Unidad 1.3. Inicios y evolución de la IA.

El nacimiento de la IA

> *Texto en la placa conmemorativa de la conferencia de Dartmouth (1956): "Primera utilización del término "Inteligencia Artificial". Fundando la Inteligencia Artificial como una disciplina de investigación para proceder sobre la base de la conjetura de que cada aspecto del aprendizaje o cualquier otra característica de la inteligencia puede, en principio, describirse tan precisamente que se puede construir una máquina para simularlo".*

A partir de los años 50 se produce una evolución explosiva de la IA aunque con altibajos, también llamados inviernos de IA.

El resumen de tal proceso se presenta en esta imagen y se desarrolla a continuación:

Copyright 2025

Desde la histórica **conferencia de Dartmouth en 1956,** que marcó el inicio formal de la Inteligencia Artificial (IA) como campo de investigación, hasta el año 2011, se han producido importantes avances en la evolución de la IA. Los impulsores de la conferencia de Dartmouth fueron: J**ohn Mcarthy (Dartmoth College), Marvin L. Minsky (MIT), Nathaniel Rochester (IBM, y Claude Shannon (Laboratorios Bell)**. Cada uno de ellos era líder en áreas cruciales para el nacimiento de la IA. McCarthy en la formulación de conceptos de IA, Minsky en el aprendizaje de máquinas y redes neuronales, Rochester en hardware y arquitectura de computadoras, y Shannon en teoría de la información. Sus conocimientos y colaboraciones fueron fundamentales para establecer una base sólida que permitió el desarrollo y crecimiento de la IA como una disciplina de investigación.

A partir de ese momento, y creciendo sobre los cimientos de las teorías, desarrollos conceptuales y tecnológicos acumulados hasta ese momento, se fueron sucediendo logros técnicos y tecnológicos que han desembocado en la irrupción aparentemente imparable de la IA en la actualidad.

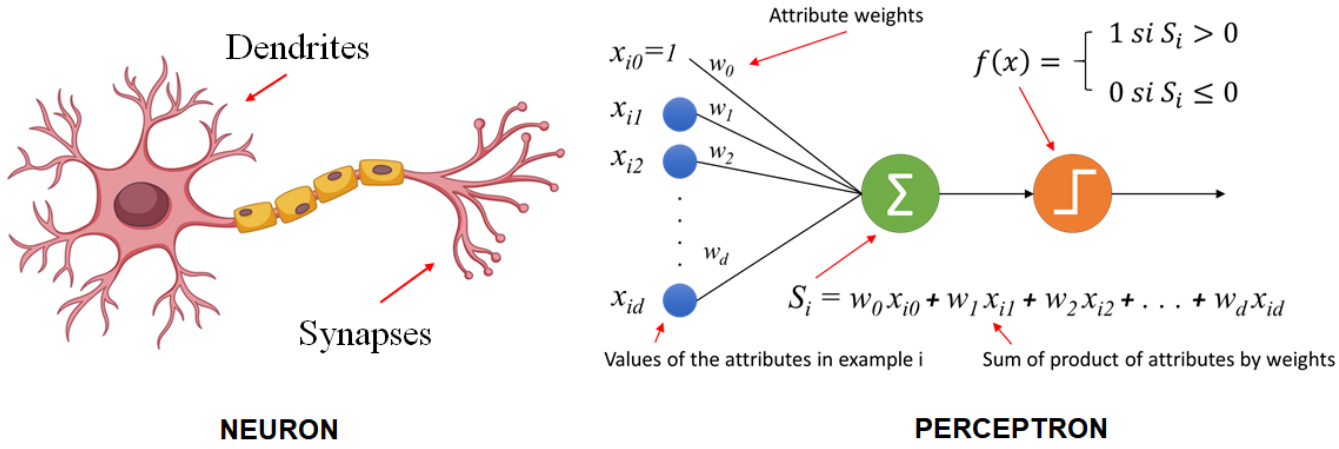
Uno de los hitos clave fue la invención del **perceptrón por Frank Rosenblatt en 1957**. El perceptrón fue uno de los primeros modelos de aprendizaje automático inspirados en las redes neuronales del cerebro humano. El perceptrón, era una evolución de la neurona McCulloch-Pitts, si bien mejoraba a esta al introducir el concepto de aprendizaje automático. Mientras que la neurona

Copyright 2025

de McCulloch-Pitts tenía pesos fijos, el perceptrón permitía el ajuste de los pesos en función de los datos de entrada a través de un proceso de aprendizaje simple, lo que permitía que el modelo aprendiera a clasificar diferentes patrones de datos automáticamente. Esta capacidad de aprender de los datos fue un avance significativo hacia el desarrollo de algoritmos más avanzados de aprendizaje automático y redes neuronales.

