

El despertar de una inteligencia global

- Prólogo
- La inteligencia artificial hasta la primera década del siglo XXI
- El verano de la inteligencia artificial.
- Complementos a la IA y su relevancia en el campo de la medicina
- En el ámbito médico
- Epílogo
- Bibliografía

Prólogo

Excmo. Sr. presidente;

Excmos. Sres. presidentes de honor;

Excmas. e Ilmas. autoridades;

Muy Ilmas. e Ilmos. Sras. y Sres. académicos;

querida familia, compañeros, amigos.

Es un gran honor dirigirme a ustedes en el Paraninfo de nuestra Universidad, que ha acogido tan magníficos eventos, después de haber sido invitado y haber aceptado formar parte de la Real Academia de Medicina de Salamanca. Un honor que me llena de orgullo y que acepto con inmensa gratitud.

Por ello, agradezco en primer lugar a la Junta de Ilmos. Académicos Numerarios el haberme concedido su confianza, permitiéndome estar ante ustedes hoy para leer este discurso que me incorpora de lleno a su seno como académico numerario.

Muchísimas gracias.

Este nombramiento es para mí tanto un reto como un desafío y un estímulo para los que espero estar a la altura y, sobre todo, no defraudar a los que han creído en mí e impulsado mi candidatura. Por ello, tengo que agradecer a los académicos de número, Ilmos. Sres. don Francisco Santiago Lozano Sánchez, don Alberto Gómez Alonso y don Eugenio Miguel Ángel Santos de Dios, por haber avalado la presentación de mi candidatura. Al primero tengo también que agradecerle la contestación a mi discurso de ingreso y ser el impulsor de mi entrada en la Academia.

Aunque mi incorporación se lleva a cabo a título individual, yo estoy aquí gracias a la labor de todos los que han trabajado y trabajan conmigo en el grupo de investigación de la Universidad, en el centro tecnológico que hemos creado y en las numerosas instituciones con las que he colaborado y sigo haciéndolo.

Desde que tengo uso de razón me he sentido atraído por la medicina y es algo que siempre ha estado muy presente entre mis intereses intelectuales. Aunque al final y en el último momento me decanté por la informática, cuando conseguí el grado de doctor tuve la sensación de estar un poquito más cerca

de mi sueño de infancia. ¡No era médico, pero sí doctor! Hoy y en este momento, y salvando las distancias, estoy dando un paso más en la dirección correcta. Muchas gracias de nuevo. Tengo la certeza de que a partir de ahora mi vida investigadora estará aún más vinculada a la medicina y podré seguir aprendiendo con los que más saben de estos temas.

Aunque me gusta mucho planificar y adelantarme al futuro, tiendo a cambiar mis planes con cierta frecuencia por impulsos, imponderables o nuevas oportunidades, siempre siguiendo mi instinto y, sobre todo, después de analizar las distintas alternativas. Una de estas oportunidades me ha llevado a estar aquí y a incorporándome a la Real Academia de Medicina de Salamanca y, con ello, a relanzar una línea de investigación que tantas satisfacciones nos ha dado.

Para mí, todo esto comenzó con una beca Erasmus que me llevó hasta una universidad escocesa, en la que seguí con mis estudios de informática y aprendí algo de inglés. Aunque no fue esto lo más importante que aconteció. Allí conocí a mi mujer, lo mejor que me ha pasado junto con algo que llegó después, el nacimiento de mis tres hijos, a los que agradezco todo el apoyo que me dan cada día. Volviendo a la parte académica, hay que

decir hasta que no llegué al Laboratorio Oceanográfico de Plymouth no supe qué era la investigación y lo que suponía. Ese año trabajando junto a investigadores de primera línea me empujó a volver a la Universidad para terminar la carrera y hacer el doctorado. Durante varios años, tuve el privilegio de estudiar y trabajar en paralelo en varios centros de investigación, al mismo tiempo que cree una empresa coincidiendo con el despegue de lo que hoy es Internet. Ahora que tengo la oportunidad, voy a agradecer a los doctores Brian Lees (me enseñó a ser metódico, un gran valor en la cultura británica) y Colin Fyfe (tenía los dones de la oportunidad y eficiencia) por dirigirme una tesis en Glasgow. También a mi compañero de departamento, profesor y director de la tesis doctoral que defendí en Salamanca, el Dr. Rafael García Bermejo-Giner (la constancia, la dedicación y el detalle en las pequeñas cosas), gracias Coti. Cómo olvidarme de mi jefe y director del Laboratorio Oceanográfico de Plymouth, Jim Aiken (la persona más lista que he conocido, especializada en sacar petróleo de las piedras), y de Gordon, compañero de clase en Glasgow. Estando muy lejos de ser brillante como estudiante, quizás más bien cerca del extremo opuesto, fue el

primero que me enseñó cómo se consigue, con una idea sutil y siendo de los primeros en lanzarla, ganar tanto dinero en pocos años como la mayoría de los más brillantes consiguen amasar en toda su vida. Debo dar las gracias también a mis padres y hermano por apoyarme en todas mis iniciativas y estar siempre ahí.

Después de pasar dos años en la universidad de Vigo, donde fui subdirector del Departamento de Informática, tuve la oportunidad de incorporarme al Departamento de Informática y Automática de nuestra Universidad en junio del año 2000.

Mi relación con la inteligencia artificial (IA) comenzó hace unos treinta años cuando, por primera vez, escuché hablar de este término y me preocupé por él con el único objetivo, en aquel entonces, de aprobar la asignatura de tercero. Fue unos años después cuando comencé a trabajar en un proyecto en el que se propuso utilizar la inteligencia artificial para resolver un problema ligado al mundo de la oceanografía para la Marina Real Británica. Justo en un momento en el que la apuesta por la IA aún no era clara y se seguía trabajando con modelos rudimentarios, aunque en evolución constante. El hecho de tener ya una cierta edad y de que esta disciplina sea

relativamente joven me ha permitido conocer de la mano de los pioneros en este campo cómo han ido evolucionando las cosas y ser partícipe de todo en las últimas tres décadas. Por ello, haré una pequeña revisión histórica sobre los orígenes de la IA y de lo que ha cambiado en ella, sobre todo en los últimos tres lustros.

La inteligencia artificial hasta la primera década del siglo XXI

La inteligencia artificial es un campo de la informática y la tecnología que se ocupa del desarrollo de sistemas informáticos que pueden realizar tareas que suelen requerir inteligencia humana, como el aprendizaje, la toma de decisiones, la resolución de problemas, la percepción y el lenguaje natural (Russell y Norvig, 2014). La IA se basa en algoritmos y modelos matemáticos complejos que permiten a los sistemas informáticos procesar grandes cantidades de datos y encontrar patrones y relaciones en ellos, lo que les hace aprender y mejorar con el tiempo.

La IA se aplica en una amplia gama de áreas, desde la

automatización de procesos industriales hasta la atención médica y la conducción autónoma de vehículos. También se utiliza en aplicaciones de *software*, como el reconocimiento de voz y de imágenes, la traducción automática y los *chatbots*. A modo de resumen, podemos decir que la IA se puede asociar a la capacidad de las máquinas para imitar y mejorar la inteligencia humana y realizar tareas que antes solo podían ejecutar las personas.

La noción de crear máquinas capaces de demostrar un comportamiento que pueda ser considerado inteligente se podría remontar a épocas anteriores al surgimiento de la ciencia de la computación. En varias culturas mitológicas se mencionan seres creados a partir de materia inanimada. En todo caso, y aunque se llevaron a cabo algunos trabajos pioneros antes de la década de 1950, los orígenes de la inteligencia artificial como disciplina científica se relacionan, por lo general, con los estudios del científico británico Alan Turing a principios de esa década. En uno de sus artículos más famosos (Searle, 2008; Turing, 1950), Turing abordó la cuestión central de la inteligencia artificial, ¿pueden las máquinas pensar?, y propuso una serie de criterios prácticos para determinar si una

máquina está realmente dotada de inteligencia. Poco después fue John McCarthy quien acuñó el término *inteligencia artificial* en 1956 y contribuyó al desarrollo del lenguaje de programación Lisp, que para muchos ha sido nuestra puerta de entrada en la IA (McCarthy *et al.*, 2006). Él, junto con otros, como Marvin Minsky, del MIT, Lotfali A. Zadeh (Universidad de Berkeley, en California) o John Holland (Universidad de Michigan), han sido los pioneros. Con todos ellos, hemos podido convivir en algunos congresos del AAAI y el IEEE, a los que como estudiante de doctorado asistía con regularidad. De sus manos han salido tendencias, modelos y algoritmos que han generado escuelas y sobre las que nosotros hemos podido construir sistemas que han supuesto verdaderos avances en campos como la medicina.

Así aparecieron ramas de trabajo como:

- La lógica simbólica, que surgió en los años 50 y se centra en la representación y manipulación del conocimiento simbólico y la lógica formal (Carroll, 1958). Pero el conocimiento que tenemos sobre muchos de los problemas actuales es limitado y pronto se observó que otras alternativas eran

necesarias.

- Los sistemas expertos, que utilizan reglas y conocimientos específicos de un dominio para simular la toma de decisiones de un experto humano. Son el máximo exponente de los sistemas simbólicos. Uno de los primeros sistemas expertos, y que se convirtió en un referente, fue el MICYN, desarrollado por Edward Shortliffe a principios de los años 70 en la Universidad de Stanford. Fue escrito en Lisp y nació inspirado en Dendral, otro sistema experto que tuvo cierto éxito a finales de los años 60. Su principal función consistía en el diagnóstico de enfermedades infecciosas de la sangre; además, MICYN era capaz de «razonar» el proceso seguido para llegar a estos diagnósticos y de recetar medicaciones personalizadas a cada paciente (según su estatura, peso, etc.). Un gran avance, pero con grandes problemas de mantenimiento, en gran medida por no existir algo como Internet.
- Las redes neuronales se basan en la idea de imitar el funcionamiento del cerebro humano (Corchado *et*

al., 2000). Esta rama de la IA utiliza algoritmos que permiten que las máquinas aprendan a través de ejemplos, sin necesidad de una programación explícita. Son el máximo exponente de los sistemas conexionistas y, en contraposición a los sistemas expertos, no necesitan partir de un conocimiento simbólico previo. En este caso, son los datos los que facilitan la reorganización de las redes y la generación de conocimiento. El problema de las redes que se crearon en este periodo inicial era la escalabilidad. Trabajaban bien con problemas pequeños, pero tenían dificultades para hacerlo con grandes cantidades de datos. Además, casi todas eran una caja negra.

- La lógica difusa, también conocida como lógica borrosa, es una rama de la inteligencia artificial que se desarrolló a finales de los años 60 (Zadeh, 2008). A diferencia de la lógica clásica, que se basa en valores binarios (verdadero o falso), la lógica difusa permite la representación y manipulación de la incertidumbre y la vaguedad en el razonamiento.

Esta tecnología supuso un antes y después en el desarrollo de sistemas de control, por ejemplo, y facilitó la posibilidad de desarrollar mecanismos para resolver problemas reales con gran eficiencia.

