



Escuela técnica superior de ingeniería (ICAI)

AVANCES Y DESAFÍOS DE LA INTELIGENCIA ARTIFICIAL

Autor: Álvaro Serrahima de Bedoya

Director: Josefina Bengoechea Fernández

MADRID | Marzo 2022

ÍNDICE

1. Introducción	7
Objetivos	8
Estructura y Metodología	8
Marco Teórico	9
2. Historia de la IA	11
Inicio.....	11
Deep learning	12
Primer Invierno de la IA.....	13
Boom de 1980	13
Segundo Invierno de la IA.....	14
Resurgir de la IA	14
3. Introducción a la inteligencia artificial: qué es y cómo funciona.	17
4. Estado actual de la ciencia	20
4.1. Principales implicaciones.....	20
Factores determinantes.....	20
Ventajas del machine learning	22
Cambios producidos por el ML	23
4.2. Aplicaciones.....	24
Percepción.....	24
Cognición.....	29
4.3. Principales desafíos y limitaciones	33
4.4. Normativa vigente y dilemas éticos	42
4.4.1. Desafíos regulatorios	42
4.4.2. Estado actual de la legislación.....	45
5. Futuro de la IA:.....	49
5.1. ¿Qué podemos esperar de la IA en el futuro?	49
5.1.1. Posibles avances tecnológicos	49
5.1.2. Futuras aplicaciones	52
5.2. Riesgos	60
5.2.1. Riesgos a corto plazo.....	60

5.2.2. Riesgos a largo plazo.....	64
6. Conclusiones.....	67
Bibliografía.....	69

Resumen:

El campo de la IA crece a pasos agigantados y cada día surgen nuevos y mejores avances en materia. Las mejoras en tecnología, los avances en computación y la mayor disponibilidad de datos han impulsado la IA a nuevos niveles. En este estudio, se pretende dar una visión exhaustiva al estado actual de la ciencia y a los futuros desafíos que presenta. Para ello, el alcance de este trabajo incluye el pasado de la tecnología, los motores de su desarrollo, las aplicaciones, así como las principales limitaciones tanto tecnológicas como socioeconómicas a las que se enfrenta. Adicionalmente, se explora su futuro, incluyendo posibles nuevas aplicaciones y potenciales riesgos. La principal contribución de este trabajo es reunir el estado actual del arte como ayuda para futuros estudios.

Palabras clave:

Inteligencia artificial, deep learning, machine learning

Abstract:

The field of AI is growing at a fast rate and new and better advances in the field emerge every day. Improvements in technology, advances in computer science, and the increased availability of data have pushed AI to new levels. This study aims to give an exhaustive vision of the current state of science and the future challenges it presents. To this end, the scope of this work includes the past of the technology, the drivers of its development, the applications, as well as the main technological and socioeconomic limitations it faces. Additionally, it explores the future of IA, including possible new applications and potential risks. The main contribution of this work is to gather the current state of the art as an aid for future studies.

Key words:

Artificial Intelligence, Deep Learning, Machine Learning

1. INTRODUCCIÓN

Ha existido, a lo largo de la historia de la humanidad, una cierta fascinación del ser humano por la inteligencia. Es para nosotros un concepto primordial, y para muchos la esencia de la propia vida. Cuando miramos al espacio, nos preguntamos si existirá más allá de nuestro planeta algún signo de vida inteligente. Apreciamos nuestra inteligencia, la estudiamos, la medimos, y desde hace menos de un siglo, intentamos replicarla.

Cuando comenzaron las primeras máquinas, se creyó haber encontrado la misma esencia de la inteligencia. Se vio, en la lógica de la programación, la oportunidad de construir máquinas capaces de mostrar inteligencia, y a su vez, surgió la creciente visión del cerebro como un ordenador muy avanzado. Surgió así lo que se conoce como inteligencia artificial. Décadas más tarde, tras numerosos avances tanto en computación como en el campo de la IA, la visión idealizada y simplificada de lo que significa replicar la inteligencia humana ha dejado paso a otra más realista, consciente de la verdadera complejidad que supone alcanzar tal reto.

Es complicado asignar una definición exacta al concepto de inteligencia artificial, ya que para ello en primer lugar sería necesario tener una definición clara de qué es la inteligencia. Cuando una persona adulta realiza una tarea simple, como hacer una suma fácil, no se considera que sea muy inteligente. Si lo hace, en cambio, un niño de un año, o un chimpancé, se considera como una muestra de inteligencia avanzada. Por tanto, el concepto de inteligencia es relativo, y de igual forma ocurre con la definición de IA. Según la RAE, se define IA como “Disciplina científica que se ocupa de crear programas informáticos que ejecutan operaciones comparables a las que realiza la mente humana, como el aprendizaje o el razonamiento lógico”. Sin embargo, esta definición deja lugar a mucha ambigüedad. Podría decirse que cualquier programa informático construido desde la creación de la computación sigue tal razonamiento lógico. Por ejemplo, existen desde hace años programas capaces de multiplicar números o guardar grandes cantidades de datos, de forma similar o mejor que lo haría un ser humano. Sin embargo, no podría decirse que una calculadora o un disco duro pertenecen al grupo de inteligencia artificial.

Un humano es capaz de entrar en una habitación y reconocer en un instante cada uno de los objetos a su alrededor, las personas que lo conforman, la canción que suena y decidir qué quiere hacer a continuación, todo ello de manera casi instintiva y sin esfuerzo. Ningún programa a día de hoy es

capaz de imitar un comportamiento similar y, sin embargo, se puede afirmar que existen algoritmos que entran dentro de la categoría de inteligencia artificial. Por lo tanto, quizás una definición más precisa del término es aquella descrita por Elaine Rich que dice: “La inteligencia artificial es el estudio de cómo hacer que los ordenadores hagan cosas en las que, en este momento, las personas hacen mejor” (1983). Esta definición, a pesar de haber sido formulada hace casi cuatro décadas, sigue ajustándose mucho al concepto que se tiene hoy de la IA. Y es que, desde sus inicios, la IA es una ciencia que ha ido evolucionando a medida que avanzaba, rompiendo sus propios paradigmas y reinventándose. Se trata de una tecnología que crece muy rápidamente y que tiene cada vez más aplicaciones en sectores muy distintos de la sociedad, especialmente desde la última década. Es, por tanto, complicado tener una visión clara y actualizada del punto en que se encuentra, al día con los últimos avances, las últimas regulaciones, y las proyecciones de futuro próximo y venidero.

OBJETIVOS

Para desarrollar la investigación, es necesario establecer qué objetivos perseguimos, lo que nos ayudará a alcanzar la cuestión principal. Como ya hemos mencionado anteriormente, el propósito final de este proyecto es, a través de una minuciosa investigación y un profundo análisis, ofrecer una visión extensa sobre la situación actual de la ciencia de la inteligencia artificial, así como de las previsiones de futuro. Por lo tanto, para alcanzarlo será necesario atender a:

- Entender la inteligencia artificial y sus orígenes,
- Describir el ámbito de actuación de la misma y sus principales aplicaciones a día de hoy,
- Examinar el impacto de la IA en la sociedad y desarrollar las principales limitaciones a las que se enfrenta en la actualidad,
- Estudiar el recorrido regulatorio de la ciencia en los últimos años y revisar el marco regulatorio existente en la Unión Europea y sus objetivos a futuro,
- Analizar las áreas de mayor potencial de crecimiento y
- Explorar los posibles riesgos que el desarrollo de la IA conllevar en el futuro.

ESTRUCTURA Y METODOLOGÍA

La metodología utilizada en el trabajo se ha dividido en cinco apartados que siguen un orden cronológico: en una primera instancia se expone la historia de la IA, seguido de una introducción

al concepto de inteligencia artificial, tras esto se realiza un estudio exhaustivo del estado actual de la ciencia, y se sigue con una exploración del futuro de la tecnología. Por último, se exponen las conclusiones del trabajo.

Es relevante conocer el pasado de la IA para comprender en profundidad cuales son los motores del desarrollo de la ciencia en el presente. Por ello, en el primer apartado sobre la historia de la IA se pretende poner en contexto el estado actual de la ciencia, resaltando los periodos de mayor y menor desarrollo y las causas de esos picos.

Durante el trabajo se nombran distintos términos como machine learning, deep learning o redes neuronales, que a menudo suelen confundirse y mezclarse. En el segundo apartado se pretende introducir el concepto de IA y aclarar la diferencia entre estos conceptos.

Una vez comprendido el pasado de la IA y los conceptos más relevantes, se realiza un análisis empírico exhaustivo de su estado actual, explorando los principales motores de su desarrollo, las aplicaciones más relevantes y los desafíos y limitaciones a los que se enfrenta.

Por último, se exploran las expectativas de crecimiento de la IA, estudiando las expectativas de inversión en cada sector y las previsiones de crecimiento pronosticadas por los expertos. Para complementar, se abordan los riesgos que acompañan a un mayor desarrollo de estos sistemas reuniendo las preocupaciones de distintos expertos y organismos.

El trabajo finaliza realizando una conclusión final, en la que se recapitula el estado actual de la IA, resumiendo las principales aplicaciones, limitaciones y líneas generales de legislación y meditando sobre los futuros riesgos y la trascendencia que estos últimos pueden tener sobre el futuro desarrollo de la IA.

MARCO TEÓRICO

Debido a que se trata de un área novedosa y que se encuentra en la actualidad en constante cambio y desarrollo, existen numerosos artículos informativos que exponen la actualidad de la inteligencia artificial. Sin embargo, debido precisamente a lo rápido que se desarrolla, es primordial utilizar fuentes recientes puesto que los artículos pasan a estar rápidamente desactualizados.

En lo que respecta a aspectos generales de la IA, existen numerosos libros y artículos escritos sobre el tema puesto. El artículo “A Brief History of artificial Intelligence: On the Past, Present, and Future of artificial Intelligence” ha servido como base para desarrollar un esqueleto de la historia de la IA, que posteriormente se ha ido completado más en detalle con otras fuentes como “The History of artificial Intelligence” de Harvard.

Para la segunda parte del trabajo de la IA se ha decidido por ofrecer una explicación breve y concisa para poder entender los distintos conceptos que aparecen a lo largo del trabajo. Para ello se han utilizado diversas fuentes entre las que destacan “What is machine learning” de IBM, que ofrece un resumen conciso de los principales aspectos.

La revista de Harvard cuenta con un número extenso de artículos de gran valor académico sobre la IA que han servido de gran ayuda. Entre ellos destacan “artificial intelligence, for real.” Y “artificial intelligence for the real world”, que ofrecen información de gran interés sobre el estado actual de la IA y “artificial intelligence - The Revolution Hasn’t Happened Yet”, enfocada en el futuro de la ciencia. Posteriormente, se ha completado la información con otros artículos académicos y fuentes más especializados en cada uno de los distintos sectores de desarrollo de la IA.

La Universidad de Standford lanzó en 2014 su investigación “One Hundred Year Study on artificial Intelligence(AI)” donde, a través de un panel de expertos, estudia la influencia de la IA en la sociedad. Desde su lanzamiento, han salido dos reportes: “artificial Intelligence and life in 2030” en 2016 y “Gathering Strength, Gathering Storms” de 2021, que pretenden dar una visión en profundidad del impacto de la IA en la gente y la sociedad mundial.

Por último, en lo referente a las regulaciones sobre la IA, la autora Inga Ulricane, junto con otros autores, ha escrito varios libros analizando la regulación de la inteligencia artificial como “The Global Politics of artificial Intelligence” y “The Routledge Handbook of European Integrations”, ambos muy recientes y que ofrecen un análisis muy completo sobre el estado de la legislación sobre la IA. Además de estos libros, la otra fuente principal han sido los comunicados y documentos publicados por la Comisión Europea donde exponen sus propuestas de regulación y su visión de futuro sobre el manejo de la IA en los próximos años.

Para concluir, se han obtenido datos y estadísticas relevantes de la IA a partir de diversas fuentes como EZmarketing, Fortune Insights, Statista o Emergen Research, entre otras.

2. HISTORIA DE LA IA

INICIO

Es difícil establecer un punto exacto de aparición de la inteligencia artificial, ya que su definición es ambigua y ha cambiado desde su aparición. Sin duda un punto clave en su comienzo fue el desarrollo de la máquina denominada *The Bombe* por Alan Turing en los años 40, y que es considerada como el primer ordenador electromecánico. Gracias a esta máquina fue posible descifrar el código Enigma usado por la armada alemana en la Segunda Guerra Mundial, una tarea casi imposible para incluso las mentes más brillantes. Este hito marcó un punto de partida en lo que se refiere a las posibilidades y la inteligencia de estas máquinas.

Similarmente, podríamos establecer el punto de partida de la inteligencia artificial, al menos conceptualmente, en la literatura de ciencia ficción escrita por Isaac Asimov en los años 40. Este autor, en un relato llamado *Circulo Vicioso (Runaround)*, contaba la historia de un robot inteligente dominado por las Tres Leyes de la Robótica: (1) un robot no hará daño a un ser humano ni, por inacción, permitirá que un ser humano sufra daño; (2) un robot debe cumplir las órdenes dadas por los seres humanos, a excepción de aquellas que entren en conflicto con la primera ley; y (3) un robot debe proteger su propia existencia en la medida en que esta protección no entre en conflicto con la primera o la segunda ley (Haenlein & Kaplan, 2019). Este relato supone un antes y un después en la concepción humana de los robots y, aunque es mera ficción, fue una de las grandes inspiraciones para la búsqueda de robots inteligentes. A día de hoy, incluso, y aunque no existen aún robots como los que se plantean en el libro, puede aún verse la influencia de las Tres Leyes de la Robótica.

No fue, en cambio, hasta 1956 cuando se acuñó por primera vez el término “inteligencia artificial” de la mano de Marvin Minsky y John McCarthy, los cuales organizaron un encuentro de estudio denominado Conferencia de Dartmouth (en inglés: Dartmouth Summer Research Project on artificial Intelligence), con el objetivo de crear máquinas que simularan la inteligencia humana. Entre los participantes se encontraban personas como Nathaniel Rochester, inventor del primer ordenador comercial, el IMB 701, o Claude Shannon, matemático fundador de la teoría de la información.

El desarrollo de esta conferencia marcó el inicio de un periodo de más de dos décadas de gran proliferación de la inteligencia artificial. Durante este tiempo se produjeron avances como el programa ELIZA, creado por Joseph Wizenbaum en 1966, capaz de simular conversaciones con humanos, o el Solucionador General de problemas, desarrollado por el premio nobel Herbert Simond, capaz de resolver problemas simples. Este periodo se caracteriza además por una creciente inversión por parte de los distintos países en inteligencia artificial. Entre ellas destaca la Agencia de Proyectos de Investigación Avanzados de Defensa del Gobierno de EEUU, conocida por sus siglas en ingles DARPA. Los avances con ELIZA convencieron al DARPA de invertir en esta tecnología, con un interés especial en el desarrollo de programas de transcripción y traducción de lenguaje hablado.

DEEP LEARNING

Durante estos mismos años se habría venido dando, al mismo tiempo, el inicio silencioso de una tecnología que no tomaría protagonismo hasta 2015, el aprendizaje profundo o deep learning (DL). Sus inicios se remontan a 1943, con la publicación de “A Logical Calculus of the Ideas Immanent in Nervous Activity”, de Walter Pitts y Warren McCulloch. En este documento se presenta un modelo matemático que imita el comportamiento de una neurona a través de una serie de cálculos matemáticos denominados lógica de umbral. Este hito es considerado como el inicio del estudio de las redes neuronales artificiales. Tras esto se siguieron otros eventos importantes. En 1957 Frank Rosenblatt idea el Perceptron, un algoritmo de clasificación que contaba con ciertas capacidades de aprendizaje. En 1960 se crea el primer modelo de retropropagación de errores de la mano de Henry J. Kelley, y una versión más simple de esto dos años más tarde por Stuart Dreyfus. El primer intento de creación de algoritmos de deep learning sucede en 1965 de la mano de Alexey Grigoryevich Ivakhnenko y Valentin Grigor'evich Lapa, quienes crean el primer perceptron multicapa usando modelos con funciones de activación polinómicas. Ivakhnenko es considerado por algunos como el padre del deep learning.

A pesar de estos avances, la investigación en el campo se estancó a partir de 1969. El detonante de este estancamiento fue principalmente el libro publicado por Marvin Minsky y Seymour Papert llamado Perceptrons. En él, los autores mostraban que el perceptrón de Rosenblat no sería capaz de resolver funciones complicadas como la XOR, al menos no con el poder de procesamiento con el que se contaba en la época.

PRIMER INVIERNO DE LA IA

La publicación del libro *Perceptrons* marcó un cambio en la tendencia optimista que se vivía sobre la IA hasta el momento al mostrar predicciones bastante pesimistas sobre la tecnología. Sin embargo, no fue hasta 1973 cuando comenzaron a ser evidentes las grandes limitaciones con que contaban aún los programas de inteligencia artificial. El proyecto de DARPA para la creación de un algoritmo de transcripción de lenguaje no dio los resultados esperados. Era evidente que los ordenadores de la época no contaban con el poder de procesamiento necesario para tareas de aquel calibre. Esto, unido a la falta de avances en el campo de la IA, hizo que las esperanzas sobre las posibilidades de esta tecnología decayeran y los gobiernos estadounidense y británico redujeron enormemente la inversión, lo que dio lugar a una etapa comúnmente conocida como el invierno de la IA, caracterizada por la falta de inversión y de interés en el campo. La razón del fracaso fue el enfoque que se le dio, ya que se trató de imitar la inteligencia artificial mediante la mecanización de reglas de razonamiento para tomar decisiones. Si bien estos sistemas pueden dar lugar a buenos resultados en determinadas áreas con alta formalización, cuentan con muchas limitaciones en áreas que no permiten ese nivel de formalización, tal como el lenguaje natural o el reconocimiento facial. Para estos casos, son necesarios sistemas capaces de interpretar datos y aprender de ellos. Por otro lado, esta decaída de interés en la ciencia también estuvo muy fomentada por la creación de expectativas poco realistas durante los años anteriores, que de la mano del boom que se estaba produciendo en el interés en el campo llevaron a marcarse objetivos imposibles. Por ejemplo, el propio Marvin Minsky, nombrado anteriormente, que dijo en 1970 en una entrevista en la revista LIFE que “Dentro de tres a ocho años, tendremos una máquina con la inteligencia general de un ser humano promedio” (LIFE Magazine, 1970). Sin embargo, cuando comenzó a dilucidarse la verdadera dificultad de la tarea que supone imitar la inteligencia humana, el interés cayó.

BOOM DE 1980

A partir de los años 80 hubo un resurgimiento de interés en el campo de la mano de la creación de los sistemas expertos, programas que imitaban el proceso de decisión de humanos expertos. Estos sistemas funcionaban preguntando a un experto cómo actuar ante cualquier situación, y luego este conocimiento era trasladado a no expertos. Durante varios años hubo una gran inversión en esta tecnología, sobretodo en Japón, donde lanzaron su proyecto de Sistemas de Computadoras de Quinta Generación (FGCS por sus siglas en inglés). Se esperaba que esta generación de

ordenadores pudiera ser capaz de contar con razonamiento elemental, traducir lenguajes, recibir órdenes por voz, y ser, en general, más fáciles de usar que los ordenadores de la época. Sin embargo, a pesar de los más de 400 millones de dólares invertidos en esa década, la tecnología acabó por no proporcionar los resultados esperados, y mostró ser cara de mantener y con tendencia a fallos (Anyoha, 2017). Como resultado de esto, hubo de nuevo una recaída del interés en la IA, dando lugar al segundo invierno de la inteligencia artificial.

SEGUNDO INVIERNO DE LA IA

Durante los años 90 hubo, de nuevo, un descenso importante de la inversión pública en IA. Muchos, incluso, evitaban incluir las palabras “inteligencia artificial” en sus proyectos ya que este término pasó a ser visto con pesimismo, y solía recibir poca inversión. A pesar de todo esto, la última década del siglo XX reunió una serie de avances importantes.

