



Universidad de Valladolid



ESCUELA DE INGENIERÍAS  
INDUSTRIALES

Máster en Ingeniería Industrial

**MÁSTER EN INGENIERÍA INDUSTRIAL**  
**ESCUELA DE INGENIERÍAS INDUSTRIALES**  
**UNIVERSIDAD DE VALLADOLID**

**TRABAJO FIN DE MÁSTER**

**UTILIZACIÓN DEL MACHINE LEARNING**  
**EN LA INDUSTRIA 4.0**

Autor: D. Alberto Maisueche Cuadrado  
Tutor: D. Pedro Sanz Angulo

Valladolid, septiembre, 2019



## RESUMEN Y PALABRAS CLAVE

### Resumen:

Los mayores crecimientos económicos vienen impulsados por grandes innovaciones tecnológicas, como la máquina de vapor, la electricidad o el motor de combustión interna. Las empresas, por su parte, tratan de aprovechar estas revolucionarias tecnologías para crear nuevos modelos de negocio y generar altos beneficios con mínimo coste.

Actualmente, nos encontramos en la cuarta revolución industrial, donde una de las tecnologías más importantes es la inteligencia artificial. En concreto, el aprendizaje automático o *Machine Learning* surge como un subcampo de la inteligencia artificial que da a las computadoras la habilidad de aprender sobre algo para lo que no han sido explícitamente programadas.

En el presente Trabajo Fin de Máster se introducen los fundamentos del *Machine Learning* bajo el contexto de la Industria 4.0, se explican los diferentes tipos de problemas que es capaz de resolver y se exponen casos reales de aplicación en la industria.

### Palabras clave:

Industria 4.0; Inteligencia Artificial; Aprendizaje Automático; Aprendizaje Profundo; Fábrica Inteligente

### Abstract:

The greatest economic growths are driven by major technological innovations, such as the steam engine, electricity or the internal combustion engine. The companies, meanwhile, try to take advantage of these revolutionary technologies to create new business models and generate high profits with minimal cost.

We are currently in the fourth industrial revolution, where one of the most important technologies is the Artificial Intelligence. In particular, the Machine Learning emerges as a subfield of artificial intelligence that gives computers the ability to learn about something for which they have not been explicitly programmed.

The present master's thesis introduces the fundamentals of Machine Learning in the context of Industry 4.0, explains the different types of problems it can solve, and exposes its real applications in the industry.

### Key words:

Industry 4.0; Artificial Intelligence; Machine Learning; Deep Learning; Smart Factory



## **AGRADECIMIENTOS**

*A Pedro, mi tutor de TFM, por sus ánimos y sugerencias a lo largo de todo el proyecto.*

*A mi familia y amigos, por todo el apoyo recibido en mi paso por la universidad.*



# Índice general

|   |    |
|---|----|
| <b>Introducción</b> .....   | 1  |
| Antecedentes.....   | 1  |
| Motivación.....   | 2  |
| Objetivos.....  | 3  |
| Alcance .....   | 3  |
| Estructura de la memoria .....  | 4  |
| <br>  |    |
| <b>1 LA INDUSTRIA 4.0</b> .....   | 7  |
| 1.1. ¿Qué es la Industria 4.0?.....   | 7  |
| 1.2. Las Revoluciones Industriales .....                                      | 8  |
| 1.3. Tecnologías habilitadoras .....  | 10 |
| 1.3.1. <i>Cloud Computing</i> .....   | 10 |
| 1.3.2. <i>Internet Of Things</i> .....  | 11 |
| 1.3.3. <i>Fabricación aditiva</i> .....                                       | 11 |
| 1.3.4. <i>Big Data</i> .....  | 12 |
| 1.3.5. <i>Inteligencia Artificial</i> .....                                   | 13 |
| 1.3.6. <i>Robots autónomos y colaborativos</i> .....                          | 13 |
| 1.3.7. <i>Realidad aumentada y realidad virtual</i> .....                     | 14 |
| 1.3.8. <i>Blockchain</i> .....  | 14 |
| 1.4. Un entorno inteligente .....   | 15 |
| 1.4.1. <i>Smart City</i> .....  | 16 |
| 1.4.2. <i>Smart Grid</i> .....  | 16 |
| 1.4.3. <i>Smart Factory</i> .....   | 17 |
| 1.4.4. <i>Smart Logistics</i> .....   | 17 |
| 1.4.5. <i>Smart Consumer</i> .....  | 18 |
| 1.5. Estado actual de la digitalización .....                                 | 18 |
| 1.5.1. <i>Digitalización en Europa</i> .....                                  | 18 |
| 1.5.2. <i>Digitalización de la industria española</i> .....                   | 21 |
| 1.5.3. <i>Principales barreras en la digitalización de la industria</i> ..... | 23 |
| <br>  |    |
| <b>2 EL MACHINE LEARNING</b> .....  | 27 |
| 2.1. ¿Qué es el Machine Learning? .....                                       | 27 |
| 2.1.1. <i>Conceptos previos</i> .....   | 27 |
| 2.1.2. <i>Historia del Machine Learning</i> .....                             | 30 |
| 2.2. Fundamentos del Machine Learning .....                                   | 31 |
| 2.2.1. <i>El aprendizaje</i> .....  | 31 |
| 2.2.2. <i>El error y los problemas de ajuste</i> .....                        | 34 |
| 2.2.3. <i>Etapas en un proyecto de Machine Learning</i> .....                 | 36 |
| 2.3. Aprendizaje supervisado.....   | 38 |
| 2.3.1. <i>Regresión Lineal</i> .....  | 38 |
| 2.3.2. <i>Regresión Polinomial</i> .....                                      | 40 |
| 2.3.3. <i>Regresión Logística</i> .....                                       | 41 |
| 2.3.4. <i>Support Vector Machine</i> .....                                    | 43 |

|          |  |           |
|----------|--|-----------|
| 2.3.5.   | <i>K-Nearest Neighbors</i> .....                                     | 44        |
| 2.4.     | Aprendizaje no supervisado.....                                      | 44        |
| 2.4.1.   | <i>K-Means</i> .....   | 45        |
| 2.4.2.   | <i>Análisis de Componentes Principales</i> .....                     | 46        |
| 2.4.3.   | <i>Singular Value Decomposition</i> .....                            | 47        |
| 2.4.4.   | <i>Análisis de Componentes Independientes</i> .....                  | 49        |
| 2.5.     | Aprendizaje reforzado .....  | 50        |
| 2.5.1.   | <i>Q-learning</i> .....  | 51        |
| 2.5.2.   | <i>SARSA</i> .....   | 52        |
| 2.6.     | Aprendizaje profundo .....   | 53        |
| 2.6.1.   | <i>Convolutional Neural Networks</i> .....                           | 55        |
| 2.6.2.   | <i>Recurrent Neural Networks</i> .....                               | 57        |
| 2.6.3.   | <i>Generative Adversarial Networks</i> .....                         | 58        |
| <b>3</b> | <b>APLICACIONES DEL MACHINE LEARNING EN LA INDUSTRIA</b> .....       | <b>61</b> |
| 3.1.     | La importancia del Machine Learning en la industria .....            | 61        |
| 3.2.     | Machine Learning aplicado en producción, fabricación y calidad ..... | 63        |
| 3.3.     | Machine Learning aplicado en logística .....                         | 67        |
| 3.4.     | Machine Learning aplicado en mantenimiento.....                      | 74        |
| 3.5.     | Machine Learning aplicado en los negocios.....                       | 76        |
| 3.6.     | Machine Learning aplicado en marketing y gestión de clientes .....   | 78        |
| 3.7.     | Machine Learning aplicado en ergonomía.....                          | 82        |
| 3.8.     | Machine Learning aplicado en seguridad .....                         | 85        |
| 3.9.     | Machine Learning aplicado en recursos humanos .....                  | 87        |
| 3.10.    | Limitaciones del Machine Learning en la industria.....               | 89        |
|          | <b>CONCLUSIONES Y LÍNEAS FUTURAS</b> .....                           | <b>91</b> |
| 1.       | Conclusiones.....  | 91        |
| 2.       | Líneas futuras.....  | 94        |
|          | <b>BIBLIOGRAFÍA</b> .....  | <b>95</b> |







# Introducción

## Antecedentes

La demanda de productos personalizados y la alta competitividad mundial hacen que la digitalización y la informatización de los procesos industriales sea una necesidad. Para sobrevivir, las empresas deben adoptar soluciones tecnológicas e innovadoras en sus modelos de negocio. Esto es lo que se denomina “transformación digital”, y supone un reto en todos los sectores, tanto para pequeñas como para grandes empresas.

Debido al impacto social y económico originado por estas transformaciones en todo el mundo, se habla de una Cuarta Revolución Industrial. La llegada de la Primera Revolución Industrial se produjo en el siglo XVIII con la invención de la máquina de vapor y el desarrollo de la mecánica, dando lugar a sistemas de transporte como el barco o el tren, y a la sustitución de algunas tareas manuales por máquinas.

En el siglo XIX, la electricidad y los motores de combustión interna inauguraron la Segunda Revolución Industrial. Esto permitió, entre otros, la producción en serie de automóviles y nuevos sistemas de comunicación como la radio. Por otro lado, el *Internet*, la automatización y la electrónica fueron los detonantes de la Tercera Revolución Industrial en el siglo XX.

La Cuarta Revolución Industrial, bajo el nombre de Industria 4.0, se centra en la obtención y aprovechamiento de los datos para facilitar la toma de decisiones y conseguir así una ventaja competitiva. Se trata de construir fábricas conectadas y autónomas, es decir, inteligentes. Para ello, han surgido una serie de tecnologías que comunican el mundo físico y el virtual.

Entre las tecnologías habilitadoras de la Industria 4.0 se encuentran el *Internet Of Things*, el *Big Data*, el *Cloud Computing*, la fabricación aditiva, la realidad virtual, la inteligencia artificial, los robots colaborativos, etc. Aunque cada una de ellas contribuye de una forma a la creación de una cadena de valor más potente y digitalizada, la aplicación conjunta de estas tecnologías es lo que verdaderamente marca la diferencia.

Sin embargo, las organizaciones son las encargadas de estudiar e identificar las tecnologías que más beneficios pueden aportar en función de sus necesidades para así aprovechar las oportunidades que brinda la Industria 4.0. Sin este análisis previo, las empresas corren el peligro de morir en el intento. Además del desconocimiento sobre las posibilidades de estas tecnologías, existen otro tipo de barreras que impiden la transformación digital, como la resistencia al cambio y la falta de talento tecnológico.

La Industria 4.0 es objeto de numerosas investigaciones puesto que tendrá un impacto estratégico importante a largo plazo. Estas investigaciones proponen soluciones tecnológicas, retos y nuevas formas de gestionar la implementación de las tecnologías habilitadoras tanto en la industria como fuera de ella.

## Motivación

La motivación del presente documento se debe a la relevancia que está adquiriendo una de las tecnologías habilitadoras en particular: la inteligencia artificial. La inteligencia artificial es una rama de las ciencias de la computación que se encarga de imitar capacidades humanas, como la lógica y el razonamiento, por medio de algoritmos.

Existen subcampos dentro de la inteligencia artificial que se centran en ciertas partes del sistema cognitivo humano. Por ejemplo, la robótica imita los movimientos de los humanos, la visión artificial se encarga de la percepción del entorno y el procesamiento del lenguaje natural se ocupa de la comunicación con los humanos.

Sin embargo, el subcampo más revolucionario en los últimos años es el *Machine Learning* o aprendizaje automático, encargado del reconocimiento de patrones y del aprendizaje. Entre sus múltiples aplicaciones se encuentra la detección de *malware* en programas y documentos, la identificación de células cancerígenas, los sistemas de recomendaciones, la predicción de la demanda, el mantenimiento predictivo, los *chatbot* de atención al cliente, el vehículo autónomo, etc. Es decir, el *Machine Learning* se aplica siempre y cuando haya datos de los que se quieran extraer patrones, relaciones, tendencias y predicciones.

Uno de los términos de moda dentro del *Machine Learning* son las redes neuronales, gracias a las cuales se representan distintos niveles de abstracción por medio de capas, asemejándose al funcionamiento del cerebro humano. Estas redes se engloban bajo el nombre de *Deep Learning* o aprendizaje profundo, siendo esta la rama del *Machine Learning* con más interés en la actualidad debido a su gran utilidad frente a problemas complejos.

Dado que se introducen diversos conceptos relacionados con esta tecnología, surge el problema de no saber las diferencias entre unos y otros. Aunque las personas sepan de la llegada de la inteligencia artificial a sus vidas, en muchas ocasiones se confunden términos y se les atribuyen capacidades que realmente no tienen.

Algunas empresas desean implantar *Machine Learning* en sus negocios sin saber lo que se necesita para ello, ni las consecuencias que esto implica. Incluso sin haber aplicado esta tecnología, otras empresas afirman ante los medios de comunicación que sus procesos y productos están basados en inteligencia artificial y redes neuronales, solo porque saben que se trata de una tendencia tecnológica atractiva.

Conviene conocer los distintos tipos de *Machine Learning* y su funcionamiento, así como sus limitaciones, para poder tener una visión más amplia y clara sobre las oportunidades que ofrece esta tecnología a la sociedad. Al fin y al cabo, el *Machine Learning* está presente en todos los sectores y en todas las áreas de una organización.

El desarrollo del *Machine Learning* promete causar un impacto similar a lo que supuso en su día el *Internet*. Por lo tanto, el interés personal hacia las aplicaciones industriales de esta tecnología es lo que ha impulsado la realización de este Trabajo Fin de Máster.

## Objetivos

Con este documento se pretenden resolver los problemas planteados anteriormente. Por lo tanto, el objetivo principal de este trabajo es dar a conocer los aspectos fundamentales del *Machine Learning* y sus posibilidades en el sector industrial bajo el ecosistema de la Industria 4.0.

Este objetivo general se divide, a su vez, en varios objetivos secundarios. El primero de ellos consiste en comprender la necesidad, las características y las tecnologías habilitadoras de la Industria 4.0. El segundo objetivo es conocer la situación de España frente a otros países en cuanto a la digitalización de las empresas, así como los principales problemas que frenan la transformación digital.

El tercer objetivo es establecer un marco teórico sobre el *Machine Learning*, donde se explique su funcionamiento y los diferentes tipos de algoritmos empleados según el problema considerado. El cuarto y último objetivo es mostrar las posibilidades del *Machine Learning* en la industria por medio de ejemplos reales de aplicación en las distintas áreas de una empresa industrial como la producción, la logística, el mantenimiento, la seguridad o los recursos humanos.

## Alcance

Para alcanzar los objetivos mencionados se ha realizado previamente una búsqueda bibliográfica sobre el estado actual de la industria, la importancia de los datos y el rol de cada una de las tecnologías emergentes. Esta búsqueda incluye noticias, artículos de revistas científicas, informes de consultoras tecnológicas y otros documentos de empresas líderes en esta transformación digital.

A continuación, se ha elaborado una guía sobre el *Machine Learning* que resume los aspectos más importantes de dicha tecnología. En ella, se aclaran los diferentes conceptos que giran en torno al *Machine Learning* así como su relación con otros campos como el Big Data y el *Data Mining*. También se muestran los orígenes y los hitos del *Machine Learning* a lo largo de la historia, y que dan lugar a la tecnología que se conoce actualmente.

Además, esta guía explica los mecanismos de aprendizaje que se utilizan para hacer que las máquinas aprendan a partir de los datos: el aprendizaje supervisado, el aprendizaje no supervisado, el aprendizaje reforzado y el aprendizaje profundo. Para cada uno de ellos, se exponen, de manera didáctica, los algoritmos más utilizados.

Para la realización de esta guía se han consultado libros y artículos de expertos en este campo. También se han seguido cursos *on-line* de *Machine Learning* gracias a plataformas virtuales de formación como Coursera, cuyo fundador Andrew Ng es uno de los mayores profesionales en el sector de la inteligencia artificial.

Finalmente, se han estudiado las aplicaciones industriales del *Machine Learning* tomando como referencia tanto a las grandes empresas tecnológicas como a las startup especializadas en esta tecnología. Se han consultado también numerosas investigaciones científicas y tecnológicas que proponen soluciones innovadoras a través del uso del *Machine Learning*.

## Estructura de la memoria

Este documento se divide en tres capítulos. El primero de ellos se centra en el contexto de la Industria 4.0, explicando lo que es y en qué consiste. Se expone la necesidad de una Cuarta Revolución Industrial y se hace referencia a los orígenes de las Revoluciones Industriales precedentes. Se comenta la utilidad de las tecnologías habilitadoras, entre las que se encuentran el *Cloud Computing*, el *Internet Of Things*, la fabricación aditiva, el *Big Data*, la inteligencia artificial, los robots autónomos, los robots colaborativos, la realidad aumentada, la realidad virtual y el *Blockchain*.

En este primer capítulo también se describe la creación de un entorno conectado como consecuencia de la Industria 4.0, que incluye ciudades, transportes, fábricas y consumidores inteligentes. Por último, se compara el estado de digitalización de la industria española con el del resto de países europeos, y se detectan las principales barreras que impiden esta transformación.

El capítulo dos trata sobre el *Machine Learning*: su historia, el aprendizaje de las máquinas, los principales problemas en la construcción de modelos y los pasos a seguir para implementar una solución basada en *Machine Learning*. Se habla sobre el aprendizaje supervisado y los modelos de clasificación y regresión, sobre el aprendizaje no supervisado y los modelos de *clustering* y reducción de dimensionalidad, sobre el aprendizaje reforzado y sobre el aprendizaje profundo y los tipos de redes neuronales.

El capítulo tres incluye la importancia del *Machine Learning* en la industria así como algunas aplicaciones de esta tecnología en la industria. En concreto, se analiza la utilidad del *Machine Learning* en producción y calidad, en logística, en mantenimiento, en los negocios, en el marketing y la gestión de clientes, en ergonomía, en seguridad y en recursos humanos. Por último, se comentan las limitaciones principales de esta tecnología en la industria actual.

Este documento finaliza con la exposición de las conclusiones obtenidas tras la investigación y las dificultades encontradas durante su realización. Se comentan también algunas posibles líneas futuras que surgen a raíz de este trabajo. Finalmente, se adjunta la bibliografía consultada.







# 1 LA INDUSTRIA 4.0

Actualmente, el sector industrial se caracteriza por una competitividad cada vez más intensa puesto que el cliente demanda productos personalizados de alta calidad en un tiempo de producción muy reducido. La solución para llevar a cabo este cometido con la máxima eficiencia consiste en la unión entre el mundo real y el virtual. Por tanto, es necesaria una transformación digital hacia un nuevo modelo de industria denominado Industria 4.0 (Siemens España, 2019).

En este primer capítulo se introduce el concepto de Industria 4.0, definiendo este término y mostrando la evolución de la industria a partir de la primera revolución industrial. A continuación, se describen las tecnologías habilitadoras que la caracterizan, entre las que se encuentran el Internet de las Cosas o el *Big Data*. Como resultado de la aplicación de estas tecnologías, se presenta el nuevo modelo de fábrica dentro de un ecosistema inteligente. Finalmente, se comenta el estado actual de la industria junto con las barreas que frenan el desarrollo de esta cuarta revolución industrial.

## 1.1. ¿Qué es la Industria 4.0?

El término Industria 4.0 hace referencia a la cuarta revolución industrial que, impulsada por la transformación digital, supone un cambio cualitativo en la forma en que las empresas gestionan la cadena de valor. Concretamente, Industria 4.0 es una iniciativa estratégica introducida por el gobierno alemán en la Feria de Hannover de 2011 cuyo objetivo es la transformación de la industria a través de la digitalización y la explotación del potencial que ofrecen las nuevas tecnologías (Rojko, 2017).

La definición original de Industria 4.0 se encuentra en el documento titulado “*Industrie 4.0 – Smart Manufacturing for the Future by Germany Trade and Invest*” (Germany Trade & Invest, 2014) presentado por la agencia del Gobierno Federal de Alemania para el Comercio Exterior e Inversiones. En él, se dice que este término se refiere a la evolución tecnológica desde sistemas embebidos hasta sistemas ciber-físicos, representando una cuarta revolución industrial liderada por el Internet de las Cosas, los datos y los servicios. Los productos pasan a comunicarse con las máquinas, y los procesos y las decisiones se descentralizan con la interacción entre el mundo real y el virtual (i-SCOOP, 2019).

No es sorprendente que este término provenga de Alemania ya que tradicionalmente ha apoyado el desarrollo y la innovación del sector industrial. Con esta estrategia, Alemania consigue mantener su posición de liderazgo como uno de los países más competitivos en maquinaria, electrónica y automoción. Esta es una de las razones por las que en 2018 Alemania obtuvo el tercer puesto en el ranking de competitividad global, publicado por el Foro Económico Mundial, donde se mide cómo un país usa sus recursos para proveer a sus habitantes de un alto nivel de prosperidad (Diario Expansión, 2019).

No obstante, conceptos similares surgieron prácticamente al mismo tiempo en otros países. General Electric propuso el concepto de Internet Industrial en Norte América en 2012,

cubriendo áreas como la generación y distribución de energía, la atención sanitaria y el transporte, además de la fabricación. Francia se centró en el concepto *Industrie du futur*, combinando industria y ciencia para permitir que tanto las pequeñas como las grandes compañías francesas pudieran beneficiarse de las nuevas tecnologías. *Made in China 2025* fue la iniciativa lanzada por el Ministerio Chino de Industria y Tecnologías de la Información en 2015 para adaptar la Industria 4.0 que había propuesto Alemania a las necesidades de China. Esta iniciativa se centra en la robótica y en el desarrollo sostenible con el objetivo de llegar a ser la superpotencia industrial mundial para 2049 (Rojko, 2017).

La Industria 4.0 surgida en Alemania describe un futuro escenario industrial denominado Plataforma Industria 4.0, caracterizado por la disponibilidad en tiempo real de información relevante procedente de la conexión entre personas, objetos y sistemas. La plataforma busca soluciones a problemas relacionados con los efectos de la Industria 4.0 en el futuro mediante la normalización y el establecimiento de un marco jurídico. Los negocios, la ciencia y la política trabajan de la mano para hacer que la Industria 4.0 sea una realidad, digitalizando los procesos productivos y consiguiendo nuevos productos y servicios (Friedrich-Ebert-Stiftung, 2016).

La Industria 4.0, además de ser una consecuencia natural de la llegada de las nuevas tecnologías y la digitalización, proporciona una nueva forma de incrementar el beneficio industrial de las empresas. Los costes de producción disminuyeron con la llegada de la filosofía *lean* y la producción *Just-In-Time*, así como con la externalización de la producción a países con mano de obra barata. Ahora, la Industria 4.0 ofrece un amplio abanico de soluciones para seguir con este propósito. De hecho, se espera que en los próximos años las empresas que adopten con éxito este enfoque podrán reducir sus costes hasta un 40% (Laguens, 2018).

## 1.2. Las Revoluciones Industriales

La cuarta revolución industrial viene representada por la Industria 4.0 debido a que supone un cambio brusco en el ámbito social, económico y tecnológico de la sociedad. Con el fin de comprender mejor este cambio y lo que conlleva, es necesario remontarse al pasado de la industria. A continuación, se comentan los hechos más relevantes de cada una de las cuatro revoluciones industriales y su impacto en la sociedad. En la *Figura 1. 1* se muestra la evolución de la industria y sus principales tecnologías desde el trabajo manual propio de principios del siglo XVIII hasta la casi completa interconexión de los sistemas automatizados de hoy.

La Primera Revolución Industrial surgió en Gran Bretaña en la segunda mitad del siglo XVIII, cuando la economía deja de lado la agricultura y la artesanía para centrarse en la industria. Aparecen nuevas fuentes de energía y la máquina de vapor se convierte en el elemento más representativo de esta revolución, al permitir el transporte por tren y por barco y, en consecuencia, una mejora del comercio nacional e internacional (Chaves Palacios, 2004).

Las máquinas de hilar y tejer, por su parte, consiguieron aumentar la producción reduciendo tiempos de fabricación y especializando la mano de obra. Sin embargo, la focalización en el aumento de la producción hizo que las condiciones de los trabajadores fueran deplorables, teniendo que trabajar largas jornadas en ambientes húmedos y sucios. Como consecuencia, aparecieron los movimientos obreros para la defensa de los intereses de los trabajadores (Chaves Palacios, 2004).

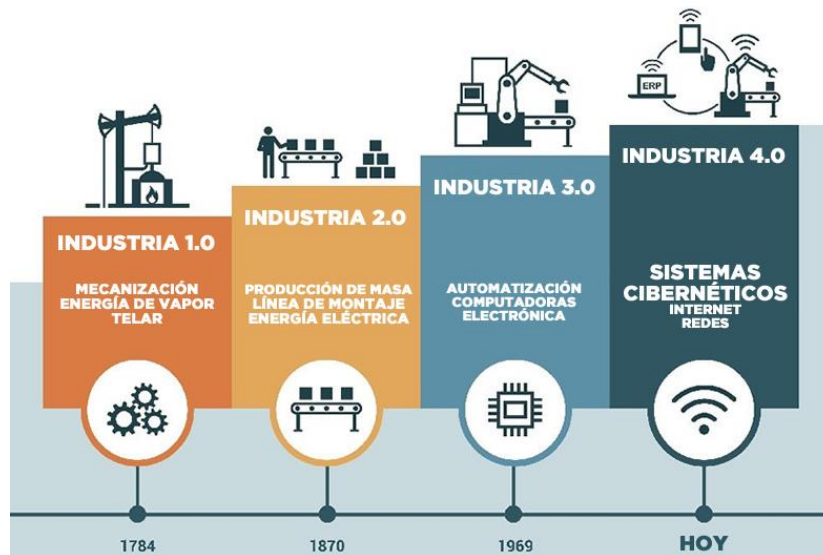


Figura 1. 1. Las revoluciones industriales en el tiempo. Recuperado de <http://www.soy mujerrural.com/mujeres-en-la-ciencia/revolucion-4-0/>

La Segunda Revolución Industrial se produjo en la segunda mitad del siglo XIX y principios del siglo XX, trayendo consigo: nuevas fuentes de energía como el petróleo, el gas o la electricidad; nuevos sistemas de transporte como el avión y el automóvil; y nuevas formas de comunicación como el teléfono y la radio. Se consideró como la primera globalización ya que el impacto de esta revolución llegó a casi toda Europa, Estados Unidos y Japón (Uriarte, 2017).

Todo ello afectó a la manera de organizar la educación, el sistema científico, el trabajo, las empresas, el mercado, etc. En las fábricas se adoptó la producción en serie de Taylor y la cadena de montaje de Ford, naciendo así las primeras multinacionales. El inconveniente de este tipo de fabricación era la poca personalización de los productos y la baja flexibilidad de los procesos (Uriarte, 2017).

La Tercera Revolución Industrial, también llamada Revolución Digital, marcó el comienzo de la Era de la información a mediados del siglo XX. Con la subida del precio del petróleo, se reorientó la tecnología para conseguir una mayor productividad empleando la menor energía y mano de obra posibles. Como resultado, esta revolución industrial se caracterizó por la introducción de la microelectrónica y la automatización, aunque el mayor avance tecnológico se produjo en las comunicaciones con el uso de las computadoras digitales, los teléfonos móviles y la explotación de Internet (Riquelme, 2018).

La biotecnología y los avances médicos permitieron la producción en masa de medicamentos, así como los descubrimientos metalúrgicos permitieron la construcción de naves espaciales y aviones. El empleo de la energía nuclear como nueva fuente de energía trajo el problema de la radiación y su uso militar. En cuanto a los trabajadores, los derechos laborales mejoraron las condiciones de trabajo y la mano de obra pasó a ser mucho más especializada (Riquelme, 2018).

Hoy en día nos encontramos sumergidos en la Cuarta Revolución Industrial, la cual implica un alto grado de digitalización y automatización de las fábricas a través de sistemas inteligentes, sensores y técnicas de procesamiento de datos. La generación masiva de estos datos requiere de un adecuado tratamiento para así obtener resultados fiables y robustos que ayuden a las

empresas en la toma de decisiones. Herramientas como la simulación, la impresión 3D, el Big Data o la Realidad Aumentada son algunas de las denominadas tecnologías habilitadoras, pues sirven para alcanzar este nuevo modelo de industria conectada.

Sin embargo, este cambio no solo se basa en las máquinas y en la tecnología, sino que tiene como pilar fundamental a las personas (LKS, 2019). Por un lado, los trabajadores se centran en las actividades que estrictamente generan valor dejando que las máquinas inteligentes realicen las tareas de no-valor añadido (toma de datos, gestión de incidencias, logística, etc.). Por otro lado, el cliente está presente en todas las etapas de la cadena de valor, definiendo el producto y aumentando así la flexibilidad de la producción.

El concepto de Industria 4.0 va más allá de la fábrica, aunque en su origen se refiriera únicamente a la industria manufacturera. Las mismas tecnologías empleadas en los procesos productivos se adaptan para satisfacer las necesidades de otros sectores. En el ámbito de la salud, los médicos pueden diagnosticar y monitorizar la salud de los pacientes, acceder a su historial y compartirlo con otros profesionales, usar la impresión 3D para órganos y huesos, mejorar la calidad de las imágenes y pruebas diagnósticas, e incluso disponer de máquinas capaces de prevenir ciertas enfermedades con mayor precisión (Siemens España, 2015).

Otro ejemplo es el sector financiero, que se encuentra en una situación inestable debido a las nuevas formas de pago como las criptomonedas o el uso masivo de la compra *on-line*. Diferentes tecnologías se emplean para garantizar la seguridad y la facilidad de las transacciones para el usuario o incluso para recopilar datos sobre los clientes y decidir si concederles un préstamo o no.

### 1.3. Tecnologías habilitadoras

Las tecnologías habilitadoras, o habilitadores digitales, son aquellos elementos y herramientas que posibilitan la digitalización de la industria. No todas estas tecnologías tienen el mismo nivel de desarrollo, pero sí en todas ellas se están realizando numerosas investigaciones para obtener el máximo beneficio en diversos ámbitos. Es importante destacar que no es necesario implementar todas y cada una de ellas en todas las fábricas, sino que lo interesante es saber cuáles son las más apropiadas para cada caso. A continuación, se describen algunas de las tecnologías habilitadoras de la Industria 4.0 y los principales beneficios que aportan.

#### 1.3.1. Cloud Computing

El *Cloud Computing*, también llamado *La Nube*, consiste en el suministro de recursos informáticos bajo demanda a través de Internet. Surge de la necesidad que tienen las empresas de disponer rápidamente de servicios informáticos flexibles y potentes, pagando únicamente lo que se utiliza (IBM, 2019).

Hay tres tipos principales de ofertas de *cloud computing*. La Infraestructura como servicio (IaaS, *Infrastructure-as-a-Service*) proporciona acceso a almacenamiento, servidores y redes; siendo los usuarios quienes utilizan sus propias aplicaciones y plataformas. El segundo tipo es

la Plataforma como servicio (PaaS, *Platform-as-a-Service*), que además de ofrecer infraestructura, proporciona un entorno en la nube donde desarrollar y testear aplicaciones. Por último, el Software como servicio (SaaS, *Software-as-a-Service*) consiste en una suscripción de un software basado en *cloud* de un proveedor, accesible a través de la web (IBM, 2019).

