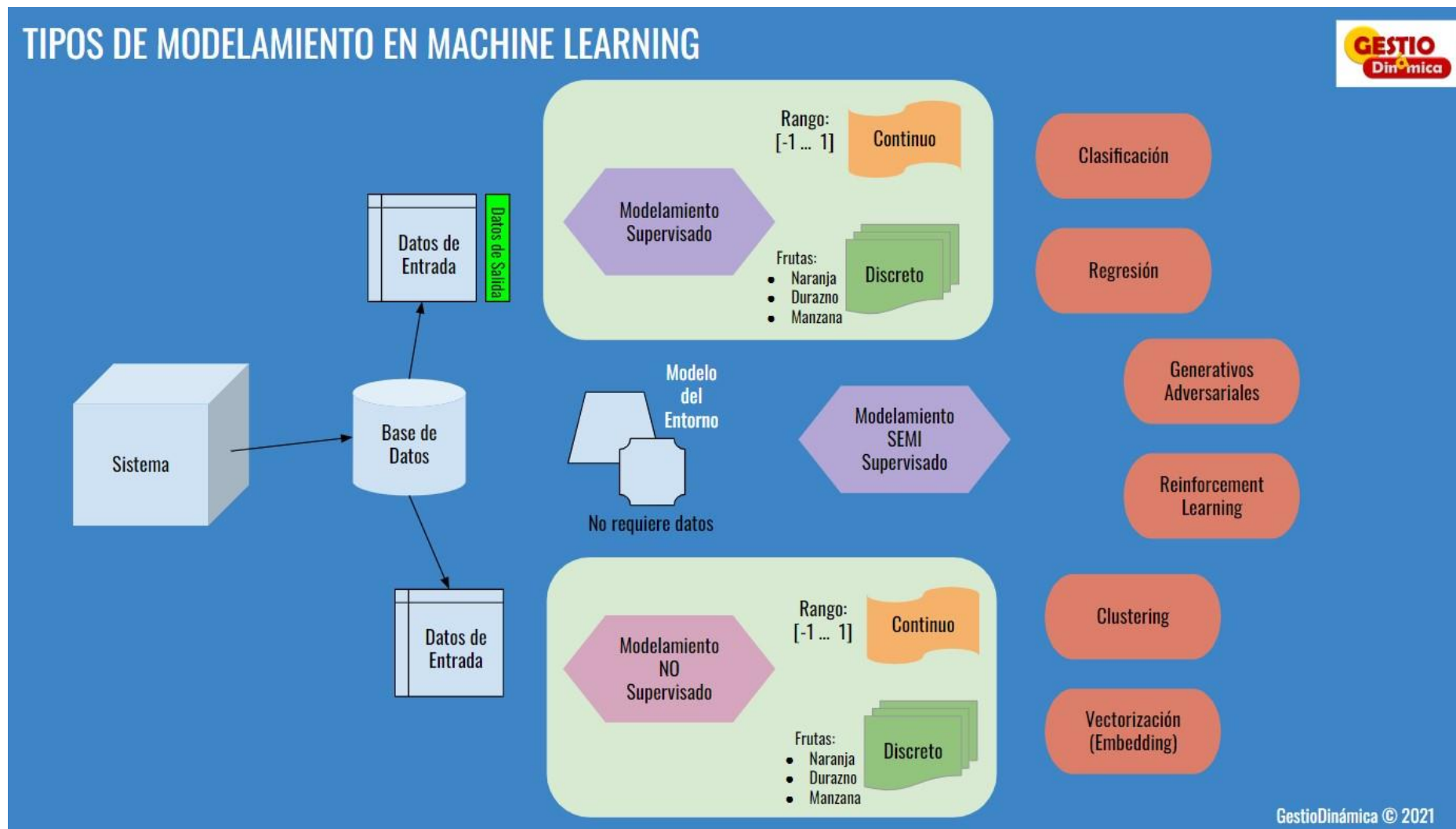
- Susceptible
- Infectado
- Recuperado

Instrucciones:

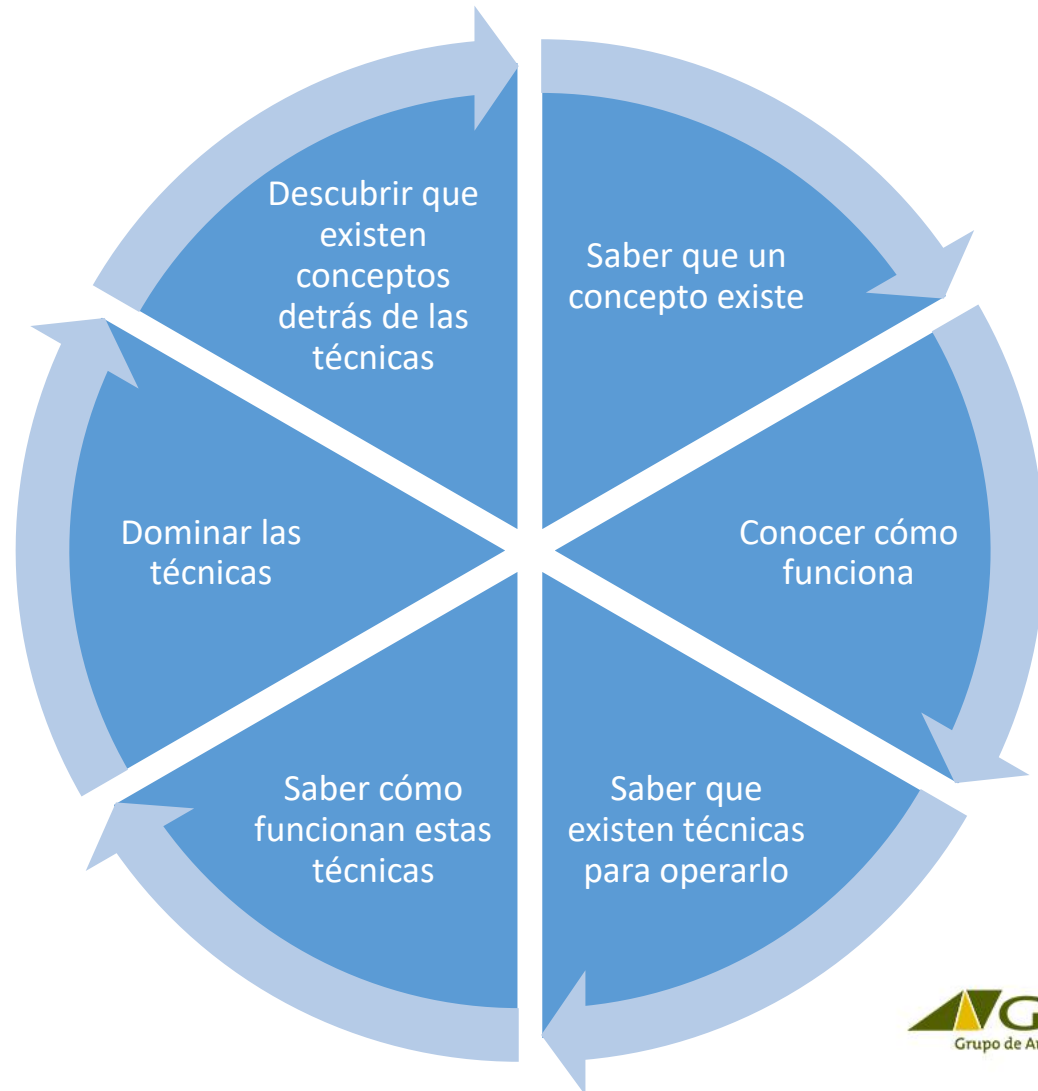
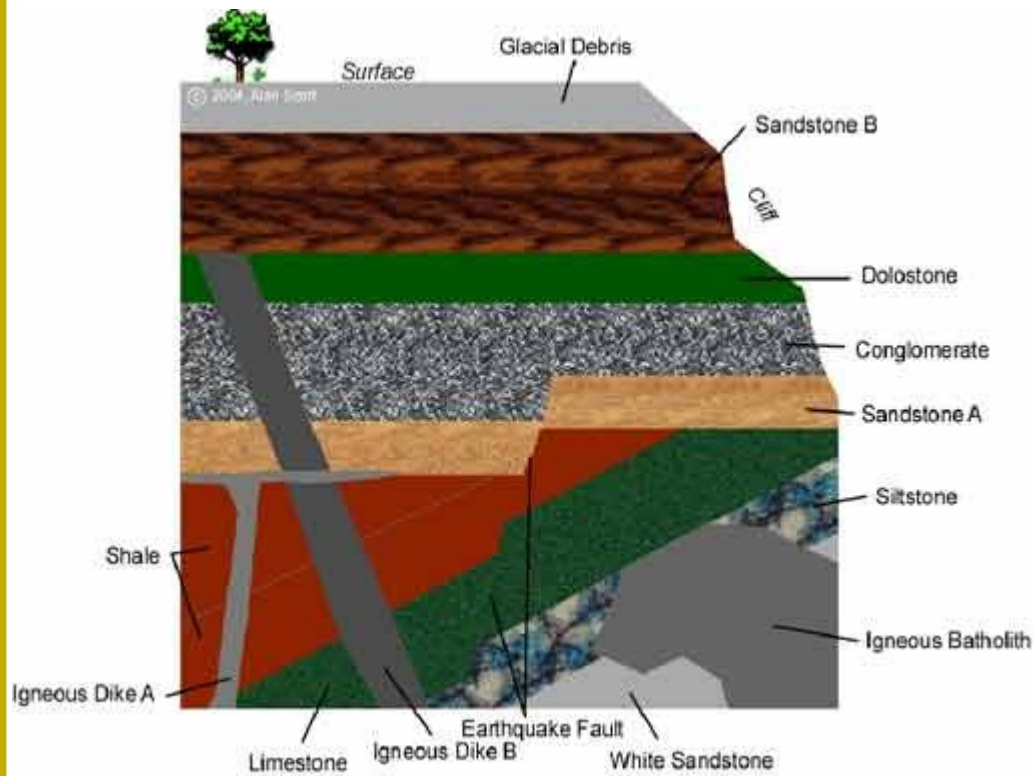
- Video: 12 min
- Opinión: Formulario
- Tipo: Individual



