



Curso de posgrado  
Técnicas de aprendizaje automático – machine learning

# Tema 1. Introducción

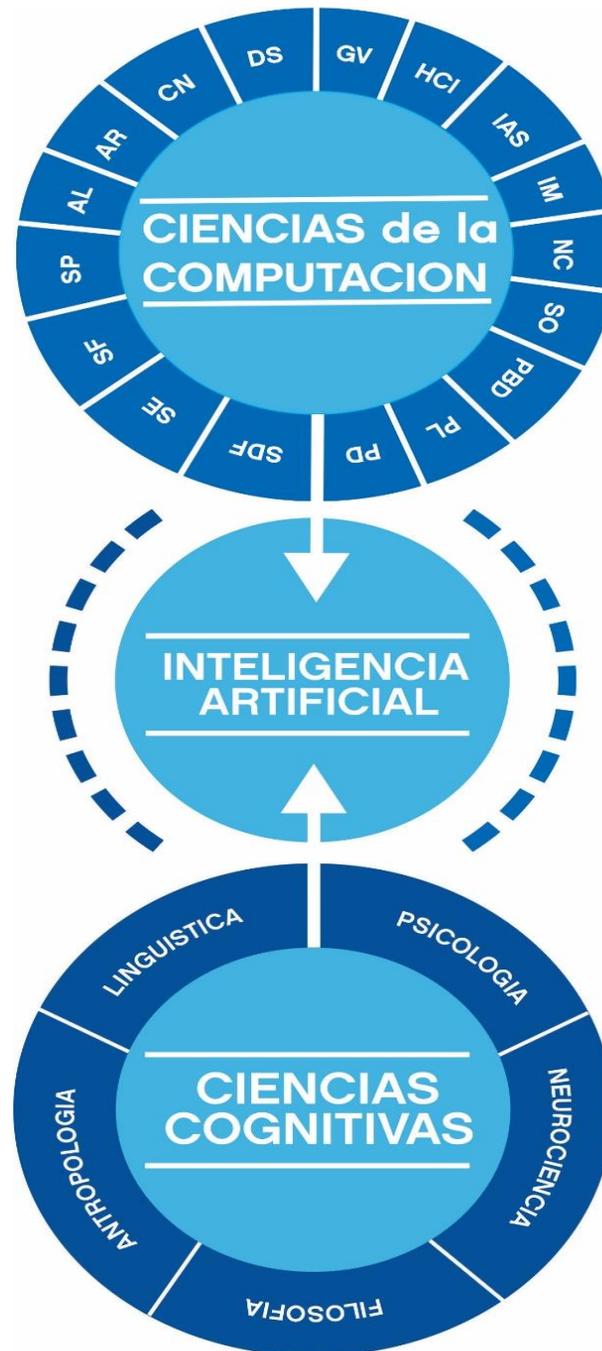
Prof. Dra. Sonia I. Mariño  
[simarinio@yahoo.com](mailto:simarinio@yahoo.com)

2023

# Indice

Introducir en

- antecedentes de la IA
- los conceptos de la IA
- el conocimiento y el aprendizaje en la IA
- las definiciones de IA, ML, RN, DL
- las metodologías en ML / CD
- los perfiles profesionales en IA
- la Ética en IA
- algunos eventos que tratan la IA



AL - Algoritmos y Complejidad,  
 AR - Arquitectura y Organización,  
 CN - Ciencias Computacionales,  
 DS - Estructuras discretas,  
 GV - Gráficos y visualización,  
 HCI - Interacción Hombre-Computador,  
 IAS - Garantía y seguridad de la información,  
 IM - Gestión de la información,  
 IS - Sistemas Inteligentes,  
 NC - Redes y comunicaciones,  
 SO - Sistemas operativos,  
 PBD - Desarrollo basado en la plataforma,  
 PD - Computación paralela y distribuida,  
 PL - Lenguajes de programación,  
 SDF - Fundamentos de Desarrollo de Software,  
 SE - Ingeniería de Software,  
 SF - Fundamentos de Sistemas,  
 SP - Cuestiones sociales y práctica profesional

Primeros Conceptos (Décadas de 1940-1950): Alan Turing y la Máquina de Turing.

- Concepto de IA y preguntas fundamentales.
- Desarrollo de las primeras ideas y teorías.

Nacimiento de la IA (Década de 1950): John McCarthy, "padre" de la inteligencia artificial.

- La conferencia de Dartmouth y el término "inteligencia artificial".
- Programas como el "Juego de Damas".

La Edad de las Reglas (Década de 1960-1970):

- Desarrollo de sistemas basados en reglas.
- Lenguajes de programación como LISP.

Auge y Caída (Década de 1980-1990):

- Avances en visión por computadora y procesamiento del lenguaje natural.
- Invierno de la IA: falta de resultados concretos.

Renacimiento de la IA (Década de 2000):

- Aprendizaje automático y algoritmos de IA.
- Desarrollo de redes neuronales y deep learning.

La Era del Big Data (Década de 2010):

- Avances en IA impulsados por grandes cantidades de datos.
- Aplicaciones en reconocimiento de voz, recomendaciones y más.

IA en la Vida Diaria (Década de 2020):

- Asistentes virtuales, automóviles autónomos, hogares inteligentes.

El Futuro de la IA (Década de 2030 y .....):

- Posibles avances como IA general, superinteligencia, ética de la IA.
- Impacto en sociedad, empleo y humanidad en general.

strategies which can be used to achieve this goal. The availability of your content has downloaded

creates a user-friendly commercial environment but also plays a leading role in increasing the sale of the products of the company to a larger extent. Jeff Bezos (CEO, Amazon) ideology of **Similar content being viewed by others**



## 2 Generations of recommender system

### 2.1 First-generation recommender system

Following are the different decision-making methods used in first-generation recommender system: content-based filtering, collaborative filtering and hybridized methods (Shvarts et al. 2017).

Cosine similarity, Euclidean distance and Pearson correlation coefficient (PCC) are the different similarity measures used by these methods (Billsus and Pazzani 2000; Adomavicius and Tuzhilin 2005). PCC has a major demerit of overestimating user preference when the user has rated only a few items identically (Ma et al. 2007).

### 2.2 Second-generation recommender system

Sections	Figures	References
<a href="#">Abstract</a>		
<a href="#">Introduction</a>		
<a href="#">Generations of recommender system</a>		
<a href="#">Data types used by recommender system</a>		
<a href="#">Knowledge discovery from database</a>		



youtube recomendaciones de videos academicos

- Videos
- Libros
- Noticias
- Imágenes
- Shopping
- Maps
- Vuelos
- Finance

Cerca de 4.420,000 resultados (0.54 segundos)

### Videos

-  **Recomendaciones Filmación Videos Académicos**  
YouTube - UPB BOLIVIA  
14 nov 2018
-  **Consejos para hacer un video académico efectivo**  
YouTube - UPB BOLIVIA  
11 oct 2019
-  **Tips para grabar un video académico desde casa**  
YouTube - Universidad Sergio Arboleda  
28 may 2021
- 3 momentos clave en este video
-  **Consejos para crear videos educativos sencillos pero con ...**  
YouTube - Elena Prieto  
22 abr 2020
- 10 momentos clave en esta video

inteligencia artificial



Introducción a la inteligencia artificial  
José María Girón Sierra



Chat GPT para DOCENTES: Inteligencia artificial al servicio de la educación  
Ediciones Rodio



El mito de la Inteligencia Artificial  
Erik J. Larson



Inteligencia artificial  
Carlos Pazos



I Jornadas sobre Inteligencia Artificial, Control y Sistemas Expertos: Alcalá de Henares, 1 al 4 de julio de 1996  
Control y Sistemas Expertos Jornadas sobre Inteligencia Artificial



MindfulDataAI, plugin para ChatGPT  
Contribuye actualizar la información brindada por  
CGPT a partir de búsquedas de Google

Modelos de lenguaje de gran tamaño (LLM), IA generativa



Genera 20 reviews de All But Thesis Syndrome en Scholar en el periodo 2020 a 2023



Lamento informarte que no tengo acceso a información en tiempo real ni a bases de datos específicas como Scholar. Además, no puedo generar reseñas de eventos futuros o publicaciones que no existen en mi base de datos actual, que se encuentra actualizada hasta septiembre de 2021. Sin embargo, puedo ayudarte a comprender qué es el "All But Thesis Syndrome" y proporcionarte información general sobre el tema. Si tienes alguna pregunta específica o necesitas ayuda con otro tema, no dudes en preguntar.



Lamentablemente, no tengo la capacidad de realizar análisis de sentimientos en tiempo real ni generar tablas en esta interfaz de texto. Sin embargo, puedo proporcionarte un análisis de sentimientos generalizado de los ejemplos anteriores para que puedas crear la tabla tú mismo:

Review	Sentimiento
1	Positivo
2	Neutral
3	Emotivo
4	Positivo
5	Neutro

Este análisis de sentimientos se basa en la tonalidad general de los textos ficticios proporcionados anteriormente. Ten en cuenta que los sentimientos son subjetivos y pueden interpretarse de diversas formas. Si tienes acceso a herramientas de procesamiento de lenguaje natural, podrías realizar análisis de sentimientos más precisos y detallados.

# Algoritmos LLM (Large Language Model) .



- ChatGPT, GPT (Transformador preentrenado generativo) de OpenAI
- Sydney de Microsoft construido a partir de ChatGPT para Bing.
- Bard, LaMDA de Google, (modelo de lenguaje para aplicaciones de diálogo)

## ChatGPT →

Interact with our flagship language models in a conversational interface

## DALL·E →

Create realistic images and art from a description in natural language

## API →

Integrate OpenAI models into your application or business

✦ Soy Bard, un colaborador creativo y útil. Tengo limitaciones y a veces me equivoco, pero tus comentarios me ayudarán a mejorar.

¿No sabes por dónde empezar? Puedes intentar lo siguiente:

[Quiero empezar a leer más ¿Qué puedo hacer?](#)

[Quiero hacerme las uñas en casa ¿Qué diseños son fáciles de hacer?](#)

También puedes retomar tu actividad donde la dejaste:

[realiza una síntesis de la evolución de la "Inteligencia Art...](#)

 Es posible que tus conversaciones de Bard sean procesadas por revisores humanos con fines de calidad. No ingreses información sensible.

[Más detalles](#)

[Entendido](#) [No volver a mostrar](#)

[Novedades de Bard \(google.com\)](https://google.com)

Bard Experimental

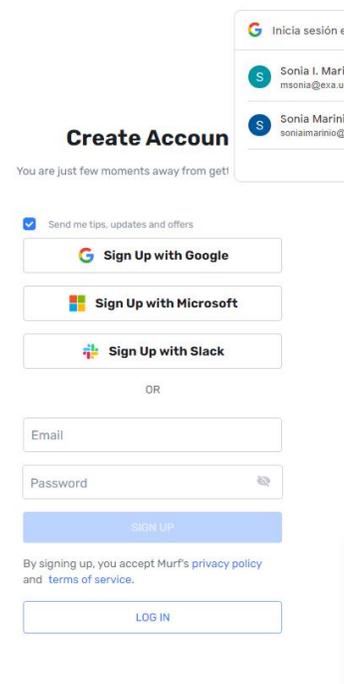
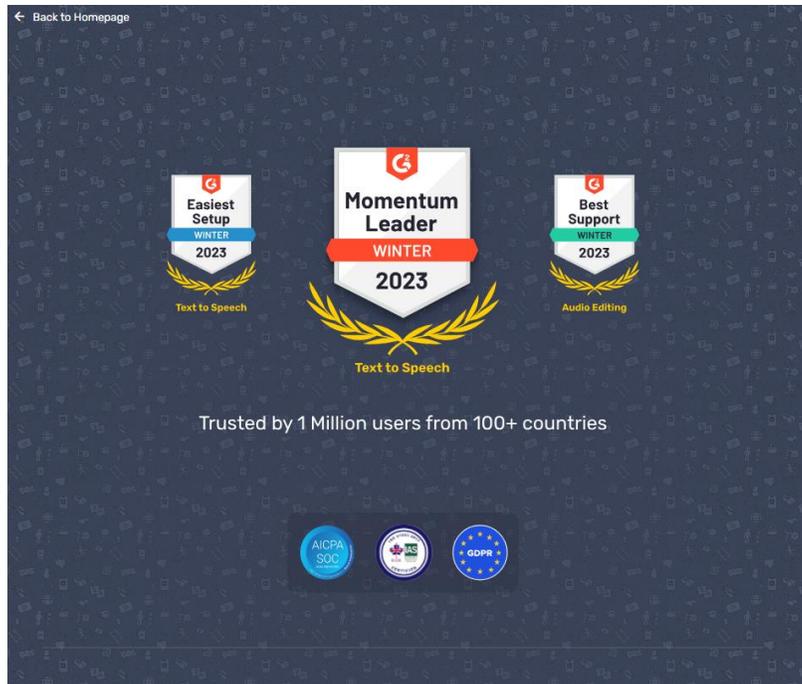
Aplican algoritmos de procesamiento de lenguaje natural (NLP) y Transformers.

