

POSGRADO INTERNACIONAL
QUINTA EDICIÓN

INTELIGENCIA ARTIFICIAL Y DERECHO

ENFOCADO EN IA GENERATIVA Y CHATGPT

GLOSARIO



.UBAderecho



IALAB

POSGRADO INTERNACIONAL
QUINTA EDICIÓN

INTELIGENCIA ARTIFICIAL Y DERECHO

ENFOCADO EN IA GENERATIVA Y CHATGPT

APRENDIZAJE AUTOMÁTICO

El aprendizaje automático o aprendizaje de máquina es el subconjunto de inteligencia artificial (IA) que se centra en desarrollar sistemas que aprenden a responder asertivamente o mejoran el rendimiento, en función de los datos con los que se entrenan. Existe una gran paleta de tipos de aprendizaje, desde el supervisado hasta el No Supervisado. Para todos los casos la supervisión se refiere a un proceso automático que realizan los algoritmos de aprendizaje en base a los datos proporcionados (1).



APRENDIZAJE NO SUPERVISADO

En esta modalidad de aprendizaje, el algoritmo de entrenamiento descubre por sí mismo similitudes que encierran los datos. Se aplican para: reconocimiento o asociación basados en la memoria, descubrir características importantes, agrupar datos según sus similitudes, entre otros. Para todos los casos, la principal distinción a simple vista es que el dataset usado NO contiene etiquetas; aunque ésta no es la única característica distintiva, suele ser útil para identificarlos rápidamente. Además, como los algoritmos de entrenamiento para el aprendizaje se basan en medir distancias, los datasets deben ser numéricos (2).

APRENDIZAJE PARCIALMENTE SUPERVISADO

Comprende todos los modos de aprendizaje automático que NO son ni totalmente Supervisado ni completamente No Supervisado.

APRENDIZAJE POR REFUERZO

(APRENDIZAJE PARCIALMENTE SUPERVISADO)

Técnica de Machine Learning en la que el modelo aprende a actuar en un entorno mediante un proceso de prueba y error, en función de las recompensas que reciba por sus acciones. De este modo, el modelo debe recoger la relación entre las decisiones tomadas y las recompensas consiguientes (3).

APRENDIZAJE PROFUNDO

Pertenece a la familia de Redes Neuronales Artificiales de Aprendizaje Supervisado. Incluye en sus procesos módulos que descubren por sí mismos características importantes que encierran los datos. A diferencia de los algoritmos tradicionales de aprendizaje automático, los sistemas de aprendizaje profundo pueden mejorar su rendimiento manipulando un gran número de parámetros que se ajustan automáticamente durante su entrenamiento. Una vez que las máquinas han conseguido suficiente experiencia mediante el aprendizaje profundo, pueden ponerse a trabajar para realizar tareas específicas (4).

El aprendizaje profundo es un método de aprendizaje de representación con múltiples capas de neuronas o elementos. Permite que los modelos computacionales que se componen de múltiples capas de procesamiento aprendan representaciones de datos con múltiples niveles de abstracción, así como también descubre una estructura intrincada en grandes conjuntos de datos mediante el uso del algoritmo de retropropagación para indicar cómo una máquina debe cambiar sus parámetros internos que se utilizan para calcular la representación en cada capa a partir de la representación en la capa anterior (5).

APRENDIZAJE SUPERVISADO

Dinámica de entrenamiento propia de Machine Learning. Se proporciona al ordenador datos etiquetados presentado en parejas la información de input (características) y de output (etiquetas). De este modo, son capaces de descubrir los patrones de comportamiento. Se usa en aplicaciones de regresión y de clasificación (6).

En el aprendizaje supervisado, el algoritmo de aprendizaje entrena con un conjunto de datos del cual ya conoce los resultados que la red debería entregarle, así el sistema evoluciona hasta que aprende y su señal de salida o respuesta es aceptablemente similar a la esperada para cada ejemplo de entrenamiento al que se enfrenta y para cada ejemplo de prueba (7).

AUTOMATIZACIÓN

Se aplica para optimizar o simplificar tareas previsibles, mecánicas, estandarizadas o rutinarias. En los procesos de automatización frecuentemente se utilizan técnicas o herramientas de IA.