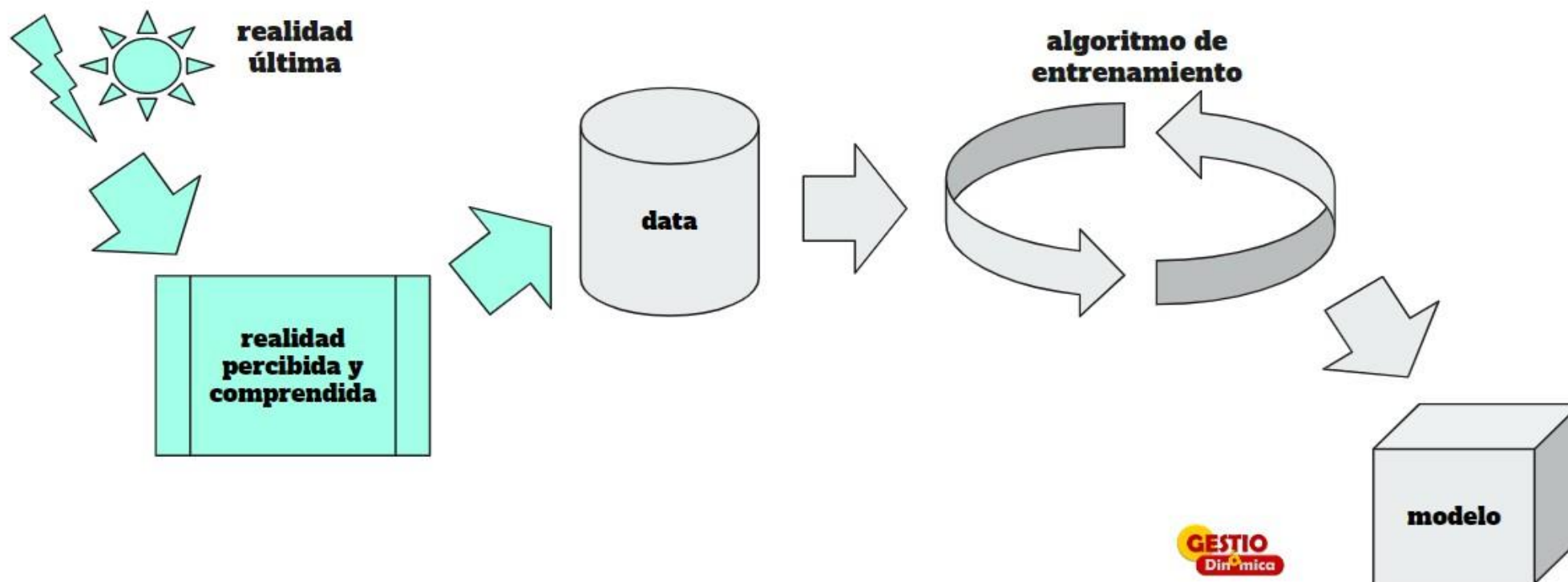
Los tipos de modelos para machine learning



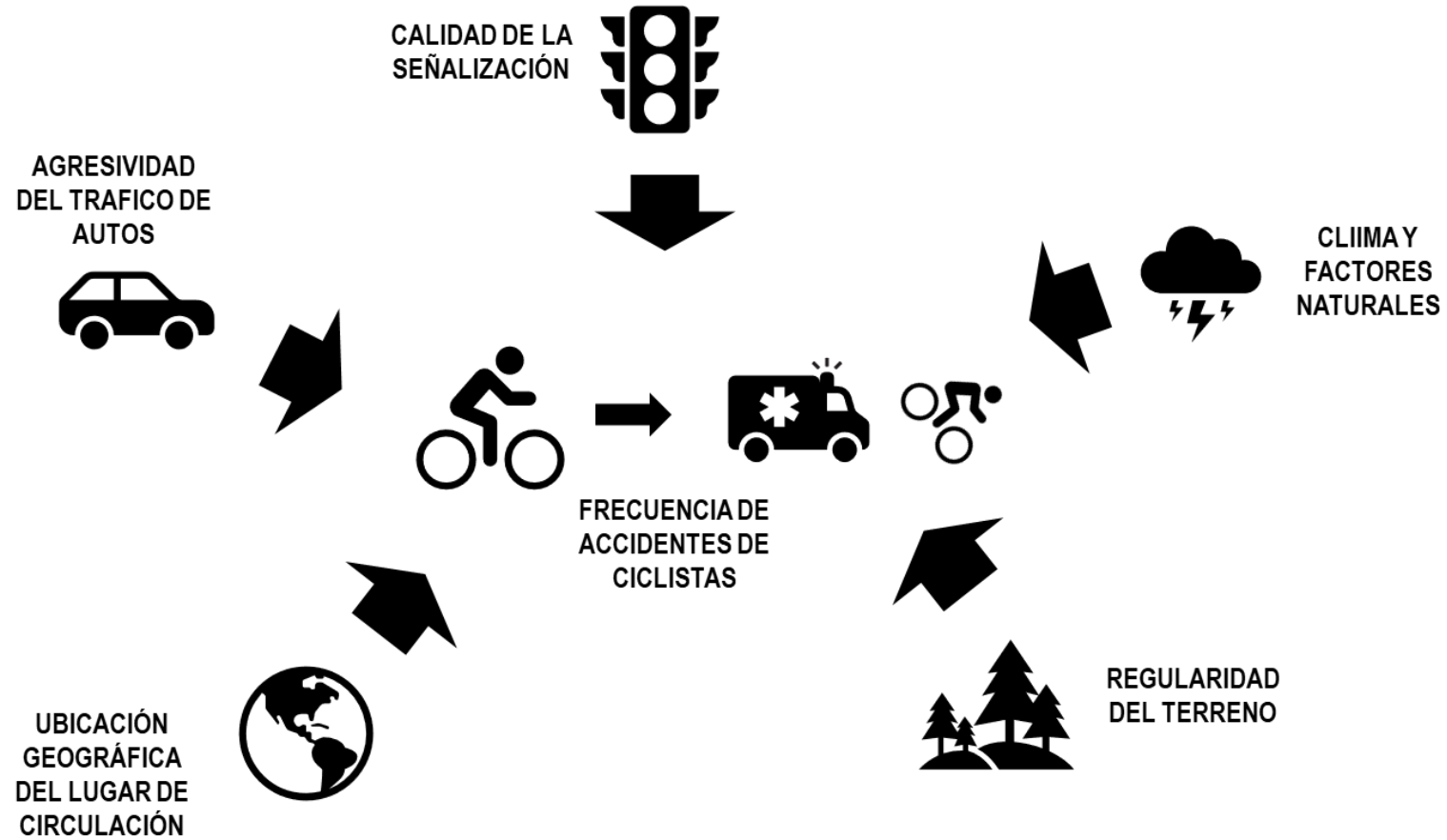
Las capas de acceso al conocimiento (capas cognitivas)



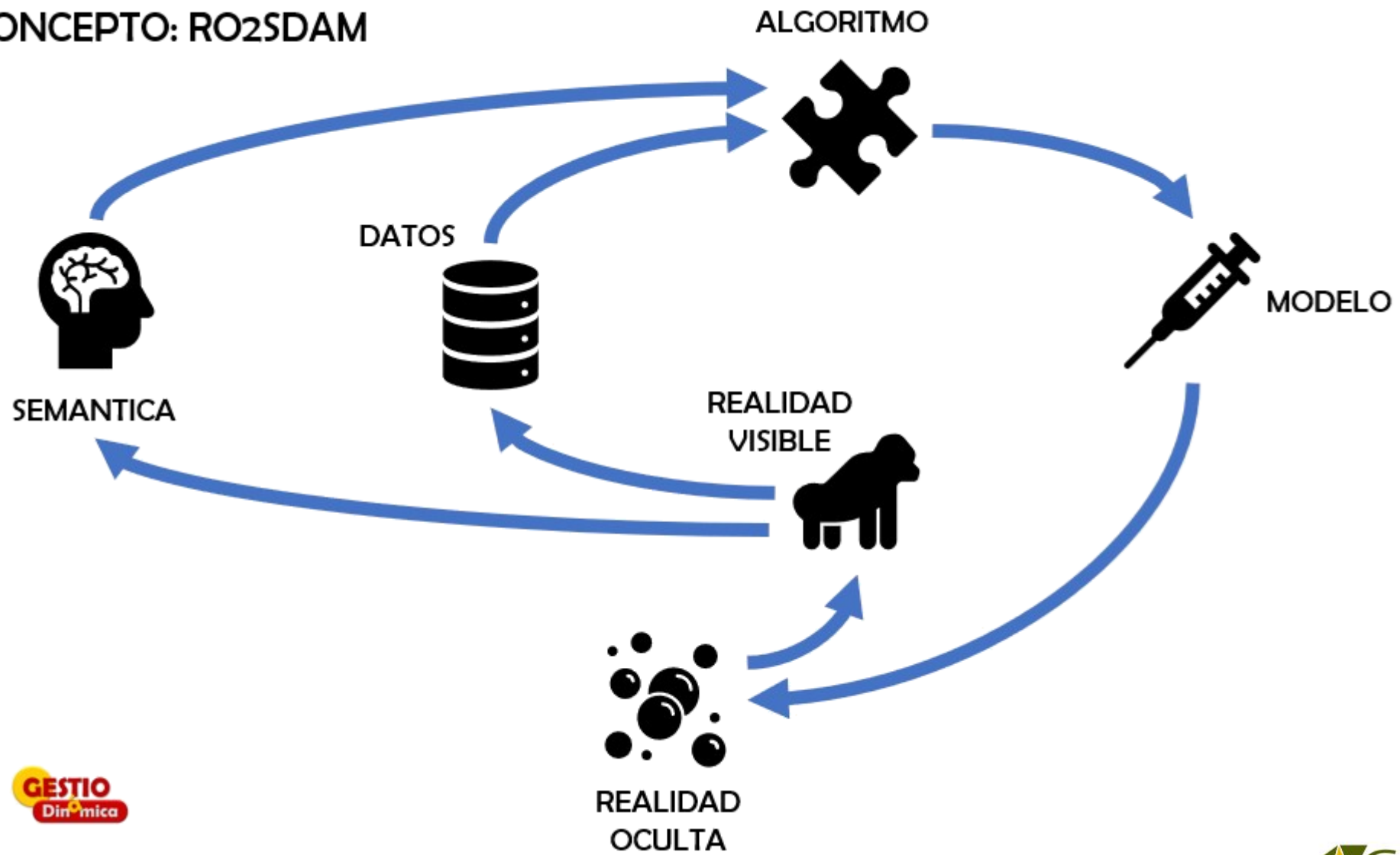
La forma de modelar la realidad se ha hecho más compleja



Modelamiento Conceptual Preliminar



CONCEPTO: RO2SDAM



Disciplinas vinculadas a la gestión de datos

GESTIÓN

- Cómo organizar los procesos como un todo.
- Cómo hacerlos viables con los actores vinculados.

INGENIERÍA

- Cómo diseñar los procesos con los datos
- Cómo hacer los procesos efectivos y vinculados a resultados.

DATOS



ARQUITECTURA

- Cómo definir la funcionalidad de un sistema de datos
- Cómo garantizar los accesos adecuados para los procesos.