- El procesamiento del lenguaje natural (PLN) es la rama que se enfoca en la comprensión y generación de lenguaje natural por parte de las máquinas (Nadkarni *et al.*, 2011). Esta área se desarrolló para facilitar la interacción entre humanos y computadoras. Aunque se puede considerar una rama de la IA, el PLN implica el uso de una amplia variedad de tecnologías que han evolucionado considerablemente hasta nuestros días. Así, pueden desarrollar procesos de tokenización del análisis morfológico, sintáctico y semántico para la generación de lenguaje natural o el aprendizaje automático. Los progresos del siglo pasado han dado lugar al extraordinario avance que vivimos en nuestros días.
- Los algoritmos genéticos son una técnica de optimización basada en la teoría de la evolución

biológica (Reeves, 2010). Esta técnica se utiliza para resolver problemas de optimización en los que se busca encontrar la mejor solución posible de entre un conjunto de posibles soluciones. Los algoritmos genéticos surgieron en la década de 1960 de la mano de científicos como John Holland, que se interesó por aplicar los principios de la selección natural y la evolución biológica a los problemas de optimización en el campo de la inteligencia artificial.

- La visión por computadora se centra en el análisis y procesamiento de imágenes y videos por parte de las máquinas (Maes *et al.*, 2019). Esta rama de la IA se aplica en áreas como la robótica, la seguridad, el reconocimiento facial y de imágenes médicas, etc.
- La robótica trabaja el diseño, construcción y operación de robots para ejecutar tareas específicas. La IA se utiliza en la robótica para permitir que los robots aprendan y se adapten a diferentes situaciones.
- Los sistemas multiagentes, que aparecieron ya en la

década de los 80, se enfocan en el diseño y desarrollo de sistemas en los que múltiples algoritmos inteligentes interactúan entre sí para lograr objetivos comunes (Bajo y Corchado, 2009). Esta área se ha utilizado en aplicaciones como el comercio electrónico y la gestión de la cadena de suministro en gran medida por la oportunidad que surgió con Internet para crear sistemas inteligentes distribuidos.

- Las máquinas sociales surgieron en la segunda mitad del siglo pasado cuando aparecieron programas como ELIZA, que simulaba una conversación terapéutica con un psicólogo. Desde entonces, los sistemas de máquinas sociales han evolucionado significativamente, con avances como Siri, Alexa y otros asistentes virtuales que utilizan inteligencia artificial para comprender y responder a las preguntas y comandos de los usuarios. Estos sistemas de máquinas sociales tienen una amplia variedad de aplicaciones, desde asistentes virtuales hasta *chatbots* y robots de servicio (Hendler y Mulvehill, 2016). Su objetivo es interactuar con las

personas de manera natural y fluida y su éxito se mide, en gran medida, por su capacidad para comprender el lenguaje humano y generar respuestas relevantes y útiles.

Todas estas ramas se dividen en subramas y estas, en otras, de forma que hoy en día el nivel de especialización es intenso. Además, en este sentido, los grupos de investigación suelen ser bastante radicales y puristas en lo que respecta a su línea de investigación. Este nivel de especialización es muy importante para el avance y la mejora constante de los modelos de resolución de problemas, pero tiene el inconveniente de que limita a los grupos a trabajar en problemas relativamente concretos.

Mis comienzos en el ámbito laboral marcaron mi trayectoria investigadora en una dirección concreta y diferenciada de lo que suele ser normal en mi área. Comencé siendo programador en un laboratorio oceanográfico para trabajar en el desarrollo de un sistema inteligente distribuido en un momento en el que Internet era casi desconocido y, desde luego, muy alejado de los sistemas de comunicación militares. Este fue uno, si no el primero, de los sistemas que requería el

uso de IA en el Laboratorio Oceanográfico de Plymouth. Tanto el resto de mis compañeros y nuestros jefes como yo tuvimos que ponernos al día en todo lo relacionado con la IA para comenzar a trabajar y resolver un problema complejo y de cierta dimensión. Aquellos eran tiempos en los que se disponía de años para completar un proyecto, en los que el mundo iba aún a otro ritmo. Este hecho me permitió asistir a numerosos congresos, especialmente en Estados Unidos, y experimentar con la práctica totalidad de los sistemas que he mencionado. Digo esto porque, a diferencia de mis compañeros, yo opté por no especializarme y dedicarme a conocer en profundidad las diferentes alternativas y a trabajar en su fusión para construir sistemas híbridos.

La mayor parte de los sistemas complejos se ven afectados por múltiples elementos, generan o están relacionados con múltiples fuentes de datos, evolucionan con el tiempo y, en la mayoría de los casos, existe cierto conocimiento experto. En este sentido, parece claro que el uso combinado de sistemas simbólicos capaces de modelar el conocimiento junto con técnicas conexionistas que analicen datos a diferentes niveles o procedentes de diferentes fuentes puede ofrecer soluciones

globales. No resulta difícil encontrar problemas de este tipo en el ámbito de la medicina, donde el modelado del conocimiento es tan importante como el análisis de los datos de los pacientes por sí solos. Un ejemplo de fusión de modelos fue la plataforma Gene-CBR para el análisis genético. Utilizaba, por un lado, el marco metodológico proporcionado con un sistema de razonamiento basado en casos junto con varias redes neuronales y sistemas difusos (Díaz *et al.*, 2006). Este sistema se construyó para facilitar el análisis del mieloma con datos que nos proporcionó el servicio de Hematología del Hospital Universitario. Siempre he tenido una especial preocupación por explicar las soluciones que ofrecían los sistemas que he gestionado. Aunque muchas de las redes neuronales son cajas negras, hay excepciones y de algunas se puede extraer el conocimiento que almacenan en forma de, por ejemplo, sistemas borrosos. La construcción de este tipo de modelos neurosimbólicos ha sido una constante en mi trayectoria laboral, lo que, a la postre, nos ha posicionado a nivel nacional e internacional y permitido impulsar numerosos proyectos, sobre todo ahora que la inteligencia artificial explicativa se ha convertido en una necesidad.

La década de 80 fue un período de gran avance para la inteligencia artificial. Los investigadores de todo el mundo trabajaron muy duro para desarrollar nuevos métodos y técnicas que permitieran a las máquinas, por ejemplo, procesar y comprender el lenguaje natural, crear sistemas que pudieran «ver» o reconocer objetos en el mundo físico. A finales del siglo pasado, el interés se relajó.

Todo esto sucedió en una época de grandes cambios, con el despegue de Internet, en un momento en el que el mundo estaba acercándose a un nuevo siglo. En un contexto en el que casi todos los de mi edad que estábamos dedicados a la informática tuvimos la oportunidad de «ganar dinero» haciendo un trabajo fácil y rutinario que permitió que los sistemas informáticos incorporaran algunos dígitos más a las variables que almacenaban fechas para evitar que llegara el colapso general. Este hecho, junto con las limitaciones del *hardware*, el desinterés de la industria por la IA y la falta de ideas rompedoras, contribuyó al comienzo de una etapa de paralización de la IA, lo que llamamos el invierno de la IA.

El verano de la inteligencia artificial

Pero después de un invierno hay un verano y este llegó a primeros de siglo, con la aparición de lo que llamamos el aprendizaje profundo o las redes convolucionales. Una gran idea que supuso un cambio radical en la forma de tratar la información (Bengio, 2009), una nueva forma de tratar los datos y el conocimiento y de resolver problemas de forma elegante y limpia, aunque con el uso de la fuerza bruta que proporcionan las GPU (unidad de procesamiento gráfico) actuales y la gran cantidad de información almacenada en nuestras nubes y espacios de datos.

Las redes neuronales que utilizan modelos de aprendizaje profundo surgieron a mediados de la década de los 2000 y han experimentado un gran avance en los últimos años. Estas redes utilizan técnicas de aprendizaje automático de forma un tanto distinta a como se había planteado en sus orígenes. A diferencia de otros modelos, tienen múltiples capas ocultas que permiten extraer características y patrones de los datos de entrada de manera cada vez más compleja y abstracta.

Este enfoque ha permitido logros significativos en áreas como la visión por computadora, el procesamiento del lenguaje natural, el reconocimiento de voz, la robótica y la toma de decisiones en general. Algunos ejemplos notables de sistemas de inteligencia artificial profunda incluyen AlphaGo, de Google, que derrotó al campeón mundial de Go en 2016, y GPT-3, de OpenAI, un modelo de lenguaje natural que puede generar textos muy convincentes.

Las redes neuronales convolucionales (CNN) son un tipo de red neuronal que utiliza un aprendizaje profundo cuyo funcionamiento se basa en la aplicación de operaciones de convolución a los datos de entrada. La estructura básica de una CNN consiste en varias capas, que se dividen en tres tipos principales: capas de convolución, capas de *pooling* y capas totalmente conectadas. Las capas de convolución son las encargadas de aplicar un conjunto de filtros a la imagen de entrada para extraer características relevantes, como bordes, esquinas, texturas, etc. Cada filtro es una matriz que se desliza sobre la imagen de entrada, multiplicando sus valores correspondientes y sumándolos para obtener un valor de salida en una nueva matriz de características.

Las capas de *pooling* se utilizan para reducir la dimensión de la matriz de características obtenida en las capas de convolución. Esto se logra mediante la aplicación de una operación de agrupación (por ejemplo, el máximo o la media) en una ventana deslizante de la matriz de características.

Por último, las capas totalmente conectadas se utilizan para clasificar la imagen a partir de las características extraídas por las capas anteriores. Estas capas son similares a las de una red neuronal convencional y consisten en neuronas que reciben entradas de todas las neuronas de la capa anterior y producen salidas que se utilizan para determinar la clase de la imagen.

Durante el entrenamiento, la red ajusta los valores de los filtros y de los pesos de las neuronas para minimizar la diferencia entre las predicciones y las etiquetas de las imágenes de entrenamiento. Una vez que la red está entrenada, se puede utilizar para clasificar nuevas imágenes con una alta precisión.

Esto supone un gran avance y un cambio de paradigma, que, en gran medida, minimiza la necesidad de utilizar los sistemas híbridos con los que hemos trabajado durante décadas, ya que permite el análisis de un problema global desde distintas perspectivas. Estos modelos suponen un antes y un después y

harán que todos demos una vuelta de tuerca a nuestra forma de trabajar. Los ordenadores automatizaron procesos y escalaron la economía de una forma extraordinaria y fueron el elemento precursor de la tercera revolución industrial (marcada por la aparición del primer microprocesador en 1971). Entre la primera revolución industrial y la tercera pasaron dos siglos. Entre la tercera y la cuarta, impulsada por la llegada de Internet a los procesos de fabricación, pasaron poco más de cuarenta años. Ahora ya se habla del comienzo de la quinta revolución, solo dos décadas después, gracias a nuestra capacidad para crear sistemas mediante la convergencia de tecnologías digitales, físicas y biológicas con el uso de estos nuevos modelos de creación de conocimiento. Si vivíamos en un mundo que avanzaba rápido, ahora tenemos que prepararnos para una nueva aceleración. Aquellos que cojan este tren verán cómo sus oportunidades de negocio, de generación de valor y servicio aumentan exponencialmente en los próximos años. Es nuestra oportunidad, como universidad y como ciudad. Se están abriendo nuevas y grandes oportunidades para la medicina de la mano de estos nuevos modelos.

