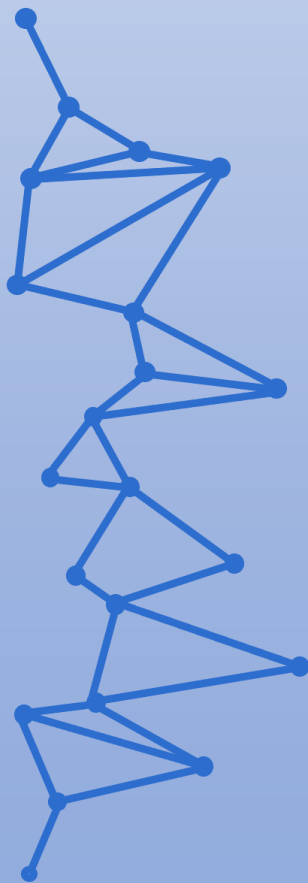




Curso de Especialización de Inteligencia Artificial y Big Data (IABD)



Modelos de Inteligencia Artificial

UD02. Tipos de IA. Escuelas y clasificaciones.
Resumen.

JUAN ANTONIO GARCIA MUELAS

Modelos de Inteligencia Artificial UD02. Tipos de IA. Escuelas-Clasificaciones

Aunque aún no existan aplicaciones para alguna clase de las definidas para la Inteligencia Artificial, podemos clasificarlas:

Tipos de Inteligencia Artificial según distintas clasificaciones



Esta primera **clasificación se basa en las tareas que ayuda a ejecutar**. Es una aproximación bastante simple, con dos opciones:

- ✓ **IA Débil**. También conocida como **IA estrecha**, se define como la IA racional que **se centra en una tarea estrecha (problemas concretos y acotados)**. Es **limitada** e incapaz de adaptarse o asumir cambios.

Desde el punto de vista de esta clasificación (por tarea a resolver), las **características** de la Inteligencia Artificial débil son:

- **Ya existen en la vida real:** Como hemos comentado, los **asistentes virtuales**, programas como **Watson o Alpha Go** que vimos en la unidad anterior.
- Se orientan a **resolver problemas muy concretos**: El programa que "sabe" jugar al Go, no sabe hacer otra cosa. Ni tiene posibilidades de aprender a jugar a otra cosa, por muy similar que sea.
- **Son reactivas: No tienen iniciativa**, es necesario que se desencadene la acción que tienen programada para que se inicie su rutina. En el ejemplo del asistente virtual, tiene programado responder cuando le preguntas, y por tanto nunca tomará la iniciativa de ofrecerte nada sin que tú lo actives previamente.
- **No son flexibles: Colapsan si se encuentran en un caso no previsto** en su programación.
- Quedan **limitadas por lo que programa un humano**: Es el humano quien programa lo que "tiene que pensar" la máquina. Si el humano no programa deja sin considerar ciertas opciones o posibles situaciones, la IA nunca será capaz de suplirlo o aprenderlo sobre la marcha por sí misma.
- Se programan **con pocas redes neuronales**: Hablaremos más adelante sobre las redes neuronales. Por el momento es suficiente entender que el nivel de

computación compleja que requieren este tipo de Inteligencias Artificiales es menor que otros casos.

- **No razonan, solo computan:** No tienen en cuenta ningún factor moral, contextual, circunstancial, emocional... que a un humano le haría reaccionar de manera diferente. La máquina está programada para alcanzar tal objetivo o funcionar de tal manera, y así lo hará sin "entender" lo que está haciendo. Por tanto: no tiene conciencia.
- **Aprenden a base de ejemplos:** Necesita conocer muchos ejemplos de lo que tiene que hacer (la base de datos), con todas las variantes posibles. Por ejemplo, en la máquina que juega al Go, se la "entrenó" con 50 millones de partidas de dicho juego.
- **Son repetitivas:** No se cansan nunca, son implacables, siempre la misma rutina.
- **No salen de su marco de trabajo:** Y esto supone que pueden ocuparse de tareas mecánicas, repetitivas, "aburridas" para sustituir al humano mejorando rendimiento y precisión. Pero **necesitará siempre una supervisión humana** que vaya decidiendo cómo adaptar el programa a las cambiantes circunstancias.

Según se mire la Inteligencia Artificial débil podría llegar a ser peligrosa, al no tomar en consideración todo un contexto amplio, ni seguir las reglas sociales, éticas... ejecuta las tareas con eficacia y contundencia, ni evaluar 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, o si quien la programa pretende, precisamente, causar mal.

- ✓ **IA Fuerte.** La Inteligencia Artificial Fuerte (IAF) o **general (IAG)**, desde el punto de vista de la tarea, sería **aquella que iguala o excede la inteligencia humana promedio**.