CIENCIA

- Descubrir el comportamiento de los datos
- Explorar las estructuras que presentan

GestioDinámica©2022

S5

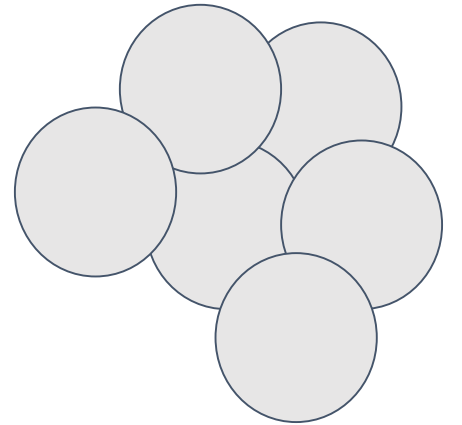
Modelamiento: Escuelas,
Enfoques, Lógicas, Algoritmos

Los límites de la estadística clásica



- Dado de 6 lados:
 - ¿cada lado igual probabilidad?
- Si es un experimento controlado, tiene a cumplirse.
- Si tiene factores externos
 - temperatura irregular del aire, la flexibilidad de los materiales, en número de personas en el lugar, entre otros.
- Definitivamente las personas no toman decisiones en base a probabilidades clásicas
 - Kahneman & Tverski (1975)

Teorema del Límite Central

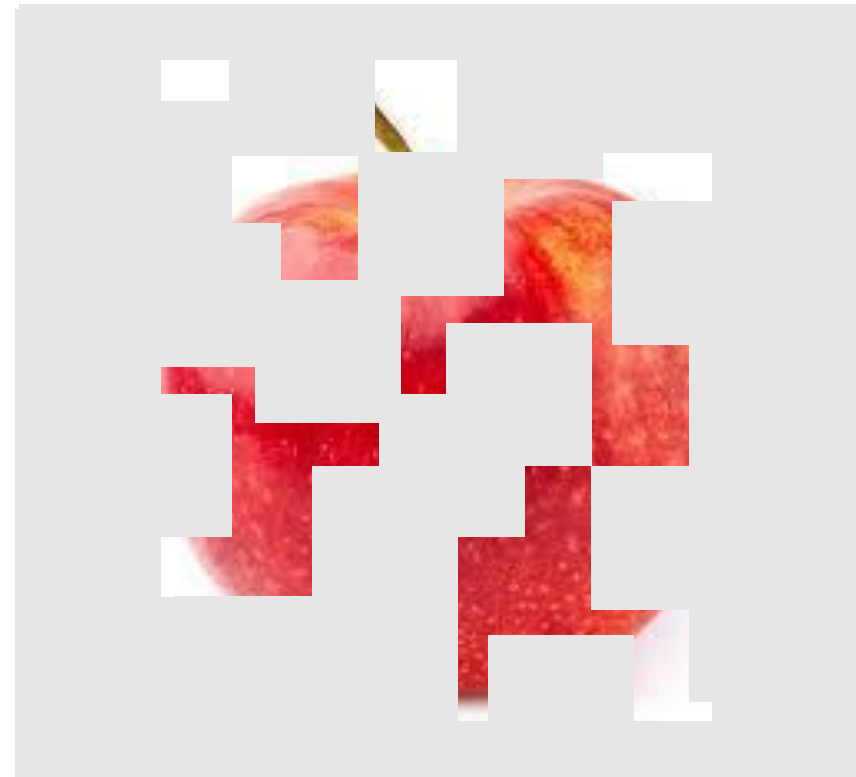


Modelo de Monte Carlo: Intuición

Cuando tenemos la
información completa



Cuando podemos tener
información sólo de ciertas
áreas aleatorias



Método de Monte Carlo

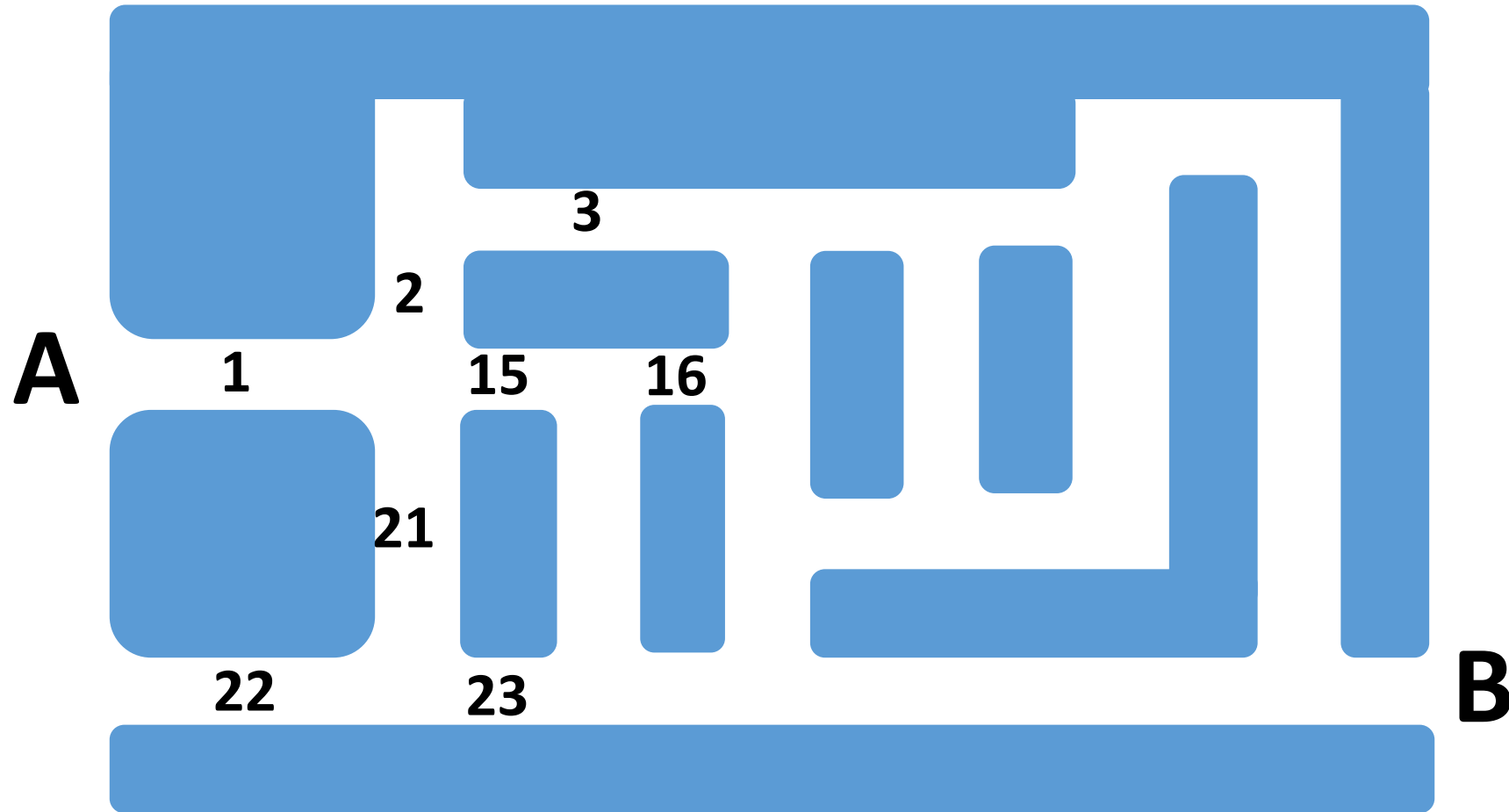
Definición

- Dado un sistema
- Se sabe que existen soluciones finitas
- No se tienen soluciones matemáticas reales
- Se generan datos aleatorios del sistema
- Se registran los resultados y se detectan patrones
- Estos patrones tienen una alta probabilidad de ser equivalentes a un patrón deducido

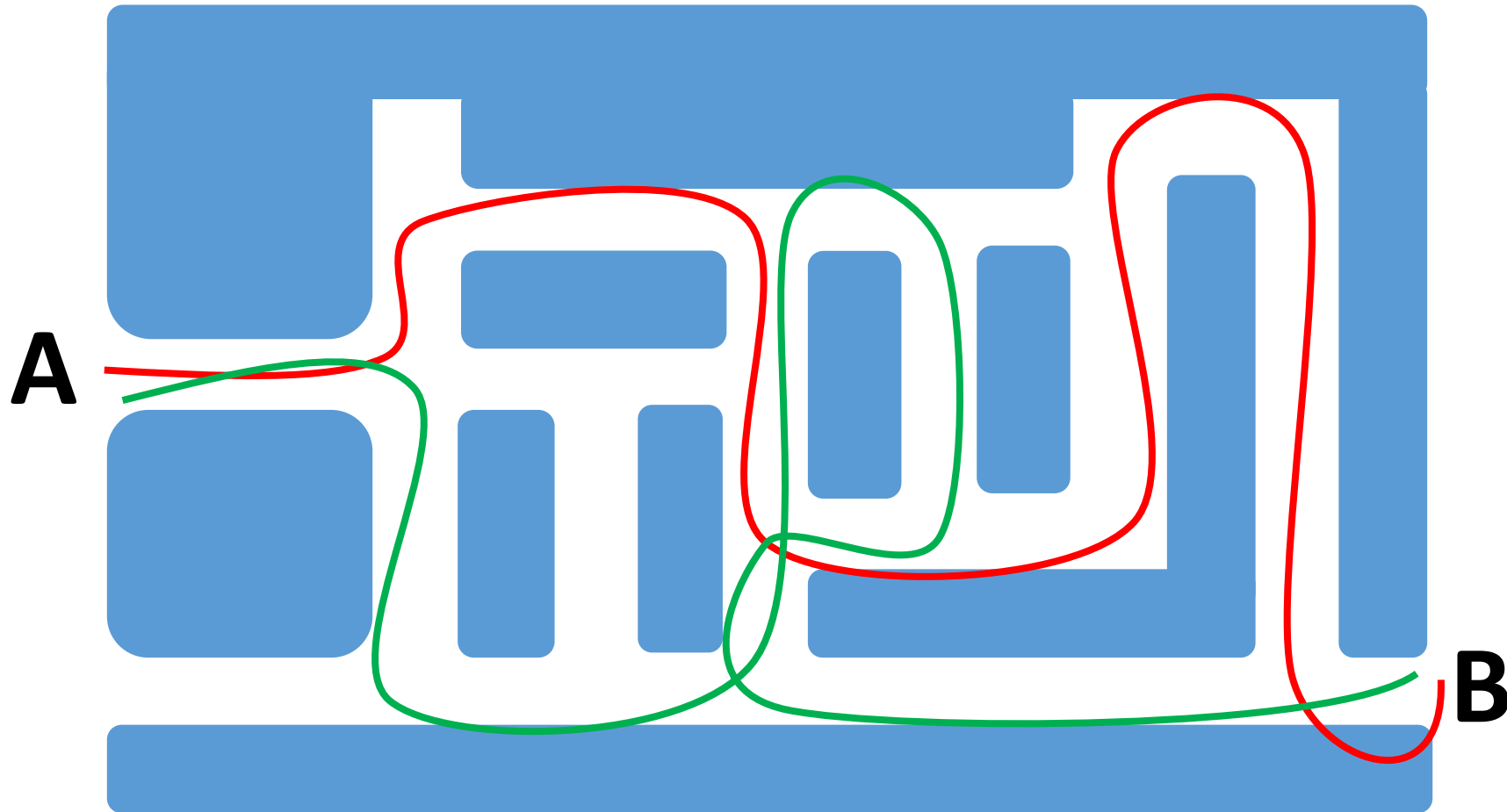
Ejemplo

- Sistema: Tránsito de vehículos en un sistema vial
- Físicamente (matemáticamente) se puede definir qué rutas son mejores (más rápidas, más seguras)
- No se tiene una fórmula que revele la ruta óptima
- Se generan escenarios aleatorios de rutas
- Se elige la ruta óptima de muchas rutas simuladas