La quinta revolución industrial llega impulsada por la

nueva era tecnológica que estamos experimentando y está propiciada por la convergencia de tecnologías digitales, físicas y biológicas. Esta revolución se caracteriza por la creciente automatización y digitalización de los procesos de producción y la creación de nuevos modelos de negocio centrados en el uso del dato y la nueva inteligencia artificial.

A diferencia de las revoluciones industriales anteriores, que se caracterizaron por la automatización de procesos específicos y la producción en masa, la nueva revolución está impulsada por la convergencia de tecnologías que permiten la personalización, la producción bajo demanda y la creación de soluciones a medida para las necesidades individuales de los consumidores, las empresas, los pacientes, etc. Aquí estamos hablando de medicina personalizada, de medicina de precisión, de toma de decisiones en tiempo real y de la hospitalización a domicilio segura, entre otras cosas.

Como ejemplo, ChatGPT han entrado en nuestras vidas para cambiarlas y casi no nos hemos enterado. Algunos solo han oído hablar de él, otros lo han utilizado en alguna ocasión y muchos otros ya están haciendo negocios y generando valor con esta tecnología. La capacidad que esta herramienta tiene

para generar texto, como este que estoy leyendo, es extraordinaria, pero esto solo es la punta del iceberg. Ya se está utilizando para crear sistemas de atención al cliente, el análisis de datos médicos y la toma de decisiones empresariales.

Pero ChatGPT no deja de ser el primero de los sistemas de este tipo que se ha abierto un hueco en el mercado. Hay muchos otros sistemas de procesamiento de lenguaje natural que pueden generar respuestas coherentes y precisas a partir de preguntas y comentarios de los usuarios. Algunos de estos sistemas son:

- GPT-3: Es uno de los modelos más grandes y avanzados de NLP hasta la fecha, desarrollado por OpenAI, la misma organización que desarrolló ChatGPT.
- BERT: Es un modelo de lenguaje natural basado en la atención, desarrollado por Google, que se utiliza en tareas de procesamiento de lenguaje natural, como la clasificación de texto y la comprensión de texto.
- XLNet: Es un modelo de lenguaje natural que utiliza

una técnica de entrenamiento de permutación de entrada y salida para capturar mejor las dependencias a larga distancia entre las palabras en una oración.

- T5: Es un modelo de lenguaje natural desarrollado por Google que se basa en la arquitectura Transformer y se puede utilizar para una amplia gama de tareas de NLP.
- RoBERTa: Es un modelo de lenguaje natural basado en la arquitectura Transformer que utiliza un enfoque de entrenamiento continuo para mejorar el rendimiento en tareas de NLP.
- Bedrock: Es un servicio de inteligencia artificial que permitirá a los usuarios crear modelos generativos a partir de AI21 Labs, Anthropic, Stability AI y Amazon.

Pero esto no es todo. Tendremos que viajar a Asia para conocer los sistemas más revolucionarios, grandes y disruptivos, que en breve serán una realidad también aquí. No sé si han oído hablar de Wu Dao (五道, que significa «cinco

caminos» en chino). Si no lo han hecho, ya se lo presento yo. Es, posiblemente, el modelo de procesamiento de lenguaje natural más grande y avanzado del mundo, desarrollado por Alibaba en China, con un total de 1,75 billones de parámetros. No conozco otro igual.

Al igual que GPT-3, Wu Dao se entrena en una amplia variedad de tareas de lenguaje natural, incluyendo la generación de texto, la traducción, el resumen de texto o la conversión de texto a voz, entre otros. Además, Wu Dao también se puede utilizar en tareas específicas de la industria, la fabricación y, sobre todo, en la atención médica y sanitaria.

Wu Dao ya se utiliza en el sector educativo y médico y en la mejora de procesos comerciales, logísticos e industriales como nunca se habría podido hacer antes con ninguna otra herramienta. Esta tecnología nos impulsará a velocidad de crucero y su uso y dominio será no solo una oportunidad, sino más bien una necesidad.

Tecnología como esta permitirá el desarrollo de sistemas de diagnóstico mucho más precisos basándose en evidencias e historiales clínicos, el uso más generalizado de la telemedicina, de sistemas de control de pacientes crónicos en sus domicilios,

etc. En esta dirección, y a diferentes niveles, están viendo la luz numerosas propuestas de gran interés en el ámbito médico como, por ejemplo, Transformers, Autoencoders, Multimodal Self-Supervised Learning, Deep Energy-Based Generative Models, Generative Models and Symmetries, Variational Prototype Inference, Deep Generative Models, Contrastive Learning Approach for Variational Autoencoder Priors o Causal Inference Multi-Agent Reinforcement Learning.

Por centrarnos en un problema concreto, los diagnósticos médicos modernos se basan en gran medida en el análisis de múltiples modalidades de imagen, en particular para el diagnóstico diferencial. Sin embargo, para utilizar aprendizaje automático supervisado es necesario anotar adecuadamente una gran cantidad de datos. La anotación de expertos de datos de pacientes multimodales a escala no es trivial; es costosa, requiere mucho tiempo y está asociada con riesgos relacionados con la protección de datos. Incluso las herramientas de *software* semiautomáticas pueden no reducir lo suficiente los costos de anotación. En consecuencia, la falta de anotaciones de datos es uno de los principales impedimentos para las aplicaciones de aprendizaje automático en imágenes

médicas. Al mismo tiempo, las canalizaciones modernas de aprendizaje profundo están aumentando en profundidad, complejidad y requisitos de memoria, creando un cuello de botella computacional adicional. Para resolver estos problemas se está trabajando con diferentes propuestas como las mencionadas antes:

- **Multimodal Self-Supervised Learning:** El aprendizaje multimodal autosupervisado ofrece una solución viable cuando los datos de formación etiquetados son escasos (Li *et al.*, 2023). En estos enfoques, las señales de supervisión se derivan de los propios datos, normalmente mediante aprendizaje no supervisado. Los modelos obtenidos mediante autosupervisión facilitan un ajuste fino eficiente en la tarea objetivo final, reduciendo la carga de anotación manual. Los métodos autosupervisados utilizan el contexto espacial como señal de supervisión para aprender representaciones eficaces de los datos. Considerar la multimodalidad desde una perspectiva anatómica es fundamental porque, por ejemplo, las diferencias en las propiedades físicas de

órganos y tejidos se traducen de forma complementaria.

- **Deep Energy-Based Generative Models:** Trabajos recientes han comenzado a demostrar que los algoritmos de RM rápida basados en la metodología del aprendizaje profundo (*deep learning*, DL) podrían proporcionar una flexibilidad aún mayor sin comprometer la calidad de la imagen (Xu, 2022). Sin embargo, estos métodos son, a veces, poco prácticos porque el requisito de una gran cantidad de datos de entrenamiento no siempre es factible. Los modelos de aprendizaje profundo no supervisado, como los modelos generativos profundos, son flexibles y robustos precisamente porque no requieren datos de etiquetas y la correspondencia estricta entre los conjuntos de entrenamiento y prueba. Entre ellos, los modelos basados en energía (EBM) que utilizan la función de energía como modelo generativo han proporcionado una nueva perspectiva, con iteración implícita y muestreo sin gradiente, ya que controlan la diversidad en los modelos probabilísticos y el colapso de modos en los GAN. Es decir, la EBM es un modelo no guiado que se aproxima mejor a la

densidad de datos que otras redes generativas. A la vez, la EBM muestra algunas ideas únicas del aprendizaje autoadversarial al considerar la función de energía como un discriminador o generador que atribuye energía a múltiples regiones de datos.

- **Generative Models and Symmetries:** Las simetrías pueden hacer que los modelos de *machine learning* (ML) sean más eficientes con los datos y capaces de extrapolar a nuevas regiones del espacio de entrada (Yanget *et al.*, 2023). La clave está en diseñar modelos que tengan en cuenta explícitamente y puedan explotar las simetrías presentes en el problema. Teniendo esto en cuenta, las redes generativas hamiltonianas (HGN) capaces de aprender la dinámica hamiltoniana ofrecen una forma de predecir el comportamiento futuro de un sistema a partir de su estado actual en el espacio de fases (es decir, su posición y momento para los sistemas newtonianos clásicos y su posición y momento generalizados de forma menos concreta). La mecánica hamiltoniana induce una dinámica con varias propiedades útiles, que también lo son para los sistemas de aprendizaje automático. Por ejemplo, capturar

la evolución temporal reversible del estado podría ser útil para los agentes que intentan explicar cómo sus acciones producen efectos; recuperar una variedad abstracta de baja dimensión con trayectorias que preserven varias propiedades está muy relacionado con problemas abiertos en el aprendizaje de representaciones, redes con gran potencial en el ámbito médico y biológico.

- Variational Prototype Inference: El desafío crítico en la segmentación semántica de pocas tomas es la escasez de datos anotados para cada categoría de objeto a segmentar (Huang *et al.*, 2022). Por lo tanto, inferir fielmente la clase de objetos a partir de imágenes de soporte es clave para guiar la segmentación de objetos en la imagen de consulta. Al modelar el prototipo de clase como una distribución en lugar de un vector determinista, la incertidumbre causada por imágenes de soporte limitadas se puede manejar mejor y se mejora la generalización para manejar grandes variaciones de objetos dentro de la clase. Esto permite una mejor segmentación de imágenes médicas. Un autocodificador variacional (VAE) proporciona una forma probabilística de describir una observación en el espacio

latente. Por lo tanto, en lugar de construir un codificador que genere un valor único para describir cada atributo de estado latente, formulamos nuestro codificador para describir una distribución de probabilidad para cada atributo latente.

- **Deep Generative Models:** Los sistemas de detección de tumores pulmonares basados en CNN suelen tener dos fases. Una fase *offline* emplea un modelo CNN profundo entrenado utilizando un conjunto de imágenes CXR clasificadas (datos de entrenamiento). Una fase en línea toma una imagen de pulmón CXR y determina si contiene tumores o no (Strokach y Kim, 2022). Los sistemas basados en CNN se han aplicado con éxito al problema de la detección y clasificación de tumores pulmonares, por ejemplo. Sin embargo, los problemas relacionados con los datos han restringido la amplia aplicación clínica del aprendizaje profundo. La creación de conjuntos de datos de entrenamiento anotados a gran escala en diversas poblaciones, desde enfermedades comunes a raras, centros médicos y protocolos de adquisición, sigue siendo un obstáculo importante para el desarrollo de un sistema de

aprendizaje profundo en medicina. Además, la eficacia clínica de los modelos supervisados de aprendizaje profundo solo se ha validado en pacientes seleccionados con riesgo de una sola enfermedad o enfermedades específicas. Por lo tanto, este enfoque no puede garantizar que el aprendizaje profundo pueda manejar condiciones nuevas o nunca vistas. Los modelos generativos profundos aprenden a capturar la distribución de datos objetivo; por lo tanto, pueden detectar datos anómalos que se desvían de la distribución objetivo sin conocimiento previo de las anomalías.

- **Contrastive Learning Approach for Variational Autoencoder Priors:** El aprendizaje contrastivo es una poderosa clase de métodos de aprendizaje de representación visual no supervisados (Jyoti *et al.*, 2022). Aprende extractores de características minimizando la distancia entre representaciones de pares positivos o muestras que son similares en algún sentido y maximizando la distancia entre representaciones de pares negativos o muestras que son diferentes en algún sentido. El aprendizaje contrastivo se puede aplicar a imágenes sin

etiquetar haciendo que los pares positivos contengan aumentos de la misma imagen y los pares negativos contengan aumentos de diferentes imágenes. Esto se puede aplicar a codificadores automáticos variacionales.

