

Tipos de Inteligencia Artificial

El estado actual de la evolución de esta tecnología comprende un campo tan amplio con tantas ramas, que es bastante complicado poder atender a una única clasificación. Las diferentes clasificaciones que se pueden encontrar dependen de la visión con la que aborda dicho concepto. Las hay más filosóficas, más técnicas, y según su aplicación. A lo largo de este tema vamos a conocer las más comunes y aceptadas.

Categorías de la Inteligencia Artificial

Stuart J. Russell y Peter Norvig, investigadores informáticos, publicaron en 1995 su libro “Artificial Intelligence: A Modern Approach”, que se ha convertido en un libro de texto fundamental en cientos de universidades a nivel mundial y que ya va por su cuarta edición. En él, comienzan por plantear cuatro categorías básicas para sistemas de inteligencia artificial. Ellos parten de un enfoque de génesis del acto inteligente, del origen y proceso por el cual se llega al comportamiento inteligente.

Una primera categoría está configurada por sistemas que piensan como humanos. Estos sistemas tratan de emular el pensamiento humano. Son sistemas que deben resolver tareas como la toma de decisiones, la resolución de problemas y el propio paradigma del aprendizaje.

Otra categoría estaría formada por sistemas que actúan como humanos, es decir, sistemas que deben imitar el comportamiento humano, aunque no el pensamiento que conduce hacia dicho comportamiento. Es una aproximación más práctica. Sería el caso de la robótica y sistemas de actuadores en el mundo físico.

En un salto hacia el razonamiento como algo con entidad propia, más allá de cómo lo hacemos los humanos, se plantea la categoría de sistemas que piensen racionalmente. Sería la forma más perfecta de inteligencia artificial, pues se basaría en el cumplimiento exacto de las leyes del razonamiento lógico. Sería el caso de los sistemas expertos. Los sistemas dentro de esta categoría requieren muchísima capacidad computacional, y es prácticamente imposible llegar a contar con un sistema de este tipo para un campo de aplicación general. Solo son posibles para campos muy especializados y acotados.

Por último, estaría la categoría de sistemas que actúan racionalmente. Estos sistemas tratan de emular de forma racional el comportamiento humano. Un ejemplo de este tipo de sistemas son los agentes inteligentes.

El último capítulo del libro de Russell y Norvig tiene como título una pregunta: “¿Qué pasaría si lo conseguimos?” y plantea algunas reflexiones sobre cómo afectaría a la motivación y el trabajo de los profesionales del campo de la computación y la inteligencia artificial el conseguir realmente una inteligencia artificial superior, que fuese más capaz que el ser humano en todo (luego hablaremos del concepto de inteligencia artificial general).

Dos tipos de inteligencia artificial según las tareas a resolver

En una aproximación simple y basada en los objetivos que tenemos de cara a que la inteligencia artificial nos ayude a ejecutar y resolver tareas de nuestros procesos cotidianos, se distinguen dos tipos de IA:

Inteligencia Artificial Fuerte

También conocida como **Inteligencia artificial** general o IAG es la **inteligencia artificial** que iguala o excede la **inteligencia** humana promedio, es decir, la **inteligencia** de una máquina que puede realizar con éxito cualquier tarea intelectual de cualquier ser humano.

Este tipo de IA es la que sería capaz de analizar cualquier situación y deducir el conjunto de acciones más adecuado para dicha situación y contexto. Lo mismo sabría conducir un coche, que resolver una ecuación matemática o mantener una conversación sobre un tema concreto.

Actualmente no existe este tipo de IA, pero ciertamente, los esfuerzos de todos los agentes del sector de los que hemos hablado ya en este tema, van fuertemente dirigidos a conseguir este paradigma. Al menos, se está trabajando en conseguir una IAF en el campo de los asistentes virtuales. Si ya quisiésemos llevarlo al ámbito físico, sería necesario contar también con la robótica y mecatrónica.

Inteligencia Artificial Débil

También conocido como **IA estrecha**, se define como la inteligencia artificial racional que se centra típicamente en una tarea estrecha. La inteligencia de la IA débil es limitada.

Los asistentes virtuales son un buen ejemplo de la inteligencia estrecha. Siri opera dentro de un rango limitado previamente definido, no hay ninguna inteligencia genuina, sin conciencia, sin vida, a pesar de ser un ejemplo sofisticado de IA débil.