Seymour Papert, que trabajó junto con Marvin Minsky y ambos fueron coautores de un libro muy influyente llamado "Perceptrons" en 1969, que exploraba las limitaciones y las capacidades de los perceptrones y las redes neuronales. Su trabajo junto con Minsky fue crucial para entender y analizar las capacidades y limitaciones de estos modelos, lo cual fue fundamental en el desarrollo temprano de la inteligencia artificial y el aprendizaje automático.

> *"El perceptrón es el embrión de una computadora que será capaz de hablar, caminar, ver, escribir, reproducirse y ser consciente de su existencia" Seymour Papert*



NEURON **PERCEPTRON**

Comparativa gráfica entre una neurona y el perceptrón. https://inteligenciafutura.mx/english-version-blog/blog-06-english-version

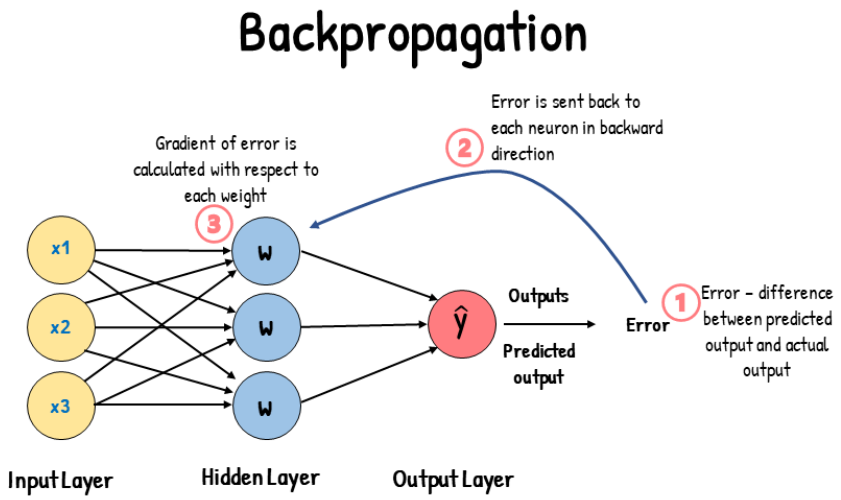
---

Ciclos de esperanza y escepticismo: Los Altibajos de la IA (1970-2010)

Sin embargo, a pesar de sus prometedoras capacidades, pronto surgieron limitaciones y desafíos que llevaron a un declive en la investigación de la IA, conocido como el **"primer invierno de la IA".** La citada publicación "Perceptrons" en la que también se mencionaban las limitaciones del mismo anticipó ese declive.

En el "primer invierno de la IA", que tuvo lugar durante la década de 1970 y principios de la década de 1980, los avances en la IA se estancaron debido a la falta de resultados prácticos y las expectativas excesivas. El financiamiento se redujo y muchos investigadores abandonaron el campo.

Si bien, aún sin financiación a gran escala, los trabajos y desarrollos teóricos siguieron produciéndose y en 1974 el matemático y economista **Paul Werbos** publicó su tesis doctoral en la que introdujo el concepto del algoritmo de **retropropagación (backpropagation)** que es fundamental para entrenar **redes neuronales multicapa**. Este algoritmo se convirtió en una base crucial para el desarrollo del **aprendizaje profundo**, facilitando el entrenamiento eficaz de redes neuronales y contribuyendo al avance de la inteligencia artificial.



Esquema gráfico del funcionamiento del algoritmo de retropropagación.

https://www.analyticsvidhya.com/blog/2023/01/gradient-descent-vs-backpropagation-whats-the-difference/

A grandes rasgos el algoritmo de retropropagación introduce el posible error cometido en la predicción realizada por el algoritmo respecto al resultado esperado, permitiendo a la red neuronal ajustar los pesos de los cálculos que se realizan en las capas de neuronas con el objetivo de afinar cada vez mas el resultado final.