- Transformer Networks: Una diferencia importante en comparación con las CNN es que las redes de transformadores tienen un fuerte mecanismo de atención inherente (Tran *et al.*, 2023). Mientras que en una CNN una neurona dada solo toma entradas de neuronas adyacentes en la capa anterior, en un transformador las entradas pueden tomarse de cualquier parte de la capa anterior. Las características utilizadas para el procesamiento están determinadas por un mecanismo de atención aprendido. Dependiendo de los datos de entrada, cada capa puede seleccionar ciertas partes de la entrada para su posterior procesamiento. Esto es posible debido a que un mecanismo de red neuronal bastante complicado imita la idea de clave, consulta y valor de las bases de datos tradicionales, donde se recupera un valor de una tabla en función de una consulta que coincide con una clave. Este concepto

de atención muestra que un transformador puede modelar con facilidad relaciones locales y distantes en los datos presentados, lo cual es un desafío para las CNN, ya que siempre tienen un contexto muy local.

- **Causal Inference Reinforcement Learning:** La causalidad se refiere a la relación entre una causa y su efecto; es decir, se basa en el supuesto de que todo efecto surge debido a una causa específica o a un conjunto de causas (Cohen *et al.*, 2020). La inferencia causal puede aplicarse para predecir resultados de pacientes basados en variables como edad, historial médico, síntomas y exámenes de laboratorio. Estos modelos pueden ayudar a identificar factores de riesgo y guiar la toma de decisiones clínicas. También puede utilizarse para identificar biomarcadores, que son indicadores biológicos de una enfermedad o condición. Esto podría incluir la identificación de patrones genéticos, proteínas o metabolitos que estén relacionados con la aparición y progresión de enfermedades.

Aunque todo parece muy prometedor, aún hay que resolver problemas relacionados con la protección de datos, la

capacidad de explicar las soluciones, el coste de las infraestructuras, la formación de personal asistencial y la concienciación de los pacientes y sus familias, entre otros. De especial relevancia son los aspectos éticos de las soluciones que generamos.

La IA tiene el potencial de cambiar fundamentalmente la forma en que vivimos y trabajamos, pero también plantea importantes desafíos éticos. En términos de privacidad y seguridad, la IA plantea una serie de riesgos en el ámbito médico. Con la capacidad de analizar grandes cantidades de datos, la IA puede ser utilizada para recopilar y analizar información personal y de comportamiento, lo que plantea importantes preocupaciones de privacidad. Además, la IA también puede ser utilizada para fines malintencionados, como el sesgo en el dato y la manipulación en la toma de decisiones.

Para abordar estos desafíos, los investigadores de IA están trabajando en nuevas técnicas de seguridad y privacidad de los datos. Estos incluyen el cifrado, la anonimización y la minimización de datos y la protección contra el uso indebido por parte de terceros.

Para abordar estos desafíos, los investigadores de IA están

trabajando en nuevas técnicas de automatización ética y justa. Estas incluyen técnicas de modelado y selección de datos que minimizan el sesgo, así como el desarrollo de marcos éticos y legales para la IA.

La explicabilidad del dato (o «explicabilidad de la IA») se refiere a la capacidad de un sistema de inteligencia artificial para explicar cómo se llega a una determinada predicción, clasificación o resultado (Van der Velden *et al.*, 2022). En otras palabras, la explicabilidad de la IA se refiere a la transparencia y comprensibilidad del proceso de toma de decisiones de un sistema de IA.

La explicabilidad del dato es especialmente importante en situaciones en las que se toman decisiones críticas basadas en el análisis de datos, como es el caso de la atención médica, la seguridad pública y la toma de decisiones financieras. La falta de explicabilidad de la IA puede generar desconfianza en el sistema y hacer que los usuarios se sientan incómodos.

En esta línea se está trabajando mucho y ya existen técnicas para mejorar la explicabilidad de la IA, incluyendo la visualización de datos, el análisis de características y el uso de técnicas de aprendizaje automático interpretable. En todo caso,

la inteligencia artificial requiere de técnicas y herramientas complementarias para poder ofrecer sus servicios de forma razonable, sostenible, ética y generalizada. Técnicas que presentaremos y que constituyen, junto con la IA, el núcleo de la transformación tecnológica que supone la quinta revolución industrial.

Complementos a la IA y su relevancia en el campo de la medicina

La resolución de problemas con inteligencia artificial requiere expertos humanos con conocimientos sobre los temas a tratar, acceso a fuentes de datos, sistemas de almacenamiento, plataformas que faciliten el acceso a algoritmos y datos, modelos que faciliten la explicación de los resultados obtenidos, mecanismos de trazabilidad, garantías sobre la fiabilidad de los resultados y la protección del dato, estrategias de simulación, arquitecturas distribuidas que permitan la toma de decisiones y una regulación adecuada, entre otras muchas cosas. Para todo ello ya hay soluciones que están apoyando y

encumbrando a la IA como elemento disruptivo en la transformación digital actual, con la que estamos viendo nacer una nueva generación de sistemas con inteligencia global como nunca habíamos visto. Estas herramientas y conceptos son, entre otros, *big data*, lagos de datos, plataformas específicas de IA, *blockchain*, aprendizaje federado, gemelo digital, IA fiable, *edge computing* y *sandbox* de IA. Revisamos ahora los elementos que caracterizan estas disciplinas.

- *Big data*. Con este término nos referimos a grandes conjuntos de datos que son demasiado complejos y masivos para ser procesados por herramientas y métodos tradicionales de procesamiento de información. Estas cantidades de gigabytes pueden generarse por una amplia variedad de fuentes, como sensores, redes sociales, transacciones financieras, registros médicos y mucho más. El *big data* se caracteriza por tres factores principales conocidos como las tres V: volumen, velocidad y variedad. El volumen se refiere a la cantidad masiva de datos generados; la velocidad, a la rapidez con la que se generan los datos; y la variedad, a la diversidad de

tipos de datos que se utilizan. La capacidad de procesar y analizar grandes conjuntos de datos se ha vuelto cada vez más importante en una variedad de campos, desde la investigación científica y la medicina hasta el comercio minorista y la industria manufacturera. El análisis de *big data* puede proporcionar información valiosa y conocimientos que antes eran imposibles de obtener. Para manejar el *big data* se utilizan herramientas y tecnologías específicas, como sistemas de almacenamiento y procesamiento distribuidos, *software* de análisis y visualización de datos, algoritmos de aprendizaje automático, entre otros.

- Lagos de datos. Los términos *datos de lagos* o *data lakes* y *data spaces* se refieren a conceptos similares en el mundo del *big data*, pero con algunas diferencias clave. Un *data lake* es un gran repositorio de datos crudos, no estructurados y sin procesar que se almacena de forma distribuida. Los datos se guardan tal como se generan, sin procesamiento previo ni estructuración, lo que significa que pueden

incluir una gran cantidad de información de diversos tipos, desde transacciones financieras hasta tuits de redes sociales. Los datos de un *data lake* se pueden procesar y analizar para extraer información valiosa utilizando herramientas y tecnologías de análisis de *big data*. Por otro lado, un *data space* es una plataforma de *big data* que proporciona un espacio de trabajo virtual y colaborativo para que los usuarios almacenen, procesen y compartan datos. A diferencia de un *data lake*, un *data space* puede incluir tanto datos estructurados como no estructurados y generalmente ofrece herramientas integradas para la visualización y análisis de datos. Además, los *data spaces* suelen contar con herramientas de colaboración y de gestión de proyectos que permiten a los usuarios trabajar juntos en el procesamiento y análisis de datos. Los datos de lagos o *data lakes* se centran en el almacenamiento de grandes cantidades de datos no estructurados, mientras que los *data spaces* se enfocan en proporcionar una plataforma colaborativa y herramientas de análisis de datos para

procesar tanto datos estructurados como no estructurados.

- Plataformas de IA. Existen numerosas plataformas importantes para construir sistemas de inteligencia artificial, cada una con sus propias fortalezas y debilidades. Algunas de las más importantes son:
 - TensorFlow es una plataforma de código abierto desarrollada por Google que se utiliza para la construcción de sistemas de IA. Es altamente escalable y puede utilizarse para la creación de redes neuronales, modelos de aprendizaje profundo y otras aplicaciones de la IA.
 - PyTorch es otra plataforma de código abierto para el desarrollo de sistemas de IA. Se utiliza para la creación de modelos de aprendizaje profundo, redes neuronales y otras aplicaciones de la IA. PyTorch es conocida por ser fácil de aprender y utilizar, lo que la hace muy popular entre los desarrolladores de IA.

- Keras es una plataforma de IA de alto nivel que se utiliza para la construcción de redes neuronales. Es fácil de utilizar y está diseñada para ser modular, lo que significa que los desarrolladores pueden utilizar las piezas que necesitan para construir sus propios sistemas de IA.
- Microsoft Azure es una plataforma de nube que proporciona una amplia gama de herramientas y servicios para el desarrollo de sistemas de IA. Incluye herramientas para el procesamiento de lenguaje natural, la visión por computadora, el aprendizaje automático y más.
- IBM Watson es una plataforma de IA desarrollada por IBM que proporciona una amplia gama de herramientas y servicios para el desarrollo de sistemas de IA. Incluye herramientas para el procesamiento de lenguaje natural, la visión por computadora, el aprendizaje automático y más.
- Deepint.net es una plataforma que utiliza

tecnologías avanzadas de procesamiento de lenguaje natural y de aprendizaje automático para analizar y extraer información valiosa de grandes conjuntos de datos no estructurados. Esto permite a los usuarios comprender mejor las tendencias, patrones y sentimientos en el contenido de los medios sociales, noticias y otras fuentes de información en línea, sin necesidad de programación.

- *Blockchain*. También conocido como cadena de bloques, es una tecnología de registro distribuido que permite registrar transacciones y asegurar la integridad de los datos de manera descentralizada. En esencia, es una base de datos descentralizada y segura que se utiliza para almacenar y compartir información sin la necesidad de un intermediario de confianza. Cada bloque en la cadena de bloques contiene una serie de transacciones. Una vez que se agrega un bloque a la cadena, la información que contiene no puede ser modificada sin alterar todos los bloques posteriores en la cadena. Esto garantiza la

inmutabilidad y la integridad de la información almacenada en la cadena de bloques. La tecnología *blockchain* se popularizó por su uso en la creación de criptomonedas como el bitcoin, ya que proporciona un sistema seguro y descentralizado para la transferencia de valor. Sin embargo, la tecnología *blockchain* también se está utilizando en otras áreas, como la gestión de identidad digital, la trazabilidad de productos y la automatización de contratos inteligentes.