Gracias al Cloud Computing no se necesitan grandes infraestructuras informáticas, sino que se alquila el acceso a un espacio en red por un tiempo determinado. Esto hace que las empresas que prestan estos servicios se encarguen tanto de la seguridad de los datos como del mantenimiento de las redes y equipos. La idea es poder centralizar los datos en una nube fácilmente accesible a través de diferentes dispositivos y compartir información con terceros (ACAN, 2018).

### 1.3.2. Internet Of Things

El Internet de las Cosas (IoT, *Internet Of Things*) representa una red de objetos y dispositivos interconectados que pueden recibir y emitir datos a través de Internet (HBM, 2019). Es un impulsor esencial para la innovación orientada al cliente, la optimización y automatización basada en datos, la transformación digital y los modelos de negocio en todos los sectores. Por ello, se considera que el IoT es una tecnología base para la implementación de la Industria 4.0.

El crecimiento del número de dispositivos conectados es exponencial. En 2015, el número de dispositivos conectados fue de cinco mil millones; en 2017 superó los ocho mil millones y actualmente nos encontramos con una previsión de entre veinte y treinta mil millones para 2020 (i-SCOOP, 2019). Esto nos da una idea del gran número de datos a tratar por dichos dispositivos y de la implicación de las empresas para obtener el mayor beneficio de ellos.

El Internet de las Cosas se ha desarrollado con mayor fuerza en las fábricas bajo el nombre de Internet Industrial de las Cosas (IIoT, *Industrial Internet Of Things*), frente al *Consumer IoT* que se desarrolla fuera de ellas. El IIoT se define como el trabajo conjunto entre máquinas, computadoras y personas que llevan a cabo operaciones industriales usando analítica de datos avanzada para el beneficio empresarial (i-SCOOP, 2019).

El IoT se fundamenta en la tecnología RFID (*Radio Frequency Identification*, identificación por radiofrecuencia), ya que permite asignar a cada producto o dispositivo un código que sirve como identificador único (HBM, 2019). Los objetos que están conectados al IoT suelen llevar sensores capaces de detectar condiciones del mundo real y actuadores con los que pueden ejecutar acciones. En resumen, el IoT focaliza en cada dispositivo la información y la toma de decisiones para después compartir esos datos en la red a través de Internet.

### 1.3.3. Fabricación aditiva

La fabricación aditiva es una técnica empleada para crear objetos físicos a partir de diseños digitales mediante la deposición de capas sucesivas de material. Existen diferentes tecnologías en función del material empleado, de la forma en la que se aplica y de la fuente de energía utilizada (i-SCOOP, 2019).

Esta tecnología se sustenta principalmente en dos máquinas. La primera es el escáner 3D, que es un dispositivo empleado para obtener una copia digital tridimensional de un objeto físico. La segunda es la propia impresora 3D, encargada de leer ese diseño digital y de añadir el material por capas hasta fabricar el producto final (ACAN, 2018).

Comparado con las técnicas de fabricación sustractivas (las que arrancan material sobrante) y conformativas (las que usan moldes y preformas), la fabricación aditiva permite la obtención rápida de piezas más ligeras y de geometrías complejas, a un menor coste y con un mayor grado de flexibilidad (ACAN, 2018).

Por todo ello, su uso se ha propagado rápidamente en sectores como el aeronáutico o el de la automoción, donde el peso de las piezas juega un papel importante. También en el campo de la salud, donde se elaboran prótesis personalizadas según las necesidades de cada paciente. Otros sectores como la moda, el deporte, la alimentación o la arquitectura se ven cada vez más influenciados por esta tecnología y el gran abanico de posibilidades que ofrece (ACAN, 2018).

#### 1.3.4. Big Data

El *Big Data* es el conjunto de datos cuyo tamaño (volumen), complejidad (variabilidad) y velocidad de crecimiento (velocidad) dificultan su captura, gestión, procesamiento o análisis mediante tecnologías y herramientas convencionales, dentro del tiempo necesario para que sean útiles (Power Data, 2019). La línea que marca si un conjunto de datos se puede tratar como *Big Data* no está firmemente definida, aunque la mayoría de los profesionales indican que estos conjuntos de datos van desde los 30 Terabytes ( $10^{12}$  bytes) hasta varios Petabytes ( $10^{15}$  bytes) (Power Data, 2019).

Una gran parte de los datos obtenidos por las tecnologías actuales son de tipo no-estructurados, lo que dificulta su tratamiento. Mediante el empleo de algoritmos y de bases de datos relacionales (por ejemplo, los *Enterprise Resource Planning*), estos datos se combinan con otros datos más estructurados con el fin de facilitar la extracción de conclusiones. La velocidad de generación de datos es abismal, por lo que es crucial que tanto el ritmo de recogida de datos como la velocidad de respuesta obtenida sean igual de elevados. Con esto se pretende disponer de la información correcta en el momento correcto.

Además de las 3V principales (volumen, variabilidad y velocidad), se incluyen otras dos para completar la definición de *Big Data*. Por un lado, la veracidad de los datos, encargada de aportar el grado de confianza de los datos y de eliminar los poco fiables. Por otro lado, el valor, pues si el resultado de este análisis no está alineado con los objetivos, no será útil para la toma de decisiones (Prometeus Global Solutions, 2019).

Existen tantos ejemplos de aplicación de *Big Data* como fuentes de datos de las que interese obtener cierta información. El uso de esta tecnología va desde la predicción de patrones para mejorar la atención de los clientes hasta el campo de la logística, pasando por el mantenimiento predictivo, la detección de fraudes o la publicidad.

### 1.3.5. Inteligencia Artificial

La Inteligencia Artificial es la habilidad que tiene una computadora para presentar las mismas capacidades que un humano a nivel de procesamiento de información, aprendizaje y toma de decisiones. Por lo tanto, su objetivo es hacer frente a problemas complejos imitando la lógica y el razonamiento humano a través de algoritmos. Para aplicar dichos algoritmos se emplea una herramienta llamada *Machine Learning* (Aprendizaje Automático), que gracias a técnicas estadísticas permite que las máquinas aprendan con la experiencia (ACAN, 2018).

El Aprendizaje Automático tiene cuatro variantes principales. En el aprendizaje supervisado, se entrena a la máquina con datos etiquetados a modo de ejemplo, aportando así una guía de aprendizaje. Por otra parte, en el aprendizaje no supervisado los datos no están etiquetados, por lo que la máquina busca similitudes y forma grupos en función de estas. El aprendizaje reforzado se basa en la prueba y error, de forma que la máquina aprende a llevar a cabo una tarea según las consecuencias de sus decisiones pasadas. Por último, el *Deep Learning* (Aprendizaje Profundo) se inspira en el funcionamiento del cerebro humano para extraer capacidades como visión, reconocimiento de patrones o control moto-sensorial mediante el empleo de redes neuronales artificiales (Branco, 2018).

La vida de las personas resulta más fácil gracias a la Inteligencia Artificial. Libera de tareas repetitivas o complejas de forma rápida, mejora la calidad del transporte y la seguridad, posibilita pronósticos de mercado y persigue una buena experiencia de usuario mediante *chatbots* y sistemas de recomendaciones, entre otros (Luna, 2018). A lo largo de este Trabajo Fin de Máster se exponen con mayor detalle tanto los aspectos técnicos de la herramienta de *Machine Learning* como sus aplicaciones en la mayoría de los sectores.

### 1.3.6. Robots autónomos y colaborativos

Los robots autónomos son aquellos que pueden operar con un alto grado de autonomía. Estos robots pueden obtener información del entorno, moverse por él y tomar sus propias decisiones sin intervención humana. Sus funciones se desarrollan normalmente en tareas peligrosas o en operaciones de suministros ya que no aportan valor añadido.

El tipo de robot autónomo más conocido es el AGV (*Automated Guided Vehicle*, vehículo de guiado automático), que consiste en una base con ruedas controlada por software y guiada por medio de bandas magnéticas o láseres. Se encargan de realizar el transporte de materiales entre las diferentes partes de la fábrica, siguiendo una ruta predeterminada, sin interrupciones y sin la intervención humana. Además, son capaces de controlar el tráfico y de dirigirse a las zonas de carga cuando es necesario (Oliva, 2018).

Por otro lado, los robots colaborativos (también llamados *cobots*) son aquellos que trabajan junto a los humanos en el mismo entorno de fabricación sin necesidad de vallados u otros elementos de seguridad. Están diseñados con unas características técnicas que garantizan la seguridad del trabajador, entre las que se encuentran los sensores, los contornos redondeados o el control de fuerza (ACAN, 2018).

La colaboración entre humanos y robots consiste en que los robots mejoran la productividad de los humanos realizando aquellas tareas monótonas, repetitivas y poco ergonómicas. Los trabajadores aportan mayor flexibilidad y destreza, mientras que los robots ofrecen más fuerza y precisión. Algunos de los beneficios industriales que proporcionan los *cobots* son la flexibilidad en la automatización de partes específicas de las líneas de producción, el manejo de piezas grandes y su fácil programación (ACAN, 2018).

### 1.3.7. Realidad aumentada y realidad virtual

La realidad aumentada consiste en combinar contenido físico con contenido digital para crear una realidad mixta en tiempo real (ACAN, 2018). Se obtiene un entorno de visualización interactivo que proporciona información adicional a la hora de realizar una tarea con el fin de mejorar el rendimiento humano. Para poder acceder a dicho entorno es necesario un soporte digital como una Tablet o unas gafas de realidad aumentada (Oliva, 2018).

Por ejemplo, un técnico de mantenimiento puede apuntar a cualquier elemento del que se disponga información y acceder a ella *in-situ*, así como detectar rápidamente cuáles son los componentes que fallan (Oliva, 2018). También se usa para la formación y como guía de trabajo o apoyo para la realización de ciertas operaciones en las líneas de producción. Además del industrial, otros sectores como la medicina, el deporte o la industria de los videojuegos, se están beneficiando de esta tecnología gracias a su facilidad de uso y a la mejora en la experiencia de usuario (Neosentec, 2017).

La realidad virtual, a diferencia de la realidad aumentada, crea un entorno enteramente virtual que transporta al usuario a un lugar tridimensional existente o imaginario. En función de la aplicación, el usuario puede interactuar o no con los elementos que se encuentren en dicho entorno (ACAN, 2018). La posibilidad de crear entornos imaginarios hace que la simulación sea su principal aplicación, sin embargo, un entrenamiento a base de simulaciones puede no dar buenos resultados en la realidad.

En el campo de la enseñanza, esta tecnología trae numerosas ventajas. La realidad virtual atrae la atención de los usuarios y ayuda a mejorar la comprensión de conceptos abstractos o complejos. Además, lo que se aprende mediante experiencias proporciona conocimientos más duraderos que cuando se escuchan de otras personas. Así, se consigue una educación más variada y personalizada (Universia, 2017).

### 1.3.8. Blockchain

El *Blockchain*, o cadena de bloques, es una tecnología que permite intercambiar información digital protegida criptográficamente entre miembros de una misma plataforma. Cada dato queda registrado sin que pueda ser modificado o borrado a menos que todos los integrantes de la red lo permitan. Las transacciones de esta red se copian en bloques y se enlazan mediante códigos criptográficos que conectan cada bloque con el anterior para que sea seguro y difícil de falsificar (Izertis, 2018; ACAN, 2018).



Se caracteriza por tratarse de una red descentralizada, transparente, segura e incorruptible que elimina los intermediarios en los intercambios de valor. Por otra parte, las criptomonedas como el *Bitcoin* son el medio digital de intercambio más utilizado ya que aseguran las transacciones financieras y verifican la transferencia de activos de forma descentralizada sobre un sistema *blockchain* (Izertis, 2018).

El sector más afectado por el momento es el financiero y el asegurador debido a la facilidad para agilizar ciertas operaciones, mientras que en el ámbito industrial se emplea habitualmente para garantizar la trazabilidad total de los productos a lo largo de la cadena de suministro (Izertis, 2018; ACAN, 2018).

## 1.4. Un entorno inteligente

El resultado de emplear conjuntamente las tecnologías explicadas en el apartado anterior en un ámbito industrial es lo que se conoce como *Smart manufacturing* o producción inteligente. Este sistema productivo digitalizado y conectado es el sustento de las fábricas inteligentes o *Smart factories* que caracterizan la Industria 4.0. Las fábricas inteligentes llevan más allá el concepto de automatización, haciendo que las máquinas funcionen prácticamente sin intervención humana. Los sistemas propios de estas fábricas pueden aprender y adaptarse en tiempo real a las situaciones del entorno que les rodea con el objetivo de mejorar el rendimiento total.

Sin embargo, las fábricas inteligentes no son capaces de obtener el máximo beneficio por sí mismas, sino que necesitan estar sumergidas en un entorno inteligente como el que se muestra en la *Figura 1. 2*.

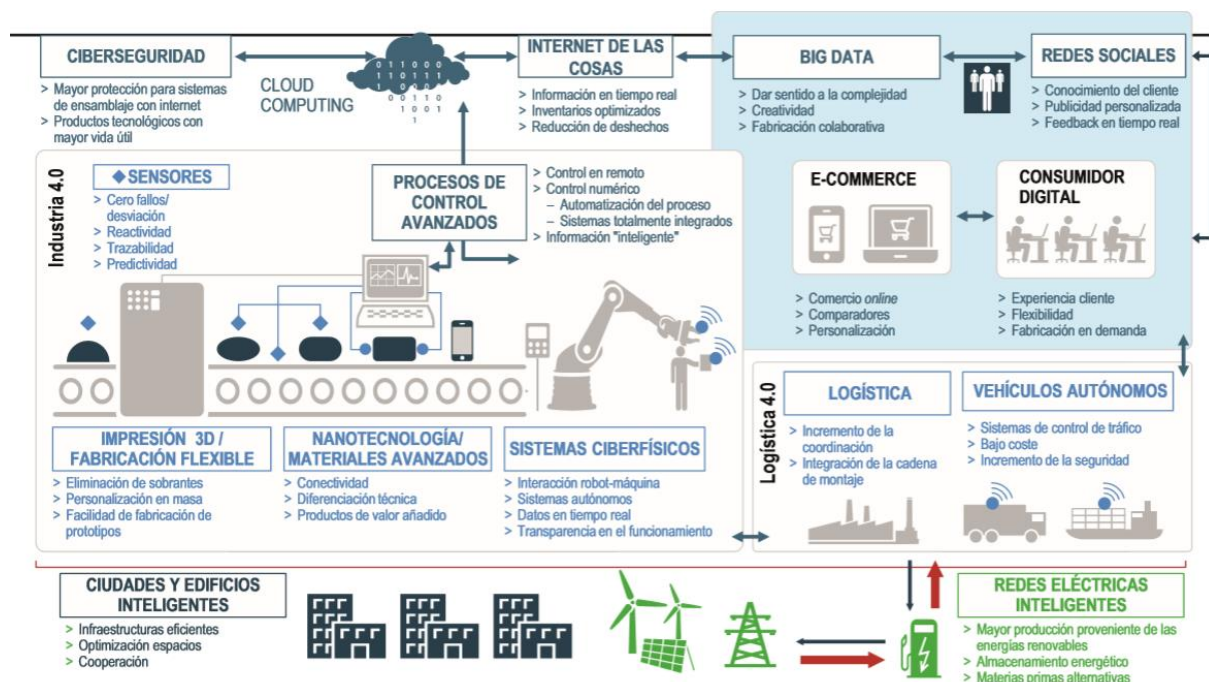


Figura 1. 2. Ecosistema inteligente (Roland Berger, 2016)

Este entorno es posible gracias a la interconexión entre las ciudades, las redes de energía, las fábricas, la logística y los consumidores, a través de las nuevas tecnologías.

#### 1.4.1. Smart City

Una ciudad inteligente es aquella cuyo desarrollo está basado en la sostenibilidad y en la tecnología para conseguir unos servicios públicos eficaces, innovadores, seguros y de alta calidad. Así, un complejo urbano es calificado de inteligente en función de las inversiones realizadas en educación, infraestructuras de energía, tecnologías de comunicación y transporte, que aseguran una calidad de vida elevada, un desarrollo económico-ambiental sostenible y un buen aprovechamiento del tiempo de los ciudadanos (García, 2013).

Tokio, Londres, Singapur, Barcelona, Ámsterdam, Oslo o Nueva York son los principales ejemplos de ciudades inteligentes por llevar a cabo numerosos proyectos *smart*. Desde la mejora del flujo del tráfico hasta la construcción de viviendas equipadas con sistemas de ahorro energético, pasando por alumbrados eficientes mediante energía solar.

Sin embargo, ciudades de menor envergadura como Valladolid también están poniendo en práctica propuestas como la Ruta Ríos de Luz, que consigue impulsar el turismo a través de la innovación aplicada al alumbrado público y ornamental, mejorando también la eficiencia energética hasta un 85%. Los usuarios que realicen esta ruta tienen acceso a una aplicación para smartphones con juegos de realidad aumentada entre sus servicios. Otro proyecto de éxito es la Cúpula del Milenio, propuesta de regeneración urbana que incluye el mayor número de medidas ambientales, consiguiendo el máximo índice de sostenibilidad según el *Green Building Council España* (eSMARTCITY, 2014).

#### 1.4.2. Smart Grid

Las redes eléctricas inteligentes o *Smart grid*, son el resultado del desarrollo conjunto en ingeniería eléctrica y las tecnologías de la información y comunicación, para realizar un uso eficiente de la energía eléctrica bajo un solo sistema de gestión. De esta forma se optimiza la producción y distribución de electricidad con el fin de equilibrar mejor la oferta y la demanda entre productores y consumidores (AEC, 2019).

Los medidores inteligentes son sistemas de medición eficiente del consumo eléctrico que envían información al sistema central para poder ajustar el uso de la red eléctrica. Gracias a estos sistemas, el consumidor dispone de una facturación detallada en tiempo real que ofrece la posibilidad de conectarse y desconectarse de la red si así se desea (Endesa Educa, 2014).

Las energías renovables han hecho que los flujos de energía de red eléctrica sean bidireccionales, pues los consumidores también pueden ser productores a través de la misma red. Esta bidireccionalidad es posible gracias la tecnología digital que controla las necesidades del consumidor, ayudando a ahorrar energía, reducir costes e incrementar la usabilidad y la transparencia (Endesa Educa, 2014).

### 1.4.3. Smart Factory

El proceso principal de las fábricas inteligentes es la conversión de lo digital a lo físico en un sistema de fabricación reconfigurable. Estos sistemas son capaces de adaptar sus componentes de hardware y software para cumplir con los requisitos del mercado en cuanto a tipo y cantidad de producto a fabricar. Por ello, las máquinas en las fábricas inteligentes son sistemas ciberfísicos, es decir, sistemas físicos integrados con componentes de comunicación. Pueden tomar sus propias decisiones basadas en algoritmos de *Machine Learning*, capturar datos en tiempo real y aprender de comportamientos exitosos en experiencias pasadas (Rojko, 2017).

Los productos en estas fábricas también son inteligentes, pues incluyen sensorica avanzada que a través de *WiFi* recogen datos en tiempo real sobre localización, estado de producto y condiciones del entorno. Tienen el control sobre sus caminos logísticos dentro del sistema productivo y son capaces de optimizarlos. Además, durante su ciclo de vida, monitorean su propio estado para alertar sobre posibles fallos futuros, facilitando así las labores de mantenimiento (Rojko, 2017).

En la Industria 4.0, todos los elementos cuentan con una identidad virtual que es almacenada en la nube de datos. Dicha identidad incluye información del producto, documentos, modelos en 3D, estado del producto, histórico de datos, medidas, etc. La interoperabilidad y a la conectividad entre estos elementos es posible gracias al IIoT, que crea flujos de información entre máquinas, productos y componentes (Rojko, 2017).

Sin embargo, otro tema importante es la colaboración entre humanos y máquinas, necesaria en ciertas tareas que no han podido ser totalmente automatizadas. Ejemplos de esta colaboración son los robots colaborativos en las líneas de fabricación o las interfaces basadas en realidad aumentada para el mantenimiento industrial.

### 1.4.4. Smart Logistics

La transformación digital ha revolucionado el concepto de transporte y la manera de gestionarlo. La logística inteligente implica una optimización y mejora en todos los aspectos de la cadena de suministro mediante la unificación de los procesos donde los agentes pueden trabajar de manera colaborativa. Gracias a los avances tecnológicos es posible, por ejemplo, controlar en tiempo real y desde un dispositivo móvil una flota de vehículos y seleccionar la mejor ruta de reparto en función del tráfico (Transportes Callizo, 2017).

La gestión de almacenes 4.0 es producto de la fusión de las herramientas *Just-In-Time* con las tecnologías 4.0, lo que incrementa los indicadores de productividad, reduce el número de movimientos y optimiza el espacio del almacén. Sistemas de *picking* y *kitting* automáticos junto con el empleo de AGV, se traduce en flujos logísticos internos automatizados y sincronizados con producción (OBS Business School, 2019).

Otras tecnologías clave en el ámbito logístico son los etiquetados inteligentes mediante radiofrecuencia (RFID), la sensorización de los contenedores, el *Blockchain* para certificar la trazabilidad de los productos, los vehículos autónomos, la simulación 3D para las mejoras en

el *layout*, el empleo de drones para el transporte de paquetes o el *Big Data* para la recolección y análisis de datos (García Samartino, 2018).

Este nuevo modelo aporta numerosas ventajas tanto al cliente como al proveedor. Algunas de ellas son la gestión automatizada de rutas, la geolocalización de productos, la reducción de costes, la mejora de la seguridad en la carretera y la mayor trazabilidad y transparencia de los procesos colaborativos. Los flujos constantes de información unidos a un buen sistema de gestión de clientes facilitan la toma de decisiones para prever las necesidades del consumidor y mejorar su satisfacción (Transportes Callizo, 2017).

#### 1.4.5. Smart Consumer

Los consumidores inteligentes son consumidores hiperconectados que, antes de realizar una compra, emplean la tecnología para investigar sobre un producto o servicio con el objetivo de hacer un balance global sobre los beneficios que esa compra le puede ocasionar (Yupcharge, 2019). Para ello, hacen uso de plataformas *online* como redes sociales o foros donde consultar las valoraciones de otros usuarios y comparar precios. Además, tienden a realizar las compras *online* debido a la comodidad y a la variedad de opciones que brinda el comercio electrónico.

Las empresas, como respuesta ante la dificultad de atraer a este nuevo modelo de consumidor, deben formar parte de estas comunidades y resolver rápidamente cualquier problema de cualquier consumidor. Dado que las personas son una fuente inagotable de datos, las empresas necesitan sistemas de *Big Data* para establecer patrones y predecir el comportamiento de los futuros consumidores (Esquivel, 2018).

Por otra parte, la publicidad juega un papel fundamental debido a que, al igual que los productos, es mucho más personalizada. Tras visitar una página *web*, los datos del consumidor quedan registrados de forma que esta persona recibe publicidad sobre ese producto mientras realiza otras consultas. En resumen, el control ha pasado a ser del consumidor, quien está altamente informado y forma parte de los procesos con el objetivo de obtener una buena experiencia de compra y un producto o servicio personalizado.

### 1.5. Estado actual de la digitalización

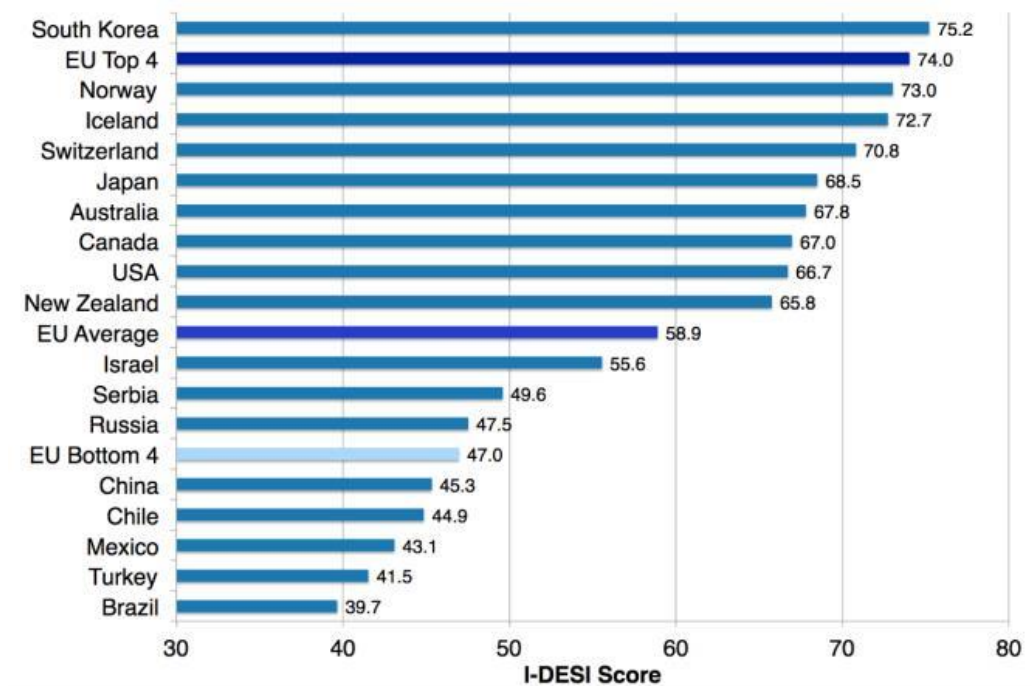
En este apartado se expone la situación actual en la que se encuentra la digitalización tanto en Europa como en España, haciendo hincapié en la industria española y en sus principales problemas frente a la digitalización.

#### 1.5.1. Digitalización en Europa

A nivel europeo, el Índice de la Economía y la Sociedad Digitales (DESI, *Digital Economy and Society Index*) refleja mediante cinco indicadores la evolución de los estados miembros de la

Unión Europea en cuanto a competitividad digital. Esos indicadores son la conectividad, el capital humano con habilidades digitales, el uso de los servicios de internet por los ciudadanos, la integración de las tecnologías digitales en los negocios y los servicios públicos digitales. Del mismo modo, el Índice Internacional de la Economía y la Sociedad Digitales (I-DESI, *International Digital Economy and Society Index*) emplea dichos indicadores para comparar Europa con otros países del mundo.

En primer lugar, se muestra el I-DESI 2016 (*Figura 1. 3*), que recoge la media de los resultados de las cinco dimensiones en economías de todo el mundo.



*Figura 1. 3. Resultados del I-DESI en 2016 para diferentes países (Comisión Europea, 2018).*

Los resultados posicionan a Corea del Sur en la cabeza de la clasificación, seguido del top 4 europeo y de Noruega. La media europea sobrepasa ligeramente la media global, así como los cuatro últimos de Europa se mantienen por delante de países como China, México y Brasil. Estados Unidos, Japón y Canadá son economías potentes que logran una posición superior a la media europea pero aún lejos de alcanzar a los líderes europeos. En la *Figura 1. 4* se observa la evolución que ha experimentado Europa, Corea del Sur, Japón, Estados Unidos y China entre 2013 y 2016.

En lo que se refiere a Europa, la tendencia hacia la digitalización ha sido prácticamente lineal. No obstante, los países europeos peor posicionados de la clasificación han dado un salto recientemente, al igual que China, con la que cada vez guarda menos diferencia. El escenario inicial de 2013 presentaba puntuaciones muy similares para Japón y Corea del Sur, y una buena diferencia entre los líderes europeos y Estados Unidos. Cuatro años más tarde, Corea del Sur y Japón demuestran que saben adaptarse mejor al mundo digitalizado y conectado que Estados Unidos, e incluso mejor que Europa en el caso de Corea.

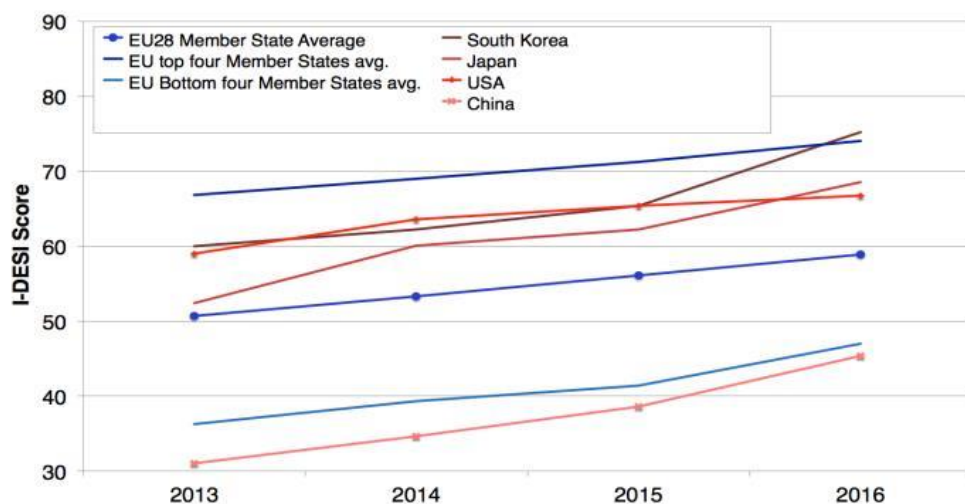


Figura 1. 4. Evolución del I-DESI entre 2013 y 2016 para diferentes países (Comisión Europea, 2018).

En segundo lugar, se presentan los resultados del DESI 2018, estableciendo una clasificación entre los 28 miembros de la Unión Europea sometidos a este estudio (Figura 1. 5)

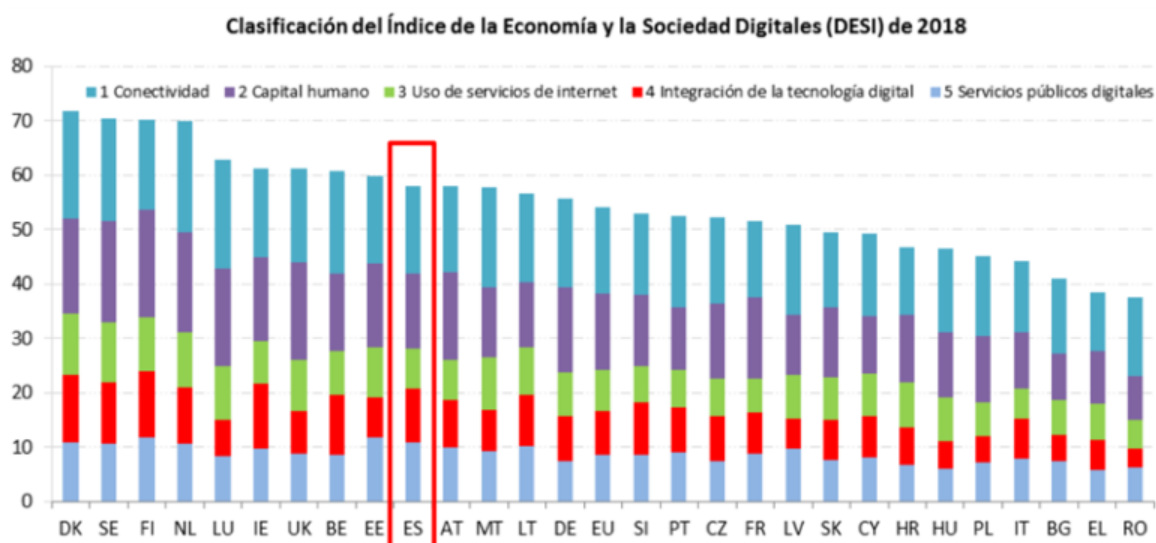


Figura 1. 5. Resultados del DESI 2018 para diferentes países de la UE (Comisión Europea, 2018).

