



1.1 ¿Qué es la inteligencia artificial?

1.1.1 Definición de IA

La Inteligencia Artificial (en adelante, IA) se refiere a un campo de estudio y desarrollo tecnológico enfocado en la creación de sistemas y máquinas capaces de realizar tareas que, cuando son ejecutadas por humanos, requieren inteligencia. Estas tareas incluyen el razonamiento, el aprendizaje, la creatividad y la planificación. Las máquinas equipadas con IA pueden percibir su entorno, interactuar con él, resolver problemas y tomar decisiones de manera autónoma con el fin de alcanzar los objetivos específicos que se les propone. Estos sistemas funcionan recibiendo datos de su entorno físico o digital, procesándolos mediante algoritmos avanzados y actuando en consecuencia para llevar a cabo tareas o responder a situaciones de manera eficiente (Kaplan y Haenlein, 2019).

Aunque la tecnología de IA no es nueva, ha experimentado una evolución significativa en los últimos años. Este avance ha estado impulsado por los nuevos desarrollos en procesamiento de datos, aprendizaje automático y en las capacidades computacionales de los ordenadores actuales (Russell y Norvig, 2021). Estamos presenciando cómo esta evolución transforma numerosas áreas de nuestras vidas y como abre nuevas posibilidades en múltiples campos como en la industria, la salud, la educación, el transporte y las finanzas (Goodfellow et al., 2016; Haenlein y Kaplan, 2019).

1.1.2. Historia y evolución de la IA

La historia de la inteligencia artificial es un camino de creatividad e ingenio tecnológico para alcanzar un deseo que ha perseguido al ser humano a lo largo de la historia: crear máquinas que amplifiquen y superen nuestras limitaciones. La IA busca crear esas máquinas a partir de la replicación de nuestras propias habilidades cognitivas. Este camino comenzó y ha discurrido paralelo al desarrollo de la computación (Russell y Norvig, 2021).

La idea de máquinas pensantes se remonta a la antigüedad, pero el concepto de inteligencia artificial tal como lo conocemos hoy comenzó a tomar forma en el siglo XX. Un hito teórico importante fue la propuesta realizada por el matemático británico Alan Turing, quien en 1950 publicó un artículo seminal en el que



planteaba la cuestión de si las máquinas podrían pensar. En este trabajo, Turing desarrolló un método para comprobarlo, conocido como el "Test de Turing". Este test pretende determinar si una máquina puede exhibir un comportamiento inteligente indistinguible del de un ser humano. Turing abrió la puerta a la posibilidad teórica de las máquinas "pensantes" (Turing, 1950).

En 1956, el término "inteligencia artificial" fue acuñado oficialmente en la Conferencia de Dartmouth, un encuentro organizado por John McCarthy, Marvin Minsky, Nathaniel Rochester y Claude Shannon. Esta conferencia es considerada el nacimiento oficial de la IA como un campo de estudio académico. Los investigadores de esa época eran optimistas y creían que la creación de una máquina inteligente era inminente (McCarthy et al., 2006).

Durante las décadas de 1960 y 1970, la investigación en IA recibió un apoyo considerable, tanto del sector privado como de agencias gubernamentales. En estos años, se desarrollaron los primeros programas de IA que podían resolver problemas de álgebra, demostrar teoremas geométricos y aprender a jugar ajedrez (Russell y Norvig, 2021). El proyecto de ajedrez culminó con el desarrollo de "Deep Blue" de IBM, un sistema que eventualmente derrotó al campeón mundial Gary Kasparov en 1997 (Campbell et al., 2002).

Sin embargo, a fines de la década de 1970 y durante los años 80, la IA pasó por períodos de derrotismo y reducción de financiación, conocidos como los "inviernos de la IA". Estos períodos fueron resultado de las altas expectativas no cumplidas, donde los procesos de procesamiento del lenguaje natural, la comprensión de las imágenes y el aprendizaje automático resultaron ser más complejos de lo esperado (Crevier, 1993).

