



Artículo especial

Inteligencia Artificial aplicada a los cuidados (I): origen y conceptos generales



Artificial Intelligence applied to care (I): origin and general concepts

Reyes Grangel^{a,*}, María Victoria Ibáñez Gual^b, Amelia Simó Vidal^b,
Juan Carlos Amengual Argudo^a y Óscar Belmonte Fernández^a

^a Grupo de Investigación en Aprendizaje Automático para Entornos Inteligentes. Departamento de Lenguajes y Sistemas Informáticos, Universitat Jaume I, Castellón de la Plana, España

^b Grupo de investigación Análisis de formas 2D y 3D, Aprendizaje Estadístico en Inteligencia Artificial, Estereología. Departamento de Matemáticas, IMAC, Universitat Jaume I, Castellón de la Plana, España

INFORMACIÓN DEL ARTÍCULO

Palabras clave:

Enfermería
Inteligencia Artificial
Toma de decisiones
Algoritmos

Keywords:

Nursing
Artificial Intelligence
Decision Making
Algorithms

R E S U M E N

Desde la aparición de ChatGPT en noviembre de 2022, la Inteligencia Artificial (IA) ha ganado una notable popularidad, extendiéndose a múltiples ámbitos de la vida diaria y profesional. Sin embargo, la IA abarca mucho más que la inteligencia generativa, incluyendo técnicas que permiten analizar datos y apoyar la toma de decisiones, especialmente en contextos como el sanitario. Este auge hace necesario un esfuerzo de transparencia y explicabilidad de los modelos predictivos y de las técnicas utilizadas, con el fin de generar confianza entre los profesionales sanitarios que deberán aplicarlas.

En este artículo se ofrece un resumen de las principales técnicas predictivas de IA aplicables al proceso asistencial en enfermería. Se describen métodos centrados en los datos (aprendizaje automático, redes neuronales profundas e IA generativa), así como técnicas centradas en análisis de procesos (monitoreo predictivo de procesos). Cada técnica se ilustra mediante ejemplos prácticos del ámbito sanitario, con el objetivo de facilitar su comprensión. Finalmente, se presentan las principales limitaciones de estas técnicas y sus perspectivas de futuro.

A B S T R A C T

Since the emergence of ChatGPT in November 2022, Artificial Intelligence has gained significant popularity, extending to multiple areas of daily and professional life. However, Artificial Intelligence encompasses much more than Generative Intelligence, including techniques that allow data analysis and support decision-making, especially in contexts such as healthcare. This boom necessitates efforts to ensure transparency and explainability of the predictive models and techniques used, to build trust among the healthcare professionals who must apply them.

This article provides a summary of the main predictive Artificial Intelligence techniques applicable to the nursing care process. Data-centric methods are described (Machine Learning, Deep Neural Networks and Generative Artificial Intelligence), as well as process-centric techniques (Predictive Process Monitoring). Each technique is illustrated with practical examples from the healthcare field, with the aim of facilitating their understanding by healthcare professionals. Finally, the main limitations of these techniques and their prospects are presented.

Introducción

El término Inteligencia Artificial (IA) está ahora de moda y se aplica en varias disciplinas científicas, pero siendo estrictos, la IA es un campo de desarrollo e investigación dentro de las ciencias de la computación en el que se lleva trabajando con rigurosidad desde mediados de los años 50.

* Autor para correspondencia.

Correo electrónico: grangel@uji.es (R. Grangel).

<https://doi.org/10.1016/j.enfcli.2026.502486>

Recibido el 28 de noviembre de 2025; Aceptado el 9 de febrero de 2026

Disponible en Internet el 11 de marzo de 2026

1130-8621/© 2026 Los Autores. Publicado por Elsevier España, S.L.U. Este es un artículo Open Access bajo la licencia CC BY-NC-ND (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

Podemos definirla como una disciplina que persigue el desarrollo de algoritmos que permitan modelar ciertos comportamientos de la inteligencia humana con el fin de poder aplicarlos en la construcción de sistemas y aplicaciones informáticas para que la imiten e incluso puedan aprender y mejorar con su uso cotidiano. En este sentido, la IA es muy general y abarca varios campos, como por ejemplo el aprendizaje automático (*Machine Learning*), siendo el que ha tenido más éxito en el desarrollo de sistemas y aplicaciones de IA (fig. 1).

Este artículo comprende varias secciones: en primer lugar, la introducción presenta una breve reseña histórica para, a continuación, establecer una taxonomía de los tipos de IA. En la segunda sección, se explican los fundamentos del aprendizaje automático, haciendo hincapié en su base matemática. Dentro del campo de investigación del aprendizaje automático, se resaltarán las redes neuronales en la tercera sección por sus resultados y se hablará de su evolución con el empleo de múltiples capas y conexiones en lo que se conoce como *deep learning* o aprendizaje profundo. En la cuarta sección se explican los fundamentos que han permitido el desarrollo de la IA generativa y los modelos conversacionales como ChatGPT, que permiten generar nuevos conceptos e ideas a partir de lo ya aprendido. En la quinta sección se abordan conceptos fundamentales de la minería de procesos, un campo relacionado con la minería de datos que permite abordar la monitorización inteligente de procesos (p. ej., en un hospital, los diferentes «caminos» que puede seguir un paciente desde su ingreso hasta su alta) para poder hacer predicciones en función de diferentes estados. Finalmente, se presentan algunas conclusiones y se establecen las limitaciones actuales de la IA y cuáles son las perspectivas para futuros desarrollos.

Historia

El objetivo de la IA es tan antiguo como el de crear máquinas que muestren las mismas destrezas que las personas, pero el término IA no se acuñó hasta 1956, cuando un grupo de investigadores, principalmente matemáticos, se reunió en Dartmouth, EE. UU., con el objetivo de describir con precisión las características de la inteligencia humana para poder construir máquinas que pudiesen utilizarlas.

Desde esta primera reunión hasta nuestros días, la IA ha experimentado un gran avance debido, principalmente, a (1) resultados decisivos de la investigación (como el avance en el reconocimiento automático de voz y el desarrollo de asistentes virtuales) (2) el aumento de la potencia de los ordenadores y (3) la disponibilidad actual de una enorme cantidad de datos digitalizados. Sin embargo, este crecimiento no ha sido lineal. Como si de una montaña rusa se tratara, la investigación en IA ha atravesado etapas de fascinación y entusiasmo desbordante, impulsadas por avances prometedores, seguidas de periodos de olvido y desencanto, a menudo causados por expectativas poco realistas.

A día de hoy, nos encontramos en un momento de altas expectativas, propiciado por los increíbles avances de la IA: en 2016, un programa desarrollado por Alphabet (Google) derrota al campeón de Go, Lee Sedol; a finales de 2022, OpenAI lanza ChatGPT, un chatbot con capacidades conversacionales cercanas a las humanas; en 2024, Demis Hassabis, John Jumper y David Baker reciben el Nobel de Química por la predicción de las estructuras de las proteínas mediante IA; ese mismo año, Geoffrey Hinton y John Hopfield reciben el Nobel de Física por sus trabajos en IA. Solo el tiempo dirá si estas expectativas marcan un punto de inflexión o una nueva vuelta en la montaña rusa de la IA.