Según esta clasificación, los líderes europeos en cuanto a digitalización son Dinamarca, Suecia, Finlandia y los Países Bajos. A continuación, se encuentran Luxemburgo, Irlanda, Reino Unido, Bélgica y Estonia. España ocupa el décimo puesto de la clasificación de entre los 28 estados analizados y figura como uno de los países europeos que más ha progresado en los últimos cuatro años, junto a Irlanda y Chipre. Por el contrario, Portugal es el que menos ha progresado en el último año y Rumanía continúa con la puntuación más baja de la clasificación (Comisión Europea, 2018). En la Figura 1. 6 se muestra el rendimiento relativo de España para cada una de las cinco dimensiones.

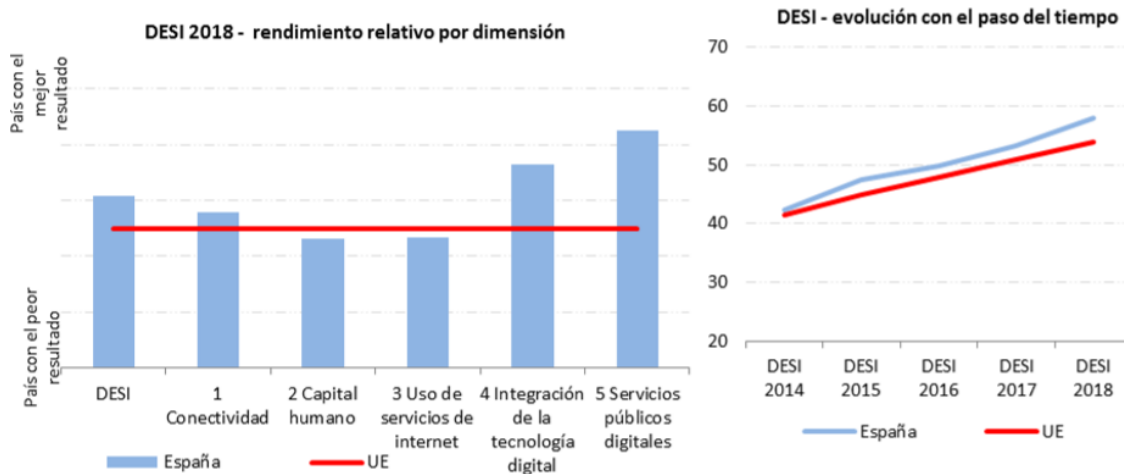


Figura 1. 6. Rendimiento relativo y evolución del DESI 2018 para España (Comisión Europea, 2018).

Los resultados de España son buenos en conectividad, debido a la disponibilidad e implantación de redes de banda ancha rápidas tanto fijas como móviles. El capital humano español sigue por debajo de la media europea a pesar de la mejora respecto a años anteriores. De hecho, la mitad de los ciudadanos españoles carece de competencias digitales básicas y una quinta parte aún no está en línea. El mercado laboral español demanda perfiles profesionales especialistas en las TIC, pero en menor medida que la media europea. El aprovechamiento en el uso de los avances tecnológicos por parte de las empresas españolas es elevado, incluyendo las redes sociales, las facturas electrónicas, los servicios *cloud* y el comercio electrónico. Finalmente, España destaca en el ámbito de los servicios públicos digitales al disponer de una gran cantidad de datos abiertos y servicios en línea (Comisión Europea, 2018).

### 1.5.2. Digitalización de la industria española

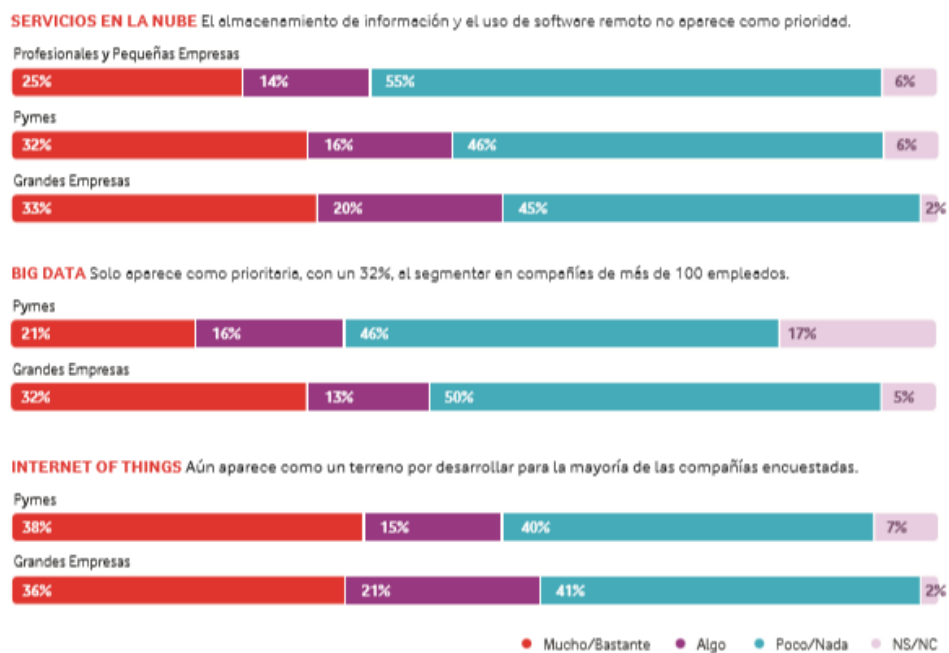
El DESI da una idea de cómo es la economía y la sociedad española en comparación con el resto de los países, pero sin profundizar lo suficiente en la situación actual de la industria. En cambio, un estudio realizado por *Advice Strategic Consultants S.L.* y presentado por el Observatorio Vodafone de la Empresa, muestra el estado de digitalización de las empresas y Administraciones Públicas españolas. Entre abril y junio de 2017 se realizaron 44 entrevistas en profundidad y 2.406 entrevistas telefónicas a altos mandos de empresas e instituciones, distinguiendo entre Profesionales y Pequeñas Empresas (1-9 empleados), PYMES (10-99 empleados), Grandes Empresas (más de 100 empleados) y Administraciones Públicas.

De entrada, más del 70% de las empresas creen estar en un nivel medio o avanzado en cuanto a digitalización, y al menos 2 de cada 3 afirma haber invertido en recursos digitales en los dos últimos años. El 10% de las grandes empresas aún no ha comenzado el proceso de digitalización y, además, no tiene un plan para hacerlo, mientras que en las pequeñas empresas se eleva hasta el 25%. Esto se debe, en parte, a que la digitalización no es una de las cinco principales

prioridades o preocupaciones para la mayoría de las empresas (Observatorio Vodafone de la Empresa, 2017).

Por otra parte, las organizaciones españolas no están abordando el cambio digital de forma planificada, sino que buscan soluciones espontáneas según unas necesidades concretas en un momento concreto. El porcentaje de pequeñas y medianas empresas que afirma tener un plan definido no supera el 20%. En cambio, prácticamente la mitad de las grandes empresas disponen de un plan definido y con presupuesto asignado (Observatorio Vodafone de la Empresa, 2017).

Todas las empresas coinciden en los beneficios que esperan obtener de la digitalización: aumentar la eficiencia de los procesos, mejorar la eficiencia en la gestión y el ahorro de tiempo junto con la accesibilidad ágil a la información. Para llegar a conseguir dichos beneficios, las empresas que invierten en lo digital priorizan las inversiones en seguridad y en conectividad. Sin embargo, las inversiones en tecnologías como los servicios en la nube, *Big Data* e *Internet Of Things* aún son escasas y solo se presentan como prioritarias en un 30% de las empresas (Observatorio Vodafone de la Empresa, 2017), tal y como se muestra en la *Figura 1. 7*.



*Figura 1. 7. Importancia para las empresas en inversiones en nuevas tecnologías (Observatorio Vodafone de la Empresa, 2017)*

En cuanto a la relación con el cliente, un alto porcentaje de empresas no ven necesario apostar por las nuevas tecnologías para mejorar su posicionamiento en marketing digital. En este caso, cuanto más grande es la empresa, más se percibe esta necesidad.

Otra cuestión que se plantea es si las empresas llevan a cabo este proceso por ellas mismas o con ayuda de proveedores externos. El 70% de las empresas cree que necesitará ayuda externa en mayor o menor medida, normalmente por parte de consultoras y empresas de informática o telecomunicaciones (Observatorio Vodafone de la Empresa, 2017).



Las Administraciones Públicas consideran que la digitalización es una prioridad muy alta, debido a la necesidad de aplicar la Ley de la Administración Electrónica y también porque las instituciones están muy concienciadas en este aspecto. La importancia de los procesos digitales para estas instituciones mejora la eficiencia en la gestión tanto interna como de servicios prestados a la ciudadanía (Observatorio Vodafone de la Empresa, 2017).

La relación con los ciudadanos ha mejorado en los últimos dos años según el 70% de las administraciones encuestadas gracias a la calidad y a la rapidez de la atención prestada con ayuda de las nuevas tecnologías. Al igual que para las empresas, los servicios digitales prioritarios son la seguridad y la conectividad, dejando a un lado la analítica y el *Big Data* (Observatorio Vodafone de la Empresa, 2017).

La principal conclusión que se obtiene de este estudio es que las empresas más grandes son las más digitalizadas. Normalmente disponen de personal más cualificado, de una mayor cantidad de dinero enfocada a proyectos de digitalización y de un plan definido con el que obtener unos beneficios concretos. Las Administraciones Públicas van por buen camino y están muy centradas en mejorar la relación con los ciudadanos, solo que aún existen grandes diferencias en cuanto a digitalización en función del tipo de administración.

### 1.5.3. Principales barreras en la digitalización de la industria

La *Figura 1. 8* representa una clasificación de barreras percibidas en diferentes tipos de empresas a la hora de llevar a cabo la transformación digital.

#### TOP 5 DE BARRERAS PERCIBIDAS

| Profesionales y Pequeñas Empresas                            | Pymes  | Grandes Empresas  |
|--|--|---|
| Coste elevado de la implementación <b>15%</b>                | Coste elevado de la implementación <b>16%</b>  | Coste elevado de la implementación <b>20%</b>               |
| Desconocimiento de la oferta de nuevas tecnologías <b>5%</b> | Formación de los empleados <b>6%</b>           | Resistencia de los empleados a la implementación <b>13%</b> |
| Falta de tiempo para implementación <b>5%</b>                | Ausencia de personal cualificado <b>5%</b>     | Ausencia de personal cualificado <b>12%</b>                 |
| Formación de los empleados <b>5%</b>                         | Falta de tiempo para implementación <b>5%</b>  | Falta de tiempo para la implementación <b>9%</b>            |
| Ausencia de personal cualificado <b>4%</b>                   | Problemas con la conexión a Internet <b>5%</b> | Gestión del cambio / adaptación <b>8%</b>                   |

*Figura 1. 8. Barreras para las empresas españolas en la transformación digital (Observatorio Vodafone de la Empresa, 2017)*

Se detectan tres barreras principales a las que se enfrenta la transformación digital en España: la resistencia al cambio por parte tanto de los directivos como de los empleados; el coste asociado a la implantación de estos sistemas digitales y a la reorganización de los procesos; y la ausencia de personal cualificado en herramientas 4.0.

El cambio organizacional surge de la necesidad de romper con el equilibrio existente para llegar a convertirlo en algo beneficioso. Este proceso involucra fuerzas que se oponen al planteamiento del cambio, por lo que las organizaciones deben saber cómo minimizarlas. No existe una situación de cambio única, si no que los factores que condicionan un cambio y las herramientas para llevarlo a cabo difieren significativamente de un caso a otro (López, Restrepo, & López, 2013).

La transición entre un estado inicial y un estado final es un momento frágil donde aún permanecen las desventajas originales y no se perciben los nuevos beneficios esperados. Por ello, una mala gestión del cambio puede resultar en situaciones muy negativas para la organización como los sobrecostos, los retrocesos a las viejas prácticas pasado un tiempo, un mal ambiente de trabajo o la desmotivación del personal (López, Restrepo, & López, 2013).

Normalmente, la resistencia al cambio se debe a una falta de información, pues es muy difícil aceptar un cambio si no se sabe en qué consiste ni cuáles serán los impactos recibidos. Sin embargo, aun teniendo información las personas pueden no poder o no querer cambiar. El “no poder” se relaciona con la falta de habilidades requeridas para la nueva situación, con la falta de recursos, con el tipo de cultura organizacional o con la dificultad de trabajar en equipo. El “no querer” surge cuando las personas consideran que el cambio no les conviene o que les va a hacer salir de su zona de confort bien por la incertidumbre que presentan los cambios o por la necesidad de hacer un sobreesfuerzo para alcanzar la nueva situación (López, Restrepo, & López, 2013).

Por lo tanto, con el fin de evitar esta resistencia al cambio, es necesario cultivar una cultura empresarial sólida que desarrolle una mentalidad abierta a las nuevas iniciativas y que tenga unos objetivos alineados con los intereses tanto de la empresa como de los trabajadores.

El siguiente aspecto problemático de la transformación digital es el elevado coste que supone un rediseño de los procesos actuales hacia la industria inteligente. Antes de nada, los empresarios tienen que plantearse la evolución de su negocio y estudiar las tendencias del mercado, es decir, realizar una reflexión estratégica. Tras haber enfocado el modelo de negocio hacia la dirección deseada, conviene saber las nuevas tecnologías que mejor se adaptan a las necesidades prioritarias de la empresa. No es necesario aplicar todas las tecnologías para obtener buenos resultados, es más, adquirir tecnologías que la empresa no es capaz de explotar hace que el proceso sea aún más difícil.

Cuando la organización no dispone de personal especializado para llevar a cabo este tipo de proyectos digitales se recurre a ayuda externa. Estas empresas externas son habitualmente consultoras tecnológicas o *start-up* especializadas en una o varias tecnologías. El problema actual está ocasionado por el desconocimiento general sobre el funcionamiento de los proyectos de transformación digital, es decir, cliente y proveedor entran en un conflicto de intereses. El proveedor vende a un precio elevado su producto tecnológico como aparente solución a todos los problemas de la empresa, mientras que el cliente desconfía de estas *start-up* debido a su reciente aparición en el mercado y a su ausencia de casos de éxito.

Ante esta situación, en la que unos no saben venderse y otros no saben lo que comprar, una solución que se plantea es que los proveedores aporten muestras de prueba de software o la realización de un proyecto a pequeña escala con el fin de que los directivos se acostumbren a estas tecnologías, vean resultados y cojan confianza. Sin embargo, es importante dejar claro

que con poco presupuesto es posible realizar grandes cambios en el funcionamiento de la organización, siempre y cuando las personas estén involucradas y las necesidades críticas estén identificadas.

Por último, la falta de profesionales cualificados limita el desarrollo de los planes de digitalización. La automatización de las nuevas formas de producción hace que se reformulen muchos puestos de trabajo, y con ello la demanda de perfiles profesionales (Blanco, Fontrodona, & Poveda, 2017). En la *Figura 1. 9* se muestra el porcentaje de personal especializado en digitalización en función del tamaño de la empresa.



*Figura 1. 9. Presencia de trabajadores especializados en digitalización (Observatorio Vodafone de la Empresa, 2017)*

Actualmente, la formación profesional de la rama industrial no se adecúa lo suficiente a las tecnologías de la Industria 4.0, aunque ya muchas titulaciones incorporen herramientas de simulación o impresoras 3D en las aulas. Además, las empresas siguen echando en falta el desarrollo de las competencias transversales durante la formación, como por ejemplo los idiomas, la adaptabilidad a los cambios o la proactividad. El objetivo es reestructurar estos ciclos formativos en función de las necesidades de las empresas y de hacerlos más atractivos, en particular para las mujeres, que representan solamente el 5% de las matriculaciones (Blanco, Fontrodona, & Poveda, 2017).

En cuanto a la formación universitaria, las carreras STEM (*Science, Technology, Engineering and Mathematics*) siguen siendo las más demandadas. El inconveniente es que la demanda crece a un ritmo superior al de personas graduadas. Un estudio de Adecco (2016) muestra que en 2015 en Europa hubo un déficit de 365.000 trabajadores en el ámbito de las TIC. Teniendo en cuenta que hay un 60% menos de estudiantes de ingeniería informática de las que demanda el mercado, el déficit puede incrementarse hasta los 756.000 trabajadores en 2020 (Blanco, Fontrodona, & Poveda, 2017). En concreto, España es uno de los países con más desajuste entre formación y necesidades de las empresas.

A pesar de que la demanda de ciertos perfiles tecnológicos aumenta, otros están destinados a desaparecer. Surge así el efecto sustitución, por el que se destruyen puestos de trabajo en determinados sectores para ser sustituidos por sistemas automatizados. También existe el efecto

complementariedad, donde la automatización complementa las tareas del trabajador, haciendo que este sea más productivo (Blanco, Fontrodona, & Poveda, 2017).

Las habilidades requeridas son diferentes, por lo que los profesionales poco cualificados que realizan tareas simples y repetitivas se sustituirán por especialistas en software o mecatrónica. Un aspecto interesante es que no solo las tareas repetitivas pasarán a ser automatizadas, sino que con los avances en Inteligencia Artificial y Analítica avanzada se puede conseguir automatizar la conducción de vehículos o la realización de diagnósticos médicos, por lo que el impacto negativo aumenta (Blanco, Fontrodona, & Poveda, 2017).

Si no se innova en políticas de empleo y en formación, las tasas de desempleo serán cada vez mayores. Al fin y al cabo, la transformación digital es una necesidad, pero al mismo tiempo una oportunidad para invertir en beneficios asegurados y aumentar la competitividad de las empresas.

## 2 EL MACHINE LEARNING

Los seres humanos necesitan buscar patrones para comprender el mundo que les rodea y anticipar ciertos comportamientos. El aumento en el número de interacciones entre usuarios y sistemas está generando grandes volúmenes de datos de los cuales se esperan obtener patrones que permitan predecir resultados o comportamientos con el mínimo error. Para lograrlo, se emplean diferentes técnicas computacionales y en los últimos años están destacando aquellas que se engloban bajo el nombre de *Machine Learning* (Aprendizaje Automático), un área de la inteligencia artificial que es cada vez más importante en todos los sectores (Orozco, 2018).

Con frecuencia se emplean indistintamente conceptos como ciencia de datos, minería de datos, aprendizaje automático o inteligencia artificial (Álogos, 2018). En este capítulo se explican las diferencias entre estos términos, así como los fundamentos del *Machine Learning*. También se presentan los diversos tipos de aprendizaje en función del problema que se quiere resolver y algunos de los algoritmos empleados en cada caso.

### 2.1. ¿Qué es el Machine Learning?

El *Machine Learning* es una técnica asociada a la detección automática de patrones relevantes dentro de un conjunto de datos. En los últimos años, se ha convertido en una herramienta muy común en prácticamente todas las tareas que requieren extraer información a partir de grandes cantidades de datos (Shalev-Shwartz & Ben-David, 2014). En el día a día, estamos rodeados de tecnología basada en *Machine Learning*: el filtrado de correos electrónicos, los sistemas de recomendaciones, la detección facial y el reconocimiento del habla del smartphone, la previsión del tiempo atmosférico o la consulta del tráfico en la carretera. También se emplea en otros ámbitos como la medicina, el marketing, la logística o el mantenimiento de equipos industriales.

Debido a la complejidad de todas estas aplicaciones, un ser humano no es capaz de programar una serie de especificaciones concretas para la realización de dichas tareas, sino que tiene que dotar a las propias computadoras con la habilidad de aprender de la experiencia y de adaptarse a las nuevas situaciones (Shalev-Shwartz & Ben-David, 2014). Por otro lado, existen diferentes tipos de aprendizaje: el aprendizaje supervisado, el aprendizaje no supervisado, el aprendizaje reforzado y el aprendizaje profundo. Cada uno de ellos tiene unas características y unos algoritmos determinados, por lo que se utilizarán unos u otros dependiendo del problema.

#### 2.1.1. Conceptos previos

Arthur Samuel, pionero en *Machine Learning* e inteligencia artificial, definió el *Machine Learning* como “el campo de estudio que da a las computadoras la habilidad de aprender sobre algo para lo que no han sido explícitamente programadas” (McClendon & Meghanathan, 2015). Allá por la década de 1950, Arthur Samuel diseñó un programa informático que aprendía a jugar a las damas. Él no era un buen jugador de damas, pero consiguió que el programa jugara

miles de veces contra sí mismo hasta obtener la experiencia suficiente como para saber qué posiciones del tablero eran buenas y cuáles eran malas (Ng, 2012).

Tom Mitchell, también pionero del *Machine Learning*, propuso en 1997 que un problema de aprendizaje bien planteado se define de la forma siguiente: “un programa aprende de la experiencia E con respecto a una tarea T y a un rendimiento P, si el rendimiento medido por P en la tarea T mejora con la experiencia E”. Tomando como ejemplo el programa de aprendizaje del juego de las damas: la tarea T es el jugar a las damas, el rendimiento P se mide con el porcentaje de juegos ganados contra oponentes, y la experiencia E es el jugar contra sí mismo a modo de entrenamiento (Mitchell, 1997).

Otras definiciones de *Machine Learning* están enfocadas desde un punto de vista matemático, algorítmico o estadístico. Sin embargo, para englobar dichas perspectivas, el *Machine Learning* se define como “un conjunto de técnicas dentro del ámbito de la inteligencia artificial que utiliza métodos estadísticos para la búsqueda de patrones a partir de los cuales se crean máquinas inteligentes capaces de aprender y tomar decisiones en base a datos empíricos obtenidos de diversas fuentes” (Álogos, 2018).

Analizando esta última definición, el *Machine Learning* forma parte de lo que se denomina inteligencia artificial. La inteligencia es la habilidad de pensar y entender en lugar de hacer algo por instinto o automáticamente. A su vez, pensar es la actividad de usar el cerebro a la hora de considerar un problema o para crear una idea. No se especifica si la habilidad de pensar es propia únicamente de alguien o si, por el contrario, algo que posea un órgano que le permita aprender, entender, resolver problemas y tomar decisiones puede tener también esta habilidad. Por lo tanto, el objetivo de la inteligencia artificial como ciencia es hacer que las máquinas realicen tareas que requirieran inteligencia de ser realizadas por un humano (Negnevitsky, 2002).

Los subcampos de la inteligencia artificial son aquellos que se especializan en imitar una parte concreta del sistema cognitivo humano y de sus comportamientos: el *Machine Learning* es el encargado del reconocimiento de patrones y del aprendizaje; los sistemas expertos se ocupan de la toma de decisiones; los sistemas de búsqueda y planificación eligen la secuencia más adecuada de pasos a seguir para resolver un problema; el reconocimiento del habla y el procesamiento del lenguaje natural analizan y tratan la información recibida a través del lenguaje humano para posibilitar la comunicación; la percepción (como la visión artificial) permite conocer y sentir lo que hay alrededor; y la robótica transforma las decisiones en movimientos (Russell & Norvig, 2010).

La inteligencia artificial, a su vez, forma parte de la ciencia de datos (*Data Science*). La ciencia de datos se encarga de analizar y revelar con datos lo que hay detrás de fenómenos naturales, humanos y sociales desde un punto de vista multidimensional, flexible y dinámico (Hayashi, 1998). Este término se usó en la década de 1960 como sinónimo de ciencia computacional, y no fue hasta 2001 que se la consideró una disciplina independiente que amplía la estadística con avances informáticos de datos.

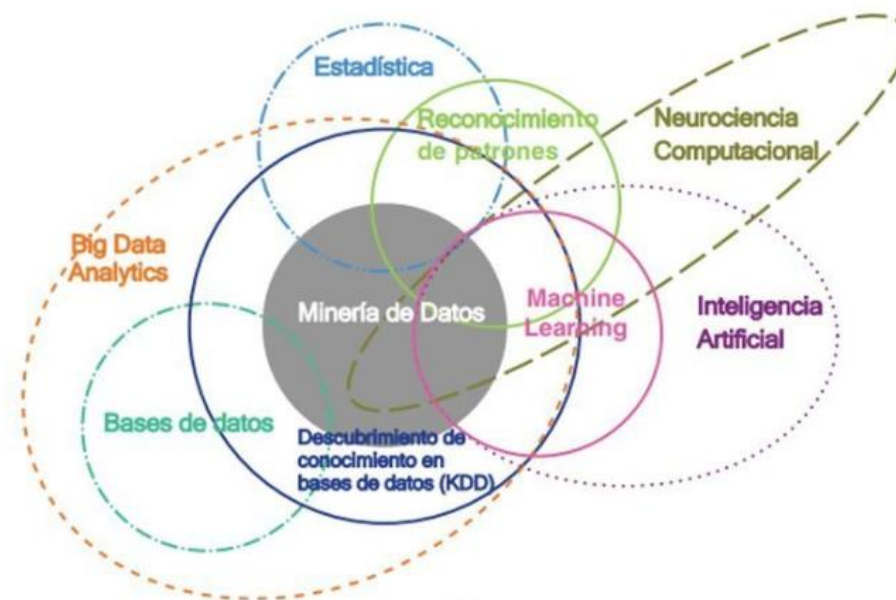
Al fin y al cabo, el objetivo de la ciencia de datos es hacer que los datos sean útiles, es decir, que influyan en acciones del mundo real. Para ello, se establecen tres fases: la búsqueda de inspiración en los datos, de la que se encarga la minería de datos y la analítica; la toma de

decisiones, basada en la estadística; y la creación de modelos, empleando algoritmos de *Machine Learning* (Kozyrkov, 2018).

Conviene diferenciar los conceptos de *Big Data* y de minería de datos (*Data Mining*). El *Big Data* se refiere a la obtención y almacenamiento de grandes conjuntos de datos mientras que el *Data Mining* se refiere al análisis profundo de los patrones en los datos. Por lo tanto, el *Big Data* necesita que el *Data Mining* aporte conocimiento y valor a los datos recogidos. Sin embargo, el *Data Mining* puede trabajar con cantidades pequeñas de datos, por lo que no es necesariamente dependiente del *Big Data* (Márquez, 2016).

La estadística se viene utilizando para aprender de los datos desde hace más de un siglo, solo que ahora existe la potencia de cálculo suficiente para tratar ficheros de datos de forma masiva y automática. Las habilidades estadísticas se requieren cuando, tras haber seleccionado los datos de los cuales se puede extraer información relevante, se desean tomar decisiones de alta calidad y bajo riesgo. No existe una técnica más inteligente que otra para llevar a cabo esta tarea, sino que cada uno emplea de forma inteligente las técnicas que mejor conoce en cada problema (Aluja, 2001).

A modo de resumen, la *Figura 2. 1* representa la ubicación del *Machine Learning* y su relación con otros campos. Se observa que el *Machine Learning* es una parte fundamental de la inteligencia artificial y que comparte técnicas con el *Data Mining* bajo el campo de la neurocomputación. Respecto al reconocimiento de patrones nace un subcampo del *Machine Learning* llamado *Deep Learning* (aprendizaje profundo), el cual emplea redes neuronales artificiales para aprender de forma no supervisada mediante varias capas jerárquicas.



*Figura 2. 1. El Machine Learning y su relación con otros campos (Villén, 2019).*

## 2.1.2. Historia del Machine Learning

El *Machine Learning* se fundamenta en las matemáticas, la estadística y la computación. Los desarrollos en computación no llegaron hasta la década de 1940, pero antes de ese momento hubo varios métodos y técnicas matemáticas y estadísticas que sirvieron de base para lo que hoy se conoce como *Machine Learning*: el Teorema de Bayes, el ajuste por Mínimos Cuadrados y las Cadenas de Márkov.

El Teorema de Bayes (1763) calcula la probabilidad de un evento A dado otro evento B a partir de la probabilidad del evento B dado A (Hayes, 2019). Esto quiere decir que, por ejemplo, si se tienen tres líneas de producción y se conoce la probabilidad de que un producto sea defectuoso habiéndose fabricado en la línea de producción A, es posible conocer la probabilidad de que un producto defectuoso provenga de la línea de producción A.

El ajuste por Mínimos Cuadrados (1805) es una técnica de optimización matemática que busca la función continua que mejor se aproxime a una serie de datos según el criterio de mínimo error cuadrático (Abdi, 2003). Por otra parte, la cadena de Márkov (1913) es un tipo de proceso estocástico discreto en el que la probabilidad de que ocurra un evento solo depende del evento inmediatamente anterior (Oxford Dictionaries, 2019).

En 1950, Alan Turing publicó un artículo titulado ‘Computación e Inteligencia’ en el que se preguntaba si las máquinas podían pensar (González, 2018). Para ello propuso el *imitation game*, una prueba consistente en una conversación escrita entre una persona y una máquina por la cual la computadora se consideraba inteligente si la persona no era capaz de distinguir si estaba hablando con una persona o con una máquina (Marr, 2016).

La primera red neuronal artificial fue creada por Marvin Minsky y Dean Edmonds en 1951 bajo el nombre de SNARC (*Stochastic Neural Analog Reinforcement Computer*), y su propósito era imitar a una rata a la hora de aprender a salir de un laberinto. En 1952, Arthur Samuel escribió un programa diseñado para aprender a jugar a las damas a partir de sus propios errores. Más tarde, en 1957, Fran Rosenblatt diseñó el *Perceptron*, una red neuronal que reconocía caracteres de forma similar al de los sistemas visuales biológicos (Marr, 2016). Pronto se vieron algunas de las limitaciones del *Perceptron*, ya que su arquitectura de una única neurona solo permitía la resolución de problemas lineales sin la posibilidad de extender el método de aprendizaje a otras redes más complejas.

Tras la publicación del libro *Perceptrons: an introduction to computational geometry* (Minsky & Papert, 1969), que explicaba matemáticamente las limitaciones del perceptrón, hubo un corte repentino en la financiación de proyectos de inteligencia artificial y, en particular, de aquellos que involucraban redes neuronales. Este periodo de más de quince años se conoce como el invierno de la inteligencia artificial, que acabó gracias a la publicación *Learning representations by back-propagating errors* (Rumelhart, Hinton, & Williams, 1986) por la cual se mostraba experimentalmente el algoritmo para conseguir que una red neuronal autoajustara sus parámetros a partir de sus errores: *Backpropagation* (Santana, 2018).

Diez años más tarde, en 1996, el computador *DeepBlue* de IBM ganó una partida de ajedrez al vigente campeón mundial (Marr, 2016). Al año siguiente, la nueva versión *DeeperBlue* consiguió vencer todas las partidas (González, 2018). En 2006, Geoffrey Hinton presentaba el



concepto de *Deep Learning* o aprendizaje profundo, basado en el concepto de *Backpropagation*, que permite que los computadores distingan objetos y textos procedentes de imágenes y vídeos (Marr, 2016).

La compañía británica DeepMind Technologies, fundada en 2010 y adquirida por Google en 2014, logró desarrollar una red neuronal capaz de aprender a jugar a videojuegos analizando únicamente el comportamiento de los píxeles en la pantalla. En 2015, DeepMind creó un algoritmo llamado AlphaGo, que venció a jugadores profesionales en el juego de mesa chino Go, considerado el más complejo y difícil de anticipar (BBC, 2019).