Método de Monte Carlo



Método de Monte Carlo



Teoría de Redes Complejas (Grafos)

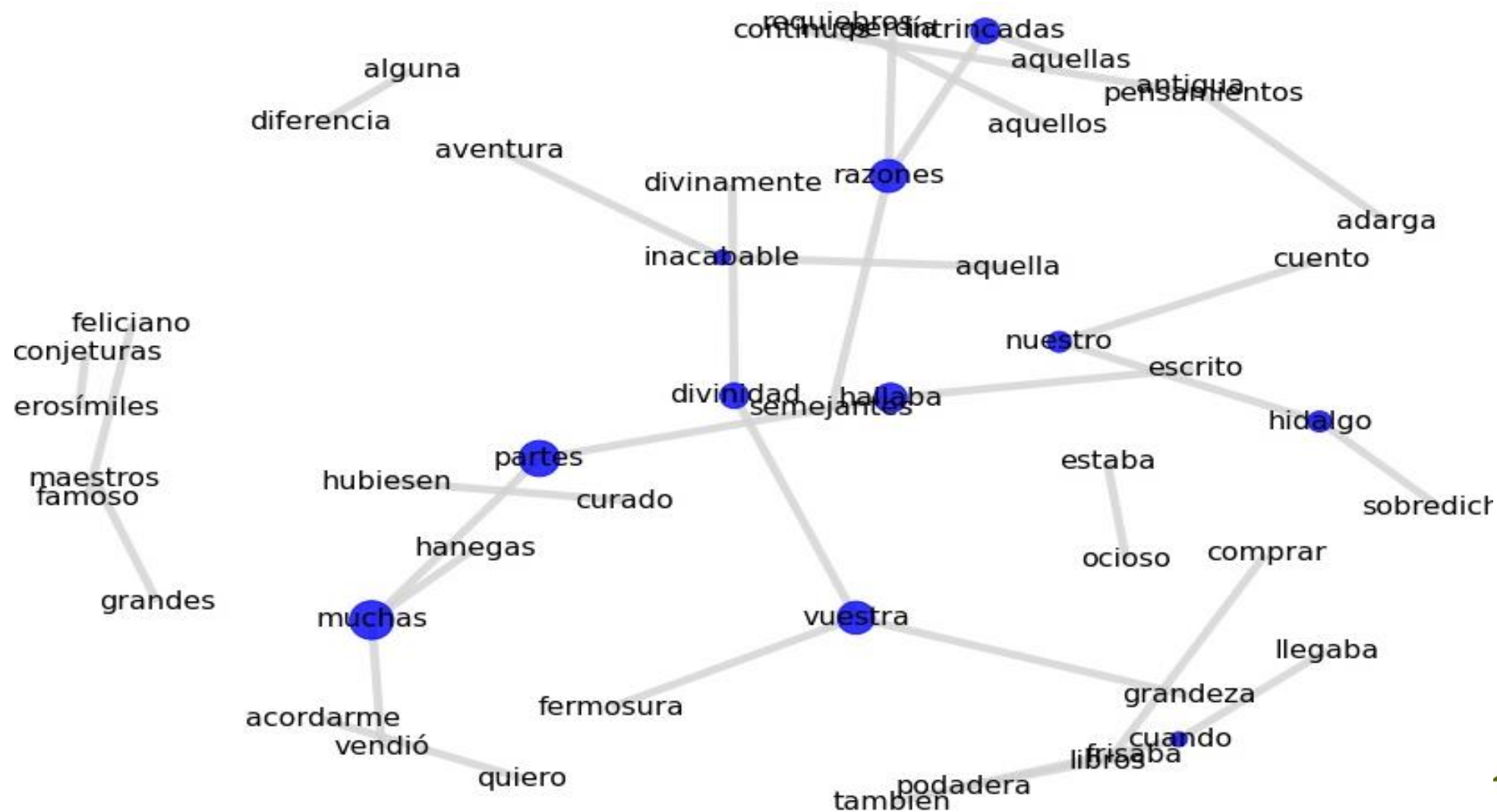
En un lugar de la Mancha, de cuyo nombre no quiero acordarme, no ha mucho tiempo que vivía un hidalgo de los de lanza en astillero, adarga antigua, rocín flaco y galgo corredor.

Una olla de algo más vaca que carnero, salpicón las más noches, duelos y quebrantos los sábados, lentejas los viernes, algún palomino de añadidura los domingos, consumían las tres partes de su hacienda.

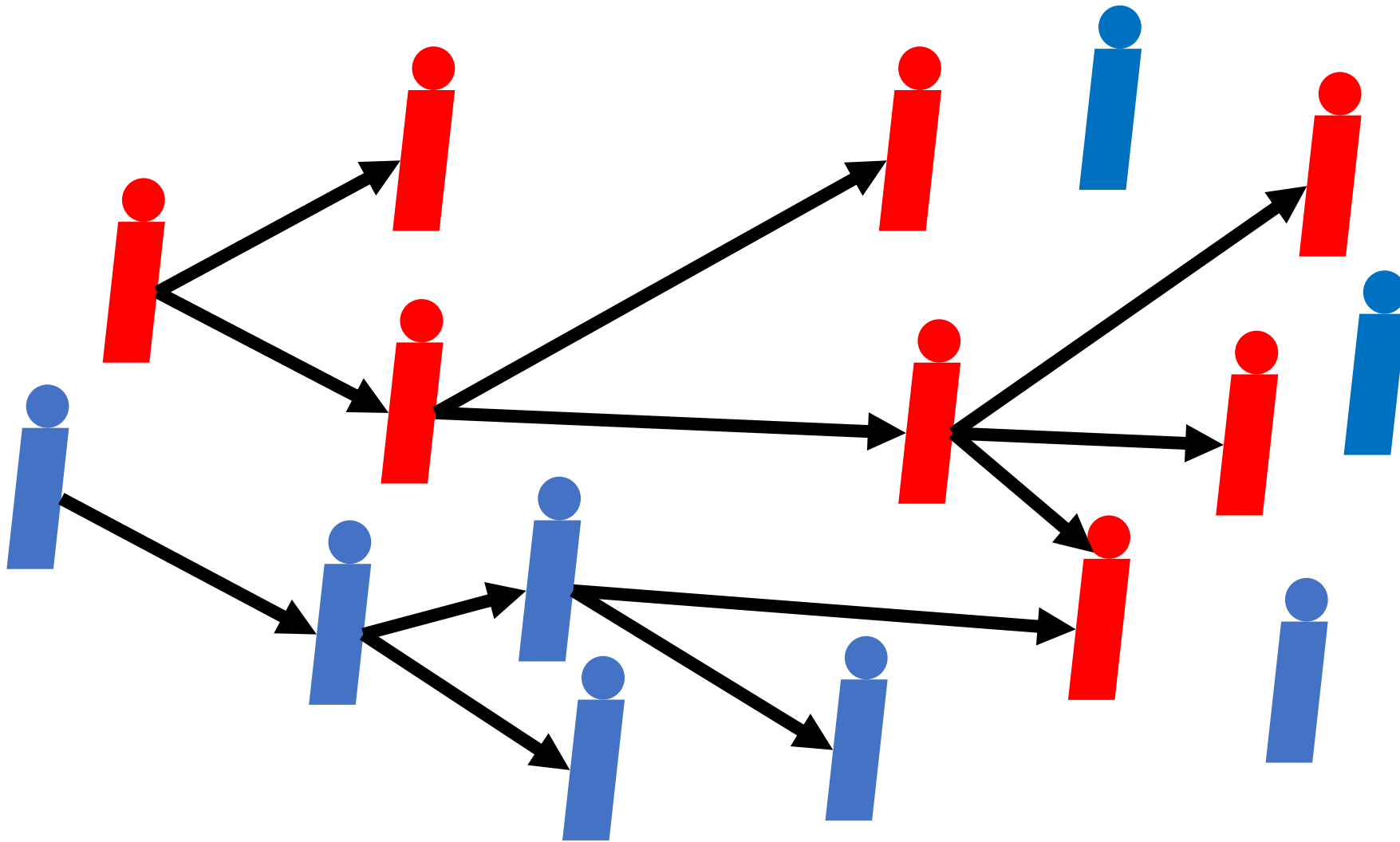
El resto della concluían sayo de velarte, calzas de velludo para las fiestas con sus pantuflos de lo mismo, los días de entre semana se honraba con su vellori de lo más fino.

Tenía en su casa una ama que pasaba de los cuarenta, y una sobrina que no llegaba a los veinte, y un mozo de campo y plaza, que así ensillaba el rocín como tomaba la podadera.

Teoría de Redes Complejas (Grafos)