Tipos de IA y ejemplos

Utilizando un enfoque generalista, podemos decir que existen tres tipos de IA:

1. **IA estrecha (IAE) o débil.** Permite desarrollar aplicaciones orientadas a resolver una tarea específica en un contexto o dominio concreto, es decir, que no pueden emplearse para realizar tareas fuera de su especialización. Ejemplos: aplicaciones de reconocimiento facial o

análisis de tomografías computarizadas, resonancias magnéticas o ecografías para la detección de tumores. El aprendizaje automático, incluyendo el supervisado, no supervisado y por refuerzo, es una de las técnicas fundamentales para desarrollar aplicaciones de IAE.

2. **IA general (IAG) o fuerte.** Se trata de una IA teórica cuyo objetivo es simular o incluso superar la capacidad intelectual humana. En el futuro, podría comprender, aprender y aplicar conocimiento a tareas diversas, como lo haría una persona. Para ello, emplearía aprendizaje automático y razonamiento para abordar distintos problemas.
3. **Superinteligencia Artificial (SIA).** Es el nivel hipotéticamente más alto y abstracto que podría alcanzar la IA: una «meta inteligencia» capaz de abarcar todos los aspectos de la inteligencia humana, incluyendo creatividad, resolución de problemas y conocimiento general. Totalmente teórica, la SIA plantea importantes cuestiones éticas y de seguridad, ya que podría desarrollar formas de aprendizaje y comprensión más allá de las capacidades humanas.

En la actualidad, el aprendizaje automático es el principal campo de desarrollo e investigación en IAE. Sus algoritmos permiten que los sistemas mejoren su rendimiento en tareas específicas a medida que procesan más datos. Mientras la IA abarca un enfoque más amplio sobre la inteligencia de las máquinas, el aprendizaje automático se centra en cómo estas aprenden a partir de los datos.

Aprendizaje automático

El término aprendizaje automático fue definido en los años 50 por Arthur Samuel¹ como la capacidad de las máquinas de «aprender» o mejorar su desempeño con la experiencia. Así pues, engloba todas las técnicas de la IA basadas en el uso de datos para aprender y tomar decisiones. Estos datos pueden ser de muchos tipos: tablas numéricas, imágenes, textos, etc. Para hacer más sencillas las explicaciones, en este apartado hablaremos únicamente de datos numéricos (peso, tensión arterial, temperatura, etc.) y/o categóricos (sexo, grupo sanguíneo, raza, etc.).

Un algoritmo de aprendizaje automático es un modelo matemático o computacional que, a partir de datos de entrada (características o predictores) X , genera una salida o respuesta Y . El modelo depende de un conjunto de parámetros (o pesos) desconocidos w que el algoritmo va ajustando a partir de los datos (aprende). Si denotamos por Y^{\wedge} las salidas obtenidas por el modelo e Y a las salidas esperadas, el proceso de aprendizaje consiste en encontrar los valores de w que minimicen una función de error E (también llamada de ajuste o pérdida):

$$\text{Min}_w E(Y_1 Y^{\wedge}) \quad (1)$$

En otras palabras, un algoritmo de aprendizaje automático ajusta sus parámetros a partir de los datos de entrada para que las respuestas que genera se acerquen lo máximo posible a las respuestas correctas, minimizando el error entre ambas. Según su objetivo, los algoritmos de aprendizaje automático pueden clasificarse en supervisado o no supervisado.

1. Un algoritmo de aprendizaje supervisado emplea un conjunto de datos de entrenamiento compuesto por entradas X_e y salidas (outputs) conocidas de antemano (datos etiquetados). A partir de ellos, el algoritmo ajusta el modelo que luego se aplica a nuevos datos de entrada X cuyas salidas son desconocidas (no etiquetados), con el fin de predecirlas. Habitualmente, el conjunto de entrenamiento se divide en dos partes: una para ajustar el modelo (entrenamiento) y otra para evaluar su capacidad predictiva (validación). En problemas supervisados, una función de error común es la suma de los cuadrados de las diferencias entre la salida de la red Y^{\wedge} y los valores reales Y_e . En la figura 2 podemos ver un esquema. Los modelos de aprendizaje supervisado se clasifican en:

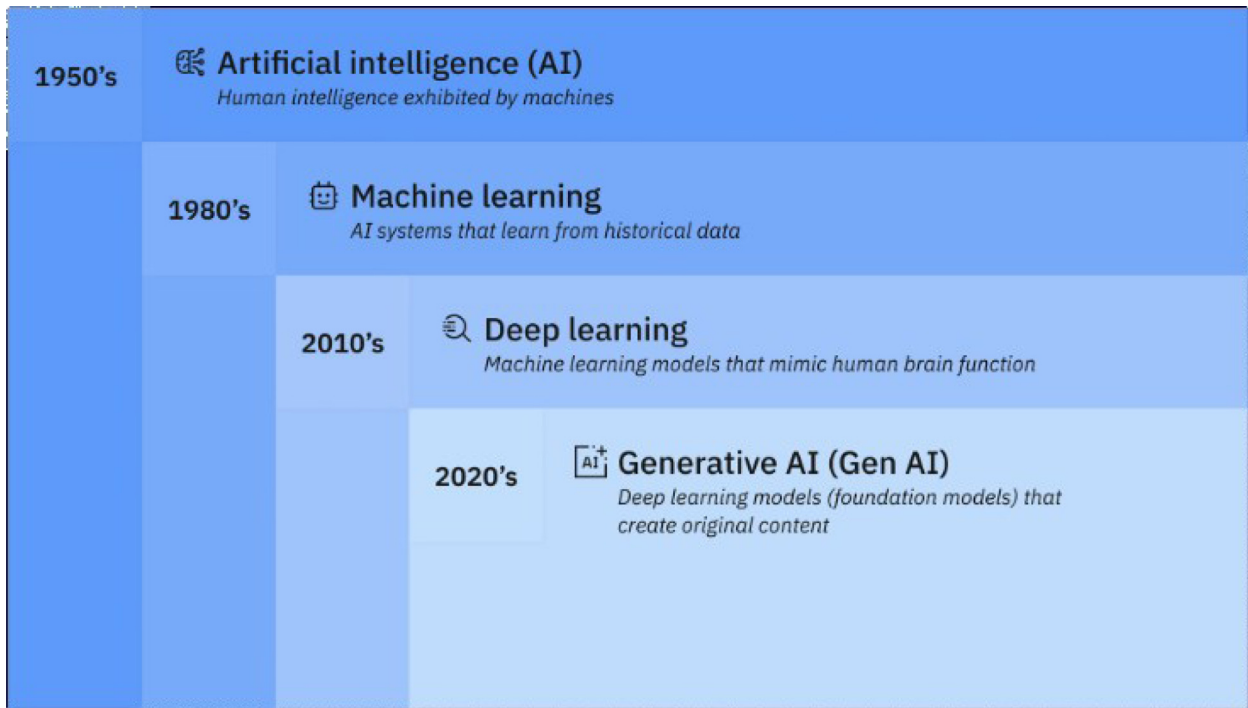


Figura 1. Esquema de la evolución de la IA.