Uno de estos avances fue Deep Blue, el programa de ordenador que fue capaz de ganar al campeón del mundo de ajedrez Garry Kasparov en 1997. No fue ni mucho menos una victoria aplastante, sino que el resultado final fue de un 3.5-2.5 a favor de Deep Blue. Además, no puede decirse que la victoria se debiese a un gran avance en la inteligencia de los ordenadores, sino que su fuerza se basaba principalmente en su gran poder de cálculo (poder calcular con antelación muchas jugadas en adelante), y también en su extensa base de datos donde se recopilaba información sobre la teoría existente sobre el juego (aperturas, partidas jugadas por otros campeones, conceptos, etc.). Aun así, a pesar de que a nivel tecnológico no supuso un avance predominante en la tecnología, fue un caso muy mediático y que para muchos supuso la primera victoria de una máquina sobre un humano y el inicio de una era marcada por la superioridad de los ordenadores. Desde entonces los motores de ajedrez han avanzado mucho y, a día de hoy, no existe ninguna duda acerca de su superioridad sobre cualquier jugador humano. Durante este periodo también se dieron avances como el software de reconocimiento de voz de Dragon Systems, que fue implementado en Windows, o el robot de reconocimiento e imitación de emociones Kismet desarrollado por Cynthia Breazeal.

RESURGIR DE LA IA

No fue hasta la segunda década del siglo XXI cuando comenzaron a darse nuevos progresos relevantes en el campo. En 2011 Apple introdujo su asistente virtual Siri, al que siguieron los

asistentes de Google en 2012 con GoogleNow y de Microsoft en 2014 con Cortana. Tras esto, en 2016 se produjo un nuevo hito histórico similar a la victoria de Deep Blue sobre Kasparov, esta vez en un juego mucho más complejo denominado Go, en el que previamente los algoritmos creados solo habían alcanzado nivel amateur. Este juego, debido a su estructura, da lugar a un número posible de movimientos mucho más grande que el ajedrez, lo que hacía creer que un ordenador nunca podría ser capaz de vencer a un humano. En esta ocasión, un programa creado por Google denominado AlphaGo consiguió vencer al campeón del mundo Lee Sedol, marcando un nuevo hito. Las habilidades del equipo de Google se demostraron nuevamente al año siguiente cuando un equipo de ingenieros desarrolló AlphaZero, un motor de ajedrez con un enfoque completamente distinto al DeepBlue de 1997. AlphaZero, en vez de utilizar extenso cálculo y teoría, utilizaba redes neuronales y machine learning. Con ello, y a pesar de que no se le introdujo ninguna documentación sobre partidas pasadas o teoría desarrollada sobre el juego, simplemente a través del entrenamiento autónomo, jugando partidas contra sí mismo, AlphaZero fue capaz de vencer al motor de ajedrez más potente de aquel entonces, Stockfish8 en una competición de gran cobertura mediática. Uno de los elementos más sorprendentes de estas partidas fue el modo de juego empleado por AlphaZero ya que, al no estar influido por partidas previas, mostró un estilo de juego muy diferente al existente, con numerosas innovaciones. El nuevo programa venció a Stockfish en un campeonato disputado en 100 partidas, ganando con un determinante resultado de 28 victorias, 72 empates y 0 derrotas. Con este hito se demostró el potencial que tenían los modelos de aprendizaje por redes neuronales. AlphaZero no siguió desarrollándose, puesto que fue un proyecto puntual, pero desde entonces los módulos de ajedrez han incluido esta tecnología a sus modelos, combinándola con la ya existente. Esto ha resultado en motores mucho más poderosos y, además, más originales. Su influencia en el juego entre humanos también ha sido determinante, mostrando nuevas formas de entender el juego. A día de hoy, los jugadores trabajan a diario con estos motores para practicar, y analizar las partidas y posiciones que juegan con el fin de ver sus errores y aprender de ello.

Desde entonces, en los años recientes la IA ha tomado gran predominancia, asentándose en cada vez más sectores. El auge de la tecnología en la sociedad ha permitido que el número de datos disponibles se multiplique exponencialmente gracias al Big Data, que unido a una capacidad de procesamiento mucho mayor a la de las décadas pasadas ha incrementado significativamente las posibilidades que ofrece la IA.

Se llega así al momento actual de la IA, un momento marcado, como se expondrá a continuación, por una gran proliferación de la IA en la casi totalidad de los sectores de la economía. Un momento en el cual la IA ha pasado de ser concebida como una ciencia puramente académica a dar el salto a la aplicación práctica, mostrando una pequeña parte del potencial que tiene de revolucionar la industria, los negocios, la vida de la gente e incluso la cultura.

3. INTRODUCCIÓN A LA INTELIGENCIA ARTIFICIAL: QUÉ ES Y CÓMO FUNCIONA.

Antes de comenzar a estudiar el estado actual de la inteligencia artificial es necesario comprender qué es, cómo funciona, y cuáles son las diferencias entre inteligencia artificial, machine learning, deep learning y redes neuronales, ya que son conceptos que aparecen en numerosas ocasiones a lo largo del trabajo.

La inteligencia artificial es un desarrollo científico por el cual una máquina es diseñada para llevar a cabo acciones similares a las de la inteligencia humana. El principal objetivo de esta ciencia es hacer que los ordenadores hagan aquello que los humanos son capaces de hacer. Según su aplicación, se puede dividir la tecnología en dos tipos. Por un lado, aquellas máquinas diseñadas para especializarse en una tarea específica se les denomina IA aplicada o débil, y es la más común. Un ejemplo de ello sería un módulo de ajedrez o un coche autónomo. Por otro lado, si la inteligencia tiene la capacidad de realizar cualquier acción que se le asigne se denomina IA generalizada o robusta. A día de hoy, no existen algoritmos de IA que cumplan estas condiciones.

El machine learning (ML) es un subcampo de la IA que reúne a aquellas aplicaciones capaces de hacer predicciones y aprender a partir de datos sin estar explícitamente programado. Para ello, hacen uso de métodos estadísticos para buscar patrones y relaciones desconocidos entre los datos, los cuales sirven de entrenamiento para el algoritmo. A través de este entrenamiento, los algoritmos de machine learning son capaces de aprender y proporcionar clasificación y predicción de nuevos datos, creando de este modo nuevo conocimiento.

El proceso a partir del cual operan los algoritmos de machine learning puede dividirse en tres partes. En primer lugar, a través del proceso de decisión, el algoritmo hace predicciones o clasificaciones sobre un conjunto de datos. Para ello los algoritmos buscan patrones desconocidos en los datos que asistan en la toma de decisiones. En segundo lugar, la función de error juzga el desempeño del proceso de decisión, según si sus predicciones o clasificaciones han sido correctas o incorrectas. En tercer lugar, el proceso de optimización del modelo se encarga de ajustar los pesos del modelo para reducir las discrepancias entre las estimaciones y los datos reales. Este proceso de tres partes se lleva a cabo de forma reiterada, mejorando así los resultados del proceso.

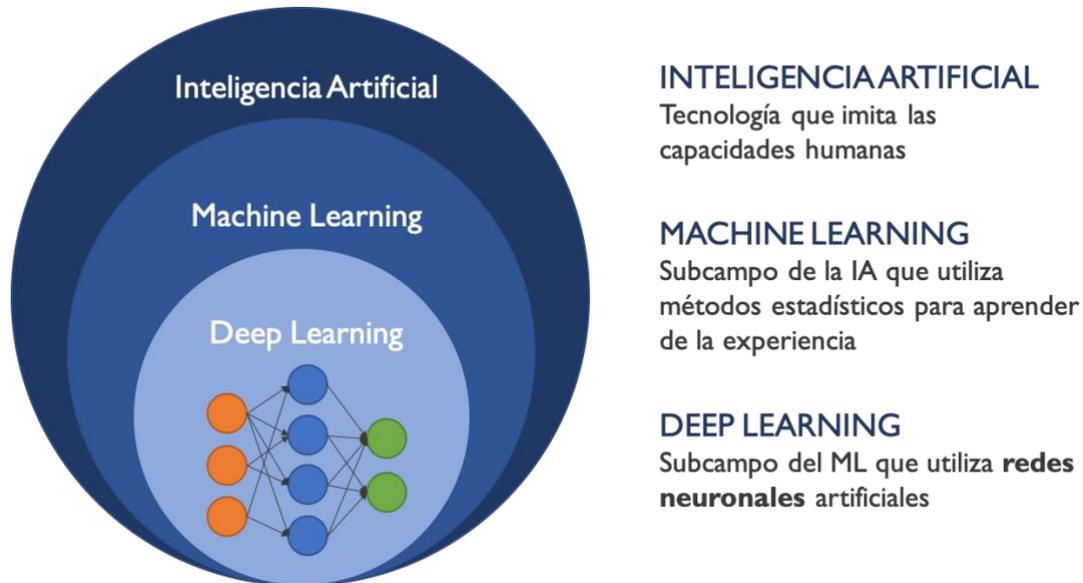
Los algoritmos de ML suelen clasificarse en tres grupos: supervisado, no supervisado, y reforzado. El aprendizaje supervisado incluye la mayoría de aplicaciones actuales, y se utiliza aquello que ya se ha aprendido en el pasado sobre datos nuevos para hacer predicciones futuras. En este tipo, se utiliza una serie de datos previamente clasificados como muestra de entrenamiento. El programa, a partir de estos datos, se entrena buscando relaciones que le permitan clasificar o predecir correctamente los datos, y utiliza las clasificaciones existentes para evaluar sus errores y mejorar. Un ejemplo de esto sería un algoritmo encargado de dividir una serie de correos en correo deseado y no deseado, partiendo como muestra de un conjunto de emails previamente clasificados. El aprendizaje no supervisado, por otro lado, no necesita de la existencia previa de datos clasificados. Este tipo de aprendizaje no busca clasificar la salida, sino que explora relaciones y organizaciones internas ocultas dentro de los datos. Finalmente, en el aprendizaje reforzado el algoritmo aprende a base de prueba y error qué acciones proporcionan una mayor recompensa. El objetivo del algoritmo será por tanto maximizar la recompensa obtenida durante un periodo de tiempo. A través de probar al azar, el programa buscará por tanto la solución óptima que da lugar a mejores resultados (Ongsulee, 2017). Una aplicación común que sigue este ejemplo son los algoritmos programados para videojuegos. Mediante este método, el programa puede aprender a dominar cualquier videojuego sin que sea necesario explicarle siquiera las nociones básicas de cómo jugarlo, tan solo a través de la prueba y el error y la búsqueda de recompensas.

Es común que los términos machine learning y deep learning se confundan y usen de manera indiscriminada, ya que ambas tecnologías son muy similares, sin embargo, no son exactamente lo mismo. Mientras que ambos son subcampos de la inteligencia artificial, el deep learning es, de hecho, un subcampo del machine learning. En concreto, deep learning se refiere a un tipo de ML donde una parte del proceso está automatizada, lo que permite el uso de conjuntos de datos más grandes y menos estructurados, sin la necesidad de tanta intervención humana. Podría decirse que el DL es un tipo escalable de ML, y entraría dentro del grupo de aprendizaje no supervisado.

El deep learning utiliza redes neuronales artificiales, que consisten en una red de nodos formada por varias capas: una capa de entrada, una de salida, y una o más capas intermedias. Se le denominan “neuronales” porque estas capas simulan el funcionamiento de las neuronas, ya que, de igual modo que el cerebro funciona a través del paso de información entre las conexiones entre neuronas, las redes neuronales funcionan mediante conexiones entre los nodos de las distintas capas los cuales cuentan cada uno con un peso que influye en el resultado final. Además de los

pesos, los nodos cuentan con un peso de barrera. Si el peso del nodo supera la barrera entonces el nodo se activa, mandando datos a la siguiente capa.

Figura 1: Esquema Explicativo IA.



Fuente: Elaboración propia.

Los algoritmos de DL tienen ventaja sobre las generaciones anteriores de machine learning en que son capaces de usar conjuntos de datos mucho más grandes para mejorar sus predicciones. Mientras que los enfoques típicos de ML mejoran con conjuntos de datos, su mejora siempre se frena en un punto, a partir del cual ya no avanzan. Los algoritmos de deep learning, en cambio, no muestran señales de frenar su aprendizaje, sino que siguen progresando cuanto mayor es el volumen de datos que se les da como entrenamiento. Trabajar con más datos requiere, como era de esperar, mayor capacidad de procesamiento. Por suerte, a día de hoy existe una cantidad inmensa de datos digitales, lo que abre un gran abanico de posibilidades para el desarrollo de esta tecnología. Algunos sistemas de deep learning han sido entrenados para utilizar 36 millones de ejemplos o más (Wu, 2016). Sin embargo, a pesar del gran éxito que está teniendo esta tecnología, especialmente en el campo del tratamiento de imágenes, lo cierto es que el deep learning cuenta con dos limitaciones principales. La primera es que necesita de conjuntos de datos extremadamente grandes para proporcionar buenos resultados. Y la segunda es la baja interpretabilidad de sus decisiones, que hace difícil sacar conclusiones y que convierte a los algoritmos de inteligencia artificial en una “caja negra”.

4. ESTADO ACTUAL DE LA CIENCIA

4.1. PRINCIPALES IMPLICACIONES

Las innovaciones tecnológicas han sido durante los últimos siglos los principales motores del crecimiento económico. Nuevos inventos, como el motor de combustión, han llevado al desarrollo explosivo de la economía y, a su vez, al desarrollo de nuevas innovaciones tecnológicas.

La inteligencia artificial, en concreto el machine learning, es para muchos la ventana a un nuevo universo de innovaciones tecnológicas. Al igual que otras innovaciones anteriores como la electricidad o el internet, el machine learning tiene el potencial de pasar a considerarse una tecnología de propósito general (GPT por sus siglas en inglés), es decir, una tecnología capaz de alterar un sistema económico y producir cambios en los factores sociales. Si bien aún es pronto para saberlo, la inteligencia artificial es una tecnología con la capacidad de incrementar la productividad en desarrollo e investigación. Podría ser, por tanto, la base para el comienzo de una Cuarta Revolución Industrial (Crafts, 2021).

FACTORES DETERMINANTES

Si se echa la vista atrás se puede observar que, durante muchos años desde el inicio de la investigación en IA, a pesar de la gran inversión por parte de muchos países, los avances han sido siempre muy lentos y muy por debajo de las expectativas marcadas en el momento, lo que llevó a muchos a adoptar una visión muy pesimista sobre las verdaderas posibilidades de esta tecnología y a los gobiernos a dejar de invertir en ella. Sin embargo, a día de hoy el desarrollo se está dando a un ritmo acelerado, y son pocos los sectores donde la IA no esté mostrando resultados. Existen tres factores principales que han colaborado a que hoy en día la IA esté avanzando tanto: mejores algoritmos, mejor hardware y mayor cantidad de datos.

El primero de los factores es la mejora en los algoritmos de IA. Durante los últimos años han surgido nuevas formas de aprendizaje tales como el deep learning supervisado y el aprendizaje reforzado, que han permitido que sea posible entrenar algoritmos con mayor cantidad de datos sin que alcancen pronto un estado donde más datos no mejoran el resultado. Además, conceptos como los algoritmos de open source (algoritmos que están disponibles públicamente y que cualquiera puede investigar, modificar y mejorar) favorecen el intercambio de conocimiento sobre la ciencia

a gran escala. A día de hoy, los nuevos avances en el área se trasladan muy rápidamente al resto del mundo.

El segundo factor de influencia es la mejora en el hardware disponible. Tal como se ha explicado, los nuevos algoritmos son capaces de aprender utilizando bases de datos mucho mayores, pero para eso también es necesario un poder de procesamiento mayor. Durante las últimas décadas la capacidad de los microprocesadores ha aumentado a un ritmo alineado con la ley de Moore, duplicándose cada dos años. Esta mejora en el poder de procesamiento ha abierto las puertas a un nuevo mundo de oportunidades donde la aplicación de IA no era anteriormente viable. Además, no solo la capacidad de procesamiento de los ordenadores es mayor, sino que ahora existen numerosas alternativas al uso de procesamiento privado que consisten en el aprovechamiento del procesamiento en la nube. Aun así, a pesar de que todos los algoritmos cada vez requieren de una capacidad de procesamiento mayor, todo apunta a que este crecimiento se frenará en los próximos años, lo que frenaría el desarrollo de la IA. Una posible solución a esto es un tipo de chip llamado Unidad de Procesamiento Gráfico (GPU por sus siglas en inglés) que funciona especialmente bien con el tipo de cálculos necesarios en las redes neuronales. Algunas compañías incluso, como Google, han desarrollado su propio chip optimizado específicamente para este uso, llamado Unidad de Procesamiento Tensorial (TPU en inglés) (Brynjolfsson & McAfee, 2017)

Por último, el tercer factor que ha influenciado en gran medida el rápido crecimiento de la IA es la mayor disponibilidad de datos surgida a raíz de la revolución que se está dando a nivel global. Durante los últimos años ha habido un boom en el uso de móviles, fotos, videos, redes sociales, sensores de todo tipo, y otros sistemas similares, que generan cada minuto cantidades enormes de datos digitalizados. El ratio de creación de datos a día de hoy es significativamente mayor al que era hace unos años, y una gran parte de esos datos se sube a internet. Hoy, más que nunca, los datos se han convertido en un factor clave para muchas de las empresas que operan en internet. Algunas, como Twitter o Facebook, ofrecen sus servicios gratuitamente y es a través de los datos de donde sacan sus principales beneficios, tanto a través de campañas de publicidad dirigidas específicamente al público, como haciendo ellos mismos uso de los datos, o directamente con su venta a terceros. La gestión de los datos por parte de estas compañías y su efecto en la privacidad de las personas es uno que está generando gran debate a día de hoy.

Estos tres factores han contribuido a que, desde 2020, la IA se encuentre en auge y sea una de las inversiones más atractivas. De 2010 a 2020, la inversión global en startups sobre IA ha pasado de 1.300 millones de dólares a más de 40.000 millones, con un crecimiento anual cercano al 50%. En lo que respecta al sector académico, el número de trabajos publicados sobre el tema se ha multiplicado por doce desde inicios de siglo, así como el número de conferencias y de participantes interesados sobre el tema. Es tal el interés mostrado en la materia que más de un 22% de los estudiantes del doctorado de informática se especializan en este tema (Toosi, Bottino, Saboury, Siegel, & Rahmim, 2021).

VENTAJAS DEL MACHINE LEARNING

Lo que hace al machine learning una herramienta tan atractiva es la capacidad que tiene de mejorar sus resultados sin la necesidad de que un humano le diga cómo. Esto es especialmente útil por dos razones principales. En primer lugar, debido a la posibilidad de aplicarlo en aquellos ámbitos donde no puede expresarse la mejora a través de una serie concreta de reglas, ya sea por su complejidad o por el desconocimiento de ello. La mente humana es capaz de llevar a cabo tareas muy complejas para las que no es capaz de dar explicación a cómo lo hace. A este hecho se le denomina la paradoja de Polanyi, por el filósofo y matemático Michael Polanyi, el cual describió que las capacidades humanas y su conocimiento acerca del modo en que funciona el mundo están más allá de su comprensión explícita. Este efecto tiene un impacto determinante, no solo en lo que podemos comunicar entre nosotros, sino en nuestra habilidad de transferir ese conocimiento a las máquinas. Un ejemplo claro de ello es el reconocimiento facial. Si bien cualquier ser humano es capaz de identificar caras y expresiones, difícilmente es capaz de explicar que elementos concretos le han llevado a tales conclusiones. En tareas como esta, la inteligencia artificial juega un papel clave puesto que es capaz de aprender por su cuenta, lo que permite automatizar tareas que hasta ahora no era posible automatizar.