Dinámica de Sistemas



Dinámica de Sistemas

$$630 - 480 = 150$$



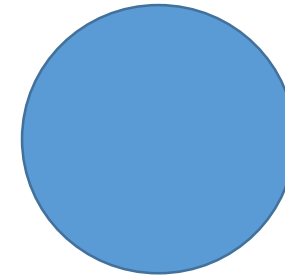
480



630

$$\frac{630}{480} = 1.31$$

$$807 - 630 = 170$$



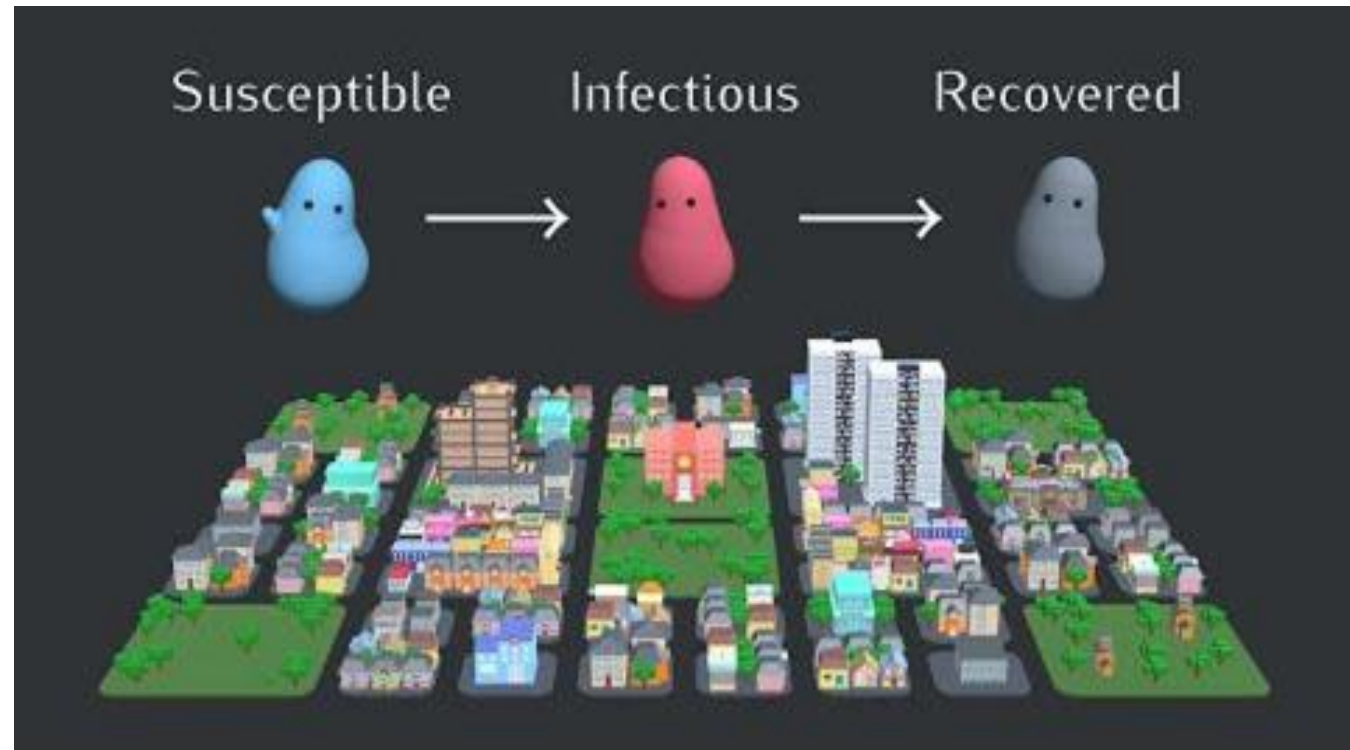
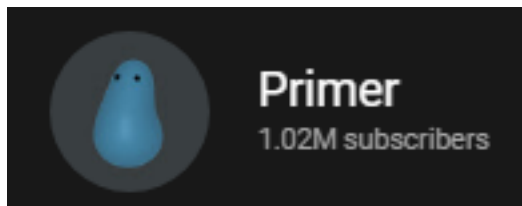
807

$$\frac{807}{630} = 1.28$$

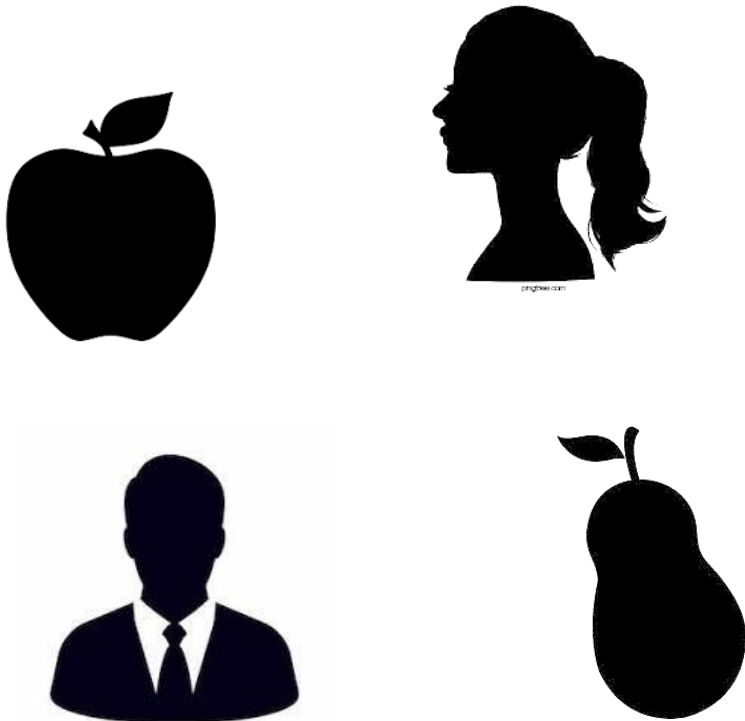
Dinámica de Sistemas

MODELO SIR:

- Susceptible
- Infectado
- Recuperado



Estadística Bayesiana



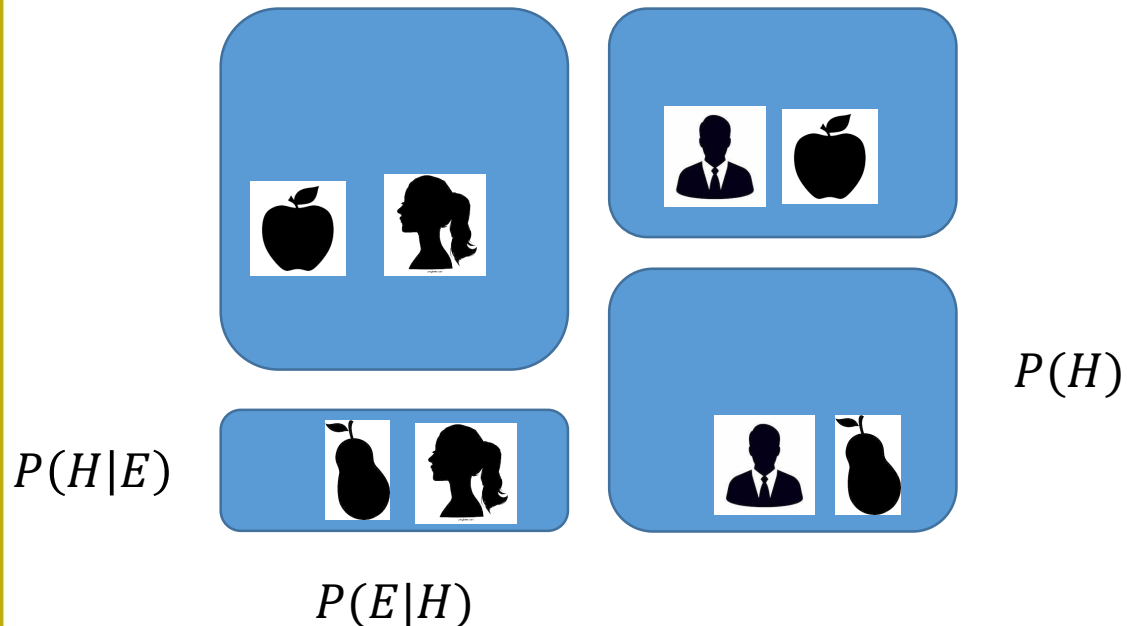
- Existe un grupo de consumidores para el cual quiero estimar su preferencia de productos
- No tengo datos sobre la preferencia de sus productos
- Pero tengo datos sobre las características de los consumidores (edad, sexo, profesión)
- Sé que hay cierta relación entre sus características y sus patrones de consumo
- Aplicando esta relación, puedo predecir el consumo en función de la probabilidad bayesiana

Estadística Bayesiana

$$P(H|E) = \frac{P(H) * P(E|H)}{P(E)}$$

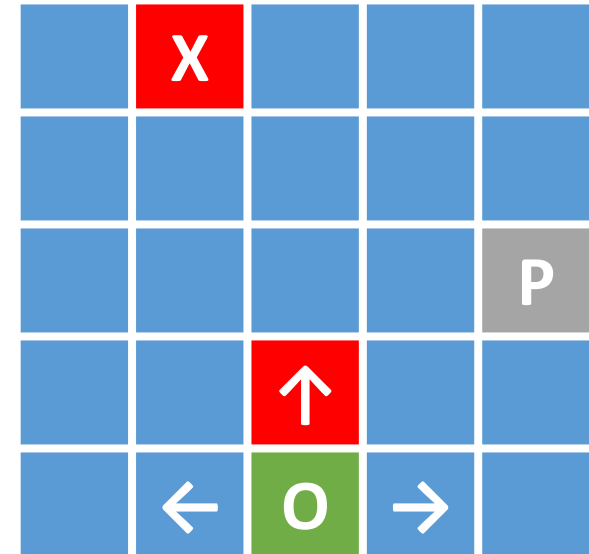
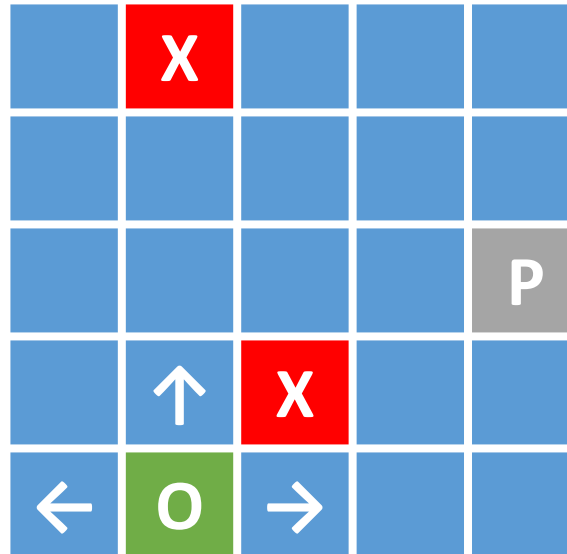
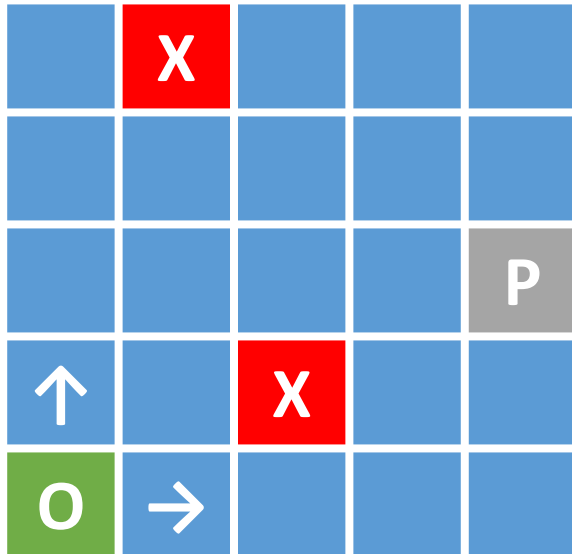
Diagram illustrating the components of Bayes' Theorem using fruit and person icons:

- $P(H|E)$: Posterior probability, represented by a pear and a person icon.
- $P(H)$: Prior probability, represented by a pear icon.
- $P(E|H)$: Likelihood, represented by a person and a pear icon.
- $P(E)$: Marginal probability, represented by a person icon.