CAJAS BLANCAS

Conjunto de técnicas que se utilizan para obtener predicciones, automatizaciones, clasificaciones o detecciones inteligentes sin el riesgo de “inexplicabilidad” del razonamiento que tienen los sistemas de “caja negra”, ya que permiten conocer el razonamiento subyacente detrás de sus resultados. Los resultados a los que se arriban son auditables, trazables, explicables e interpretables y ello resulta muy beneficioso para comprender la dinámica del tratamiento automatizado (8).

CAJAS NEGRAS

Tipos de Sistemas Inteligentes de aprendizaje automático en los cuales no es posible proporcionar explicabilidad del funcionamiento del modelo entrenado (9), de modo tal que sea fácilmente comprensible al cerebro humano.

CHATBOTS

En el nivel más básico, un chatbot es un programa informático que simula y procesa conversaciones humanas (ya sea escritas o habladas), permitiendo a los humanos interactuar con dispositivos digitales como si se estuvieran comunicando con una persona real. Los chatbots pueden ser tan sencillos como programas rudimentarios que responden a consultas sencillas con una respuesta de una sola línea, o bien tan sofisticados como los Agentes Conversacionales que pueden aprender y evolucionar para ofrecer niveles de personalización cada vez mayores a medida que reúnen y procesan información (10).

CICLO DE VIDA DE LA IA

Ciclo de desarrollo de un sistema de Inteligencia Artificial que incluye las siguientes fases: i) Planeación y diseño, ii) recolección, conocimiento y preparación de datos, iii) modelado, iv) evaluación e implementación y vi) monitoreo. Estas fases suelen tener lugar de manera iterativa y no son necesariamente secuenciales (11).

CLASIFICACIÓN

Técnica de Aprendizaje Supervisado cuyo objetivo es predecir a qué clase pertenece un elemento individual dentro de un conjunto amplio de elementos (12). La clase es identificada con una etiqueta, el conjunto de elementos es el dataset que se suministra al algoritmo inteligente y el elemento individual es un ejemplo de ese dataset.

CONTROL HUMANO

Refiere a la capacidad de que intervengan seres humanos durante el ciclo de diseño del sistema, en el seguimiento de su funcionamiento y durante todos los ciclos de decisión del sistema. En otras palabras, busca supervisar la actividad global del sistema de IA (incluidos, desde un punto de vista más amplio, sus efectos económicos, sociales, jurídicos y éticos), así como la capacidad de decidir cómo y cuándo utilizar el sistema en una situación determinada. Esto puede incluir la decisión de no utilizar un sistema de IA en una situación particular, establecer niveles de discrecionalidad humana durante el uso del sistema o garantizar la posibilidad de ignorar una decisión adoptada por un sistema. El control humano ayuda a garantizar que un sistema de IA actúe tal cual como fue previsto para su uso por el ser humano (13).

DATOS ABIERTOS

Aquellos datos digitales que son puestos a disposición con las características técnicas y jurídicas necesarias para que puedan ser usados, reutilizados y redistribuidos libremente por cualquier persona, en cualquier momento y en cualquier lugar. En más detalle, los principios definidos por La Carta Internacional de Datos Abiertos son que los datos sean abiertos por defecto, oportunos y exhaustivos, accesibles y utilizables, comparables e interoperables, para mejorar la gobernanza y la participación ciudadana (14).

DATOS PERSONALES

Datos personales significa "cualquier información relacionada con una persona física identificada o identificable ("sujeto de datos"); Una persona física identificable es aquella que puede ser identificada, directa o indirectamente, en particular por referencia a un identificador, como un nombre, un número de identificación, datos de ubicación, un identificador en línea o uno o más factores específicos a lo físico, fisiológico, identidad genética, mental, económica, cultural o social de esa persona física" (15).

FINE-TUNING O AJUSTE FINO

Es una técnica de entrenamiento que consiste en la reutilización de arquitecturas de redes neuronales predefinidas y preentrenadas. Se trata de un proceso en el que se realiza un "ajuste fino" de algunas capas de la red para obtener las salidas deseadas. Es decir, se ajustan ligeramente ciertas representaciones del modelo preentrenado para que este resulte más relevante en el problema que se desea resolver. Así, se evita realizar el entrenamiento de la red desde cero.