## Plataformas para creadores de contenidos

<https://lovo.ai/>

plataformas de IA de voz más completas y fáciles de usar para producir voces que se asemejen a la humana. Su generador con Inteligencia Artificial «Genny» permite convertir texto a voces y añadirlas a la edición de video en la misma herramienta.

<https://murf.ai/studio/login>



Pruebas de conceptos  
 asignatura  
 Inteligencia Artificial - 2021



**Chatbot – Y. Aquino-L. Ledezma**

# CHATBOTCITO ASISTENTE VIRTUAL DE TELEGRAM

## EL CHATBOT PARA TU UNIVERSIDAD

ChatBotcito esta totalmente programado en Python y disponible en Telegram.

ChatBotcito te permite atender las consultas de los estudiantes realizadas en lenguaje natural.

Utiliza a ChatBotcito para obtener información sobre Mesas de Exámenes, Páginas Oficiales, Sitios de interés, entre otras.

Comunicación clara y efectiva al servicio de los estudiantes y solamente se necesita tener un telegram.

**@CHATBOTCITO\_BOT**

**Chatbot – S. Caselli, G. Escalante**

# IA en Ciencias de la Computación

- Ciencia que intenta crear programas/algoritmos que imiten el comportamiento, el razonamiento y la comprensión humana, que sea capaz de aprender, reconocer y pensar.
- Tecnología que favorece la creación de artefactos para resolver problemas de las personas

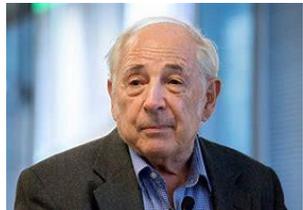
# Test de Turing



Alan Turing (1950), propuso un test, una máquina debe desarrollar la habilidad de comportamiento inteligente equivalente a una persona



## Test de la habitación China (Searle)

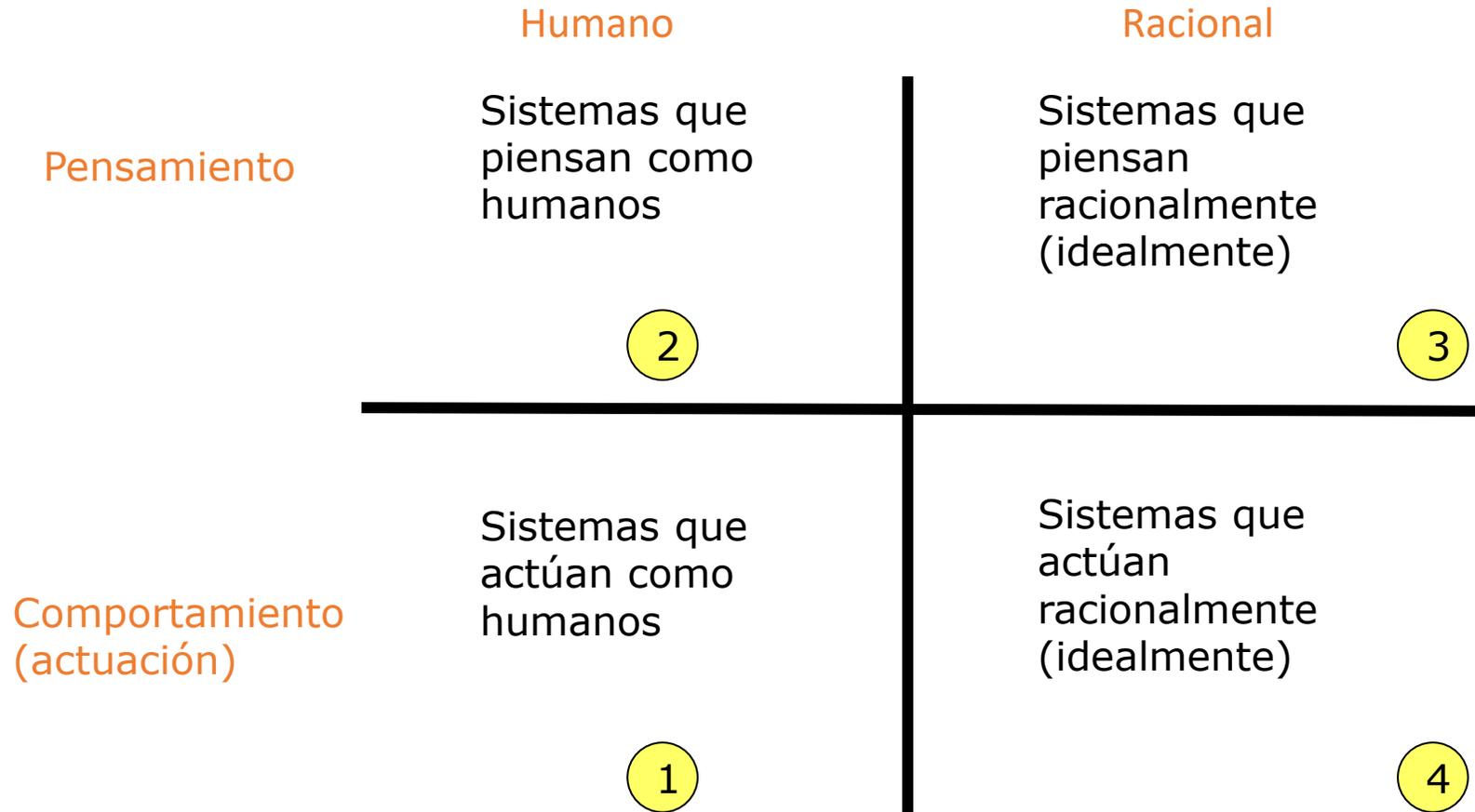


John Searle (1980), planteó que la capacidad de manipular ordenadamente un conjunto de símbolos no necesariamente implica que exista una comprensión o un entendimiento lingüístico de esos símbolos.

# Categorías de IA

- Inteligencia Artificial débil, Artificial Narrow Intelligence (ANI)
  - Resolución de problemas específicos
- Inteligencia Artificial fuerte / general, Artificial General Intelligence (AGI)
  - Resolver diferentes problemas y eventualmente aprender a resolver nuevos problemas.
- Super Inteligencia Artificial, Artificial Super Intelligence (ASI)
  - Desarrollar comportamiento / conciencia similar a la humana en lugar de simularla

# Definiciones de IA



Pensamiento

→ (Lógica) razonamiento para llegar a conclusiones.

Rusell & Norvig, 2004

Comportamiento

→ (Acción) acciones para llegar a objetivos.

# Definiciones de IA

Sistemas que piensan como humanos

“La interesante tarea de lograr que las computadoras piensen... *Máquinas con mente*, en su amplio sentido literal” (Haugeland, 1985)

“La automatización de] actividades que vinculamos con procesos de pensamiento humano, actividades tales como la toma de decisiones, resolución de problemas, aprendizaje...” (Bellman, 1978)

Sistemas que actúan como humanos

“El arte de crear máquinas con capacidad de realizar funciones que realizadas por personas requieren inteligencia” (Kurzweil, 1990)

“El estudio de cómo lograr que las computadores realicen tareas que, por el momento, los humanos hacen mejor” (Rich y Knight, 1991)

“El estudio de las facultades mentales mediante el uso de modelos computacionales” (Carniak y McDermott, 1985)

“El estudio de los cálculos que permiten percibir, razonar y actuar” (Winston, 1992)

“Un campo de estudio que se enfoca a la explicación y emulación de la conducta inteligente en función de procesos computacionales” (Schalkoff, 1990).

“La rama de la ciencia de la computación que se ocupa de la automatización de la conducta inteligente” (Luger y Stubblefield, 1993).

Sistemas que piensan racionalmente (idealmente)

Sistemas que actúan racionalmente (idealmente)

# Paradigmas de la IA

- Paradigma Simbólico
- Paradigma Subsimbólico o conexionista
- Abordajes híbridos.

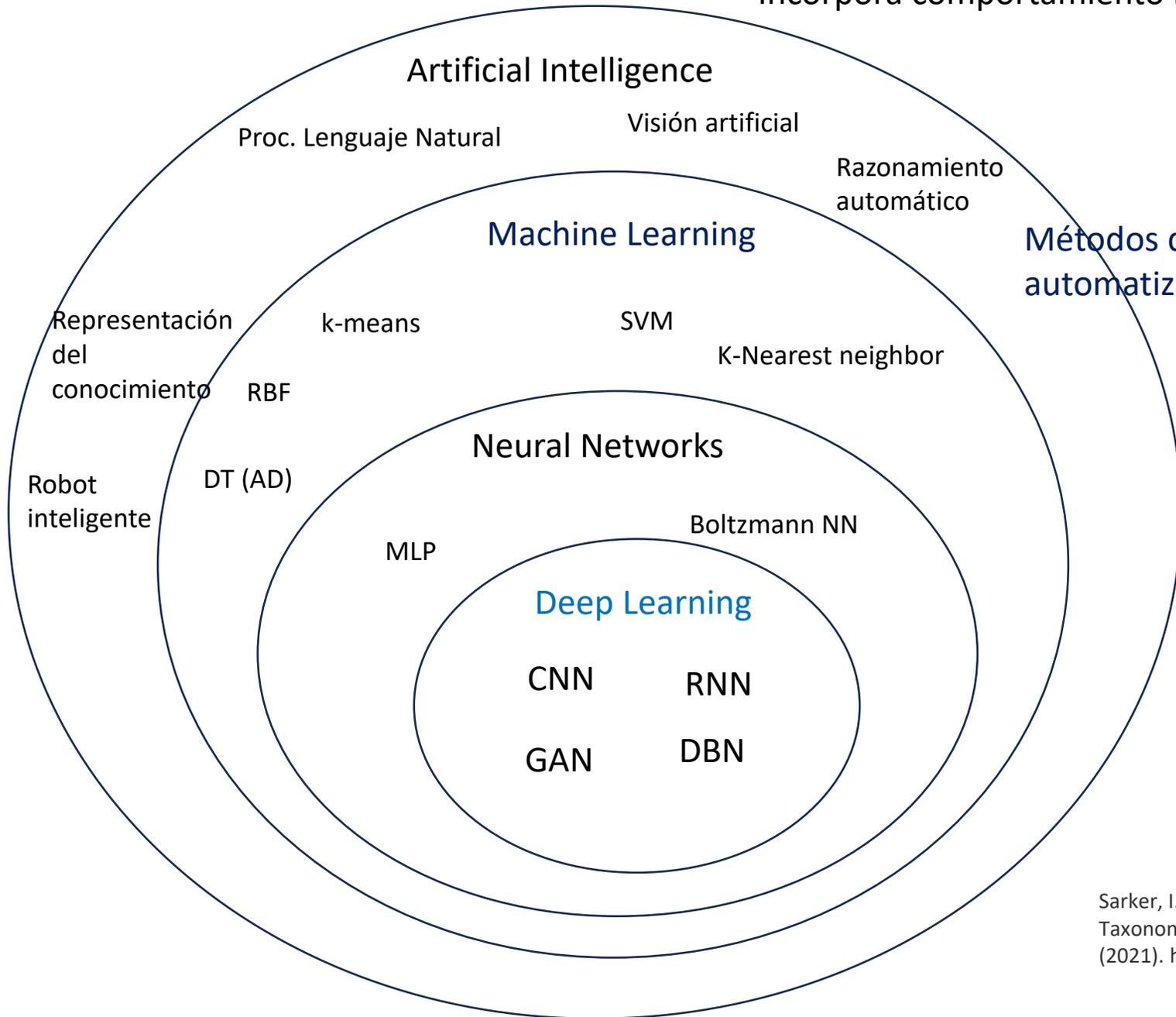
# Analogías de la naturaleza y de la IA

- Neuronas → Red Neuronal
- Evolución Natural → Computación evolutiva,  
Algoritmo genético
- Experiencia → Sistema Experto
- Razonamiento → Lógica Difusa
- Enfriamiento de metales → Recocido Simulado
- Hormigas → Colonia de Hormigas
- bandada de pájaros o .  
movimiento de bancos de peces → Optimización de enjambre de  
partículas (PSO)