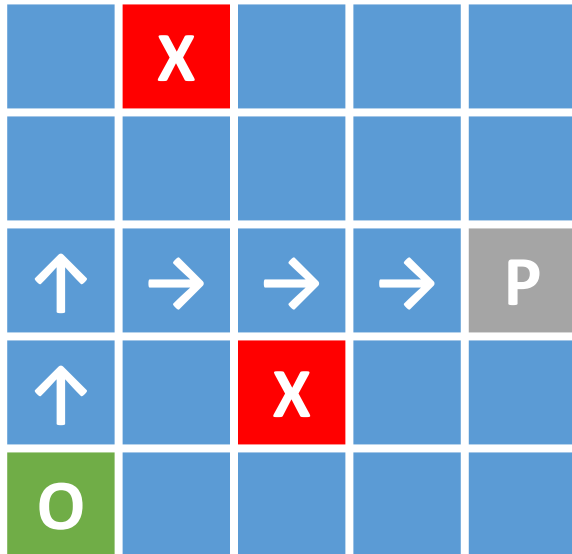


- Dado un sistema con dos factores
- Se quiere predecir el estado del factor final
- Pero no se tienen datos del factor final
- Se toman datos del factor secundario
- Se registran datos técnicos o históricos de la relación entre el factor final y el factor secundario
- Se aplica esa relación como una estimación del estado del factor final en función de una muestra de datos del factor secundario

Inteligencia Artificial: Reinforcement Learning

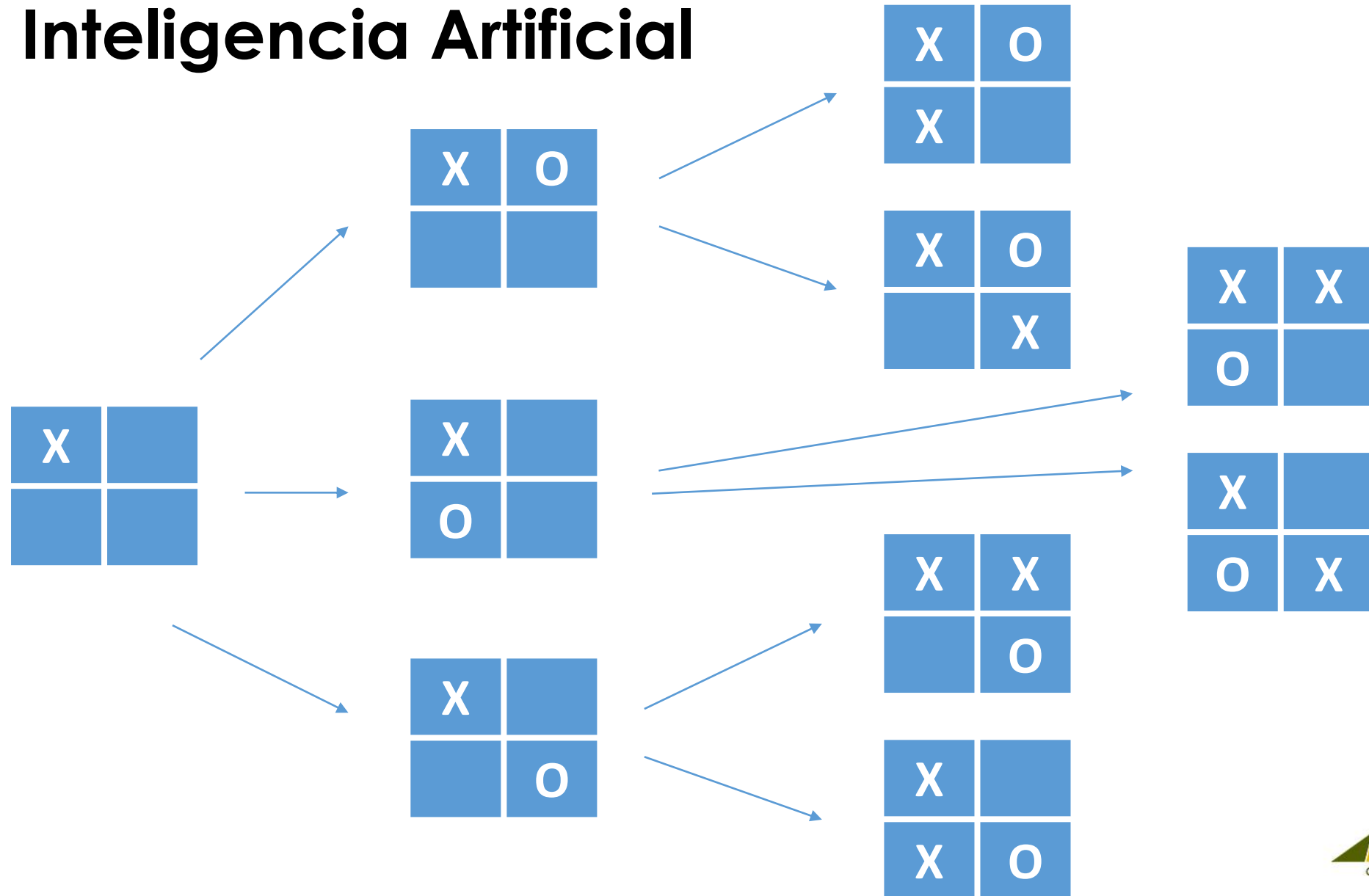


Inteligencia Artificial: Reinforcement Learning

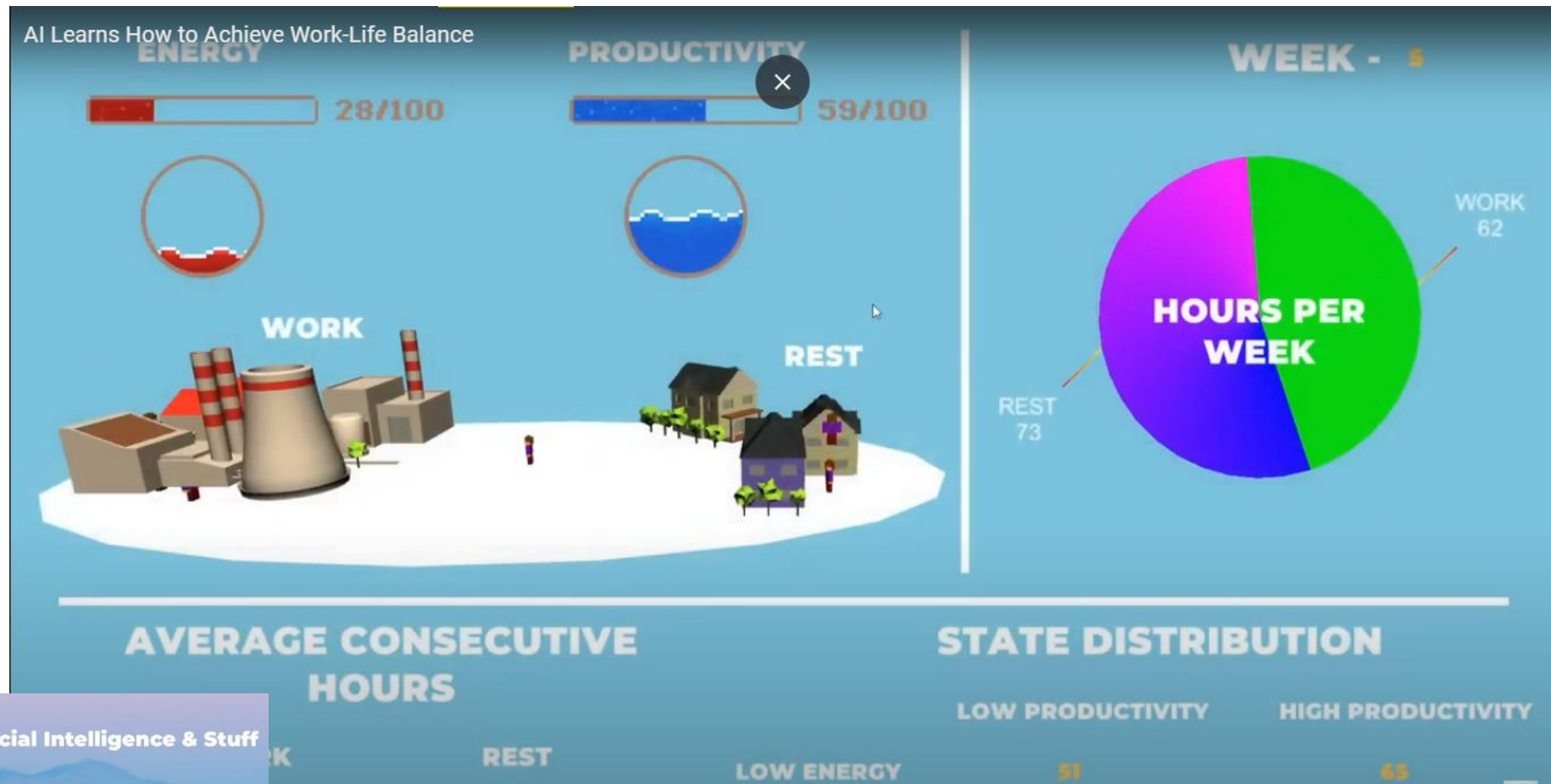


- Existe un agente
- Existe un entorno
- Existe un estado inicial de este agente
- Existe una colección de acciones que permiten un cambio de estado
- Se define un concepto de éxito final
- Se construye un sistema de recompensas locales (por cada cambio de estado), que no son conocidas inicialmente
- Se simulan muchos episodios
- Se va optimizando secuencias de acciones
- Se descubre la mejor secuencia de acciones para la mayor recompensa