GRANDES MODELOS DE LENGUAJE O LLM

Es un modelo de lenguaje que consta de una red neuronal con muchos parámetros, entrenados en grandes cantidades de texto sin etiquetar mediante aprendizaje autosupervisado o aprendizaje semisupervisado. Estos modelos pueden ser utilizados en una amplia gama de tareas relacionadas con el procesamiento del lenguaje natural, como la creación de textos, el resumen de documentos o la traducción de idiomas.

HUMAN IN THE LOOP

Sugiere que los humanos están involucrados activa e involucradamente en la supervisión. Los humanos retienen el control total e IA solo proporciona recomendaciones. Las decisiones no se pueden tomar sino con intervención humana.

HUMAN OVER THE LOOP

La supervisión humana se limita a monitorear y supervisar a la IA, con la habilidad de tomar control cuando el modelo de IA encuentra eventos inesperados o no deseados. Permite a los humanos ajustar hiperparámetros durante la operación del algoritmo. Ej: sistema de recomendación de ruta por GPS en el cual el humano puede modificar parámetros que lleve al algoritmo a modificar la ruta.

INTELIGENCIA ARTIFICIAL DÉBIL O RESTRINGIDA

Sistemas de IA en los que el aprendizaje se especializa en una única labor o ejecución de una única tarea, logrando una precisión igual o superior a la de un ser humano (16).

INTELIGENCIA ARTIFICIAL FUERTE O GENERAL

IA que logra generalizar el proceso de aprendizaje aplicándolo a distintas tareas con creatividad y conciencia propia (17).

INTELIGENCIA ARTIFICIAL GENERATIVA

Algoritmos de aprendizaje automático que permiten crear contenido nuevo a partir de los datos existentes, como música, videos, fotografías o texto. El objetivo es obtener producciones que sean indistinguibles de aquellas creadas por los humanos.

INTELIGENCIA ARTIFICIAL "SUPERINTELIGENCIA"

Tipo de Inteligencia Artificial que lograría superar la capacidad cognitiva del ser humano en todos los aspectos (18).

MODELOS DE APRENDIZAJE AUTOMÁTICO

Tipos de modelos de aprendizaje automático que se entrenan para reconocer determinados tipos de patrones (19).

MODELOS DE LENGUAJE

Son modelos de aprendizaje automático que funcionan prediciendo cual será la siguiente palabra dentro de una oración, teniendo en cuenta el contexto de las palabras anteriores. Para esto se basan en una descripción probabilística de los fenómenos del lenguaje.

PROCESAMIENTO DEL LENGUAJE NATURAL (PLN)

Es el campo de conocimiento de la Inteligencia Artificial que se ocupa de investigar de la investigar la manera de comunicar las máquinas con las personas mediante el uso de lenguas naturales, como el español, el inglés o el chino (20).

PROMPT

Mecanismo que permite que los humanos indique a la inteligencia artificial aquello que desea lograr cuando no es posible mostrar un ejemplo de referencia. Por ejemplo, cuando se desea a generación de una imagen con inteligencia artificial y no se puede mostrar un ejemplo, la persona utiliza las palabras para describir lo que quiere.

Prompt engineering: Técnicas que permiten guiar los modelos en el proceso de generación de respuestas que tengan sentido.

RECONOCIMIENTO DE PATRONES

Disciplina encargada de identificar figuras, reconocer formar o leer patrones, con la finalidad de recopilar información sobre el objeto que se está estudiando y asignarlo a un grupo o clase.

Reconocimiento: Disciplina encargada de identificar figuras, reconocer formas o leer patrones, con la finalidad de recopilar información sobre el objeto que se está estudiando y asignarlo a un grupo o clase.

Patrón: Objeto que presenta cierto comportamiento y/o características ya conocidas.

Clase: Grupo de objetos que comparten los mismos atributos y comportamientos.

Atributo: Cualquier medida o cualidad extraíble. Constituye los parámetros, factores, variables y comportamientos.

RED ADVERSARIA GENERATIVA

Técnica de inteligencia artificial generativa que utiliza dos redes neuronales para generar nuevos datos que se asemejan a los datos de entrada. Una red genera datos falsos, mientras que la otra intenta distinguir entre los datos falsos y los reales. Compiten entre sí: la red generadora trata de producir datos que no se pueden distinguir de los datos reales, y la red discriminadora trata de identificar correctamente qué datos son falsos y cuáles son reales.