# Estrategias diversas / definiciones

- Inteligencia Artificial y sus tecnologías como alternativas de análisis predictivos y prescriptivos
- Aprendizaje automático (AA/ML): aplicación de algoritmos y heurísticas para convertir muestras de datos en información capaz de generalizar comportamientos e inferencias para un conjunto más amplio de datos. Algoritmos estadísticos y de la computación
- Ciencia de Datos: campo interdisciplinario, involucra métodos científicos, procesos y sistemas para extraer conocimiento, analizar y aportar a la toma de decisiones
- Deep Learning, comprendidos en el AA. Se basa en RNA. DL, se refiere al uso de múltiples capas en la red. Los métodos utilizados pueden ser supervisados, semisupervisados o no supervisados.
- Simulación, modelos para experimentar ante distintas situaciones, diversas técnicas
- Business Intelligence (BI), inteligencia de negocios, ofrece ventajas competitivas, permite conocer, con alta precisión, su estado actual
- Business Analytics (BA), analítica de negocios, construye inferencias de futuros comportamientos del mercado y llevar a cabo acciones proactivas en función de análisis predictivos y prescriptivos
- Big Data, se explica a través de las V : volumen, velocidad, variedad y veracidad. // + valor
- Explotación de la información: sub-disciplina Informática, aporta a la Inteligencia de Negocio las herramientas para transformar información en conocimiento. Trata la búsqueda de patrones y de regularidades en datos
- Otras ??

Incorpora comportamiento humano e inteligencia a maquinas o sistemas



Métodos que aprenden de datos o experiencia, automatización en la construcción de modelos

Computación utilizando RN multi-layer

# Conocimiento y Aprendizaje en IA

# Conocimiento en IA

Conocimiento tácito, implícito, explícito

- Tácito, permanece en un nivel "inconsciente", desarticulado. Se implementa y ejecuta de manera mecánica.
- Implícito, se sabe que se dispone y se utiliza habitualmente
- Explícito, se sabe que se dispone, se comparte. Se transmiten de manera estructurada.

Conocimiento declarativo, procedimental, heurístico

- Procedimental o interpretativo, es el tipo de conocimiento en el que se aclara cómo hacer algo para resolver un problema dado.
- Declarativo o descriptivo, es el tipo de conocimiento que cuenta el conocimiento básico sobre algo. Enfatiza qué hacer para resolver un problema dado.
- Heurístico. Tipo de conocimiento especial para resolver problemas complejos. Consiste en estrategias, criterios, métodos o procesos

# Aprendizaje en IA

Simon propuso que el aprendizaje implica:

*...cambios en el sistema, los que se adaptan para permitir llevar a cabo la misma tarea o tareas a partir de las mismas condiciones de un modo más eficiente y eficaz cada vez.*

El aprendizaje responde a diversos fenómenos.

- Perfeccionar una habilidad.
- Adquirir conocimiento.

# Modalidades de Aprendizaje

Siguiendo a Rusell y Norvig (2004):

- Aprendizaje memorístico
- Aprendizaje mediante consejos
- Aprendizaje en la resolución de problemas
  - Aprendizaje por medio de ajustes de parámetros
  - Aprendizaje con macro-operadores
  - Aprendizaje mediante troceado
- Aprendizaje a partir de ejemplos: inducción
- Aprendizaje basado en explicaciones
- Descubrimiento
- Aprendizaje basado en RNA y AG

# Modalidades de aprendizaje

- Aprendizaje supervisado
- Aprendizaje no supervisado
- Aprendizaje semi-supervisado
- Aprendizaje por refuerzo

# Aprendizaje Supervisado

Algunos algoritmos más utilizados en ML

- Árboles de decisión.
- Clasificación de Naïve Bayes.
- Regresión por mínimos cuadrados.
- Regresión Logística.
- Support Vector Machines (SVM).
- Métodos “Ensemble” (Conjuntos de clasificadores).

Se disponen de datos “etiquetados” y sus correspondientes valores de salida

El algoritmo aprende a través de un entrenamiento con un conjunto de datos históricos / conocido. En procesos posteriores puede predecir o clasificar para proponer solución a un problema

# Aprendizaje no Supervisado

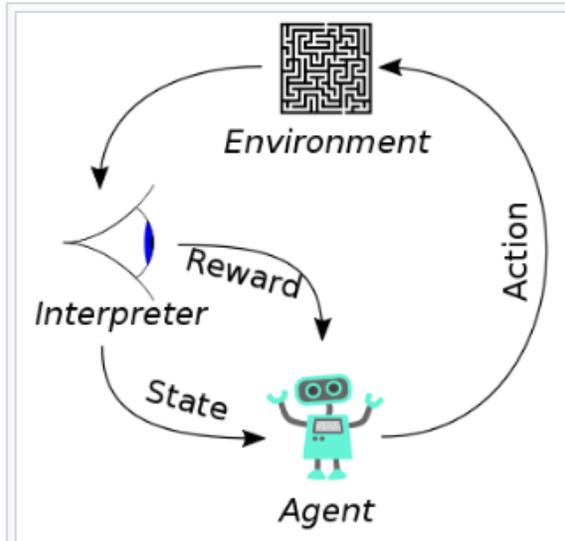
- Clustering
- Reducción de dimensionalidad.

Se carece de datos “etiquetados” para el entrenamiento y se desconoce los datos de salida correspondientes a cada instancia o input.  
Algoritmos exploratorios,  
Finalidad: agrupar los datos con similitudes

# Aprendizaje semi-supervisado

- Combina algoritmos de AS para etiquetar puntos de datos con etiquetas conocidas, y algoritmos de ANS para agrupar puntos de datos.
- Se aplica a problemas costosos en tiempo o procesamiento.
- Ej. Algoritmo: Deep Belief Networks (DBN) – o Redes de Creencia Profunda-, compuestas de redes simples denominadas Restricted Boltzmann Machines (RBS) [entrenamiento no supervisado de manera secuencial], y se continua con entrenamiento supervisado.
- Ej. Etiquetar algunas personas en fotos y aplicar procesos de reconocimiento a posteriori

# Aprendizaje por refuerzo



The typical framing of a Reinforcement Learning (RL) scenario: an agent takes actions in an environment, which is interpreted into a reward and a representation of the state, which are fed back into the agent.

Aprendizaje que mejora la respuesta del modelo usando un proceso de retroalimentación.

Aprendizaje a partir de la observación del mundo en que interviene. Su información de entrada es el feedback o retroalimentación que obtiene del mundo exterior como respuesta a sus acciones.

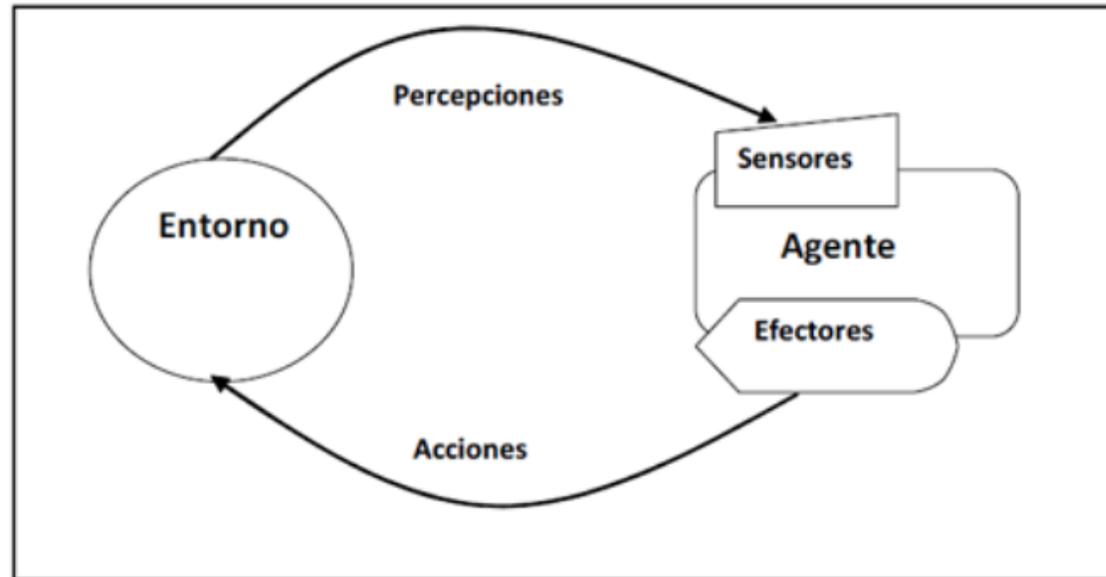
El aprendizaje se basa en ensayo-error.