OpenAI es una compañía sin ánimo de lucro creada en 2015 y dedicada a la investigación en inteligencia artificial para el bien de la humanidad. En 2017, esta compañía despunta los avances en agentes conversacionales o *chatbots*, así como en el mundo de los videojuegos (Marr, 2016).

Actualmente, el término *singularity* está en boca de todos los expertos. Este concepto es el resultado del crecimiento exponencial de las capacidades cognitivas de la inteligencia artificial hasta el punto de igualar o superar las del cerebro humano. Así, la inteligencia artificial crearía versiones mejoradas de sí misma, llegando a tener a los humanos como esclavos. Ciertos organismos estudian la probabilidad de que este escenario ocurra en un futuro y, aunque no sea muy probable, sigue suponiendo un gran riesgo para la raza humana (BBC, 2019).

## 2.2. Fundamentos del Machine Learning

En este apartado se presentan los fundamentos para comprender la forma en la que las máquinas aprenden. Para ello, se explican los conceptos de función de coste y de método del descenso del gradiente. A continuación, se muestran los principales errores o problemas que surgen a la hora de realizar un modelo de *Machine Learning*: la varianza (*variance*), el sesgo (*bias*), el sobreajuste (*overfitting*) y el subajuste (*underfitting*). Por último, se explican los pasos a seguir para desarrollar e implementar una solución con *Machine Learning*.

### 2.2.1. El aprendizaje

El aprendizaje en un sistema de *Machine Learning* consiste en el ajuste de los parámetros de un modelo en función de los datos recibidos. Este conjunto de datos recibe el nombre de *data set*, el cual contiene variables tanto independientes como dependientes. Las variables independientes (*features*) son aquellas columnas del *data set* usadas por el algoritmo para generar un modelo que prediga lo mejor posible las variables dependientes. Por otro lado, las variables dependientes (*labels*) son las columnas del *data set* resultado de una correlación entre variables independientes, por lo que deben ser predichas por el modelo implementado (Esteve, 2017).

La programación clásica se entiende como un conjunto de reglas diseñadas por una persona con el fin de obtener unas respuestas que cumplan dichas reglas a partir de unos datos de entrada.

El *Machine Learning*, en cambio, se encarga de obtener el conjunto de reglas más efectivas que relacionan los datos de entrada con las respuestas que esperamos obtener a través del aprendizaje (Figura 2. 2). La ventaja del *Machine Learning* es que estas reglas pueden aplicarse a diferentes datos de entrada para producir respuestas que han sido generadas automáticamente por lo que el sistema aprendió y no por las instrucciones generadas por un humano (Chollet, 2017).

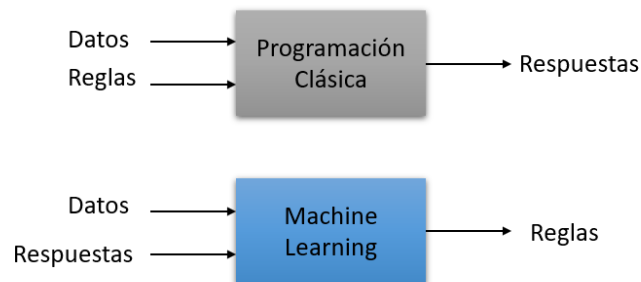


Figura 2. 2. La programación clásica frente al Machine Learning. Elaboración propia.

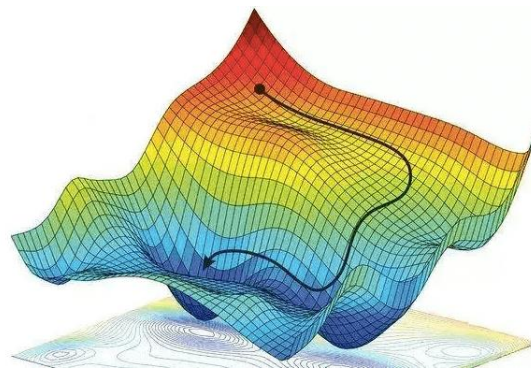
Por ejemplo, si se quisiera diseñar un algoritmo de *Machine Learning* para predecir el precio de una vivienda en una determinada ciudad, la variable dependiente (lo que queremos obtener o predecir) sería el precio de la vivienda; mientras que las variables independientes serían todos aquellos factores que influyen en el precio de la vivienda como la superficie, el número de habitaciones, la distancia al centro de la ciudad, etc.

El modelo debe estar lo suficientemente ajustado a los datos de entrada, pero también debe tener la suficiente consistencia como para dar un buen resultado ante la introducción de datos diferentes. Para ello, el *data set* se divide en dos subconjuntos de datos: los datos de entrenamiento (*training data*), que corresponden aproximadamente al 80% del *data set*; y los datos de test (*test data*), que corresponden al 20% restante y se emplearán para medir la calidad del modelo tras el entrenamiento (Esteve, 2017).

Una vez se tienen los datos se necesita establecer una hipótesis, es decir, encontrar una ecuación que se aproxime lo mejor posible al comportamiento real del fenómeno que se está modelizando. Esta ecuación es la que relaciona los datos de entrada y los parámetros del modelo con la salida. Aquello que no se mide no se puede mejorar, así que el siguiente paso consiste en encontrar el error en la predicción y tratar de minimizarlo (Santana, 2017).

La función de coste es la encargada de recopilar el error entre la variable dependiente que se quiere determinar y la hipótesis, en función de los parámetros del modelo. Para minimizar una función se recurre a igualar su derivada a cero. En algunos casos es sencillo encontrar la fórmula que, introduciendo los datos de entrada y de salida, proporcione el mejor valor de los parámetros para reducir el error al mínimo. Sin embargo, con otros modelos y otras funciones de coste no siempre se va a poder encontrar el mínimo de la función de coste de forma analítica (Santana, 2017). Por esta razón, lo habitual es emplear métodos iterativos que permitan minimizar poco a poco el error, como por ejemplo el método del Descenso del Gradiente (*Gradient Descent*).

La diversidad de modelos y funciones de coste en el *Machine Learning* obliga a encontrar soluciones para las funciones no-convexas, es decir, para aquellas que tienen más de un mínimo. El método del Descenso del Gradiente aprovecha el cálculo de la derivada para encontrar los mínimos locales, ya que la derivada indica el valor de la pendiente en un punto determinado. Puesto que lo que se desea es llegar al punto mínimo, lo lógico es avanzar en la dirección en la que la pendiente es mayor. De ahí que los pasos a seguir por este método sean los siguientes: localizar la mayor pendiente en la posición actual, avanzar en esa dirección una distancia determinada y pararse en esa nueva posición (Santana, 2018). Se repite el proceso iterativamente hasta la convergencia. En la *Figura 2. 3* se representa el método gráficamente.



*Figura 2. 3. Representación gráfica del Descenso del Gradiente. Recuperado de: <https://easyai.tech/en/ai-definition/gradient-descent/>*

Matemáticamente, el gradiente es una generalización de la derivada a funciones de más de una variable. En concreto, el gradiente en un punto de una función de varias variables es el conjunto ordenado de las derivadas parciales de esa función en ese punto. Da información sobre lo que varía la función por cada unidad que varía cada variable en el punto considerado (Redondo, Melchor, & Quintela, 2012). Se representa mediante el operador nabla  $\nabla$ . Por ejemplo, el gradiente de una función de tres variables se designa de la siguiente forma (Ecuación 2.1.):

$$\text{grad}f = \nabla f = \left( \frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}, \frac{\partial f}{\partial z} \right) \quad (2.1)$$

Un gradiente positivo indica la dirección por la que avanzar si se quisiera encontrar un máximo (Ascenso del Gradiente). En este caso, se toman pasos proporcionales al negativo del gradiente (Ecuación 2.2.), ya que el objetivo es encontrar un mínimo de la función (Descenso del Gradiente). El ratio de aprendizaje  $\alpha$  es un factor que define cuánto afecta el gradiente a la actualización de los parámetros en cada iteración, en otras palabras, indica lo que se avanza en cada paso (Santana, 2018).

$$\theta := \theta - \alpha \nabla f \quad (2.2)$$

Para valores bajos del ratio de aprendizaje, el algoritmo resulta ineficiente al necesitar muchas iteraciones para converger. Para valores altos el algoritmo no converge debido a que, entre una iteración y otra, se recorre un paso demasiado grande como para que el punto se introduzca en la zona de mínimo coste. La correcta configuración de este ratio es fundamental para hacer que el algoritmo funcione correctamente (Santana, 2018).

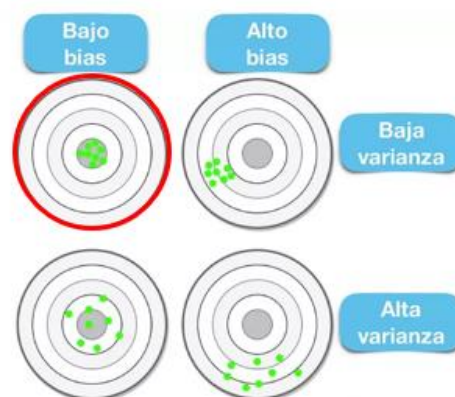
### 2.2.2. El error y los problemas de ajuste

La precisión y la capacidad de generalizar son aspectos clave a la hora de realizar un modelo de *Machine Learning* solo que, desgraciadamente, es imposible conseguir que un modelo esté libre de errores por completo. Comprender las principales fuentes de error ayuda a prevenir dos de los problemas más habituales en el ajuste de parámetros: el sobreajuste (*overfitting*) y el subajuste o falta de ajuste (*underfitting*) (González, 2018).

Los errores principales en la predicción de un modelo y que están asociados al algoritmo empleado son la varianza y el sesgo (*bias*). Existe un tercer tipo de error asociado a la cantidad de ruido de los datos de entrada, sobre el cual no se puede actuar porque es irreducible e independiente de lo bueno que sea el modelo (González, 2018).

El error de varianza está relacionado con el grado en el que la función objetivo cambia según los datos de entrenamiento proporcionados. Existen algoritmos que se ven muy influenciados por pequeñas variaciones en los datos de entrada, mientras que otros son más consistentes y absorben mejor estos cambios (González, 2018).

El error de sesgo es la diferencia entre los valores reales y la predicción que espera el modelo. Hay algoritmos más flexibles que otros a la hora de adaptarse a una distribución determinada de datos de entrada. Un alto sesgo implica un aprendizaje más rápido y sencillo pero un menor rendimiento predictivo (González, 2018). El error total es una combinación de varianza y sesgo, por lo que para minimizarlo es imprescindible lograr una baja varianza y un bajo sesgo. En la *Figura 2. 4* se muestran las cuatro combinaciones posibles entre varianza y sesgo.



*Figura 2. 4. Combinaciones entre los errores de varianza y sesgo (González, 2018).*

Sin embargo, la estrecha relación entre la varianza y el sesgo hace que disminuir uno de ellos implique aumentar el otro. Así, los algoritmos de baja varianza y alto sesgo resultan más simples y rígidos, como por ejemplo la Regresión Lineal o la Regresión Logística. Por otro lado, los algoritmos de alta varianza y bajo sesgo como los Árboles de Decisión o los *Support Vector Machine* son más complejos y flexibles, pero más inconsistentes (González, 2018).

En definitiva, para construir un buen modelo real es necesario llegar al equilibrio entre la varianza y el sesgo. Esto se consigue eligiendo un nivel de complejidad adecuado (Figura 2.5), donde un aumento del sesgo equivalga a una reducción de la varianza (González, 2018).

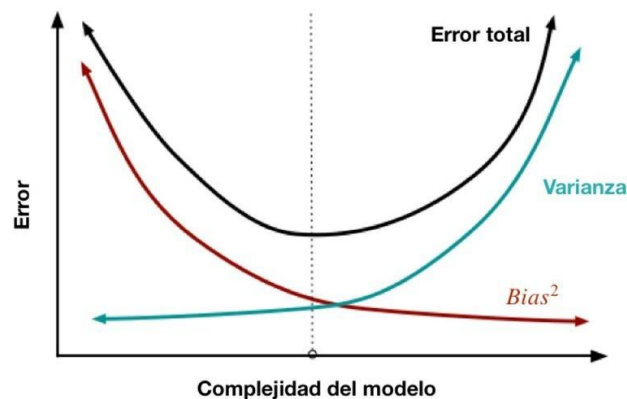


Figura 2.5. Nivel de complejidad del modelo en función del error (González, 2018).

El subajuste (*underfitting*) se refiere a un modelo con un nivel de complejidad muy bajo que no tiene la precisión suficiente como para alcanzar un ajuste adecuado debido a su alto sesgo. Puede ocurrir cuando el conjunto de datos de entrenamiento no es suficiente, o cuando se utiliza un modelo lineal para ajustar datos no lineales. Por otra parte, el sobreajuste (*overfitting*) se produce cuando el nivel de complejidad es elevado y, por lo tanto, el modelo no tiene la capacidad de generalizar su comportamiento ante diferentes datos de entrada. Sucede cuando el modelo recoge el ruido de los datos de entrenamiento y, en consecuencia, aumenta mucho su varianza (Santana, 2019). En la Figura 2.6 se muestran ejemplos de subajuste, de sobreajuste y de ajuste óptimo para los problemas clásicos de regresión y clasificación.

El problema más común es el del sobreajuste (*overfitting*), para el cual existen dos tipos de soluciones. La primera consiste en reducir el número de variables independientes (*features*), ya sea manualmente mediante la observación de aquellas variables que podrían desecharse, o bien a través de métodos que realizan esta tarea automáticamente. Sin embargo, se corre el riesgo de perder información realmente útil para el comportamiento del modelo. Por eso, la segunda opción es la regularización, gracias a la cual se mantiene el número de variables a costa de reducir el valor de los parámetros que afectan a las variables. Este método funciona bien cuando existe un gran número de variables y todas ellas contribuyen un poco a la hora de hacer la predicción (Ng, 2012).

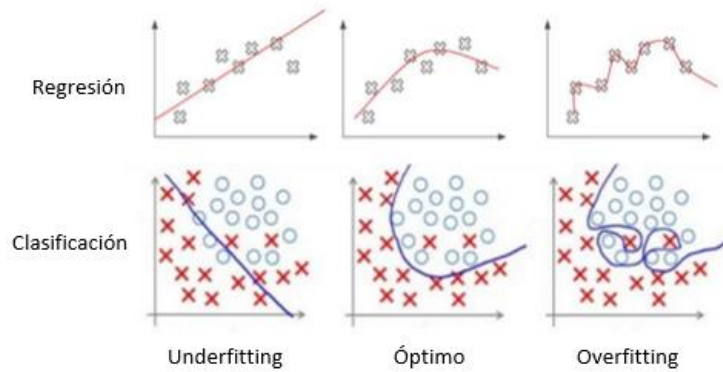


Figura 2. 6. Distintos ajustes para los problemas de regresión y clasificación. Recuperado de: <https://www.slideshare.net/AlyOsama2/deep-learning-overfitting-underfitting-and-regularization>

### 2.2.3. Etapas en un proyecto de Machine Learning

Una vez comprendidos los principios básicos del *Machine Learning*, es conveniente conocer el procedimiento para dar solución a problemas reales. Un proyecto de *Machine Learning* no se centra únicamente en elegir un modelo y entrenarlo, sino que, como todo proyecto, cuenta con una serie de etapas o pasos a seguir para aumentar sus probabilidades de éxito. A continuación, se describen ocho etapas genéricas para llevar a cabo un proyecto de *Machine Learning*.

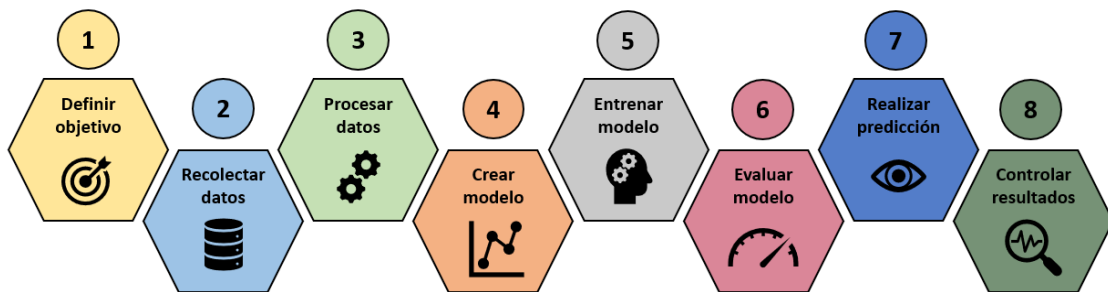


Figura 2. 7. Etapas genéricas para llevar a cabo un proyecto de Machine Learning. Elaboración propia.

La primera etapa consiste en entender el problema que se quiere resolver, ya que gran parte de las decisiones tomadas a lo largo del proyecto dependerán de lo bien que se haya comprendido el contexto (Sánchez, 2019). Para ello es necesario definir unos objetivos claros, específicos, medibles y alcanzables en un periodo determinado de tiempo. Es una fase tediosa en la que se requiere mucha información por parte de expertos en el sector sobre el que se va a trabajar (Martínez, 2018). En este punto se clasifica el problema (supervisado, no supervisado, etc.) e incluso se elige el tipo de modelo que se va a entrenar (regresión, clasificación, agrupamiento, etc.) (Sánchez, 2019).

En la segunda etapa se define la cantidad y el tipo de datos necesarios, así como el origen de dichos datos. La calidad de los datos que alimentan la máquina tiene un impacto directo en el

funcionamiento del modelo (Bagnato, 2017). En algunos casos, no se dispone de todos los datos necesarios para resolver el problema, por lo que se compran a entidades externas. Aun así, en otros muchos casos no se puede acceder a los datos deseados porque ni siquiera existen (Velogig, 2018).

El tratamiento de los datos una vez recolectados corresponde a la etapa número tres. El objetivo principal de esta etapa es visualizar y analizar cuáles son las variables que mejor representan aquello que se quiere predecir (Sánchez, 2019). Aquí se observan algunas de las correlaciones entre las variables independientes y las dependientes de forma que, en caso de tratarse de un número elevado de variables, pueden aplicarse métodos de regularización. Los datos necesitan un formato determinado para poder ser procesados por la máquina de la manera más sencilla posible (Martínez, 2018). Una vez hecho esto, se divide el *data set* en los subconjuntos de *training data* y *test data*.

Aunque en el primer paso ya se conoce el tipo de problema, es en la etapa cuatro donde se define por completo el modelo que mejor se ajusta al problema: regresión lineal, árboles de decisión, red neuronal, k-vecinos más cercanos, etc. Esta etapa es crucial pero no suele ser complicada dado que los algoritmos pueden encontrarse en bibliotecas predeterminadas.

La quinta etapa se dedica al entrenamiento del modelo a partir de los datos de entrenamiento. Los parámetros se ajustan automáticamente por el algoritmo seleccionado a medida que se entrena el modelo (Bagnato, 2017). En proyectos nuevos se busca entrenar modelos sencillos y en poco tiempo simplemente para validar que el problema puede ser solucionado con *Machine Learning*, ya que no siempre se puede (Sánchez, 2019).

En la etapa número seis se verifica la precisión del modelo mediante la introducción de los datos de test, que son datos que la máquina aún no conoce. Un 50% de precisión es insuficiente para validar un modelo, pues indica que la mitad de las veces fallará. En cambio, alcanzar un 90% de precisión proporciona una confianza suficiente. También se evalúan los errores que hacen que el modelo no generalice bien con el fin de elegir la solución más conveniente: adquirir más datos, usar un modelo más simple, usar uno más complejo, comprender mejor el problema, etc. A esta etapa también se la conoce como *Parameter Tuning* (configuración de parámetros), pues consiste en ajustar los parámetros del modelo para mejorar los resultados obtenidos (Bagnato, 2017).

Cuando se alcanza el nivel de error deseado, el modelo queda validado y puede pasarse a la penúltima etapa, que es la unión entre la simulación y el mundo real. Se trata de integrar el modelo en un sistema real con el que pueda comunicarse. Más que un problema de ciencia de datos es un problema de desarrollo de *software*, ya que consiste en entender cómo cambian los procesos una vez se tenga la predicción (Sánchez, 2019).

Por último, la etapa número ocho pone fin al proceso con la monitorización de los resultados. Es necesario asegurar que el modelo aporta un alto valor predictivo y, lo más importante, que cumple con los objetivos marcados en la primera etapa (Sánchez, 2019). Sin este *feedback*, con el paso del tiempo es posible que el modelo no de los resultados que se esperaba.

## 2.3. Aprendizaje supervisado

En el aprendizaje supervisado, el agente observa pares de datos de entrada y salida a modo de ejemplo para aprender una función que modele la salida según la entrada (Russell & Norvig, 2010). Por lo tanto, en los datos utilizados para construir el modelo se encuentra la información que se desea predecir. Existen dos tipos de problemas en el aprendizaje supervisado: los problemas de regresión y los problemas de clasificación. Fundamentalmente, un modelo de regresión predice una cantidad continua mientras que un modelo de clasificación predice una etiqueta (Management Solutions España, 2018).

A continuación, se explican brevemente algunos de los algoritmos de regresión y de clasificación más empleados. Entre los algoritmos de regresión destacan la regresión lineal y la regresión polinomial; y entre los algoritmos de clasificación destacan la regresión logística, *Support Vector Machine* y *K-Nearest Neighbors*. Otros algoritmos conocidos son los Árboles de Decisión, *Random Forest* y la clasificación de Naïve Bayes.

### 2.3.1. Regresión Lineal

La Regresión Lineal es una técnica ampliamente utilizada en el campo del Aprendizaje Supervisado debido a su sencillez y a su gran utilidad. Consiste en predecir una variable dependiente y en función de una o varias variables independientes  $x$ , mediante el trazado de la línea recta que mejor se ajusta al conjunto de datos proporcionados (Tusell, 2011). La Ecuación 2.3 representa la hipótesis de este modelo.

$$h(x) = b_0 + b_1x_1 + b_2x_2 + b_3x_3 \dots \quad (2.3)$$

Los coeficientes  $b_i$  son los parámetros del modelo que se ajustan en la etapa de entrenamiento a medida que se introducen más datos. El objetivo del modelo es minimizar la función de coste, es decir, hacer que la diferencia entre el valor de salida real ( $y$ ) y el valor de la predicción ( $h$ ) sea mínima. En este caso, la función de coste es el Error Cuadrático Medio, que mide la distancia al cuadrado entre cada punto y la vertical que lo une con la recta de regresión (Ecuación 2.4.).

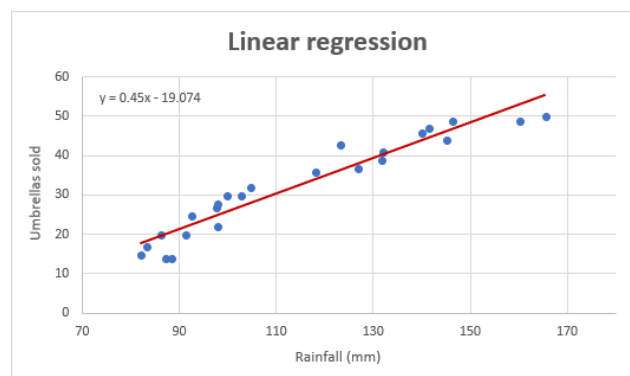
$$J(a, b) = \frac{1}{2m} \sum_{i=1}^m (h(x^{(i)}) - y^{(i)})^2 \quad (2.4)$$

Para minimizar esta función de coste se emplea el algoritmo del Descenso del Gradiente, que es un proceso iterativo que reduce poco a poco el error hasta encontrar un mínimo en la función de coste (ver apartado 2.2.1).

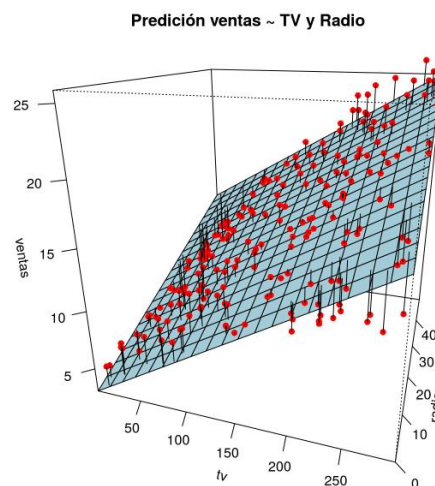


La Regresión Lineal puede ser simple si solo existe una variable independiente, o múltiple si existe más de una. Este modelo es rápido y robusto, pero para asegurar su buen funcionamiento debe haber cierta relación lineal entre la entrada y la salida (González, 2018).

Un ejemplo de Regresión Lineal simple es predecir el número de paraguas que se venderán en función de la cantidad de lluvia según el histórico del año anterior (*Figura 2. 8*). Un ejemplo de Regresión Lineal múltiple es la predicción de las ventas de un producto en función del dinero invertido en publicidad de TV y en publicidad de radio (*Figura 2. 9*). En este último caso, al haber dos variables independientes, es posible su representación mediante el ajuste de un plano en lugar de una recta.



*Figura 2. 8. Ejemplo de regresión lineal simple. Predicción de venta de paraguas en función de la lluvia. Recuperado de: <https://www.ablebits.com/office-addins-blog/2018/08/01/linear-regression-analysis-excel/>*



*Figura 2. 9. Ejemplo de regresión lineal múltiple. Predicción de las ventas de un producto en función del dinero invertido en publicidad de TV y en publicidad de radio. Recuperado de: [https://rpubs.com/Joaquin\\_AR/226291](https://rpubs.com/Joaquin_AR/226291)*

### 2.3.2. Regresión Polinomial

La Regresión Polinomial trata de encontrar un polinomio de grado  $n$  que se ajuste a la distribución de datos mediante una curva. Es útil cuando la Regresión Lineal no es capaz de ajustarse lo suficiente a los datos debido a algún tipo de no linealidad entre ellos (González, 2018). La Ecuación 2.5 representa la hipótesis para una Regresión Polinomial de tercer grado y de una variable, mientras que la Ecuación 2.6 representa la hipótesis para una Regresión Polinomial de segundo grado y de dos variables.

$$h(x) = b_0 + b_1x_1 + b_2x_1^2 + b_3x_1^3 \quad (2.5)$$

$$h(x) = b_0 + b_1x_1 + b_2x_1^2 + b_3x_2 + b_4x_2^2 + b_5x_1x_2 \quad (2.6)$$

Los modelos polinomiales ganan mucha flexibilidad con respecto a los lineales, ya que en función del grado del polinomio se obtienen curvas más o menos ajustadas. Sin embargo, aumentar demasiado el grado del polinomio trae consigo el problema del sobreajuste (*overfitting*) ya que, aunque se reduzca mucho el error para los datos iniciales, el modelo pierde su capacidad para generalizar ante nuevos datos de entrada (González, 2018).

Por ejemplo, la demanda horaria de energía en función de la temperatura exterior sigue un comportamiento no lineal que se ajusta con un polinomio de grado cuatro (Figura 2. 10).

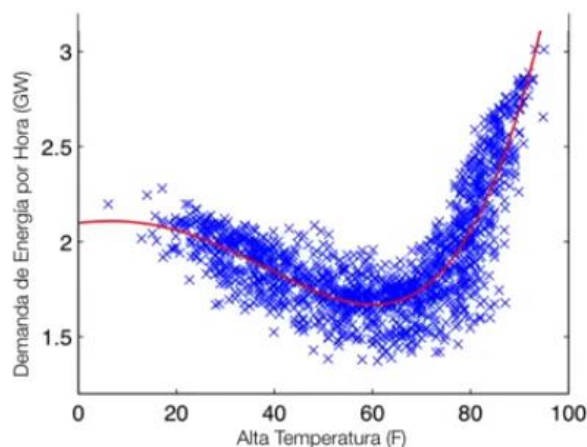


Figura 2. 10. Ejemplo regresión polinomial. Predicción de la demanda horaria de energía en función de la temperatura exterior. (González, 2018)

### 2.3.3. Regresión Logística

La Regresión Logística es un algoritmo de clasificación empleado para predecir el resultado de una variable categórica binaria en función de las variables independientes. Los objetivos de este modelo son: determinar la existencia o ausencia de relación entre una o más variables independientes y una variable dependiente dicotómica (que solo admite dos categorías que definen situaciones opuestas); medir el signo de dicha relación y predecir la probabilidad de que se produzca el suceso  $Y=I$  en función de los valores de las variables independientes (Salas, 1996).

La función logística es aquella que, para cada individuo, halla la probabilidad (P) de que se presente el efecto en cuestión. Para comprender el origen de esta función es necesario explicar el concepto de *Odd*, que es la razón entre la probabilidad de que un suceso ocurra y la probabilidad de que ese suceso no ocurra (Ecuación 2.7) (Fuiza & Rodríguez, 2000).

$$Odd = \frac{P}{1 - P} \quad (2.7)$$

La función *Logit* (Ecuación 2.8) es una transformación logarítmica sobre el *Odd* que convierte los valores de probabilidad del rango [0,1] en valores dentro del rango  $[-\infty, \infty]$  (Román, 2019).

$$Logit(P) = \ln \frac{P}{1 - P} \quad (2.8)$$

La función *Logit* se puede representar de forma lineal como una función similar a la empleada en Regresión Lineal Múltiple (Ecuación 2.9) (Román, 2019).

$$z = w^T x = w_0 + w_1 x_1 + \dots + w_m x_m \quad (2.9)$$

De esta forma, se obtiene que la inversa de la función *Logit* es la función logística buscada (función hipótesis), llamada función *sigmoide* (Ecuación 2.10) y representada en la *Figura 2.11*.

$$\phi(z) = \frac{1}{1 + e^{-z}} = h(x) = \frac{1}{1 + e^{-w^T x}} \quad (2.10)$$

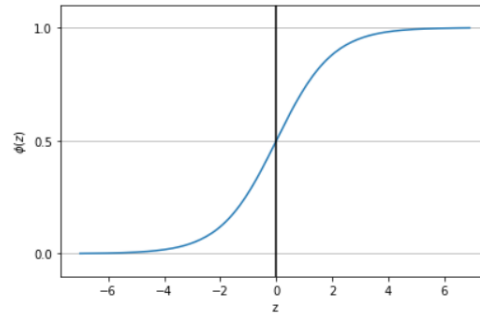


Figura 2. 11. Función sigmoide. (Román, 2019)

Esta función se encarga de tomar valores reales y transformarlos en un valor del rango [0,1], lo que indica la probabilidad de que una muestra pertenezca a la clase 1 dadas las características  $x$  parametrizadas por los pesos  $w$  (Román, 2019). Un cuantificador traduce la probabilidad en una salida binaria (Ecuación 2.11).

$$y = \begin{cases} 1 & \text{si } z \geq 0 \\ 0 & \text{para el resto} \end{cases} \quad (2.11)$$

Este modelo se utiliza cuando se desea investigar si una o varias variables explican una variable dependiente que toma un carácter cualitativo. Esto sucede a menudo en medicina, ya que se intenta responder a preguntas formuladas en base a la presencia o ausencia de una determinada característica que no es cuantificable, sino que representa la existencia o no de un efecto de interés (Fuiza & Rodríguez, 2000).

Por ejemplo, la Regresión Logística en medicina se emplea para predecir si un tumor es benigno o maligno en función de la edad del paciente y del tamaño del tumor (Kapur, 2016). El límite de decisión lo marca  $z = w^T x$ , por lo que se varía el grado de cada término para conseguir la curva que mejor se ajuste a los datos (Figura 2. 12).