- Aprendizaje federado. El aprendizaje federado es una técnica de aprendizaje automático que permite entrenar modelos de manera distribuida y colaborativa sin la necesidad de compartir datos de entrenamiento entre diferentes entidades. En lugar de compartir los datos de entrenamiento entre los diferentes nodos de la red, el aprendizaje federado permite que cada nodo entrene un modelo local utilizando sus propios datos y luego se combinan para crear un modelo global. Esta técnica se utiliza en situaciones en las que la privacidad y la seguridad

de los datos son una preocupación, como en la atención médica y la banca, donde la privacidad de los datos del paciente y del cliente es crítica. En lugar de enviar datos sensibles a una ubicación central para el procesamiento, el aprendizaje federado permite que los modelos se entrenen de forma descentralizada, manteniendo la privacidad de los datos en todo momento. El proceso de aprendizaje federado se lleva a cabo en tres etapas principales: selección del modelo, entrenamiento descentralizado y agregación del modelo. Durante la selección del modelo, se selecciona un modelo base que se utilizará en cada uno de los nodos de la red. Durante el entrenamiento descentralizado, cada nodo entrena su modelo local utilizando sus propios datos. Por último, durante la agregación del modelo, se combinan los modelos de todos los nodos para crear un modelo global que es utilizado por toda la red. El aprendizaje federado es una técnica prometedora que se está utilizando cada vez más en diferentes sectores para mejorar la privacidad y la seguridad de los datos

en aplicaciones de aprendizaje automático.

- **Gemelo digital.** Un gemelo digital, también conocido como *digital twin* en inglés, es una réplica digital de un objeto, sistema o proceso del mundo real. El gemelo digital es una representación virtual detallada y en tiempo real de un objeto físico o proceso que permite su monitorización, análisis y simulación a través de tecnologías digitales. El objetivo del gemelo digital es proporcionar una representación precisa y detallada de un objeto o proceso en el mundo real, lo que permite a los usuarios analizar y experimentar con él sin afectar al objeto o proceso físico en sí. Esto permite a los usuarios probar diferentes escenarios y cambios en el gemelo digital antes de aplicarlos en el mundo real. El gemelo digital se utiliza en una amplia variedad de aplicaciones, como la monitorización de la salud de equipos en tiempo real, la planificación de la producción, la simulación de procesos complejos, la gestión de activos y la optimización de la cadena de suministro. Además, el gemelo digital se está

convirtiéndose en una parte importante de la industria 4.0 y la fabricación inteligente. Esta tecnología permite a los usuarios analizar y experimentar con el gemelo digital sin afectar al objeto o proceso físico en sí, lo que ofrece un mayor control y optimización de los procesos y sistemas del mundo real.

- IA fiable. La inteligencia artificial fiable se refiere a la capacidad de un sistema de inteligencia artificial para funcionar de manera segura, predecible y confiable en todas las situaciones. Esto significa que debe ser capaz de tomar decisiones adecuadas y precisas en diferentes entornos y circunstancias y de garantizar que los resultados sean consistentes y confiables. La inteligencia artificial fiable es importante porque, a medida que los sistemas de inteligencia artificial se vuelven más avanzados, también se vuelven más complejos y difíciles de entender. Esto hace que sea crucial garantizar que los sistemas de inteligencia artificial sean capaces de operar de manera confiable y segura, sobre todo en situaciones críticas, como en la atención médica o la

conducción autónoma de vehículos. Para garantizar que la inteligencia artificial sea fiable se utilizan técnicas de verificación, validación y certificación de sistemas de inteligencia artificial. Esto incluye la implementación de pruebas rigurosas y exhaustivas para garantizar que funcione de manera predecible y confiable en diferentes entornos y situaciones. Además, se están desarrollando normas y directrices de seguridad y ética. Estas normas y directrices también guían el desarrollo de sistemas de inteligencia artificial para que sean transparentes y comprensibles, lo que permitirá a los usuarios entender cómo funciona el sistema y cómo se toman las decisiones.

- *Edge computing*. Es una arquitectura de red que permite procesar y almacenar datos de forma descentralizada y cercana al usuario, en lugar de enviarlos a servidores remotos para su procesamiento. En estas arquitecturas se utilizan dispositivos periféricos (como sensores, cámaras o *smartphones*) para procesar los datos de forma local.

Esto reduce la latencia y aumenta la velocidad de procesamiento, lo que permite una respuesta más rápida a las solicitudes del usuario y una mayor eficiencia en el procesamiento de datos. El *edge computing* se utiliza en una amplia variedad de aplicaciones, incluyendo el Internet de las cosas (IoT), la monitorización de la salud y la seguridad en tiempo real, la automatización de las fábricas, el procesamiento de datos de drones y vehículos autónomos y mucho más. Además, el *edge computing* es una parte clave de la arquitectura de la red 5G, que requiere un procesamiento de datos más rápido y cercano al usuario. El *edge computing* es una arquitectura de red innovadora que ofrece una mayor eficiencia en el procesamiento de datos y una mejor respuesta a las solicitudes del usuario. Como la cantidad de dispositivos conectados a Internet crece, el *edge computing* se está convirtiendo en una parte cada vez más importante de la infraestructura de red.

- *Sandbox* de IA: La *sandbox* de IA es un entorno de pruebas controlado donde las empresas e

instituciones pueden experimentar y probar nuevas soluciones de inteligencia artificial antes de su lanzamiento al mercado. El objetivo de la *sandbox* de IA es reducir los riesgos y los costos asociados con la implementación de nuevas soluciones de IA, al mismo tiempo que se fomenta la innovación y se garantiza la seguridad y la protección de los derechos de los usuarios. En España, la *sandbox* de IA forma parte de la estrategia nacional de IA, que fue presentada en 2019 y se encuentra bajo la supervisión de la Secretaría de Estado de Digitalización e Inteligencia Artificial. Este entorno permite a las empresas experimentar con soluciones de IA en una amplia variedad de sectores, incluyendo la salud, el transporte, la energía y el medio ambiente. A nivel europeo, España es uno de los países líderes en el desarrollo de políticas y estrategias de IA. En 2020, la Comisión Europea nos seleccionó como uno de los nueve países que liderarán el desarrollo de políticas de IA en Europa. Además, España está trabajando en el desarrollo de una estrategia de IA a nivel europeo

en colaboración con otros países europeos y la Comisión Europea.

Todas estas tecnologías, entornos de prueba o desarrollo y/o conceptos de alto nivel ofrecen un complemento extraordinario para facilitar el uso de algoritmos inteligentes. De hecho, es difícil concebir la construcción de sistemas de generación de conocimiento sin tener en cuenta algunos de los términos anteriores. Especialmente en ámbitos como el de la medicina, es inviable crear sistemas sin tener en cuenta todos los aspectos éticos que afectan a la toma de decisiones. También es imprescindible mantener la privacidad del dato y, al mismo tiempo, compartir con cuantos sea posible las buenas prácticas y los modelos de éxito. El uso de sistemas de simulación que permitan identificar, por ejemplo, cómo pueden evolucionar las epidemias o los efectos de las medidas sobre estas, podría aumentar la tasa de éxito. En este caso, los gemelos digitales pueden ser una de las mejores opciones.

En el ámbito médico

Con mi llegada a la Universidad de Salamanca, y tan pronto fue

posible, solicité la creación del Grupo de Investigación en Bioinformática, Sistemas Informáticos Inteligentes y Tecnología Educativa (BISITE). Al nacer el IBSAL también nos incorporamos como grupo de investigación. A lo largo de estas dos décadas hemos podido impulsar varios proyectos relacionados con el análisis de datos genético y las neurociencias y ahora estamos trabajando en un nuevo proyecto relacionado con la generación de vacunas de forma computacional utilizando técnicas de vacunología inversa (*reverse vaccinology*) (Moxon *et al.*, 2019). Esta sección presenta tres proyectos en desarrollo en diferentes niveles de madurez.

El primero de ellos, DeepNGS, es un proyecto relacionado con la secuenciación de alto rendimiento para la identificación de tumores y enfermedades hereditarias. Se trata de un desafío que nos está proporcionando la oportunidad de trabajar con nuevas técnicas de IA y plataformas especializadas utilizando herramientas de aprendizaje federado y de gestión de espacios de datos. Todo un reto para nuestro equipo, cuya primera versión está en evaluación en el Centro Nacional de Investigaciones Oncológicas. Es el primer paso en la creación de herramientas complementarias de transcriptómica y epigenética, que ya estamos comenzando a definir.

HERMES (Hybrid Enhanced Regenerative Medicine) es el

siguiente proyecto que presentamos, cuyo objetivo es controlar un implante híbrido (formado por un dispositivo electrónico neuromórfico y un implante de células madre) para tratar el foco epiléptico y evitar la propagación y generación de crisis epilépticas.

Finalmente, comentaré un último proyecto, con el que estamos muy ilusionados, en el que la IA también juega un papel muy importante y que incorporará muchas de las tecnologías comentadas en la sección anterior. Se trata de la creación de un gemelo digital para generar vacunas de forma computacional. El objetivo de esta estrategia, siguiendo los pasos de otros que presentaron este concepto a primeros de siglo, es partir del genoma completo del organismo patógeno en cuestión e ir seleccionando en diferentes etapas, a través de un *pipeline* o *workflow* inteligente, los elementos más adecuados que generen una respuesta inmune potente dentro de una vacuna, que pueden ser proteínas, genes completos o fragmentos denominados péptidos, contando para ello con simuladores cuánticos. Nuestro objetivo es avanzar en este campo para crear un sistema que facilite el trabajo de laboratorio y, al mismo tiempo y en paralelo, prepararnos para el momento en el que tengamos acceso a computadores cuánticos y podamos dar un salto cuantitativo con ayuda de la nueva generación de algoritmos inteligentes que tenemos a nuestra disposición.

DeepNGS:

La secuenciación del ADN, junto con el descubrimiento de su estructura de doble hélice, ha sido uno de los hallazgos más revolucionarios en la biología del siglo XX, permitiendo leer la secuencia de nucleótidos (A, C, T o G) que conforman la estructura genética de cualquier organismo vivo. Este hito ha posibilitado el surgimiento del campo de la genómica para estudiar el conjunto de todos los genes y regiones no codificantes de un organismo con los que poder comprender las bases fisiopatológicas y moleculares de multitud de procesos celulares, con implicaciones clínicas que dan paso a lo que conocemos como medicina de precisión.

Desde la publicación del método de Sanger (Frederick Sanger, en 1975) hasta nuestros días, la secuenciación del genoma ha experimentado un gran desarrollo. La comunidad científica comprendió entonces que era necesario el desarrollo de nuevas técnicas de secuenciación que permitieran un nivel más alto de rendimiento y paralelización para democratizar su uso y aplicación.

Como respuesta a esta necesidad surgieron, a mediados de la primera década del presente siglo, los primeros métodos de secuenciación de alto rendimiento, o High-Throughput Sequencing (HTS), cuya principal característica es la alta paralelización del

sistema que permite la secuenciación de miles de fragmentos de ADN simultáneamente, mejorando el rendimiento de secuencias obtenidas por ejecución. En este sentido, se puede hablar de técnicas de secuenciación de próxima generación, o Next Generation Sequencing (NGS). Este es un desarrollo muy interesante para la construcción de gemelos digitales y la aplicación de la IA en numerosas partes de sus procesos:

- Evaluación de la calidad de las lecturas secuenciadas.
- Filtrado.
- Alineamiento o mapeado respecto al genoma de referencia.
- Detección de variantes genéticas.
- Anotación e interpretación de estas según la información de las bases de datos existentes.