Fuente: IBM Corp.

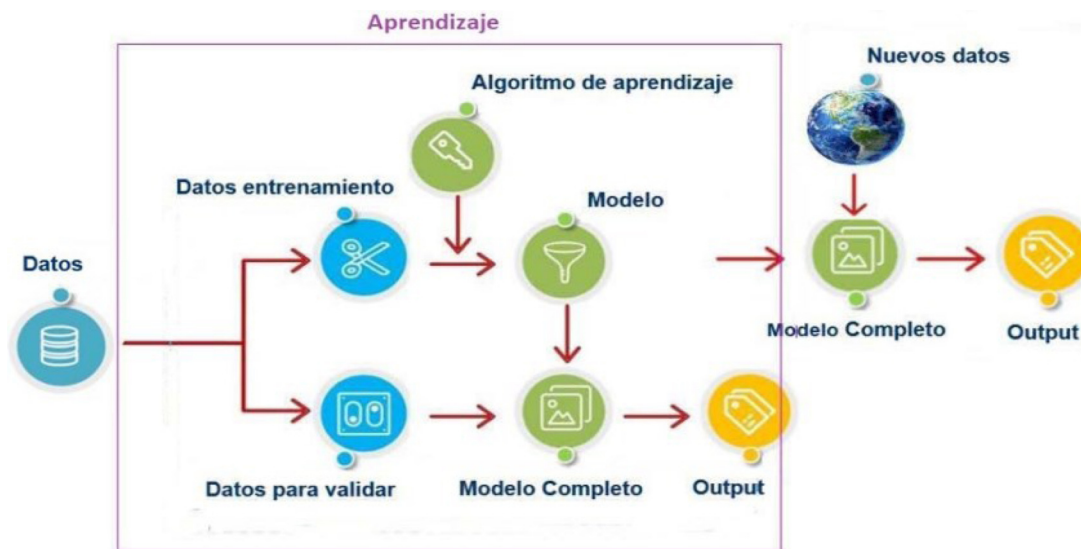


Figura 2. Algoritmo de aprendizaje supervisado.

- a. Modelos de regresión: cuando la salida del modelo, Y , es una característica numérica. En este grupo se incluyen los modelos de regresión lineal, los modelos lineales mixtos, los árboles de regresión y las redes neuronales. Por ejemplo, en Adams et al.² se emplean para predecir el grado de incumplimiento de las medidas de prevención de infecciones (Y) a partir de características del entorno doméstico y organizativo (X), usando datos de una encuesta a 353 enfermeras de atención domiciliaria.
- b. Modelos de clasificación: cuando la salida es una variable categórica con k categorías, es decir, el objetivo es asignar una nueva observación a una de dichas categorías. Modelos de clasificación populares son el análisis discriminante, la regresión logística, el método de Bayes ingenuo, las máquinas de vectores de soporte, los bosques aleatorios y las redes neuronales. Como ejemplo, Brom

- et al.³ usan estos modelos para predecir si un nuevo paciente corre riesgo de reingreso hospitalario (categórica con etiquetas: «sí» o «no») a partir de las características X incluidas en su historial clínico.
 - c. Reglas de asociación: obtienen como salida un conjunto de relaciones entre las características de entrada. Por ejemplo, Peng et al.⁴ aplican reglas de asociación a registros administrativos de salud según grupos de edad. Las reglas obtenidas revelaron asociaciones clínicas relevantes; por ejemplo, en el grupo de 20 a 44 años, entre cardiopatía isquémica crónica e infarto agudo de miocardio.
2. Un algoritmo de aprendizaje no supervisado tiene como objetivo identificar patrones, estructuras o relaciones relevantes en el conjunto de datos de entrada X sin salidas ya conocidas, es decir, «aprende» sin datos de entrenamiento etiquetados. Algunos de ellos se conocen

también como algoritmos de minería de datos. A su vez, los modelos de aprendizaje no supervisado pueden ser:

- a. Modelos de *clustering* o clasificación no supervisada. Su objetivo es formar grupos (p. ej., de pacientes) de forma que los elementos pertenecientes a un mismo grupo sean lo más similares posible en cuanto a las características de entrada X . Los modelos k -medias y jerárquicos son los más populares. Como ejemplo, An et al.⁵ los emplean para desarrollar un sistema de clasificación de pacientes ingresados en unidades de cuidados críticos, en el que las características de entrada X son la gravedad de la enfermedad y las necesidades asistenciales.
- b. Modelos de reducción de dimensión. Su objetivo es simplificar el conjunto de datos de entrada reduciendo el número de características sin perder información. Entre los más utilizados se encuentran el análisis de componentes principales (PCA) y las redes neuronales autoencoders. Como ejemplo, Wanzala et al.⁶ aplican PCA a 151 características del sistema sanitario en Kenia, incluyendo indicadores de recursos humanos, servicios, infraestructuras y finanzas (X). El análisis permitió extraer un número reducido de factores que explican la preparación y eficiencia del sistema sanitario.

Dentro de estos modelos se encuentran algunos modelos clásicos de estadística, que se diferencian del resto por considerar la incertidumbre debida a no disponer de toda la información y la influencia de las variables de entrada en el resultado. También se conocen como modelos de aprendizaje estadístico.

Entre los distintos modelos de aprendizaje automático, los más utilizados en la actualidad son las redes neuronales, debido a su notable capacidad para procesar grandes volúmenes de datos y resolver problemas de alta complejidad. Por este motivo, se dedica un apartado exclusivo a este tipo de modelos.

Redes neuronales: aprendizaje profundo

Dentro de los modelos de aprendizaje automático, una red neuronal puede definirse como un modelo inspirado en la estructura y el funcionamiento de las neuronas y sus conexiones sinápticas en el cerebro humano. Este modelo contiene una serie de nodos (las «neuronas») interconectados, organizados en una o más capas de entrada, varias capas ocultas (por lo general) y, normalmente, una capa de salida. Cada conexión tiene un peso asociado para determinar su importancia y cada nodo una función de activación que puede completarse con un sesgo (p. ej., sumar una cantidad fija al resultado de la función de activación).

Si bien el primer modelo de red neuronal conocido fue el de McCulloch y Pitts en 1943⁷, donde empleaban salidas binarias (verdadero/falso) en los nodos de la red para implementar cualquier función lógica computable, no fue hasta el año 1958 cuando se publicó el primer modelo de computación importante basado en este concepto: el perceptrón, ideado por Rosenblatt⁸ y ya capaz de clasificar patrones. Fue en este artículo donde se desarrolló la idea de utilizar funciones más complejas para realizar cálculos con grupos de neuronas (de entrada, por ejemplo) y luego utilizar funciones de activación con sesgos en la capa oculta para, finalmente, producir la salida correspondiente en cada neurona o grupo de neuronas de la capa de salida.