Inteligencia Artificial



IA: Deduce aspectos que consideramos ambiguos



Comparación: física clásica y cuántica

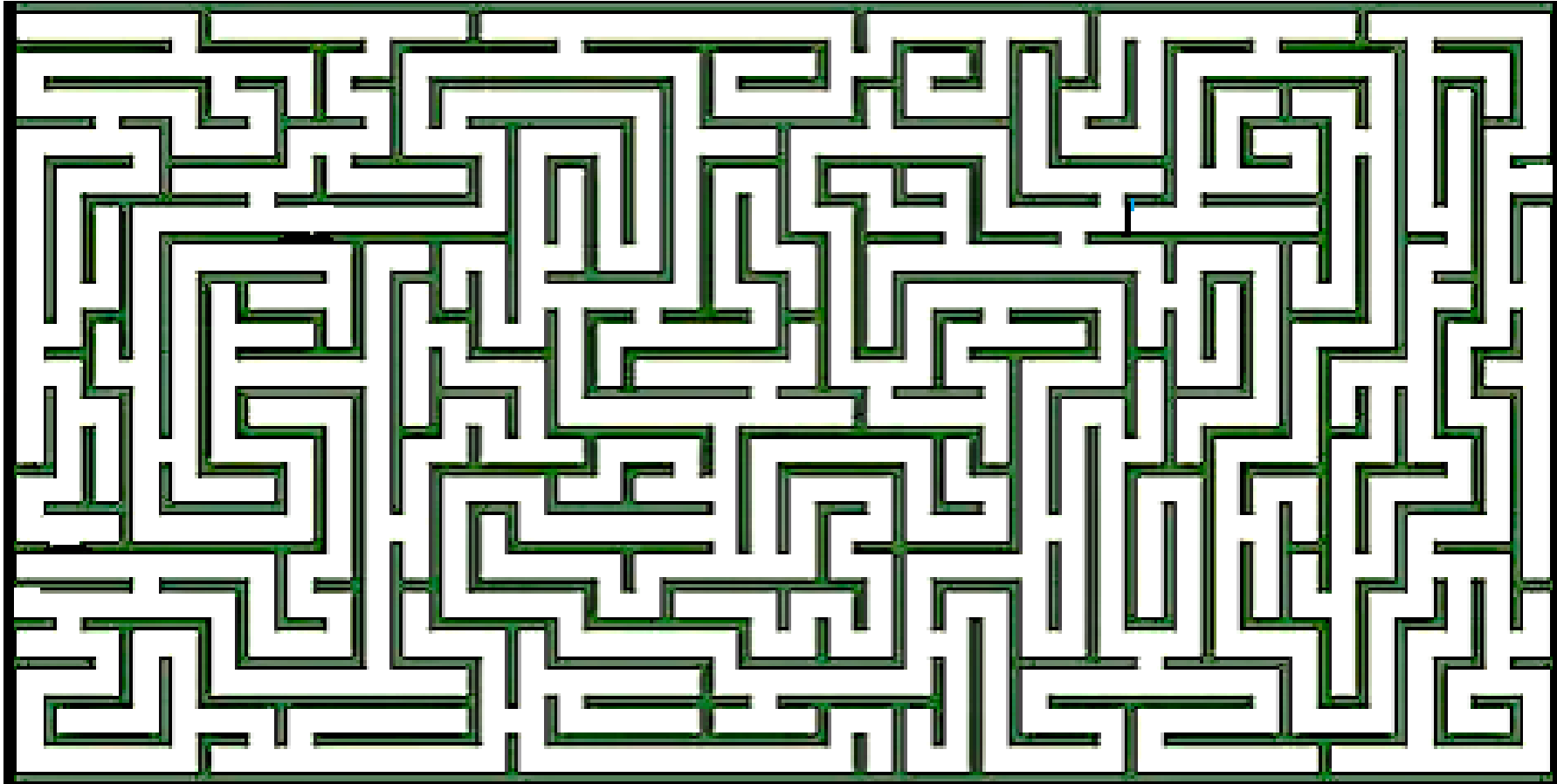
Física de Newton y vinculaciones

- Los cuerpos macroscópicos se mueven por acción y reacción
- La fuerza de la gravedad atrae a todos los cuerpos hacia el centro de la Tierra
- Los objetos se subdividen en partículas que pueden hacerse independientes una de otra
- La máxima divisibilidad de la materia es el átomo
- El tiempo y el espacio tridimensional son absolutos

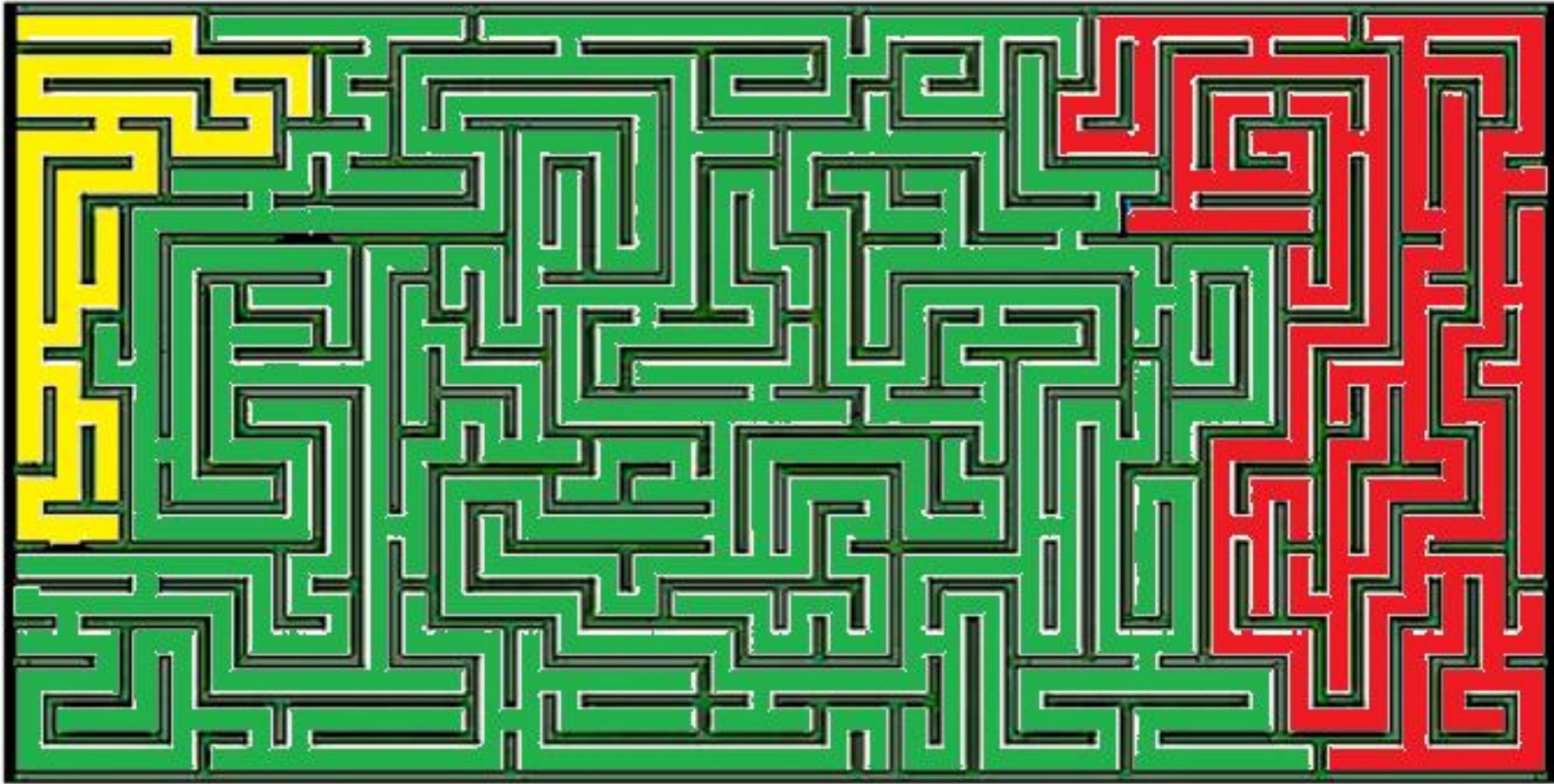
Física cuántica

- Los cuerpos macroscópicos se mueven por acción y reacción sólo bajo ciertas condiciones (no cuando están cerca de un agujero negro o cuando son, por ejemplo, superconductores o están a temperaturas muy bajas)
- La gravedad es sólo un corrector de las velocidades de diferentes puntos de un objeto
- Los elementos fundamentales de la materia no son partículas sino ondas que al interactuar generan fuerzas de diferente naturaleza
- El tiempo y el espacio tridimensional es relativo

La lógica cuántica



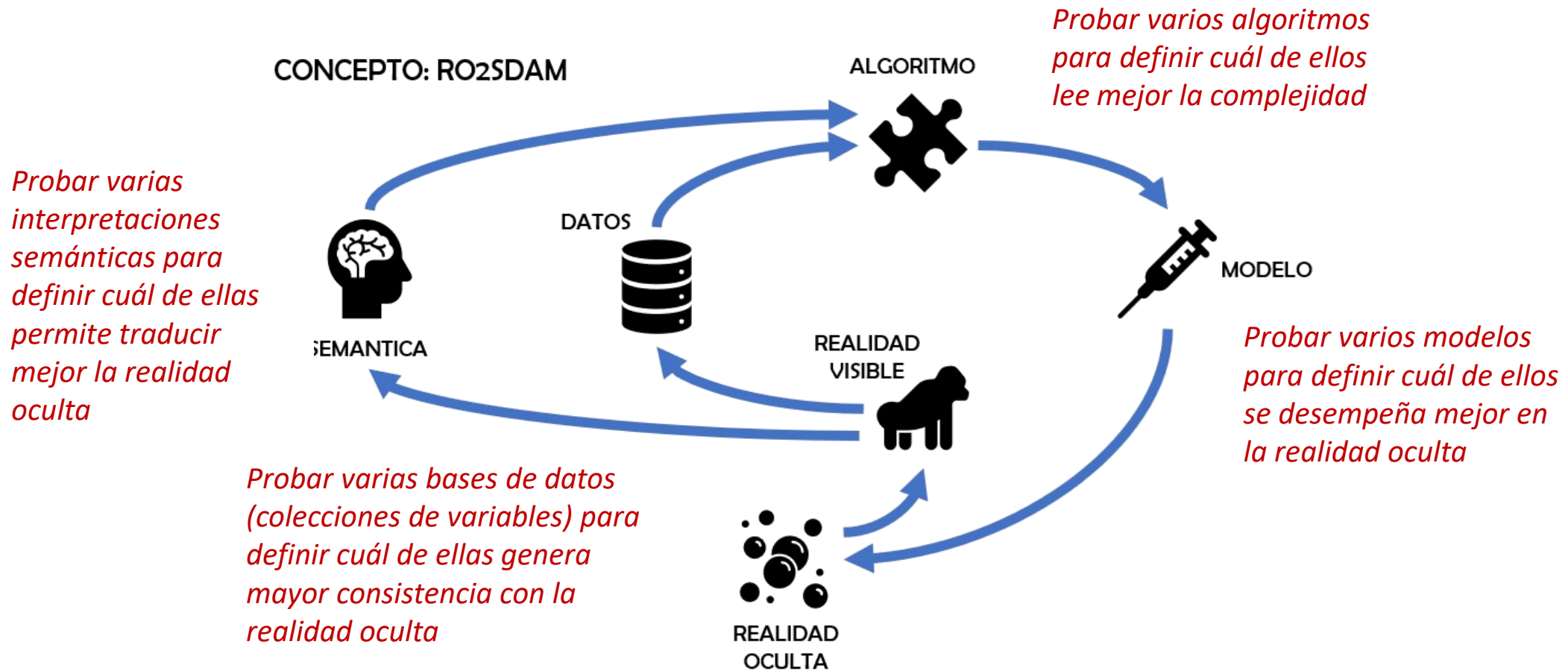
La lógica cuántica



S5

Modelamiento:
Experimentación/Simulación

La experimentación empieza desde la concepción inicial del modelo



Riesgos de la Experimentación en Modelamiento

Sobreajuste

- esto ocurre cuando un modelo es demasiado complejo y tiene demasiados parámetros, lo que hace que se ajuste muy bien a los datos de entrenamiento pero tenga un rendimiento deficiente en datos nuevos e invisibles.