Como grupo de investigación, nos propusimos crear una plataforma y un *pipeline* basado en IA y en ello estamos utilizando algunos de los algoritmos que hemos mencionado para:

- Mejorar la precisión y la eficiencia de las comprobaciones de control de calidad de los datos mediante la automatización del proceso de identificación y filtrado de lecturas de baja calidad, contaminación de adaptadores y otros posibles problemas.

- Mejorar la precisión y la velocidad del mapeo de lecturas mediante la identificación de parámetros de alineación óptimos y la optimización de los algoritmos de alineación.
- Determinación de variantes para mejorar la precisión y la sensibilidad de la llamada de variantes mediante el desarrollo de modelos de aprendizaje automático que pueden identificar las variantes verdaderas a partir del ruido de secuenciación y reducir las llamadas falsas positivas y negativas.
- Mejorar la anotación de variantes integrando grandes cantidades de datos de diferentes fuentes, como datos de genómica funcional e información de contexto genómico, para proporcionar una anotación más completa y precisa.
- Mejorar la precisión y sensibilidad del análisis de expresión diferencial mediante el desarrollo de modelos de aprendizaje automático que puedan identificar genes expresados diferencialmente con mayor precisión y solidez, sobre todo en diseños experimentales complejos.
- Identificar las vías y redes clave que están desreguladas en diferentes condiciones mediante la integración de diferentes tipos de datos ómicos y el uso de enfoques basados en redes para identificar centros y reguladores clave, entre otros.

Para todos estos procesos utilizamos desde técnicas convencionales y bien conocidas, como Random Forest o Support Vector Machines (SVM), hasta Multimodal Self-Supervised Learning, Deep Energy-Based Generative Models, Variational Prototype Inference, Deep Generative Models o Variational Autoencoder, modernos algoritmos que están suponiendo un cambio radical.

HERMES:

Ya en el ámbito de las neurociencias, formamos parte de un extraordinario proyecto europeo que utiliza IA, modelos basados en Transformers, para generar un predictor que (Hernández *et al.*, 2023), observando la actividad del córtex cerebral mediante señales de EEG (García-Retuerta *et al.*, 2020), es capaz de avisar de la ocurrencia de un ataque epiléptico minutos antes de su comienzo. Esta predicción permite la actuación temprana sobre el ataque con el objetivo de llegar a evitarlo. El modelo se entrena con un conjunto de datos etiquetados y el resultado es una clasificación binaria que indicaría si se va a producir un ataque o no. En base a esta predicción, un segundo agente de IA actúa como algoritmo de control de un dispositivo de estimulación cerebral (Canal-Alonso *et al.*, 2022) que interviene en milisegundos evitando las crisis epilépticas y permitiendo que un

implante de células madre se desarrolle en neuronas y células del sistema nervioso sanas.

Del mismo modo, aprovechamos las altas capacidades de análisis de secuencias largas que nos brindan las LSTM para realizar modelos similares que nos permitan comparar el rendimiento de estas arquitecturas. También estamos trabajando en la construcción de modelos mixtos que aprovechen las fortalezas de cada uno de estos tipos de redes neuronales.

Los trastornos cerebrales afectan a mil millones de personas en todo el mundo, siendo las enfermedades neurodegenerativas las que tienen el mayor impacto. La epilepsia y los accidentes cerebrovasculares, con una prevalencia del 11,2 por ciento en Europa, representan la afección más invalidante, superando al VIH, el cáncer y la isquemia cardiaca.

Los pacientes, familiares y sus cuidadores se enfrentan a retos diarios debido al deterioro cognitivo y físico. Los efectos secundarios de las intervenciones médicas o quirúrgicas pueden afectar aún más a la calidad de vida del paciente. Por lo tanto, los trastornos cerebrales incurables tienen consecuencias significativas en la sociedad y en la atención de la salud pública y exigen intervenciones innovadoras y urgentes.

La medicina regenerativa surge en este contexto como una rama prometedora de la ciencia de la salud que tiene como objetivo restaurar la función fisiológica del cerebro reconstruyendo o reemplazando las áreas cerebrales afectadas cuando los tratamientos canónicos han fracasado. Sin embargo, la reparación de un cerebro disfuncional es uno de los retos más difíciles en la investigación de la salud y, a pesar de los resultados alentadores, hasta la fecha somos incapaces de reconstruir eficazmente la materia cerebral. Se trata de un proyecto con grandes implicaciones éticas y que requiere un gran control sobre la IA y lo que esta determine de cara a evitar influir en el comportamiento humano. Este proyecto supone un reto al que nos hemos apuntado y en el que estamos aprendiendo mucho de otros grupos expertos en tratamiento de células madre, neurociencias y microelectrónica.

Reverse Vaccinology:

Las vacunas han transformado y revolucionado la salud pública desde su descubrimiento. En 1921 se desarrolló la primera vacuna contra la tuberculosis; en 1938, para el tétanos; en 1945 apareció la primera vacuna contra la gripe causada por influenza; en 1963, contra el sarampión, y así sucesivamente con tantas vacunas desarrolladas en el siglo XX hasta la aparición de las últimas grandes vacunas contra

el COVID-19 o el ébola. Si hay algo de lo que se ha hablado en el último lustro es de COVID y vacunas. Desde luego, parece que esta última pandemia ha acelerado los procesos de creación y distribución de estos productos.

El método tradicional de diseño de vacunas se ha basado en el cultivo controlado del organismo patógeno en el laboratorio y su posterior caracterización antigénica mediante métodos bioquímicos, serológicos y genéticos, lo cual es un proceso arduo y costoso en términos de tiempo y recursos. Por lo general, abarca un gran número de pruebas de laboratorio que pueden alargar el proceso desde los cinco hasta los quince años, sumando todas las etapas de testado preclínico en modelos animales hasta su validación en fases clínicas. Esto, unido a las innumerables complicaciones debido a factores genéticos y/o sociopolíticos, hace que el desarrollo de una nueva vacuna contra un patógeno emergente o ya establecido sea muy costoso, dificultando llegar a un nivel de protección efectiva para esa enfermedad.

En las últimas décadas ha surgido un sistema de diseño de vacunas basado en técnicas computacionales, algoritmos de aprendizaje, bases de datos con ingentes cantidades de información biológica y recursos de *hardware* optimizados, cuya importancia ya

se está haciendo notar en este campo. Se conoce como diseño de vacunas computacional a todo aquel sistema o flujo de trabajo que, mediante el uso de estas tecnologías, consigue obtener un candidato o conjunto de candidatos de vacuna para su posterior testeo de una forma rápida, automática, precisa y mucho menos costosa en tiempo y recursos. La llegada del *big data* en términos ómicos gracias a los avances en tecnologías de secuenciación de nueva generación, que permiten obtener genomas completos en apenas unas horas, ha permitido obtener en poco tiempo una gran cantidad de secuencias de organismos patógenos sobre los que trabajar e investigar para conseguir vacunas, revolucionando por completo los procesos de vacunación.

En relación con el diseño in silico de vacunas existe un concepto desarrollado por primera vez por Rino Rappuoli en el año 2000 conocido como Reverse Vaccinology (RV), o vacunología inversa, aprovechando las fortalezas de la bioinformática y la ciencia de datos. El objetivo de esta estrategia, y de nuestro proyecto actual, es partir del genoma completo del organismo patógeno en cuestión e ir seleccionando en diferentes etapas, a través de un *pipeline* o *workflow*, los elementos más adecuados que generen una respuesta inmune potente dentro de una vacuna, que pueden ser proteínas, genes

completos o fragmentos denominados péptidos.

Pues bien, siendo conocedores de estas posibilidades, y también por lo atractivo del problema y las oportunidades que ofrece en nuestro ámbito de trabajo, nos hemos decidido a crear nuestro propio *pipeline*, nuestro gemelo digital, dedicado a la creación de vacunas de forma computacional. En todas las fases del proceso de generación de estas vacunas, de forma computacional se pueden utilizar algoritmos de IA para la secuenciación del genoma (i. e. SVM, LSTM), la identificación de candidatos de vacunas (i. e. Transformes, redes bayesianas o *autoencoders*), las labores de genómica comparativa (i. e. Deep Energy-Based Generative Models, Variational Prototype Inference), el análisis de homología de secuencias (i. e. Variational Autoencoder, Multimodal Self-Supervised Learning) o la predicción de epítomos (i. e. redes neuronales de grafos (GNN) o Recurrent Networks). Además, por ejemplo, el procesamiento de lenguaje natural puede utilizarse para extraer información sobre proteínas patógenas, respuestas inmunitarias del huésped y desarrollo de vacunas a partir de publicaciones científicas.

En general, las herramientas bioinformáticas utilizadas en la fase de identificación de vacunas candidatas de la vacunología inversa permiten un enfoque más específico del diseño ~~de vacunas~~ al

identificar posibles dianas vacunales con más probabilidades de provocar una respuesta inmunitaria. Sin embargo, parece claro que el uso de herramientas bioinformáticas no está exento de limitaciones y que sigue siendo necesaria la validación experimental para confirmar la eficacia y seguridad de las vacunas resultantes.

Con nuestro trabajo, pretendemos que nuestros gemelos digitales modelen, por ejemplo, el acoplamiento molecular para identificar de forma eficiente posibles dianas vacunales que puedan unirse a receptores inmunitarios o moléculas CMH y provocar una respuesta inmunitaria.

Uno de los principales retos a la hora de utilizar la IA en la fase de identificación de candidatos a vacunas de la vacunología inversa es la calidad y exhaustividad de los datos. Los datos genómicos y proteómicos pueden ser ruidosos e incompletos y contener errores y sesgos que pueden afectar a la precisión de los algoritmos de IA. Además, la disponibilidad de grandes conjuntos de datos de alta calidad sigue siendo una limitación para muchos patógenos. Otro reto es la interpretación de los resultados obtenidos a partir de algoritmos de IA, que puede resultar difícil y requerir conocimientos de expertos para su validación.

En general, la vacunología inversa es un enfoque prometedor

para el diseño de vacunas que permite un enfoque más rápido y más específico para su desarrollo. Se trata de un método relativamente nuevo y un reto para nuestro equipo, en el que nos hemos embarcado y al que dedicaremos parte de nuestro tiempo durante los próximos años con el objetivo de optimizar el proceso y garantizar la seguridad y eficacia de las vacunas resultantes.

Epílogo

La inteligencia artificial ha avanzado significativamente en los últimos años y su impacto en la sociedad y la economía es innegable. La creciente adopción de la IA en numerosos ámbitos y, en especial, en el campo de la salud y la medicina en general está ofreciendo nuevas alternativas e importantes avances. Esta tendencia continuará en el futuro, impulsada por la necesidad de mejorar la productividad y la competitividad y por el aumento de la disponibilidad de datos y recursos informáticos.

Otro aspecto que destacar es la forma en la que evoluciona la IA. En particular, la investigación se está centrando en el desarrollo de sistemas de IA más adaptables, interpretables, éticos y globales.

Esto implica avanzar en la comprensión de cómo funcionan los modelos de IA y cómo se toman las decisiones, así como en el diseño de sistemas que sean capaces de adaptarse a situaciones cambiantes y de aprender con menos datos. También se están explorando nuevas formas de garantizar la privacidad y la transparencia en el uso de los datos, así como de evitar sesgos y discriminación en la toma de decisiones. Gracias a tecnologías complementarias como los espacios de datos, los gemelos digitales, el aprendizaje federado, la *blockchain* o el *edge computing* veremos una nueva generación de productos y servicios que en breve cambiarán nuestra forma de trabajar e, incluso, de relacionarnos.