El segundo factor que hace al machine learning tan útil es su gran capacidad de aprendizaje, la cual supera a la de los humanos. Los sistemas entrenados por ML son capaces de producir resultados con mayor precisión y rapidez que la de un humano. Así, gracias a esta tecnología, el conocimiento que se le puede transferir a un ordenador ya no está limitado por el propio conocimiento del que lo programa, sino que puede ser aprendido de forma independiente por la máquina. Esto se vio presente, por ejemplo, en la batalla entre módulos de ajedrez donde

AlphaZero, que había aprendido a jugar al ajedrez únicamente por machine learning, sin conocimientos humanos previos sobre el juego, venció a un módulo de ajedrez que contaba con toda la teoría y conocimientos humanos sobre el juego conocidos hasta el momento.

CAMBIOS PRODUCIDOS POR EL ML

Con estas ventajas en mente, el machine learning está calando cada vez más en distintos sectores de la economía, produciendo cambios en su modo de funcionamiento. Estos cambios pueden clasificarse en tres grupos: cambios en las tareas, cambios en los procesos de negocio y cambios en el modelo de negocio. (Brynjolfsson & McAfee, 2017)

En primer lugar, los cambios en las tareas incluyen aquellos que implican la introducción de sistemas de IA para ocupar o asistir en puestos anteriormente ocupados por humanos. Gracias a la IA, puede revolucionarse la concepción de muchos de los puestos de trabajo que existen actualmente, dejando aquellas tareas que se puedan automatizar a un algoritmo y permitiendo que los humanos se centren en otras más complejas. Esto, como se verá más adelante, no implica necesariamente la pérdida de puestos de trabajo por humanos, sino que muy a menudo los algoritmos necesitan de la presencia complementaria de un humano, y su implicación rara vez sustituye completamente su trabajo. Un ejemplo de ello son los algoritmos que asisten en la detección de células de cáncer, donde el programa es capaz, por medio del tratamiento de imagen, de estudiar un número muy grande de fotografías para seleccionar aquellas donde existen indicios de células cancerígenas. Este proceso no sustituye al de un médico, sino que le libera de esa tarea repetitiva y le deja más tiempo para emplearlo en otras tareas más importantes.

El segundo grupo reúne los cambios en los procesos de negocio, que se refiere a una redefinición en el modo en el que se llevaba a cabo un procedimiento gracias a la introducción de la IA, lo cual puede ocurrir de dos maneras. Por un lado, es posible maximizar el modo de operación de un negocio a través de sistemas de machine learning. Si bien esto ya era posible con los algoritmos tradicionales, el ML ofrece un grado más amplio de flexibilidad. Este es el caso de Amazon, que utiliza algoritmos para optimizar su gestión del inventario. Por otro lado, gracias a las nuevas posibilidades que ofrece la inteligencia artificial, es posible realizar cambios sustanciales en la operación de un proceso. De nuevo Amazon es un ejemplo de ello, ya que desde hace unos años comenzó a implementar robots en sus centros y son los que se encargan de una gran parte de la logística.

Por último, el tercer tipo de cambio la IA puede llegar a producir en un negocio es en el modelo de negocio. La IA ofrece a los negocios un nuevo abanico de posibilidades a explorar que les permiten, no solo mejorar la forma en la que operaban, sino también redefinirla. Las plataformas multimedia, por ejemplo, han cambiado su modelo de negocio en los últimos años incorporando algoritmos que estudian los datos de los usuarios, pasando de un negocio basado en proporcionar a la carta aquello que el cliente busca a ser ellos los que recomiendan al cliente contenido multimedia basado en sus preferencias. Esto, a su vez, proporciona a estos servicios una cantidad muy grande de datos sobre los gustos y preferencias de los consumidores, datos que pueden llegar a convertirse por sí mismos en una fuente muy rentable de ingresos.

4.2. APLICACIONES

En unos pocos años, la IA ha pasado de ser una tecnología en desarrollo y con un potencial muy limitado a un recurso de gran valor, siendo de gran utilidad en una variedad muy grande de sectores. Son muchas las aplicaciones que ya se están poniendo en marcha a día de hoy de la IA y es imposible nombrarlas todas. Sin embargo, es posible agrupar los principales avances en dos grupos de trabajo: la percepción y la cognición.

PERCEPCIÓN

El campo de la percepción incluye a todas aquellas aplicaciones relacionadas con la captación sensorial de información. Así, se encuentran dentro de este grupo aplicaciones relacionadas con la voz, las imágenes, los sonidos, etc. A día de hoy, algunas máquinas son capaces de igualar e incluso superar las habilidades humanas en algunos de estos ámbitos. Sin embargo, sus capacidades siguen siendo limitadas y en muchas ocasiones es más barato y mejor dejar que un humano realice la tarea. Sin embargo, a medida que la tecnología disponible consiga superar las capacidades humanas, su crecimiento se volverá mucho más acelerado y serán más las situaciones en las que estos sistemas podrán implementarse.

RECONOCIMIENTO DE VOZ

Entre las aplicaciones del campo de la percepción, una de las más presentes en el día a día actual es el reconocimiento de voz. Herramientas como Alexa, Siri o Google Assistant, usadas a diario por millones de personas, tienen detrás complejos algoritmos de IA capaces de transcribir voz a texto con márgenes de error cada vez menores. Un estudio llevado a cabo por James Landay, en

Standford, mostró que la escritura a través del reconocimiento de voz del móvil es de media tres veces más rápida que la escritura normal (2016). Además, otro estudio reciente mostró que el 52% de la gente prefiere las búsquedas de información por voz frente a la clásica búsqueda por internet, debido principalmente a su comodidad (Laricchia, 2017). El mercado global de estas herramientas ha crecido en los últimos años, con un valor estimado de 11.900 millones de dólares tan solo en EEUU en 2019, y el cual se espera que crezca hasta los 35.500 millones en 2025 (Laricchia, Smart Speaker Market Value Worldwide 2014–2025, 2022). Actualmente, la investigación en el campo del reconocimiento de voz se desarrolla en tres frentes: la mejora de la tecnología para aumentar su eficacia, el incremento de la seguridad y la privacidad, y el estudio de las motivaciones de uso de los usuarios (Zwakman, Pal, & Arpnikanondt, 2021)

RECONOCIMIENTO DE IMAGEN

Junto con el reconocimiento de voz, otra de las aplicaciones del campo de la percepción que merece una mención especial es el reconocimiento de imagen. Esta tecnología es capaz de extraer y clasificar las características de una imagen, permitiendo al ordenador saber qué es lo que está viendo. Este proceso funciona habitualmente en tres pasos: preprocesamiento de la imagen, extracción de características y clasificación. Como se expuso anteriormente, es gracias a la aparición de los algoritmos de deep learning que el reconocimiento de imagen ha alcanzado el nivel de avance que tiene a día de hoy, ya que esta tecnología permite entrenar algoritmos con grandes cantidades de datos no clasificados. La utilidad del reconocimiento de imagen es muy alta, y está siendo usada en muchos campos como la medicina, el transporte, la comunicación y la seguridad (Yinglong Li, 2022).

MEDICINA

El campo de la medicina es uno que destaca especialmente ya que es uno de los lugares donde mayor y más rápido impacto está teniendo la IA a día de hoy, con importantes avances en la detección de todo tipo de enfermedades a través del tratamiento de imágenes, así como del análisis de muestras que ayuda en la elaboración de nuevos tratamientos. Un ejemplo de ello es Enlitic, que usan redes neuronales para asistir en el diagnóstico de enfermedades mediante el análisis de imágenes médicas. O AlphaFold, desarrollado por DeepMind, que utilizó deep learning para lograr avances en un problema de varias décadas de antigüedad conocido como “problema de plegamiento de proteínas” (Toosi, Bottino, Saboury, Siegel, & Rahmim, 2021). Además de esto,

el tratamiento de imagen también puede ayudar a los médicos a la hora de tomar decisiones y asistirles durante las cirugías. La limitación que encuentran estas aplicaciones es la dificultad de encontrar grandes cantidades de datos que sirvan como entrenamiento, ya que para ello serían necesarias muchas horas de grabación y monitoreo de cirugías que pueda utilizar el algoritmo, algo que escasea y que, en caso de existir, sería complicado conseguirlo debido a la privacidad inherente a los datos médicos de los pacientes. La dificultad de obtención de datos es un problema común a muchas de las aplicaciones de IA.

En concreto, destaca el papel que ha tenido la IA en la lucha contra el COVID-19. Los avances en el campo de la medicina han permitido poner en marcha múltiples aplicaciones que utilizan IA. En lo que respecta a la detección, han tenido un papel clave, con algoritmos que detectan el COVID-19 en pruebas de pecho con hasta un 95% de precisión, u otros como Blue-dot, que ayudan a la detección temprana identificando pacientes que muestran problemas respiratorios gracias al reconocimiento facial, los movimientos, el sonido y la luz. En el tratamiento de pacientes, se han implementado algoritmos que han desarrollado caracterizaciones de tejidos y técnicas de manejo del dolor y robots manejados por IA que han operado asistiendo pacientes y entregando medicinas, evitando riesgos de contagio. Por otro lado, muchas de las clases y prácticas médicas han pasado a impartirse online con la ayuda de simuladores (Ahuja & Nair, 2021).

TRANSPORTE

En lo que respecta al transporte, el reconocimiento de imagen por IA está permitiendo el desarrollo de coches inteligentes autónomos, capaces de analizar la información que reciben a través de sus cámaras y procesarla para identificar obstáculos y peatones, leer las señales de tráfico y reaccionar a las condiciones de la carretera. En los últimos años esta habilidad ha mejorado, y el error a la hora de reconocer peatones se ha reducido de uno de cada 30 fotogramas a menos de uno de cada 30 millones (Brynjolfsson & McAfee, 2017). Estos coches no solo son capaces de procesar más rápido que un humano la información de sus sensores, sino que a través de la información disponible en la red son capaces de optimizar la conducción conociendo con antelación el estado del tráfico o del tiempo y, en el futuro, la información recibida por otros coches autónomos y otros sensores distribuidos por la ciudad.

DETECCIÓN FACIAL

También está cobrando gran importancia la detección facial. Gran parte de los móviles actuales son capaces de desbloquearse mediante reconocimiento facial, y empresas como Google y Facebook, por ejemplo, son capaces de detectar en las fotos las caras de amigos para sugerir etiquetas. Otras empresas como Affectiva, han entrenado sus algoritmos para detectar emociones. En esta misma línea, el tratamiento de imágenes también está resultado de utilidad para compañías como Aptonomy and Sanbot, que han implementado algoritmos de IA para mejorar sus sistemas seguridad y de cámaras de vigilancia. El tratamiento de imagen aplicado a cámaras de seguridad permite identificar en tiempo real lo que se observa, añadiendo una capa extra de seguridad. Este tema también ha producido controversias pues algunos temen de los riesgos para la privacidad que pueda suponer si el reconocimiento facial se usa a gran escala. En China, por ejemplo, el uso de la tecnología de reconocimiento facial se emplea de manera generalizada en servicios como la seguridad o las transacciones.

GANs

Cabe mencionar también una tecnología muy relacionada con el tratamiento de imagen y que ha supuesto uno de los mayores avances en inteligencia artificial de los últimos años, las Redes Neuronales Generativas Adversativas, comúnmente llamadas GANs por sus siglas en inglés. Esta tecnología consiste en un nuevo enfoque en el modo de entrenar las redes neuronales por el cual se contraponen dos algoritmos para que compitan entre ellos, uno de ellos encargado de generar nuevo contenido realista y un segundo encargado de clasificarlo como verdadero o falso. Ambos algoritmos compiten el uno contra el otro en un juego de suma cero, lo que permite que ambos aprendan el uno del otro y mejoren en sus respectivas tareas. El entrenamiento termina cuando el algoritmo clasificador acierta tan solo el 50% del tiempo, lo que indica que el algoritmo generador está creando ejemplos suficientemente buenos. De esta forma, es posible generar nuevo contenido original indistinguible de los datos de entrenamiento originales. Así, gracias a esta tecnología, hoy en día es posible crear imágenes muy realistas que un humano no es capaz de distinguir de una real, como por ejemplo rostros humanos de gente que nunca ha existido. Además de esto, es posible llevar los algoritmos GANs un paso más allá y especificar condiciones en su generación de contenido, lo que permite utilizarlos para aplicaciones como la transformación de texto a imagen o la transformación de imagen a imagen (Brownlee, 2019). Esto posibilita, por ejemplo, crear

imágenes o cuadros a partir de cualquier concepto que el autor imagine, incluso, con un sistema lo suficiente sofisticado, crear videos generados completamente por IA. A continuación se presenta un ejemplo el proceso de creación de una imagen a partir de un código open source que emplea la tecnología GANs:

Figura 2: Tres iteraciones del proceso de creación de una imagen con GANs a partir de las palabras “león, corona y llaves”, elementos que conforman el escudo de ICAI.



Fuente: Elaboración propia.

Además de la creación de imágenes y videos, gracias a esta tecnología surgen aplicaciones muy variadas y originales como, por ejemplo, la creación de videojuegos de mundo abierto en los que el paisaje sea generado de forma aleatoria por IA. Es tal el nivel de realismo que han alcanzado algunas de estas aplicaciones que se ha generado un debate público sobre las posibles consecuencias negativas de esta tecnología, en concreto de cómo puede afectar a aquello que creemos como cierto o no. Por ejemplo, es posible a día de hoy, a partir de una muestra inicial de una persona, crear un video realista de esa persona haciendo algo que nunca ha hecho, como por ejemplo cometiendo un delito, diciendo algo que nunca a dicho o haciendo cualquier otra cosa que sirva para ridiculizarla públicamente. A esta aplicación concreta se le suele denominar comúnmente como “deepfake” (ultrafalso). Para evitar que esta herramienta pueda ser usada con fines perversos, es necesario que haya herramientas que permitan distinguir estas imitaciones del contenido verídico.

COGNICIÓN

En contraposición con las aplicaciones del campo de la percepción, se puede determinar un segundo grupo de aplicaciones de IA que están teniendo un desarrollo importante en los últimos años, y son aquellas que operan en el campo de la cognición. En este grupo se encuentran las aplicaciones que requieren el procesamiento de datos y la creación de conocimiento, con herramientas que van desde el tratamiento del lenguaje, la resolución de problemas o las finanzas, por ejemplo.

TRATAMIENTO DEL LENGUAJE

Tal como se expuso al inicio de este documento, durante finales del siglo XX hubo numerosos intentos de crear algoritmos que comprendiesen y reprodujesen el lenguaje humano con la idea de utilizarlo en aplicaciones como la traducción y la transcripción de audios. Sin embargo, a pesar de la gran inversión, los algoritmos apenas eran capaces de reproducir algunas palabras sin coherencia e, incluso así, tardaban mucho más de lo que lo hacía un humano, por lo que se acabó por abandonar el proyecto. Sin embargo, durante los últimos años la tecnología del tratamiento del lenguaje, de la mano de la mejora en el poder de procesamiento y el aumento de datos disponibles, ha vivido una mejora notoria, lo que ha llevado al desarrollo de arquitecturas complejas con la capacidad de aprender y reproducir el lenguaje humano de formas mucho más realistas. A día de hoy, los algoritmos no solo son capaces de comprender a muy alto nivel el lenguaje humano, sino que también existen modelos con la habilidad de crear textos casi indistinguibles de los de un humano como tales como párrafos o historias. Por ejemplo, existen numerosas aplicaciones capaces de continuar párrafos de texto a partir de inteligencia artificial. Para ilustrarlo, se ha decidido a modo de ejemplo introducir el primer párrafo de este documento en un generador de texto online que funciona por redes neuronales llamado InferKit y pedirle que continúe el texto. Para ello se ha traducido primero el texto al inglés, y luego el resultado de vuelta al español, ya que el programa trabaja mejor en este idioma:

Texto introducido como base:

Ha existido, a lo largo de la historia de la humanidad, una cierta fascinación del ser humano por la inteligencia. Es para nosotros un concepto primordial, y para muchos la esencia de la propia vida. Cuando miramos al espacio, nos preguntamos si existirá más allá de nuestro planeta algún signo de vida inteligente. Apreciamos nuestra inteligencia, la estudiamos, la medimos, y desde hace menos de un siglo, intentamos replicarla.

Texto creado por inteligencia artificial:

Pero, ¿somos realmente el pináculo de la inteligencia? Si el cerebro que controla los pensamientos e impulsos de los humanos es solo uno de los muchos cerebros del universo, ¿qué nos hace tan especiales? ¿Cómo nos distinguimos de nuestros homólogos cerebrales?

Como se puede observar, incluso con una muestra muy pequeña, el algoritmo es capaz de crear un texto que no solo tiene sentido gramatical, sino que además continúa correctamente la línea de pensamiento del párrafo original. Aunque quizá alguna frase tenga un sentido difuso, difícilmente podría alguien asegurar que el segundo texto no ha sido escrito por un humano.

Otros usos de esta tecnología incluyen la traducción en tiempo real de texto, el reconocimiento de voz, aplicaciones de asistencia a la escritura o servicios de asistencia al cliente automatizados. Por ejemplo, en Singapur una aseguradora utiliza un sistema de IBM para automatizar sus reclamos. Otro ejemplo es Lumidatum, una firma que ofrece consejos sobre la mejora del servicio al cliente a partir de inteligencia artificial (Brynjolfsson & McAfee, 2017). La capacidad de estos algoritmos ha avanzado hasta el punto de que es posible crear servicios como Codex, capaz de asistir a los programadores corrigiendo errores, traduciendo entre lenguajes e incluso programando su propio código (Metz, 2021).

JUEGOS

El sector de los juegos es uno que siempre ha estado intrínsecamente ligado a la IA ya que siempre ha sido el “campo de pruebas” de sus avances. Numerosos hitos han marcado la historia del desarrollo de la IA: la victoria de ajedrez sobre Kasparov, la batalla entre los módulos Stockfish y AlphaZero, la victoria de AlphaGo sobre el campeón Mundial de Go en 2015, y recientemente los algoritmos de IA han sido capaces incluso de vencer a los humanos en partidas multijugador en

juegos como el Poker o videojuegos como StarCraft II (Michael L. Littman, 2021). Pero la inteligencia artificial no solo ha servido para derrotar a los humanos en videojuegos, sino que también ha colaborado mucho a su mejora. En el ajedrez, por ejemplo, los módulos de ordenador se utilizan mundialmente para aprender y mejorar, analizando las partidas, presentando errores y mostrando posibles alternativas. Además, sirven de gran utilidad a los maestros de ajedrez para analizar en profundidad sus jugadas y descubrir nuevas tácticas y conceptos, lo que ha revolucionado por completo el modo de jugar ajedrez en los últimos años. Por otro lado, algunos videojuegos, como No Man's Sky, son capaces, por medio de la IA, de crear de forma automática el escenario de juego con cada partida, lo que genera paisajes y ecosistemas únicos para cada jugador. Adicionalmente, la tecnología ha permitido mejorar la inteligencia de los NPCs (personajes controlados por ordenador), añadiendo variedad y complejidad a los videojuegos.