Avances como la retropropagación, el nuevo enfoque de los desarrollos que se centraban en problemas concretos, lo que resultó en lo que se dio a denominar **Sistemas Expertos**, así como la

Copyright 2025

fundamentación de los sistemas en la gestión estructurada de grandes cantidades de datos condujo a mediados de la década de 1980 a un resurgimiento en la investigación de la IA y a una vuelta de la financiación, impulsada no solo por iniciativa privada sino también por proyectos nacionales como en Japón o Estados Unidos.

Sin embargo, a finales de la década de los 80, volvió a enfriarse el entusiasmo. Las causas en esta ocasión se encontraban en un estancamiento provocado por las limitaciones tecnológicas y la falta de robustez en los sistemas existentes. Por tanto, se dejaron de lado los proyectos en curso en lo que se denomino "**segundo invierno de la IA**" y los esfuerzos volvieron a centrarse en un replanteamiento del enfoque.

Este cambio de perspectiva se materializó en varios frentes:

Copyright 2025

Copyright 2025

https://view.genial.ly/65246f2a064bf30010bfd3f4

Copyright 2025

Relacionado con este cambio de paradigma se publico el artículo:

- **"Elephants Don’t Play Chess" de Rodney Brooks (Laboatorio de IA del MIT)**

En este artículo, **Rodney Brooks** argumenta en contra del enfoque tradicional basado en la lógica simbólica y a favor de un enfoque más basado en la interacción con el entorno. Brooks propone que los sistemas de IA deberían estar más arraigados en el mundo físico y ser capaces de interactuar con él de manera efectiva, en lugar de estar confinados a tareas abstractas y descontextualizadas como jugar al ajedrez. Su trabajo promovió una visión de la IA que valora la interacción con el entorno y la capacidad de adaptarse a él, lo cual es crucial para el desarrollo de r eractuantes.

https://msujaws.wordpress.com/2010/11/18/elephants-dont-play-chess/

En contraposición a los enfoques anteriores que se centraban en la lógica simbólica y en sistemas expertos con un conocimiento codificado de manera rígida, estos nuevos enfoques adoptaron una perspectiva más flexible y basada en datos, que permitió a la IA superar el segundo invierno y entrar en una nueva era de innovación y crecimiento. Además, el apoyo continuo de la comunidad académica y de la industria, junto con avances en hardware y disponibilidad de datos, contribuyeron a superar las limitaciones que habían causado el segundo invierno de la IA.

Estos cambios en el enfoque hacia una IA más interactuante y emergente representaron un movimiento importante lejos de los enfoques anteriores, y se alinean con la tendencia general hacia la IA que es capaz de aprender y adaptarse de manera más efectiva a su entorno.

Este periodo, que abarcaría desde la década de 1990 hasta finales de la década de 2010, estuvo plagado de logros concretos de la IA que supusieron hitos importantes en la constatación de su evolución en campos que siguen desarrollándose actualmente, como ejemplos

