

FORO PARA EL FUTURO DEL EMPLEO

Participantes:

Bentolilla, Samuel – CEMFI

Jorge Sicilia – BBVA

Cardoso Lecourtois, Miguel – BBVA

Cruz Palacios, Juan Manuel – ACCIONA

Conde-Ruiz, Ignacio – UCM

Doménech, Rafael – BBVA

García López, Juan Ramón – BBVA Research

Godino Reyes, Martín – SAGARDOY

Gutiérrez, Fernando – BBVA

Martín, Juan Luis – Prosegur

Neut, Roberto Alejandro – BBVA

Sicilia, Jorge – BBVA

Suárez Zarcos, Luis – FCC

PONENTE INVITADO: Ramón López de Mantaras, *Director of the IIIA (Artificial Intelligence Research Institute) of the CSIC (Spanish National Research Council)*



ACTA DE LA SESIÓN

INTELIGENCIA ARTIFICIAL Y EFECTOS SOBRE EL MERCADO DE TRABAJO

¿Qué es la Inteligencia Artificial? Es el área de la informática que estudia el comportamiento inteligente mediante métodos computacionales.

A pesar de todo lo que se está oyendo y leyendo, la verdad es que quedan muchas cosas por hacer para que podamos hablar de “inteligencia” de verdad. Es más, primero habría que definir qué se entiende como “comportamiento inteligente”, ya que ni siquiera los neurocientíficos ni los psicólogos cognitivos saben exactamente lo que es, ni se ponen de acuerdo en una definición. Lo que sí está claro es que se trata de una actividad emergente de la actividad neuronal, pero lo que no se sabe es cómo emerge.

Con todo, la inteligencia podría describirse ampliamente como la capacidad para resolver problemas complejos, lo cual se lleva a cabo gracias a las siguientes capacidades o componentes:

- 1) La percepción (los sentidos): vista, oído, olfato, gusto y tacto.
- 2) Acción: la selección de acciones para lograr objetivos.
- 3) Capacidad de razonamiento y planificación.
- 4) Aprendizaje automático.
- 5) Comunicación: lenguaje natural y gestual. En concreto, el lenguaje gestual es tan importante como el natural ya que, entre otras cosas, nos sitúa en el contexto.

Por ejemplo, un robot debería ser capaz de comprender lenguaje gestual para poder interactuar con el entorno adecuadamente. Habría que enseñar a las máquinas a adaptarse al contexto social. La importancia del lenguaje gestual en los sistemas de IA se ve, por ejemplo, en el comportamiento de los coches autónomos ante situaciones como la de los cruces de cuatro stop típicos en EEUU. Allí se entiende que pasa el que llega primero, sin embargo, en los casos en los que varios coches llegaban al mismo tiempo, al no poder intercambiar señales gestuales, se quedaban allí parados indefinidamente.

Es por todo ello que un sistema de IA debería integrar todos estos elementos, de ahí que en IA se habla de Sistemas Integrados. Y en la actualidad existen muy pocos casos de sistemas integrados que consigan coordinar todos estos procesos de forma sencilla. Donde más podemos ver esta integración, aunque de manera parcial, es en las inteligencias artificiales corpóreas, es decir, en los robots con cuerpo (no necesariamente humanoide, sino en aquellos que tienen partes; excluyendo, por supuesto, los robots industriales de modelos fordistas).

A día de hoy una arquitectura cognitiva general capaz de integrar de forma eficiente todos los componentes que permitan implementar todas las capacidades no existe. Esto significa que todos los ejemplos que tenemos actualmente de IA se refieren a una

IA específica: sistemas que realizan una única tarea y dentro de un contexto determinado. Un ejemplo son sistemas de *deep learning* que se basan en el reconocimiento de patrones o regularidades presentes en los datos de entrenamiento. Por ejemplo, la realización de diagnósticos por imagen. La IA hace este reconocimiento de patrones de forma altamente eficaz. Sin embargo, todo lo que se salga de esos patrones, no servirá, ya que la red neuronal no lo reconocerá.

Las distintas investigaciones aplicando el *deep learning* nos muestran una gran efectividad de la IA frente a la capacidad humana en entornos muy específicos. Por ejemplo, un grupo de investigación de la Universidad de Queensland (Australia) ha aplicado *deep learning* para analizar de forma combinada imágenes médicas de rayos X, MRI y ultrasonidos y ha logrado diagnosticar el cáncer de mama mejor que los médicos.