Aunque nos enfrentamos a ciertos riesgos, las ventajas que trae la IA son caras y se están manifestando de forma constante en nuestro día a día. Los nuevos modelos globales, capaces de resolver una amplia variedad de proyectos, cambiarán nuestra forma de trabajar y, en gran medida, muchos ámbitos de nuestra vida. Es importante que la sociedad y los gobiernos estén atentos a los riesgos, los desafíos y el potencial que nos brinda la IA y establezcan modelos de regulación adecuados. Hoy, la regulación de la IA es fragmentada y varía entre países y sectores. En el futuro, se requerirán marcos regulatorios más amplios y consistentes que aborden cuestiones como la privacidad, la

seguridad, la transparencia y la responsabilidad. Es importante que estas regulaciones sean diseñadas de manera que no limiten la innovación y el progreso en la IA, pero, al mismo tiempo, aseguren que la tecnología se utilice de manera adecuada, sobre todo en el ámbito de la medicina.

El futuro de la IA es prometedor y nos esperan grandes cambios a corto y medio plazo de los que todos seremos partícipes. Para mí, hoy se abre un nuevo camino, el de la Real Academia de Medicina de Salamanca, uno más de los que pasearé a diario esperando aportar a la RAMSA todo lo que sé, y siempre con ayuda del grupo de investigación que coordino para estar a la altura de este nombramiento.

Referencias

- Bajo, J., & Corchado, J. M. (2009). Thomas: Practical applications of agents and multiagent systems. In *Bio-Inspired Systems: Computational and Ambient Intelligence: 10th International Work-Conference on Artificial Neural Networks, IWANN 2009, Salamanca, Spain, June 10-12, 2009. Proceedings, Part I* 10 (pp.

512-513). Springer Berlin Heidelberg.

- Bengio, Y. (2009). Learning deep architectures for AI. *Foundations and trends[®] in Machine Learning*, 2(1), 1-127.
- Canal-Alonso, Ángel, Casado-Vara, R., Castellano, O., Herrera-Santos, J., Gonçalves, J., Márquez-Sánchez, S., ... Corchado, J. M. An affordable implantable vagus nerve stimulator system for use in animal research. Royal Society. <https://doi.org/10.1098/rsta.2021.0010>
- Carroll, L. (1958). *Symbolic logic and the game of logic* (Vol. 1). Courier Corporation.
- Cohen, A. O., Nussenbaum, K., Dorfman, H. M., Gershman, S. J., & Hartley, C. A. (2020). The rational use of causal inference to guide reinforcement learning strengthens with age. *npj Science of Learning*, 5(1), 16.
- Corchado, J. M., Díaz, F., Borrajo, L., & Fernández, F. (2000). *Redes neuronales artificiales. Un enfoque práctico*. Servicio de Publicacións da Universidade de Vigo.
- Díaz, F., Fernández-Riverola, F., & Corchado, J. M. (2006). gene-CBR: A Case-Based Reasoning Tool for Cancer Diagnosis Using Microarray Data Sets. *Computational Intelligence*, 22(3-4), 254-

268.

- García-Retuerta, D., Canal-Alonso, A., Casado-Vara, R., Rey, A. M., Panuccio, G., Corchado, J. M. (2021). Bidirectional-Pass Algorithm for Interictal Event Detection. In: Panuccio, G., Rocha, M., Fernández-Riverola, F., Mohamad, M., Casado-Vara, R. (eds). Practical Applications of Computational Biology & Bioinformatics, 14th International Conference (PACBB 2020). PACBB 2020. Advances in Intelligent Systems and Computing, vol 1240. Springer, Cham. https://doi.org/10.1007/978-3-030-54568-0_20
- Hendler, J., & Mulvehill, A. M. (2016). Social machines: the coming collision of artificial intelligence, social networking, and humanity. Apress.
- Hernández, M., Canal-Alonso, Á., de la Prieta, F., Rodríguez, S., Prieto, J., Corchado, J. M. (2023). Machine Learning and Deep Learning Techniques for Epileptic Seizures Prediction: A Brief Review. In: Fernández-Riverola, F., Rocha, M., Mohamad, M. S., Caraiman, S., Gil-González, A. B. (eds). Practical Applications of Computational Biology and Bioinformatics, 16th International Conference (PACBB 2022). PACBB 2022. Lecture Notes in Networks and Systems, vol 553. Springer, Cham.

https://doi.org/10.1007/978-3-031-17024-9_2

- Huang, X., Hu, Y., Luo, X., Han, J., Zhang, B., & Cao, X. (2022). Boosting Variational Inference with Margin Learning for Few-Shot Scene-Adaptive Anomaly Detection. *IEEE Transactions on Circuits and Systems for Video Technology*.
- Jyoti, A., Schwing, A., Kautz, J., & Vahdat, A. (2022). A Contrastive Learning Approach for Training Variational Autoencoder Priors. *NEURIPS*.
- Li, H., Ke, Q., Gong, M., & Drummond, T. (2023). Progressive Video Summarization via Multimodal Self-supervised Learning. In *Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision* (pp. 5584-5593).
- Maes, F., Robben, D., Vandermeulen, D., & Suetens, P. (2019). The role of medical image computing and machine learning in healthcare. *Artificial intelligence in medical imaging: opportunities, applications and risks*, 9-23.
- McCarthy, J., Minsky, M. L., Rochester, N., & Shannon, C. E. (2006). A proposal for the dartmouth summer research project on artificial intelligence, august 31, 1955. *AI magazine*, 27(4), 12-12.

- Moxon, R., Reche, P. A., & Rappuoli, R. (2019). Reverse vaccinology. *Frontiers in immunology*, 10, 2776.
- Nadkarni, P. M., Ohno-Machado, L., & Chapman, W. W. (2011). Natural language processing: an introduction. *Journal of the American Medical Informatics Association*, 18(5), 544-551.
- Reeves, C. R. (2010). Genetic algorithms. *Handbook of metaheuristics*, 109-139.
- Russell, S. J., & Norvig, P. (2014). *Artificial intelligence: a modern approach*, Harlow.
- Searle J. R. (2008) *Mind, language and society: Philosophy in the real world*. Basic books.
- Shortliffe, E. H. (1974). MYCIN: a rule-based computer program for advising physicians regarding antimicrobial therapy selection. Stanford Univ Calif Dept of Computer Science.
- Strokach, A., & Kim, P. M. (2022). Deep generative modeling for protein design. *Current opinion in structural biology*, 72, 226-236.
- Tran, M. H., Gomez, O., & Fei, B. (2023, April). A video transformer network for thyroid cancer detection on hyperspectral histologic images. In *Medical Imaging 2023: Digital and*

Computational Pathology (Vol. 12471, pp. 32-41). SPIE.

- Turing A. M. (1950) “Computing machinery and intelligence”. In: *Mind* 59.236, pp. 433–460.
- Van der Velden, B. H., Kuijf, H. J., Gilhuijs, K. G., & Viergever, M. A. (2022). Explainable artificial intelligence (XAI) in deep learning-based medical image analysis. *Medical Image Analysis*, 102470.
- Xu, Y. (2022). *Deep Energy-Based Generative Modeling and Learning*. University of California, Los Angeles.
- Yang, J., Walters, R., Dehmamy, N., & Yu, R. (2023). Generative Adversarial Symmetry Discovery. arXiv preprint arXiv:2302.00236.
- Zadeh, L. A. (2008). Is there a need for fuzzy logic? *Information sciences*, 178(13), 2751-2779.

DISCURSO DE CONTESTACION

del

Excmo. Dr. D. Francisco S. Lozano Sánchez

Académico de Número de la Real Academia de Medicina de
Salamanca

Excmo. Sr. Presidente,
Excmos Sres. Presidentes de Honor,
Excmas e Ilmas. Autoridades,
Excmos e Ilmos. Sras. y Sres. Académicos,
Señoras y Señores.

Mis primeras palabras deben ser de agradecimiento a la Junta de Gobierno de esta Real Academia de Medicina de Salamanca, por haberme designado para contestar al discurso de ingreso del nuevo académico. Conjuntamente debo manifestar mi satisfacción por tan honroso encargo, especialmente por la relación que me une con el beneficiario.

Decía, el sagaz escritor y dramaturgo madrileño, Enrique Jardier Poncela “Todo discurso se articula en dos partes: el exordio y el incordio”. Estén tranquilos, pues quien les habla

ha estructurado su discurso de contestación de forma más convencional.

La tradición dicta que el discurso de ingreso de un nuevo miembro en cualquiera de las Reales Academias de España sea respondido por otro académico, que primero, objective las razones que llevaron a su admisión en la institución, glosando sus méritos y, en segundo lugar valore y en ciertos casos discuta los contenidos científicos de su disertación.

Una *laudatio* o elogio siempre debe resultar fácil de hacer en la medida en que, por definición, se realiza en honor de quien reúne infinitamente más méritos y es bastante más conocido que quien intenta su alabanza.

Es con todo frecuente que el padrino aproveche la situación para poner de manifiesto la estrechísima relación que le une con el homenajeadado, astuta forma de sugerir que si se ha merecido su atención, uno tendrá también parecidos méritos. No abusaré de esa tendencia: al fin de cuentas sólo conozco al Prof. Corchado desde hace unas pocas décadas, cuando realizó para el departamento de cirugía del profesor Gómez Alonso un sistema, entonces avanzado, de historia clínica, que a día de hoy está totalmente obsoleto y por tanto abandonado; pero,

además, sería tremendamente injusto con la tarea encomendada, si intentara apoderarme de una milésima parte de sus merecimientos.

Gran parte de mi azoramiento deviene del hecho de que he oído y leído suficientes laudatios sobre el Prof. Corchado, nuestro recipiendario, como para no sentirme oprimido por el intento de hacerle justicia, pues han expuesto sus méritos personas de mucha más relevancia que quien ahora les habla. No obstante, intentaré resumirles su amplio y brillante *Curriculum Vitae*, si bien seguramente me dejaré numerosos aspectos por mencionar, que espero me perdone el Dr. Corchado.

Juan Manuel Corchado Rodríguez nace en Salamanca un 15 de mayo de 1971 (es decir tiene 52 años recientemente cumplidos). Actualmente es Catedrático del área de Ciencias de la Computación e Inteligencia Artificial de la Universidad de Salamanca. Doctor en Ciencias de la Computación por la Universidad de Salamanca y Doctor en Inteligencia Artificial por la *West University* de Escocia. Ha sido Vicerrector de Investigación y Director del Parque Científico de la Universidad de Salamanca. Elegido dos veces como Decano de la Facultad de Ciencias de la citada Universidad.

Dirige el grupo de investigación BISITE, acrónimo que significa “Bioinformática, Sistemas Inteligentes y Tecnología Educativa”. También es director del *IOT Digital Innovation Hab*. Presidente del *AIR Institute*, y líder del proyecto en el consorcio *DIGIS3*. Corchado también es profesor visitante en el Instituto Tecnológico de Osaka, profesor visitante en la Universidad de Malasia Kelantan y miembro del *Advisory Group Online Terrorist Propaganda de la European Counter Terrorism Centre*, más conocida como la EUROPOL.