FINANZAS

Exceptuando al sector de la tecnología, los servicios financieros son el sector donde más se está invirtiendo en servicios externos de IA, y se espera que durante los próximos años este gasto se incremente (Citi GPS, 2018). Existen dos razones principales por las que este sector está creciendo tanto. Por un lado, los intermediarios financieros generan grandes cantidades de ingresos, lo que supone un incentivo grande a la implementación de mejoras. Por otro lado, debido a la digitalización del sector, existen una gran cantidad de datos digitalizados que facilitan el entrenamiento de modelos mucho más precisos. Esto hace que existan, en todos los aspectos de las finanzas, una potencial oportunidad para la implementación de la IA. De este modo, las aplicaciones de IA en las finanzas transcurren a lo largo de todo el proceso financiero incluyendo el trato al cliente, el manejo de riesgos, la gestión de carteras, la infraestructura de pagos, la seguridad y la supervisión monetaria y legal. Además, la alta presión regulatoria traducida en altos costes de cumplimiento y sanciones ha impulsado a las instituciones a instaurar servicios de IA que aseguren el cumplimiento de la reglamentación. Muchas compañías como Credit Suisse utilizan bots para el cumplimiento de consultas, y otras como JP Morgan los utilizan para revisar préstamos comerciales, lo que equivale a 360.000 horas de trabajo de abogados al año. Esta compañía también utiliza IA para ejecutar operaciones de renta variable a la máxima velocidad (Zetsche, Arner, Buckley, & Tang, 2020). Otras empresas como PayPal utilizan estos algoritmos

para prevenir el lavado de dinero, y decenas de compañías a lo largo del mundo utilizan bots comerciales para decidir qué operaciones llevar a cabo (Brynjolfsson & McAfee, 2017)

OPTIMIZACIÓN

Históricamente muchas compañías han utilizado ordenadores para tratar de mejorar y optimizar la implementación de sus servicios. Los algoritmos de IA amplían esta oportunidad permitiendo su aplicación de manera mucho más flexible. Por ejemplo, Google utilizó su algoritmo DeepMind de machine learning para optimizar la refrigeración de sus centros de datos un 15% extra de lo que su equipo de técnicos fue capaz. De igual modo, Amazon utiliza ML para optimizar su inventario y mejorar su recomendación de productos a consumidores. Otro ejemplo es Infinite Analytics, que desarrolló un sistema para predecir si un usuario haría clic en un anuncio y otro para mejorar el proceso de búsqueda de los clientes, el cual resultó en un incremento del beneficio anual de 125 millones de dólares. Incluso algunas ciudades están empezando a implementar estos algoritmos para optimizar las luces de semáforo atendiendo a las condiciones del tráfico (Brynjolfsson & McAfee, 2017).

ROBÓTICA

El campo de la robótica es uno de los escaparates más distinguidos de la inteligencia artificial. En los últimos cinco años ha habido un progreso considerable en este sector gracias al ML, la mejora en la computación y la disponibilidad de sensores más avanzados. Esto ha permitido que existan ya a día de hoy robots autónomos con alta agilidad y movilidad que pueden emplearse tanto en la industria como domésticamente. El mayor exponente en la actualidad es la compañía de diseño de robótica Boston Dynamics, que ha producido robots como Atlas, un humanoide capaz de saltar, correr, bailar, operar en terrenos complicados e incluso hacer volteretas. Su otro gran proyecto es Spot, un robot cuadrúpedo de tamaño similar a un perro que está siendo usado en zonas de construcción. En China, por otro lado, algunos restaurantes han implementado robots que ayudan a cocinar y servir comida (Michael L. Littman, 2021). Sin ir más lejos, en España una cadena de restaurantes llamada Amazonia Chic opera unos robots que entregan la comida a las mesas. Sin embargo, exceptuando aplicaciones concretas, la industria de la robótica aún se encuentra en desarrollo y la mayoría de aplicaciones están aún siendo diseñadas.

OTRAS APLICACIONES

Existen muchas otras aplicaciones de IA que ya están siendo puestas en marcha. Una de ellas son los sistemas de recomendación, que recomiendan a los usuarios productos basados en los datos recopilados sobre ellos a través de sofisticados sistemas de machine learning. Este sistema es muy popular a día de hoy y está siendo implementado por un gran número de compañías como Youtube, Netflix o Amazon. Estos sistemas van incluso más allá, y pueden servir a las compañías no solo para recomendar productos sino para obtener información importante sobre qué características de sus productos los hacen más deseados. Así por ejemplo Netflix podría crear una película siguiendo los parámetros que sabe que son más buscados por sus usuarios.

Otra herramienta destacable es la implementación de IA en el Internet de las Cosas (IoT), permitiendo recoger grandes cantidades de datos de diferentes dispositivos interconectados para mejorar su utilidad. De este modo podrían surgir ciudades y hogares inteligentes comprendidos por gran cantidad de dispositivos y sensores enlazados y operados por un sistema avanzado de IA.

Finalmente, la IA tiene la oportunidad de facilitar mucho el día a día de la gente. En esta línea han surgido servicios como el pago inteligente, que permite a los usuarios pagar sin necesidad de carteras, dinero físico ni contraseñas, solo a través de la voz y el reconocimiento facial, un servicio que ya ha implementado KFC en algunos restaurantes con un sistema llamado Alipay Pay. Muy similarmente, Amazon está trabajando en la implementación de supermercados inteligentes llamados “Just Walk Out”, donde los clientes pueden entrar, coger productos y llevárselos sin necesidad de pagar, escanearlos o indicarlo de ningún modo, y estos son automáticamente cargados a sus cuentas.

4.3. PRINCIPALES DESAFÍOS Y LIMITACIONES

Tal como ha pasado numerosas veces a lo largo de la historia, la inteligencia artificial tiende a generar unas expectativas demasiado optimistas. Este optimismo, en parte, beneficia al desarrollo de la ciencia puesto que favorece la inversión. Sin embargo, también suele llevar a grandes decepciones. Para sentar bases de lo que es posible y lo que no, muchos intentan dilucidar cuales son los factores que limitan el avance de la inteligencia artificial. En el pasado, se pensó durante años que estas limitaciones estarían relacionadas con todo lo que implicase aquellas cosas que comúnmente son concebidas como “humanas”: emociones, socialización, pintura, música, etc. Sin

embargo, en los últimos años esta idea ha cambiado a medida que aparecen nuevos algoritmos capaces de operar sorprendentemente bien en estos frentes. Desde algoritmos capaces de detectar emociones humanas, otros capaces de simular conversaciones completas casi indistinguibles de las de un humano, a incluso IAs como MuseNet, capaces de componer canciones. Estos hitos han demostrado que no son estas las áreas donde la inteligencia artificial tiene más dificultades. Lo cierto es que, en cualquier campo en el que uno pueda obtener grandes cantidades de datos de entrenamiento las IAs parecen operar especialmente bien. Si uno recopila una cantidad suficientemente grande de canciones, conversaciones, cuadros, fotografías de emociones..., lo más probable es que un algoritmo entrenado con esos datos sea capaz de crear por sí mismo un resultado similar. Por lo tanto, si esas no son las limitaciones de la IA, ¿cuáles son?

GENERALIZACIÓN

Un error común que se comete a la hora de valorar las capacidades de la IA es creer que un alto conocimiento en un área muy específica de trabajo confiere un grado de inteligencia general sobre el ámbito que lo reúne. La realidad es que, al menos en la actualidad, las aplicaciones que utilizan inteligencia artificial son muy limitadas. Si bien muchos sistemas son capaces de ofrecer resultados muy superiores a las de cualquier humano, lo cierto es que su entrenamiento es, por norma general, muy especializado. Un algoritmo de inteligencia artificial puede, por ejemplo, estar entrenado para detectar con un alto grado de precisión expresiones humanas a partir de imágenes de caras y, sin embargo, no saber qué implica cada una de esas emociones, qué las causa ni mucho menos ser capaz de replicarlas por sí mismo. Los algoritmos que existen a día de hoy operan entrenándose para tareas muy específicas por lo que, aunque existen algoritmos capaces de realizar numerosas tareas distintas, aún estamos lejos de alcanzar una máquina con verdadera inteligencia a lo largo de un espectro amplio de campos.

Por otro lado, los algoritmos de IA son entrenados estadísticamente a partir de conjuntos de datos de entrada. Es posible que un algoritmo trabaje bien con todos aquellos datos que se parezcan al conjunto original pero que produzca resultados erróneos si el dato de entrada difiere mucho de este conjunto. No es posible asegurar que el sistema dará buenos resultados en cualquier situación. Es más, mientras que los humanos suelen cometer errores, estos errores suelen contenerse siempre dentro de unos límites de lo que es lógico. Con una máquina, en cambio, no es posible asegurar el grado de error que puede llegar a cometer ante una situación desconocida. Además, no siempre es

posible obtener conjuntos de datos lo suficientemente grandes como para entrenar correctamente al algoritmo. Por ejemplo, en el caso de un algoritmo que asista en operaciones de corazón, es muy probable que no exista una gran cantidad de datos de este tipo. Debido a las dificultades de los algoritmos de ML de trabajar con datos distintos a los de entrenamiento, sería difícil entrenar un sistema capaz de operar bien ante las diversas complicaciones que puedan surgir.

Dar solución a estas limitaciones es uno de los objetivos clave en la investigación actual. Para ello, se está trabajando en formas de hacer que los algoritmos puedan trabajar con muestras de datos más pequeñas. Una de las formas en que esto podría alcanzarse es si se consigue que los modelos de machine learning sean capaces de generalizar o transferir lo que han aprendido de su entrenamiento a otras tareas similares. Ser capaz de extrapolar conocimiento es para muchos una muestra distintiva de inteligencia, ya que requiere la capacidad de adaptarse a nuevas circunstancias unido a un conocimiento amplio del mundo. La capacidad de generalizar, por tanto, no solo ampliaría las posibilidades de la IA, sino que añadiría robustez a los algoritmos, permitiéndoles adaptarse mejor a nuevas situaciones.

PROFUNDIDAD COMPRENSIVA

Algunos expertos, como el investigador Roger Shank, opinan que una de las formas en que se puede evaluar la verdadera inteligencia de un sistema es a través de su habilidad para contar historias. En un experimento, Sam Bowman, un profesor de la Universidad de Nueva York, trató de entrenar a una IA proporcionándole docenas de guiones de obras de teatro para ver si la IA sería capaz de crear la suya propia (Frick, 2017). El resultado, una novela llamada Sunspring, era a primera vista similar a una de las obras originales, con actores, frases y líneas de texto que tenían sentido gramatical, lo cual es sorprendente teniendo en cuenta que la IA no contaba con información previa sobre el lenguaje, sino que basó todo su conocimiento las obras de teatro proporcionadas. Sin embargo, la obra carecía de cualquier tipo de narrativa, y no era más que un compendio de frases sin sentido en su conjunto, mostrando que la IA es muy buena a la hora de imitar, pero que falla aun en alcanzar un grado de comprensión más amplio. (Frick, 2017). Esto demostró la limitación que aún sufren estos algoritmos para generar ideas y conceptos a alto nivel.

Al igual que con las historias, ocurre algo similar con las IA centradas en el procesamiento del lenguaje, como los chatbots, los cuales intentan llevar a cabo conversaciones realistas con humanos. Estos algoritmos son capaces de generar cantidades grandes de texto que son difíciles de distinguir

de las de un humano, con frases que tienen perfecto sentido gramatical y que generalmente contestan correctamente a la frase formulada por el humano. Sin embargo, aun cometen muchos errores que desafían las normas convencionales de las conversaciones humanas y el sentido común, y tienen muchas dificultades con conversaciones largas que requieren un hilo de pensamiento más complejo. En este sentido, aun nadie ha conseguido desarrollar un algoritmo que supere el test de Turing, un test que consiste en que una máquina mantenga una conversación fluida con un humano manteniendo un nivel de profundidad sintáctico cultural y contextual que la haga indistinguible de la de un humano.

PENSAR “FUERA DE LA CAJA”

Los avances recientes han demostrado el gran potencial que tiene la IA en dar solución a un gran número de retos planteados. La gran flexibilidad de estos algoritmos hace creer que podrían ser capaces en el futuro de sustituir a los humanos en todos los ámbitos implementados. Sin embargo, una de sus limitaciones más determinantes radica, no en su habilidad de dar solución a los retos planteados, sino en su capacidad de plantear nuevos retos y objetivos. Tal y como se trató antes, la IA no deja de ser un sistema especializado para una tarea. No se trata realmente de una inteligencia con un conocimiento genérico. Aunque sean capaces de llevar a cabo tareas complejas que los sistemas tradicionales no podían, siguen estando limitadas a objetivos muy específicos. Funcionan bajo unos parámetros y siguiendo unas directrices, por lo que no son capaces de plantear nuevas soluciones fuera del marco de posibilidades con el que han sido entrenadas. Si bien responden bien a los retos planteados, no son buenas a la hora de plantear nuevos retos, de diseñar o inventar nuevas soluciones o de averiguar cuál debe ser el nuevo proyecto u oportunidad a explorar a continuación. Es posible que este sea uno de los horizontes que más tardará en alcanzar la IA, si es que lo consigue, ya que para ello se necesita, por un lado, una comprensión profunda e interrelacionada de la realidad, y por otro, la capacidad de trabajar “fuera de la caja”, de superar los paradigmas con los que ese sistema ha sido creado y entrenado. La IA, de momento, no se encuentra cerca de alcanzar ninguno de estos dos objetivos.

CAJA NEGRA

Es precisamente esa “caja” la causante de uno de los rompecabezas que más preocupa a día de hoy a los expertos en inteligencia artificial y es que, de igual modo que los humanos en ocasiones no somos capaces de explicar las razones que nos llevan a ciertas conclusiones o el modo en el que

hacemos algunas cosas, algo parecido ocurre a menudo con la IA. Las redes neuronales pueden llegar a tener cientos de miles de conexiones, y nodos a partir de los cuales lleva a cabo su trabajo. Cuando se entrena, el peso y las conexiones entre nodos varía, por lo que al final cada una de las conexiones afecta un poco al resultado final. Todo esto hace que resulte muy complicado entender las razones subyacentes a un resultado mostrado por la IA, lo que convierte todo lo que ocurre dentro de una red neuronal en una “caja negra”.

Las consecuencias de la inexplicabilidad de las decisiones de la IA son muchas y suponen un gran reto al desarrollo de la tecnología. Por un lado, se limita la posibilidad de que los humanos aprendan de estos algoritmos ya que, aunque podemos aprender de los resultados, para entenderlo el razonamiento sería necesario estudiar todos y cada una de los nodos y conexiones que componen el algoritmo, algo virtualmente imposible. Además, el desconocimiento produce desconfianza en la validez de sus resultados. Muchas empresas pueden ser reticentes a delegar decisiones a un algoritmo sin entender, al menos ampliamente, las razones bajo las cuales toma sus decisiones. Además, ya ha ocurrido en el pasado que un algoritmo desarrolle inclinaciones racistas o sexistas. Sabiendo esto, es difícil asegurar que los algoritmos nuevos que se desarrollan no cometen el mismo error, ya que es imposible conocer su razonamiento. Por ejemplo, una empresa que quiera implementar un algoritmo que asista en el proceso de contratación no puede conocer si este algoritmo está discriminando por razones de sexo, raza o religión. Esto supone, además, un reto a nivel legislativo ya que dificulta enormemente asegurar la imparcialidad de los algoritmos y exigir responsabilidades ante los fallos que estos cometan. Por último, conocer en más profundidad el razonamiento de un modelo ayudaría a corregir los posibles errores que surjan. Por todas estas razones se está trabajando en el desarrollo de modelos y aplicaciones que permitan visualizar y comprender mejor la toma de decisión de los algoritmos.

CONFIANZA

Son muchos los que visualizan un mundo donde la IA se encuentre presente en todas y cada una de las facetas de la vida humana. Desde coches autónomos, ciudades inteligentes, robots, superordenadores... Sin embargo, para que todo esto sea viable, no solo es necesario que la IA sea capaz de realizar estas tareas, sino que también hace falta que la población acepte estos cambios. La población es uno de los principales stakeholders involucrados, por lo su opinión debe ser tomada en cuenta a la hora de desarrollar la IA. Además, es importante reconocer el poder que la

opinión pública tiene sobre el desarrollo de una tecnología, ya que si el sentimiento en contra es fuerte supondrá una gran barrera en su desarrollo, como ha pasado ya, por ejemplo, con la energía nuclear o la comida genéticamente modificada.

De momento, las opiniones sobre la tecnología están divididas. Un estudio realizado en 2019 en 142 países mostró que las opiniones cambian mucho según la zona geográfica. En Asia del Este el 59% creía que la IA ayudaría a la sociedad y un 11% que la dañaría. En Latinoamérica, en cambio, las opiniones eran opuestas, siendo los porcentajes de 26% y 49%, respectivamente. En Europa las opiniones están muy equilibradas, con un 38% que lo ve beneficioso y un 43% dañino (Neudert, Knuutila, & Howard, 2020). Otro estudio realizado en 2020 mostró que los hombres, los jóvenes y las personas con mayor educación son más propensos a valorar positivamente la inteligencia artificial (Johnson & Tyson, 2020).

Los factores que influyen en la desconfianza en la inteligencia artificial son varios. Para algunos, simplemente el hecho de que la IA no sea humana hace que no sean merecedores de confianza a la hora de tomar decisiones. También influye el hecho de que los algoritmos no pueden ser responsables de las decisiones. No pueden ser despedidos, ni castigados, ni multados, por lo que no sufren las consecuencias de tomar malas decisiones. Además, como se ha tratado antes, de momento su funcionamiento interno es una caja negra, por lo que es complicado entender su razonamiento, por lo que seguir sus decisiones requiere de confianza ciega. Esto, unido al hecho de que su comportamiento es mucho menos predecible que el de un algoritmo convencional, afecta muy negativamente a su confianza. Por otro lado, su entrenamiento está muy basado en los datos de entrada, por lo que son propensos a cometer errores y a desarrollar discriminación por ciertos grupos. Para superar estas barreras, los expertos están trabajando en el desarrollo de algoritmos más robustos y transparentes, que colaboren a mejorar la confianza de la población en la tecnología.

Los humanos no somos perfectos ni mucho menos, cometemos errores e irresponsabilidades. Sin embargo, mientras que estamos muy acostumbrados a que un humano cometa un fallo, cuando lo comete una máquina su repercusión es mucho mayor en nuestra confianza. Esto se debe, por un lado, a que no somos capaces de empatizar con el algoritmo. Solemos esperar que los ordenadores den resultados con total precisión, algo que no esperamos de un humano. Por otro, incluso si la máquina comete muchos menos errores que un humano, su error puede darse en algo que para un

humano puede ser muy obvio, causando un impacto mayor. Un ejemplo de ello es lo que ocurre a menudo con los coches autónomos. Es común ver en las noticias o en redes sociales muestras de accidentes o fallos que cometen estos coches, lo cual genera a la gente inseguridad sobre la fiabilidad de sus algoritmos. Sin embargo, cada día tienen lugar en el mundo miles de accidentes de coche causados por errores humanos. Los coches autónomos no se cansan, ni se emborrachan, ni se despistan, ni se duermen al volante y su capacidad de reacción es mayor. A pesar de ello, las encuestas realizadas muestran que alrededor de un 70% de la gente no se sentiría cómodos conduciendo en un vehículo completamente automatizado, cifra que incluso ha aumentado desde 2016. (Advocates for Highway and Auto Safety, 2018). Esto muestra que, a menudo, la gente se siente más segura con los errores humanos que con los de la máquina, aunque los segundos sean menores. Implementar en la sociedad la confianza en la IA supone un reto para el desarrollo de la tecnología y en su implementación en la vida cotidiana.

RELACIONES HUMANAS

A pesar de los avances conseguidos en los ámbitos del reconocimiento facial y de expresiones y el desarrollo de algoritmos capaces de simular conversaciones casi indistinguibles a las de un humano, hay ciertos aspectos del tipo social que son difíciles de entrenar y por tanto de aprender. Estos son aquellas habilidades sociales más complejas que requieren de un entendimiento a alto nivel de los motores del comportamiento social. Estas habilidades serían, por ejemplo, la capacidad de persuadir, motivar o inspirar. Aquellos puestos de trabajo en los que el factor social juega un papel crucial se encuentran por tanto aún fuera del rango de posibilidades de la IA. Un buen profesor, por ejemplo, además de impartir los conocimientos de la asignatura, debería ser capaz de conectar emocionalmente con sus estudiantes, de transmitirles pasión, curiosidad y sobretodo de inspirarles.