TOKEN

En el contexto de modelos de lenguaje grandes o LLMs, un token es una unidad básica de información para el procesamiento de texto. Estos tokens, por lo general, corresponden a palabras o frases, aunque también pueden ser signos de puntuación o números. La transformación del texto en cadenas de tokens, permite a los LLMs procesar los prompts de los usuarios y generar las respuestas correspondientes.

Un modelo como GPT-3.5 podría manejar 4.096 tokens o alrededor de 8.000 palabras. El número de tokens procesados en una entrada crece a medida que se entrenan LLMs de más complejidad.

TRANSFORMADOR PREENTRENADO GENERATIVO

Red neuronal profunda que permite procesar grandes cantidades de datos de texto y generar nuevo texto similar a los datos de entrada con los que se entrenó.

REFERENCIAS

1. Definición basada en el texto Corvalán Juan Gustavo, Díaz Dávila Laura, Simari Gerardo, Inteligencia Artificial: Bases conceptuales para comprender la revolución de las revoluciones, publicado en Tratado de Inteligencia Artificial y Derecho, dirigido por Corvalán Juan Gustavo, Tomo 1, Thomson Reuters, 2021.
2. Definición basada en el texto Corvalán Juan Gustavo, Díaz Dávila Laura, Simari Gerardo, Inteligencia Artificial: Bases conceptuales para comprender la revolución de las revoluciones, publicado en Tratado de Inteligencia Artificial y Derecho, dirigido por Corvalán Juan Gustavo, Tomo 1, Thomson Reuters, 2021.
3. Basado en Programa en línea de Machine Learning, Instituto Tecnológico de Massachussets, dictado el 23 de abril de 2019 al 18 de junio de 2019.
4. Definición basada en el texto Corvalán Juan Gustavo, Díaz Dávila Laura, Simari Gerardo, Inteligencia Artificial: Bases conceptuales para comprender la revolución de las revoluciones, publicado en Tratado de Inteligencia Artificial y Derecho, dirigido por Corvalán Juan Gustavo, Tomo 1, Thomson Reuters, 2021.
5. Basada en Yann LeCun, Yoshua Bengio & Geoffrey Hinton
<https://www.cs.toronto.edu/~hinton/absps/NatureDeepReview.pdf>
6. Basado en Programa en línea de Machine Learning, Instituto Tecnológico de Massachussets, dictado el 23 de abril de 2019 al 18 de junio de 2019.
7. Definición basada en el texto Corvalán Juan Gustavo, Díaz Dávila Laura, Simari Gerardo, Inteligencia Artificial: Bases conceptuales para comprender la revolución de las revoluciones, publicado en Tratado de Inteligencia Artificial y Derecho, dirigido por Corvalán Juan Gustavo, Tomo 1, Thomson Reuters, 2021.

8. Definición basada en el texto Corvalán Juan Gustavo, Díaz Dávila Laura, Simari Gerardo, Inteligencia Artificial: Bases conceptuales para comprender la revolución de las revoluciones, publicado en Tratado de Inteligencia Artificial y Derecho, dirigido por Corvalán Juan Gustavo, Tomo 1, Thomson Reuters, 2021.

9. Definición basada en MIT Technology Review, Douglas Heaven, Caja negra vs. de cristal: la IA que funciona contra la que se explica, 7 de febrero de 2020, disponible en: <https://www.technologyreview.es/s/11839/caja-negra-vs-de-cristal-la-ia-que-functiona-contra-la-que-se-explica>

10. Definición basada en ¿Qué es un chatbot? Oracle, disponible en: <https://www.oracle.com/ar/chatbots/what-is-a-chatbot/>

11. Definición basada en Felipe González, Teresa Ortiz y Roberto Sánchez, IA responsable, BID, 2020, disponible en: <https://publications.iadb.org/publications/spanish/document/IA-Responsable-Manual-tecnico-Ciclo-de-vida-de-la-inteligencia-artificial.pdf>

12. Basado en Programa en línea de Machine Learning, Instituto Tecnológico de Massachussets, dictado el 23 de abril de 2019 al 18 de junio de 2019.