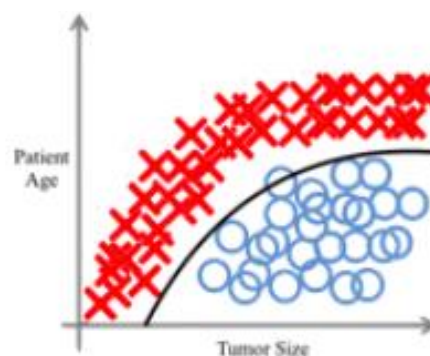
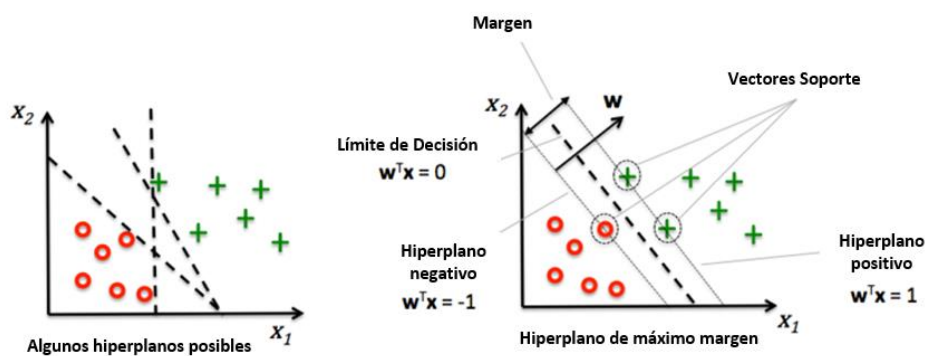


Figura 2. 12. Ejemplo regresión logística. Clasificación de un tumor en benigno o maligno en función de la edad del paciente y del tamaño del tumor (Kapur, 2016).

### 2.3.4. Support Vector Machine

El algoritmo de vectores de soporte (*Support Vector Machine*) es un clasificador discriminatorio cuyo objetivo es definir un hiperplano en un espacio N-dimensional (siendo N el número de variables independientes) que maximice la distancia entre datos de ambas clases (Gandhi, 2018). Es decir, el hiperplano es el límite de decisión que ayuda a clasificar los datos, por lo que a cada lado del hiperplano se le atribuye una clase.

Los vectores de soporte son aquellos que se encuentran más cerca del hiperplano, influenciando así la posición y la orientación de este. Encontrar un hiperplano que separe los datos no suele ser una solución única. Sin embargo, la solución óptima es la que maximiza la distancia entre dos líneas paralelas situadas simétricamente a cada lado del límite de decisión de forma que los vectores de soporte estén contenidos en ellas. Esta distancia se conoce como margen. Aquellos modelos con gran margen reducen el error de generalización, mientras que los modelos con poco margen tienden a ser menos propensos al *overfitting* (Gandhi, 2018). En la *Figura 2. 13* se representa el hiperplano escogido de entre todos los hiperplanos posibles, así como los diferentes elementos que se han comentado.



*Figura 2. 13. Representación gráfica de Support Vector Machine (Román, 2019)*

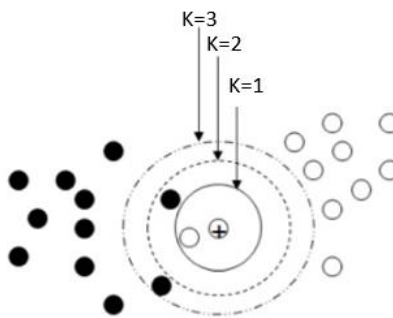
En la Regresión Logística el valor lineal de salida se transforma en un valor del rango  $[0,1]$  a través de la función *sigmoide*, donde para valores mayores de 0.5 se asigna la clase 1 y para valores menores la clase 0. En el algoritmo de vectores de soporte, si la salida lineal es mayor que 1 se identifica con una clase y si es menor que -1 se identifica con la otra (Gandhi, 2018). Por lo tanto, el rango que actúa como margen es  $[-1,1]$ .

Para problemas de clasificación no lineales se emplean técnicas de mapeo y kernelización. Estas técnicas reorganizan los datos a través de funciones matemáticas, ya sea en el mismo plano o en planos de dimensiones mayores, para que adopten una configuración que los haga separables (Román, 2019). Esta es una de las grandes ventajas de los vectores de soporte frente a otros modelos, junto con su importante precisión a costa de poca potencia de computación (Gandhi, 2018). Además, estos principios se pueden emplear para resolver problemas de regresión (*Support Vector Regression*).

### 2.3.5. K-Nearest Neighbors

El algoritmo de los K-vecinos más cercanos (*K-Nearest Neighbors*, K-NN) es un algoritmo de clasificación que etiqueta a cada dato en función de la etiqueta que tengan los datos más próximos a él. Para ello, se define la variable  $k$ , que corresponde al número de vecinos más cercanos que se escogen para llevar a cabo la clasificación (Ruiz, 2017). En función de esta variable se tienen unas predicciones u otras.

En el ejemplo de la *Figura 2. 14*, para  $k=1$  el algoritmo clasifica al elemento como blanco; para  $k=2$  el algoritmo necesita un criterio para clasificar puesto que hay un vecino de cada color; y para  $k=3$  el algoritmo clasifica al elemento como negro, ya que dos de los tres elementos cercanos son negros (voto mayoritario).



*Figura 2. 14. Ejemplo de K-NN. Clasificación en función del parámetro K (Ruiz, 2017).*

A diferencia de otros algoritmos de aprendizaje supervisado, K-NN no aprende con los datos de entrenamiento, sino que el aprendizaje se produce con los datos de test una vez se han memorizado los datos de entrenamiento. Estos algoritmos se conocen como algoritmos vagos (*lazy algorithms*), y permiten resolver simultáneamente varios problemas dado que la función objetivo se aproxima localmente para cada elemento. De esta forma, el algoritmo no aprende de un modelo, sino que utiliza los datos para generar una respuesta solo cuando se solicita una predicción. Como desventaja, el coste computacional es muy elevado debido al almacenamiento de todos los datos de entrenamiento (Román, 2019).

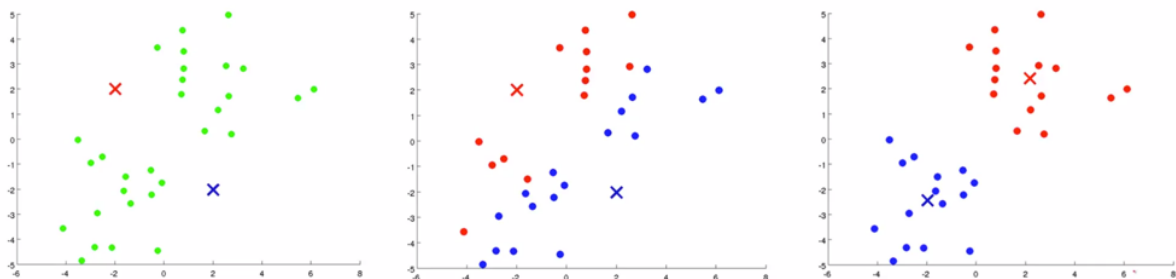
## 2.4. Aprendizaje no supervisado

En el aprendizaje no supervisado, el agente detecta los patrones existentes en los datos de entrada sin necesidad de observar cuál es la salida (Russell & Norvig, 2010). Por lo tanto, el objetivo es extraer información significativa de los datos de entrada, los cuales carecen de etiqueta y cuya estructura es desconocida. Existen dos tipos de problemas en el aprendizaje no supervisado: el agrupamiento o *clustering*, y la reducción dimensional. El agrupamiento se encarga de crear conjuntos de objetos de características similares, mientras que la reducción dimensional busca información redundante en los datos para reducir el número de variables y así mejorar el rendimiento computacional (Román, 2019).

Entre los algoritmos de agrupamiento se encuentran el *K-Means*, el Agrupamiento Jerárquico, el GMM (*Gaussian Mixture Models*), el *K-Modes* o el DBSCAN (*Density-Based Spatial Clustering of Applications with Noise*). Por otra parte, algoritmos como el Análisis de Componentes Principales (*Principal Component Analysis*, PCA), el Análisis de Componentes Independientes (*Independent Component Analysis*, ICA), el Análisis Discriminante Lineal (*Linear Discriminant Analysis*, LDA) o la Descomposición en Valores Singulares (*Singular Value Decomposition*, SVC) son algoritmos de reducción dimensional. A continuación, se explican algunos de los algoritmos mencionados.

### 2.4.1. K-Means

El algoritmo *K-Means* es el algoritmo de agrupamiento más popular. Se trata de un algoritmo iterativo de varias etapas. La primera consiste en definir la variable  $K$ , que es el número de clústeres. Después se escogen aleatoriamente  $K$  datos del set, denominados centroides. A cada dato se le asigna el centroide más próximo, obteniendo así  $K$  clases. A continuación, el centroide se desplaza hacia el punto que corresponda a la media de las distancias entre cada dato y su centroide. En esta nueva posición se vuelve a realizar la asignación entre el centroide y los puntos más cercanos. El proceso iterativo finaliza cuando el centroide apenas se mueve entre una iteración y otra, llegando así a la convergencia (Chandupatla, 2019). Una representación del proceso se muestra en la *Figura 2. 15*.



*Figura 2. 15. Explicación gráfica en varias etapas del algoritmo K-Means. (Ng, 2012)*

La función de coste que determina lo bueno que es este algoritmo se denomina *Distortion cost function*, y corresponde a la distancia media entre cada dato y el centroide asignado. En ocasiones, observando la distribución de los datos es fácil intuir el número de clústeres. Sin embargo, lo habitual es encontrar un conjunto de datos poco agrupados a simple vista o incluso de altas dimensiones. Por lo tanto, para escoger un valor de  $K$  adecuado suele emplearse el Método del Codo (*Elbow Method*). Este método consiste en representar el valor de la función de coste para cada valor de  $K$ , y escoger como  $K$  óptimo el codo de la función (*Figura 2. 16*) (Chandupatla, 2019).

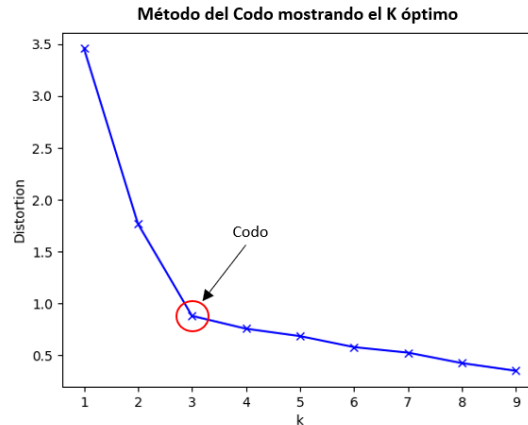


Figura 2. 16. Método del codo para la elección del  $K$  óptimo (Chandupatla, 2019).

Este algoritmo es rápido, robusto y simple cuando los datos son muy distintos o pueden separarse de forma lineal. En cambio, es muy sensible al valor de  $K$  y puede resultar poco efectivo cuando los datos están superpuestos o contienen mucha cantidad de ruido (González, 2018). Algunos ejemplos de aplicación del algoritmo  $K$ -Means pueden ser la segmentación del mercado o el agrupamiento de palabras con definiciones similares para la mejora de los motores de búsqueda (Chandupatla, 2019).

#### 2.4.2. Análisis de Componentes Principales

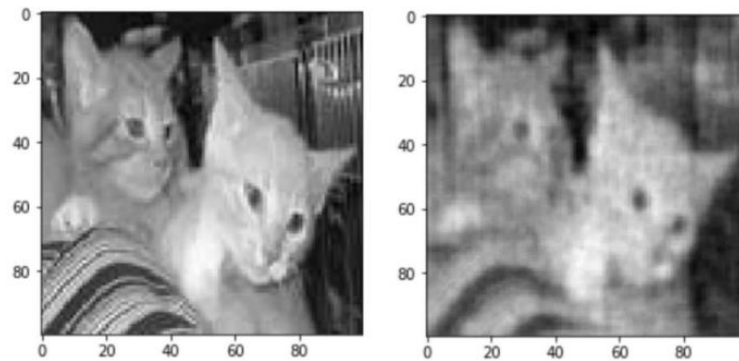
Dado un conjunto de datos, el Análisis de Componentes Principales (*Principal Component Analysis*, PCA) trata de encontrar los variables independientes más significativas que representen dicho conjunto de datos (Peres-Neto, Jackson, & Somers, 2004). Cuando la cantidad de variables y de datos es muy elevada, el rendimiento de los modelos disminuye debido al coste computacional que supone tratar con algunos datos que realmente no aportan información relevante. Por esta razón, se necesitan algoritmos de reducción dimensional (Chandupatla, 2019).

Matemáticamente, PCA reduce las  $n$  dimensiones iniciales en  $k$  dimensiones mediante la búsqueda de  $k$  vectores en los que proyectar los datos, minimizando el error de proyección. La minimización del error de proyección retiene los datos de mayor varianza, ya que son aquellos que juegan un papel importante en la determinación de la etiqueta de salida. Por el contrario, una baja varianza indica que el valor de esa variable no influye demasiado en la etiqueta que se intenta predecir (Chandupatla, 2019).

Esto no quiere decir que los datos de partida se ordenen de mayor a menor varianza y se eliminen los de varianza más baja, ya que de esa forma se estaría perdiendo gran cantidad de información, sino que los  $k$  vectores nuevos deberán contener la máxima varianza posible (Chandupatla, 2019). En realidad, se transforma el conjunto original de  $n$  variables en otro conjunto más pequeño de  $k$  nuevas variables no correlacionadas que se denominan “componentes principales” (Marín, 2013).



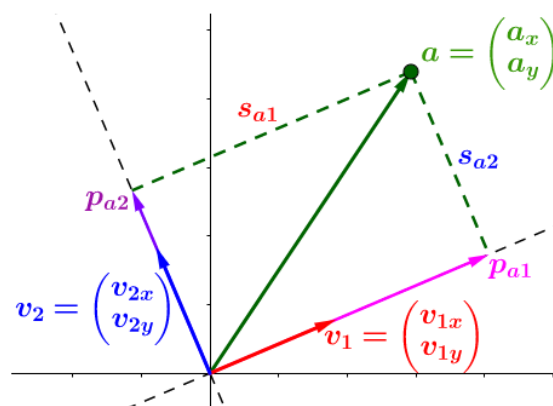
Este algoritmo es de gran utilidad en la compresión de imágenes, donde cada pixel corresponde a una variable. En el ejemplo de la *Figura 2. 17*, el objetivo es retener la información de los pixeles que contienen la etiqueta que se desea predecir, por lo que todos los pixeles que no contienen gatos son inútiles. El tiempo de computación de una red neuronal que trabaje con imágenes se reduce considerablemente gracias al PCA como etapa de preprocesamiento (Chandupatla, 2019).



*Figura 2. 17. Ejemplo de aplicación de PCA (Chandupatla, 2019).*

### 2.4.3. Singular Value Decomposition

En álgebra lineal, la Descomposición en Valores Singulares (*Singular Value Decomposition*, SVD) es un método de factorización de una matriz real o compleja usado para la reducción de dimensiones. Se basa en el principio de descomposición de vectores en sus ejes ortogonales (*Figura 2. 18*), de manera que cualquier vector  $a$  puede expresarse mediante dos variables: el vector unitario que indica la dirección de proyección ( $v_i$ ) y la longitud de la proyección ( $s_{ai}$ ). En SVD se extiende esta conclusión a muchos vectores y en todas las dimensiones (Abdullatif, 2019).



*Figura 2. 18. Descomposición de un vector en dos ejes ortogonales (Abdullatif, 2019).*

La proyección del vector  $a$  en la dirección de los vectores unitarios proporciona la longitud de las proyecciones, según la Ecuación 2.12 (Abdullatif, 2019).

$$a^T \cdot v_i = (a_x \ a_y) \cdot \begin{pmatrix} v_{ix} \\ v_{iy} \end{pmatrix} = s_{ai} \quad (2.12)$$

Empleando matrices se llega a la Ecuación 2.13, donde  $A$  es la matriz de puntos,  $V$  la matriz ortogonal de descomposición de ejes y  $S$  la matriz de longitud de las proyecciones (Abdullatif, 2019).

$$A = S V^{-1} = S V^T \quad (2.13)$$

En la Ecuación 2.14, la matriz  $S$  se descompone en las matrices  $U$  y  $\Sigma$  (Abdullatif, 2019).

$$S = U \Sigma \quad (2.14)$$

Donde  $U$  es la matriz de longitudes unitarias de proyección y  $\Sigma$  incluye los valores de  $\sigma_i$ , los cuales aumentan si los puntos se aproximan al eje del correspondiente  $\sigma_i$ , tal y como indica la Ecuación 2.15 (Abdullatif, 2019).

$$\sigma_i = \sqrt{(s_{ai})^2 + (s_{bi})^2} \quad (2.15)$$

Finalmente, la expresión del método SVD es el mostrado en la Ecuación 2.16 (Abdullatif, 2019).

$$A = U \Sigma V^T \quad (2.16)$$

La principal aplicación de este método es el de escoger unos vectores unitarios (matriz  $V$ ) que coincidan con los componentes principales del conjunto de datos (matriz  $A$ ). Así, si se conoce la línea de máxima varianza, pueden proyectarse los puntos en ella, como se representa en la *Figura 2. 19* (Abdullatif, 2019). El SVD también se aplica en los sistemas de recomendaciones encontrando similitudes entre usuarios y productos, solo que realmente no es capaz de dar una explicación al usuario de por qué ha recomendado ese producto (Huang, 2018).

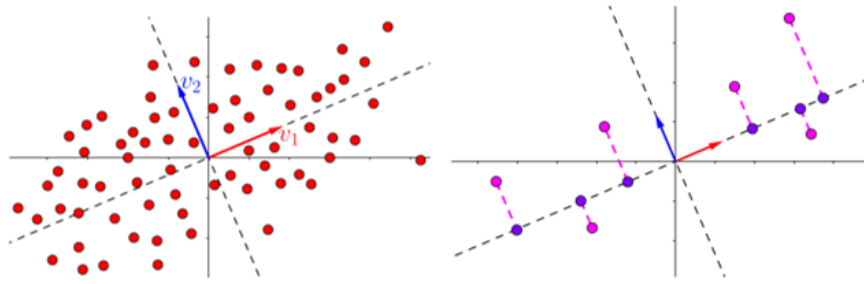


Figura 2. 19. Representación gráfica del algoritmo SVD. Proyección de los puntos sobre la línea de máxima varianza (Abdullatif, 2019).

#### 2.4.4. Análisis de Componentes Independientes

El Análisis de Componentes Independientes (*Independent Component Analysis*, ICA) es una técnica computacional y estadística empleada para revelar factores ocultos dentro de un conjunto de variables, medidas o señales. El objetivo es descomponer una señal multivariable en subcomponentes aditivos asumiendo que son señales no Gaussianas y estadísticamente independientes. Las variables encontradas por ICA se denominan “componentes independientes” (Hyvärinen, Karhunen, & Oja, 2004).

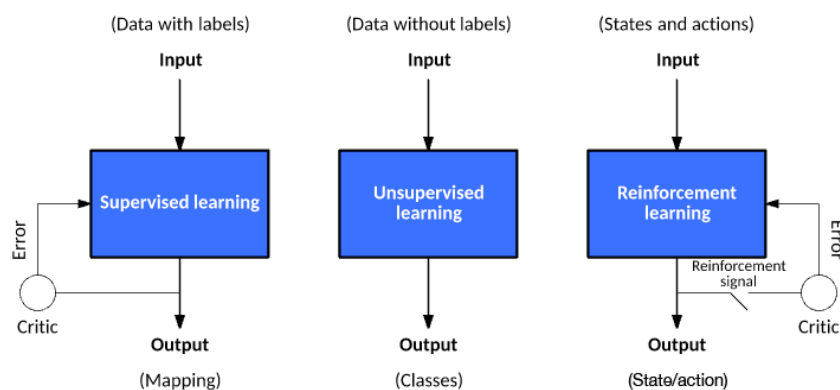
Esta técnica de reducción dimensional, extensión del PCA, es mucho más potente y funciona bien en casos donde los métodos clásicos fallan por completo (Hyvärinen, Karhunen, & Oja, Independent Component Analysis, 2004). Esto se debe a que PCA solo puede imponer independencia hasta el segundo orden y, en consecuencia, definir direcciones ortogonales (Comon, 1994). La restricción fundamental en ICA es que los componentes independientes sean no Gaussianos. La distribución multivariable de dos variables Gaussianas independientes es simétrica y, por lo tanto, no contiene información sobre las direcciones además de la ortogonalidad. No obstante, en la mayoría de las teorías estadísticas se asume que las variables aleatorias siguen distribuciones Gaussianas, impidiendo la aplicación de ICA (Hyvärinen & Karhunen, 2000).

El ICA es un tipo particular de separación ciega de fuentes (*blind source separation*), pues obtiene las señales procedentes de  $n$  fuentes a partir de las mezclas recogidas por  $m$  sensores. Un ejemplo muy común en su aplicación en audio es el *cocktail party problem*, que ocurre cuando varios locutores hablan simultáneamente. La idea es centrarse únicamente en un locutor para conseguir extraer su voz y eliminar el ruido de fondo u otras conversaciones. Este proceso se puede modelar como una mezcla lineal seguido de un filtrado de fuentes, de lo que se encarga el ICA mediante la recuperación ciega de dichas fuentes suponiendo que las señales originales son independientes (Aldecoa, Carné, & Monte, 2003).

## 2.5. Aprendizaje reforzado

El aprendizaje reforzado (*reinforcement learning*) es una forma de aprendizaje automático basado en un sistema de recompensas y castigos en el que un agente busca las decisiones óptimas para obtener la máxima recompensa tanto a corto como a largo plazo. La recompensa mide la contribución de la acción en el objetivo final. Este agente necesita saber si la decisión que ha tomado es buena o mala, pero en lugar de tener un supervisor como en el aprendizaje supervisado, este aprende de su propia experiencia a partir de las interacciones con el entorno. El marco que define la interacción entre el agente y el entorno es un Proceso de Decisión de Markov (*Markov Decision Process*, MDP), por el cual cada estado depende únicamente del estado anterior (Russell & Norvig, 2010).

A diferencia del aprendizaje no supervisado, el aprendizaje reforzado trata de maximizar la función de recompensa (*reward function*) en lugar de encontrar ciertos patrones ocultos en un conjunto de datos no etiquetados. Además, estas recompensas no tienen por qué ofrecerse instantáneamente, sino que pueden llegar de forma retardada (Caminero, 2018). En la *Figura 2. 20* se representan esquemáticamente los tres tipos de aprendizaje: supervisado, no supervisado y reforzado.



*Figura 2. 20. Representación esquemática de los tres tipos de aprendizaje. Recuperado de: <https://developer.ibm.com/articles/cc-models-machine-learning/>*

El principal desafío del aprendizaje reforzado es conseguir el equilibrio entre la exploración y la explotación (*exploration and exploitation dilema*). Para obtener altas recompensas el agente tiende a repetir aquellas acciones que han resultado ser positivas. Sin embargo, para poder descubrir dichas acciones se requiere probar aquellas que no se han probado anteriormente. Por lo tanto, el agente explota el conocimiento actual al mismo tiempo que explora distintas opciones (Caminero, 2018).

Cada algoritmo de aprendizaje reforzado debe seguir una política (*policy*) para decidir qué decisión tomar en función del estado (*state*) en el que se encuentre. No obstante, esta política puede no seguirse en la etapa de aprendizaje. Aquellos algoritmos cuya regla de actualización (*update rule*) realiza la acción que traerá el máximo beneficio a pesar de que la política actual restrinja dicha acción, se denominan algoritmos *off-policy*. Por el contrario, los algoritmos *on-policy* son aquellos que siguen estrictamente la política (Zychlinski, 2019).

Cuando un agente intenta optimizar su política, puede hacerlo desde dos enfoques diferentes. Siguiendo un enfoque *Model-Based*, el agente calcula las probabilidades de éxito de todos los estados antes de comenzar, con el fin de actuar en función de dichas probabilidades y modelando así el entorno. Con un enfoque *Model-Free*, el agente obtiene información sin previo conocimiento del entorno a partir de prueba y error (Zychlinski, 2019).

### 2.5.1. Q-learning

*Q-learning* es un algoritmo *off-policy* de aprendizaje reforzado que busca la mejor acción dado el estado actual. Se considera *off-policy* porque la función aprende de las acciones que están fuera de la política actual, como por ejemplo las acciones aleatorias. El objetivo es aprender la política que maximice la recompensa total. La Q de *Q-learning* viene de *quality*, que en este caso mide lo útil que ha sido una acción para ganar una recompensa futura (Sutton & Barto, 2018; Violante, 2019).

El algoritmo crea lo que se denomina *Q-table* o matriz (Figura 2. 21), que representa los estados frente a las acciones. Esta matriz se inicializa con todos los valores a cero, y con cada episodio se van actualizando los Q-valores de todas las posibles parejas estado-acción. De este modo, la matriz se convierte en una referencia para que el agente seleccione la mejor acción basándose en los Q-valores (Violante, 2019).

| Q-table  | Acción 1 | Acción 2 | Acción 3 |
|----------|----------|----------|----------|
| Estado 1 | 0        | 10       | 0        |
| Estado 2 | 0        | 0        | -1       |
| Estado 3 | -5       | 0        | 0        |
| Estado 4 | 3        | 0        | 12       |

Figura 2. 21. Ejemplo de Q-table o matriz. Elaboración propia.

En este punto, el agente puede interactuar con el entorno de dos formas. La primera consiste en la explotación del conocimiento adquirido, por lo que la siguiente acción será la que tendrá un Q-valor máximo. La segunda consiste en la exploración, donde el agente actúa aleatoriamente para añadir nuevos estados a la tabla. El parámetro  $\epsilon$  equilibra el porcentaje de acciones de explotación y de exploración (Huang, 2018; Violante, 2019).

La Ecuación 2.17 representa la regla de actualización de este algoritmo (Huang, 2018).

$$Q(s_t, a_t) \leftarrow Q(s_t, a_t) + \alpha[r_{t+1} + \gamma \max_a Q(s_{t+1}, a) - Q(s_t, a_t)] \quad (2.17)$$

Esta regla ajusta los Q-valores basándose en la diferencia entre los nuevos valores descontados y los antiguos valores. Se descuentan los nuevos valores con el factor de descuento ( $\gamma$ ) y se ajusta el tamaño del paso con la tasa de aprendizaje ( $\alpha$ ). La recompensa ( $r$ ) es el valor recibido tras completar una acción en un determinado estado. Por lo tanto, la siguiente acción se escoge para maximizar el Q-valor del siguiente estado en lugar de seguir la política actual (Huang, 2018; Violante, 2019).

Una versión más compleja y popular es la del *Deep Q-Network*, la cual reemplaza la matriz estado-acción por una red neuronal para poder trabajar con cantidades masivas de datos y realizar tareas más complejas. Una de las aplicaciones de estos algoritmos es la de hacer que los computadores aprendan a jugar a videojuegos incluso mejor que los humanos (Violante, 2019).

### 2.5.2. SARSA

SARSA (*State-Action-Reward-State-Action*) es un algoritmo de aprendizaje reforzado que aprende una política basada en un Proceso de Decisión de Markov, al igual que el *Q-learning*. La diferencia es que SARSA es un algoritmo *on-policy*, es decir, los Q-valores aprendidos siguen siempre la política actual. La máxima recompensa para el estado siguiente no es necesariamente usada para actualizar los Q-valores. En lugar de eso, se selecciona una nueva acción y una nueva recompensa usando la misma política que determinó la acción original (Huang, 2018; Sutton & Barto, 2018).

La Ecuación 2.18 representa la regla de actualización de este algoritmo (Huang, 2018).

$$Q(s_t, a_t) \leftarrow Q(s_t, a_t) + \alpha[r_{t+1} + \gamma Q(s_{t+1}, a_{t+1}) - Q(s_t, a_t)] \quad (2.18)$$

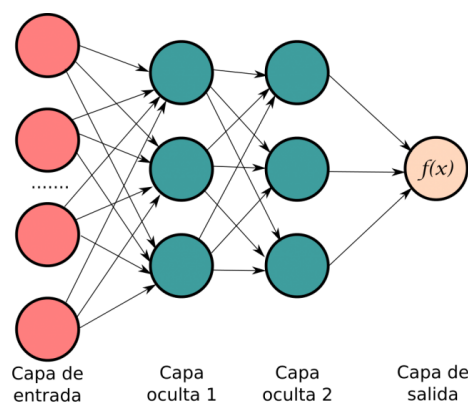
Esta regla depende del estado actual, de la acción actual, de la recompensa obtenida, del estado siguiente y de la acción siguiente. Esta quintupla ( $s, a, r, s', a'$ ) da el nombre al algoritmo. Tener en cuenta la acción siguiente significa que se requiere almacenar la información durante más tiempo antes de actualizar los valores (Huang, 2018).

La convergencia del método depende de la naturaleza de la política, de su dependencia con Q y de la tasa de aprendizaje. SARSA converge con probabilidad 1 a una política óptima si todos los pares estado-acción son visitados un número infinito de veces. En general, SARSA tiene una convergencia más rápida, pero es más propenso a quedarse atascado en un mínimo local. En cambio, *Q-learning* tiene un mejor desempeño final, pero necesita más tiempo de aprendizaje. Una combinación de ambos es el *Backward Q-learning*, que consigue aumentar la velocidad de aprendizaje al mismo tiempo que mejora el rendimiento final (Wang, Li, & Lin, 2013).

## 2.6. Aprendizaje profundo

En los últimos años ha emergido un área específica del *Machine Learning* llamada *Deep Learning* o Aprendizaje Profundo, que se basa en algoritmos de aprendizaje en múltiples niveles de representación y de abstracción con el fin de modelar relaciones más complejas entre los datos. Los niveles se corresponden con distintos niveles de conceptos, donde los más altos vienen definidos por los más bajos de forma jerárquica. Esta idea de representaciones sucesivas por capas es lo que da el nombre de “profundo” a este tipo de aprendizaje, siendo la profundidad del modelo el número de capas que contiene (Deng & Yu, 2013).

En *Deep Learning*, estas representaciones por capas forman una Red Neuronal (*Neural Network*) que, al fin y al cabo, es un conjunto estructurado de neuronas (*Figura 2. 22*). Estos términos provienen de la neurobiología ya que estas redes se inspiran en el funcionamiento del cerebro. Las capas están formadas por un número determinado de neuronas, donde cada neurona recibe información de la capa anterior por medio de estímulos externos a través de sus conexiones de entrada. Las neuronas realizan cálculos internamente y devuelven un valor de salida que se transmite a las neuronas de la capa siguiente. En la última capa, la salida es la predicción buscada (Chollet, 2017).