De igual modo, en aquellos puestos de trabajo que operan de cara al público, suele valorarse mucho la capacidad de generar simpatía. Esto es especialmente necesario en trabajos del sector de los cuidados como los asistentes de cuidados especiales, enfermeras, médicos, etc. Existe, en el campo de la medicina, un gran impulso por implementar cada vez más elementos de inteligencia artificial. Sin embargo, a pesar de que estos avances tienen el objetivo de mejorar el servicio tanto para pacientes como para empleados, muchos temen que esto pueda llevar a una deshumanización del

servicio, que resultaría últimamente en un empeoramiento significativo en el cuidado de los pacientes.

Por otro lado, en el ámbito de los negocios, los buenos empresarios no se caracterizan solo por su precisión a la hora de analizar datos o de tomar decisiones, sino que a menudo necesitan contar con la habilidad de motivar a sus empleados, de entender sus necesidades emocionales y, sobretodo, de inspirarles a trabajar por el objetivo propuesto. Además, en las negociaciones, las habilidades humanas suelen jugar un papel crucial. Entre ellas, la confianza que uno sea capaz de generar tanto en su equipo como en la otra parte negociadora pueden ser determinantes en el desempeño de la tarea. Con esto en mente, parece muy improbable que en los próximos años ningún algoritmo pase a sustituir humanos en puestos que requieran motivar las personas, inspirarlas o generar confianza con ellas tales como la gestión de una empresa, de un equipo de personas, en puestos de negociación, etc. En cambio, el potencial en estos puestos parece apuntar más a una coordinación entre máquina y humano, donde la máquina proporcione datos, recomendaciones y análisis que ayuden al humano a desarrollar mejor su tarea.

COORDINACIÓN HUMANO-MÁQUINA

Es por ello que, al menos de momento, la gran mayoría de aplicaciones de machine learning a día de hoy implican una combinación de algoritmo y humano. Es poco común el caso en el que la implementación de IA en un puesto de trabajo implica la sustitución completa de actividad humana. En cambio, lo que suele ocurrir es que se pone a la máquina a cargo de alguna parte muy concreta de un proceso para automatizarla, dejando a los humanos a cargo del ámbito general del proceso. Este enfoque suele ser mucho más productivo para las empresas ya que las máquinas operan mejor cuando están muy especializadas mientras que los humanos son capaces de combinar distintas habilidades y ofrecer una visión más genérica.

Aquellas empresas que, por lo tanto, quieran incorporar los sistemas de IA en sus procesos deben ser conscientes de estas limitaciones. Ser capaz de adaptar una empresa para aprovecharse de los beneficios que proporciona esta tecnología puede suponer una gran ventaja competitiva. En el futuro, se espera que cada vez más empresas adopten procesos que mezclen las aportaciones de los algoritmos de ML en los trabajos humanos. Sin embargo, no todos los puestos de trabajo se verán igualmente afectados por esto. Atendiendo a las limitaciones anteriormente expuestas, los puestos que menos riesgo corren son aquellos puestos que impliquen tanto la creación de nuevas

ideas y proyectos, y la utilización de habilidades sociales complejas. En la intersección de estas dos habilidades se encuentra la labor de líder, alguien capaz de comenzar nuevos proyectos y de comunicar y persuadir a otros para que les sigan, una tarea que será difícil que una máquina lo reemplace. En un mundo con cada vez más presencia de IA, aquellos puestos que sean más complicados de automatizar serán aquellos puestos que más se demanden y que se adaptarán mejor a los cambios que vengan.

LIMITACIONES FÍSICAS

Además de las limitaciones subyacentes a la propia inteligencia artificial y los retos que esto plantea para la tecnología, también existen una serie de limitaciones físicas que dificultan su implementación práctica.

Una de estas limitaciones es el poder de computación de los ordenadores, el cual, aunque ha ido aumentando en un orden similar a lo predicho por la ley de Moore, duplicándose cada 2 años, muestra cada vez más señales de frenarse. Los algoritmos más modernos, con el uso de deep learning, requieren cada vez de ordenadores cada vez más potentes para manejar cada vez cantidades mayores de datos. Por ello, a no ser que se produzca un avance significativo en la tecnología de los ordenadores pronto, o que en cambio se encuentre la forma de entrenar algoritmos precisos, pero más simples, es muy probable que esto acabe suponiendo un factor limitante en el desarrollo de la inteligencia artificial.

Por otro lado, también es necesario que aquellos que quieran llevar a cabo soluciones por medio de IA tengan acceso al software y hardware necesarios para implementarlo. Por suerte para muchas empresas, estos recursos son cada día más accesibles para el público general gracias a empresas como Google, Amazon, Microsoft o Salesforce, que alquilan y venden algoritmos e infraestructura de ML.

Por último, además de los recursos tecnológicos necesarios para implementar sistemas de IA en una empresa, también son necesarios recursos humanos. Del mismo modo que las aplicaciones de la IA crecen cada día, crece el número de expertos en materia necesarios para implementar estas soluciones. A día de hoy existe una demanda de expertos en ciencia de datos y ML superior a la oferta. Existe, por tanto, una necesidad de que más gente, especialmente los jóvenes, se interesen por esta especialización y elijan carreras que cubran esta demanda. Como alternativa para paliar este déficit de oferta, también existen algunos recursos como Udacity o Coursera que ofrecen

cursos online sobre diversas materias tanto a nivel introductorio como más avanzado. Por otro lado, algunas empresas están optando por ofrecer sus propias sesiones de entrenamiento sobre estas tecnologías a sus empleados.

4.4. NORMATIVA VIGENTE Y DILEMAS ÉTICOS

El auge de la inteligencia artificial y la mejora de los algoritmos han hecho que cada vez esté más presente en distintos ámbitos de la sociedad y más integrada en las relaciones humanas. A pesar de ello, el marco regulatorio existente hasta la fecha es muy escaso y poco actualizado. Distintos países alrededor del planeta están comenzando a tomar los primeros pasos para asegurar un desarrollo seguro de la ciencia. Existen aún, sin embargo, muchas incógnitas sobre la dirección y el alcance que tendrán estas medidas y quedan muchos dilemas éticos sin cubrir.

4.4.1. DESAFÍOS REGULATORIOS

DEFINICIÓN DE LA IA

Una de las razones por las que la regulación de la IA es una tarea compleja es que ni siquiera es fácil definir qué abarca el propio concepto de inteligencia artificial. Con el objetivo de regularla, es necesario definir claramente qué es la inteligencia artificial y qué elementos incluye: el código, la arquitectura del sistema operativo, los chips que lo hacen funcionar, la transmisión de datos... Además, no todos los sistemas de IA son iguales ni tienen la misma inteligencia o aplicaciones. Ante esto, existe un reto muy complicado por parte de los organismos reguladores a la hora de definir esta tecnología y comenzar a poner en marcha medidas que pasa por entenderla en profundidad.

LUCHA A CONTRARRELOJ

El desarrollo de leyes y regulaciones requiere habitualmente periodos de tiempo largos. Hasta ahora, esto no ha sido un problema mayor ya que la mayor parte de los avances tecnológicos han seguido también mejoras progresivas y escalonadas, haciendo fácil a los reguladores prever su impacto y anticipar las leyes. Con la IA, en cambio, es difícil anticipar el impacto y los problemas que pueden surgir a partir de su desarrollo. Mientras que los avances tecnológicos habituales están limitados por variables como periodos largos de diseño, construcción, logística, implementación etc., que ralentizan su desarrollo y su impacto, en el caso de la IA esta ralentización está mucho menos presente. Los algoritmos de IA tienen la capacidad de aprender y mejorar en segundos de

computación, lo que hace que legislarla se convierta en una carrera a contrarreloj contra los nuevos avances y aplicaciones que vayan surgiendo.

Si bien existe un factor limitante en la tecnología y el poder de computación, lo que impide a día de hoy que exista una inteligencia artificial muy avanzada es en gran parte la falta de conocimientos que tenemos sobre el área, y no tanto por la tecnología necesaria para ello. Por ello, es difícil predecir si en algún momento ocurrirá un avance brusco en el nivel de inteligencia de las IA que haga que la regulación presente quede atrasada. Algunos predicen que el momento en que la IA alcance definitivamente el nivel de inteligencia de un humano está cerca de llegar. Ray Kurzweil, científico e inventor estadounidense famoso por sus predicciones acertadas en el pasado, cree que esto ocurrirá en 2029, momento en el cual la IA será capaz de superar el test de Turing y por tanto ser considerada similar en inteligencia a los humanos. Para 2045, Kurzweil cree que los humanos habremos alcanzado la “singularidad”, haciendo referencia a que nos habremos fusionado con la IA para multiplicar nuestra inteligencia (Barfield & Pagallo, 2018). Si esas predicciones se cumplen, será necesario que los países comiencen ya a prepararse y a trabajar en las medidas del futuro.

RESPONSABILIDAD JURÍDICA

En el derecho, es fundamental saber asignar la responsabilidad de una persona sobre un acto para poder determinar la correspondiente condena. Con el avance de la IA comienzan a surgir algunos algoritmos a los que se les asigna la capacidad de tomar decisiones, o incluso otros capaces de programar su propio código a través del aprendizaje. Ante esto, surgen dudas no solo judiciales sino también éticas sobre cómo debe asignarse la responsabilidad en caso de que el autor de los hechos sea un algoritmo de IA. En el caso de un robot convencional automatizado, que opera bajo un conjunto de ordenes específicas, como un brazo automatizado, si ocurre algún accidente, la responsabilidad de sus acciones puede asignarse estudiando si el robot tenía piezas defectuosas, si cumplía con las normas de seguridad o si su programación fue negligente. En cambio, en el caso de un algoritmo de IA, este proceso se hace más complicado ya que, a través del entrenamiento, el sistema es capaz de realizar acciones que no se encuentren dentro de lo inicialmente programado. Por ejemplo, en octubre de 2020 Nabla, una compañía de atención médica, estaba probando un servicio de atención al cliente basado en el generador de texto por IA GPT-3 cuando en un momento dado el algoritmo le dijo a un paciente que debía suicidarse (Daws, 2020). Casos como

este son solo un ejemplo de los muchos problemas inesperados que pueden surgir en el entrenamiento de estos programas, lo cual hace complicado asignar responsabilidades.

DISCRIMINACIÓN

En la misma línea que el punto anterior, otro de los riesgos que implica el entrenamiento de un modelo de inteligencia artificial es el de discriminación. Si bien los algoritmos de IA no son discriminatorios de base, pueden llegar a serlo según el método y los datos que se usen para entrenarlo. Así, por ejemplo, se dio en 2017 el caso de un bot de IA para chatear llamado Tay que tuvo que ser suspendido porque comenzó a mostrar comportamientos machistas, sexistas y racistas. Este problema no vino a causa de que los programadores lo hubiesen diseñado así, sino que la IA aprendía a través de conversaciones online con cualquier persona que decidiese hablar con el algoritmo, por lo que surgió de forma independiente, probablemente debido a que los usuarios de internet mostraron ese tipo de comportamiento a la máquina. En casos como este, es complicado asignar quién es el responsable de los posibles daños causados. Un caso similar ocurrió en EEUU, donde se quiso implementar una IA diseñada para asistir en la toma de decisiones judiciales, pero que acabó mostrando resultados racistas. Otro ejemplo es el de un algoritmo médico desarrollado por Optum en 2019, y también en un sistema de detección de melanomas que fallaba más en casos de piel oscura (Michael L. Littman, 2021). Esto muestra, una vez más, la necesidad de introducir ciertas regulaciones o garantías sobre los métodos de entrenamiento de los algoritmos para asegurar la objetividad e imparcialidad de los resultados.

Para tratar de solventar estos problemas, tanto de asignación de responsabilidades como de discriminación, existe una presión cada vez mayor por hacer que los algoritmos de IA sean lo más transparentes y comprensibles posible, permitiendo en última medida que pueda darse explicación al razonamiento detrás todas las decisiones que tome el algoritmo. Que los algoritmos dejen de ser cajas negras, creen muchos, colaboraría en gran medida a incrementar la seguridad de estos algoritmos y la confianza. Organismos como la Casa Blanca, el Comité Selecto de inteligencia artificial de la Cámara de los Lores o el Parlamento Europeo han declarado en distintas ocasiones su interés en que eso ocurra (BUITEN, 2019). Adicionalmente, un aumento de la transparencia colaboraría al uso justo y ético de los datos. El Reglamento General de Protección de Datos recoge que, cuando se hacen perfiles sobre personas, los sujetos tienen el derecho de recibir información significativa sobre la lógica involucrada (General Data Protection Regulation, 2016).

4.4.2. ESTADO ACTUAL DE LA LEGISLACIÓN

Con el fin de comenzar a desarrollar las regulaciones que deberán estar presentes en un futuro con cada vez mayor presencia de IA, algunos países ya han puesto en marcha algunos mecanismos para estudiar posibles respuestas y su impacto. Por ejemplo, en Japón llevan desde 2003 estudiando la regulación de la convivencia entre humanos y robots a través de zonas especiales de para el testeo y desarrollo de robótica llamadas Zonas RT (Weng, 2015). En lo que respecta a los coches autónomos, numerosos países llevan habilitando desde hace años zonas para el estudio empírico de esta tecnología. Suecia, por ejemplo, financió el primer proyecto a gran escala para la conducción del vehículo eléctrico, juntando coches autónomos en carreteras públicas con condiciones normales, y del mismo modo países como Alemania o EEUU han habilitado zonas donde realizar estos experimentos (Barfield & Pagallo, 2018). La propia Comisión de Asuntos Jurídicos del Parlamento Europeo remarcó en 2016 la necesidad de hacer pruebas con robots en condiciones de vida reales.

La UE ha estado en los últimos años quedándose atrás en lo que se refiere al desarrollo y regulación de IA. El reporte “10 Imperativos para Europa en la Era de la IA y la Automatización” (McKinsey & Company, 2017) mostró que la inversión exterior en IA en Europa fue de tan solo 3 a 4 miles de millones de dólares en 2016, en comparación con los 8 a 12 de China o los 23 de Norteamérica. Sin embargo, se espera que esta diferencia disminuya, ya que desde 2018 la Comisión Europea en ha remarcado su ambición de convertirse en la región líder en IA. Su objetivo consiste en facilitar los impactos positivos que pueda traer la tecnología mientras que se mitigan sus posibles riesgos. La Comisión Europea también ha remarcado que por el momento existe una brecha grande entre lo que se desea alcanzar con la IA y lo que realmente es capaz de hacer a día de hoy. (Ulnicane, 2022).

Para sus primeros pasos en la regulación de la IA, la UE se ha inspirado en las medidas tomadas por otros países y construyó en 2017 un reporte reuniendo un resumen de las medidas internacionales tomadas en otras zonas del mundo. Además, ha surgido un interés creciente en la creación de un centro de investigación similar al CERN, pero enfocado a la IA, donde expertos en la tecnología puedan discutir los fundamentos y las soluciones de manera abierta (European Commission, 2017). En este mismo año se publicó “Civil Law Rules on Robotics”, el primer documento de regulación en IA de la UE y uno de los primeros mundialmente. En este documento

se enfatiza en la legislación sobre responsabilidades de daños causados por robots, aunque también se tocan otros temas como problemas éticos, seguridad, inversión en investigación, innovación y educación. En esta resolución, el Parlamento Europeo propuso que los robots autónomos más avanzados deberían tener el grado de “persona electrónica”, con los derechos y obligaciones que esto supone, incluida la posibilidad de reclamar derechos de autoría sobre un trabajo, pero también la responsabilidad ante daños a terceros. Esto causó una gran controversia, y fue criticado por un gran número de expertos de ser una medida “sin sentido” y que “sobreevalúa las capacidades reales incluso de los robots más avanzados”, acusándola de tener una visión distorsionada por la ciencia ficción y el sensacionalismo (Robotics Open Letter, 2017).

Siguiendo este intento de colocarse como líderes en IA, la Comisión Europea aboga por que los distintos países miembros trabajen juntos con una estrategia en común, basada en tres pilares principales. En primer lugar, impulsar la capacidad tecnológica e industrial europea, fijando un objetivo de 20.000 millones de euros de inversión al año. En segundo lugar, prepararse para los cambios socioeconómicos traídos por la IA, principalmente la pérdida de trabajos, a través de la modernización en educación y entrenamiento y la anticipación y adaptación de cambios. En tercer lugar, asegurar un marco ético y legal que beneficie a la sociedad en su conjunto y que “ponga el poder de la IA al servicio del desarrollo humano” (European Commission, 2018). Este plan de colaboración conjunta tomó finalmente forma a partir de la Declaración de Cooperación en inteligencia artificial, firmada por los Estados Miembros con el objetivo de trabajar juntos en el impulso de la industria y la tecnología de la IA. (European Commission, 2018). Entre las actividades propuestas en la declaración se incluye la creación de espacios de almacenamiento de datos comunes, el impulso y la retención de talento tecnológico en Europa, la inversión en infraestructura y la intención de desarrollar un marco regulatorio con flexibilidad suficiente como para impulsar el desarrollo.

En lo que respecta a la ética, la Comisión Europea insta a perseguir la “ética por diseño”, que quiere decir que los principios éticos estén imbuidos en los productos desde el principio del diseño de estos. Para asegurar que el desarrollo de la IA sigue la dirección adecuada, la Comisión Europea publicó en 2019 “Directrices éticas para una IA fiable” (Comisión Europea, 2019), un conjunto de pautas desarrolladas por expertos independientes de la industria. En este documento se enfatiza la importancia de colocar la IA como un medio al servicio del ser humano, y valora el papel que

puede jugar esta en la mejora del bien común, el progreso y el desarrollo humano. En el documento se menciona las tres componentes que debe reunir la tecnología para ser fiable: ser lícita, ética y robusta. Para ello, se nombran siete requerimientos esenciales: 1) Acción y supervisión humanas, 2) Solidez técnica y seguridad, 3) Gestión de la privacidad y de los datos, 4) Transparencia, 5) Diversidad, no discriminación y equidad, 6) Bienestar social y ambiental y 7) Rendición de cuentas.

Desde el inicio de su presidencia al frente de la Comisión Europea el 1 de diciembre de 2019, Ursula von der Leyen prometió poner en marcha una legislación para las implicaciones humanas y éticas de la IA. En esta línea, en febrero de 2020 la Comisión Europea publicó el libro blanco “Sobre la inteligencia artificial - un enfoque europeo orientado a la excelencia y la confianza” (Comisión Europea, 2020), donde presenta una serie de opciones de regulación de la tecnología, así como formas seguras de invertir en ella. Para ello, este documento propone la división de las distintas aplicaciones según su riesgo, de forma que pueda ponerse en marcha una regulación proporcional al balance riesgo/beneficio que se plantea. Aquellas aplicaciones que, o bien se den en un sector de alto riesgo como la salud o el transporte, o bien sean propensas a tener consecuencias legales como daños o muertes, entrarían dentro de la categoría de “alto riesgo”. Estas aplicaciones estarían sujetas a una regulación más estricta, y tendrían requerimientos tales como supervisión, robustez o precisión de información. Para el resto de aplicaciones, se propone un etiquetado voluntario que reconozca las medidas de seguridad que un producto cumple.

Con este papel, la Comisión Europea establece una visión clara sobre la inteligencia artificial en Europa. Por un lado, es una visión basada en dos pilares fundamentales: la excelencia, tanto en el diseño como en la aplicación de la tecnología, y la confianza, clave en el desarrollo y la inversión. Por otro lado, esta visión comprende que la IA puede aportar grandes avances en el desarrollo humano, pero teniendo en cuenta siempre las dos caras de la moneda: los riesgos y los beneficios. Por último, la Comisión Europea centra el objeto del desarrollo alrededor del ser humano, remarcando que el objetivo del desarrollo debe ser siempre la mejora del bienestar, asegurando siempre el cumplimiento de los derechos fundamentales y éticos del ser humano.