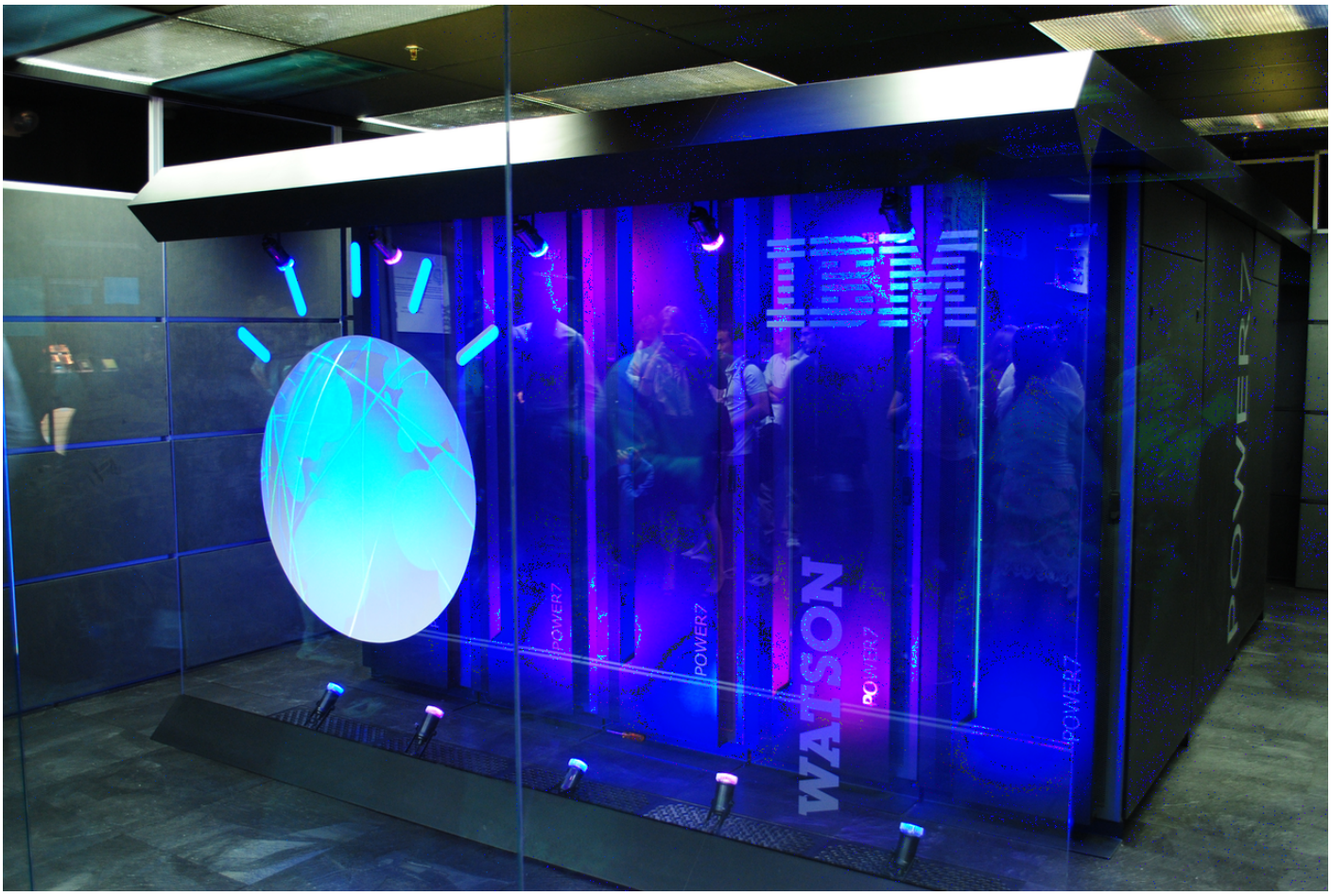
- **Estrategias de juego y simulación:** En una serie de seis partidas de ajedrez, la computadora **Deep Blue de IBM** derrotó al entonces campeón mundial **Garry Kasparov.** Esta victoria marcó la primera vez que una máquina derrotó a un campeón mundial de ajedrez en condiciones de torneo, y fue un hito significativo en la demostración de la capacidad de las máquinas para realizar tareas que tradicionalmente se consideraban dominio exclusivo de la inteligencia humana.

- **Reconocimiento de voz: Dragon NaturallySpeaking** fue uno de los primeros programas comerciales de reconocimiento de voz, presentado en 1997. Permitió a los usuarios hablar directamente a sus computadoras y convertir sus palabras en texto. Todavía se comercializa en la actualidad.

- **Conducción autónoma: e**l **DARPA Grand Challenge** (2004-2007) fue una serie de desafíos, organizados por la Agencia de Proyectos de Investigación Avanzada de Defensa (DARPA) de EE.UU. que incentivaron el desarrollo de vehículos autónomos. El desafío de 2005 fue notable porque cinco vehículos completaron un recorrido en el desierto de 131 millas, con el equipo de Stanford llamado **"Stanley"** llegando primero.

- **Procesamiento de lenguaje natural (NLP): WordNet** (1995) una base de datos léxica para el inglés que ha sido fundamental para muchas aplicaciones de NLP. Ayuda en la relación semántica entre palabras. También en este campo **IBM Watson,** sistema de IA capaz de responder a preguntas en lenguaje natural y con capacidad de aprendizaje automático venció contra campeones humanos en el juego de televisión "**Jeopardy!**".