Fue presidente de la asociación *IEEE Systems, Man and Cybernetics*, y coordinador académico del Instituto Universitario de Investigación en Arte y Tecnología de la Animación de la Universidad de Salamanca, e investigador en diversas universidades del Reino Unido y España.

En la actualidad compagina toda esta ingente actividad con la dirección de diferentes programas de Master en nuestra Universidad, como el de “Seguridad en Internet”, o el de “Diseño e impresión en 3D”.

Corchado desarrolla además trabajos, en numerosos países, en proyectos relacionados con la Inteligencia Artificial (IA), o el *Machine Learning*. entre otros muchos.

Por la amplitud y nivel de sus aportaciones, mencionaré sólo superficialmente, su participación en congresos y reuniones, así como sus contribuciones en revistas científicas. Parece suficiente referir que posee 114 artículos en revistas del primer cuartil (Q1). Numéricamente su Scholar H Index es de 106 (34.885 citas, 20.459 desde 2017), un Index i10 de 599 (445 desde 2017), y un Scopus H Index de 50 (9.219 citas - 686 artículos citados). En resumen, según los rankings al uso, es uno de los principales investigadores de la Universidad española y por tanto de las salmantinas

Conjuntamente ha dirigido 42 tesis doctorales en los últimos 10 años. Sus proyectos de investigación competitivos, de toda índole, en los que es investigador principal (IP), son incontables. A pesar de su juventud posee 4 sexenios de investigación y uno de transferencia. Ya les indiqué con anterioridad que esta *laudatio* sería muy fácil de realizar.

Es significativo también su trabajo como editor-jefe de tres importantes revistas de su especialidad: *Advances in*

Computing and Artificial Intelligence Journal, Oriental Journal of Computer Science and Technology, y Electronics.

Su actividad profesional se extiende por numerosos países de los cinco continentes, y debo confesarles que en este último año casi siempre que hablamos por teléfono le pillo trabajando en Qatar, Arabia Saudi, Singapur o Kazagistán. Querido amigo me salen caras tus llamadas al móvil.

Pero, ¿Cómo comenzó este magnífico historial académico y profesional?. Como el mismo ha referido en su discurso, todo comenzó con una beca Erasmus que lo llevó a una universidad escocesa. Después de finalizar sus estudios en la USAL, se trasladó a la Universidad de Vigo, donde fue subdirector del departamento de informática. En el año 2000, se incorpora al departamento de informática y automática de la USAL; aquí ha pasado sus dos décadas más productivas.

Dice Corchado “llevo 30 años de relación con la IA, cuando en tercero de carrera escuché por vez primera este término”; esa afirmación remarca cómo muy pronto le sugestionó y atrapó la IA. Posteriormente, como el mismo sigue refiriendo, conoce bien esta especialidad “de la mano de los pioneros”, siendo

participe de la evolución de esta disciplina en las últimas tres décadas.

Finalizó su amplio y brillante curriculum vitae, comentando cómo recientemente la prensa local, se hacía eco de su nombramiento como patrono de la **Fundación AstraZeneca**, para junto con otros profesionales sanitarios e investigadores de reconocido prestigio, puedan acercar el conocimiento científico a la sociedad. Objetivo este que compartimos en nuestra Academia.

Después de haber escuchado el excelente discurso del Prof. Corchado titulado “El despertar de una inteligencia global”, he de seguir cumpliendo con las normas protocolarias contestando al trabajo expuesto.

El nuevo académico nos ha presentado un tema muy novedoso con claridad, precisión, concisión, corrección y sobre todo con convicción, entusiasmo y acción. Como dijo uno de los componentes de los Beatles, Paul McCartney: “Pensar globalmente, actuar localmente”.

Todo comienza por un título novedoso y sugestivo. La expresión IA fue creada por el informático e inventor

estadounidense John McCarthy en 1956 y comenzó a usarse en español en los años sesenta del pasado siglo. Pues bien, cincuenta años después, la expresión IA ha sido elegida palabra del año 2022 por la Fundeu-RAE; los motivos de su elección han sido su importante presencia en los medios de comunicación durante los últimos doce meses, así como en el debate social. En este sentido recordar, que las anteriores expresiones ganadoras fueron confinamiento (2020) y vacuna (2021); no creo precisen mayor explicación.

El Dr. Corchado comienza su discurso, como no puede ser de otra manera, presentado los conceptos y la historia de la IA. Define la IA como un campo de la informática y la tecnología, es decir se refiere a sistemas o máquinas que imitan la inteligencia humana para realizar tareas y que pueden mejorar iterativamente a partir de la información que recopilan.

Sobre la historia de la IA, se centra en lo que denomina “primavera de la IA” y señala cómo estamos inmersos en la quinta revolución industrial con la adopción de tecnologías, tales como la IA, el aprendizaje automático, el internet de las cosas, la robótica avanzada, la impresión 3D y la nanotecnología. Manifiesta claramente que esta nueva

revolución transformará el mundo que actualmente conocemos.

En este contexto la prestigiosa revista *Science* acaba de publicar el primer mapa de conexiones neuronales del cerebro de la larva de la mosca de la fruta (*Drosophila melanogaster*), que parece ser que abre la puerta a modelos computacionales de IA.

En esa línea, la IA en medicina, es el uso de modelos de aprendizaje automático para buscar datos médicos y descubrir conocimientos que ayuden a mejorar los resultados de salud. Así, gracias a los avances en ciencias de la computación e informática, la IA se está convirtiendo en una parte integral de la moderna atención médica. Actualmente, las funciones más comunes de la IA, en entornos médicos, son el apoyo a la toma de decisiones clínicas y el análisis de imágenes.

Precisamente los apartados del discurso más importantes para nosotros, son aquellos que el Dr. Corchado dedica, incluso con ejemplos, a la relación IA con la Medicina. La medicina personalizada, la medicina de precisión, la monitorización y hospitalización a domicilio, etc., tendrán un importante desarrollo en los próximos años.

En este sentido, la cirugía vascular, mi especialidad, no escapa a este desarrollo, como lo muestran diferentes publicaciones al respecto y donde destaca una revisión sistemática publicada en *Annals Vascular of Surgery*, que refiere la existencia de una importante literatura, en crecimiento exponencial, del uso de la IA en cirugía vascular. La revisión recoge 249 publicaciones, con un 63% de estudios en los últimos 5 años. El enfoque de la IA en cirugía vascular se centra fundamentalmente en la enfermedad carotídea (47%) y la aorta abdominal (20%). La aplicación de la IA está relacionada con métodos de diagnóstico (37%) y modelos predictivos (22%). Las redes neuronales y las máquinas de vectores de soporte componen la mayoría de los métodos de IA en la literatura vascular. Es relevante destacar que el 59% de las publicaciones aparecieron en revistas con un alcance relacionado con la informática o la ingeniería y, que sólo el 9% de estudios se publicaron en revistas específicas de cirugía vascular. Por todo ello es importante que los cirujanos vasculares aprecien el potencial y limitaciones de la IA. Además, a medida que aumente su uso, existe la necesidad de médicos y cirujanos con experiencia en métodos de IA que puedan optimizar su transición a la práctica diaria.

Volviendo al discurso de ingreso, no olvida Corchado resaltar los problemas por resolver de la IA, como aquellos relacionados con la protección de datos y sobre todo los desafíos éticos de la IA, tales como la privacidad, la seguridad o el impacto de la IA en el mercado laboral y en la sociedad en general. Es tranquilizante escucharle cuando nos dice que se está trabajando para cuidar todos estos aspectos, utilizando las denominadas técnicas de automatización ética y justa; que en el caso de la medicina pasan también por las denominadas “Buenas Prácticas”.

En nuestra opinión, cuestionarnos el desarrollo que puede llegar a tener la IA, ventajas y desventajas, es clave para saber hacia dónde tenemos que ir en este desarrollo tecnológico. Quizás influenciado por este momento, demasiadas veces leo o escucho en los medios de comunicación el impacto mediático que tiene la IA; no se cansan estos medios en describir las “luces y sombras de la IA” y señalar/resaltar tópicos tales como “las máquinas dominarán el mundo”.

Personalmente me tranquiliza la opinión de Isaac Asimov cuando dice “Si el conocimiento puede crear problemas, no es con ignorancia con la que podremos resolverlos”. Por ello, es importante dejar patente que la IA no debe tomar decisiones

sobre los humanos sin que este proceso haya sido supervisado por personas.

En este sentido, el nuevo Código de Deontología Médica, aprobado muy recientemente por Consejo General de Médicos, incluye nuevos capítulos en ámbitos como la Telemedicina, Big Data e Inteligencia Artificial. Respecto a la IA, el artículo 86.1 dice literalmente “Los datos de salud extraídos de grandes bases de datos sanitarias o los sistemas robóticos, pueden servir de ayuda en la toma de decisiones clínicas y sanitarias, pero no sustituyen a la obligación que el médico tiene de utilizar los métodos necesarios para la buena práctica profesional”.

Después de presentar al Prof. Corchado (a partir de este momento Académico de Número en la especialidad de Inteligencia Artificial, Sección 7: Ciencias Afines) y comentado su Discurso de Ingreso me permito la licencia de dirigirme a los padres del nuevo académico para felicitarles por su hijo, al que hoy tenemos el privilegio de acogerle entre nosotros como compañero de Academia.

También deseamos transmitir nuestro afecto y felicitación a tu esposa y tres hijos, a quienes a buen seguro les has sustraído un montón de horas (y añadiría, y las que les quedan).

Finalmente, Profesor Dr. D. Juan Manuel Corchado Rodríguez, reciba de todos los miembros de esta Corporación, así como en el mío propio, la más afectuosa felicitación y que su incorporación nos sirva como modelo para un mayor esplendor de los principios con que se constituyó esta Academia.

Muchas gracias por su atención.

He dicho.

DISCURSO DE CLAUSURA

Francisco S. Lozano Sánchez

Presidente de la RAMSA

La RAMSA está de enhorabuena. Esta tarde no solo incorporamos a un nuevo académico, sino que debido a su perfil científico y profesional, fortalecemos un área que cada vez tendrá mayor importancia en las Academias Médicas.

Como muchos de ustedes conocen, las Academias Médicas están formada por secciones. La mayoría de secciones son clásicas (medicina o cirugía y sus especialidades, etc.), existiendo una sección denominada “ciencias afines”, que en nuestra opinión cobra cada día mayor dimensión y protagonismo.

De todos es conocido que la moderna medicina crece, no solo por sus propias áreas de conocimiento, sino por aquellas como la que hoy incorporamos. No descubrimos nada si decimos que los avances científicos y la evolución tecnológica aportan ventajas e inconvenientes a la sociedad en general y a nuestros pacientes en particular.

Estos cambios pueden modificar la forma de entender la relación médico-paciente. Afortunadamente muchos médicos, e incluso residentes, todavía llevan y usan el estetoscopio y lo que supone no perder el contacto con la medicina tradicional. Precisamente una de las misiones de la Academia es vigilar y analizar estos cambios, para ello precisamos contar con recursos humanos maduros y expertos en los temas a evaluar.

Hoy la RAMSA ha dado un paso en este sentido.

Muchas gracias por su atención.