*Figura 2. 22. Arquitectura básica de una red neuronal. Recuperado de: <https://naukas.com/2015/12/09/acertando-quinielas-redes-neuronales/>*

Sin embargo, para llegar a una buena predicción, la red debe ser capaz de minimizar el error entre la predicción y el valor esperado mediante el autoajuste de los parámetros del modelo. La neurona asigna un peso ( $w$ ) a cada una de sus variables de entrada ( $x$ ) para reflejar la intensidad con la que cada variable afecta a la salida (Ecuación 2.19) (Villanueva, 2019).

$$w_0x_0 + w_1x_1 + w_2x_2 + \dots + w_Nx_m \quad (2.19)$$

Tras calcular la combinación lineal de los pesos y las entradas, se aplica una función de activación ( $\phi$ ) cuyo objetivo es distorsionar la salida ( $y$ ), es decir, hacerla no lineal para que la red pueda resolver problemas no lineales (Ecuación 2.20) (Villanueva, 2019).

$$y = \phi(w_0x_0 + w_1x_1 + w_2x_2 + \dots + w_Nx_m) \quad (2.20)$$

Algunas de las funciones de activación más usadas en redes neuronales son la función escalón, la función *sigmoide*, la función rectificadora lineal y la función tangente hiperbólica (Figura 2. 23) (Villanueva, 2019).

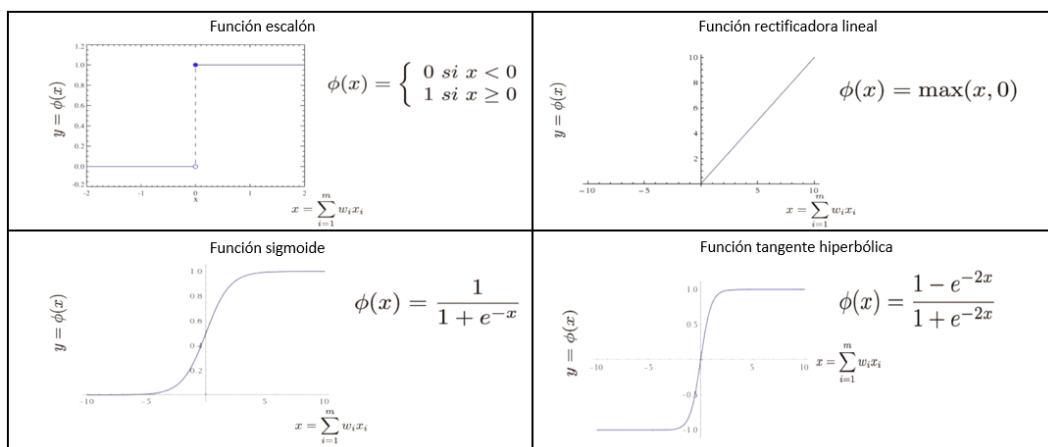


Figura 2. 23. Funciones de activación más usadas (Villanueva, 2019).

Una representación esquemática de la neurona se muestra en la Figura 2. 24.

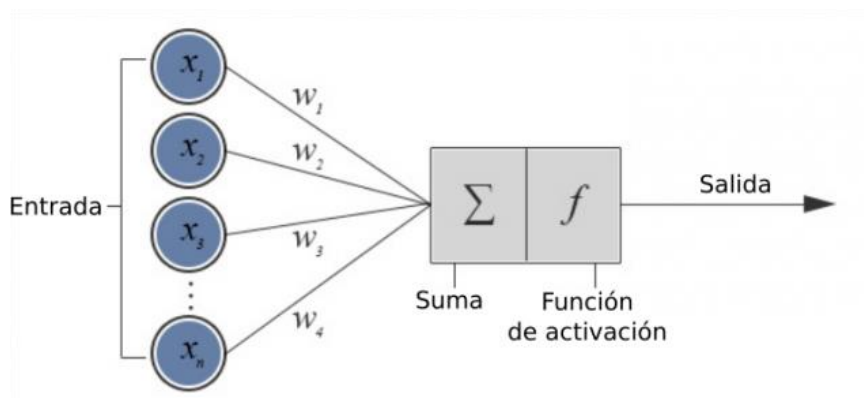


Figura 2. 24. Representación esquemática de una neurona. Recuperado de: <https://naukas.com/2015/12/09/acertando-quinielas-redes-neuronales/>



*Backpropagation* (propagación hacia atrás) es el mecanismo por el cual las redes neuronales aprenden, transmitiendo información sobre el error desde las últimas capas hasta las primeras y modificando así sus parámetros. Por lo tanto, la red neuronal sabe el peso o la influencia que ha tenido cada neurona en ese resultado erróneo. El problema de este mecanismo es que el error se va diluyendo a medida que se propaga entre capa y capa, así que en redes muy profundas solo se entrenan las últimas capas mientras que las primeras apenas sufren cambios (Chollet, 2017; Deng & Yu, 2013).

La ventaja del aprendizaje en *Deep Learning* es que permite que todas las capas del modelo aprendan al mismo tiempo, en lugar de sucesivamente. Cuando el modelo ajusta un parámetro, el resto de los parámetros que dependen de él se ajustan automáticamente. De esta forma, todo está supervisado bajo una única señal de realimentación (Chollet, 2017).

Otras propiedades importantes que hacen del *Deep Learning* una revolución en el mundo de la inteligencia artificial son la simplicidad, la escalabilidad y la versatilidad. La simplicidad se debe a la automatización del *feature engineering*, esto es, el proceso de creación de características que tanto tiempo y esfuerzo supone en las etapas previas al desarrollo de un modelo. La escalabilidad permite trabajar con conjuntos de datos de diferentes tamaños, mientras que la versatilidad hace que algunos modelos ya entrenados se puedan emplear para otros propósitos (Chollet, 2017).

El *Deep Learning* se emplea para una gran cantidad de aplicaciones, como por ejemplo la identificación de clientes potenciales para una empresa, el análisis de imágenes y datos médicos para predecir enfermedades, la detección y prevención de amenazas en tiempo real en el ámbito de la ciberseguridad, la generación de textos y el procesamiento del lenguaje, o la identificación de varios objetos en imágenes y vídeos. En el Capítulo 3 de este documento se profundiza sobre las aplicaciones de esta tecnología en la industria.

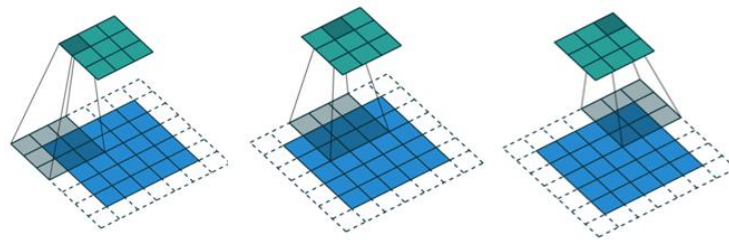
Existen múltiples tipos de redes neuronales, cada una diseñada para unos usos específicos y distintos niveles de complejidad. Escoger un tipo u otro depende también del tipo de datos con los que se va a trabajar. No obstante, el objetivo de todas ellas es encontrar patrones en los datos para llevar a cabo una tarea concreta como puede ser la clasificación, el agrupamiento o la predicción de un valor (Dormehl, 2019). A continuación, se explican tres tipos de redes neuronales: las *Convolutional Neural Networks*, las *Recurrent Neural Networks* y las *Generative Adversarial Networks*.

### 2.6.1. Convolutional Neural Networks

Una Red Neuronal Convolutiva (*Convolutional Neural Network*, CNN) es un algoritmo de *Deep Learning* empleado normalmente para clasificar imágenes y agruparlas en función de sus características. Las CNN realizan el reconocimiento óptico de caracteres (*Optical Character Recognition*, OCR) para digitalizar texto y posibilitar el procesamiento del lenguaje natural. Se aplican también en la identificación de sonidos a través de espectrogramas (Krizhevsky, Sutskever, & Hinton, 2012; Saha, 2018).

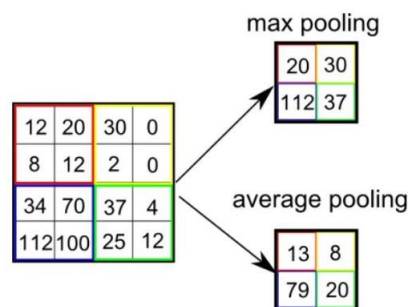
Los avances en visión artificial aplicados a los coches autónomos, a los drones, a la seguridad o a los diagnósticos médicos han aumentado mucho en los últimos años gracias al reconocimiento de imágenes a través de las CNN. La ventaja principal de las CNN frente a otros sistemas de identificación de imágenes es la capacidad de aprender los filtros necesarios para pre-procesar la imagen (*automated feature engineering*), en lugar de realizarlo manualmente (Fernández, 2018; Saha, 2018).

El objetivo de una CNN es reducir las imágenes para facilitar su procesado sin perder información crítica. Para ello, lo primero es aplicar capas convolutivas para detectar los patrones en los datos que proporciona la imagen. Un filtro en las primeras capas convolutivas de la red se encarga de extraer las características de menor nivel, mientras que en las capas posteriores se extraen características de mayor nivel con el fin de obtener una red que permita comprender la totalidad de la imagen. Pueden aplicarse varios filtros, donde cada uno de ellos detecta tipos diferentes de patrones. El paso o *stride* es la cantidad de píxeles que recorre un filtro en cada iteración. El número de filtros empleados equivale a la profundidad de la CNN (Krizhevsky, Sutskever, & Hinton, 2012; Saha, 2018). En la *Figura 2. 25* se observan tres iteraciones de un filtro de 3x3x1 de paso dos sobre una imagen de 5x5x1.



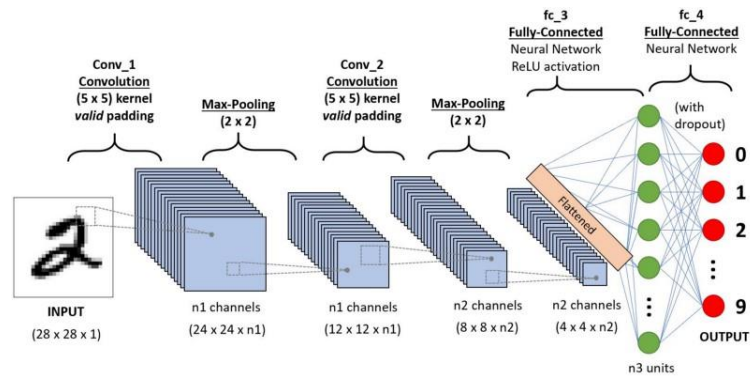
*Figura 2. 25. Recorrido de un filtro (verde) sobre una imagen (azul) en una CNN (Saha, 2018)*

Además de las capas convolutivas, existen otras capas llamadas *Pooling*, que son las responsables de reducir dimensionalmente el tamaño de la imagen. Estas capas transforman una matriz inicial en una matriz de menor tamaño siguiendo una regla determinada. Cuando la regla es crear una matriz con los valores máximos contenidos en el filtro, se habla de *Max Pooling*; mientras que, si se toma el valor promedio de los valores contenidos en el filtro, se trata de *Average Pooling*. Suele preferirse el *Max Pooling* por su mejor capacidad para eliminar ruidos (Fernández, 2018; Saha, 2018). Un ejemplo de ambos métodos se observa en la *Figura 2. 26*.



*Figura 2. 26. Ejemplo de Max Pooling y Average Pooling (Saha, 2018).*

Añadir una capa totalmente conectada (*fully-connected layer*) sirve para aprender combinaciones no lineales de las características gracias a la aplicación de una función de activación, habitualmente ReLU. Tras el entrenamiento, el modelo es capaz de distinguir cuáles son las características dominantes y clasificar así la imagen (Saha, 2018; Simard, Steinkraus, & Platt, 2003). La *Figura 2. 27* representa un ejemplo del proceso completo.

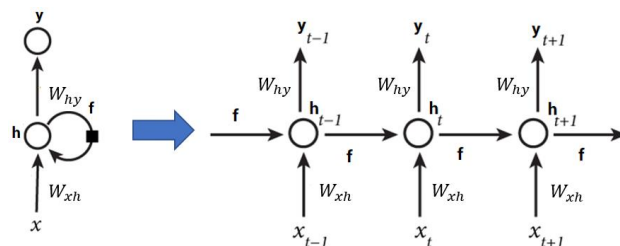


*Figura 2. 27. Proceso completo del funcionamiento de una CNN (Saha, 2018)*

## 2.6.2. Recurrent Neural Networks

La predicción de datos secuenciales supone un problema en *Machine Learning*, ya que habitualmente se analizan casos aislados como una imagen o un carácter que se desea clasificar. Sin embargo, lo que una persona entiende al ver una película o al mantener una conversación, se basa en conocimientos anteriores que van adquiriendo sentido con cada instante que transcurre. Las Redes Neuronales Recurrentes (Recurrent Neural Networks, RNN) son un tipo de algoritmo de Deep Learning que se encarga de resolver este problema mediante el procesamiento de datos secuenciales (Goodfellow, Bengio, & Courville, 2016).

Las RNN, en lugar de tener una estructura por capas como las redes neuronales convencionales, poseen una estructura cíclica donde el output de un estado pasa a ser un input del estado siguiente (*Figura 2. 28*). Esto permite detectar dependencias temporales, es decir, dotar a la red de memoria (estado oculto o *hidden state*) para que retenga la información sobre lo calculado en la etapa anterior. De esta forma, utiliza los mismos parámetros para cada entrada, reduciendo así la complejidad de la red (Zaremba, Sutskever, & Vinyals, 2015).



*Figura 2. 28. Ejemplo de una RNN, equivale a un conjunto de neuronas normales concatenadas. Recuperado de: <https://www.analyticsvidhya.com/blog/2017/12/introduction-to-recurrent-neural-networks/>*

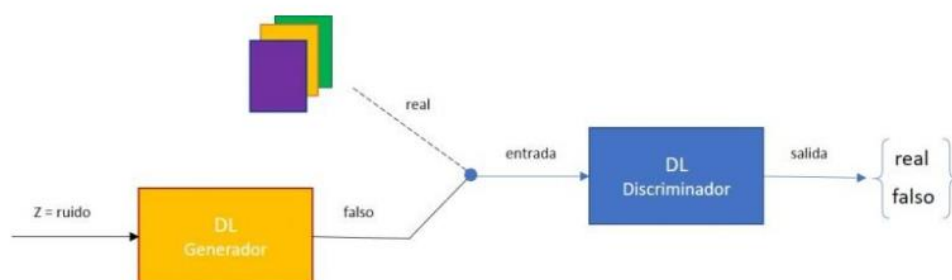
La desventaja de guardar tanta cantidad de información es que, en muchas ocasiones, es difícil identificar qué información de los pasos anteriores es relevante para el estado actual. Este problema se denomina desvanecimiento de gradiente (*vanishing gradient*), y se soluciona con un tipo de arquitectura de red: las LSTM (*long short-term memory*). Una red LSTM es capaz de decidir qué información es retenida y cuál es olvidada (Garbisu, 2016).

Estas redes emplean su propia versión de backpropagation, llamada *backpropagation through time* (BTT), que aplica el principio de *backpropagation* en los diferentes instantes de tiempo como si se tratase de capas de profundidad sucesivas. El número de repeticiones de este proceso viene dado por  $K$ , llegando la memoria de la red más atrás en el tiempo con mayores valores de  $K$  (Garbisu, 2016; Mozer, 1989).

Las aplicaciones de las RNN se centran en el procesamiento de secuencias de texto e imagen. Se emplean en sistemas de reconocimiento de voz, en la comprensión y generación de textos y conversaciones, en traductores automáticos e incluso en la detección de palabras clave que expresen opiniones o sentimientos determinados (Mozer, 1989; Zaremba, Sutskever, & Vinyals, 2015).

### 2.6.3. Generative Adversarial Networks

Una Red Generativa Adversaria (*Generative Adversarial Network*, GAN) es un tipo de algoritmo de *Deep Learning* capaz de aprender a generar nuevos conjuntos de datos, habitualmente imágenes. Estas redes están formadas, a su vez, por dos redes: una generadora y otra discriminadora. La generadora realiza un proceso llamado deconvolución, pues genera una imagen a partir de un vector de números, es decir, al contrario que las CNN. Por otro lado, la discriminadora es la que se alimenta con imágenes reales para clasificar la imagen generada como verdadera o falsa (Goodfellow, 2014; Martínez, 2019). El esquema de una GAN se muestra en la *Figura 2. 29*.



*Figura 2. 29. Esquema de una GAN (Martínez, 2019)*

Las redes generadora y discriminadora son adversarias, ya que su mecanismo de aprendizaje es como un juego en el que ambas compiten por mejorar. La red generativa aprende las razones por las cuales la red discriminadora ha rechazado su imagen, y actúa en consecuencia para generar imágenes más realistas. El problema de las GAN es que, además de generar imágenes de baja resolución, el aprendizaje es lento e inestable (Goodfellow, 2014; Martínez, 2019).

No obstante, en la publicación de NVIDIA titulada “*Progressive Growing of GANs for Improved Quality, Stability, and Variation*” (Karras, Aila, Laine, & Lehtinen, 2018), se demuestran los resultados de aplicar un aprendizaje adaptativo a las GAN. Esto supone comenzar el aprendizaje generando imágenes de baja resolución e ir aumentando la calidad de la imagen a medida que la red va adquiriendo cierto nivel de conocimiento. En consecuencia, la imagen realiza un ajuste más fino en las últimas fases del aprendizaje, reduciendo así los tiempos y el poder de cómputo. Algunos ejemplos de rostros generados automáticamente por esta GAN se muestran en la Figura 2. 30.



Figura 2. 30. Ejemplos de rostros generados por una GAN (Karras, Aila, Laine, & Lehtinen, 2018)

En un estudio realizado por la Universidad de Berkeley (Zhu, Park, Isola, & Efros, 2018) se presenta un tipo de GAN llamado CYCLEGAN, cuyo objetivo es transformar cualquier imagen de un estilo a otro. Mediante aprendizaje no supervisado, se alimenta a la red con imágenes de un mismo estilo. La red detecta las características de ese estilo y lo aplica a las imágenes que se introduzcan posteriormente en la red. Para ello se crea un ciclo formado por dos GAN: la primera traduce la imagen al estilo deseado, y la segunda realiza la traducción en sentido inverso para verificar que el resultado es válido Figura 2. 31.

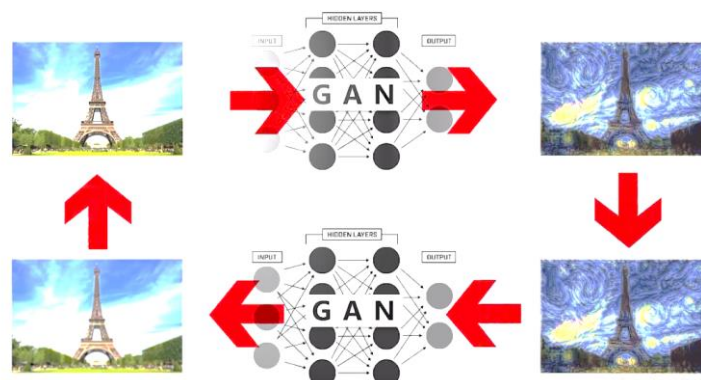


Figura 2. 31. Transformaciones en una CYCLEGAN (Santana, 2019)



## 3 APLICACIONES DEL MACHINE LEARNING EN LA INDUSTRIA

La Industria 4.0 introduce las tecnologías de la información en la industria con el objetivo de llevar a cabo una transformación digital que las haga inteligentes. Gracias al uso de los sistemas ciber físicos, del *Internet Of Things* (IoT), del *Cloud Computing* y del *Big Data*, es posible extraer y almacenar una gran cantidad de datos. Sin embargo, se requiere la aplicación de algoritmos de *Machine Learning* para crear modelos que den valor a esos datos y faciliten la toma de decisiones.

Debido al aumento del poder computacional y al abaratamiento del *hardware*, las investigaciones y las aplicaciones relacionadas con el *Machine Learning* en la industria son cada vez más numerosas y desafiantes. En este capítulo se presentan algunas razones por las que el *Machine Learning* es una de las tecnologías más relevantes en el desarrollo de la cuarta revolución industrial. También se muestra cómo se está aplicando el *Machine Learning* en las diferentes áreas de una empresa industrial como la logística, el mantenimiento o los recursos humanos. Finalmente, se exponen algunas de las limitaciones actuales de dicha tecnología.

### 3.1. La importancia del Machine Learning en la industria

El objetivo de toda industria es ofrecer productos o servicios de alta calidad con el mínimo coste. Los fabricantes están alcanzando este propósito con más éxito que nunca gracias a las fábricas inteligentes, en las cuales una de las tecnologías principales es la inteligencia artificial y, en particular, el *Machine Learning*. Unido a la facilidad de obtención y almacenamiento de datos, esta tecnología tiene un impacto directo en la mejora de la eficiencia de los sistemas productivos, la calidad de los productos y la seguridad de las personas (Seebo, 2019).

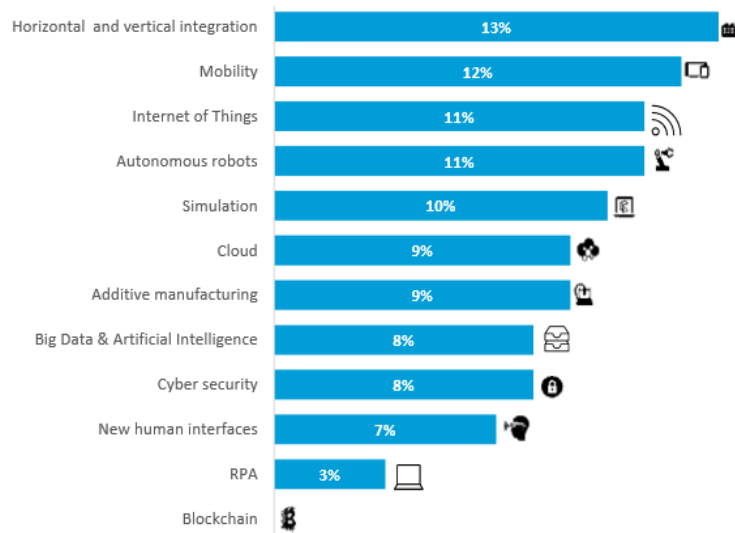
Un artículo de El País (Martín, 2017) afirmó que el aprendizaje automático sería una tecnología clave en 2018 por el hecho de cambiar la forma en que vivimos y trabajamos con el mismo impacto que supuso la llegada de *Internet*. El rol de las personas en ciertos trabajos se adapta de manera que el aprendizaje automático de soporte a los humanos y los haga más eficientes. De hecho, según (Gartner, 2017), el mayor beneficio se obtiene con el trabajo colaborativo entre humanos y máquinas en lugar de con cualquiera de ellos por separado.

Las empresas que aprovechen el poder del aprendizaje automático se verán altamente beneficiadas, ganando una ventaja competitiva frente al resto. Su habilidad para reducir los tiempos de ejecución y toma de decisiones, junto a la mejora en la efectividad de todos los procesos de todos los sectores hace que ningún negocio pueda permitirse el lujo de dejar pasar esta oportunidad (Martín, 2017).

Así como el *Big Data* y el *Cloud Computing* son tecnologías muy potentes y necesarias para la toma y el almacenamiento de datos, el *Machine Learning* es la herramienta verdaderamente industrializada que transforma dichos datos en valor para la empresa (Martín, 2017). Además,

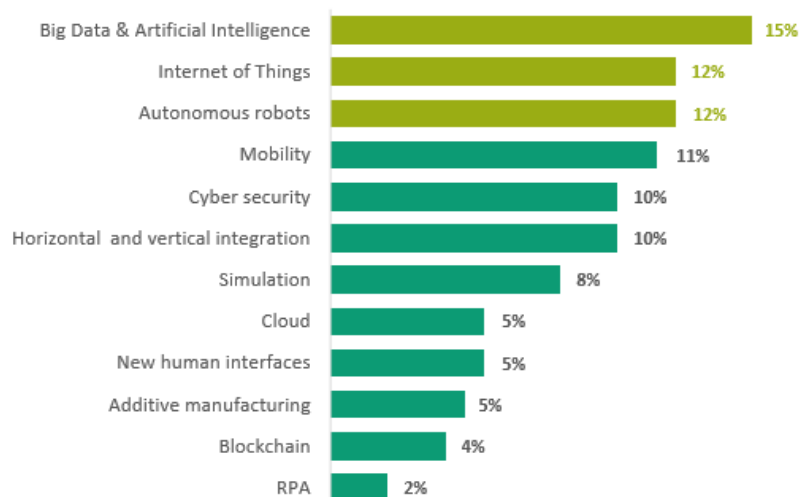
su uso se extiende a cualquier situación o proceso donde se deseen identificar tendencias, realizar predicciones, detectar anomalías, etc.

Según el *Estudio Smart Industry 4.0* realizado por (everis Spain, 2018), solo el 8% de las empresas disponían en 2018 de aplicaciones industriales relacionadas con el *Big Data* y la inteligencia artificial (*Figura 3. 1*).



*Figura 3. 1. Paradigmas más punteros actualmente en las empresas (everis Spain, 2018).*

No obstante, las empresas son conscientes de la cantidad de datos disponibles en sus procesos, y por ello el 15% de las empresas muestra gran interés en focalizar sus esfuerzos e inversiones hacia estas tecnologías (*Figura 3. 2*).



*Figura 3. 2. Tecnologías en las que las empresas invertirán más en los próximos años (everis Spain, 2018).*



Entre los beneficios directos de la aplicación del *Machine Learning* en las fábricas se encuentra la reducción de costes mediante mantenimiento predictivo. De esta manera, disminuye el nivel del inventario y la cantidad de chatarra gracias al mejor cuidado de las máquinas. Con *Machine Learning* también se consiguen optimizar los flujos de producción y de transporte, así como focalizar la atención en el cliente para responder rápido ante los cambios en el mercado (Seebo, 2019).

Además de las numerosas ventajas que aporta el *Machine Learning* en la industria a nivel estratégico y operacional, apostar hoy por este tipo de soluciones implica, a su vez, aspectos positivos transversales. Uno de ellos está relacionado con la rápida evolución de la inteligencia artificial y la correspondiente escasez de profesionales en el área de la ciencia de datos. La alta demanda de este perfil tecnológico hace que las universidades, las plataformas de educación *on-line* y las empresas atraigan a estudiantes motivados para formarles ellos mismos según unas necesidades tecnológicas específicas (Brynjolfsson & McAfee, 2017).

Otro aspecto fundamental es la posibilidad de elegir si comprar o alquilar *hardware* para el desarrollo de *Machine Learning*. Empresas como Google, Microsoft y Amazon disponen de una infraestructura de *Machine Learning* accesible desde *La Nube*. De esta forma, las empresas pueden ajustarse a sus necesidades y comprar únicamente lo que saben que van a utilizar (Brynjolfsson & McAfee, 2017).

Por último, el acceso a un conjunto de datos apropiado para alimentar los algoritmos de *Machine Learning* suele ser un problema para muchas empresas. Esto se debe a que, en ocasiones, las empresas piensan que solo se pueden obtener mejoras considerables con un número masivo de datos. Sin embargo, un conjunto suficiente de datos no es ni tan complicado de obtener ni tan económicamente elevado mientras se haya identificado y analizado correctamente el problema (Brynjolfsson & McAfee, 2017).

### 3.2. Machine Learning aplicado en producción, fabricación y calidad

La complejidad de los sistemas productivos y la demanda de productos personalizados, hacen necesaria la introducción de soluciones flexibles que puedan hacer frente a los problemas de un entorno cambiante liderado por la incertidumbre. El término *Intelligent Manufacturing Systems* (IMS, Sistemas Inteligentes de Producción) hace referencia a la nueva generación de sistemas productivos que utilizan los resultados de las investigaciones en inteligencia artificial para resolver los problemas ligados a la falta de información (Monostori, 2003).

De hecho, la relación entre inteligencia y aprendizaje es clara, puesto que las tecnologías deseadas en la industria son aquellas que no solo generan y almacenan información, sino que, además, la convierten en conocimiento para apoyar y facilitar las tareas de los humanos (Monostori, 2003).

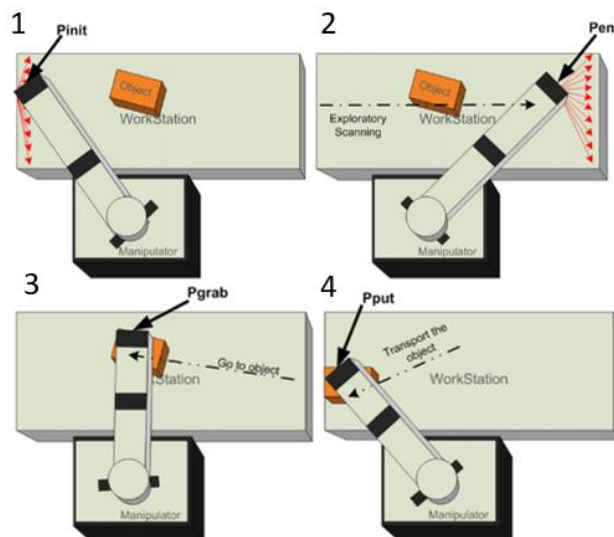
En el sector metalúrgico existe una amplia variedad de estudios relacionados con la monitorización de los procesos de mecanizado y con la predicción de la composición de los productos. En (Laha, Ren, & Suganthan, 2015) se demuestra experimentalmente que el algoritmo de *Support Vector Regression* es el que mejor predice el límite elástico en la producción de acero a partir de sus materias primas. Debido a la complejidad de las transferencias

de calor y masa, así como a las reacciones químicas propias del proceso, la naturaleza de las variables de estudio es extremadamente no lineal.

En el corte de metal por láser, (Tsai, Li, & Chen, 2008) proponen una combinación de redes neuronales y algoritmos genéticos para encontrar los valores óptimos de corriente, frecuencia y velocidad de corte del proceso según la calidad deseada. En el proceso de torneado de acero, (Pontes, de Paiva, Belestrassi, Ferreira, & da Silva, 2012) obtienen un modelo de redes neuronales de base radial para predecir la rugosidad media. (Çaydas & Hasçalik, 2008) llevaron a cabo un estudio similar aplicado al corte por chorro de agua, combinando redes neuronales con modelos de regresión para predecir la rugosidad superficial a partir de la velocidad del chorro, la presión y la distancia de aplicación; donde se demostró que la presión era el parámetro más significativo.

La visión artificial es una disciplina científica muy utilizada en la industria, ya que permite a los robots realizar tareas como el control de calidad y la inspección y reconocimiento de objetos. Los sistemas de visión también sirven para determinar la posición y la orientación de los objetos de forma que un robot pueda manipularlos. Sin embargo, las cámaras que normalmente se emplean para este propósito están muy influenciadas por la luz del entorno industrial.

En (Pinto, Rocha, & Moreira, 2013) se describe una solución basada en un telémetro láser 2D (Laser Range Finder, LRF) que crea una imagen 3D en escala de grises en función de la distancia al objeto. Después, se emplean algoritmos de *Machine Learning* para realizar el reconocimiento y la clasificación de los objetos. Finalmente, un brazo robótico se dirige al objeto y lo transporta al lugar correspondiente (*Figura 3. 3*). En el estudio se comparan los resultados de tres algoritmos: KNN, SVM y redes neuronales; siendo SVM el que mejor generaliza. Este sistema puede utilizarse para situar los objetos en máquinas de mecanizado o en elementos de transporte, reduciendo así las labores manuales.