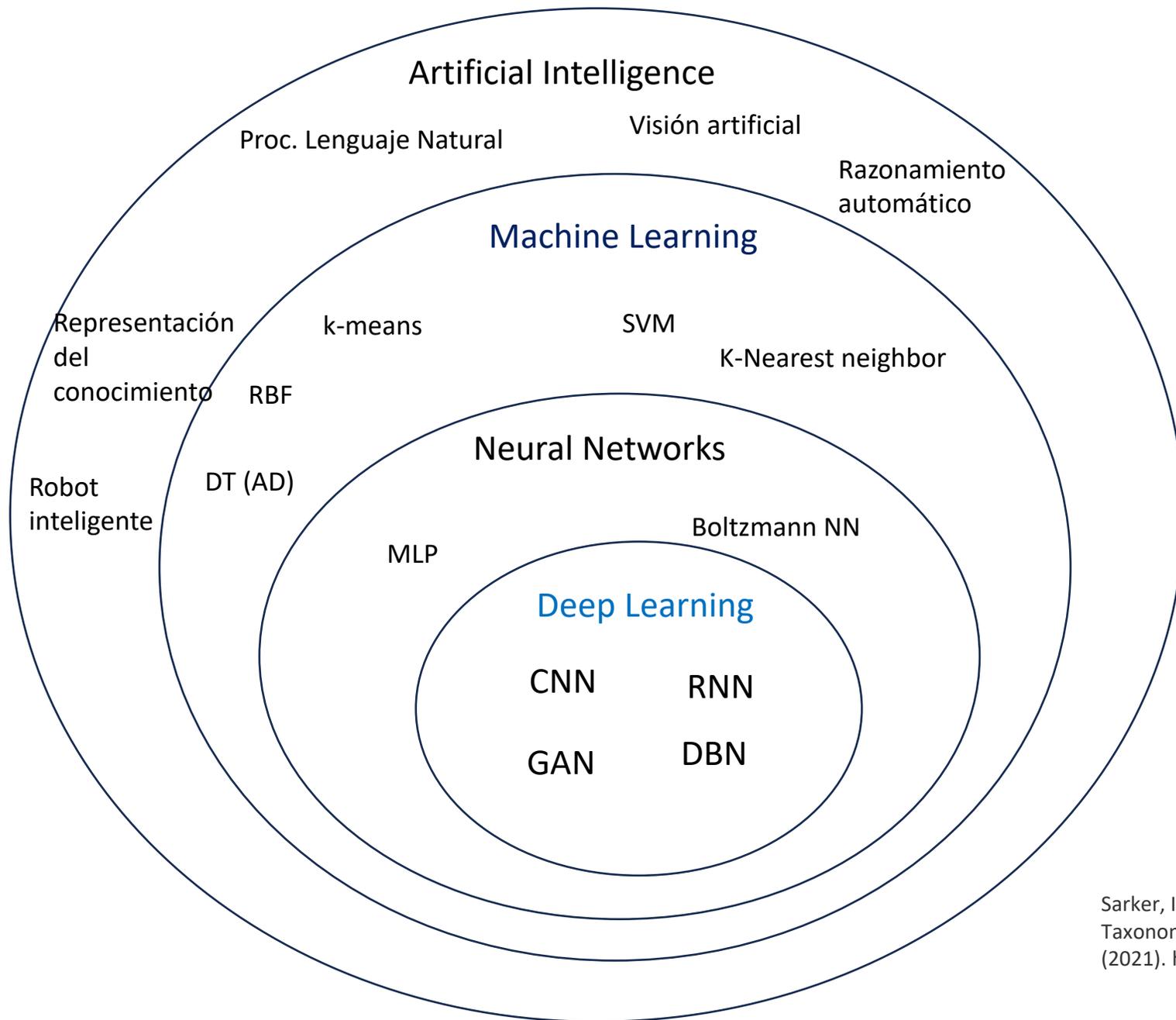
Puede ser entendido como AS

Algoritmos: Q-learning, Deep Q-learning

Aplicaciones: entretenimiento, salud, otros

# Agentes y sistemas inteligentes





- Multilayer Perceptrons (MLPs)
- Convolutional Neural Networks (CNNs)
- Recurrent Neural Networks (RNNs)
- Long Short Term Memory Networks (LSTMs)
- Generative Adversarial Networks (GANs)
- Radial Basis Function Networks (RBFNs)
- Self Organizing Maps (SOMs)
- Deep Belief Networks (DBNs)
- Restricted Boltzmann Machines( RBMs)
- Autoencoders
- LLM (Large Language Model)
- Otros

# Otros terminos relacionados

En RN / NN

El modelo discriminativo

**Enfoque ascendente** en el que los datos fluyen desde la capa de entrada a través de las capas ocultas hasta la capa de salida. Se utilizan en el entrenamiento supervisado para problemas como la clasificación y la regresión.

El modelo generativo,

**Enfoque descendente**. Se utilizan en problemas de distribución probabilística y de pre-entrenamiento no supervisado.

Dada la entrada  $x$  y la etiqueta  $y$ :

- modelo discriminativo aprende la distribución de probabilidad  $p(y|x)$ .
- modelo generativo aprende la probabilidad conjunta de  $p(x,y)$ , a partir de la cual  $P(y|x)$  se puede predecir [20].

En general:

Con datos etiquetados -> enfoques discriminativos, entrenamiento efectivo,

Sin datos etiquetados -> adoptar enfoque generativo

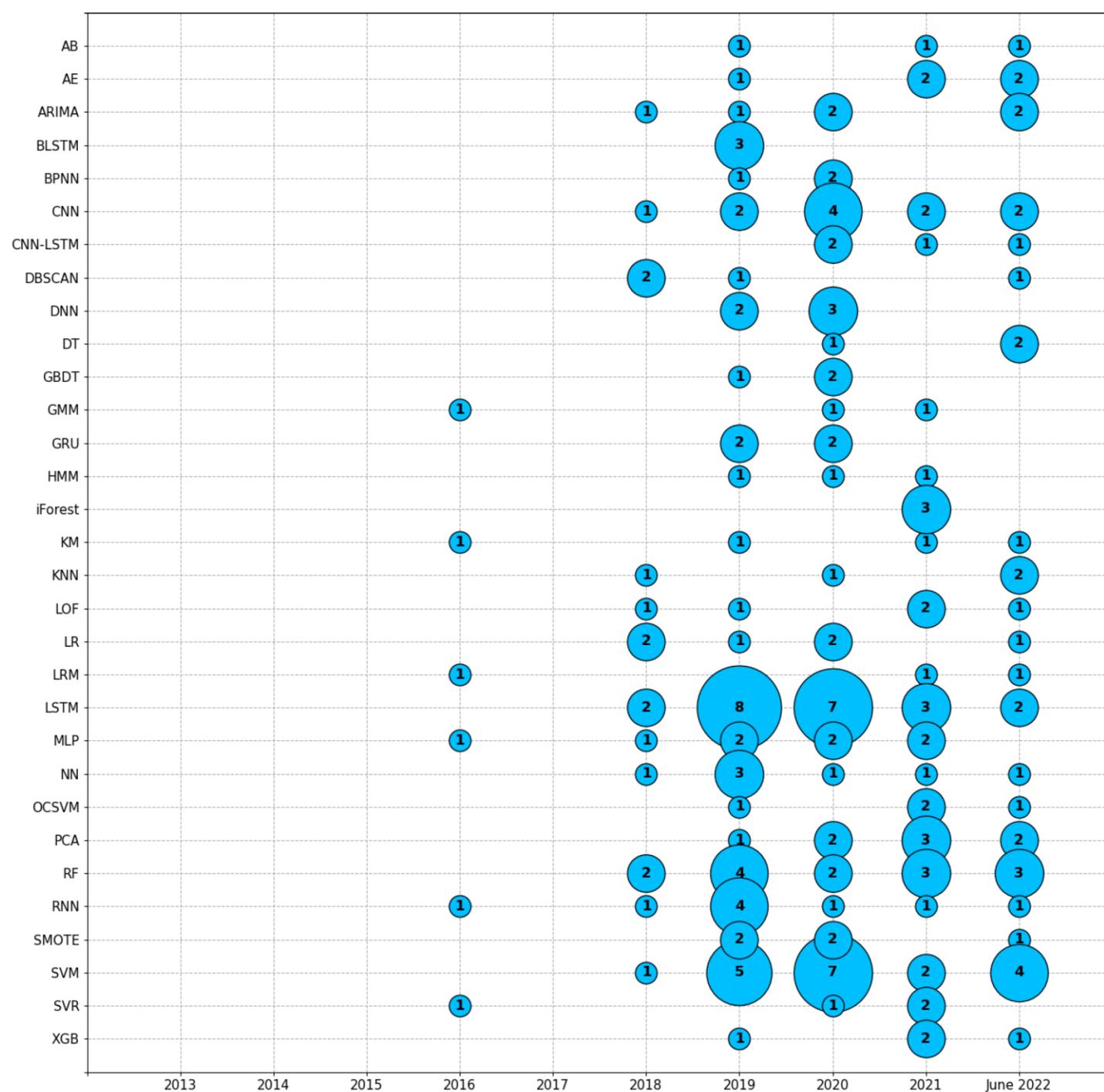
**Figure 5** shows how many times a data science method appeared over the years of publication. Long short-term memory (LSTM) networks were the method that most appeared in the corpus, with 22 occurrences. Then, support vector machine (SVM) had 19 occurrences. Next, the random forest (RF) method appeared 14 times. The years 2019, 2020, and 2021 presented the highest concentration of data science methods.

Fuente datos

ACM, IEEE, Scopus, Springer, and Wiley

: Arruda, H.M.; Bavaresco, R.S.; Kunst, R.; Bugs, E.F.; Pesenti, G.C.; Barbosa, J.L.V. Data Science Methods and Tools for Industry 4.0: A Systematic Literature Review and Taxonomy. *Sensors* 2023, 23, 5010.

<https://doi.org/10.3390/s23115010>



# BloombergGPT: A Large Language Model for Finance

Shijie Wu<sup>1,\*</sup>, Ozan İrsoy<sup>1,\*</sup>, Steven Lu<sup>1,\*</sup>, Vadim Dabravolski<sup>1</sup>, Mark Dredze<sup>1,3</sup>, Sebastian Gehrmann<sup>1</sup>, Prabhanjan Kambadur<sup>1</sup>, David Rosenberg<sup>2</sup>, Gideon Mann<sup>1</sup>

<sup>1</sup> Bloomberg, New York, NY USA

<sup>2</sup> Bloomberg, Toronto, ON Canada

<sup>3</sup> Computer Science, Johns Hopkins University, Baltimore, MD USA

## Abstract

The use of NLP in the realm of financial technology is broad and complex, with applications ranging from sentiment analysis and named entity recognition to question answering. Large Language Models (LLMs) have been shown to be effective on a variety of tasks; however, no LLM specialized for the financial domain has been reported in literature. In this work, we present BLOOMBERGGPT, a 50 billion parameter language model that is trained on a wide range of financial data. We construct a 363 billion token dataset based on Bloomberg’s extensive data sources, perhaps the largest domain-specific dataset yet, augmented with 345 billion tokens from general purpose datasets. We validate BLOOMBERGGPT on standard LLM benchmarks, open financial benchmarks, and a suite of internal benchmarks that most accurately reflect our intended usage. Our mixed dataset training leads to a model that outperforms existing models on financial tasks by significant margins without sacrificing performance on general LLM benchmarks. Additionally, we explain our modeling choices, training process, and evaluation methodology. We release Training Chronicles (Appendix C) detailing our experience in training BLOOMBERGGPT.

[2303.17564.pdf \(arxiv.org\)](https://arxiv.org/pdf/2303.17564.pdf)

09/05/2023

TABLE 2. DNN network comparison table.

Network Type	Architecture	Network Model	Training Type	Training Algorithm	Implementation Sample	Common Application	Popular Dataset Sample	DL Framework (sample)
Feedforward Neural Network	CNN	Discriminative	Supervised	Gradient Descent based Backpropagation	Siamese Network, Deep CNN	Image recognition/classification	MNIST	TensorFlow, Caffe, Theano, Torch, Deeplearning4j, Microsoft Cognitive Toolkit, Keras, MXNet, PyTorch
	Residual Network	Discriminative	Supervised	Gradient Descent based Backpropagation	Deep ResNet; HighwayNet; DenseNet	Image recognition	ImageNet	TensorFlow, PyTorch, Keras
	Autoencoder	Generative	Unsupervised	Backpropagation	Sparse Autoencoders, Variational Autoencoders	Dimensionality Reduction; Encoding	MNIST	TensorFlow, Deeplearning4j, Keras
	Adversarial Networks	Generative & Discriminative	Unsupervised	Backpropagation	Generative Adversarial Network	Generate realistic fake data; Reconstruction of 3D models; Image improvement	CIFAR10	TensorFlow, Keras
	RBM	Generative with Discriminative finetuning	Unsupervised	Gradient Descent based Contrastive divergence	Deep Belief Network; Deep Boltzmann Machine	Dimensionality Reduction; Feature learning; Topic modeling	MNIST	TensorFlow, Deeplearning4j, Keras, MXNet, Theano, Torch
Recurrent Neural Network	LSTM	Discriminative	Supervised	Gradient Descent & Backpropagation through Time	Deep RNN, Gated Recurrent Unit (GRU), Neural Machine Translation (NMT)	Natural Language Processing; Language Translation	MNIST Stroke Sequence	TensorFlow, Caffe, Theano, Torch, Deeplearning4j, Microsoft Cognitive Toolkit, Keras, MXNet, PyTorch
Radial Basis Function NN	RBF Network	Discriminative	Supervised and Unsupervised	K-means Clustering; Least Square Function	Radial Basis Function NN	Function approximation ; Time series prediction	Fisher's Iris data set	TensorFlow
Kohonen Self Organizing NN	Nodes arranged in hexagonal or rectangular grid	Generative	Unsupervised	Competitive Learning	Kohonen Self Organizing NN	Dimensionality Reduction; Optimization problems; Clustering analysis	SPAMbase	TensorFlow

# Recursos

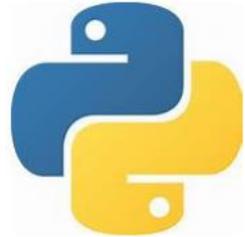
## Keras

Neural network library



## Python

Software



Python is a high-level, interpreted, general-purpose programming language. Its de... +



<https://scikit-learn.org/stable/index.html>

## TensorFlow



TensorFlow is a free and open-source software library for machine learning and artificial intelligence. It can be used across a range of tasks but has a particular focus on training and inference of d...