El investigador de IA, **Ben Goertzel**, indicó en su blog en el 2010 que Siri era "muy estrecho y frágil" evidenciada por resultados molestos si se le hacían preguntas fuera de los límites de la aplicación.

Algunos comentaristas piensan que la IA débil podría ser peligrosa. En 2013 **George Dvorsky** declaró: "La IA estrecha podría noquear nuestra red eléctrica, dañar plantas de energía nucleares, causar un colapso económico a escala global, desviar los vehículos autónomos y a los robots..."

Una IA débil, al no tomar en consideración todo un contexto amplio, ni seguir las reglas sociales, éticas,... ejecuta las tareas para las que se le ha entrenado con eficacia y contundencia. No evalúa las consecuencias como lo hacemos los humanos, considerando un espectro amplio de efectos y relaciones. Por eso, es una opción incompleta, inestable y peligrosa si no se utiliza con prudencia.

Escuelas de pensamiento

Por otro lado, también podemos encontrar, en el ámbito de la inteligencia artificial moderna, dos escuelas de pensamiento:

- Inteligencia artificial convencional
- Inteligencia artificial computacional

Estas dos escuelas difieren en la ciencia que hay tras los procesos que siguen para llegar a los resultados esperados. Pero, con los avances que se están dando en los recursos que utiliza la segunda, muchas de las aplicaciones que tenía la primera, están siendo llevadas al campo computacional.

Inteligencia artificial convencional

Se conoce también como IA simbólico-deductiva. Está basada en el análisis formal y estadístico del comportamiento humano ante diferentes problemas:

- Razonamiento basado en casos: Ayuda a tomar decisiones mientras se resuelven ciertos problemas concretos y, aparte de que son muy importantes, requieren de un buen funcionamiento.
- Sistemas expertos: Infieren una solución a través del conocimiento previo del contexto en que se aplica y ocupa de ciertas reglas o relaciones.
- Redes bayesianas: Propone soluciones mediante inferencia probabilística.
- Inteligencia artificial basada en comportamientos: Esta inteligencia contiene autonomía y puede auto-regularse y controlarse para mejorar.
- Smart process management: Facilita la toma de decisiones complejas, proponiendo una solución a un determinado problema al igual que lo haría un especialista en dicha actividad.

Esta rama de la IA ha sido la que ha proporcionado la mayoría de algoritmos que conocemos como “automatización”, y básicamente se sirven de sistemas con reglas condicionales y estadística avanzada.

Inteligencia computacional

La Inteligencia Computacional (también conocida como IA subsimbólica-inductiva) implica desarrollo o aprendizaje interactivo (por ejemplo, modificaciones interactivas de los parámetros en sistemas de conexiones). El aprendizaje se realiza basándose en datos empíricos, utilizando métodos computacionales inspirados en procesos de la naturaleza, que permiten alcanzar soluciones aptas a problemas complejos que los modelos tradicionales no pueden resolver por no existir una solución analítica, por no contar con todos los parámetros necesarios o porque el problema es en sí estocástico y precisa de una aproximación envolvente en vez de convergente.

Esta corriente será la que impulse lo que conocemos como “Aprendizaje Automático” o “Machine Learning”.

Algunas técnicas de esta escuela son:

- Máquina de vectores soporte: sistemas que permiten reconocimiento de patrones genéricos de gran potencia.
- Redes neuronales: sistemas basados en redes de unidades de computación lineal para simular computación no lineal
- Modelos ocultos de Markov: aprendizaje basado en dependencia temporal de eventos probabilísticos.
- Sistemas difusos: técnicas para lograr el razonamiento bajo incertidumbre
- Computación evolutiva: también conocidos como algoritmos genéticos, aplica conceptos inspirados en la biología, tales como población, mutación y supervivencia del más apto para generar soluciones sucesivamente mejores para un problema.

La clasificación de Hintze

En noviembre de 2016, Arend Hintze, profesor de la Universidad de Michigan e investigador en el campo de la inteligencia artificial, escribió un artículo “Understanding the four types of AI, from reactive robots to self-aware beings”, en el que sintetizaba toda la evolución de los últimos desarrollos y avances en materia de inteligencia artificial para aportar una clasificación más realista y concreta para los tipos de entidades que existen o que se aspira a crear.

Según esta nueva clasificación, tenemos cuatro tipos de IA:

1. Máquinas reactivas

Los tipos más básicos de sistemas de IA son puramente reactivos. No tienen la capacidad de formar recuerdos. Tampoco pueden utilizar experiencias pasadas en las que basar las decisiones actuales.