En comparación con la Inteligencia Artificial débil, las **características** de la fuerte son:

- **No existe** en la realidad.
- **Resolverán problemas abiertos:** Deberían poder abarcar múltiples posibles tareas, distintas unas de otras (reparar una puerta, ir a recoger a los niños del colegio, regar las plantas, darte conversación...).
- **Serán proactivas:** En función de la misión u objetivo que tenga, y de las circunstancias, iniciará cualquier tipo de rutina sin esperar a que un humano se lo pida o esté pendiente.
- Serán **flexibles:** Podrán encontrar similitudes entre algo que conocen y algo que se le parezca un poco. Por ejemplo, aunque inicialmente solo haya sido programada para saber andar, será capaz de aprender a correr sin necesidad de intervenir en su programa.
- **Se autoprogramarán:** Serán capaces de detectar sus propios límites y se regularán a sí mismas para no excederlos.
- Usarán **muchas redes neuronales:** Y además podrán entrar en conflicto entre ellas en algunas ocasiones. Esto quiere decir que necesitarán una capacidad de almacenaje de información y cómputo que aún hoy no hemos llegado a alcanzar.
- **Imitarán el comportamiento humano:** Serán **capaces de razonar**, y por tanto, de alcanzar **algún tipo de consciencia**.
- **Aprenderán como las personas:** Podrán recordar datos, observar nuevas situaciones y encontrar relaciones entre diferentes acciones. Esto quiere decir que si saben jugar al ajedrez y "observan" el juego de las damas, podrán aprender a jugar a las damas basándose en lo que saben sobre jugar al ajedrez.

- Serán capaces de **aprender nuevas tareas**: Modificarán la tarea o cómo realizan la tarea para adaptarse a las circunstancias.
- Serán capaces de **adaptarse a nuevos escenarios**: Podrán adaptarse a cambios y nuevas situaciones para seguir cumpliendo su objetivo.

En el ámbito de la Inteligencia Artificial más moderna podemos encontrar **dos escuelas de pensamiento**:

- ✓ Inteligencia Artificial **Convencional**, o simbólico-deductiva. Basada en el **análisis formal y estadístico**. Es la que **ha proporcionado la mayoría de los algoritmos** que conocemos como “**automatización**”, y básicamente se sirven de **sistemas con reglas condicionales y estadística avanzada**. Ej: Sistemas expertos, redes bayesianas, Smart process management...
- ✓ Inteligencia Artificial **Computacional**, o subsimbólica inductiva. Implica desarrollo y aprendizaje interactivo. Es la corriente que **impulsó el “Machine Learning”**. Algunas técnicas son: Máquina de vector soporte, redes neuronales, Modelos ocultos de Markov, Sistemas difusos o Computación evolutiva.

Tipos de Inteligencia Artificial según Hintze

En 2016 publicó un artículo en el que sintetizaba la evolución y avances de la IA, aportando una clasificación más realista y concreta. Es por tanto la tipología más reciente.



Para saber más

Stuart J. Russell y Peter Norvig, investigadores informáticos, publicaron en **1995** su libro “[Artificial Intelligence: A Modern Approach](#)”, que se ha convertido en el libro de texto fundamental en cientos de universidades a nivel mundial (ya lleva varias ediciones publicadas).

Plantean cuatro categorías básicas partiendo de un enfoque de génesis del acto inteligente, del origen y proceso por el cual se llega al comportamiento inteligente.

- **Sistemas Cognitivos**: Piensan como humanos, **intentan emular el proceso humano** → Proceso de toma de decisiones, resolución de problemas, y el propio paradigma del aprendizaje.
- **Test de Turing**: Actúan como humanos, intentan emular el comportamiento humano (sin pasar por el pensamiento o razonamiento que conduce a dicho comportamiento) → A nivel práctico **se aplica en la robótica y sistemas de actuadores** en el mundo físico.
- **Leyes del pensamiento**: Piensan **con razonamientos**. Cumplimiento exacto de las leyes del razonamiento lógico, teniendo en cuenta todos los factores que afectan a la cuestión. **No hemos llegado a esto aún. Sería el caso de los sistemas expertos**. Solo son posibles aproximaciones para campos de investigación muy especializados y acotados.

- **Agentes inteligentes:** Actúan racionalmente (sin pasar por el proceso de razonamiento lógico).

Los dos últimos requieren una capacidad de cómputo muy importante, a veces, aún, inaccesible.

Tipos de IA según Russell y Norvig



MÁQUINAS REACTIVAS.

Son los **tipos más básicos de sistemas de IA**. Deep Blue, es capaz de ganar al ajedrez, pero no tiene concepto del pasado. Enfoca las piezas y elige un movimiento.

MEMORIA LIMITADA.

Pueden mirar hacia el pasado (Coche autónomo. Observa los parámetros de otros vehículos y en base a la memorización toma decisiones). Sin embargo, es una **información transitoria**, que no guarda durante años como puede hacer un humano.

TEORÍA DE MENTE.

No existen aún. Las máquinas de esta clase son más avanzadas. No solo **forman representaciones sobre el mundo**, también sobre **otros agentes o entidades**.

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.