En los años 60 las expectativas fueron grandes, pero los resultados escasos, debido a las limitaciones inherentes a estos modelos (no solo con redes neuronales). Esto condujo a lo que se conoce como el «invierno de la IA», un lapso prolongado de tiempo en el que no se produjeron avances destacables. En el caso de las redes neuronales, debemos remontarnos hasta mediados de los años 80 para encontrar el siguiente hito importante en su avance: el desarrollo del algoritmo de retropropagación (*backpropagation*) y el concepto de perceptrón *multicapa* por parte de Rumelhart et al.⁹. Este artículo fue clave, pues permitió entrenar redes neuronales más complejas, con muchas más conexiones y un número elevado de capas ocultas.

En la **figura 3** podemos apreciar el diseño de un perceptrón multicapa sencillo, donde tenemos la capa de entrada con cuatro nodos $X (X_1 X_2 X_3 X_4)$, una capa oculta con tres nodos, y la capa de salida con un solo nodo Y (obtenida, además, a partir de una cierta función). Se puede observar que cada nodo de una capa está conectado con cada nodo de la siguiente capa y cada conexión tiene un peso w_{ij} . La retropropagación del error (o *backpropagation*) es el algoritmo que ajusta los pesos de una red neuronal para minimizar el error entre la salida esperada y la obtenida (Ec.(1)). Lo hace iterativamente, actualizando los pesos desde la capa de salida hacia la de entrada, de ahí su nombre.

Dado el alto número de pesos, se requiere una gran cantidad de datos, y el algoritmo puede necesitar desde cientos hasta millones de iteraciones para converger, según la complejidad del problema. Así, las redes neuronales aprenden patrones y relaciones internas que les permiten realizar tareas como clasificación, reconocimiento, predicción o reducción de dimensión.

Aunque el modelo funciona como una «caja negra» debido a sus capas ocultas, el diseño de la red requiere intervención experta humana. El especialista debe definir la *arquitectura* de la red: número de nodos de entrada según cómo se procesen los datos (en imágenes, por ejemplo, agrupar píxeles por regiones de color o tonalidad), conexiones entre nodos, capas ocultas, y nodos de salida.

En los años 90 y principios de los 2000, surgieron diversas arquitecturas de redes neuronales aplicadas a nuevos campos como el procesamiento de imágenes, el procesamiento del lenguaje natural y la traducción automática. Ejemplos destacados son:

1. Las redes neuronales recurrentes (RNN). Están diseñadas para procesar datos secuenciales, incorporando a cada nodo un estado interno que permite tener en cuenta información de pasos anteriores. La salida de cada nodo no depende solo de la entrada actual, sino también de este estado, que puede reflejar múltiples instantes pasados mediante operaciones recurrentes.
2. Las redes neuronales convolucionales¹⁰. Se diseñaron, principalmente, para trabajar con matrices bidimensionales en tareas como clasificación de imágenes y reconocimiento de objetos, aplicaciones especialmente útiles en el ámbito sanitario. Por ejemplo, Chen et al.¹¹ emplean este tipo de red para segmentar imágenes de resonancia magnética en el diagnóstico del Alzheimer.
3. Los algoritmos conocidos como *autoencoders*¹². Son un tipo de red neuronal diseñada para reproducir su entrada en la salida, reduciendo el número de dimensiones en las sucesivas capas intermedias (**fig. 4**), es decir, el modelo aprende una representación eficiente y de menor dimensión de la entrada, conocida como codificación.

En la segunda década del siglo XXI, gracias a los avances en hardware —especialmente los procesadores paralelos de las tarjetas gráficas de Nvidia— y la disponibilidad de grandes conjuntos de datos, Krizhevsky et al.¹³ presentaron una red neuronal profunda con la que ganaron el *ImageNet Large Scale Visual Recognition Challenge* (ILSVRC), superando con amplio margen los métodos anteriores. Su trabajo demostró la eficacia del entrenamiento a gran escala de redes neuronales profundas con *Graphics Processing Unit* (GPU), marcando el inicio del llamado aprendizaje profundo (*deep learning*), caracterizado por arquitecturas con cientos de capas ocultas y miles de millones de parámetros. En esta época surgió la arquitectura *Transformer*¹⁴, que abordaremos en la siguiente sección por su enorme relevancia.

IA generativa: grandes modelos de lenguaje

Los avances en tareas complejas, el aumento exponencial del poder de las tarjetas gráficas, la disponibilidad de grandes volúmenes de datos y las bibliotecas de aprendizaje automático impulsaron el desarrollo del aprendizaje profundo, dando origen a sistemas de IA generativa como ChatGPT. La IA generativa es una rama dentro de la IA cuyos modelos

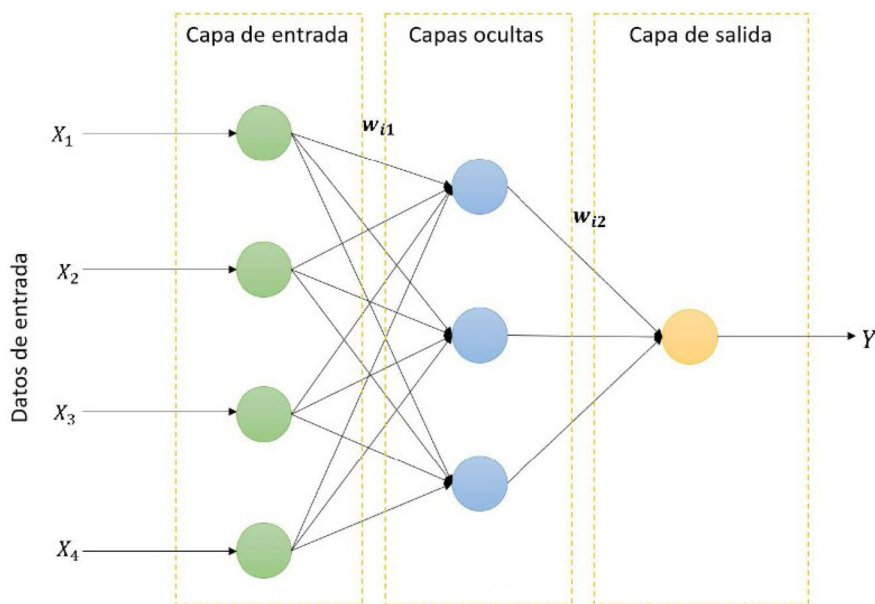


Figura 3. Modelo de red neuronal.

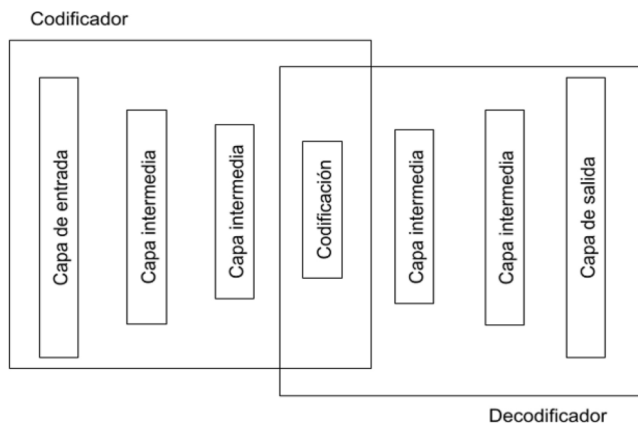


Figura 4. Esquema de la arquitectura autoencoder.

son capaces de generar nuevos datos a partir de los patrones aprendidos durante el proceso de aprendizaje.