13. Definición basada en Comisión Europea. Grupo de expertos sobre IA de alto nivel. Directrices éticas de una IA fiable, abril de 2019, disponible en: https://ec.europa.eu/commission/presscorner/detail/en/QANDA_21_1683#1

14. Definición basada en Arturo Muent Kunigami y Norma Palomino, Datos abiertos: conceptos básicos y temas claves, BID, 4 de marzo de 2019, disponible en: <https://blogs.iadb.org/conocimiento-abierto/es/datos-abiertos/>

15. Definición basada en el Reglamento General de Protección de Datos Personales.

16. Definición basada en FairLac, Marcelo Cabrol, Natalia González A, Cristina Pombo, Roberto Sánchez, Adopción ética y responsable de la Inteligencia Artificial en América Latina y el Caribe, Nota Técnica N° IDB-TN-1839, BID, enero 2020, página 8, disponible en: <https://publications.iadb.org/es/fair-lac-adopcion-etica-y-responsable-de-la-inteligencia-artificial-en-america-latina-y-el-caribe>

17. Definición basada en FairLac, Marcelo Cabrol, Natalia González A, Cristina Pombo, Roberto Sánchez, Adopción ética y responsable de la Inteligencia Artificial en América Latina y el Caribe, Nota Técnica N° IDB-TN-1839, BID, enero 2020, página 8, disponible en: <https://publications.iadb.org/es/fair-lac-adopcion-etica-y-responsable-de-la-inteligencia-artificial-en-america-latina-y-el-caribe>

18. Definición basada en FairLac, Marcelo Cabrol, Natalia González A, Cristina Pombo, Roberto Sánchez, Adopción ética y responsable de la Inteligencia Artificial en América Latina y el Caribe, Nota Técnica N° IDB-TN-1839, BID, enero 2020, página 8, disponible en: <https://publications.iadb.org/es/fair-lac-adopcion-etica-y-responsable-de-la-inteligencia-artificial-en-america-latina-y-el-caribe>

19. Definición basada en Definición de un modelo de aprendizaje automático, Microsoft, 1 de abril de 2019, disponible en: <https://docs.microsoft.com/es-es/windows/ai/windows-ml/what-is-a-machine-learning-model>

20. Procesamiento de lenguaje natural, Instituto de Ingeniería del Conocimiento, disponible en: <https://www.iic.uam.es/inteligencia/que-es-procesamiento-de-lenguaje-natural/>

Un asistente digital, también conocido como chatbot predictivo, es un programa informático avanzado que simula una conversación con las personas que lo usan, generalmente a través de Internet. Los asistentes digitales utilizan inteligencia artificial avanzada (IA), procesamiento del lenguaje natural, comprensión del lenguaje natural y aprendizaje autónomo para aprender a medida que avanzan y proporcionan una experiencia conversacional personalizada. Al combinar información histórica como las preferencias de compra, la propiedad de la vivienda, la ubicación, el tamaño de la familia, entre otros, los algoritmos pueden crear modelos de datos que identifican patrones de comportamiento y luego, refinar esos patrones a medida que se agregan datos. Al conocer el historial del usuario, sus preferencias y otra información, los asistentes digitales pueden responder preguntas complejas, proporcionar recomendaciones, hacer predicciones e incluso, iniciar conversaciones. Ver ¿Qué es un asistente digital? Oracle, disponible en: <https://www.oracle.com/ar/chatbots/what-is-a-digital-assistant/>

21. Ver más en Elena Bello, ¿Qué es un prompt en IA y cómo crear uno?, 8 de marzo de 2023, disponible en: <https://www.iebschool.com/blog/que-es-prompt-en-ia-tecnologia/>

22. Ver más en Chat GPT vs. GPT-4 ¿Imperfecto por diseño? Explorando los límites de la inteligencia artificial conversacional, Laboratorio de Innovación e Inteligencia Artificial de la Facultad de Derecho de la Universidad de Buenos Aires, año 2023, disponible en: <https://ialab.com.ar/wp-content/uploads/2023/03/ChatGPT-vs-GPT-4.pdf>

23. Ver más en Reconocimiento de patrones: qué es y cómo funciona, disponible en: <https://www.ceupe.com/blog/reconocimiento-de-patrones.html>