Otra investigación, esta vez de la universidad de Stanford, ha llevado a conseguir que una IA consiga identificar 3 indicadores nuevos para evaluar la posibilidad de que una biopsia de células mamarias pudiera resultar positiva en base a analizar grandes cantidades de datos de muestras de células mamarias cancerosas y normales. Hasta ahora contábamos con 8 indicadores que se estudiaban y ahora contamos con 11 gracias a los análisis de este sistema de IA.

Finalmente, en la Carnegie Mellon University se analizaron 133.000 pacientes de 4 hospitales de Chicago y se logró hacer una predicción de la posibilidad de infarto de los pacientes ingresados en la UCI con hasta 4 horas de antelación (frente a la antelación máxima de 30 minutos que lograban los cardiólogos).

DESAFÍOS: ¿QUÉ PODEMOS MEJORAR?

El gran desafío, como ya hemos mencionado, es el de conseguir pasar de una IA específica a una genérica. Y es que los algoritmos de aprendizaje no aprenden realmente. Las experiencias más recientes muestran claramente que la IA es, en términos generales, mejor que la humana para hacer algunas tareas muy específicas. Es más, si después de haber entrenado una red neuronal para hacer algo muy bien, luego la entrenamos para hacer otra cosa, también la aprenderá a resolver muy bien pero olvidará por completo lo anteriormente aprendido. Es decir, cuando un sistema de IA aprende una determinada tarea (por ejemplo, jugar al ajedrez) olvida lo que había aprendido previamente (por ejemplo, jugar a Go). Esto es lo que se llama “olvido catastrófico”.

Pero, ¿por qué ocurre esto con la IA y no nos pasa a nosotros los seres humanos?. Un posible explicación es que el ser humano tiene un cerebro con una cantidad de neuronas de 10^{11} y muchas más sinapsis lo que facilita acumular el aprendizaje de multitud de tareas y competencias de forma simultánea aunque nadie conoce los detalles de cómo lo hace. Para lograr eso mismo con las máquinas más avanzadas necesitaríamos millones de años de entrenamiento ya que cuanto más grande y compleja es la red más datos se necesitan para entrenarla. Para que una red neuronal artificial no olvidara sería necesario compartimentar el aprendizaje de las distintas tareas logrando una red enorme de módulos de aprendizaje separados. Pero cada módulo debería ser independiente de los demás módulos porque se ha demostrado que, si hay conectividad entre ambos, se afectan entre ellos y se pierde algo de cada uno, es decir hay un porcentaje de olvido. Por otra parte, tampoco se tiene claro si el coste de reentrenar ese porcentaje perdido es menos costoso que el de entrenarlo completamente. En principio pudiera parecer que debe ser así, pero no se sabe.

En nuestro cerebro no hay absoluta modularidad pues sabemos que varias partes del cerebro están activas al mismo tiempo e interactuando cuando llevamos a cabo alguna actividad y sin embargo no sufrimos el fenómeno del olvido catastrófico. Hay quien ha sugerido que un sistema de IA general se podría lograr mediante la unión de multitud de sistemas específicos. El problema es como lograr una arquitectura que una de forma eficiente todos los componentes específicos, es decir sin que ello reste efectividad al conjunto. Actualmente nadie sabe como lograrlo por lo que hoy por hoy por motivos de eficiencia es mejor resolver problemas concretos mediante sistemas de IA específicos

De ahí que para avanzar hacia la inteligencia artificial más general necesitamos progresos importantes no solo en prácticamente todos los componentes de la IA sino sobre todo, en su integración para mejorar su funcionamiento y acercarla a la de ser humano.

Finalmente, y como ya se dijo al principio, necesitamos una comprensión profunda del lenguaje y del reconocimiento visual. Respecto del lenguaje, el ejemplo paradigmático del avance lo encontramos en *Google Translate*, que aprende muy rápido sin necesidad de análisis profundo de la semántica, lo cual parecía impensable hace un tiempo, pero que ha sido posible gracias a la cantidad de correcciones de los usuarios que permite el análisis masivo de patrones creando una estadística a gran escala. Con todo, no podemos decir que este sistema tenga verdadera comprensión del lenguaje, sino que en lugar de analizar semánticamente el lenguaje lo que hace es detectar patrones (secuencias de palabras) en el lenguaje fuente que son recurrentemente presentes en los grandes corpus de traducciones a los que tiene acceso y reproducir la secuencia equivalente de palabras en el lenguaje de destino .