## R



R is a programming language for statistical computing and graphics supported by the R Core Team and the R Foundation for Statistical Computing. Created by statisticians Ross Ihaka and Robert Gentleman...



## PyTorch

Machine learning framework

<https://pytorch.org>

# Librerías de Python



NumPy  
(2005 -  
Presente)



matplotlib  
(2003 –  
Presente)



SciPy  
(2001 -  
Presente)



scikit-learn  
(2007 -  
Presente)



pandas  
(2008 - Presente)



TensorFlow  
(2015 - Presente)



PyTorch  
(2016 - Presente):



Django  
(2005 - Presente):



Flask  
(2010 - Presente):



FastAPI  
(2018 - Presente):

# Librerías de Python

NumPy (2005 - Presente):

- NumPy librería para la computación científica en Python. Introduce el objeto `numpy.ndarray`, una estructura de datos eficiente para el almacenamiento y manipulación de matrices n-dimensionales. Ha sido crucial para el desarrollo de otras librerías como SciPy, pandas y scikit-learn.

matplotlib (2003 - Presente):

- Matplotlib, librería de gráficos 2D que se utiliza para crear visualizaciones de datos. Ha evolucionado para incluir diversos estilos de gráficos y mejoras de rendimiento, además de interfaces para la creación interactiva de gráficos en entornos web y notebooks.

SciPy (2001 - Presente):

- SciPy se basa en NumPy. amplia variedad de algoritmos científicos, incluyendo optimización, álgebra lineal, estadísticas y procesamiento de señales. Su desarrollo ha llevado a mejoras en el rendimiento y la funcionalidad de muchas áreas de la ciencia computacional.

scikit-learn (2007 - Presente):

- scikit-learn es una librería AA. Proporciona herramientas para tareas como clasificación, regresión, clustering y selección de modelos. Mejora en número de algoritmos, la eficiencia y la usabilidad.

pandas (2008 - Presente):

- pandas es una librería de análisis de datos que ofrece estructuras de datos flexibles y eficientes, como DataFrames y Series, que simplifican la manipulación y análisis de datos tabulares. Ha continuado evolucionando con mejoras en el rendimiento y la capacidad de manejar conjuntos de datos aún más grandes.

# Librerías de Python

TensorFlow (2015 - Presente):

- TensorFlow, desarrollado por Google, es una librería de AA y DL. Ha evolucionado para incluir tanto API simbólicas (TensorFlow 1.x) como API basadas en funciones (TensorFlow 2.x), simplifica su uso y mejora la velocidad de desarrollo.

PyTorch (2016 - Presente):

- PyTorch, desarrollado por Facebook, librería AA / DL. enfoque en la programación dinámica e interfaz intuitiva

Django (2005 - Presente):

- Django, framework de desarrollo de aplicaciones web robustas y escalables. Con características para la administración de la base de datos, la autenticación y el soporte para aplicaciones RESTful.

Flask (2010 - Presente):

- Flask, framework de desarrollo web ligero y flexible. Presenta extensiones y complementos que ofrecen funcionalidades adicionales, como autenticación, manejo de formularios y API REST.

FastAPI (2018 - Presente):

- FastAPI, framework para la creación de APIs web. Fácil de usar, documentación automática y soporte para Python.

**TABLE 1. Popular deep learning frameworks and libraries.**

Framework	Institution	License	1st Release
Caffe	Berkeley Research	AI BSD / Free	2015
Microsoft Cognitive Toolkit	Microsoft	MIT License / Free	2016
Gluon	AWS and Microsoft	Open Source	2017
Keras	Individual Author	MIT License / Free	2015
MXNet	Apache Software Foundation	Apache 2.0 / Free	2015
TensorFlow	Google Brain	Apache 2.0 / Free	2015
Theano	University of Montreal	BSD / Free	2008
Torch	Ronan Collobert et al.	BSD / Free	2002
PyTorch	Facebook	BSD / Free	2016
Chainer	Preferred Networks	BSD / Free	2015
Deeplearning4j	Adam Gibson et al.	Apache 2.0 / Free	2014

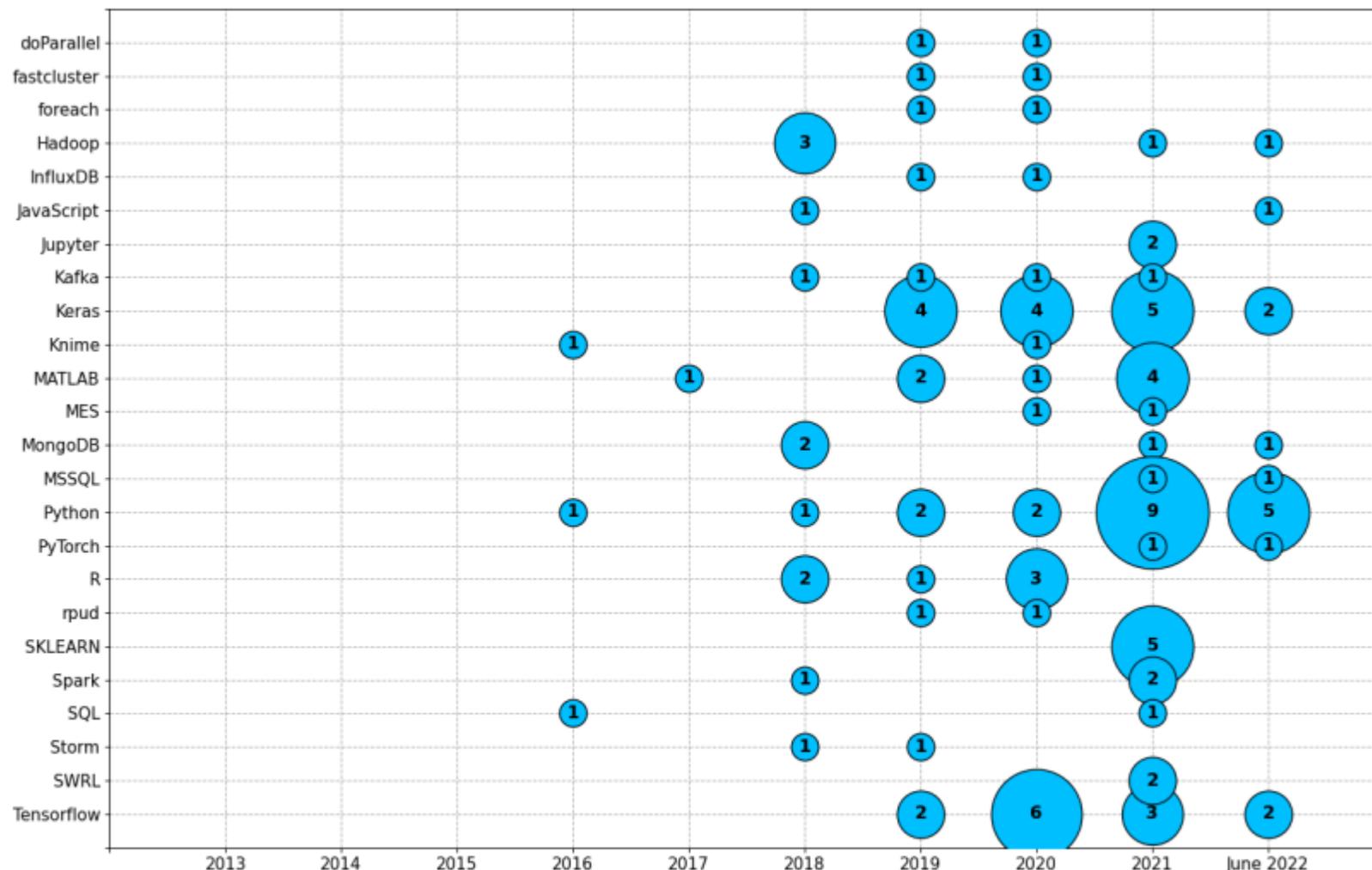
Licencia Pública General de GNU (GPL) es una licencia de software libre que se utiliza para proteger la libertad de los usuarios finales de software.

BSD: licencia Berkeley

Apache / MIT  
Licencias permisiva,s NO o exige que las obras derivadas (versiones modificadas) del software se distribuyan usando la misma licencia

Open Source, otorga a usuario acuerdo legal que define los términos y condiciones bajo los cuales el software puede ser utilizado, modificado y distribuido.

**Figure 6** shows the software tools grouped by years. The Python programming language was the most used tool, appearing in 20 papers, followed by Keras, which appeared in 15 papers, and Tensorflow which appeared in 13 articles.



Citation: Arruda, H.M.; Bavaresco, R.S.; Kunst, R.; Bugs, E.F.; Pesenti, G.C.; Barbosa, J.L.V. Data Science Methods and Tools for Industry 4.0: A Systematic Literature Review and Taxonomy. *Sensors* 2023, 23, 5010.

<https://doi.org/10.3390/s23115010>

**Figure 6.** Software tools grouped by year. The definition of each tool is in Table A3. Python was the tool with the most occurrences (20), followed by Keras (15), and Tensorflow (13). For a better visualization, only tools with more than one occurrence appear in the picture.

Fuente datos  
ACM, IEEE, Scopus, Springer, and Wiley

# Recursos. Datasets

**Registry of Open Data on AWS** 

Open Data on AWS is now available on AWS Data Exchange. The Registry of Open Data are now discoverable on AWS Data Exchange alongside 3,000+ existing data products from category-leading data providers. Explore the catalog to find open, free, and commercial data sets. [Learn more about AWS Data Exchange](#)

[Explore the catalog](#)



[+ Create](#)

[Home](#)

[Competitions](#)

[Datasets](#)

[Models](#)

Search

## Datasets

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

[+ New Dataset](#)

## [Registry of Open Data on AWS](#)

## [Find Open Datasets and Machine Learning Projects | Kaggle](#)

 [Datasets](#) [Contribute Dataset](#) [About Us](#)  [Log](#)

**Welcome to the UC Irvine Machine Learning Repository**

We currently maintain 624 datasets as a service to the machine learning community. Here, you can donate and find datasets used by millions of people all around the world!

[VIEW DATASETS](#) [CONTRIBUTE A DATASET](#)

Popular Datasets		New Datasets							
	<b>Iris</b> A small classic dataset from Fisher, 1936. One of the earliest datasets used for evaluation...		<b>Sepsis Survival Minimal Clinical Records</b> This dataset collection contains minimal health records of 110,204 admissions (primary c...						
	Classification		150 Instances		Classification		110.34K Instances		3 Attributes

## [Home - UCI Machine Learning Repository](#)

# Recursos. Datasets

## [Portal de Datos Abiertos \(ciudaddecorrientes.gov.ar\)](http://ciudaddecorrientes.gov.ar)

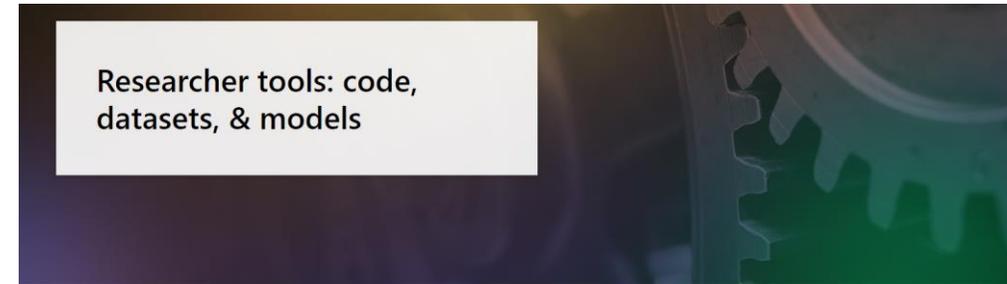
Municipalidad de **CORRIENTES** Datasets Organizaciones Acerca ▾

### Portal de Datos Abiertos

El Portal de Datos Abiertos es una iniciativa ligada a las políticas de Gobierno Abierto que sirve para almacenar y compartir las bases de datos producidas por el Municipio para facilitar el conocimiento del gobierno, fortalecer el rendimiento de cuentas y mejorar la participación ciudadana.