El interés y la inversión en IA resurgieron en los años 90 gracias a mejoras en los algoritmos y un aumento exponencial en la capacidad de procesamiento de las computadoras (Nilsson, 2010). El aprendizaje automático, y especialmente el aprendizaje profundo, empezó a dominar la investigación, permitiendo avances significativos en procesamiento de imágenes y reconocimiento de voz (LeCun et al., 2015).

Desde 2010, la IA ha experimentado un crecimiento espectacular y ha comenzado a introducirse en la industria, la medicina, el transporte, y en los ámbitos de nuestra vida cotidiana. Los asistentes virtuales como Siri, Alexa y Google Assistant



se han hecho comunes, y ya nadie prescinde de Google Maps en sus desplazamientos. Hoy en día, la IA no solo optimiza procesos industriales y servicios, sino que también ayuda en la toma de decisiones médicas, mejora las experiencias de los consumidores y transforma las prácticas en agricultura y conservación ambiental. Sin embargo, estos avances impresionantes también plantean cuestiones éticas y sociales, como la privacidad, la seguridad y el impacto en el empleo (Gunning y Aha, 2019).

1.1.3. Tipos de IA

La IA abarca una amplia gama de sistemas y tecnologías que varían en complejidad y capacidad. Para entender mejor su alcance teórico y aplicado, vamos a explorar y definir distintas categorías. Una primera clasificación en función de la capacidad sería: IA débil o estrecha, IA fuerte o general, y super inteligencia.

- **IA débil (IA estrecha):** La IA débil, también conocida como IA estrecha, se refiere a sistemas diseñados y entrenados para realizar tareas específicas sin poseer habilidades generales de resolución de problemas. Estas inteligencias artificiales operan dentro de un ámbito limitado y no pueden realizar actividades fuera de su dominio especializado. Ejemplos de esta categoría serían los asistentes virtuales como Siri y Alexa, así como sistemas de recomendación de plataformas como Netflix y Amazon (Nilsson, 2010; Kaplan y Haenlein, 2019).
- **IA fuerte (IA general):** La IA fuerte, o IA general, se refiere a sistemas que poseen capacidad cognitiva general similar a la humana, lo que les permite entender, aprender y aplicar conocimientos de manera amplia. A diferencia de la IA débil, la IA fuerte es capaz de realizar cualquier tarea intelectual que un ser humano pueda llevar a cabo. Actualmente, este es un concepto teórico sin implementación práctica, pero que para muchos es un objetivo alcanzable (Kurzweil, 2005; Goertzel, 2020).
- **IA súper inteligente:** La IA súper inteligente se refiere a un nivel de inteligencia que supera ampliamente la inteligencia humana en todos los aspectos, incluyendo la creatividad, la resolución de problemas y la toma de decisiones. En la actualidad, este tipo de IA es totalmente teórico, no existe



ninguna implementación práctica, y su viabilidad es muy discutida (Bostrom, 2014; Yampolskiy, 2020).

Otros tipos de IA según su capacidad de aprender (Inteligente y Artificial, s.f.) :

- **IA reactiva:** La IA reactiva se refiere a sistemas que pueden responder a diferentes tipos de estímulos sin tener la capacidad de usar experiencias pasadas para informar decisiones actuales. Estas inteligencias artificiales no tienen memoria y solo pueden actuar basándose en la información presente. Un ejemplo es Deep Blue de IBM, que jugaba ajedrez (Russell y Norvig, 2021; Tesauro, 2019).
- **IA con memoria limitada:** La IA con memoria limitada puede utilizar experiencias pasadas para informar decisiones futuras, aunque esta información es transitoria. Este tipo de IA almacena datos temporales para mejorar sus procesos de toma de decisiones. Ejemplos incluyen coches autónomos que observan el entorno y ajustan sus decisiones en consecuencia (Goodfellow et al., 2016; LeCun et al., 2015).
- **Teoría de la mente:** La teoría de la mente en IA se refiere a sistemas que pueden comprender y predecir las emociones, creencias y pensamientos de otras entidades. Estos sistemas son capaces de interactuar con humanos de manera más intuitiva al anticipar y responder a sus estados emocionales y cognitivos. Aunque todavía está en investigación y desarrollo, esta IA tendrá grandes aplicaciones en la interacción humano-robot (Stone et al., 2016; Schuller y Batliner, 2013).
- **Autoconciencia:** La autoconciencia en IA se refiere a sistemas que poseen una forma de autoconciencia, entendiendo su propio estado y siendo capaces de anticipar sus propios “sentimientos”. Este tipo de IA representa el nivel más avanzado de desarrollo, donde las máquinas podrían tener una conciencia similar a la humana. Actualmente es un concepto teórico y no existen tales sistemas (Tegmark, 2017; Eden et al., 2013).