Deep Blue fue una supercomputadora creada por IBM. Fue capaz de vencer al ajedrez al gran maestro internacional Garry Kasparov. Ocurrió a fines de la década de 1990 y es el ejemplo perfecto de este tipo de máquina. Puede identificar las piezas en un tablero de ajedrez y saber cómo se mueve cada una. Puede realizar predicciones sobre los mejores movimientos y elegir el mejor de todas las posibilidades.

Pero no tiene ningún concepto del pasado. Tampoco posee recuerdos de lo que ha sucedido antes. Aparte de una regla de ajedrez, Deep Blue ignora todo antes del momento presente. Todo lo que hace es enfocar las piezas del tablero en tiempo real y elegir entre los siguientes movimientos posibles.

Es importante que el usuario sepa que está tratando con una máquina en una conversación de texto o voz, y evitar crear falsas expectativas sobre lo que puede esperar de dicha conversación.

2. Memoria limitada

El Tipo II maneja máquinas que pueden mirar hacia el pasado. Los vehículos autónomos ya hacen algo parecido. Por ejemplo, observan la velocidad y dirección de otros autos.

Para que funcionen así hay que identificar objetos específicos y monitorearlos a lo largo del tiempo.

Digamos que estas observaciones se agregan a las representaciones preprogramadas para la memoria de estos coches. Se incluyen marcas de carril, semáforos y otros elementos importantes, como curvas en la carretera.

También se añaden experiencias como cuando el automóvil decide en qué momento cambiar de carril para evitar interrumpir a otro conductor o ser embestido por un automóvil cercano.

Pero estas simples piezas de información sobre el pasado son solo transitorias. No se guardan como parte de la biblioteca de experiencias del automóvil. En estos tipos de inteligencia artificial, la máquina no puede compilar la experiencia durante años, como lo hace un humano.

Entonces, ¿cómo podemos construir sistemas de IA que construyan representaciones completas, recordar sus experiencias y aprender cómo manejar situaciones nuevas?

3. Teoría de la mente

Llegamos a un punto en el que nos acercamos más a los tipos de inteligencia artificial que deseamos en un futuro. Las máquinas de la siguiente clase son más avanzadas. No solo forman representaciones sobre el mundo, también sobre otros agentes o entidades.

En psicología, esto se denomina 'teoría de la mente'. Implica la comprensión de que las personas, las criaturas y los objetos en el mundo pueden tener pensamientos y emociones que afectan a su propio comportamiento. Esto es crucial para la forma en que los humanos formamos sociedades, porque nos permite la interacción social.

Si las máquinas van a andar entre nosotros, deberán tener una comprensión sobre cómo pensamos y cómo sentimos. Además deberán llegar a saber qué esperamos y cómo queremos que nos traten. Tendrán que ajustar su comportamiento en consecuencia.

4. Autoconciencia

El paso final del desarrollo de la IA es construir sistemas que puedan formar representaciones sobre sí mismos. En última instancia, los investigadores de la IA tendrán que comprender no solo la conciencia, sino también construir máquinas que la tengan.

Los seres conscientes son conscientes de sí mismos, conocen sus estados internos y pueden predecir los sentimientos de los demás. Es probable que estemos lejos de crear máquinas que sean conscientes de sí mismas. Sin embargo, los esfuerzos se enfocan hacia la comprensión de la memoria, el aprendizaje y la capacidad de basar las decisiones en experiencias pasadas.

Este es un paso importante para entender la inteligencia humana por sí misma. Es crucial para diseñar o desarrollar máquinas que sean más excepcionales para clasificar lo que ven frente a ellas.

Los cuatro tipos de inteligencia artificial dan una idea sobre las intenciones que el hombre tiene acerca del futuro de la máquina. Puede que estemos muy lejos de la IA autoconsciente. No obstante, está claro que eso es lo que se persigue en última instancia.

Campos de aplicación de la inteligencia artificial

Actualmente hay cuatro áreas de desarrollo en las que la inteligencia artificial está jugando un papel clave y que van a revolucionar multitud de procesos, cambiando muchas de las tareas que se realizan, tanto en la vida cotidiana como en el entorno profesional, de forma más eficiente y eficaz.

Estos campos son: el reconocimiento de imagen, el procesamiento de lenguaje natural, los modelos predictivos y los modelos generativos.