

UNA BREVE MIRADA AL ESTADO ACTUAL DE LA INTELIGENCIA ARTIFICIAL

BREVE HISTORIA DE LA IA

Con el desarrollo de la lógica matemática a finales del siglo XIX y principios del siglo XX se establecieron las bases de una formalización matemática de la idea de proceso mecánico o algoritmo. Tanto la lógica matemática como la teoría de la computación han sido fundamentales en el desarrollo de la inteligencia artificial, como se describe a continuación.

LA CONFERENCIA DE DARMOUTH, 1956:

Es considerada como el punto de partida de la inteligencia artificial moderna. Esta conferencia reunió a investigadores de diversas disciplinas para discutir la posibilidad de enseñar a los computadores a pensar como seres humanos.



A finales del siglo XX e inicios del XXI, la IA se concentró en desarrollar soluciones a problemas específicos, en lo cual jugó un rol fundamental el uso de técnicas de otras disciplinas como estadística, optimización y economía.



El aprendizaje de máquina tuvo un avance fundamental en la última década, en lo que es conocido como aprendizaje profundo. La disponibilidad de mayor poder de cálculo, en particular en la forma de unidades de procesamiento gráfico o GPUs, y el acceso a grandes volúmenes de datos digitalizados permitió el avance del aprendizaje profundo, y nos ha llevado a avances fundamentales en el procesamiento de lenguaje natural, como el desarrollo de ChatGPT.

EL PRIMER TRABAJO RECONOCIDO COMO IA

La neurona artificial de McCulloch y Pitts en 1943, que es una unidad de cálculo que intenta modelar el comportamiento de una neurona.

Estimulada por las discusiones de la conferencia, la investigación en inteligencia artificial de los años 50 comienza a desarrollarse en dos corrientes:

INTELIGENCIA ARTIFICIAL SIMBÓLICA (SYMBOLIC AI)

Es una rama que se enfoca en la utilización de técnicas de lógica matemática, inferencia y búsqueda para resolver problemas complejos y realizar tareas cognitivas, como la planificación, toma de decisiones, razonamiento deductivo e interpretación de datos.

ENFOQUE CONEXIONISTA

Se basa en el uso de redes neuronales artificiales para simular el funcionamiento del cerebro humano en la realización de tareas. Ha demostrado ser efectivo en áreas como reconocimiento de patrones, procesamiento del lenguaje natural y visión por computadora, y se ha utilizado en aplicaciones como reconocimiento de imágenes, diagnóstico médico y sistemas de recomendación.

APRENDIZAJE DE MÁQUINA

El aprendizaje de máquina es un área de la inteligencia artificial en la cual se desarrollan algoritmos que pueden aprender modelos desde datos, con el objetivo de que estos modelos puedan ser usados en predicciones posteriores. Por ejemplo, podemos pensar en un modelo meteorológico para predecir si va a llover.

Los componentes esenciales del aprendizaje de máquina son entonces:

1. Los datos de entrenamiento utilizados por los algoritmos de aprendizaje para construir modelos.
2. Los algoritmos de aprendizaje, que son métodos matemáticos utilizados para entrenar y mejorar los modelos.
3. Los modelos, que son representaciones matemáticas de las relaciones entre los datos utilizados para hacer predicciones.
4. Los métodos de evaluación y validación, que implican el uso de datos de prueba para medir el desempeño del modelo y su capacidad de generalización.
5. Los métodos de optimización y ajuste, que permiten la mejora continua de los modelos.

APRENDIZAJE PROFUNDO

El aprendizaje profundo (deep learning) es una subárea del aprendizaje de máquina en el cual los modelos generados por los algoritmos de aprendizaje son redes neuronales artificiales con múltiples capas. Tiene los mismos componentes esenciales que el aprendizaje de máquina, pero se pueden distinguir algunas características particulares:

1. Se necesitan grandes volúmenes de datos, dada la complejidad de estos modelos y los resultados precisos que busca.
2. Los algoritmos de aprendizaje son esencialmente algoritmos de optimización.
3. Los modelos son redes neuronales profundas, que son arquitecturas con múltiples capas de neuronas artificiales para aprender características y patrones complejos de los datos. Es importante considerar que dado el volumen de datos y la complejidad de los modelos, se necesita de hardware especializado para realizar el aprendizaje de los modelos.

PROCESAMIENTO DE LENGUAJE NATURAL

El procesamiento del lenguaje natural es un área de la inteligencia artificial que se enfoca en la interacción entre humanos y máquinas mediante el uso del lenguaje. Estudia tareas de análisis léxico, morfológico, sintáctico y semántico. El procesamiento del lenguaje natural facilita el desarrollo de sistemas de traducción automática, generación de resúmenes de texto y chatbots, entre otras aplicaciones.



APRENDIZAJE DE REPRESENTACIONES

El PLN ha experimentado avances significativos en la última década gracias al uso de word embeddings, que son representaciones de palabras que capturan su significado. Un word embedding es una representación que mapea cada palabra en un vector numérico denso de longitud fija. Este vector representa la semántica de la palabra, lo que significa que palabras similares se asignan a vectores similares.

car: [0.25, 0.68, 0.42, -0.11]
bus: [0.22, 0.71, 0.38, -0.08]
train: [0.23, 0.67, 0.45, -0.07]
plane: [0.13, 0.54, 0.62, 0.09]

Como se puede ver, los vectores para palabras relacionadas como "car", "bus" y "train" son similares entre sí, mientras que el vector para "plane" es distinto a los otros.