Copyright 2025



Watson de IBM. De Clockready - Trabajo propio, CC BY-SA 3.0, https://commons.wikimedia.org/w/index.php?curid=15891787

- **Reconocimiento de imágenes: Scale-Invariant Feature Transform (SIFT)** es un algoritmo usado en visión artificial para extraer características relevantes de las imágenes que posteriormente pueden usarse en reconocimiento de objetos, detección de movimiento, estereopsis, registro de la imagen y otras tareas.
  El algoritmo fue publicado por primera vez por **David Lowe** en 1999 pero lo describió completamente y patentó en Estados Unidos en 2004.
  También en reconocimiento de imágenes es destacable **ImageNet** o **ImageNet Large Scale Visual Recognition Challenge (ILSVRC).** Iniciado en 2010, este desafío se convirtió en una competición anual que evalúa los algoritmos en tareas de detección de objetos y clasificación de imágenes a gran escala utilizando el conjunto de datos **ImageNet**. La competición ha tenido un impacto inmenso en el avance de los algoritmos de reconocimiento de imágenes, particularmente con el advenimiento de las **redes neuronales convolucionales (CNN).**

- **Google Brain:** fue fundado en 2011. Uno de sus primeros proyectos importantes fue un **sistema de aprendizaje profundo** que, cuando se entrenó con miles de horas de videos de YouTube, aprendió por sí mismo a reconocer gatos y otros objetos. Esta demostración fue una de las primeras en mostrar el poder y potencial del aprendizaje profundo a gran escala.

Hitos contemporáneos, de la IA a la AGI

> "*Attention is all you need.*" - Vaswani et al., "Attention Is All You Need" (2017)

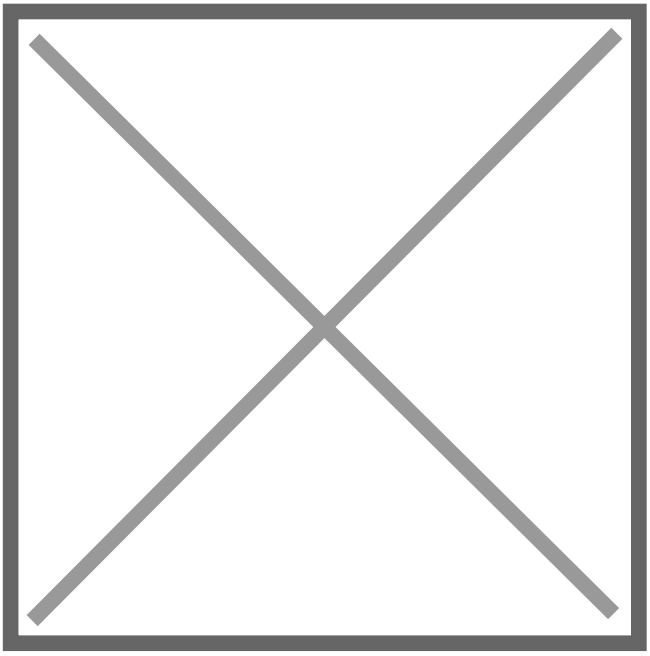
> "*Los modelos de lenguaje como ChatGPT pueden ser herramientas increíblemente poderosas, pero también conllevan una gran responsabilidad para garantizar que se utilicen de manera ética y segura.*" - Sam Altman