055 DATASETS  
012 ORGANIZACIONES CON DATOS

## [Researcher tools: code, datasets, & models - Microsoft Research](#)



An index of datasets, SDKs, APIs and other open source code created by Microsoft researchers and shared with the broader academic community. We also maintain [a collection](#) highlighting some of the tools you'll find here.

## [Datos Argentina](#)

**datos.gov.ar** Datasets Series Organizaciones APIs Acerca ▾

Ponemos a tu alcance datos públicos en formatos abiertos para que puedas usarlos, modificarlos y compartirlos. Estos datos son tuyos. Podés crear visualizaciones, aplicaciones y grandes herramientas con ellos.

036 DATASETS  
ORGANIZACIONES CON DATOS

- Agroganadería, pesca y forestación
- Asuntos internacionales
- Ciencia y tecnología
- Economía y finanzas
- Educación, cultura y deportes
- Energía
- Gobierno y sector público
- Justicia, seguridad y legales
- Medio ambiente
- Población y sociedad
- Regiones y ciudades
- Salud
- Transporte

# Práctica.



## Objetivo

- Adquirir o profundizar conocimientos concernientes a métodos de ML / AA

## Actividades

1. Constituir equipos de 3 – 4 integrantes

2. A partir de la lectura del artículo: <https://doi.org/10.1007/s42979-021-00815-1>, responder

- Establecer la relación entre los conceptos de IA, ML, DL
- ¿Se mencionan los tipos de datos manipulables por los modelos de DL?. Se asocian a algunos en particular? ¿ Que propiedades clave y dependencias en referencia a las técnicas de DL, deben contemplarse al diseñar un proyecto de DL aplicable a resolución de problemas reales?
- ¿Se mencionan términos / elementos tratados en los modelos MLP? ¿Cuáles?
- Entre las tecnologías de DL, se identifica: ¿qué tipo de aprendizaje corresponde?, ¿el algoritmo de aprendizaje que implementa? ¿el año en que se presentan o difunden? ¿la sigla asociada a cada tecnología?
- ¿Se menciona el software para el tratamiento de los datos según las distintas tecnologías resumidas en el artículo?
- ¿Se mencionan aplicaciones de las distintas tecnologías resumidas en el artículo? Conoce otras mencionadas en los artículos que indagó
- Otros aspectos que considera relevante mencionar...

3. Exponer

# Metodologías

El ciclo de vida de los sistemas de IA, comprende el estudio y aplicación de metodologías específicas según sea el paradigma en que se inscribe la propuesta de resolución del problema

- Metodologías de SBC
- Metodologías en ML / CD
- Otras, según las tecnologías elegidas

# Metodologías de minería de datos(MD)

Algunas metodologías de MD

- KDD
- *CRISP-DM*
- *SEMMA*

# Metodologías de minería de datos (MD)

## KDD

- Desarrollada en 1989, por Gregory Piatetsky-Shapiro
- Descubre patrones al procesar datos sin procesar, analizar la información en busca de datos necesarios e interpretar los resultados.

**DATOS**



# CRISP-DM - Cross Industry Standard Process for Data Mining

- Perspectiva de **metodología**, descripciones de las fases de un proyecto, las tareas en cada fase y explica las relaciones entre las tareas.
- Perspectiva de **modelo de proceso**, resumen del ciclo de minería de datos.



# CRISP-DM - Cross Industry Standard Process for Data Mining

## 1. Business Understanding

- Determine Business Objectives
- Assess Situation
- Determine Data Science Goals
- Produce Project Plan

## 2. Data Understanding

- Collect Initial Data
- Describe Data
- Explore Data
- Verify Data Quality

## 3. Data Preparation

- Select Data
- Clean Data
- Construct Data
- Integrate Data
- Format Data

## 4. Modeling

- Select Modeling Technique
- Generate Test Design
- Build Model
- Assess Model

## 5. Evaluation

- Evaluate Results
- Review Process
- Determine Next Steps

## 6. Deployment

- Plan Deployment
- Plan Monitoring & Maintenance
- Produce Final Report
- Review Project

## Abstract

CRISP-DM is the de-facto standard and an industry-independent process model for applying data mining projects. Twenty years after its release in 2000, we would like to provide a systematic literature review of recent studies published in IEEE, ScienceDirect and ACM about data mining use cases applying CRISP-DM. We give an overview of the research focus, current methodologies, best practices and possible gaps in conducting the six phases of CRISP-DM. The main findings are that CRISP-DM is still a de-facto standard in data mining, but there are challenges since the most studies do not foresee a deployment phase. The contribution of our paper is to identify best practices and process phases in which data mining analysts can be better supported. Further contribution is a template for structuring and releasing CRISP-DM studies.

© 2021 The Authors. Published by Elsevier B.V.

This is an open access article under the CC BY-NC-ND license (<https://creativecommons.org/licenses/by-nc-nd/4.0>)

Peer-review under responsibility of the scientific committee of the CENTERIS - International Conference on ENTERprise Information Systems / ProjMAN - International Conference on Project MANagement / HCist - International Conference on Health and Social Care Information Systems and Technologies 2020

*Keywords:* CRISP-DM; Literature Review; Data Mining; Process Methodology; Deployment

Table 1: CRISP-DM process model descriptions [10].

Phase	Short description
Business Understanding	The business situation should be assessed to get an overview of the available and required resources. The determination of the data mining goal is one of the most important aspect in this phase. First the data mining type should be explained (e. g. classification) and the data mining success criteria (like precision). A compulsory project plan should be created.
Data understanding	Collecting data from data sources, exploring and describing it and checking the data quality are essential tasks in this phase. To make it more concrete, the user guide describe the data description task with using statistical analysis and determining attributes and their collations.
Data preparation	Data selection should be conducted by defining inclusion and exclusion criteria. Bad data quality can be handled by cleaning data. Dependent on the used model (defined in the first phase) derived attributes have to be constructed. For all these steps different methods are possible and are model dependent.
Modeling	The data modelling phase consists of selecting the modeling technique, building the test case and the model. All data mining techniques can be used. In general, the choice is depending on the business problem and the data. More important is, how to explain the choice. For building the model, specific parameters have to be set. For assessing the model it is appropriate to evaluate the model against evaluation criteria and select the best ones.
Evaluation	In the evaluation phase the results are checked against the defined business objectives. Therefore, the results have to be interpreted and further actions have to be defined. Another point is, that the process should be reviewed in general.
Deployment	The deployment phase is described generally in the user guide. It could be a final report or a software component. The user guide describes that the deployment phase consists of planning the deployment, monitoring and maintenance.

# CRISP-DM Twenty Years Later: From Data Mining Processes to Data Science Trajectories

Fernando Martínez-Plumed, Lidia Contreras-Ochando, Cèsar Ferri, José Hernández-Orallo, Meelis Kull,  
Nicolas Lachiche, María José Ramírez-Quintana and Peter Flach

**Abstract**—CRISP-DM (CRoss-Industry Standard Process for Data Mining) has its origins in the second half of the nineties and is thus about two decades old. According to many surveys and user polls it is still the *de facto* standard for developing data mining and knowledge discovery projects. However, undoubtedly the field has moved on considerably in twenty years, with *data science* now the leading term being favoured over *data mining*. In this paper we investigate whether, and in what contexts, CRISP-DM is still fit for purpose for data science projects. We argue that if the project is goal-directed and process-driven the process model view still largely holds. On the other hand, when data science projects become more exploratory the paths that the project can take become more varied, and a more flexible model is called for. We suggest what the outlines of such a trajectory-based model might look like and how it can be used to categorise data science projects (goal-directed, exploratory or data management). We examine seven real-life exemplars where exploratory activities play an important role and compare them against 51 use cases extracted from the NIST Big Data Public Working Group. We anticipate this categorisation can help project planning in terms of time and cost characteristics.

**Index Terms**—Data Science Trajectories, Data Mining, Knowledge Discovery Process, Data-driven Methodologies.



# Metodología CRIP-DM. Ejemplo

Esta propuesta permite integrar una herramienta GIS dentro de la etapa CRISP-DM “Evaluación”. Una vez obtenidos los datos explotados resultantes de la aplicación de los algoritmos de minería de datos, éstos se visualizan a través de un GIS para lograr un análisis más profundo con la ayuda del componente espacial. A continuación, se observa en la Fig. 1, la integración de tecnología GIS en la etapa Evaluación de la metodología de explotación de información CRISP-DM.

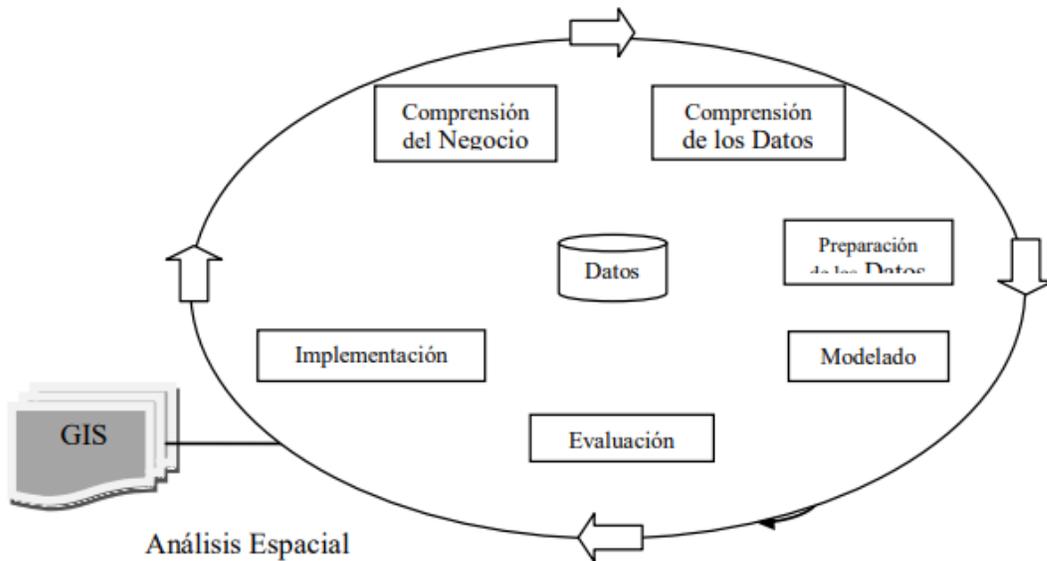


Fig. 1. Integración de tecnología GIS en CRISP-DM (Fuente propia)