Cuando estos modelos son redes neuronales profundas y los datos de entrenamiento son texto, hablamos de grandes modelos de lenguaje (LLM). Típicamente, antes de iniciar el proceso de entrenamiento se realiza un pre-proceso del texto. Parte de este pre-proceso implica limitar el número de palabras distintas con las que se va a trabajar. A este conjunto restringido de palabras se le llama corpus. Si en un determinado texto una palabra no se encuentra dentro del corpus, se sustituye por un símbolo especial cuyo significado es «desconocida».

Una vez restringido el número de palabras con el que se va a trabajar, se crea una representación numérica de cada una de las palabras contenida en el corpus, para que un algoritmo las pueda procesar. Si el corpus contiene, por ejemplo, 10.000 palabras distintas, cada una puede representarse mediante un vector de 10.000 dimensiones, donde únicamente una posición tiene el valor 1 y el resto son ceros. Esta técnica se conoce como codificación one-hot, y permite identificar de forma única cada palabra. Por ejemplo, la primera palabra sería [1,0, 0,..., 0], la segunda [0,1, 0,..., 0], y así sucesivamente.

Sin embargo, este tipo de representación es muy ineficiente, ya que requiere mucho espacio de almacenamiento y una elevada capacidad de cómputo para realizar operaciones con los vectores, por lo que el siguiente paso consiste en reducir el número de dimensiones de

los vectores que representan las palabras, sin perder «el significado» de la representación, utilizando *autoencoders*. A esta nueva representación numérica «reducida» se le denomina con el término inglés «*word embedding*».

Un método muy popular por su potencia para crear *word embeddings* es el algoritmo Word2Vec^{15,16}. Para crear la representación para una palabra, tiene en cuenta el resto de las palabras que están por delante y por detrás hasta una determinada distancia, por ejemplo, de dos palabras. Así, dentro de la frase «El paciente presenta úlceras por presión en el antebrazo izquierdo», para codificar la palabra «presión», se tendrían en cuenta las palabras «úlceras», «por», «en» y «el» y consideraría los siguientes pares de entrenamiento: (presión, úlcera) (presión, por) (presión, el) (presión, en). Este proceso se repite para todas las palabras de todas las frases del corpus, generando miles de pares. A partir de estos pares, Word2Vec entrena una red neuronal pequeña para ajustar los vectores de cada palabra. Después del entrenamiento, cada palabra se identificará con un vector numérico que recogerá su «proximidad» («relación») con otras palabras. Por ejemplo, en el resultado final la palabra «úlceras» podría estar cerca de «herida», «presión», y «dependencia» o la palabra «enfermera» podría estar cerca de «cuidados», «unidad» y «evalúa». Estos vectores pueden usarse para medir similitudes entre palabras (coseno de ángulos entre vectores); para agrupar conceptos similares (*clustering*) o también como entrada en modelos más grandes (p. ej., para clasificar diagnósticos).

En los inicios, las arquitecturas más utilizadas para trabajar con textos en aplicaciones de IA eran las redes convolucionales. Por ejemplo, en el ámbito de la enfermería, Hughes et al.¹⁷ las utilizaron para clasificar automáticamente textos clínicos a nivel de frase. También cobraron protagonismo las RNN, y en particular las *Long-Short Term Memory* (LSTM)¹⁸. Una vez entrenadas, las LSTM son capaces de generar texto coherente y de calidad aceptable: por ejemplo, Baek et al.¹⁹ las utilizan para predecir las necesidades de enfermería en pacientes hospitalizados.

Sin embargo, la aparición de la arquitectura *Transformer* ha revolucionado el panorama de las redes neuronales aplicadas a la generación de texto²⁰. La potencia de la arquitectura *Transformer* se debe a varias innovaciones que la distinguen de modelos anteriores. La primera es la creación de los «*word embedding*», donde cada palabra se representa con un vector al que se le suma otro que indica su posición en el texto. La segunda es que, al igual que los *autoencoders*, los *Transformers* tienen una parte codificadora y otra decodificadora, y en ambas se tiene en cuenta toda la secuencia al codificar, no solo las palabras

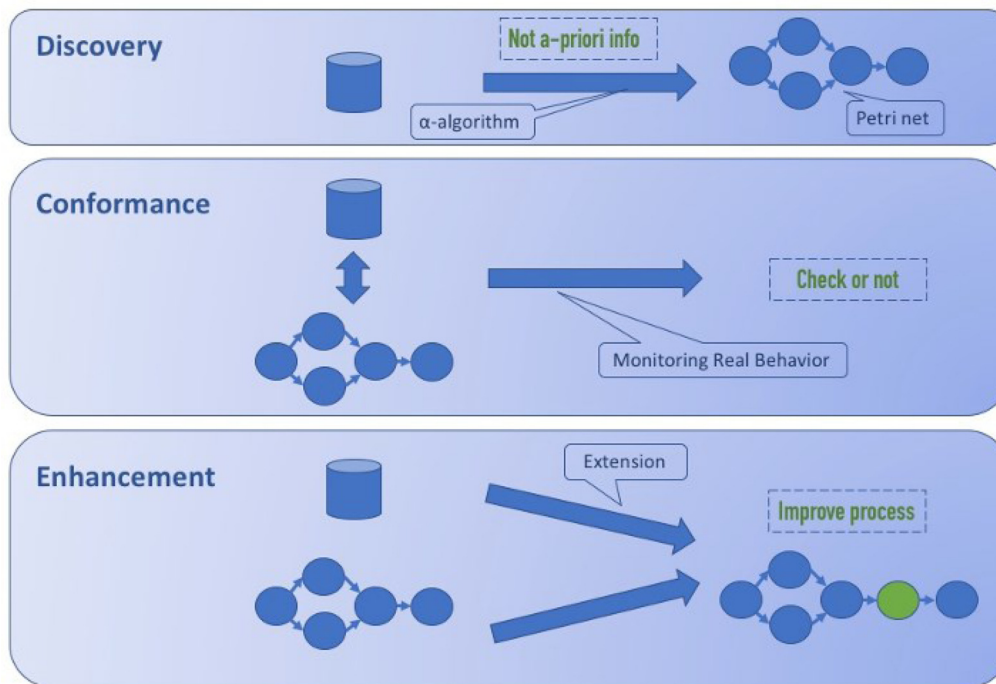


Figura 5. Tipos de minería de procesos.

cercanas; a esto se le llama auto-atención. La tercera, basada en la anterior, es que la auto-atención se realiza en paralelo múltiples veces, de modo que cada proceso atiende a diferentes aspectos del texto: uno puede detectar relaciones sujeto-verbo, otro sujeto-adjetivo, etc. Estas tres características permiten que los modelos *Transformer*, una vez entrenados, generen texto con calidad muy cercana al escrito por humanos. Un ejemplo destacado de esta arquitectura es el algoritmo ChatGPT.