Sin duda el modelo generativo de texto más conocido hoy es Chat GPT. Chat GPT es una extensión de GPT-3 que fue entrenado para responder preguntas y generar respuestas en lenguaje natural. Una capacidad que tiene Chat GPT es la de incorporar al mecanismo generativo de texto una conversación, poniendo énfasis en la retroalimentación que obtiene en la interacción con el humano. Esto le permite generar respuestas coherentes en el contexto de una conversación. A este tipo de interacción humano-máquina se le denomina chatbot. Chat GPT incorporó un tipo de entrenamiento denominado aprendizaje reforzado para mejorar la calidad de sus respuestas. El aprendizaje reforzado incorpora retroalimentación durante el proceso de entrenamiento.



APLICACIONES DE LA INTELIGENCIA ARTIFICIAL

La disponibilidad de mayor poder de cálculo y el acceso a grandes volúmenes de datos digitalizados han permitido avances fundamentales en inteligencia artificial durante las últimas décadas, incluyendo adelantos a gran velocidad en las áreas de procesamiento de imágenes y procesamiento de lenguaje natural. Estos avances han permitido su aplicación en prácticamente todos los sectores productivos, servicios y áreas del saber.

INTELIGENCIA ARTIFICIAL EN LA SALUD

En particular se espera que la inteligencia artificial tenga un rol preponderante en:

- Generación y análisis de cantidades masivas de datos digitalizados generados por los pacientes.
- Aumento de la productividad de los profesionales de la salud mediante la automatización y simplificación de tareas.
- Reducción de costos mediante el desarrollo de procesos y tecnologías más eficientes y menos costosas.
- Mejora de la accesibilidad a servicios de salud mediante la automatización y simplificación de tareas y la atención médica remota en tiempo real.



APLICACIONES DE INTELIGENCIA ARTIFICIAL EN CHILE

En Chile, las implementaciones de sistemas de inteligencia artificial han ido avanzando e integrándose en diversas áreas de la sociedad. Esto refleja el desarrollo tecnológico del país y el avance en el uso de la IA como herramienta de apoyo.

Algunos ejemplos de estas aplicaciones en nuestro país incluyen un sistema de identificación de transacciones anómalas e irregularidades implementado por el Servicio de Impuestos Internos (SII), y una aplicación con IA para sistematizar audiencias de control de detención, implementada por la Defensoría Penal Pública del Estado.

ASPECTOS ÉTICOS DE LA INTELIGENCIA ARTIFICIAL

En los últimos años, la ética se encuentra en el centro de las preocupaciones académicas y sociales de la IA, contribuyendo a la creación de principios éticos y líneas de investigación que se hacen cargo fundamentalmente de los riesgos y el impacto socio tecnológico de esta tecnología.

Debido a esta creciente preocupación por el impacto social y las consecuencias del desarrollo e implementación de la IA en distintos ámbitos de la sociedad, se han propuesto variados principios para guiar las prácticas profesionales, institucionales, y gubernamentales que sustentan la IA. Se destacan aquí 4 principios éticos principales en base a la sistematización de cuatro influyentes artículos de investigación:

REFLEXIONES FINALES

El acelerado avance de la IA en los últimos cinco años ha revolucionado el quehacer científico y también las diferentes profesiones. Es por esto que resulta no sólo necesario, sino que fundamental, incluir el conocimiento sobre este tipo de tecnologías en el aula universitaria. Este conocimiento, sin embargo, no debe limitarse a la mera aplicación o uso de diferentes herramientas de IA, ya sea por parte de los docentes como por el mismo estudiantado.

TRANSPARENCIA: Este es el principio más recurrente. Se entiende en términos generales desde una respuesta a uno de los mayores desafíos que plantea la IA desde el punto de vista de gobernanza y funcionalidad: la opacidad.

JUSTICIA: Se basa fundamentalmente en la prevención de sesgos algorítmicos. Hay muchas maneras en las cuales diferentes sesgos pueden entrar en los sistemas de IA, influyendo en entrenamientos no representativo o inclusive por sesgos cognitivos que traen los humanos en el bucle.

PRIVACIDAD: Este principio se relaciona a prácticas de protección de datos y medidas de seguridad. Se busca garantizar la privacidad de datos a través del ciclo de desarrollo de la IA, así como también proveer el derecho de control de la información a los ciudadanos.

RESPONSABILIDAD: El concepto de responsabilidad se relaciona con aspectos legales, morales, personales, colectivos, y ecosistémicos. Llama a asegurar la prevención de riesgos y daños causados tanto a individuos, como a ecosistemas. Esto implica minimizar acciones y decisiones que no tengan asociadas responsabilidades directas, tanto en procesos de desarrollo como en implementación y decisiones operacionales.

PRINCIPIOS ÉTICOS

La inclusión de la IA a las aulas debe integrar una visión crítica y reflexiva, que permita dimensionar las limitaciones y posibles riesgos que la tecnología conlleva.