Tras la publicación de este Libro Blanco, y en línea con la visión planteada por este, la Comisión Europea convocó una junta para dar respuesta a las peticiones de medidas regulatorias realizadas por el Parlamento Europeo. Con esto, el 21 de abril de 2021 publicó la propuesta de “Reglamento

del Parlamento Europeo y del consejo por el que se establecen normas armonizadas en materia de inteligencia artificial (ley de inteligencia artificial) y se modifican determinados actos legislativos de la Unión”. (Comisión Europea, 2021). Además, este documento tiene también el objetivo de armonizar la regulación existente y así promover un funcionamiento uniforme a través de los distintos países miembros. En este documento se define qué es la IA y se prohíben ciertas prácticas perjudiciales. Adicionalmente, al igual que en el Libro Blanco, se propone la clasificación de los sistemas de IA según su riesgo, en sistemas de riesgo inadmisibles, alto, limitado y mínimo, siendo los de alto riesgo aquellos que pueden suponer mayores riesgos tales como sistemas de contratación, categorización de personas físicas, riesgos para la salud o aquellos utilizados por administraciones públicas. Estos sistemas de alto riesgo deberían cumplir una serie de medidas tales como evaluaciones de riesgo, procedimientos de utilización, documentación técnica, transparencia y supervisión humana. Además, se plantean regulaciones también para los proveedores de estos sistemas tales como sistemas de gestión de calidad, bases de datos y evaluaciones de conformidad (KPMG, 2021). Esta propuesta, una vez aprobados todos los trámites del procedimiento legislativo, podría ser aprobada en la segunda mitad de 2022, aunque aún no sería obligatoria hasta pasados los dos años de periodo de transición.

5. FUTURO DE LA IA:

Son muchas las novelas, películas e historias inventadas acerca de un futuro donde una inteligencia artificial ha tomado consciencia de sí misma, volviéndose en contra de los que una vez fueron sus creadores. Este futuro distópico ha calado en nuestra cultura e imaginación, y son muchos los que a día de hoy temen que nos estemos acercando hacia un futuro así. Lo cierto es que, tal como se ha expuesto, a día de hoy la IA está aún lejos de alcanzar nada parecido a una verdadera inteligencia, y mucho menos de contar con sus propios deseos y ambiciones, y los avances que se han alcanzado no muestran signos de que eso vaya a cambiar en el futuro cercano. Aun así, surgen dudas razonables sobre cuál es el futuro de esta tecnología, sobre hacia donde avanzan los progresos en el campo y sobre los posibles problemas que pueden darse en las próximas décadas.

A continuación, se expondrá el futuro cercano de la IA, mostrando los proyectos que se están comenzando a reuniendo las distintas previsiones sobre qué se espera de la IA en el futuro. Finalmente, se analizarán los riesgos que pueden surgir a partir de estos avances y qué medidas podrían llegar a tomar las distintas instituciones gubernamentales con el objetivo de asegurar un uso seguro de la tecnología

5.1. ¿QUÉ PODEMOS ESPERAR DE LA IA EN EL FUTURO?

El potencial de la inteligencia artificial es enorme y son muchos los factores que afectarán a su desarrollo. Los expertos coinciden en que la inversión en los próximos años en IA se incrementará a un ritmo alto. En 2021 el volumen de mercado global de esta tecnología fue de 35.920 millones de dólares, y se espera que para 2028 este valor se haya incrementado hasta los 360.360 millones, con un crecimiento anual del 33.6% (Fortune Business Insights, 2021).

5.1.1. POSIBLES AVANCES TECNOLÓGICOS

Por un lado, en lo referente a la tecnología de la IA, los últimos avances con el uso de conjuntos de datos grandes y el éxito que estos han tenido aplicados con técnicas de deep learning han desplazado progresivamente a métodos de aprendizaje más tradicionales. En el futuro, se espera que la cantidad de datos disponibles crezca y por tanto que se siga dando una proliferación de esta rama de la IA. De igual modo, también se espera un desarrollo rápido de las técnicas de Procesamiento del Lenguaje Natural (PLN), debido a su potencial para realizar estudios de

mercado, entender las necesidades de los clientes, los hábitos de consumo, sus procesos de decisión y las tendencias. Un ejemplo de ello es Yellow.ai, un servicio desarrollado en India en 2021 que pone a disposición de las empresas un bot manejado por IA que “entiende la personalidad, el contexto y el sentimiento del usuario para responder con empatía, incluso con menos datos de entrenamiento.” (Yellow.ai, 2022). Se espera que más servicios como este surjan permitiendo a los negocios ofrecer un servicio más adaptado a sus clientes.

MEJORAS EN COMPUTACIÓN

El crecimiento en el uso de sistemas de IA cada vez más grandes y complejos requerirá de potentes sistemas de computación. Por ello, hay un gran interés en el desarrollo de ordenadores más eficientes capaces de hacer frente a estos retos, y también en la creación de sistemas de computación especializados para realizar este tipo de tareas. Una de las ramas por las que se está trabajando en este objetivo es la computación neuromórfica la cual, a diferencia de los ordenadores normales, funciona imitando el funcionamiento biológico del cerebro humano con el objetivo de mejorar el rendimiento del hardware. Esto se debe a que el cerebro humano es capaz de llevar a cabo tareas muy complejas consumiendo muy poca energía, todo ello sin contar con sistemas de refrigeración. Por ejemplo, el entrenamiento de un modelo de procesamiento del lenguaje natural en un ordenador consume aproximadamente 1000 kWh. Un cerebro humano, en cambio, tardaría 6 años en consumir esa misma energía realizando todas las funciones diarias normales (Marković, 2020). Por ello, la investigación en este campo busca desarrollar nuevos ordenadores que consuman menos energía y sean más eficientes. Al igual que un cerebro humano, estos ordenadores tienen un rendimiento peor a la hora de realizar grandes cálculos, pero son buenos cuando se trata de explorar simultáneamente varias soluciones en paralelo.

Por el momento, estos ordenadores aún no han producido avances sustanciales, y apenas están comenzando a ser viables económicamente, pero se esperan que en el futuro sean útiles a la hora de aplicar redes neuronales y deep learning. No se espera que estos ordenadores reemplacen a los convencionales, sino que sirvan como una alternativa para aquellas tareas que puedan ser desarrolladas mejor en este tipo de hardware. Sin embargo, una de las barreras principales a la que se enfrenta el avance de esta tecnología son las limitaciones tecnológicas y físicas a la hora de intentar imitar el cerebro humano. Son necesarias docenas de transistores para replicar cada neurona, y colocar muchas de ellas juntas, con la tecnología actual, resulta en ordenadores

demasiado grandes y con enormes pérdidas térmicas. Además, el cerebro cuenta con una media de 10,000 sinapsis por neurona, algo imposible de conseguir con la tecnología actual, sobre todo en dos dimensiones.

Por otro lado, también se prevé un crecimiento rápido de los servicios en la nube ofrecidos por proveedores como Azure Cloud Services, Einstein o IBM Cloud, en respuesta a la creciente demanda de estructuras optimizadas y de virtualización. Hasta hace poco, muchas empresas no podían permitirse la implementación de sistemas de inteligencia artificial debido a los altos costes de adquirir el software y hardware necesarios. Sin embargo, la computación en la nube permite el acceso a sistemas avanzados de inteligencia artificial a un número mucho mayor de personas, ya que permite trasladar el almacenamiento de datos y el procesamiento a la nube, sin necesidad de grandes inversiones de dinero ni de poseer de grandes conocimientos en el área de la IA. A medida que cada vez más empresas comiencen a aplicar IA en sus procesos, se espera que la demanda de los servicios en la nube se incremente (EZmarketing, 2020).

SINGULARIDAD

En la actualidad se califica la capacidad de la inteligencia artificial como débil debido a su alta especialización en tareas concretas. Durante muchos años se creyó que la idea de una IA fuerte era tan solo propia de la ciencia ficción, a siglos o más de distancia. Sin embargo, los recientes avances han resultado en el alcance de hitos que muchos expertos pensaban que no podrían conseguirse hasta dentro de varias décadas, haciendo que la idea de una IA superinteligente pase a verse más cercana. Mientras que muchos expertos siguen pensando que esto no es posible o que aún está a siglos de distancia, la mayoría de expertos coincide en que la singularidad, es decir, el momento en que la IA supere la inteligencia humana, ocurrirá durante este siglo. Una encuesta realizada mostró que el 62% de los expertos creen que esto se dará antes de 2100, y un 45% cree que ocurrirá antes de 2060 (Faggella, 2019).

Si esto llega a suceder, lo que ocurra a partir de ese punto es mucho más complicado de predecir, razón por la cual se le denomina singularidad. Potencialmente, una IA con un nivel de inteligencia superior al de un humano podría ser capaz de diseñar por sí mismo otros sistemas de IA o incluso de experimentar automejora recursiva, lo que resultaría en una estampida de inteligencia que dejaría la inteligencia humana atrás. Esta superinteligencia podría ser capaz de acabar con los mayores problemas humanos como la guerra, o la pobreza. Sin embargo, algunos expertos han

mostrado preocupación por las posibles consecuencias negativas que podrían derivar de esto (Tegmark, 2022). Analizando nuestra propia evolución, se puede observar que la razón por la que los humanos controlamos el planeta no es por nuestra velocidad, fuerza ni tamaño, sino gracias a nuestra inteligencia. Por ello, si no conseguimos alinear los intereses de la humanidad con los de la IA o desarrollar un sistema de dominio estricto sobre ella, nada asegura que vayamos a ser capaces de mantener el control.

5.1.2. FUTURAS APLICACIONES

Aunque no es posible conocer con exactitud que aplicaciones deparará al futuro de la IA, es posible hacer una estimación de aquellos sectores donde se espera que tenga un mayor impacto, así como revisar en que objetivos y tecnologías están trabajando las empresas del sector.

SALUD

El sector de la salud es uno de los que más potencial muestra en el futuro cercano. Ya es, a día de hoy uno, de los sectores donde más aplicaciones prácticas se están creando. Un reporte prevé que el mercado global de la IA en la salud crezca desde los 6.900 millones de dólares en 2021 hasta los 67.400 millones en 2027, con una tasa de crecimiento anual de 46.5% (MarketsandMarkets, 2021). Entre los factores que colaboran en este rápido crecimiento destacan el aumento significativo de los conjuntos de datos médicos y la necesidad de reducir los costes médicos y la eficiencia del servicio. Además, la situación mundial creada por el COVID-19 sirvió como oportunidad para demostrar el potencial de la IA en el sector sanitario, con la implementación de distintas aplicaciones que ayudaron a detectar la enfermedad, a tratar pacientes y a asistir al personal sanitario.

Uno de los avances esperados en el futuro cercano son algoritmos que asistan al personal médico en el diagnóstico de pacientes. Por ejemplo, se espera que las mejoras en los algoritmos ayuden a identificar enfermedades a través de imágenes y a analizar datos médicos de forma mucho más automatizada, liberando a los doctores de una gran carga de trabajo. Los futuros avances en medicina podrían permitir a los médicos contar con algoritmos que les proporcionen en tiempo real asistencia médica evaluando enfermedades y tomando decisiones médicas, todo ello teniendo en cuenta los datos médicos del paciente. También se espera un gran avance en robots que realicen cirugías o asistan a los cirujanos en ellas.

Con el aumento de la esperanza de vida a nivel global y el crecimiento del porcentaje de población envejecida existe una gran oportunidad de mercado en el sector del cuidado a las personas de edad avanzada. En los próximos quince años el número de ancianos en los EEUU crecerá alrededor de un 50% y la Oficina Nacional de Estadísticas Laborales de EEUU proyecta que las ayudas sanitarias domésticas crecerán un 38% en los próximos diez años (Stone, y otros, 2016). Por ello, se prevé un desarrollo de sistemas de monitorización de la salud que ayuden a tratar y prevenir enfermedades. Por ejemplo, la plataforma americana CarePredict ha desarrollado un dispositivo corporal capaz de detectar cambios en los patrones de conducta de las personas mayores que preceden problemas médicos como depresión, malnutrición o desmayos. El dispositivo también emite una alerta en caso de cualquier urgencia médica. En el futuro, es posible que estas herramientas colaboren a la detección temprana de enfermedades a nivel global. Y no solo prevenir enfermedades, sino que el seguimiento del progreso de estas podría llegar a realizarse a través del móvil, con aplicaciones para el seguimiento de síntomas, o asistencia personalizada en casa a través de inteligencia artificial. Por otro lado, se esperan también grandes avances en sistemas que mejoren las capacidades de las personas con dificultades, como herramientas de asistencia de la visión, el oído o la movilidad. Del mismo modo, se prevé una mejora en los robots de asistencia y cuidado personal, incluso el desarrollo de exoesqueletos que mejoren la movilidad de las personas con dificultades. Para las personas con autismo, podrían entrenarse algoritmos avanzados en interacción del lenguaje y comportamiento social.

Estos avances traerán consigo algunas cuestiones sobre las consecuencias negativas del incremento de la presencia de las máquinas en el sector médico. Muchos temen que, con el objetivo de mejorar la calidad del servicio, se incluya una presencia excesiva de máquinas, lo que acabe por sustituir el trato personal. La conexión emocional y el trato personal forman un papel esencial en el cuidado y la salud de la gente, y es por ello este es un factor que no debe dejarse de lado.

COLABORACIÓN HUMANO-MÁQUINA

Como se ha mostrado antes, la IA tiene aún ciertas limitaciones en el tipo de tareas que es capaz de realizar, por ello, existen muchas formas en las que sería posible combinar el trabajo de humanos y máquinas para potenciar mutuamente sus fortalezas.

La primera categoría donde esta combinación puede ser especialmente útil en el futuro es a la hora de sacar conclusiones. La IA puede ser en ocasiones una caja negra, es decir, a través del

entrenamiento es capaz de dar con soluciones correctas, pero no de explicar de qué modo lo hace. Es ahí donde un humano podría ser de sacar analizar y entender en profundidad los conceptos que han llevado a esos resultados, avanzando en el entendimiento en el sector. A través de mejoras en la visualización e interpretación de los resultados que proporciona la IA, es posible que en el futuro esta combinación lleve a grandes avances científicos.

La segunda categoría es la de la toma de decisiones. Mientras que la IA es mejor a la hora de tomar decisiones teniendo en cuenta extensas cantidades de datos, los humanos son mejores tomando decisiones de carácter más genérico y estratégico. Para que esta interacción sea óptima, es necesario que la IA esté diseñada teniendo en cuenta la interacción y las necesidades del usuario, y a su vez, el usuario debe ser consciente de las limitaciones del algoritmo y hasta qué punto puede asistir en la decisión. La IA puede asistir a la hora de analizar grandes cantidades de datos, ya sea resumiéndolos, organizando la información o proporcionando cálculos, predicciones o evaluaciones del riesgo.

Por último, la tercera rama donde la IA tiene gran potencial de avance es como asistente de labores humanas. Por ejemplo, a través de herramientas que proporcionan ayuda en tiempo real tales como traductores automáticos, asistentes en carretera, proporción de información personalizada, monitoreo de sensores, etc.

Sin embargo, para que todas estas aplicaciones puedan aplicarse de forma eficiente es necesario que se trabaje en mejorar la capacidad de las máquinas de colaborar con humanos, aumentando la confianza generada, mejorando la respuesta de los algoritmos a los distintos estados emocionales y facilitando la interacción y el intercambio de información.

TRANSPORTE

De entre los distintos sectores que se verán más afectados por la IA en el futuro cercano, el sector del transporte será uno en donde más pronto se producirán estos cambios. En concreto, será lugar donde gran parte de la población interactúe por primera vez con la IA. Es por ello que este sector juega un papel clave en la aceptación de este tipo de tecnologías que mezclan interacción de humanos y máquinas inteligentes.

A día de hoy, el sector automovilístico cuenta ya con varios avances en lo que respecta a la automatización. Sin embargo, existen aún muy pocos coches completamente autónomos en

funcionamiento y los que existen aún no presentan cifras lo suficientemente buenas. Además, existe cierta desconfianza en la fiabilidad de la IA por gran parte de la población. En los próximos años, se espera que los coches autónomos se desarrollen, haciéndose más seguros y accesibles a la población, tomando una presencia cada vez mayor, lo cual se espera que colabore a mejorar la opinión popular sobre la tecnología

Los accidentes de tráfico producen cada año aproximadamente 1.3 millones de muertes a escala global, y son la causa de muerte principal entre personas entre 5-29 años. (World Health Organization, 2021). La implementación de la IA en el transporte a través de coches autónomos podría reducir drásticamente el número de accidentes en carretera. Los expertos calculan que los accidentes serían entre un 0 y un 10% de los actuales (Ertel, 2018) . Además, brindaría de múltiples comodidades a la población, tales como conducción automatizada, mayor detección del entorno, sincronización entre vehículos para aumentar la eficiencia y evitar atascos, mayor control del tráfico o aparcamiento automatizado, entre otras.

La automatización del transporte no solo traería mejoras en la conducción, sino que se espera que en el futuro la IA suponga una revolución completa en el sector de la automoción y la urbanización. Potencialmente, los coches autónomos podrían cambiar la conceptualización actual del transporte privado. Por un lado, transformando el interior de los coches, actualmente diseñados en torno a la conducción, para dejar paso a entornos de entretenimiento o trabajo. Por otro lado, podrían producirse cambios en el diseño urbano para dejar lugar a elementos que faciliten la interacción con los coches autónomos, creando un intercambio de información entre coches y entorno. Este intercambio de información facilitaría tanto la conducción eficiente de los usuarios como la planificación urbana, abriendo el camino hacia una ciudad inteligente.

Además, la automatización podría traer un cambio de mentalidad en lo que respecta a la compra de vehículos. En la actualidad, los coches pasan aparcados cerca del 95% del tiempo (Barter, 2013), lo supone un desperdicio enorme de recursos. En el futuro, es posible que no sea necesario ser propietario de un coche para poder utilizar uno siempre que se desee. Los coches autónomos abren la posibilidad a los coches compartidos conducidos de forma autónoma, transformando el coche progresivamente desde un bien hacia un servicio. Un estudio realizado por McKinsey (Gao, Hensley, & Zielke, 2014) determinó que, si se llega a adoptar un uso generalizado de coches autónomos, descendería de forma significativa el número de coches requeridos a escala global, así

como de plazas de garaje y, en su conjunto, del número de recursos consumidos. Por otro lado, un estudio realizado en el Laboratorio Nacional Lawrence Berkeley mostró que podría reducirse hasta un 90% la emisión de gases de efecto invernadero por pasajero debido a la optimización del uso del vehículo causada por los coches inteligentes (Greenblatt & Saxena, 2015). Si esto llega a ocurrir, el coche individual podría llegar a transformarse en una alternativa igual o más sostenible que el transporte público convencional, el cual suele operar con baja ocupación, especialmente en áreas rurales.

EDUCACIÓN

En los últimos años se han visto grandes avances en inteligencia artificial aplicada a la educación. Esta tecnología tiene el potencial de mejorar significativamente la calidad de la enseñanza en el futuro, especialmente gracias a que permitirá un grado de personalización muy alto. En los últimos años y especialmente a partir del comienzo de la pandemia se ha visto un aumento significativo de la popularidad de la enseñanza online, ya sea por elección o necesidad. Los cursos online permiten hacer llegar el conocimiento a cualquier lugar del mundo simplemente a través del internet, y gracias a la inteligencia artificial, la calidad de estos cursos puede verse aumentada implementando herramientas sofisticadas de aprendizaje y personalización de contenido según el alumno. Los cursos online son, además de una herramienta muy útil, una gran fuente de datos que permite adaptar la enseñanza a los resultados del alumno, y que a gran escala colaborará a desarrollar mejores técnicas de aprendizaje. Algunas organizaciones como SOLAR están trabajando en algoritmos que aplican deep learning y procesamiento natural del lenguaje para analizar y desarrollar técnicas de enseñanza adaptadas al comportamiento, la implicación y los resultados de los alumnos (Society for Learning Analytics Research (SoLAR), 2022). Por otro lado, gracias a la IA será posible desarrollar simuladores avanzados de entrenamiento que simulen con gran precisión las condiciones reales, por ejemplo, para prácticas de medicina, de combate, de vuelo...