*Figura 3. 3. Escáner exploratorio para manipular objetos con visión artificial (Pinto, Rocha, & Moreira, 2013)*

La empresa ZenRobotics utiliza un robot que combina visión artificial y *Machine Learning* para clasificar cuatro mil elementos reciclables por hora (*Figura 3. 4*). Los datos son suministrados en tiempo real gracias a los sensores y a las imágenes que procesan las cámaras con el objetivo de identificar una amplia variedad de basura y envoltorios de comida y bebida. Además, este sistema es capaz de reconocer logos, marcas y formas en 3D (Gesing, Peterson, & Michelsen, 2018; Kite-Powell, 2017). Máquinas similares pueden utilizarse, por ejemplo, en operaciones logísticas internas de las empresas de mensajería, clasificando de forma rápida y fiable los distintos paquetes.



*Figura 3. 4. Máquina de reciclaje de ZenRobotics (Gesing, Peterson, & Michelsen, 2018).*

Los controles de calidad realizados por humanos corren el riesgo de ser erróneos. Sin embargo, la calidad de los productos es un aspecto imprescindible que puede ser controlado con *Machine Learning*. Además, conocer de antemano si la calidad de la producción se está deteriorando, previene el gasto innecesario de materia prima y de tiempo (Seebo, 2019).

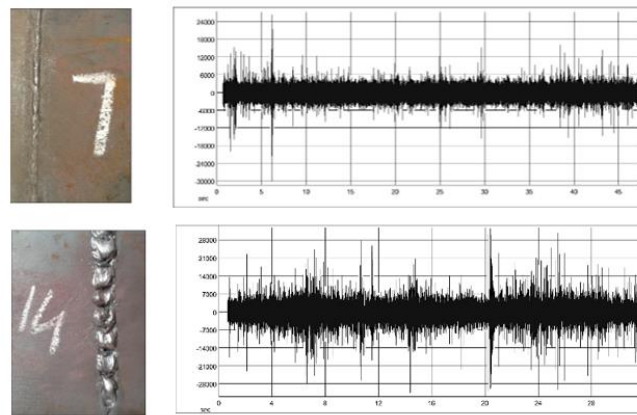
Algunos estudios relacionados con la industria textil utilizan redes neuronales para detectar los defectos relacionados con los hilos y las telas, así como los defectos de corte y de costura (Yildirim, Birant, & Alpyildiz, 2018). Por ejemplo, en (Xin, Li, Qiu, & Liu, 2012) se crea un sistema de evaluación de calidad de arrugas en las telas basado en el análisis de imágenes mediante redes neuronales.

El color es un elemento de calidad muy importante en la industria textil. En (Zhang & Yang, 2014) se compara el método de *Support Vector Machine* con el algoritmo de *Naive Bayes* en la predicción de diferencias de color entre telas teñidas. Un estudio realizado por (Golob & Osterman, 2008) propone determinar la correcta combinación de pigmentos para tintes usando redes neuronales entrenadas con mil cuatrocientas treinta muestras obtenidas de diez tintes.

La soldadura es una de las técnicas de unión más empleadas en la industria, ganando vital importancia en sectores como el aeronáutico o el de la automoción. La física de la soldadura es de naturaleza compleja y no lineal, por lo que no es fácil extraer un modelo matemático que correlacione la calidad con las diferentes variables del proceso.

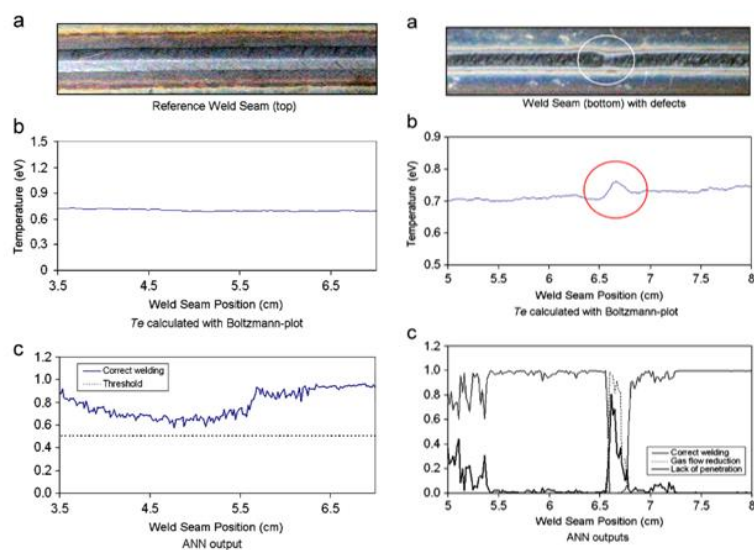
Según el experimento realizado por (Sumesh, Rameshkumar, Mohandas, & Shyam Babu, 2015), los algoritmos de *Random Forest* (Bosques Aleatorios) se pueden emplear para controlar la calidad de la soldadura a partir del sonido del arco. Los parámetros de entrada son la corriente, la tensión y la velocidad de avance; mientras que la salida es la buena o mala calidad de la

soldadura. Situando un micrófono donde se efectúa la soldadura, se obtiene un espectrograma del sonido del arco que sirve de entrenamiento al modelo (*Figura 3. 5*).



*Figura 3. 5. Comparación entre un espectrograma asociado a una buena soldadura (arriba) y uno asociado a un defecto de exceso de soldadura (abajo) (Sumesh, Rameshkumar, Mohandas, & Shyam Babu, 2015)*

Otro estudio (Mirapeix, García-Allende, Cobo, Conde, & López-Higuera, 2007) emplea el análisis espectroscópico del espectro del plasma producido durante la soldadura para monitorizar la calidad de cordón de soldadura. La estimación del perfil de temperatura electrónica ofrece una correlación entre el espectro del plasma y la calidad del proceso (*Figura 3. 6*). La detección automática del defecto y su correspondiente clasificación se consigue aplicando *Principal Component Analysis (PCA)* para comprimir la información del espectro y facilitar el procesado, seguido de una red neuronal que realiza la propia detección y clasificación.



*Figura 3. 6. Comparación entre las salidas de la red neuronal para una buena soldadura (izquierda) y una mala soldadura (derecha) (Mirapeix, García-Allende, Cobo, Conde, & López-Higuera, 2007)*

El *Machine Learning* también se aplica en la producción de energía, siendo la predicción de carga eléctrica un aspecto fundamental para conseguir una gestión eficiente de los recursos. Modelos de predicción no lineales como el *Support Vector Regression* junto con modelos de optimización como el *Modified Firefly Algorithm*, consiguen predecir a corto plazo la demanda eléctrica, como se explica en (Kavousi-Fard, Samet, & Marzbani, 2014).

En (Tüfekci, 2014) se propone una comparación de quince modelos de regresión para predecir la potencia de una central eléctrica. Esta solución sirve para evitar los errores derivados de las aproximaciones matemáticas propias de la termodinámica, y para reducir el tiempo y el coste computacional. Con el histórico de datos de temperatura ambiente, presión atmosférica, humedad relativa y presión de vapor a la salida de la turbina; se consigue un modelo basado en árboles de decisión que predice la producción de electricidad.

### 3.3. Machine Learning aplicado en logística

La cadena de suministro genera a diario una cantidad inmensa de datos tanto estructurados como no estructurados que solo pueden explotarse gracias a la inteligencia artificial. La logística se basa en redes físicas y digitales que no pueden ser optimizadas por humanos debido a su alta complejidad. Por ello, el objetivo de la inteligencia artificial es transformar los comportamientos reactivos en proactivos, los manuales en automáticos y los estandarizados en personalizados (Gesing, Peterson, & Michelsen, 2018).

La inteligencia artificial es uno de los mayores contribuidores en el desarrollo del vehículo autónomo. Para que el transporte autónomo sea aceptado, debe superar las capacidades de un humano al volante, empezando por la percepción del entorno y su habilidad para predecir cambios en él. Eso es posible gracias a la combinación de tecnologías que construyen un mapa tridimensional del entorno (DHL, 2014; Gesing, Peterson, & Michelsen, 2018).

El *Deep Learning* se encarga de procesar datos para identificar las señales de tráfico, detectar obstáculos y otros coches en la carretera, así como de cumplir con las leyes de tráfico. Estos algoritmos son imprescindibles para que el vehículo adquiera conocimiento, dado que los humanos no son capaces de programar todas las posibles situaciones que pueden ocurrir en la carretera (DHL, 2014; Gesing, Peterson, & Michelsen, 2018).

Industrias del automóvil como BMW, Daimler y Ford indican que la inteligencia artificial es un elemento crucial para el vehículo autónomo. Otras empresas como Google y Tesla ya han desarrollado sus propios vehículos autónomos a partir de sus propias técnicas de inteligencia artificial (*Figura 3. 7*). Los proveedores de componentes como Bosch y ZF ya incluyen sensores y la capacidad de gestionar datos en sus productos. Es más, plataformas de movilidad como Uber, se están asociando con las empresas para ofrecer viajes bajo demanda de forma autónoma (DHL, 2014; Gesing, Peterson, & Michelsen, 2018).

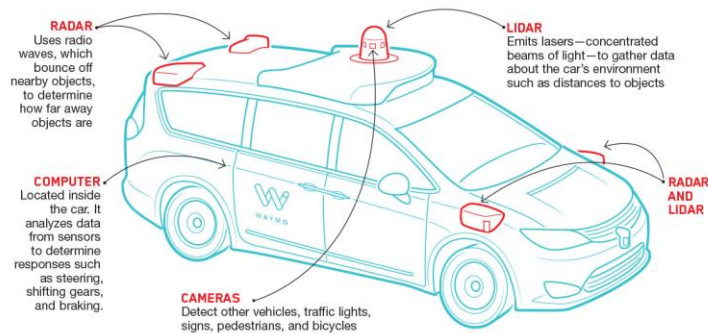


Figura 3. 7. Componentes principales de Waymo, el proyecto de vehículo autónomo de Google. Recuperado de: <https://scienceworld.scholastic.com/issues/2017-18/090417/hands-free-ride.html#1070L>

Un menor coste, una mayor eficiencia a nivel de emisiones y una menor tasa de accidentes son algunas ventajas que se obtendrían con el uso del vehículo autónomo. A pesar de la rápida evolución en el desarrollo de estos sistemas que posibilitan la conducción autónoma, los asuntos legales hacen que esta transición hacia la autonomía lleve más tiempo del deseado (DHL, 2014; Gesing, Peterson, & Michelsen, 2018).

Además, aún existe una desconfianza general hacia la conducción autónoma debido a que no todo el mundo está dispuesto a subirse a bordo de un vehículo conducido por un robot ni a pagar mucho más dinero por ese servicio. Sin embargo, según (Standard & Poor's, 2018) el vehículo autónomo en 2030 será una aplicación común para el transporte de mercancías y paquetería en los Estados Unidos. Por el contrario, en ese mismo país se estima que en 2024 existirá una falta de conductores profesionales, acelerando así la adopción del vehículo autónomo en logística.

Varios ejemplos de camiones autónomos ya están siendo utilizados. Uno de ellos es el Volvo Vera, un camión eléctrico y autónomo diseñado para el transporte de grandes volúmenes de carga entre centros logísticos. Vera no cuenta con cabina para el conductor y trabaja gracias a la monitorización de datos en *la nube*. Los sistemas de comunicación incorporados le permiten estar en contacto con el centro de control de transporte, desde donde se controlan los parámetros ligados al transporte. Vera forma parte de la flota de vehículos que opera en el puerto de Gotemburgo transportando contenedores (Figura 3. 8), llegando a las treinta y dos toneladas de carga máxima (Gutiérrez, 2019).



Figura 3. 8. Volvo Vera operando en el puerto de Gotemburgo, Suecia (Gutiérrez, 2019)

En cuanto al transporte aéreo de mercancías, Project Wing (*Figura 3. 9*) es un servicio aéreo de reparto con drones autónomos que ha obtenido recientemente la aprobación de la Aviación Australiana para poder operar comercialmente en ciertas partes del país y bajo ciertos límites, convirtiéndose así en el primer servicio de drones mensajeros del mundo (Pérez, 2019). El *Deep Learning* es el núcleo de este sistema de envíos, empleándose principalmente en la detección de obstáculos como árboles, pájaros o tendido eléctrico (X Company, 2019).



*Figura 3. 9. El dron de Project Wing llevando mercancía en Australia (Pérez, 2019)*

Amazon desplegó en 2013 un proyecto de servicio de drones similar que aún está pendiente de aprobación. Recientemente, se ha anunciado el rediseño de su dron con el fin de hacerlo más seguro y aerodinámico, esperando así agilizar la aprobación por parte de la Administración Federal de Aviación de Estados Unidos. Su objetivo con este proyecto es realizar entregas de mercancías inferiores a tres kilogramos en un radio de veinticuatro kilómetros en menos de treinta minutos (Vincent, 2019).

Los drones, además de entregar mercancías, se utilizan para otra serie de aplicaciones logísticas como la gestión de inventarios o el abastecimiento de materiales en los puestos de fabricación. La cadena Walmart se interesó inicialmente por los drones con el objetivo de realizar los envíos más rápidamente. En lugar de eso, dispone de una flota de drones que escanea los productos del almacén y detecta aquellos que están descolocados o por debajo de stock necesario gracias a una combinación de visión artificial, sensores y tecnología RFID (PINC, 2019). En (Cho, Kim, Roh, & Hwang, 2018) se diseña un sistema de detección de códigos de barras para drones usando una cámara infrarrojos y aplicando *Support Vector Machine* para la clasificación de imágenes (*Figura 3. 10*).



*Figura 3. 10. Sistema de identificación de códigos de barras en almacenes (Cho, Kim, Roh, & Hwang, 2018)*

El uso de drones en logística interna es posible debido a su alta precisión en los movimientos y a su capacidad de disminuir la velocidad en entornos de trabajo. En (Olivares, Cordova, Sepúlveda, & Derpich, 2015) se modelan las estaciones de trabajo y las rutas con algoritmos genéticos, y después se establecen los pesos transportados por cada dron en función de si retiran o entregan materiales en el puesto de trabajo. A diferencia de los AGV y las cintas transportadoras, los drones pueden trabajar en varios planos, obteniendo más flexibilidad en el diseño de rutas. Según (Duhalde, 2017), Audi está implementando drones para el transporte de piezas entre zonas de la planta, ganando agilidad y tiempo de reacción (*Figura 3. 11*). Algunas pruebas han demostrado su efectividad en situaciones de urgencia, donde un humano o un sistema de transporte terrestre hubieran fracasado.



*Figura 3. 11. Dron llevando un volante en la fábrica de Audi (Duhalde, 2017)*

El rol de los almacenes ha cambiado drásticamente debido a la complejidad y variedad de pedidos, y a la necesidad de información en tiempo real. En las operaciones logísticas tradicionales es muy común ver a personas usando carretillas y transpaletas para mover mercancías de un punto a otro. Sin embargo, estas tareas dan lugar a bajos rendimientos del almacén por no satisfacer el ritmo y los requerimientos de los consumidores (Lee, Lv, Ng, Ho, & Choy, 2017).

Los sistemas de gestión de almacenes (*Warehouse Management Systems, WMS*) son aplicaciones informáticas que sincronizan, optimizan y controlan la utilización de los recursos del almacén para ganar eficiencia, flexibilidad y competitividad. Los datos y la información se introducen en un sistema inteligente de gestión de inventarios que organiza los pedidos y resuelve los problemas de *picking* con ayuda de algoritmos de *Machine Learning* como el *clustering* (Lee, Lv, Ng, Ho, & Choy, 2017). El procesamiento del lenguaje natural (*Natural Language Processing, NPL*) se utiliza en estos sistemas con la finalidad de comunicarse con el técnico del almacén a través de dispositivos móviles y ayudarlo así en tareas repetitivas como la introducción de datos o la búsqueda de información (*Figura 3. 12*) (Gesing, Peterson, & Michelsen, 2018).





Figura 3. 12. Sistemas con NPL para el soporte en el almacén (Gesing, Peterson, & Michelsen, 2018)

Los almacenes están incorporando robots autónomos (como los AGV, *Automated Guided Vehicles*) y robots colaborativos para que las personas se dediquen a tareas de mayor valor añadido. GreyOrange es una compañía de automatización y robótica que ha desarrollado un *software* llamado GreyMatter como complemento a un WMS. GreyMatter emplea *Machine Learning* para hacer que los AGV colaboren en tiempo real y optimicen tanto sus recorridos como sus velocidades. Además, les dota de un mecanismo de autoaprendizaje para que mejoren sus capacidades con la experiencia. Por ejemplo, los AGV pueden cambiar la posición de los productos más populares para acercarlos al punto de *picking* y reducir así los desplazamientos (Figura 3. 13) (Gesing, Peterson, & Michelsen, 2018; GreyOrange, 2019).



Figura 3. 13. Empleo de AGV en la gestión de almacenes con GreyMatter (GreyOrange, 2019)

Este mismo sistema de gestión de almacenes es el que se emplea en los centros logísticos de Amazon, como por ejemplo el de Castellbisbal, que es el primer almacén robotizado de España. Las estanterías de productos se denominan *Pods*, y los AGV que las mueven se llaman *Drives*. Cada robot puede transportar hasta mil trescientos kilogramos por la superficie de trabajo, conocida como *Robotics Field*. En este almacén trabajan trescientos cincuenta *Drives* y dos mil *Pods* a una velocidad inferior a dos metros por segundo. Los productos son acercados hasta los operarios de almacén para preparar los pedidos (Figura 3. 14). En concreto, desde que una persona compra el producto hasta que el pedido está listo, solamente pasan entre quince y setenta minutos. Además, la ausencia de pasillos aumenta el aprovechamiento del espacio en un 50% (López, 2017).



*Figura 3. 14. Las estanterías se acercan a los operarios para que preparen el pedido. Recuperado de: <https://www.manutencionyalmacenaje.com/Articulos/237916-Amazon-Robotics-en-los-centros-logisticos-de-Espana.html>*

Los robots colaborativos (*cobots*) posibilitan el trabajo conjunto entre un robot y un operario en condiciones seguras. Son ligeros, móviles y llevan sistemas inteligentes incorporados, brindando un amplio abanico de aplicaciones tanto en producción como en logística. Es más, se estima que las ventas de *cobots* crecerán exponencialmente en los próximos años, llegando a un tercio de los robots vendidos en 2027 (Samaniego, 2019).

Los *cobots* son ideales para las tareas de *pick and place* (coger y colocar), de empaquetado y paletizado y de control de calidad. La empresa de automatización logística Vanderlande ha desarrollado un *cobot* llamado SIR (*Smart Item Robotics*), que incorpora visión artificial y *Machine Learning* con el fin de coger objetos y situarlos en la mejor ubicación dentro de los contenedores. Su funcionamiento y diseño evitan cualquier tipo de riesgo tanto para los productos como para las personas cercanas (*Figura 3. 15*) (Schilperoort, 2018).



*Figura 3. 15. Robot colaborativo SIR en funcionamiento junto a un trabajador (Schilperoort, 2018)*

La empresa logística DHL es un referente en la aplicación de soluciones 4.0 en el ámbito logístico. En sus almacenes de Estados Unidos han incorporado un *cobot* llamado LocusBot, creado por la empresa Locus Robotics. Su finalidad es recorrer los almacenes junto a los trabajadores, optimizando los recorridos y llevando los productos. Este *cobot* entiende varios idiomas e interactúa con los trabajadores a través de una pantalla táctil (*Figura 3. 16*) (Locus Robotics, 2019; Samaniego, 2019).



*Figura 3. 16.* Un trabajador de almacén valida la carga depositada en LocusBot. Recuperado de: <http://www.digitaltoo.com/2017/04/10/robots-dhl-localizacion-recogida-productos/>

Otra de las creaciones de DHL consiste en una herramienta de *Predictive Network Management* basada en *Machine Learning* que predice las demoras en los tránsitos de transporte aéreo de mercancías de la semana siguiente. Además, identifica los factores que provocan dichos retrasos y obtiene relaciones entre ellos para facilitar la toma de decisiones. También han desarrollado un *Predictive Demand and Capacity Planning* con el que detectan las nuevas tendencias a partir del análisis de contenido en redes sociales para predecir aumentos bruscos de demanda en determinados productos, ganando así capacidad de reacción (Gesing, Peterson, & Michelsen, 2018).

Un aspecto importante en logística es el de la optimización de rutas, ya que mejora la eficiencia de los transportes y las entregas. Ontruck es un operador español de transporte de mercancías por carretera que combina *Machine Learning* y *Business Intelligence* para realizar un cálculo dinámico de los precios de cada servicio y optimizar así las rutas y las cargas en función de la distancia a recorrer, la hora, el tráfico y las condiciones climáticas de la zona. Hasta ahora, ha conseguido reducir los kilómetros en vacío en un 10%, lo equivalente a trescientos noventa mil euros en 2018 (La Razón, 2019).

### 3.4. Machine Learning aplicado en mantenimiento

En los últimos años, el abaratamiento de los sensores y su menor tamaño han facilitado la obtención de información valiosa sobre el estado de las máquinas. En concreto, midiendo diferentes puntos y características de una máquina, es posible tener una visión casi a tiempo real sobre su estado. Gracias al *Machine Learning*, se crean modelos a partir de esos datos con el fin de detectar posibles anomalías antes de que se produzcan. Esto se denomina mantenimiento predictivo, y su aplicación es una de las formas más fiables para evitar que las máquinas fallen y perjudiquen al proceso productivo (Rodríguez, 2018).

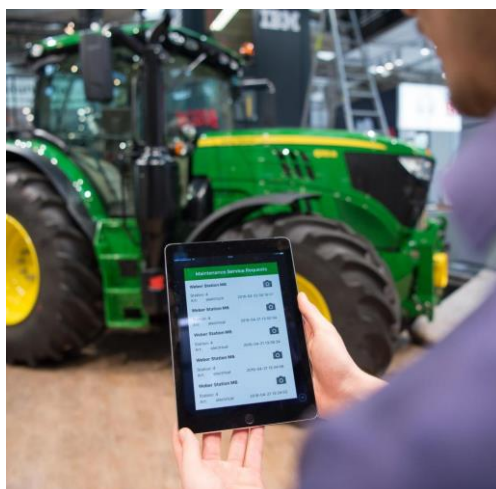
A pesar de que el concepto de mantenimiento predictivo no supone una novedad, el desarrollo tanto de las tecnologías de obtención y almacenamiento de datos como el de las aplicaciones de *Machine Learning*, han contribuido a crear una nueva perspectiva sobre este término. De este modo, la contribución de datos procedentes de muchas fuentes es tratada con algoritmos más complejos que han permitido una reducción en los costes de mantenimiento de entre un 10-15 % (Breunig, Kelly, Mathis, & Wee, 2016).

Tradicionalmente, las alertas en los procesos industriales se configuraban manualmente, por lo que se corría el riesgo de pasar por alto un falso positivo, como por ejemplo un pico exagerado en la monitorización de la temperatura de un proceso. Sin embargo, dado que los algoritmos de *Machine Learning* están alimentados y entrenados con datos de diferentes procedencias, es posible encontrar una correlación entre esa anomalía y otros parámetros del proceso, llegando así a una causa raíz antes de que se deteriore la máquina por completo (Seebo, 2019).

Un algoritmo *K-Means* de dos etapas propuesto por (Yiakopoulos, Gryllias, & Antoniadis, 2011) permite realizar el diagnóstico automático de defectos en rodamientos según su frecuencia de vibración. La principal ventaja de este algoritmo es que se trata de un tipo de aprendizaje no supervisado, así que no necesita entrenamiento a partir de ejemplos específicos. Los buenos resultados de este estudio se deben a la correcta selección de las variables frecuenciales fruto de la experiencia previa sobre los defectos de los rodamientos. En la primera etapa se decide si el rodamiento tiene defecto, y en la segunda se clasifica el tipo de defecto.

Estudios más recientes (Li, y otros, 2019) emplean estructuras profundas de SVM multicapa para diagnosticar fallos en máquinas rotativas como bombas centrífugas y rodamientos. En (Guo, Wu, Sun, Long, & Zhang, 2019) se diseña un algoritmo *Transfer Support Vector Machine* (TSVM) en el diagnóstico de fallos de impresoras 3D. Los algoritmos de transferencia de conocimiento y adaptación de dominio como el TSVM se utilizan para aprender modelos a partir de conjuntos de datos similares, reduciendo así el coste asociado a la recolección de datos etiquetados.

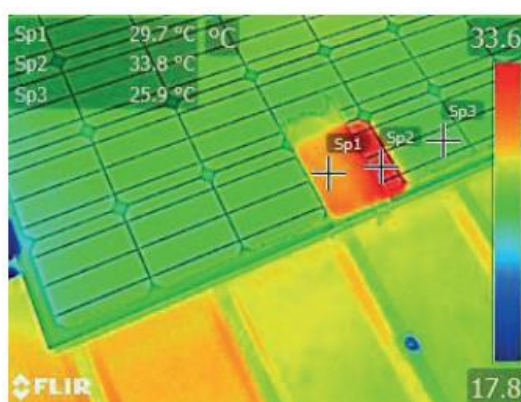
En el sector de la agricultura, John Deere emplea IBM Watson (ver apartado 3.5) y sistemas inteligentes de producción para impulsar el mantenimiento basado en inteligencia artificial. El personal de mantenimiento hace uso de dispositivos móviles conectados a sistemas de *Deep Learning* para obtener las causas de fallo de una máquina a partir de una foto, gracias a los algoritmos de reconocimiento de imágenes (Figura 3. 17). El trabajador recibe una explicación por parte de la aplicación sobre cómo arreglar dicho fallo, e incluso pueden comunicarse con comandos de voz para habilitar el trabajo en manos libres (Gesing, Peterson, & Michelsen, 2018; Lee, 2016).



*Figura 3. 17. Identificación de defectos mediante una foto tomada con un dispositivo móvil. Recuperado de: <https://www.welt.de/wirtschaft/article154838727/Das-neue-Technik-Zeitalter-folgt-dem-Prinzip-Hundefoto.html>*

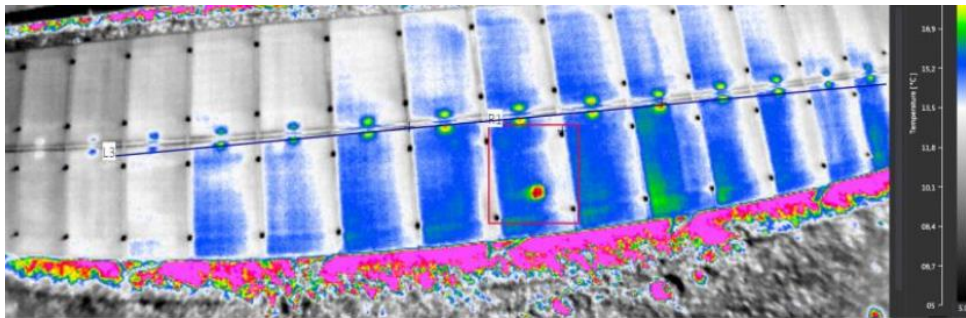
Además, el sistema permite el pedido automático de repuestos cuando identifica que algún elemento no funciona correctamente. También es capaz de encontrar el mejor momento para realizar la labor de mantenimiento en función de los horarios y la agenda del técnico (Gesing, Peterson, & Michelsen, 2018; Lee, 2016).

Las termografías son imágenes que representan la temperatura de un lugar o de un objeto mediante una escala de colores. La aplicación de cámaras termográficas en la industria es un tipo de mantenimiento predictivo. Los roces entre elementos mecánicos, las fugas de ciertos líquidos y gases, y la quemaduras producidas por exceso de electricidad son señales detectables por medio de su temperatura y que ayudan a prevenir un desastre mayor. En la *Figura 3. 18* se muestra una termografía de un panel fotovoltaico donde una de sus celdas experimenta una disipación de calor debido a un mal funcionamiento (Glavas, Vidakovic, Jersek, & Kraus, 2018).



*Figura 3. 18. Punto caliente en un panel fotovoltaico detectado con una cámara termográfica (Glavas, Vidakovic, Jersek, & Kraus, 2018)*

Las termografías aéreas surgen de la combinación de las cámaras termográficas con los drones, ayudando al personal de mantenimiento a detectar anomalías en lugares altos, de acceso complicado o incluso peligroso. Además, permiten ampliar el campo de visión desde lugares altos y reducir el tiempo de ejecución. Los sistemas de visión artificial y *Machine Learning* incorporados se encargan de la detección automática de las anomalías y de su clasificación (Droniter, 2019). En la *Figura 3. 19* se muestra lo que se vería desde la cámara termográfica acoplada a un dron sobre un conjunto de paneles fotovoltaicos.



*Figura 3. 19. Punto caliente en un panel fotovoltaico detectado con un dron (Droniter, 2019)*

Aunque el *Machine Learning* consiga analizar datos y aprender de forma extremadamente eficaz, es importante resaltar el papel de los humanos a la hora de desarrollar y mejorar las herramientas relacionadas con el mantenimiento. Tanto la experiencia del personal como los históricos de datos de las máquinas son información básica para que los sistemas puedan trabajar. Además, se necesita el *feedback* por parte de una persona experimentada para ajustar los algoritmos y validar los resultados que muestran. Al fin y al cabo, las decisiones las toman los humanos en base a predicciones realizadas por el *Machine Learning* (Seebo, 2019).

### 3.5. Machine Learning aplicado en los negocios

En el entorno empresarial existe una demanda creciente de aplicaciones inteligentes que aprovechan las técnicas del aprendizaje automático para optimizar la toma de decisiones. La multinacional estadounidense de tecnología y consultoría, IBM, desarrolla algunas de estas aplicaciones tanto para uso interno como para sus clientes. En ambos casos, el objetivo es realizar un análisis de negocio y optimizar las decisiones operativas y estratégicas rutinarias (Apte, 2010).

La forma en la que las empresas aprovechan la información ha ido evolucionando desde un método de análisis descriptivo hasta uno de análisis predictivo, o incluso prescriptivo. El análisis descriptivo de datos permite al usuario obtener una visión retrospectiva del negocio, obteniendo respuestas a preguntas como "qué sucedió", "cuántas veces" y "dónde". El análisis predictivo permite a un usuario obtener una visión prospectiva del negocio, obteniendo respuestas a preguntas como "qué podría suceder", "qué pasaría si estas tendencias continúan" y "lo que podría suceder a continuación si..." (Apte, 2010).

Por otro lado, el análisis prescriptivo permite al usuario saber cuál es el conjunto de acciones necesarias para lograr un objetivo de negocio, bajo un conjunto determinado de predicciones y restricciones. La adopción de aplicaciones tanto predictivas como prescriptivas mejoran considerablemente la competitividad de las empresas (Apte, 2010).

Dentro de IBM, las primeras aplicaciones exitosas de aprendizaje automático aparecieron en las áreas de control de calidad de fabricación y gestión del rendimiento de las computadoras a principios de los años 90. Se vio que los algoritmos de aprendizaje supervisado eran muy útiles a la hora de generar rápidamente modelos predictivos a partir de datos (Apte, 2010).

Unos años más tarde, el enfoque de IBM cambió hacia los servicios para los clientes, pudiendo acceder a cantidades mayores de datos empresariales. Las aplicaciones de *Machine Learning* en las que IBM comenzó a trabajar para otros clientes giraban en torno al área del marketing y la gestión de la relación con los clientes (Apte, 2010).