1.1.4. La IA generativa, fundamentos y aplicaciones

La inteligencia artificial generativa es una subcategoría de la IA que se centra en la creación de contenido nuevo y original, ya sea texto, imágenes, música, voz, u otras formas de expresión digital. A diferencia de otros tipos de IA que se limitan a



la automatización de tareas, la IA generativa utiliza algoritmos avanzados, como redes neuronales profundas y modelos de lenguaje, para aprender de una gran cantidad de datos y luego producir salidas que no son simplemente respuestas directas o decisiones, sino nuevas creaciones que nunca antes existieron. Estas herramientas pueden generar arte, redactar textos, componer música, o incluso simular conversaciones humanas de manera coherente y creativa, imitando así aspectos de la cognición humana relacionados con la creatividad y la innovación (Goodfellow et al., 2016; Radford et al., 2019).

A. Fundamentos de la IA Generativa

Los fundamentos de la IA generativa se basan en modelos y algoritmos que permiten a las máquinas aprender patrones, estilos, y estructuras a partir de datos existentes. Algunas de las técnicas que han dado mejores resultados son:

- **Redes neuronales profundas (Deep Learning):** Las redes neuronales profundas son un tipo de aprendizaje automático que imita la manera en que los humanos adquieren ciertos tipos de conocimientos. Estas redes son capaces de detectar y replicar patrones complejos y son fundamentales para muchas aplicaciones de IA generativa (LeCun et al., 2015).
- **Redes generativas antagónicas (en inglés, Generative Adversarial Networks o GAN):** Las GANs son un tipo de algoritmos de IA utilizados en el aprendizaje no supervisado. Funcionan poniendo dos redes neuronales, una generativa y una discriminativa, una contra la otra. Esto ayuda a mejorar la calidad y la precisión de los resultados generativos (Goodfellow et al., 2014).
- **Autoencoders variacionales (VAEs):** Los VAEs son otro tipo de modelo generativo que ayuda a reconstruir entradas (como imágenes) basándose en el aprendizaje de sus aspectos fundamentales y variaciones. Son especialmente útiles para tareas que requieren una alta fidelidad en la reconstrucción (Kingma y Welling, 2014).



B. Aplicaciones de la IA Generativa

La IA generativa ha encontrado aplicaciones en numerosos campos y continúa sorprendiéndonos con nuevas funcionalidades y capacidades:

- **Arte y creatividad:** En el campo del arte, la IA generativa puede crear obras visuales, música y literatura. Por ejemplo, plataformas como DALL-E y DeepArt utilizan GANs para generar nuevas imágenes artísticas basadas en descripciones textuales o estilos artísticos previos (Ramesh et al., 2021).
- **Diseño:** La IA generativa puede acelerar el proceso de diseño al generar automáticamente prototipos y modelos basados en especificaciones dadas. Esto es especialmente prevalente en la arquitectura y la moda, donde la rapidez y la innovación son cruciales (Park et al., 2019).
- **Entretenimiento y medios:** En la industria del entretenimiento, la IA generativa está transformando la producción de contenido, desde la creación de nuevas piezas musicales hasta guiones para películas o videojuegos, ofreciendo narrativas y diálogos en tiempo real adaptados a las preferencias del usuario (Huang et al., 2020).
- **Educación:** La IA generativa ha demostrado su capacidad para personalizar materiales de aprendizaje y tutorías adaptadas a las necesidades individuales de los estudiantes, ayudando a explicar conceptos complejos o creando ejemplos nuevos y relevantes para facilitar la comprensión (Chen et al., 2020).
- **Investigación y ciencia de datos:** En la ciencia de datos, los modelos generativos se utilizan para crear grandes conjuntos de datos sintéticos que simulan datos reales, permitiendo a los investigadores realizar experimentos y análisis sin comprometer la privacidad o la seguridad de los datos reales (Yoon et al., 2019).
- **Publicidad y marketing:** En marketing, la IA generativa puede producir contenido creativo, como anuncios, imágenes y textos personalizados que resuenen con audiencias específicas, mejorando la eficiencia de las campañas publicitarias (Gao et al., 2018).



Referencias

- Bostrom, N. (2014). *Superintelligence: Paths, Dangers, Strategies*. Oxford University Press.
- Campbell, M., Hoane, A. J. & Hsu, F. (2002). Deep Blue. *Artificial Intelligence*, 134(1-2), 57-83. [https://doi.org/10.1016/S0004-3702\(01\)00129-1](https://doi.org/10.1016/S0004-3702(01)00129-1)
- Chen, C., Li, O., Tao, D., Barnett, A., Rudin, C. & Su, J. K. (2020). This Looks Like That: Deep Learning for Interpretable Image Recognition. *Proceedings of the 33rd International Conference on Neural Information Processing Systems (NeurIPS 2019)*. <https://proceedings.neurips.cc/paper/2019/hash/adf7ee2dcf142b0e11888e72b43fcb75-Abstract.html>
- Crevier, D. (1993). *AI: The Tumultuous History of the Search for Artificial Intelligence*. Basic Books.
- Eden, A. H., Moor, J. H., Soraker, J. H. & Steinhart, E. (Eds.). (2013). *Singularity Hypotheses: A Scientific and Philosophical Assessment*. Springer.
- Gao, Y., Li, J. & Yin, H. (2018). Automated design of billboard advertisements using generative adversarial networks. *Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining (KDD '18)*.
- Goertzel, B. (2020). Artificial general intelligence: Concept, state of the art, and future prospects. *Journal of Artificial General Intelligence*, 11(1), 1-48. <https://doi.org/10.2478/jagi-2014-0001>
- Goodfellow, I., Bengio, Y. & Courville, A. (2016). *Deep Learning*. MIT Press.
- Goodfellow, I., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., ... & Bengio, Y. (2014). Generative adversarial nets. *Advances in Neural Information Processing Systems (NeurIPS 2014)*. https://proceedings.neurips.cc/paper_files/paper/2014/hash/5ca3e9b122f61f8f06494c97b1afccf3-Abstract.html
- Gunning, D. & Aha, D. W. (2019). DARPA's Explainable Artificial Intelligence (XAI) Program. *AI Magazine*, 40(2), 44-58. <https://doi.org/10.1609/aimag.v40i2.2850>
- Haenlein, M. & Kaplan, A. (2019). A brief history of artificial intelligence: On the past, present, and future of artificial intelligence. *California Management Review*, 61(4), 5-14. <https://doi.org/10.1177/0008125619864925>