Los nuevos proyectos en este campo se orientan a mejorar la educación a través del análisis de datos, lo que permitiría entender los errores más comunes realizados por los estudiantes, aportarles feedback en tiempo real, adaptar la enseñanza a cada estudiante y predecir aquellos estudiantes en riesgo de suspender. Además, con una recopilación mayor de datos se pretende comprender los procesos cognitivos detrás del aprendizaje para desarrollar nuevas técnicas de enseñanza.

APLICACIONES MILITARES

Muchos países ya tienen puesta su mira en las posibilidades que ofrece la Inteligencia artificial con fines militares. El tamaño de mercado de la IA en aplicaciones militares en 2020 fue de 6500 millones de dólares y se espera que crezca hasta los 17.630 millones en 2027 (Emergen Research, 2021). Tan solo en EEUU el Pentágono planea invertir 874 millones de dólares en 2022, casi un 50% más que en 2021 (Keller, 2021). En lo referente al entrenamiento de tropas, se espera el desarrollo de simuladores de aprendizaje avanzados de combate, de prácticas de tiro, de operaciones de campo o de vuelo. Por otro lado, gran parte de la investigación militar está destinada a la investigación en armas letales autónomas. La idea de aplicar máquinas para automatizar armas no es nueva, ya que llevan operando en docenas de países desde hace décadas, y son capaces de proteger o atacar objetivos sin intervención humana. Sin embargo, con la ayuda de la IA esta automatización podría llevarse a un nuevo nivel, generando armas de gran peligro, como, por ejemplo, drones autónomos o robots adaptados para matar. Los últimos avances en robots autónomos como los desarrollados por Boston Dynamics hacen que estas aplicaciones no parezcan tan lejanas de conseguirse. Muchos investigadores del campo de la IA y la robótica han expresado sus preocupaciones, temiendo que estos avances puedan suponer un riesgo para la estabilidad internacional (Future of life institute, 2015).

ENTRETENIMIENTO

Durante los últimos años se ha venido produciendo una progresiva digitalización del entretenimiento, transformada por el auge de las redes sociales, el contenido multimedia y los dispositivos electrónicos. Esto ha hecho que aumente considerablemente la cantidad de datos recolectados sobre los consumidores y sus preferencias, lo que abre la puerta a un gran número de oportunidades en el sector de la IA. El éxito de determinadas formas de entretenimiento es complicado de predecir puesto que varía mucho según las modas y los gustos de los consumidores. Sin embargo, las últimas tendencias sugieren que habrá determinadas aplicaciones en el ámbito del entretenimiento en las que la IA podría tener impacto.

Hasta ahora, la mayoría de avances en entretenimiento se han dado impulsados por nuevas aplicaciones de software. Sin embargo, el aumento y abaratamiento de los sensores podría traer consigo innovaciones en el hardware relacionado con el entretenimiento, como por ejemplo las gafas de realidad virtual, simuladores de videojuegos más avanzados o robots de entretenimiento.

Un ejemplo de ello es el Metaverso, desarrollado por Meta, un mundo de realidad virtual que, aunque no se espera que esté listo al menos hasta dentro de 10 años, podría revolucionar el entretenimiento tal como se conceptúa (Snider & Molina, 2022). Además, se espera que las mejoras en el procesamiento del lenguaje ayuden al desarrollo de sistemas de recreo más interactivos, capaces de dialogar y adaptarse a las emociones del consumidor.

En la generación de contenido, el aumento de datos podría permitir el desarrollo de productos de entretenimiento cada vez más personalizados y adaptados a los gustos de sus consumidores. Asimismo, la creación de contenido es algo cada vez más accesible a un coste menor gracias a herramientas de IA que permiten producir contenido de alta calidad con pocos recursos. Existen ya algunos algoritmos capaces de crear piezas musicales, generar videos 3D personalizados a partir de textos o coreografías mediante avatares. Asimismo, la tecnología permitirá disponer de herramientas de doblaje, traducción de contenido y personalización a costes muy bajos y de forma muy accesible.

FINANZAS

La implementación de la IA en la industria de las finanzas es algo que ya lleva ocurriendo desde hace unos años y se espera que siga creciendo. En 2018 se evaluó el mercado de la IA en finanzas en 17,765 millones de dólares, y se espera que alcance los 247,366 millones para 2026 (Allied Market Research, 2022), un crecimiento muy marcado.

Uno de los puntos donde se espera un desarrollo es en la comunicación entre bancos y fintecs, que se beneficiarán de una mayor disponibilidad de datos. También existe un foco importante en sistemas avanzados de seguridad con IA que permitan detectar ataques digitales e identificar actividades sospechosas en tiempo real. En la actualidad el 58% de las aplicaciones de IA en finanzas implican detección de fraude y el 41% procesos financieros y análisis (Statista, 2020). Los algoritmos de IA aplicados en las finanzas para asistir en la toma de decisiones pueden jugar un papel esencial en la industria financiera ya que proporcionan análisis y recomendaciones consistentes, una cualidad muy buscada en este sector (Otieno, 2022). Asimismo, estos algoritmos son de gran utilidad a la hora de realizar estudios de impacto de mercado y aplicaciones de puntuación de crédito, aunque los bancos deben hacer todos los esfuerzos posibles por eliminar posibles discriminaciones en sus algoritmos, además de que deben ser capaces de explicar el proceso de decisión tomado.

En lo que refiere al servicio al cliente, los bancos están migrando de forma progresiva a una atención al cliente cada vez más digital, y se espera que esta tendencia siga la misma línea, con el desarrollo de aplicaciones y servicios digitalizados y el auge de bancos completamente online. Además, gracias a los avances en PLN, se podrán poner a disposición de los clientes robo-consultores financieros que simplifiquen la experiencia del cliente.

A pesar de todo esto, los expertos de la industria no creen que los algoritmos de IA vayan a sustituir nunca a los consultores financieros humanos y defienden que la inteligencia humana es capaz de adaptarse mejor y tomar mejores decisiones. De igual modo, los clientes prefieren tener acceso a información online, pero poder contactar con un consultor humano en caso de tener un problema. Por ello, el modelo que se espera que la mayoría de empresas adopten en el futuro es uno híbrido que combine las ventajas tanto de humanos como de máquinas (Otieno, 2022)

OTROS SECTORES

Muchas ciudades han comenzado ya a implementar sistemas de IA que colaboren en la gestión y el cumplimiento de la ley. Se espera que para 2030, gracias a las mejoras en robots, detección facial, tratamiento de imágenes y la mayor recolección de datos, la presencia de estas tecnologías sea significativa, incluyendo cámaras de seguridad mejoradas, drones de vigilancia y algoritmos para detectar toda clase de crímenes. La mejora de estos servicios podría llegar incluso a la prevención de crímenes a través de la vigilancia policial predictiva. Estos avances, a pesar de aumentar la seguridad ciudadana, levantan preocupaciones acerca de las implicaciones que pueda tener sobre la privacidad y la libertad de la gente a no ser vigilada.

Otro sector que podían experimentar desarrollo en el futuro es el de la industria de la venta al público, proporcionando una experiencia de compra mejorada, incrementando la calidad del producto, y aplicando modelos a la predicción de la demanda y la gestión de inventario.

Por otro lado, también se espera un desarrollo en los sensores conectados a través de IoT, que tendrán un foco mayor en la interacción con humanos. Del mismo modo, se está trabajando en mejorar la relación de los robots con los humanos para hacer generalizable y predecible su comportamiento y mejorar su interacción con objetos, a la vez que mejoran sus capacidades de la mano de los avances en visión computacional y procesamiento del lenguaje

5.2. RIESGOS

5.2.1. RIESGOS A CORTO PLAZO

Desde hace unos pocos años, los avances en inteligencia artificial han llegado al punto en el que sus aplicaciones son capaces de ofrecer beneficios prácticos, lo que ha permitido que su desarrollo salga del puro ámbito científico para transformarse en una tecnología económicamente rentable. Esto crea un círculo vicioso de desarrollo donde las empresas invierten en investigación para desarrollar más tecnología que les proporcione mayores beneficios, y esta investigación a su vez aporta nuevas y mejores oportunidades. Durante los próximos años, las distintas empresas, movidas por el objetivo de incrementar sus beneficios, se lanzarán en una carrera por aplicar la IA a sus procesos. Ante esto, surge el riesgo de que este crecimiento no tenga en cuenta los posibles peligros que puede implicar el uso irresponsable de esta tecnología.

CONFIANZA CIEGA

A medida que la IA se desarrolla, las expectativas sobre la tecnología crecen y se genera optimismo. A su vez, existe cada vez más presión por parte de las empresas en comenzar a aplicar estos algoritmos para participar en los enormes beneficios que pueden producir. La mayoría de la gente no cuenta con formación avanzada en inteligencia artificial, por lo que sus expectativas suelen estar sesgadas por noticias que a menudo ofrecen una visión extremadamente optimista de la IA. Ya ha ocurrido varias veces en la historia de la IA que las expectativas generadas superan las verdaderas posibilidades que ofrecen estos sistemas. Por ello, uno de los riesgos a los que se enfrenta la IA en el futuro es la confianza ciega en esta tecnología, que lleva a muchos a sobrevalorar sus capacidades y a la vez infravalorar sus riesgos. La gente puede llegar a pensar que, al tratarse de una máquina es incapaz de cometer errores, cuando la realidad es que la IA también es capaz de equivocarse y su comportamiento es mucho menos predecible que el de los algoritmos que siguen reglas lógicas. Por tanto, es importante que allí donde se decida aplicar la IA se sea consciente sobre estas limitaciones, y no dejar a la ligera decisiones importantes en manos de algoritmos. Que la IA pueda aplicarse en una tarea no significa de deba hacerse.

DISCRIMINACIÓN

Al igual que con los errores, la gente tiende a sobreestimar la objetividad de los algoritmos, creyendo que los ordenadores son, por definición, objetivos. Sin embargo, si los datos de entrenamiento cuentan de por sí con ciertos prejuicios, lo más probable es que el algoritmo adopte esos mismos prejuicios. Por ello, es necesario ser cautos en el futuro a la hora de implementar IA en la asistencia a la toma de decisiones, especialmente en aquellas que afecten directamente a vidas humanas, tales como decisiones judiciales, concesión de seguros, asistencia a la contratación, puntuación del crédito social, valoración de decisiones médicas, etc., siendo conscientes del riesgo de discriminación inherente que puede existir en los resultados.

DESIGUALDAD

Por otro lado, otro de los riesgos de la implementación de algoritmos de IA a gran escala, es que esto podría agrandar las diferencias sociales en la población., por ejemplo, pronosticando peores soluciones para aquellos grupos históricamente en peor estado y por tanto estadísticamente peor valorados, causando a su vez una retroalimentación de esta discriminación. Además, existe el riesgo de afectar especialmente a de las minorías, ya que aquellos grupos menos representados en los datos de entrenamiento suelen recibir peores predicciones. De igual modo, la IA podría causar diferencias sociales a través del control de los datos. Dado que los modelos de deep learning trabajan con grandes cantidades de información, serán en el futuro aquellas compañías que consigan reunir mayores cantidades de datos las que tengan una ventaja competitiva con sus algoritmos. Ya se está viendo a día de hoy la tendencia que siguen las grandes compañías tecnológicas enfocada a reunir grandes cantidades de datos. Existe el riesgo de que se produzca un monopolio de los datos, donde unas pocas compañías reúnan la mayor cantidad de estos, limitando la competitividad del resto de empresas. El control de los datos podría ser en el futuro uno de los motores principales del crecimiento económico.

LIMITACIONES TECNOLÓGICAS

Desde los primeros pasos en la creación de algoritmos de IA hasta la fecha se han producido avances continuados en el campo tanto a nivel técnico como teórico. Se ha incrementado notablemente el conocimiento sobre esta ciencia, y este avance ha ido acompañado de grandes mejoras tecnológicas que han permitido ampliar los límites de la IA. Sin embargo, aunque hasta

ahora la ley de Moore se ha cumplido con gran precisión, todo crecimiento exponencial alcanza su límite en algún momento, y es posible que ese momento se de en el futuro cercano. A pesar de los avances, es difícil que el poder de procesamiento continúe duplicándose ya que estamos cerca de alcanzar límites tanto de costes de desarrollo como relacionados con la propia física cuántica, debido al tamaño tan minúsculo de los transistores. A no ser que se produzca algún descubrimiento tecnológico revolucionario, es posible que en las próximas décadas esta ley deje de cumplirse. Esto afecta especialmente al desarrollo de la IA ya que los algoritmos de deep learning requieren gran capacidad de procesamiento. Por el momento, existe al menos cierta seguridad en que esta ley seguirá alcanzándose hasta 2025, y algunas compañías como Intel se aventuran a pronosticar que podrán seguir cumpliéndola incluso pasado ese año (HPC wire, 2021).

Para combatir esa desaceleración, existen algunas alternativas. Por ejemplo, podrían conseguirse mejoras de rendimiento mediante posibles avances en computación neuromórfica o a través de hardware especializado, como el TPU de Google, mencionados anteriormente. Otra opción es aumentar el tamaño del modelo para permitir la entrada de más parámetros, como es el caso del GPT-3, una red neuronal que batió en 2020 el record de tamaño con 175 billones de parámetros (Michael L. Littman, 2021). Si el poder de procesamiento de los ordenadores no aumenta al mismo ritmo que se desarrolla la IA, el tiempo que se tarda en procesar estos datos aumentará y acabará siendo un factor limitante en esta tecnología.

Con todo, es necesario tener en cuenta el impacto ambiental que este procesamiento y almacenamiento de datos tiene. Un reporte de 2015 predice que para 2030 la mitad de la energía del planeta será consumida por tecnologías de computación y comunicación (Andrae & Edler, 2015).

PRIVACIDAD

Uno de las cuestiones que más debate está generando la IA es su efecto en la privacidad de las personas. Por un lado, el deep learning ha proporcionado un nuevo valor a los datos personales, haciendo que las empresas activamente trabajen por recolectar la máxima cantidad de información posible. Mucha gente se muestra contraria a que se recopile tal cantidad de información sobre su presencia en las redes, y exigen regulación en este aspecto para que las plataformas de internet no puedan recopilar información sin su consentimiento. Una medida que está ganado popularidad es el uso de pestañas de incógnito o ciertos navegadores como Brave que evitan que las empresas

obtengan información sobre la actividad online privada. Por otro lado, estos conjuntos enormes de datos suponen un objetivo de gran interés para potenciales hackers que quieran hacerse con esa información. En 2021 ya se han vivido varias filtraciones de datos personales de gran magnitud, por ejemplo, la filtración de 533 millones de números de teléfono de Facebook (Irwin, 2022). La protección de datos será un reto de gran prioridad en el futuro, especialmente en aquellos sectores donde la privacidad de los datos es más relevante, como el sector médico o el financiero, en los cuales se prevé un gran crecimiento de aplicaciones de inteligencia artificial. Irónicamente, la solución a este problema puede que venga de la mano de la propia inteligencia artificial. Un estudio prevé que para 2023 el 40% de la tecnología de protección de privacidad utilizará IA (Gartner, 2020).

Por otro lado, las mejoras en reconocimiento de imagen, el reconocimiento facial y el aumento en el número de sensores levantan preocupaciones sobre la excesiva intrusión en la privacidad personal. Muchos países ya están comenzando a implementar sistemas de seguridad avanzados que implementan reconocimiento facial y otros como China lo llevan más allá utilizándolo como método de control autoritario. Otros lugares como California, New Hampshire, y Oregón han optado en cambio por prohibir el reconocimiento facial (Kerry, 2020)

MANIPULACIÓN Y DESINFORMACIÓN

La IA permite a empresas e individuos manipular a las personas a través de dos frentes. Por un lado, a través de contenido falsificado realista que confunda, haciéndoles creer algo que no es. A este tipo de contenido se le denomina deepfake es cada vez más fácil de producir gracias a la IA, y en unos años cualquier persona podrá crearlo sin mucho esfuerzo. Desde videos audios de voz replicados a partir de la muestra de voz de una persona, o grabaciones de personas haciendo cosas que nunca han hecho. Aunque en muchas ocasiones se usan para fines de entretenimiento o humorísticos, desde su aparición en 2017 también se han usado para fines perversos como la creación de videos porno de ciertas personas, manipulación de noticias a través de videos falsificados, para engañar al electorado haciendo creer que un candidato es capaz de hablar ciertos idiomas, e incluso colaboraron a provocar un golpe de estado fallido en Gabón, haciendo creer que el presidente estaba enfermo (Cahlan, 2020). En el futuro cercano, es probable que este tipo de falsificaciones crezcan en realismo y en número debido a la facilidad de producirlos, lo que supone un reto para la veracidad de la información.

La segunda forma en que la IA puede colaborar a manipular la opinión pública es a través de la personalización de la información en los medios de comunicación. Los algoritmos de machine learning, aplicados a las redes sociales, pueden recopilar información sobre el conjunto de la población y modificar el tipo de contenido recomendado para cambiar la opinión del público sobre ciertos temas. La culpa no es necesariamente de aquellos que gestionan los algoritmos de recomendación, como con los que cuentan YouTube, Facebook o Instagram. En muchas ocasiones, estos algoritmos pueden ser manipulados a través de bots que actúan en redes sociales interactuando con cierto tipo de contenidos para modificar la actuación del algoritmo o directamente creando contenido falso, popularmente conocidas como fake news.

PÉRDIDA DE TRABAJOS

Otro de los problemas que preocupa a algunos sectores de la población es el impacto de la automatización causada por la IA en los puestos de trabajo. La automatización de puestos de trabajo es algo que lleva ocurriendo desde hace décadas y gracias a la IA es posible automatizar y digitalizar no solo procesos simples sino también aquellos de mucha más complejidad. Algunos ejemplos de ello son el desarrollo de coches autónomos que amenaza acabar con los puestos de transporte de mercancías y de personas, o tecnologías de reconocimiento de imágenes capaces de detectar patrones como células cancerígenas mucho mejor que humanos. Si bien estos avances pueden destruir algunos puestos de trabajos, también crean otros relacionados con el manejo y supervisión de esta tecnología. A su vez, aumentan la productividad lo cual puede llevar al crecimiento económico. Por lo tanto, aunque no está claro si la aplicación de estas nuevas tecnologías producirá un aumento o un descenso de los puestos de trabajo, y surgen dudas sobre si es necesario crear regulación sobre este respecto. Algunos proponen, por ejemplo, limitar la automatización en empresas, u obligar a estas a destinar parte de los ahorros en programas de entrenamiento para empleados.

5.2.2. RIESGOS A LARGO PLAZO

La mayoría de los expertos coinciden en que la inteligencia artificial superará a la inteligencia humana a lo largo de este siglo, y muchos que lo hará antes de 2060. Las opiniones acerca de las implicaciones de esto están divididas, ya que son muchos los potenciales beneficios, pero también es grande el número posibles riesgos. A diferencia de otros retos anteriores, no existe ningún precedente similar en la historia de la humanidad donde se haya creado algo con la capacidad de

ser más inteligente que nosotros, por lo que no es posible basarse en desarrollos tecnológicos anteriores como marco de referencia. Por ello, es necesario estudiar las posibles consecuencias negativas para ser capaces de prepararse y tratar de prevenirlas.