Otros modelos basados en la arquitectura *Transformer*, como RoBERTa²¹ y ClinicalBERT²², han mostrado alto rendimiento en tareas de análisis de texto clínico, incluyendo aplicaciones relevantes en enfermería. RoBERTa puede ajustarse para tareas clínicas como la clasificación de notas de enfermería, la detección de información sensible (PHI) y el análisis semántico de registros clínicos. ClinicalBERT, por su parte, es una versión de BERT preentrenada específicamente con notas médicas del conjunto MIMIC-III, lo que lo hace especialmente adecuado para comprender la terminología clínica, abreviaturas y el lenguaje natural empleado en registros de salud. Es una herramienta útil para automatizar y mejorar el análisis documental en entornos asistenciales.

Minería de procesos: monitorización predictiva de procesos

La minería de procesos surge tanto por el aumento del volumen de datos generado por sistemas informáticos, como por la necesidad de analizar datos reales de los procesos para su mejora²³. Esta disciplina conecta la Ciencia de Datos con la Ciencia de Procesos, aplicando técnicas analíticas a registros de eventos en sistemas de información, especialmente en el ámbito hospitalario para optimizar los procesos asistenciales^{24,25}. Para comprenderla, es clave considerar dos conceptos fundamentales:

- Datos de registros de eventos (*event logs data* o simplemente *event logs*): registros de los datos de un proceso con la marca de tiempo de su ejecución, *timestamp*, si es el caso.
- Modelos de procesos: representación abstracta, que puede ser textual, pero que generalmente es gráfica, de un flujo real de acciones con un propósito concreto.

En el ámbito sanitario, cada interacción registrada, como una acción administrativa o una observación clínica, se integra en un registro de eventos (*event log*) con sus respectivas marcas de tiempo (*timestamps*). Este registro documenta la ejecución de un proceso, tal como el recorrido asistencial de un paciente desde su admisión hasta el alta. Por tanto, cada paciente genera una traza que refleja su trayectoria individual dentro del modelo general del proceso.

Se pueden distinguir tres tipos de minería de procesos (fig. 5): i) para descubrir cómo se está ejecutando un proceso en la realidad a partir de los registros de eventos (modelo de proceso AS-IS); ii) para comparar modelos de procesos obtenidos a partir de registros de eventos con modelos de procesos normativos (determinados por una auditoría, por ejemplo) o deseados (modelo de proceso TO-BE); y iii) para mejorar procesos, bien mediante la extensión o la reingeniería²⁶.

El algoritmo alfa, precursor en minería de procesos, permite descubrir automáticamente el flujo real de actividades a partir de registros de eventos²⁶. Sin conocer previamente el protocolo, identifica relaciones de causalidad, paralelismo o elección entre actividades clínicas (como toma de constantes, administración de medicación o diagnóstico) y genera un modelo gráfico del proceso tal como se desarrolla en la práctica.

Por ejemplo, si tras la valoración inicial del paciente los registros muestran posibles pasos como traslado al box u observación, el algoritmo alfa identifica estas relaciones y las incorpora al modelo. A partir del conjunto de trazas, se genera una representación completa del proceso con todas sus variantes reales. En sanidad, esta información permite comprender la organización de tareas, detectar rutas frecuentes, desviaciones o cuellos de botella, y así optimizar tiempos, recursos, calidad asistencial y apoyo a la toma de decisiones clínicas mediante el análisis de patrones atípicos.

Existen otros tipos de algoritmos que o bien complementan las deficiencias del anterior, como los de la familia alfa, o que tienen otro enfoque para la generación del modelo visual del proceso: *heuristic miner*²⁷, *fuzzy miner*²⁸, *inductive miner*²⁹, *genetic miner*³⁰, etc. Por ejemplo, Özdagöglu et al.³¹ muestran la eficacia del *fuzzy miner* ayudando a identificar cuellos de botella y patrones clave en los flujos de atención para mejorar la gestión de recursos hospitalarios.

Tabla 1
Resumen de técnicas y modelos mencionados en este trabajo

Datos de IMAGEN		
Técnicas / Modelos	Aplicaciones mencionadas	Contexto
Redes neuronales convolucionales	Reconocimiento facial; análisis de TC, RM y ecografías para detección de tumores Segmentación de RM para diagnóstico de Alzheimer	General / sanitario Diagnóstico por imagen
Datos de TEXTO		
Técnicas / Modelos	Aplicaciones mencionadas	Contexto
Redes neuronales convolucionales	Clasificación automática de textos clínicos	Documentación clínica
Redes neuronales recurrentes	Generación de texto; predicción de necesidades de enfermería	Lenguaje / hospitalario
Word embeddings	Similitud de palabras y apoyo a la clasificación de diagnósticos	Texto clínico
Transformers	Generación de texto	Lenguaje natural
RoBERTa	Clasificación de notas, detección de PHI, análisis semántico	Notas de enfermería
ClinicalBERT	Comprensión de terminología clínica y análisis documental	Registros médicos
Datos de PROCESOS Y REGISTROS		
Técnicas / Modelos	Aplicaciones mencionadas	Contexto
Aprendizaje supervisado (regresión, clasificación, reglas de asociación)	Incumplimiento de prevención de infecciones; riesgo de reingreso; asociaciones clínicas en registros administrativos	Domicilio / hospital
Aprendizaje no supervisado (<i>clustering</i> , PCA)	Clasificación de pacientes UCI; evaluación sistemas sanitarios	Cuidados críticos / sistemas
Minería de procesos (<i>fuzzy miner</i>)	Descubrimiento de flujos clínicos de atención; detección de cuellos de botella	Planificación hospitalaria
Monitoreo predictivo de procesos	Predicción exámenes adicionales en el proceso de realización de pruebas médicas	Planificación hospitalaria

La integración de la IA en la minería de procesos ha originado el monitoreo predictivo de procesos, rama centrada en anticipar la evolución de la ejecución de un proceso³², con especial relevancia en el ámbito sanitario. Las predicciones pueden referirse a resultados (booleanos o categóricos), métricas cuantitativas (tiempo, coste) o la secuencia de actividades futuras, lo que permite prevenir retrasos y resultados indeseados. El procedimiento consta de dos fases: (i) entrenamiento, se construye un modelo predictivo a partir de trazas de ejecución históricas completas, y (ii) ejecución, el modelo se consulta para predecir el futuro de un caso en curso. Di Francescomarino et al.³² ilustra el monitoreo predictivo en un hospital, aplicándolo al proceso de realización de pruebas médicas. A partir de datos históricos, se predice la necesidad de exámenes adicionales, optimizando así la planificación de recursos. Este enfoque facilita la toma de decisiones en tiempo real, la detección temprana de riesgos y la personalización de la asistencia sanitaria.