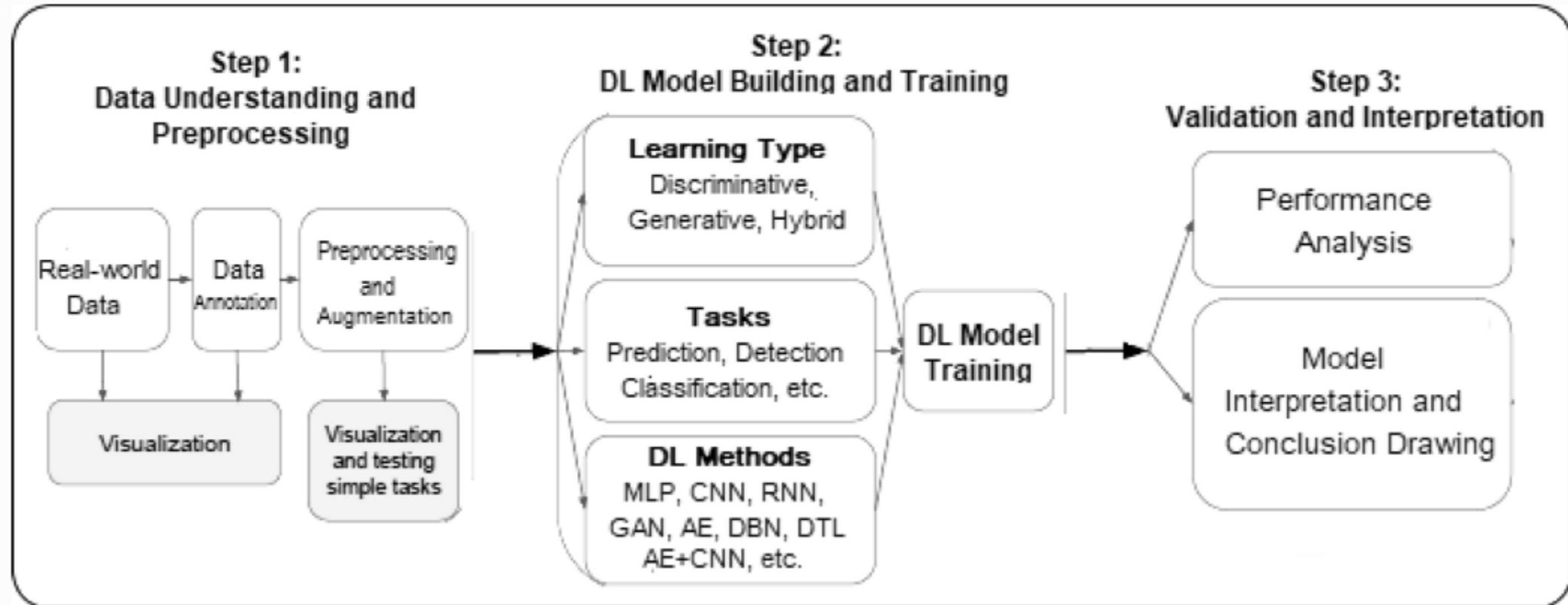
# SEMMA

- Sample, Explore, Modify, Model, and Assess.
- Es una lista de pasos secuenciales desarrollados por SAS Instituto.
- Guía la implementación de aplicaciones de minería de datos.

Sample' (muestreo)	Explore' (exploración)	Modify' (modificación):	Model' (modelado)	Assess (evaluación)
captura de datos, y eventualmente creación de tablas u otras estructuras que los contengan, Considerar el tamaño de las muestras para disponer de un conjunto de datos representativo y suficientemente para el procesamiento.	entender los datos disponibles, se aplica a través de la visualización, el agrupamiento, etc para observar las relaciones, tendencias y otra información proporcione conocimiento sobre los datos y el fenómeno subyacente.	trabajo sobre los datos para su uso en el posterior modelado. Se transforman y seleccionan o crean variables a partir de los datos , otras	determina el modelo más adecuado (el que mejor predice la o las variables de salida a partir de la o las variables de entrada), seleccionando entre las familias disponibles (redes neuronales, árboles de decisión, regresión logística, etc) y afinando el modelo en la o las opciones seleccionadas	evaluar el funcionamiento del modelo en su conjunto en cuanto a fiabilidad, utilidad etc

Fig. 4

From: [Deep Learning: A Comprehensive Overview on Techniques, Taxonomy, Applications and Research Directions](#)

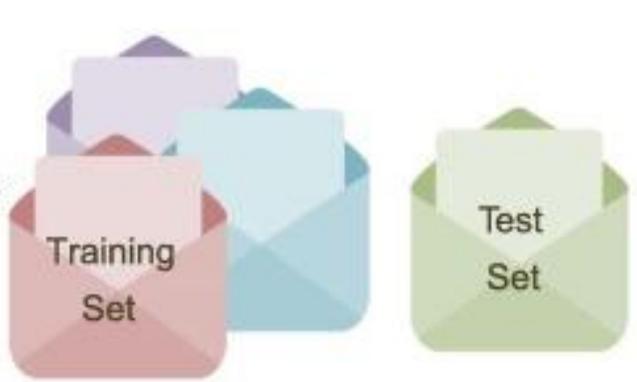


A typical DL workflow to solve real-world problems, which consists of three sequential stages (i) data understanding and preprocessing (ii) DL model building and training (iii) validation and interpretation

**(a) Dataset: Enron**



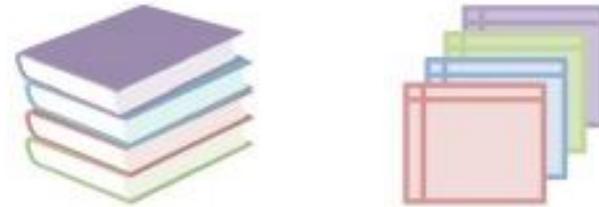
**(b) Allocate train set and test set**



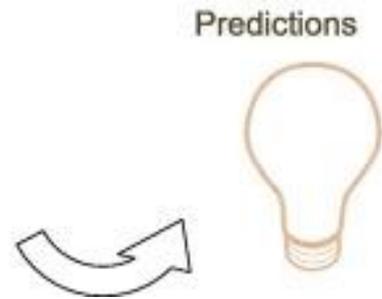
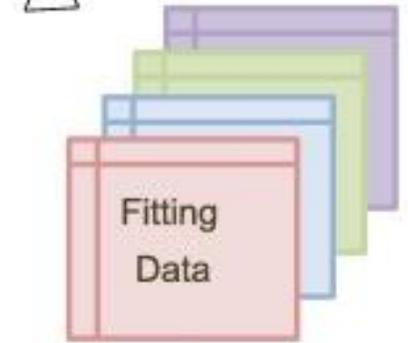
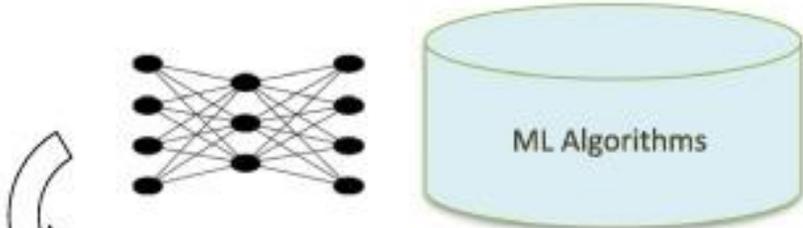
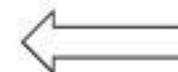
**(c) Preprocessing: Stop words, HTML Tags, Lemmatisation**



**(d) Feature extraction. Building dictionary and generating matrix**



**(e) Applying the 12 ML models for fitting the data and generating predictions**



**(f) Statistical analysis and results interpretation**



We proposed and tested a pipeline to compare and explain the classification outcomes of **12 machine learning models**. We applied the pipeline for optimising and testing the models in a spam filtering context, with lemmatisation and noise-reduction techniques as preprocessing steps. The pipeline, which we make publicly available, was developed to compare the performance of the classifiers in terms of precision, recall, F-score, and ROC curves.

MEDIA > ESTUDIANTES > SECTORES Y PROFESIONES ESTUDIANTES

## Inteligencia Artificial (IA), los trabajos del futuro

07 ene 2020 5 min



**Aparición de nuevos perfiles profesionales** que deben convivir con los ya existentes. Estos perfiles especializados abren un nuevo abanico de posibilidades innovadoras que afectan directamente a procesos y herramientas, que ven **mejorado su rendimiento**. *NLP Engineer, Image Specialist, Audio Specialist, Tools expert y Cognitive Trainer* son algunos de los nuevos puestos que ya se incluyen en los equipos de trabajo.

[¿Cómo impacta la Inteligencia Artificial en las organizaciones? \(deloitte.com\)](#)



### Machine learning engineer: un experto en resolución de problemas

Ahmed Ahres, machine learning engineer en Revolut, nos explica en qué consiste su trabajo y cómo fue su camino hasta...

09 sept 2020

[Nuevos empleos asociados a la inteligencia artificial \(welcometothejungle.com\)](#)

<https://cesi.org.ar/perfiles-it/>

Desarrollo de Software

**◆ Análisis**

Diseño

Calidad de Software

Soporte

Implementadores

Comunicación Online

Seguridad



Analista Big Data [Data Scientist]

Es el responsable de interpretar y realizar descubrimientos en base a grandes volúmenes de información.

**Nombres de cargos similares**

Experto en Big Data, Analista Data Scientist, 'Chief data officer' (CDO), Analista Digital.

Consultor BI – Business Intelligence

Comprender y analizar el contexto de negocio y procesos de las organizaciones y en base a eso diseñar e implementar mejoras. Brindar soporte a la operación.

**Nombres de cargos similares**

Analista Business Intelligence, Especialista en Business Intelligence.

# **Científico, analista, ingeniero de Datos**

## **Científico de datos**

- trata los datos recopilados , debe tener competencias para analizar requerimientos de los clientes, valorar los datos y agregar valor. Apoya toma de decisiones y en la elaboración de una estrategia coherente y razonable.

## **Analista de datos**

- Se especializa en un área de la CD.

## **Ingeniero de datos**

- responsable del funcionamiento del mantenimiento y del desarrollo de la infraestructura y de los sistemas de tratamiento de datos.

# Práctica 2. Perfil en IA

## Objetivo

- Adquirir o profundizar conocimientos concernientes a roles definidos en proyectos ML / DL / CD

## Actividades

1. Analizar las competencias, habilidades y responsabilidades definidas para los talentos asociados a tecnologías inteligentes comprendidas en la CESSI
2. Responder ¿En cuál perfil te identificas ?
3. Exponer

# Ética y Deontología

## Ética

- Estudio y reflexión sobre lo que se considera moralmente correcto o incorrecto en términos de comportamiento humano.

## Deontología

- Ética de las profesiones
- Conjunto de principios y normas éticas que guían la conducta y el comportamiento de profesionales en un campo específico



Raúl Soriano Benajes

Intento de ingeniero con algún que otro logro en el portfolio. Intento de emprendedor con alguna que otra idea que no ha resultado ser tan mala.

13 artículos

+ Seguir

27 de marzo de 2023

En la era de la inteligencia artificial (IA), el debate sobre su impacto en la sociedad y las implicaciones éticas y morales que conlleva es cada vez más relevante. ¿Cómo habrían abordado este tema filósofos de la talla de Descartes y Platón? En este artículo, nos sumergimos en un diálogo ficticio entre estos dos grandes pensadores, explorando sus ideas y perspectivas sobre la ética y la moralidad en el contexto del uso de la IA. Al hacerlo, intentamos comprender cómo sus teorías podrían iluminar nuestra propia reflexión sobre la IA y su papel en nuestra sociedad.

Publicado por



Raúl Soriano Benajes

Intento de ingeniero con algún que otro logro en el portfolio. Intento de emprendedor con alguna que otra idea que no ha resultado ser tan mala.  
Fecha de publicación: 1 día

🤖 ¿Cómo debatirían Descartes y Platón sobre la ética de la Inteligencia Artificial? 🧠🏛️

Imagina por un momento a dos de los más grandes filósofos de la historia, Descartes y Platón, discutiendo sobre la ética y la moralidad del uso de la inteligencia artificial (IA). ¿Cómo podrían sus ideas y perspectivas encajar en este debate moderno? 😊

En mi último artículo, he creado un diálogo filosófico ficticio entre estos dos pensadores, explorando sus puntos de vista sobre la IA y cómo sus teorías podrían aplicarse en el contexto actual. 📖🌐

Sumérgete en esta conversación única y reflexiona sobre las implicaciones éticas de la IA desde dos perspectivas filosóficas distintas. No te pierdas esta oportunidad de aprender y debatir con estos grandes maestros del pensamiento. 😊🎓

👉 ¡Haz clic en el enlace para leer el diálogo completo y comparte tus propias opiniones en los comentarios! 🗣️

#Descartes #Platón #InteligenciaArtificial #Ética #Moralidad #Filosofía #Debate #IA #LinkedIn

<https://www.linkedin.com/pulse/c%C3%B3mo-debatir%C3%ADan-descartes-y-plat%C3%B3n-sobre-la-%C3%A9tica-de-soriano-benajes/?originalSubdomain=es>

# Algunos aspectos éticos en IA.

The ethics of Machine Learning has become an unavoidable topic in the AI Community. The deployment of machine learning systems in multiple social contexts has resulted in a closer ethical scrutiny of the design, development, and application of these systems. The AI/ML community has come to terms with the imperative to think about the ethical implications of machine learning, not only as a product but also as a practice (Birhane, 2021; Shen et al. 2021). The critical question that is troubling many debates is what can constitute an ethically accountable machine learning system. In this paper we explore possibilities for ethical evaluation of machine learning methodologies. We scrutinize techniques, methods and technical practices in machine learning from a relational ethics perspective, taking into consideration how machine learning systems are part of the world and how they relate to different forms of agency. Taking a page from Phil Agre (1997) we use the notion of a critical technical practice as a means of analysis of machine learning approaches. Our radical proposal is that supervised learning appears to be the only machine learning method that is ethically defensible.