A pesar del miedo que pueda surgir por el hecho de crear algo con unas capacidades cognitivas superiores a nosotros, la mayoría de expertos coinciden en que tal inteligencia superinteligente es improbable que desarrolle emociones ni tampoco deseos propios. Por ello, a diferencia que la literatura de ciencia ficción convencional, no se espera que la IA pueda llegar a ser capaz de volverse malvada y suponer por ello una amenaza a la humanidad. Por el contrario, los riesgos que tal inteligencia puede plantear podrían surgir a partir de dos escenarios

USO MALÉVOLO DE LA IA

Si bien la IA no tiene intenciones malvadas, sí que las tienen algunos humanos. Por ello, en un escenario en el que la IA tiene una inteligencia muy avanzada, podría llegar a ser una herramienta de gran poder destructivo si se usa con fines malvados. Este riesgo dependerá de la accesibilidad a este tipo de inteligencia y el poder de destrucción de esta. Se trata de un riesgo que ya existe a día de hoy, como es el caso de los Sistemas de Armas Autónomos Letales (SAAL), pero cuyo riesgo crece a medida que crece la capacidad de la tecnología

USO BENÉVOLO PERO MALA EJECUCIÓN

Al igual que con un genio mágico malvado que concede deseos, pero de una forma que acaba siendo perjudicial, la IA puede llegar a resultar peligrosa si no se especifica bien aquello que se desea. Por ejemplo, un coche autónomo podría valorar que pasar por una calle peatonal atropellando gente es la forma mejor forma de cumplir su objetivo de llegar al destino lo antes posible solo porque es el camino más rápido. Es por ello que una ingeniería ambigua y mal diseñada puede llegar a muy perjudicial, incluso si se usa con objetivos benévolos. Podría darse el caso incluso, al igual que en algunas películas de ciencia ficción, en la que la IA evalúe intentos de desconectarla como impedimentos a la compleción de su objetivo, tratando por tanto de evitarlo.

En resumen, es posible que la humanidad se encuentre inmensa en una carrera entre el poder creciente de la inteligencia artificial y su capacidad para desarrollar modos de gestionarla de forma segura. Ante este reto, es prioritario trabajar en apoyar lo segundo, más que en frenar el desarrollo. En esta línea, se publicó en 2015 una carta firmada por un número grande de expertos y personas

de interés entre las que se encuentran Elon Musk, Steve Wozniak o Stephen Hawking, en la que se insta a investigar en el desarrollo de una inteligencia artificial robusta y beneficiosa, en preparación para los potenciales riesgos de una posible singularidad (Russell, Dewey, & Tegmark, 2015). En esta carta se proponen una serie de prioridades de investigación, tanto a corto como a largo plazo. A corto plazo se insta a investigar en la investigación sobre la optimización del impacto económico de la IA, en leyes y ética y en la búsqueda de una IA robusta mediante ciencia computacional. A largo plazo piden orientar la investigación a una mayor verificación de algoritmos, mayor seguridad y control.

6. CONCLUSIONES

Durante los últimos cinco años, la inteligencia artificial ha experimentado un crecimiento considerable, consiguiendo salir de lo puramente académico para integrarse en un gran número de campos, teniendo un impacto real en la sociedad y la cultura presentes. Las mejoras en los algoritmos, el aumento de la capacidad de computación y el explosivo aumento de los datos disponibles han catapultado el desarrollo de la IA, resultando en avances significativos habían sido perseguidos durante años por los expertos en el campo. El machine learning, y en especial el deep learning, han llevado las posibilidades de la IA a un nuevo nivel, alcanzando sistemas de gran manejo en tareas sofisticadas como el reconocimiento de imagen, el procesamiento del lenguaje natural, la resolución de tareas complejas y el desarrollo de robots muy avanzados. Estos progresos han convertido a esta tecnología en un potencial motor económico, lo que ha llevado a muchas empresas a invertir en el desarrollo e investigación de nuevos proyectos, creando un círculo vicioso que sin duda llevará a la inteligencia artificial a nuevos límites. Aun así, queda todavía un largo camino por recorrer, y el objetivo primordial de conseguir crear una máquina con la capacidad cognitiva de un humano está aún a décadas de distancia.

La capacidad de aprendizaje de estos nuevos sistemas está causando un impacto en la economía mundial, con importantes aplicaciones sobretodo en campos como la salud, las finanzas y el transporte. Es evidente que los crecientes éxitos en el campo están conduciendo a una sociedad mucho más impactada por la IA. El acceso a esta tecnología es cada vez mayor, y también lo son los intereses económicos. Por ello, en los próximos años, movidos estos los intereses económicos, la presión de la competencia y el optimismo generado, es muy posible que muchas empresas se precipiten a la hora de implantar algoritmos de inteligencia artificial. Nos encontramos en un punto de inflexión, ya que será el desarrollo de los próximos años lo que dé forma a la opinión pública, la cual se encuentra aún indecisa en su apoyo a esta nueva corriente. Será la capacidad de los nuevos avances en generar confianza entre la sociedad lo que con gran probabilidad marque el devenir de la ciencia. Por ello es esencial meditar sobre los riesgos y las desventajas que conlleva la aplicación de la IA, y asegurar un desarrollo próspero y seguro, adelantándose a los riesgos que se esconden bajo tanto optimismo.

La discriminación, los sesgos, la invasión de la privacidad o los deepfakes son tan solo algunos de los retos a los que se enfrenta esta nueva tecnología. En este nuevo escenario, no basta con que

sean los desarrolladores, científicos e investigadores los que decidan sobre el nuevo futuro, sino que es evidente que debe crearse un debate amplio que incluya todas las áreas de investigación humana. Es primordial sentar las bases para un correcto desarrollo de la IA, que permitan minimizar los riesgos negativos para la sociedad a la vez que potencial el crecimiento del campo. Esto requiere de soluciones tecnológicas que aseguren el florecimiento de una IA robusta, confiable y transparente, con sistemas que faciliten la detección de errores y que provean a los algoritmos de mayor seguridad y control. También será fundamental el desarrollo de herramientas que permitan entender el razonamiento detrás de cada decisión, colaborando a la confianza, la integridad y la imparcialidad del funcionamiento de los sistemas.

De cara al futuro, será preciso que se tomen una serie de medidas de gran importancia. El papel de los gobiernos durante este periodo de grandes cambios e incertidumbre será esencial. La mayoría ya han reconocido la importancia de esta nueva tecnología, y están dando los primeros pasos en el desarrollo de un marco regulatorio firme. Alrededor del mundo, la tónica predominante entre los distintos organismos gubernamentales gira en torno al desarrollo de medidas que eviten potenciales consecuencias negativas no deseadas. La Unión Europea ha dejado ya clara su intención de establecer un marco regulatorio sólido que permita el desarrollo seguro de la IA en Europa a la vez que fomente la inversión en el campo, que se ha traducido en la publicación de marcos éticos, recomendaciones, y la publicación de una proposición regulatoria que se espera que sea aprobada durante 2022. Sin embargo, los retos regulatorios son complejos, y las instituciones gubernamentales todavía están rezagadas, y hará falta de mucha investigación e inversión para seguir el ritmo a una tecnología que avanza a pasos agigantados.

Nadie sabe qué deparará el futuro de la inteligencia artificial, pero una cosa está clara, debemos hacer todo lo posible por conseguir que el florecimiento de la IA se traduzca en mejoras en la calidad de vida de la sociedad. Alinear ambos intereses debe ser el principal objetivo de todos. El futuro pasa por aprender a obtener lo mejor de las máquinas y combinarlo con lo mejor de los humanos, en una simbiosis que nos ayude a superar aquellos retos que no podemos alcanzar solos.

BIBLIOGRAFÍA

- Cahlan, S. (2020). How misinformation helped spark an attempted coup in Gabon. *Washington Post*.
- Landay, J. (2016). Smartphone speech recognition can write text messages three times faster than human typing. *Stanford News Service*, 650.
- Laricchia, F. (2017). Factors surrounding preference of voice assistants over websites and applications, worldwide, as of 2017. *Statista*.
- Laricchia, F. (2022). Smart Speaker Market Value Worldwide 2014–2025. *Statista*.
- LIFE Magazine. (20 de Noviembre de 1970). Meet Shaky, the first electronic person: The fascinating and fearsome reality of a machine with a mind of its own by Brad Darrach. *LIFE*, 58B-58D.
- Citi GPS. (2018). *THE BANK OF THE FUTURE The ABCs of Digital Disruption in Finance*. Citi GPS: Global Perspectives & Solutions.
- Comisión Europea. (2019). Directrices éticas para una IA fiable.
- Comisión Europea. (19 de Febrero de 2020). Sobre la inteligencia artificial - un enfoque europeo orientado a la excelencia y la confianza. Bruselas.
- Comisión Europea. (21 de abril de 2021). Propuesta de reglamento del Parlamento Europeo y del Consejo por el que se establecen normas armonizadas en materia de inteligencia artificial y se modifican determinados actos legislativos de la Unión. Bruselas.
- Crafts, N. (23 de 9 de 2021). Artificial intelligence as a general-purpose technology: an historical perspective. *Oxford Review of Economic Polic*, 37(3), 521–536.
- Allied Market Research. (2022). *Artificial Intelligence in BFSI Market*. Allied Market Research.
- Advocates for Highway and Auto Safety. (August de 2018). *Public Opinion Polls Show Deep Skepticism About Autonomous Vehicles*. Obtenido de Safety, Advocates for Highway and Auto: <https://saferoads.org/wp-content/uploads/2018/04/AV-Public-Opinion-Polls-4-26-18.pdf>
- Agrawal, A., Gans, J., & Goldfarb, A. (s.f.). *Agrawal, A., Gans, J., & Goldfarb, A.*

- Ahuja, V., & Nair, L. V. (2021). Artificial Intelligence and technology in COVID Era: A narrative review. *Journal of anaesthesiology, clinical pharmacology*, 37(1), 28–34.
- Andrae, A. S., & Edler, T. (Abril de 2015). On Global Electricity Usage of Communication Technology: Trends to 2030. Suecia.
- Anyoha, R. (28 de Agosto de 2017). The History of Artificial Intelligence. *SITN Science in the News*.
- Barfield, W., & Pagallo, U. (2018). *Research handbook on the law of artificial intelligence*. Edward Elgar Publishing.
- Barr, A., & Feigenbaum, E. A. (2014). *The Handbook of Artificial Intelligence: Volume 1 (Vol.1)*. Butterworth-Heinemann.
- Barr, A., & Feigenbaum, E. A. (2014). *The Handbook of Artificial Intelligence: Volume 2 (Vol. 2)*. Butterworth-Heinemann.
- Barter, P. (2013). "Cars are parked 95% of the time." Let's check.
- Brownlee, J. (Julio de 2019). *A Gentle Introduction to Generative Adversarial Networks (GANs)*. Obtenido de Machine Learning Mastery: <https://machinelearningmastery.com/what-are-generative-adversarial-networks-gans/>
- Brynjolfsson, E., & McAfee, A. (2017). Artificial intelligence, for real. *Harvard Business Review*.
- BUITEN, M. C. (Marzo de 2019). Towards Intelligent Regulation of Artificial Intelligence. *European Journal of Risk Regulation*, 10(1), 41 - 59.
- Davenport, T. H., & Ronanki, R. (2018). Artificial intelligence for the real world. *Harvard business review*, 96.
- Daws, R. (18 de Octubre de 2020). Medical chatbot using OpenAI's GPT-3 told a fake patient to kill themselves. *AI News*.
- Dick, S. (2019). *Artificial intelligence*.
- Emergen Research. (2021). *Artificial Intelligence in Military Market By Offering (Software, Hardware and Services), By Platform (Land, Sea, Air, Others), By Application*

- (*Reconnaissance and Intelligence Gathering, Surveillance and Counterinsurgency and Others*), *By Technology, and By Region Forecast to 2028*. Emergen Research.
- Ertel, W. (2018). *Introduction to artificial intelligence*. Springer.
- European Commission. (Abril de 2018). Declaration. Cooperation on Artificial Intelligence. Brussels.
- European Commission. (Noviembre de 2017). AI policy seminar: Towards an EU strategic plan for AI. Digital Transformation Monitor.
- European Commission. (Abril de 2018). *Artificial intelligence for Europe. Communication*.
- EZmarketing. (2020). How We're Expecting Artificial Intelligence to Influence Cloud Computing in 2020. *Monroy IT Services*.
- Faggella, D. (2019). When Will We Reach the Singularity? – A Timeline Consensus from AI Researchers. *Emerj*.
- Fortune Business Insights. (2021). *Artificial Intelligence (AI) Market Size, Share & COVID-19 Impact Analysis, By Component (Hardware, Software, and Services), By Technology (Computer Vision, Machine Learning, Natural Language Processing, and Others), By Deployment (Cloud, On-premises), By Industry (Healthcare, Retail, IT & Telecom, BFSI, Automotive, Advertising & Media, Manufacturing, and Others), and Regional Forecast, 2021-2028*. Fortune Business Insights.
- Frick, W. (Julio de 2017). Why AI can't write this article (yet). *Artificial Intelligence, for real*.
- Future of life institute. (2015). Autonomous weapons: an open letter from ai & robotics researchers . *Future of life*.
- Gao, P., Hensley, R., & Zielke, A. (2014). *A road map to the future for the auto industry*. . McKinsey Quarterly.
- Gartner. (2020). Gartner Says Over 40% of Privacy Compliance Technology Will Rely on Artificial Intelligence in the Next Three Years. *Gartner*.
- General Data Protection Regulation. (2016). Regulation (EU) 2016/679 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC. *OJ L119/1*.

- Greenblatt, J., & Saxena, S. (2015). Autonomous taxis could greatly reduce greenhouse-gas emissions of us light-duty vehicles. *Nature Clim. Change*, 5(9), págs. 860–863.
- Górriz, J. M., Ramírez, J., Ortíz, A., Martínez-Murcia, F. J., Segovia, F., Suckling, J., & Ferrández, J. M. (2017). *Artificial intelligence within the interplay between natural and artificial computation*. The University of Edinburgh, Edinburgh Research Explorer. Advances in data science, trends and applications.
- Haenlein, M., & Kaplan, A. (2019). A Brief History of Artificial Intelligence: On the Past, Present, and Future of Artificial Intelligence. *California Management Review*, 61, 5-14.
- HPC wire. (11 de 12 de 2021). Intel Unveils Plan to ‘Propel Moore’s Law Beyond 2025’.
- IBM. (15 de Julio de 2020). What is Machine Learning? *IBM Cloud Learn Hub*.
- Inga Ulnicane, W. K., Leach, T., Stahl, B. C., & Wanjiku, W.-G. (s.f.). Governance of Artificial Intelligence. Emerging International Trends and Policy Frames. En *The Global Politics of Artificial Intelligence*. 2022: Chapman and Hall/CRC.
- Irwin, L. (2022). Data breaches and cyber attacks in 2021: 5.1 billion breached records. *IT Governance*.
- Johnson, C., & Tyson, A. (Diciembre de 2020). People globally offer mixed views of the impact of artificial intelligence, job automation on society. *Pew Research Center*.
- Jordan, M. I. (2019). Artificial Intelligence—The Revolution Hasn’t Happened Yet. *Harvard Data Science Review*.
- Keller, J. (Junio de 2021). Pentagon eyes \$874 million for AI and machine learning. *Military+Aerospace Electronics*.
- Kerry, C. F. (2020). Protecting privacy in an AI-driven world. *Brooklins*.
- KPMG. (2021). Nuevas obligaciones en camino para los entornos de IA: La Ley de Inteligencia Artificial.
- McCarthy, J. (2007). *What is artificial intelligence?*
- McKinsey & Company. (Septiembre de 2017). 10 Imperatives for Europe in the Age of AI and Automation.

- MarketsandMarkets. (2021). *Artificial Intelligence in Healthcare Market by Offering (Hardware, Software, Services), Technology (Machine Learning, NLP, Context-aware Computing, Computer Vision), Application, End User and Geography - Global Forecast to 2027*. MarketsandMarkets.
- Marković, D. M. (28 de Julio de 2020). Physics for neuromorphic computing. *Nat Rev Phys*, 2, 499–510.
- Metz, C. (9 de Septiembre de 2021). A.I. Can Now Write Its Own Computer Code. That’s Good News for Humans. *The New York Times*.
- Michael L. Littman, I. A.-V. (September de 2021). Gathering Strength, Gathering Storms: The One Hundred Year Study on Artificial Intelligence (AI100) 2021 Study Panel Report. CA, EEUU: Stanford University.
- Neudert, L.-M., Knuutila, A., & Howard, P. N. (7 de Octubre de 2020). Global Attitudes Towards AI, Machine Learning & Automated Decision Making: Implications for Involving Artificial Intelligence in Public Service and Good Governance . *Oxford Comission on AI & Good Governance*.
- Ongsulee, P. (2017). *Artificial Intelligence, Machine Learning and Deep Learning*. Bangkok, Thailand: Faculty of Science, Siam University.
- Otieno, N. (1 de Febrero de 2022). The future role of AI in finance. *World Finance*.
- Pfeifer, R., & Scheier, C. (2001). *Understanding Intelligence*. Cambridge, Massachusetts, London: The MIT Press.
- Radu, R. (2021). Steering the governance of artificial intelligence: national strategies in perspective. *Policy and Society*, 40(2), 178–193.
- Rich, E. (1983). *Artificial Intelligence*. McGraw-Hill.
- Robotics Open Letter. (2017). *OPEN LETTER TO THE EUROPEAN COMMISSION ARTIFICIAL INTELLIGENCE AND ROBOTICS*. Obtenido de Robotics OpenLetter: <http://www.robotics-openletter.eu/>
- Russell, S., Dewey, D., & Tegmark, M. (2015). Research Priorities for Robust and Beneficial Artificial Intelligence. *AI Magazine*.

- Russell, S., Dewey, D., & Tegmark, M. (2015). Research priorities for robust and beneficial artificial intelligence, Open letter. *Future of life*.
- Snider, M., & Molina, B. (2022). *Everyone wants to own the metaverse including Facebook and Microsoft. But what exactly is it?* Obtenido de USA Today: <https://eu.usatoday.com/story/tech/2021/11/10/metaverse-what-is-it-explained-facebook-microsoft-meta-vr/6337635001/>
- Society for Learning Analytics Research (SoLAR). (2022). *About SoLAR: What is Learning Analytics?* . Obtenido de solarsearch: <https://www.solaresearch.org/about/what-is-learning-analytics/>
- Statista. (2020). *AI use cases in financial services industry worldwide as of 2020*. Statista.
- Stone, P., Brooks, R., Brynjolfsson, E., Calo, R., Etzioni, O., Hager, G., & Teller, A. (2016). *Artificial intelligence and life in 2030: the one hundred year study on artificial intelligence*. Stanford University.
- Tegmark, M. (2022). Benefits and risks of Artificial Intelligence. *Future of life*.
- Toosi, A., Bottino, A. G., Saboury, B., Siegel, E., & Rahmim, A. (Octubre de 2021). A brief history of AI: How to prevent another winter(a critical review). *PET Clin*.
- Ulnicane, I. (2022). Artificial intelligence in the European Union: Policy, ethics and regulation. *En The Routledge Handbook of European Integrations* (Vol. 14). Taylor & Francis.
- Weng, Y. S. (2015). Intersection of “Tokku” Special Zone, Robots, and the Law: A Case Study on Legal Impacts to Humanoid Robots. *Int J of Soc Robotics* , 7, 841–857.
- World Health Organization. (Julio de 2021). *Road traffic injuries*. Obtenido de World Health Organization: <https://www.who.int/news-room/fact-sheets/detail/road-traffic-injuries>
- Wu, Y. S. (2016). *Google's neural machine translation system: Bridging the gap between human and machine translation*. arXiv preprint arXiv:1609.08144.
- Yellow.ai. (2022). *Yellow.ai*. Obtenido de <https://yellow.ai>
- Yinglong Li. (2022). Research and Application of Deep Learning in Image Recognition. *Wuhan University*.

Zetsche, D., Arner, D., Buckley, R., & Tang, B. (Febrero de 2020). Artificial Intelligence in Finance: Putting the Human in the Loop. *CFTE Academic Paper Series: Centre for Finance, Technology and Entrepreneurship, 2020/006*.

Zwakman, D. S., Pal, D., & Arpnikanondt, C. (2021). Usability Evaluation of Artificial Intelligence-Based Voice Assistants: The Case of Amazon Alexa. *SN Computer Science* , 2-28.