Las técnicas de monitoreo predictivo se clasifican según sus fundamentos en: enfoques basados en modelos y enfoques basados en aprendizaje automático³². Las primeras usan modelos explícitos, como sistemas de transición, y codifican las trazas mediante técnicas como codificación booleana, de frecuencia o basada en cargas útiles recientes. Las segundas construyen modelos implícitos a partir de datos históricos, y se dividen en: (1) predicción de resultados, mediante clasificación y, a veces, agrupamiento; (2) predicción de valores numéricos, basada en regresión para estimar métricas como tiempo o costo; y (3) predicción de eventos futuros, anticipando la secuencia y atributos de las próximas actividades.

Limitaciones actuales y perspectivas futuras

En la [tabla 1](#) se presentan los ejemplos de aplicaciones de IA mencionados a lo largo del texto a modo de resumen.

En general, las limitaciones a las que se enfrentan todas las técnicas de aprendizaje automático y monitoreo predictivo de procesos son: (1) posibles sesgos en los datos de entrenamiento (p. ej., si hay menos datos

por sexo o población)³³; (2) sobreajuste³⁴: el modelo se ajusta demasiado a los datos de entrenamiento y con datos nuevos su precisión cae drásticamente; (3) falta de robustez: pequeñas perturbaciones en los datos de entrada provocan cambios drásticos de la salida³⁵. Todas plantean retos para la investigación en IA y su posible aplicación en la práctica clínica.

Entre las principales debilidades en el ámbito sanitario se encuentra la dificultad de interpretación clínica de los modelos, especialmente cuando carecen de transparencia. Precisamente, los métodos basados en redes neuronales no son explicables, en general, ya que es difícil tanto interpretar sus resultados como su funcionamiento interno. Por ejemplo, mecanismos de poda de parámetros y nodos pueden conducir de manera contraintuitiva a producir mejores resultados. Además, es necesario tener en cuenta consideraciones éticas y normativas de protección de datos.

Sin embargo, la popularización de los LLM, y el esfuerzo llevado a cabo por empresas e investigadores, ha mejorado sustancialmente las capacidades de interpretación de la IA, por ejemplo, incluyendo citas en el texto generado por los LLMs donde se puede comprobar la veracidad de las afirmaciones realizadas por estos modelos. La previsión es que estos modelos en particular y la IA en general se acerquen cada vez más al objetivo de lo que se ha llamado IA explicable.

Los modelos de aprendizaje automático tradicionales siguen siendo muy usados en el ámbito sanitario, sobre todo cuando se trabaja con conjuntos de datos pequeños y bien estructurados. Sin embargo, el aumento de datos no estructurados y heterogéneos ha hecho cada vez más necesario recurrir a redes neuronales y técnicas de aprendizaje profundo.

Cabe destacar que las redes neuronales profundas y los LLM requieren un gran volumen de datos para su entrenamiento, así como demandan altas capacidades de cómputo, que se traducen en ingentes cantidades de energía para entrenar estos modelos. En el futuro, se espera poder entrenarlos con menos datos y menor consumo de tiempo y recursos.

En cuanto al monitoreo predictivo de procesos hay que subrayar la baja calidad y heterogeneidad de los datos clínicos disponibles en ocasiones, lo cual se traduce en registros de eventos incompletos o no estructurados, dificultando así la construcción de modelos predictivos fiables. También representa un reto su integración en los flujos asistenciales reales.

Asimismo, la existencia de una marca temporal en cada uno de los registros es de gran ayuda en la construcción de modelos de monitoreo predictivo de procesos. La tecnología puede ayudar a insertar estas marcas temporales, por ejemplo, en el registro mediante dispositivos electrónicos que inserten estas marcas automáticamente, o el uso de recordatorios para que los añada el personal de enfermería.

Financiación

Este trabajo está parcialmente financiado por la Universitat Jaume I a través de la red temática multidisciplinar de colaboración en investigación e innovación *Nursing and healthcare system innovation through artificial Intelligence Applications* (NursIA). La entidad financiadora no ha intervenido en el diseño, desarrollo ni redacción del manuscrito.

Consideraciones éticas

El presente estudio se basa en el análisis de literatura científica publicada. No ha implicado la participación de personas ni el uso de datos clínicos o personales, por lo que no ha requerido la evaluación ni la aprobación de un Comité de Ética de la Investigación.

Consentimiento informado

Por la naturaleza del estudio no se incluye la participación de sujetos ni el uso de datos individuales, no ha sido necesario obtener consentimiento informado.

Protección de datos

Este trabajo no ha utilizado datos personales ni información identificable de pacientes.

Conflicto de intereses

Los autores declaran no tener ningún conflicto de intereses.

Uso de inteligencia artificial

Los autores declaran que se han utilizado herramientas de inteligencia artificial generativa únicamente como apoyo en tareas de mejora lingüística. Los autores han revisado y validado íntegramente el contenido final y asumen plena responsabilidad sobre el mismo.

Autoría

Todos los autores cumplen los criterios internacionales de autoría científica. Han contribuido de forma sustancial a la concepción del trabajo, análisis de la literatura, redacción y revisión crítica del manuscrito, y han aprobado la versión final para su publicación.

IA generativa

La Inteligencia Artificial se está incorporando al ámbito de enfermería, pero existe dispersión conceptual y confusión entre técnicas, especialmente más allá de la IA generativa.

Ofrece una síntesis estructurada de fundamentos y técnicas de Inteligencia Artificial, desde el *Machine Learning* hasta técnicas de monitorización predictiva de procesos, aplicables al proceso asistencial en

enfermería, con ejemplos, énfasis en explicabilidad y limitaciones para facilitar su comprensión.

Agradecimientos

Este trabajo está parcialmente financiado por la Universitat Jaume I a través de la red temática multidisciplinar de colaboración en investigación e innovación *NursIA – Nursing and healthcare system innovation through artificial Intelligence Applications*.