# Algunos aspectos éticos en IA

Decision-making on numerous aspects of our daily lives is being outsourced to machine learning (ML) algorithms and artificial intelligence (AI), motivated by speed and efficiency in the decision process. ML approaches—one of the typologies of algorithms underpinning artificial intelligence—are typically developed as black boxes. ***The implication is that ML code scripts are rarely scrutinised; interpretability is usually sacrificed in favour of usability and effectiveness.*** Room for improvement in practices associated with programme development have also been flagged along other dimensions, including inter alia fairness, accuracy, accountability, and transparency. In this contribution, the production of guidelines and dedicated documents around these themes is discussed. The following applications of AI driven decision-making are outlined: (a) risk assessment in the criminal justice system, and (b) autonomous vehicles, highlighting points of friction across ethical principles. Possible ways forward towards the implementation of governance on AI are finally examined.

# Algunos aspectos éticos en IA.

Un uso confiable de IA pasa por abogar que la misma sea inclusiva, responsable y explicable (XAI) <sup>3</sup> buscando transparentar los modelos y lograr que los resultados sean entendibles incluso para interesados no tecnológicos. En este sentido, las recomendaciones de AI4IA (2020) son:

- Los aspectos de ética de la IA tienen que ser accesibles, además de considerar el rol de las comunidades y la diversidad cultural en el diseño, desarrollo y entrenamiento.
- La ética de la información ya figura en los programas de estudio, pero se recomienda incluir también **aspectos de codificación de algoritmos**, y a su vez los autores de los algoritmos deben recibir formación en ética de la información.
- La ética, la transparencia, la dignidad humana y los derechos de los niños deben ser promovidos e implementados desde el inicio del desarrollo de un sistema de IA.
- La necesidad de disminuir la brecha digital a todos los niveles se torna indispensable. El concepto de FAIR aplicado a la IA, o lo que se podría denominar una IA inclusiva, consiste en no discriminar por características sensibles como el género, la etnia, la religión, la discapacidad o la orientación sexual.

Como los modelos de IA se diseñan en una cultura que reproduce sesgos, es necesario testarlos para no excluir y violar los derechos humanos. El movimiento del software libre promueve un software que respeta la libertad de los usuarios y la comunidad. A grandes rasgos, significa que los usuarios tienen la libertad de ejecutar, copiar, distribuir, estudiar, modificar y mejorar el software. Este movimiento fue inspirador de otros de apertura del conocimiento como el acceso abierto

FAIR tiene el significado de justo, equitativo, sin sesgos. En IA, FAIR tiene el sentido de prevenir sesgos hacia determinados subgrupos y mitigar su discriminación

Temésio Vizoso, S. G. (2022). Reflexiones sobre la inteligencia artificial y la bibliotecología. *Palabra Clave (La Plata)*, 11(2), e159.  
<https://doi.org/10.24215/18539912e159>

# Algunos aspectos éticos en IA.

La Inteligencia artificial y el Machine Learning entran en el panorama de una medicina de elevada complejidad con la promesa de mejorar la toma de decisiones médicas y poder recomendar un tratamiento personalizado. Es imprescindible conocer el funcionamiento de los algoritmos para poder adecuarlos en la práctica clínica, conocer sus limitaciones y riesgos. Los principales conflictos éticos detectados están en relación a la privacidad y confidencialidad, la fiabilidad y su impacto en la relación medico paciente (responsabilidad y confianza).

# Algunos aspectos éticos en IA.

## 8 Ethics, Limitations, and Implications

The rapid development and adoption of large language models have been accompanied by a rigorous conversation about the ethics, uses, and limitations of these models. For a more complete treatment of these topics, we direct the reader to Bommasani et al. (2021); Bender et al. (2021); Birhane et al. (2022); Weidinger et al. (2021, 2022). We discuss issues that are directly relevant to the development of BloombergGPT.

### 8.1 Ethical Use

Finance is a sensitive area for technology, and ensuring accurate, factual information is crucial for our products, our clients, and the firm's reputation in the marketplace. (...). Moreover, we conduct our research, development, and deployment of NLP and AI systems in accordance with all applicable regulations.

### 8.2 Openness

An ongoing debate in the community concerns how LLMs should be released, if at all. While models that are not publicly available cannot be fully evaluated by the community, distributing models can lead to nefarious purposes. Especially for a model like BloombergGPT, which is trained on a significant amount of press releases, news articles, and filings, a release carries a high risk for abuse through imitation....

BloombergGPT: A Large Language Model for Finance  
Shijie Wu, , Ozan Irsoy, , Steven Lu, , Vadim Dabravolski ,  
Mark Dredze , Sebastian Gehrmann , Prabhanjan Kambadur ,  
David Rosenberg , Gideon Mann

[2303.17564.pdf \(arxiv.org\)](https://arxiv.org/pdf/2303.17564.pdf)

09/05/2023

# Algunos aspectos éticos en IA

Estela es una solución avanzada de digital compliance que ayuda a las empresas a cumplir con las regulaciones y normativas vigentes en materia de protección de datos y seguridad de la información. Esta plataforma ha sido desarrollada por un equipo de expertos en ciberseguridad y legaltech, con el objetivo de proporcionar una solución integral y eficiente para el cumplimiento normativo en el ámbito digital.

Se basa en tecnologías de vanguardia, como la inteligencia artificial y el análisis de big data, para analizar y gestionar grandes volúmenes de datos de manera rápida y precisa. Esto permite a las organizaciones identificar posibles riesgos de seguridad, detectar vulnerabilidades y tomar medidas preventivas para proteger sus activos digitales y la información de sus clientes.

[Estela busca revolucionar el digital compliance empresarial | Silicon](#)

# Práctica 3. Algunos aspectos éticos en IA.

## Objetivo

- Adquirir o profundizar conocimientos en torno a la Ética en la IA

## Actividades

1. Constituir equipos de 3 – 4 integrantes
2. Seleccionar uno de los artículos recomendados
3. Alguno de los aspectos mencionados se implementa en su ámbito laboral / profesional ?, Podría implementarse en el corto plazo en su ámbito laboral / profesional.
4. Exponer

## **Simposios de las 51 JAIIO**

AGRANDA - Simposio Argentino de Ciencia de Datos y GRANdes DATos

ASAI - Simposio Argentino de Inteligencia Artificial

ASSE - Simposio Argentino de Ingeniería de Software

CAI - Congreso Argentino de AgroInformática

CAIS - Congreso Argentino de Informática y Salud

EST - Concurso de Trabajos Estudiantiles

IETF Day - Taller del Grupo de Trabajo de Ingeniería de Internet/Argentina

JUI - Jornadas de Vinculación Universidad - Industria

SACS - Simposio Argentino de Ciberseguridad

SAEI - Simposio Argentino de Educación en Informática

SAHTI - Simposio Argentino de Historia, Tecnologías e Informática

SAIV - Simposio Argentino de Imágenes y Visión

SID - Simposio Argentino de Informática y Derecho

SIE - Simposio de Informática en el Estado

SIIIO - Simposio Argentino de Informática Industrial e Investigación Operativa

STS - Simposio Argentino sobre Tecnología y Sociedad

[Simposios de las 51 JAIIO | 51 JAIIO \(sadio.org.ar\)](http://sadio.org.ar)

## **Simposios de las 52 JAIIO**

[AGRANDA - Simposio Argentino de Ciencia de Datos y GRANdes DATos](#)

[ASAI - Simposio Argentino de Inteligencia Artificial](#)

[ASSE - Simposio Argentino de Ingeniería de Software](#)

[CAI - Congreso Argentino de AgroInformática](#)

[CAIS - Congreso Argentino de Informática y Salud](#)

[EST - Concurso de Trabajos Estudiantiles](#)

[IETF Day - Taller del Grupo de Trabajo de Ingeniería de Internet/Argentina](#)

[JUI - Jornadas de Vinculación Universidad - Industria](#)

[SACS - Simposio Argentino de Ciberseguridad](#)

[SAEI - Simposio Argentino de Educación en Informática](#)

[SAHTI - Simposio Argentino de Historia, Tecnologías e Informática](#)

[SAIC - Simposio Argentino en Ingeniería en Computación](#)

[SAIV - Simposio Argentino de Imágenes y Visión](#)

[SID - Simposio Argentino de Informática y Derecho](#)

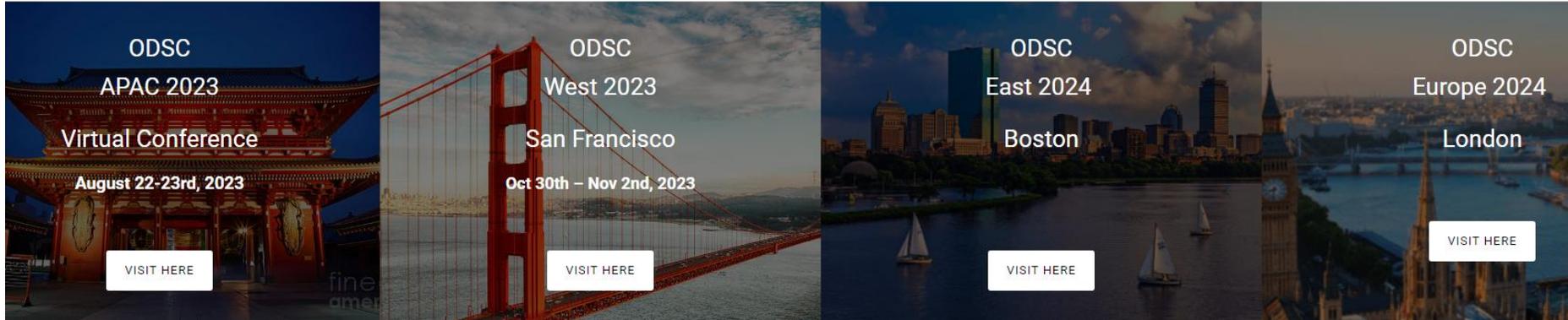
[SIE - Simposio de Informática en el Estado](#)

[SIIIO - Simposio Argentino de Informática Industrial e Investigación Operativa](#)

[STS - Simposio Argentino sobre Tecnología y Sociedad](#)

[Simposios de las 52 JAIIO | 52 JAIIO \(sadio.org.ar\)](#)

# Eventos



[Open Data Science Conference - Deep Learning & ML, Predictive Analytics, AI \(odsc.com\)](https://odsc.com)

# Eventos

[Semana Internacional de Datos y SciDataCon 2023 - Consejo Internacional de Ciencias](#)

[\(council.science\)](#)

# Referencias

S. I. Mariño, UNNE, 2019 - 2023.

N. J. Nilsson. Inteligencia Artificial: Una Nueva Síntesis, 2001, Ed McGraw-Hill.

S. Russell, y P. Norvig, Inteligencia Artificial: un enfoque Moderno, 2004 <http://aima.cs.berkeley.edu/>

S. Russell, y P. Norvig, Artificial Intelligence, A Modern Approach, Prentice–Hall, 3rd Edition, 2010

Artículos de la disciplina