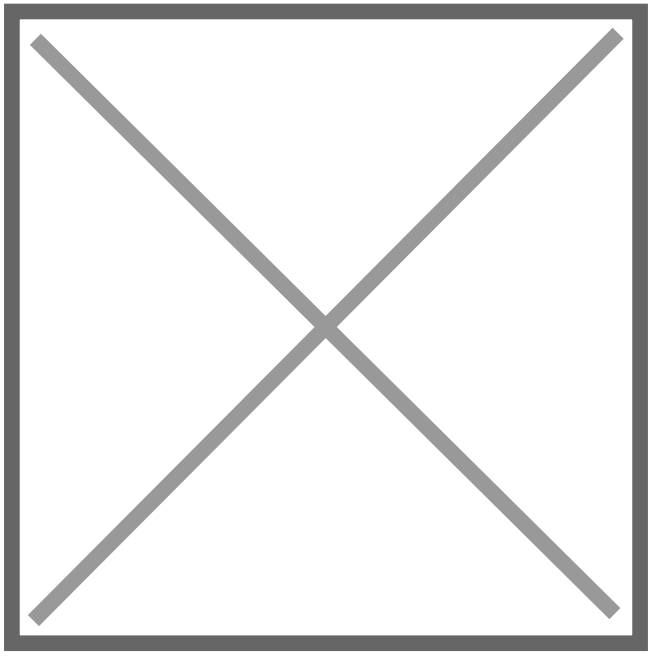
Durante el periodo de 2011 a 2023, ha habido un notable avance en las tecnologías de Inteligencia Artificial, particularmente en el desarrollo de nuevos **algoritmos**, evolucionadas **redes neuronales ,** así como revolucionarias **arquitecturas**. Estas tecnologías han sido fundamentales para el procesamiento de imágenes, el procesamiento del lenguaje natural y otras tareas de IA. Aquí se presenta una visión general de su evolución:

En cuanto a las **redes neuronales** destacan:

- Las **redes neuronales generativas adversariales (GAN)**, la red consta de dos modelos, uno generador y otro discriminador.
  El discriminador debe analizar las muestras generadas por el generador y tratar de discernir si son reales o sintéticas, retroalimentando el aprendizaje.

Copyright 2025





A la izquierda una GAN "trabajando" para crear el retrato de ***Edmond Belamy***, a la derecha. Esta obra fue subastada en **Christie's** y se convirtió así en la primera casa de apuestas en subastar una obra generada por una IA.

El **Generador** (de la red GAN) produjo imágenes basadas en un conjunto de datos de 15.000 retratos pintados entre los siglos XIV y XX, y el **Discriminador** intentó diferenciar las obras creadas por el hombre y las generadas por IA.

"El objetivo es engañar al Discriminador haciéndole creer que las nuevas imágenes son retratos de la vida real", dice Caselles-Dupré, cofundador de *Obvious,* colectivo artístico promotor de la iniciativa. "Entonces tenemos un resultado". Edmon Belamy es uno de los 11 retratos que representan a la familia ficticia Belamy.

- Las **redes neuronales convolucionales (CNN),** citadas anteriormente: Las CNN experimentaron un gran avance en el procesamiento de imágenes y la visión por computadora. En 2012, el modelo **AlexNet** revolucionó el campo al ganar la competición ImageNet, superando con éxito a los enfoques tradicionales. Desde entonces, se han propuesto y mejorado diversas arquitecturas de CNN, como VGGNet, InceptionNet y ResNet, que han superado los límites de precisión en tareas de clasificación, detección y segmentación de objetos en imágenes.

- Las **redes neuronales LSTM (Long Short-Term Memory)** son un tipo de red neuronal recurrente (RNN) que han sido fundamentales en el procesamiento del lenguaje natural y

Copyright 2025

en la generación de texto. Las LSTM resuelven el problema de las RNN tradicionales al permitir el almacenamiento y el acceso a información a largo plazo, lo que las hace más efectivas en el procesamiento de secuencias largas. Estas redes han mejorado significativamente la capacidad de modelar y generar texto coherente y natural.

Además de las redes neuronales la evolución de las **arquitecturas**, entendidas como el diseño general de la estructura de un modelo, ha representado un salto fundamental en el campo de la inteligencia artificial y en concreto el procesamiento de lenguaje natural (NLP):

- La **arquitectura encoder-decoder**, se ha utilizado para la extracción de características y la representación eficiente de datos. Esta arquitectura permite comprimir la información relevante de una entrada en una **representación latente de menor dimensión**, lo que facilita el procesamiento y análisis posterior.

- Los **transformers,** son una arquitectura de redes neuronales que ha revolucionado el procesamiento del lenguaje natural y otras tareas de secuencia. Introducidos en 2017, los transformers se basan en el **mecanismo de atención** para capturar relaciones entre elementos en una secuencia. Esta arquitectura ha demostrado un rendimiento sobresaliente en tareas como la traducción automática, la generación de texto y el procesamiento del lenguaje natural.