- Huang, A., Cahuantzi, R. & Pfenning, F. (2020). Real-time interactive text-based game generation with recurrent networks. *Proceedings of the 19th International Conference on Autonomous Agents and MultiAgent Systems (AAMAS '20)*.
- Inteligente y Artificial. (n.d.). Tipos de inteligencia artificial. Inteligente y Artificial. Retrieved July 11, 2024, from <https://inteligenteyartificial.com/tipos-de-inteligencia-artificial/#:~:text=%2D%20M%C3%A1quinas%20reactivas,-Es%20un%20tipo&text=Un%20ejemplo%20t%C3%ADpico%20de%20este,mediados%20de%20los%20a%C3%Blos%20ochenta>
- Kaplan, A. & Haenlein, M. (2019). Siri, Siri, in my hand: Who's the fairest in the land? On the interpretations, illustrations, and implications of artificial intelligence. *Business Horizons*, 62(1), 15-25. <https://doi.org/10.1016/j.bushor.2018.08.004>
- Kingma, D. P. & Welling, M. (2014). Auto-encoding variational bayes. *Proceedings of the International Conference on Learning Representations (ICLR 2014)*. <https://doi.org/10.48550/arXiv.1312.6114>
- Kurzweil, R. (2005). *The Singularity is Near: When Humans Transcend Biology*. Viking Press.
- LeCun, Y., Bengio, Y. & Hinton, G. (2015). Deep learning. *Nature*, 521(7553), 436-444. <https://doi.org/10.1038/nature14539>
- McCarthy, J., Minsky, M., Rochester, N. & Shannon, C. E. (2006). A proposal for the Dartmouth Summer Research Project on Artificial Intelligence, August 31, 1955. *AI Magazine*, 27(4), 12-14. <https://doi.org/10.1609/aimag.v27i4.1904>
- Nilsson, N. J. (2010). *The Quest for Artificial Intelligence: A History of Ideas and Achievements*. Cambridge University Press.
- Park, J., Woo, S., Choi, I. & Kweon, I. S. (2019). BAM: Bottleneck Attention Module. *Proceedings of the British Machine Vision Conference (BMVC 2019)*. <https://doi.org/10.48550/arXiv.1807.06514>
- Radford, A., Wu, J., Child, R., Luan, D., Amodei, D. & Sutskever, I. (2019). Language models are unsupervised multitask learners. *OpenAI Blog*, 1(8), 9. <https://insightcivic.s3.us-east-1.amazonaws.com/language-models.pdf>
- Ramesh, A., Pavlov, M., Goh, G., Gray, S., Voss, C., Radford, A., ... & Sutskever, I. (2021). Zero-shot text-to-image generation. *Proceedings of the 38th*



- International Conference on Machine Learning* (ICML 2021).
<https://proceedings.mlr.press/v139/ramesh21a.html>
- Russell, S. & Norvig, P. (2021). *Artificial Intelligence: A Modern Approach* (4th ed.). Pearson. <https://thuvienso.hoasen.edu.vn/handle/123456789/8967>
- Schuller, B. & Batliner, A. (2013). *Computational Paralinguistics: Emotion, Affect, and Personality in Speech and Language Processing*. Wiley-IEEE Press.
- Stone, P., et al. (2016). *Artificial Intelligence and Life in 2030. One Hundred Year Study on Artificial Intelligence: Report of the 2015-2016 Study Panel*. Stanford University. <https://doi.org/10.48550/arXiv.2211.06318>
- Tegmark, M. (2017). *Life 3.0: Being Human in the Age of Artificial Intelligence*. Knopf.
- Tesauro, G. (2019). Temporal difference learning and TD-Gammon. *Communications of the ACM*, 38(3), 58-68. <https://www.csd.uwo.ca/~xling/cs346a/extra/tdgammon.pdf>
- Turing, A. M. (1950). Computing machinery and intelligence. *Mind*, 59(236), 433-460. https://doi.org/10.1007/978-1-4020-6710-5_3
- Yampolskiy, R. V. (2020). Unpredictability of AI: On the impossibility of accurately predicting all actions of a smarter agent. *Journal of Experimental & Theoretical Artificial Intelligence*, 32(1), 1-12. <https://doi.org/10.1142/S2705078520500034>
- Yoon, J., Jarrett, D. & van der Schaar, M. (2019). Time-series generative adversarial networks. *Advances in Neural Information Processing Systems* (NeurIPS 2019). <https://proceedings.neurips.cc/paper/2019/hash/c9efe5f26cd17ba6216bbe2a7d26d490-Abstract.html?ref=https://githubhelp.com>