Bibliografía

- Samuel AL. Some studies in machine learning using the game of checkers. *IBM J Res Dev*. 1959;3:210–229, <http://dx.doi.org/10.1147/rd.33.0210>.
- Adams V, Song J, Shang J, et al. Infection prevention and control practices in the home environment: examining enablers and barriers to adherence among home health care nurses. *Am J Infect Control*. 2021;49:721–726, <http://dx.doi.org/10.1016/j.ajic.2020.10.021>.
- Brom H, Carthon JMB, Ikeaba U, Chittams J. Leveraging electronic health records and machine learning to tailor nursing care for patients at high risk for readmissions. *J Nurs Care Qual*. 2020;35:27–33, <http://dx.doi.org/10.1097/NCQ.0000000000000412>.
- Peng M, Lee S, D'Souza AG, Doktorchik CTA, Quan H. Development and validation of data quality rules in administrative health data using association rule mining. *BMC Med Inform Decis Mak*. 2020;20:75, <http://dx.doi.org/10.1186/s12911-020-1089-0>.
- An R, Chang GM, Fan YY, Ji LL, Wang XH, Hong S. Machine learning-based patient classification system for adult patients in intensive care units: a cross-sectional study. *J Nurs Manag*. 2021;29:1752–1762, <http://dx.doi.org/10.1111/jonm.13284>.
- Wanzala MN, Oloo JA, Nguka G, Were V. Application of principal component analysis to assess health systems capacity using cross sectional data in rural western Kenya. *Am J Public Health Res*. 2019;7:27–32. Disponible en: <http://pubs.sciepub.com/ajphr/7/1/5>.
- McCulloch WS, Pitts W. A logical calculus of the ideas immanent in nervous activity. *Bull Math Biophys*. 1943;5:115–133, <http://dx.doi.org/10.1007/BF02478259>.
- Rosenblatt F. The perceptron: a probabilistic model for information storage and organization in the brain. *Psychol Rev*. 1958;65:386–408, <http://dx.doi.org/10.1037/h0042519>.
- Rumelhart DE, Hinton GE, Williams RJ. Learning representations by back-propagating errors. *Nature*. 1986;323:533–536, <http://dx.doi.org/10.1038/323533a0>.
- LeCun Y, Boser B, Denker JS, et al. Backpropagation applied to handwritten zip code recognition. *Neural Comput*. 1989;1:541–551, <http://dx.doi.org/10.1162/neco.1989.1.4.541>.
- Chen X, Li L, Sharma A, Dhiman G, Vimal S. The application of convolutional neural network model in diagnosis and nursing of MR imaging in Alzheimer's disease. *Interdiscip Sci*. 2022;14:34–44, <http://dx.doi.org/10.1007/s12539-021-00450-7>.
- Kramer MA. Nonlinear principal component analysis using autoassociative neural networks. *AIChE J*. 1991;37:233–243, <http://dx.doi.org/10.1002/aic.690370209>.
- Krizhevsky A, Sutskever I, Hinton GE. ImageNet classification with deep convolutional neural networks. *Commun ACM*. 2017;60:84–90, <http://dx.doi.org/10.1145/3065386>.
- Vaswani A, Shazeer N, Parmar N, et al. Attention is all you need. *NeurIPS 2024*. 2017;1:1–15, <http://dx.doi.org/10.48550/arXiv.03762>.
- Mikolov T, Chen K, Corrado G, Dean J. Efficient estimation of word representations in vector space. En: *Proceedings of the Workshop at ICLR*. Scottsdale: ICLR; 2013.
- Mikolov T, Sutskever I, Chen K, Corrado G, Dean J. Distributed representations of words and phrases and their compositionality. *NeurIPS*. 2013.
- Hughes M, Li I, Kotoulas S, Suzumura T. Medical text classification using convolutional neural networks. *Stud Health Technol Inform*. 2017;235:246–250.
- Hochreiter S, Schmidhuber J. Long short-term memory. *Neural Comput*. 1997;9:1735–1780, <http://dx.doi.org/10.1162/neco.1997.9.8.1735>.
- Baek Y, Han K, Jeon E, Yoo HY. Prediction of nursing need proxies using vital signs and biomarkers data: application of deep learning models. *J Clin Nurs*. 2024;34:4121–4131, <http://dx.doi.org/10.1111/jocn.17612>.
- Cho HN, Jun TJ, Kim YH, et al. Task-specific transformer-based language models in health care: scoping review. *JMIR Med Inform*. 2024;12:e49724, <http://dx.doi.org/10.2196/49724>.
- Laios A, Kalampokis E, Mamalis ME, et al. RoBERTa-assisted outcome prediction in ovarian cancer cytoreductive surgery using operative notes. *Cancer Control*. 2023;30:10732748231209892, <http://dx.doi.org/10.1177/10732748231209892>.
- Huang K, Altaoar J, Ranganath R. ClinicalBERT: modeling clinical notes and predicting hospital readmission. *arXiv*. 2020. Disponible en: 10.48550/arXiv.1904.05342.
- van der Aalst WMP. Data scientist: the engineer of the future. En: Mertins K, Bénabou F, Poler R, Bourrières JP, eds. *Enterprise Interoperability VI*. Cham: Springer; 2014:13–26, http://dx.doi.org/10.1007/978-3-319-04948-9_2.
- Aversano L, Iammarino M, Madau A, Pirlo G, Semeraro G. Process mining applications in healthcare: a systematic literature review. *PeerJ Comput Sci*. 2025;11:e2613, <http://dx.doi.org/10.7717/peerj-cs.2613>.
- van der Aalst WMP. *Process mining: data science in action*. 2 ed. Heidelberg: Springer; 2016, <http://dx.doi.org/10.1007/978-3-662-49851-4>.

26. van der Aalst WMP, Weijters AJMM, Maruster L. Workflow mining: discovering process models from event logs. *IEEE Trans Knowl Data Eng.* 2004;16:1128–1142, <http://dx.doi.org/10.1109/TKDE.2004.47>.
27. Weijters AJMM, Medeiros AKA, van der Aalst WMP. *Process mining with the heuristics miner algorithm*. Eindhoven: TU Eindhoven; 2006.
28. Günther CW, van der Aalst WMP. Fuzzy mining – adaptive process simplification based on multi-perspective metrics. En: Alonso G, Dadam P, Rosemann M, eds. *Business Process Management*. Berlin: Springer; 2007:328–343, http://dx.doi.org/10.1007/978-3-540-75183-0_24.
29. Leemans SJJ, Fahland D, van der Aalst WMP. Discovering block-structured process models from event logs - A constructive approach. En: Colom JM, Desel J, eds. *Application and Theory of Petri Nets and Concurrency*. Berlin: Springer; 2013:311–329, http://dx.doi.org/10.1007/978-3-642-38697-8_17.
30. de Medeiros AKA, Weijters AJMM, van der Aalst WMP. Genetic process mining: an experimental evaluation. *Data Min Knowl Discov.* 2007;14:245–304, <http://dx.doi.org/10.1007/s10618-006-0061-7>.
31. Özdağoğlu G, Damar M, Erenay FS, Damar HT, Himmetoğlu O, Pinto AD. Monitoring patient pathways through process mining via fuzzy miner. *BMC Med Inform Decis Mak.* 2025;25:199, <http://dx.doi.org/10.1186/s12911-025-03016-5>.
32. Di Francescomarino C, Ghidini C. Predictive process monitoring. En: van der Aalst WMP, Carmona J, eds. *Process Mining Handbook*. Cham: Springer; 2022, http://dx.doi.org/10.1007/978-3-031-08848-3_10.
33. O'Connor S, Booth RG. Algorithmic bias in health care: opportunities for nurses to improve equality in the age of artificial intelligence. *Nurs Outlook.* 2022;70:780–782, <http://dx.doi.org/10.1016/j.outlook.2022.08.004>.
34. Gygi JP, Kleinstein SH, Guan L. Predictive overfitting in immunological applications: pitfalls and solutions. *Hum Vaccin Immunother.* 2023;19:2251830, <http://dx.doi.org/10.1080/21645515.2023.2251830>.
35. Ghaffari Laleh N, Chen RJ, Diao JA, et al. Adversarial attacks and adversarial robustness in computational pathology. *Nat Commun.* 2022;13:5711, <http://dx.doi.org/10.1038/s41467-022-33266-0>.