Estos avances técnicos dieron pie a la consecución de determinados logros, algunos de los cuales se consideran momentos clave en la historia de la IA, algunos de ellos son:

- **AlphaGo - AlphaZero:** en 2014 Google presenta su sistema de IA llamado **Google DeepMind** y lanza su primer gran proyecto de IA conocido como **AlphaGo.** AlphaGo derrota a **Lee Sedol**, campeón mundial de Go, un juego extremadamente complejo y estratégico, demostrando la capacidad de la IA para superar a los mejores expertos humanos en juegos de mesa. Alpha Go había sido entrenado mediante una combinación de **aprendizaje supervisado** y aprendizaje por refuerzo, utilizando una gran cantidad de datos de partidas humanas.

Copyright 2025

Posteriormente, DeepMind subió la apuesta y en 2017 lanzó **AlphaGo Zero**. A diferencia de AlphaGo, que se entrenó en juegos jugados por humanos, AlphaGo Zero fue entrenado, a partir del único conocimiento de las reglas del juego, desde cero, solo mediante el **aprendizaje de refuerzo**. A través de la autoevaluación y la retroalimentación de las partidas jugadas contra ella misma, AlphaGo Zero ajustaba sus estrategias y mejoraba su rendimiento. AlphaGo Zero se convirtió en el mejor jugador del mundo en tres días, superando la versión anterior de AlphaGo. Este evento mostró el potencial de los sistemas de aprendizaje profundo para superar incluso a los mejores jugadores humanos sin la necesidad de datos de entrenamiento previos.

- **GPT (Generative Pre-trained Transformer):** primera versión del modelo de lenguaje de OpenAI, desarrollado por el investigador de IA **Andrej Karpathy,** demuestra una capacidad asombrosa para generar texto coherente y de alta calidad.

Los modelos basados en la arquitectura Transformer han mejorado significativamente la capacidad de procesamiento de los modelos de lenguaje natural. Antes de la introducción de los Transformers, los modelos de lenguaje natural más populares eran los modelos basados en RNN (Redes Neuronales Recurrentes), que se enfrentaban a problemas de "**desvanecimiento de gradiente" y "memoria a corto plazo**". Estos problemas hacían que los modelos de lenguaje basados en RNN fueran difíciles de entrenar y limitaban su capacidad para comprender contextos más amplios y a largo plazo.

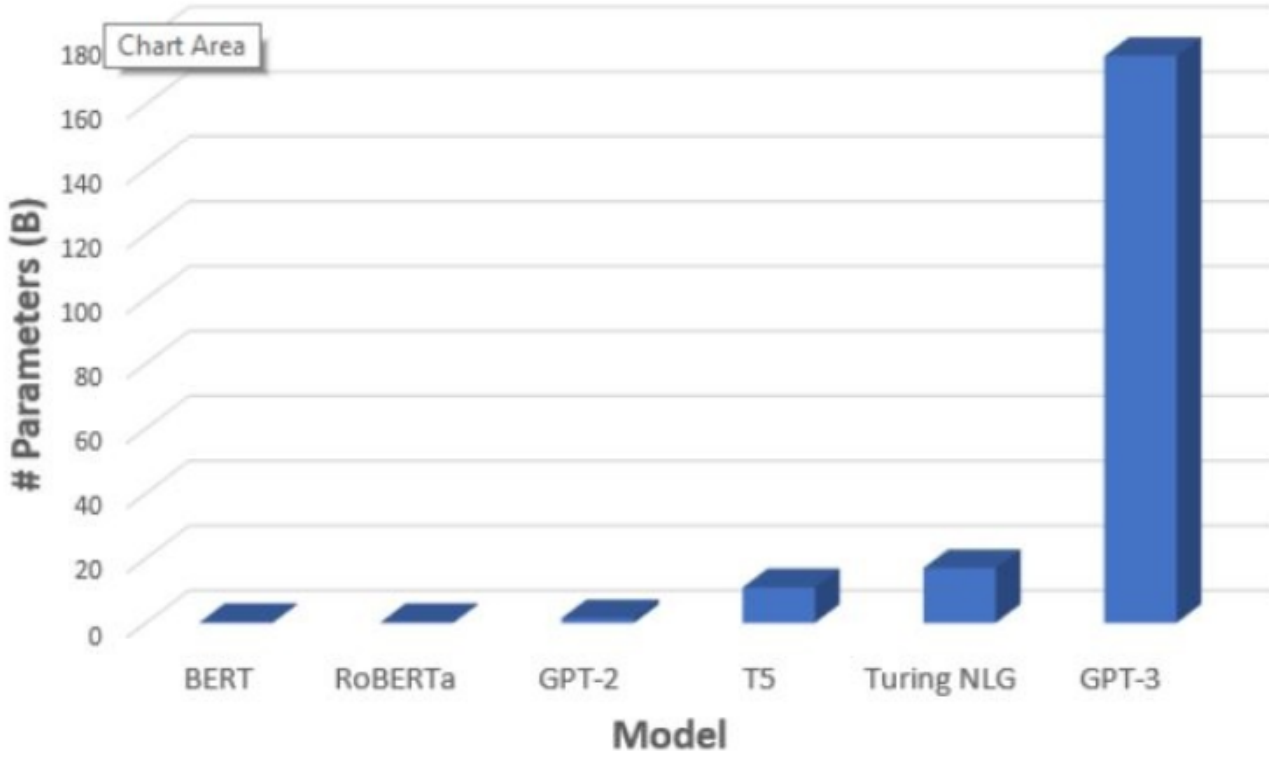
En contraste, los modelos que utilizan Transformers, como GPT (Generative Pretrained Transformer) o BERT de Google (Bidirectional Encoder Representations from Transformers) entre