Ajuste insuficiente

- esto ocurre cuando un modelo es demasiado simple y no puede capturar los patrones subyacentes en los datos, lo que lleva a un rendimiento deficiente tanto en los conjuntos de entrenamiento como de prueba.

Sesgo

- esto puede ocurrir si los datos utilizados para entrenar el modelo no son representativos de la población del mundo real, lo que lleva a predicciones que son sistemáticamente incorrectas para ciertos grupos.

Consideraciones éticas

- los modelos de aprendizaje automático pueden perpetuar y amplificar los sesgos existentes en los datos, lo que genera preocupaciones éticas sobre su uso en ciertos contextos. Es importante considerar cuidadosamente estos riesgos y tomar medidas para mitigarlos en el diseño experimental.

Tipos de Aprendizaje Computacional

Supervisado

- Con etiquetas
- Depende del conocimiento existente

No Supervisado

- No depende del conocimiento existente
- No requiere etiquetas
- Difícil de interpretar

Semi Supervisado

- Utiliza unas cuantas etiquetas solamente, pero podrían estar equivocadas
- Se basa en el conocimiento existente plasmado en un escenario

Auto Supervisado

- Genera sus propias etiquetas con un no supervisado inicial
- Podría arrastrar errores iniciales que lo encierran

Transfer Learning

- Necesita un alto volumen de datos, tarda en entrenarse
- No sirve para casos ad-hoc aunque puede hacerse fine-tuning para una parte de los casos

Ejemplos de objetos de experimentación en el modelamiento

Pruebas A/B:

- Comparar el rendimiento de dos modelos o configuraciones diferentes en una misma base de datos

Validación cruzada:

- Dividir los datos en pliegues, entrenar el modelo en algunos de los pliegues y evaluar su rendimiento en los pliegues restantes.
- Ayuda a prevenir el sobreajuste y a estimar el error de generalización del modelo.

Optimización de hiperparámetros:

- Búsqueda en cuadrícula o la búsqueda aleatoria para optimizar los valores de los hiperparámetros de un modelo, con el fin de mejorar su rendimiento.

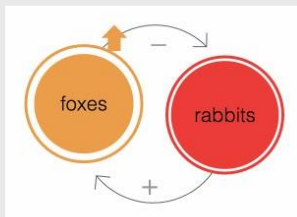
S5

**Modelamiento:
Dinámica de Sistemas**

Instrumentos prácticos de dinámica de sistemas



In a world filled with ever-more-complex technological, sociological, ecological, political & economic systems... a tool to make interactive simulations may not be that much help. But it can certainly try.



play with simulations

It's the ancient, time-honored way of learning: messing around and seeing what happens. Play with simulations to ask "what if" questions, and get an intuition for how the system works!

NetLogo

Home
Download
Help
Resources
Extensions
FAQ
NetLogo Publications
Contact Us
Donate

Models:
Library
Community
Modeling Commons

Beginners Interactive
NetLogo Dictionary
(BIND)
NetLogo Dictionary

User Manuals:
Web
Printable
Chinese
Czech
Farsi / Persian
Japanese
Spanish
(intro)
(tutorial #1) (#2)
(#3)
(guide)
(dictionary)

Donate

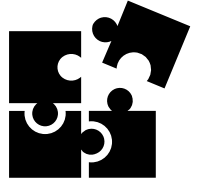
NetLogo is a multi-agent programmable modeling environment. It is used by many hundreds of thousands of students, teachers, and researchers worldwide. It also powers [HubNet](#) participatory simulations. It is authored by [Uri Wilensky](#) and developed at the [CCL](#). You can download it free of charge. You can also try it online through [NetLogo Web](#).

What can you do with NetLogo? Read more [here](#). Click [here](#) to watch videos.

Join mailing lists [here](#).

Download NetLogo **Go to NetLogo Web**

NetLogo comes with a large library of sample models. Click on some examples below.



Resumen: Modelamiento y Simulación

- Los **modelos** son imperfectos, pero se puede saber qué modelo es mejor que otro en función de qué modelo está más cerca a la **realidad**
- La **simulación** es un instrumento para **anticipar** la realidad para conocer algo sobre los aspectos que no están incluidos en un modelo
- El **modelamiento** manual es laborioso, el **modelamiento automático** requiere más cuidados pero es menos laborioso para el equipo humano
- El **modelamiento automático** puede usar **datos** como fuente para sus estructuras, a diferencia de agentes humanos que usan **conceptos**
- La **simulación** automática reduce la incertidumbre sin necesidad de que el equipo humano aprenda, pero el equipo humano pierde una parte de la experiencia del aprendizaje

S5

Python: Intro

Los datos como fuente de los modelos computacionales

Los **datos** se han incrementado, pero además se han **diversificado** y no necesariamente se presentan estructurados.

La ampliación de captura de datos significa también un reto de cómo **estructurarlos**.



Elementos básicos de Python

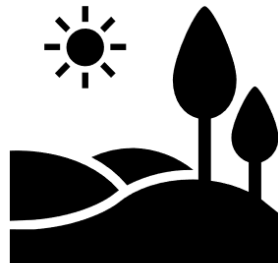


LIBRERÍA



Elegimos un dispositivo que
ya está preparado para
hacer imágenes

COMANDO



Queremos compartir con
alguien un paisaje

VALOR



Enviamos esa imagen
a la persona que
queremos compartir



Elementos básicos de Python

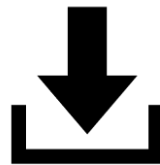


LIBRERÍA



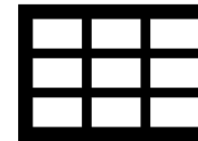
Usamos la librería de Pandas que administra la lectura e importación de archivos

COMANDO



Queremos importar un archivo

VALOR



Convertimos ese archivo en una variable con al cual podemos operar



Plataformas y Arquitecturas para Programar en Python

Plataformas
operativas

Google
colab


ANACONDA

Arquitecturas de Redes Neuronales

 **TensorFlow**

 **PyTorch**

Lenguaje de
Programación



python

TPU

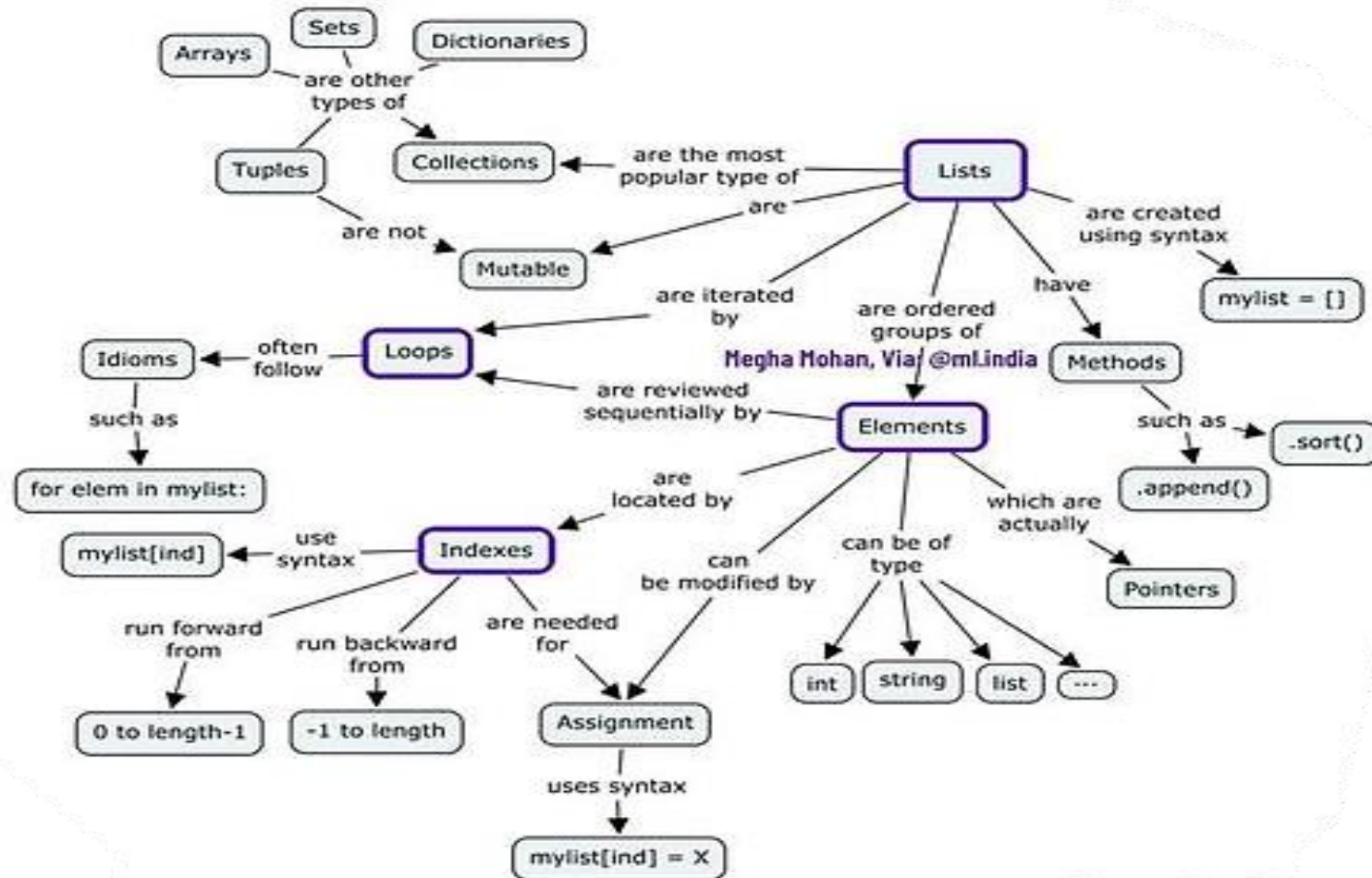
GPU

CPU

Hardware para Redes Neuronales



Estructuras para Programar en Python



Librerías típicas en Python

Pandas - used for data analysis	NumPy - multidimensional arrays
SciPy - algorithms to use with numpy	Matplotlib - data visualization tool
HDFS - used to store and manipulate data	PyTables - used for managing HDF5 datasets
Jupyter - research collaboration tool	IPython - powerful shell
HDFS - C/C++ wrapper for Hadoop	Pymongo - MongoDB driver
SQLAlchemy - Python SQL Toolkit	Redis - Redis' access libraries
pyMySQL - MySQL connector	Scikit-learn - used for machine learning algorithms
Theano - deep learning with neural networks	Keras - high-level neural networks API
Lasagne - build and train neural networks in Theano	Bokeh - data visualization tool
Seaborn - data visualization tool	Dask - data engineering tool
Airflow - data engineering tool	Luigi - data engineering tool
Elasticsearch - data search engine	SymPy - symbolic math
PyBrain - algorithms for ML	Pattern - natural language processing



¡Muchas gracias!