Copyright 2025

otros, utilizan una arquitectura basada en **atención** que les permite comprender relaciones entre diferentes partes de un texto de manera más efectiva, permitiendo una mejor comprensión del contexto y sin problemas de desvanecimiento de gradiente. Además, los Transformers permiten un entrenamiento más rápido y escalable que los modelos RNN, lo que ha llevado a mejoras significativas en la precisión de los modelos de lenguaje natural.

Esto ha permitido también la inclusión de muchos más parámetros en el modelo, lo que ha llevado a mejoras significativas en la precisión y la capacidad de comprensión del modelo.

La tercera versión del modelo, GPT-3, por ejemplo, tiene 175 mil millones de parámetros, mientras que los modelos de lenguaje más grandes, como BERT, con 110 millones de parámetros, o Turing NLG que tiene 17 mil millones.



La inclusión de más parámetros en el modelo permite a GPT-3 comprender mejor el contexto y generar textos más coherentes y naturales, lo que ha llevado a avances significativos en las tareas de procesamiento del lenguaje natural, como la traducción automática, la respuesta a preguntas y la generación de texto.

- **Alphafold:** es uno de los logros mas significativos conseguidos por la IA hasta la fecha. La empresa DeepMind resolvió el problema del plegamiento de las proteinas usando aprendizaje automático para predecir la estructura 3D de las proteinas que a su vez determina su funcionalidad.

Copyright 2025

AlphaFold, desarrollado por DeepMind, es un sistema de inteligencia artificial que ha revolucionado la predicción de la estructura 3D de las proteínas, lo que determina a su vez la funcionalidad de las m[illegible]nas, al lograr una precisión sin precedentes usando aprendizaje automático. Esto tiene un g[illegible]n impacto en la investigación biomédica y farmacéutica al acelerar el desarrollo de tratamientos y terapias médicas. También marca un hito importante en la inteligencia artificial al demostrar c[illegible]mo puede abordar con éxito problemas complejos y multidisciplinarios en la ciencia.

En la imagen se aprecia la forma de dos proteínas distintas, en cada una de ellas, en color verde la forma real de la proteína y en azul la predicción realizada por Alphafold. Mas información:

https://alphafold.ebi.ac.uk/

Estos son solo algunos de los hitos importantes que han ocurrido en el campo de la IA a partir de 2011. La IA continúa avanzando rápidamente, con nuevos descubrimientos, aplicaciones innovadoras y desafíos éticos y sociales que surgen a medida que la tecnología evoluciona.

Copyright 2025

Revision #33
Created 2 October 2023 17:48:29 by Pedro López
Updated 27 November 2023 19:05:27 by Luis Hueso

Copyright 2025