IBM fue desarrollando nuevas soluciones de *Machine Learning* a medida que los problemas empresariales se hacían más complejos, haciendo frente a desafíos como la insuficiencia de ejemplos etiquetados y la alta o baja dimensionalidad. Esto ha dado lugar a innovadores proyectos en áreas como la detección de fraude, las redes sociales, la gestión de la delincuencia, la gestión de la salud o la gestión de la sostenibilidad (Apte, 2010).

Actualmente, el sistema empresarial de inteligencia artificial más utilizado es IBM Watson, que debe su nombre al primer CEO de IBM, Thomas John Watson. Se trata de un sistema de computación cognitiva que realiza las tareas de observar, interpretar, evaluar y decidir en base a los datos proporcionados, en lugar de guiarse por aspectos emocionales como hacen los humanos. El objetivo de Watson es posibilitar la comunicación natural entre humanos y computadoras a través de distintos procesos para que los humanos puedan hacer preguntas y obtener respuestas justificadas (IBM, 2016).

Según (IBM, 2016), “*IBM Watson se ha convertido en una tecnología comercial, accesible a través de la nube y que cuenta con clientes en 17 sectores distintos y 30 países del mundo, entre ellos España, donde gracias a CaixaBank ha aprendido español*”. Por ejemplo, en el sector petrolífero, IBM Watson está ayudando a Repsol a analizar miles de documentos, noticias e imágenes sísmicas para comparar escenarios hipotéticos y así obtener mejores resultados en la exploración y producción de hidrocarburos.

Por otro lado, gracias al *Machine Learning* pueden crearse agentes que aprendan estrategias de negociación, de forma que den soporte en las negociaciones o incluso posibiliten su completa automatización. El principal desafío de la negociación se debe a que cada parte tiene información privada e ignora las estrategias de negociación de la otra parte, dificultando así un posible acuerdo común. Para ayudar a los humanos a encontrar soluciones que beneficien a ambas partes, se proponen agentes que de forma autónoma logran este objetivo (Oliver, 1996).

Un estudio realizado por (Chen & Weiss, 2015) muestra un sistema capaz de modelar a los oponentes sin conocimiento previo. Para ello emplea dos técnicas: *Discrete Wavelet Transformation* (DWT) y procesos Gaussianos. Este sistema consiste en agentes automáticos que trabajan en negociaciones bilaterales multi-problema en entornos complejos. De esta forma, se consigue aprender la estrategia del oponente para decidir los siguientes movimientos, así como para proponer ofertas con alta probabilidad de aceptación.

En (Carbonneau, Kersten, & Vahidov, 2008) se presenta un modelo basado en redes neuronales entrenadas con Regularización Bayesiana, que predice las ofertas y contraofertas de la otra parte durante el proceso de negociación electrónica. Se demuestra que este sistema es capaz de aportar estrategias muy interesantes al negociador a pesar del ruido en los datos. Sin embargo, su capacidad de generalizar no es suficiente, ya que el sistema ha sido entrenado en contextos similares de negociación electrónica.

La efectividad de los diferentes mecanismos automáticos de negociación depende de las características del problema, por lo que no hay un mecanismo único que funcione bien para todos los casos. En (Aydogan, Marsa-Maestre, Klein, & Jonker, 2018) se propone una solución de *Machine Learning* para elegir el mecanismo adecuado a partir del comportamiento de varios mecanismos sobre varios escenarios.

### 3.6. Machine Learning aplicado en marketing y gestión de clientes

Con inteligencia artificial se consiguen productos y servicios personalizados para los consumidores. El 79% de los usuarios de internet son activos en redes sociales, lo que supone una gran fuente de datos para aquellas compañías que deseen conocer mejor los intereses de la población (Gesing, Peterson, & Michelsen, 2018; Williams, 2017). La explotación de dichos datos requiere algoritmos de *Machine Learning* que emplean el procesamiento del lenguaje natural para identificar los sentimientos y las opiniones de las personas.

Con esa información, las organizaciones personalizan la publicidad en función de los gustos de cada persona. Solo el 2% de las personas que compran *on-line* realizan la compra en su primera visita a la *web*, por lo que esta publicidad personalizada es una buena forma de atraer de nuevo a los posibles clientes (Gesing, Peterson, & Michelsen, 2018; Williams, 2017). Algoritmos de *Machine Learning* ajustan la ubicación de los anuncios en la *web* en base a la predicción de que un consumidor haga *clic* o no (Brynjolfsson & McAfee, 2017).

Otra aplicación del *Machine Learning* en el comercio electrónico son los sistemas de recomendación de productos. Estos sistemas sugieren productos a personas con perfiles similares en función de las compras que han realizado anteriormente. Por ejemplo, el 30% de los beneficios de Amazon se debe a productos vendidos gracias a su sistema de recomendación (Gesing, Peterson, & Michelsen, 2018; Rejoiner, 2019).

Infinite Analytics es una *startup* del MIT (*Massachusetts Institute of Technology*) especializada en la aplicación de inteligencia artificial en el comercio electrónico. Con *Machine Learning* desarrollan varias herramientas para las empresas, entre las que se encuentran los motores de búsqueda y recomendación, la comparación de productos con los de los competidores, la obtención de atributos a partir de imágenes de productos, la monitorización de campañas de *marketing* digital e incluso un motor de comparación entre el estilo de los productos y el estilo de los *influencers* sociales (Infinite Analytics, 2019).

La gestión de la relación con los clientes (*Customer Relationship Management*, CRM) comprende todas aquellas actividades que tienen el objetivo de mejorar la atención al cliente y maximizar el conocimiento sobre los clientes para entender en profundidad sus necesidades y anticiparse a ellas. En (Lariviere & Van den Poel, 2005) se aplica la técnica de los Bosques

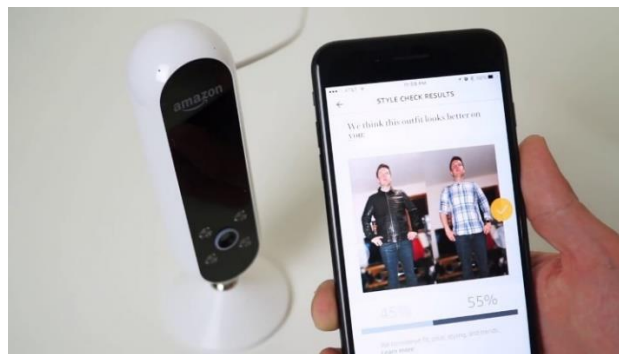


Aleatorios (*Random Forests*) para comprender mejor tres pilares fundamentales del CRM: cuándo será la próxima compra, cuándo se está alejando el cliente y la evolución de la rentabilidad del cliente. Una base de datos recopila información histórica sobre los clientes y, mediante estos algoritmos, se extraen las relaciones entre las variables para predecir los tres aspectos mencionados.

El *Machine Learning* también sirve para automatizar la atención al cliente mediante *chatbots* y agentes virtuales que pueden establecer conversaciones de distintos niveles de complejidad con los clientes. Autodesk, empresa líder en software de ingeniería y diseño por ordenador, utiliza el servicio de IBM Watson Conversation con el que realiza unas treinta mil interacciones al mes con un tiempo medio de respuesta a las consultas de cinco minutos (Gesing, Peterson, & Michelsen, 2018; IBM, 2019).

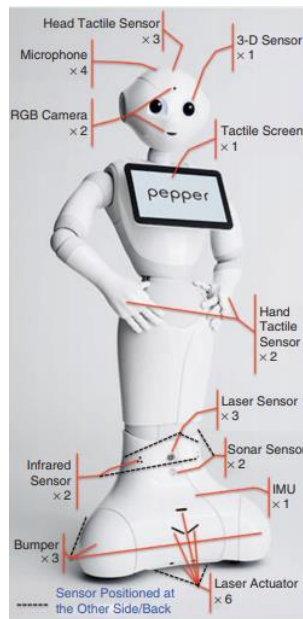
El BlueBot de la aerolínea KLM nació como una aplicación móvil para tratar automáticamente consultas básicas. Actualmente, los algoritmos de *Machine Learning* incorporados hacen que este *chatbot* ofrezca una atención más personalizada en las consultas, además de dar soporte en la búsqueda y compra de vuelos (Gesing, Peterson, & Michelsen, 2018; KLM Royal Deutch Airlines, 2019).

Echo Look es el *chatbot* de Amazon especializado en la comparación de *outfit*. Los clientes indecisos que no sepan qué prendas de ropa les sientan mejor, pueden enviar un par de fotos al experto en estilo de Amazon para recibir una sugerencia por su parte (Figura 3. 20). Dicha sugerencia personalizada se realiza en función del análisis de colores y del estilo de la prenda, teniendo en cuenta las últimas tendencias (Venkatesh, 2018).



*Figura 3. 20. Echo Look, el chatbot de Amazon que ayuda a elegir la ropa. Recuperado de: <https://www.america-retail.com/supply-chain/supply-chain-amazon-echo-look-la-camara-fashionista-ya-esta-la-venta/>*

La atención al cliente en las tiendas también se ha visto afectada con el desarrollo del *Machine Learning* y la robótica. Por ejemplo, Softbank Robotics ha creado a Pepper, un robot humanoide que interactúa con los clientes para ofrecerles información, guiarles por la tienda y resolver sus preguntas. Los sensores y el *Machine Learning* de este robot (Figura 3. 21) se encargan de analizar las expresiones y el tono de voz de los cliente para reconocer sentimientos y ofrecer así una experiencia de compra (Kumar & Gelin, 2018).



*Figura 3. 21. Sensores de Pepper (Kumar & Gelin, 2018)*

Tally es el robot desarrollado por Simbe Robotics que, a diferencia de Pepper, se encarga de monitorizar el inventario y el precio de los productos en las tiendas. El sistema de visión artificial que lleva incorporado detecta los espacios libres en las estanterías, lanza una orden de reposición a los trabajadores y prioriza las alertas. También identifica aquellos productos descolocados o con precio incorrecto (*Figura 3. 22*). Esto mejora la experiencia de los clientes, asegurando la existencia de los productos en todo momento (Simbe Robotics, 2019; Underwood, 2019).



*Figura 3. 22. Tally realizando el stock check en un supermercado. Recuperado de: [https://www.upi.com/Top\\_News/US/2017/08/07/Tally-the-robot-could-be-the-stock-boy-of-the-future/1161502148250/](https://www.upi.com/Top_News/US/2017/08/07/Tally-the-robot-could-be-the-stock-boy-of-the-future/1161502148250/)*

Lowe's es una distribuidora de productos para la mejora del hogar, la construcción y el bricolaje, que con la colaboración de Fellow Robots han creado LoweBot (Gesing, Peterson, & Michelsen, 2018). Este robot es una combinación de los dos anteriores, ya que además de controlar la existencia de los productos en tienda, interactúa con los clientes (en diferentes

idiomas, si es necesario) y les guía hasta el producto buscado (*Figura 3. 23*). Los algoritmos de *Machine Learning* de este robot identifican patrones de compra de forma que los trabajadores puedan comprender qué mercancías se venden más, en qué momento del día y en qué temporada del año (Lowe's Companies, 2016; Song, 2017).



*Figura 3. 23. LoweBot guiando a un cliente (Lowe's Companies, 2016)*

Sorprendentemente, la inteligencia artificial es capaz de controlar toda una tienda, y no únicamente la atención al cliente. En Estados Unidos ya se han abierto supermercados inteligentes sin cajeros donde se puede realizar la compra sin esperas. Amazon abrió su primera tienda Amazon Go en Seattle en 2017. A través de la aplicación de Amazon Go, los clientes escanean un código en la entrada de forma que, a partir de ese momento, su actividad en la tienda queda registrada (*Figura 3. 24*) (Polacco & Backes, 2018).



*Figura 3. 24. Entrada y salida de la tienda Amazon Go a través de la aplicación. Recuperado de: <https://www.muycanal.com/2018/12/04/amazon-go-grandes-establecimientos>*

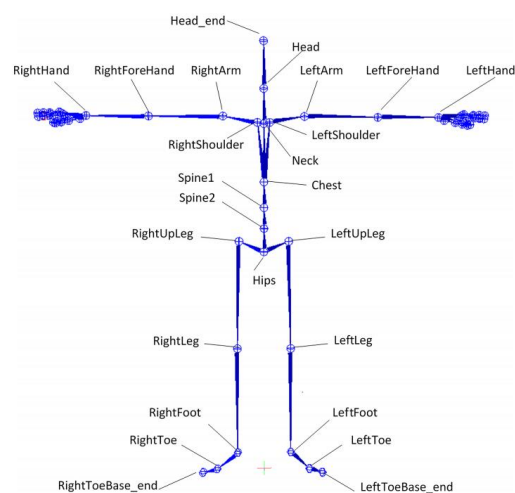
Las tecnologías principales utilizadas en las tiendas de Amazon Go son las mismas que las del vehículo autónomo: los sensores, la visión artificial y el *Deep Learning*. Las estanterías detectan los productos cogidos o cambiados de sitio, así como la persona que lo hizo. Cuando un cliente sale por la tienda se le carga automáticamente el importe en su cuenta de Amazon, evitando colas o cualquier otro tipo de espera (Polacco & Backes, 2018).

### 3.7. Machine Learning aplicado en ergonomía

En cuanto a las condiciones de trabajo en los puestos de producción y montaje, los sistemas de análisis de movimientos (*Motion Analysis Systems*, MAS) se utilizan para crear informes detallados sobre el rendimiento tanto productivo como ergonómico del trabajador. Esto se consigue con un *hardware* llamado MOCAP (*Motion Capture* o capturadora de movimientos) integrado con un *software* basado en redes neuronales especializadas en el tratamiento de imágenes y vídeos (Bortolini, Faccio, Gamberi, & Pilati, 2018; Pilati & Regattieri, 2018).

La adopción de las tecnologías MOCAP en la industria ha adquirido gran importancia con la llegada de las fábricas inteligentes. En su origen, los MOCAP se diseñaron para reconocer movimientos en los videojuegos, pero ahora se han visto las ventajas de utilizar el mismo principio para estudiar y mejorar las actividades manuales (Bortolini, Faccio, Gamberi, & Pilati, 2018; Pilati & Regattieri, 2018).

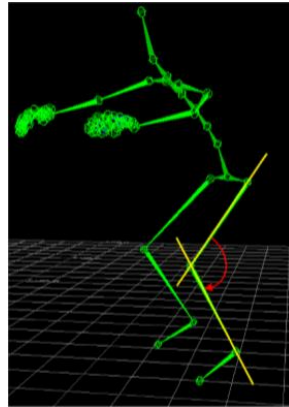
Esta tecnología modela el esqueleto humano y analiza la posición de cada uno de los puntos representados en la *Figura 3. 25*. La red neuronal se alimenta de todos los datos relativos a las características físicas del trabajador y a las propiedades de las herramientas y equipos, así como a su distribución espacial en el puesto de trabajo (Bortolini, Faccio, Gamberi, & Pilati, 2018; Pilati & Regattieri, 2018). De esta forma, el sistema reconstruye un entorno virtual en 3D del puesto de trabajo, diferenciando entre piezas, productos, muebles, etc.



*Figura 3. 25. Modelo del esqueleto humano para analizar con MAS (Bortolini, Faccio, Gamberi, & Pilati, 2018)*

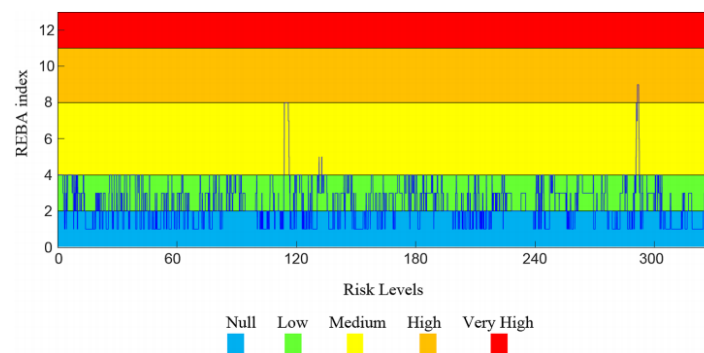
Se definen unos KPI (*Key Performance Indicators*, Indicadores Clave de Rendimiento) que serán evaluados por el sistema MAS de forma automática durante la ejecución de la tarea productiva. Algunos de los KPI más relevantes son la distancia recorrida, la distribución de las manos en el puesto, el porcentaje de tiempo de valor añadido en el tiempo de ciclo o la postura corporal en base a ciertos índices ergonómicos (Bortolini, Faccio, Gamberi, & Pilati, 2018; Pilati & Regattieri, 2018).

La perspectiva ergonómica engloba un análisis de todos los movimientos del trabajador durante la actividad, así como la evolución de los ángulos de las articulaciones (*Figura 3. 26*), con la finalidad de detectar las malas posturas y evitar posibles lesiones. Otros índices ergonómicos normalmente analizados son el OWAS (*Ovako Working Body Assessment*), el REBA (*Rapid Entire Body Assessment*), el NIOSH (*National Institute for Occupational Safety and Health index*) y el EAWS (*European Assembly Worksheet*) (Bortolini, Faccio, Gamberi, & Pilati, 2018; Pilati & Regattieri, 2018).



*Figura 3. 26. Ángulo de la rodilla medido con MAS (Bortolini, Faccio, Gamberi, & Pilati, 2018)*

En (Bortolini, Faccio, Gamberi, & Pilati, 2018) se lleva a cabo un estudio del REBA. En la *Figura 3. 27* se observa que el riesgo del trabajador es bajo o nulo en la mayoría del periodo de tiempo analizado. Sin embargo, surgen dos picos con un riesgo medio o alto que pueden suponer un problema de seguridad y salud para el trabajador. Por otro lado, en la *Figura 3. 28* se representa el REBA promedio para cada parte del cuerpo, siendo posible evaluar las articulaciones críticas.



*Figura 3. 27. Evolución dinámica del REBA y clasificación del riesgo para la salud (Bortolini, Faccio, Gamberi, & Pilati, 2018)*

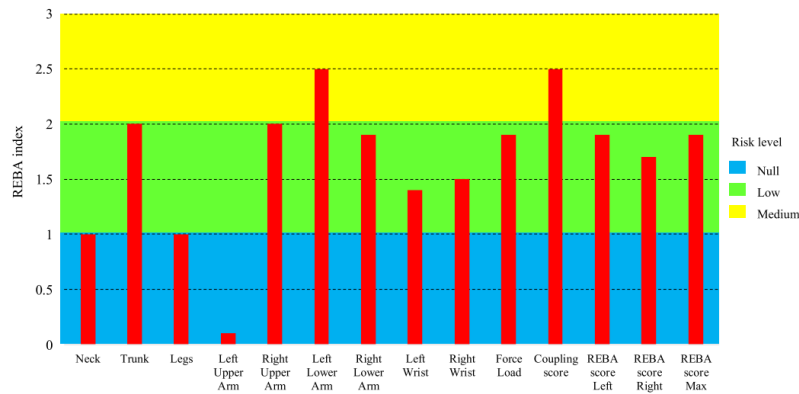


Figura 3. 28. Promedio del REBA según la zona corporal (Bortolini, Faccio, Gamberi, & Pilati, 2018)

Otra forma de aplicar el *Machine Learning* en ergonomía se consigue con los sensores equipados en los *smartphone*. Estos sensores proporcionan información sobre la localización y los movimientos que realiza el usuario, como por ejemplo el cuentapisos. A diferencia de otros dispositivos, los *smartphone* son baratos, fáciles de usar y no necesitan gran mantenimiento. En particular, se utilizan para la prevención de enfermedades musculoesqueléticas (MSD, *Musculoskeletal Disorders*) asociadas a malas posturas. Cuando las MSD son causadas en el entorno laboral, se las denomina *Work-related Musculoskeletal Disorders* (WMSD) (Nath, Chaspari, & Behzadan, 2018).

Estas enfermedades afectan a un gran número de trabajadores, desde personal de oficina hasta el sector de la construcción, dando lugar a bajas de larga duración que derivan en pérdidas económicas. Una solución para prevenir las WMSD consiste en la PtD (*Prevention through Design*, prevención a través del diseño), la cual se encarga de asegurar un diseño ergonómico de tareas, herramientas y entornos de trabajo. Para llevar a cabo un PtD se necesita un análisis previo de los factores de riesgo existentes, lo que implica una toma de datos relacionados con las condiciones de trabajo (Nath, Chaspari, & Behzadan, 2018).

Los métodos tradicionales de toma de datos consumen demasiado tiempo y entorpecen la labor de los trabajadores. Por ello, en (Nath, Chaspari, & Behzadan, 2018) se diseña un sistema automático de procesamiento de datos recogidos con sensores de *smartphone*, que calcula el riesgo ergonómico asociado a tareas de esfuerzo excesivo. A continuación, se hace uso de algoritmos de *Machine Learning* como el SVM, que reconocen y clasifican las actividades realizadas por el usuario y extraen información sobre la duración y la frecuencia de dichas actividades.

Algunas aplicaciones comentadas en apartados anteriores mejoran al mismo tiempo la ergonomía de ciertas tareas. Por ejemplo, el uso de *cobots* (ver apartado 3.3) libera a los humanos de actividades extenuantes y aumenta la precisión de las operaciones. También los drones (ver apartado 3.4) evitan que los humanos tengan que acceder a lugares en altura o de difícil acceso.

### 3.8. Machine Learning aplicado en seguridad

El aumento del uso de *Internet*, tanto en la vida social como en la laboral, cambia por completo la forma en que las personas aprenden y trabajan. Sin embargo, el número de ciberataques aumenta de la misma manera. En el periodo de transformación digital en el que se encuentran las empresas, un incidente tecnológico de este tipo puede poner fin a la continuidad de la empresa. Por ello, surge la ciberseguridad como un conjunto de tecnologías y procesos diseñados para proteger a las redes, a las computadoras, a los programas y a los datos, de los posibles ataques y amenazas (Xin, Kong, Liu, & Chen, 2018).

Los potentes algoritmos de *Machine* y *Deep Learning* en ciberseguridad se emplean principalmente para el análisis de *malware* y la detección y prevención de intrusos. El desarrollo de estos algoritmos viene impulsado por la necesidad de anticiparse a un ciberataque y de restringir el acceso a los archivos o programas infectados (Handa, Sharma, & Shukla, 2019).

Los sistemas de seguridad de redes cuentan normalmente con cortafuegos, antivirus y sistemas de detección de intrusos (*Intrusion Detection Systems*, IDS). En particular, los IDS sirven para determinar e identificar comportamientos no autorizados como el uso, la copia o la destrucción de ciertos archivos o programas (Xin, Kong, Liu, & Chen, 2018).

Existen tres tipos de IDS: los basados en firmas, los basados en anomalías y los híbridos. Los IDS basados en firmas comparan las firmas de los ataques con las firmas registradas, por lo que solo detectan ataques ya conocidos. Los administradores deben actualizar la base de firmas manualmente para que los sistemas detecten los nuevos ataques sufridos. Por ejemplo, no pueden identificar los ataques de día cero, que son aquellos que ejecutan código dañino sobre aplicaciones nuevas aprovechando las vulnerabilidades propias de sus primeros días en el mercado (Xin, Kong, Liu, & Chen, 2018).

Los IDS basados en anomalías identifican cualquier desviación en el comportamiento habitual de la red, haciendo frente a los ataques de día cero. Sin embargo, algunos comportamientos nuevos que no son necesariamente una amenaza son catalogados como tal, dando lugar a falsas alarmas. Por lo tanto, la solución se encuentra en los IDS híbridos, ya que reducen la tasa de falsas alarmas y detectan mejor las intrusiones. Por ello, la mayoría de las herramientas de *Machine Learning* y *Deep Learning* son híbridas (Xin, Kong, Liu, & Chen, 2018).

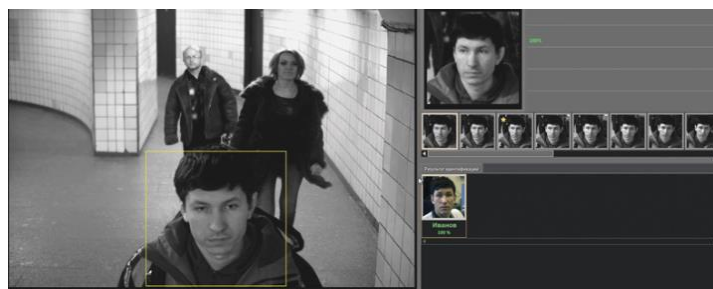
En (Yu, Long, & Cai, 2017) se propone un *dilated convolutional autoencoder* (DCAE) para la detección de intrusos en redes, que combina las ventajas de los *autoencoders* y las CNN. El modelo aprende automáticamente las características esenciales de un conjunto grande de datos procedentes de redes infectadas, llegando a una tasa de acierto superior al 98%. Otros estudios emplean modelos como el SVM (Saxena & Richariya, 2014), el KNN (Shapoorifard & Shamsinejad, 2017) o los árboles de decisión combinados con algoritmos genéticos (Azad & Kumar, 2015) con el mismo propósito.

Tecnologías similares a MAS (ver apartado 3.7) se utilizan para la videovigilancia de almacenes y comercios. AI Guardman es una cámara basada en *Machine Learning* que asegura una videovigilancia sin intervención humana. Incluye técnicas de reconocimiento facial y de detección de comportamientos sospechosos (*Figura 3. 29*), además de un sistema de alertas que avisa al responsable al momento (Dar, 2018).



*Figura 3. 29. Reconocimiento de movimientos sospechosos de AI Guardman. Recuperado de: <https://siesa.com.ar/japon-estrena-una-camara-de-seguridad-avanzada-que-funciona-con-inteligencia-artificial/>*

La empresa española líder en productos y aplicaciones electrónicas de seguridad, Hommax Sistemas, ya comercializa cámaras con aplicaciones de *Deep Learning* como el conteo de multitudes, la búsqueda de personas, el reconocimiento facial e incluso el reconocimiento de vehículos mal estacionados (*Figura 3. 30*) (Hommax Sistemas, 2019).



*Figura 3. 30. Sistema de búsqueda de personas empleando Deep Learning (Hommax Sistemas, 2019)*

El equipamiento de estas cámaras de seguridad en drones permite ampliar el campo de visión en la videovigilancia de grandes superficies como parques, terrenos agrícolas y naves industriales. Pueden programarse para realizar inspecciones rutinarias y automáticas, o bien ser pilotados manualmente (*Figura 3. 31*). Los sistemas de reconocimiento facial detectan la presencia de intrusos, los busca y los localiza. Además, al no ser sistemas fijos, hace que sea más complicado evadirlos o destruirlos (Prevent Security Systems, 2019).



*Figura 3. 31. Videovigilancia con drones. Recuperado de: <https://areaurbana.com/el-drone-que-garantiza-la-seguridad-publica/>*



### 3.9. Machine Learning aplicado en recursos humanos

Revisar currículums, tanto en papel como en plataformas de empleo, supone un tiempo muy elevado y, por lo tanto, un coste asociado también elevado. El procesamiento de datos permite extraer y clasificar a un conjunto de candidatos según su experiencia profesional, sus habilidades personales y técnicas, o sus expectativas salariales. Este método se está convirtiendo en una práctica especialmente popular y valiosa para encontrar a los candidatos pasivos, es decir, aquellos que están contentos con su trabajo actual y suponen un reto para los reclutadores (Van Pay, 2018).

Amazon desplegó en 2014 una herramienta de contratación basada en *Machine Learning* a partir de la cual se puntuaba a los candidatos de una a cinco estrellas, al igual que ellos pueden hacer con sus productos. La idea es analizar automáticamente un conjunto de currículums para obtener un *ranking* con los mejores, quienes pasarían a la siguiente fase del proceso. Sin embargo, examinando los currículums para un puesto de desarrollador de *software*, se detectó que el sistema no puntuaba a los hombres y a las mujeres por igual (Dastin, 2018; Kobie, 2018).

Dado que el sistema se entrenó con datos de aplicantes de los últimos diez años, y que el sector informático aún cuenta con mayoría masculina, él mismo aprendió a preferir a los hombres. A pesar de las modificaciones realizadas en el programa, nada garantizaba la ausencia de otro tipo de discriminación. Aunque los reclutadores defendían que solo utilizaban la herramienta para obtener sugerencias de contratación, el proyecto se acabó abandonando. No obstante, Amazon sigue trabajando en el desarrollo de motores de reclutamiento más justos (Dastin, 2018; Kobie, 2018).

LinkedIn, la mayor red profesional del mundo, ofrece a los empleadores una clasificación de los candidatos de la plataforma en función de su adecuación al puesto ofertado. Ni siquiera los altos cargos de estas empresas confían por completo en la inteligencia artificial para esta aplicación, pues defienden que la tecnología no está preparada aún (Dastin, 2018; Kobie, 2018).

Empresas como HireVue realizan vídeo entrevistas donde se analiza la forma de hablar y las expresiones faciales de los candidatos, dejando a un lado el problema de los sistemas examinadores de currículums. El formato consiste en una serie de preguntas que el candidato debe responder ante una cámara, permitiendo a los reclutadores entrevistar a un gran número de personas al mismo tiempo (Dastin, 2018; Kobie, 2018; Van Pay, 2018).

Existen *software* más completos enfocados a los procesos de selección, los cuales pueden ser programados en función de las necesidades de las empresas. Por ejemplo, VERA (Figura 3.32) es un robot virtual creado por dos expertos en recursos humanos con el objetivo de simplificar las tareas repetitivas. VERA busca y filtra información sobre los candidatos en distintas plataformas de empleo, detectando si un mismo candidato está inscrito en varias plataformas. Cada mes, VERA realiza cuarenta mil llamadas telefónicas y envía treinta y siete mil correos para obtener y gestionar la información de los candidatos. VERA entrevista a un grupo de candidatos que considera adecuados para el puesto y, después de ello, elabora una lista de candidatos para que el personal de recursos humanos termine con el proceso de selección (Jumbo, 2019).



*Figura 3. 32. Robot virtual VERA para la realización de entrevistas on-line. Recuperado de: <https://www.bloomberg.com/news/articles/2018-03-28/this-ai-software-aims-to-do-90-percent-of-hr-s-recruiting-work>*

Un aspecto importante es la falta de empatía e interacción humana que caracteriza a estos sistemas de inteligencia artificial. Con ellos, se pierde la oportunidad de mantener conversaciones personales con los candidatos durante el proceso, siendo el trato personal un punto muy favorable en muchos candidatos (Van Pay, 2018). AIRA (*Artificial Intelligence Recruitment Assistant*), a diferencia de VERA, ofrece una experiencia más agradable y transparente del proceso gracias a la detección de actitudes y emociones (Jumbo, 2019).

Una fase en los procesos de selección de algunas empresas consiste en demostrar habilidades como la destreza, la resolución de problemas, la agilidad mental o la toma de decisiones, mediante juegos *on-line*. El sistema detecta el perfil de cada candidato con los datos recopilados y, en caso de superar esa fase, se les ofrece una entrevista *on-line* en la que una inteligencia artificial evalúa su comportamiento ante diferentes situaciones reales (Jumbo, 2019).

Entre las ventajas que se consiguen con la aplicación de *Machine Learning* y otras tecnologías en los procesos de selección, se encuentra la variedad y el volumen de perfiles analizados, así como la rapidez para detectar si un perfil se adecúa al puesto de trabajo. Las investigaciones actuales se centran en mejorar el reconocimiento de cualidades como el humor o la ironía durante las entrevistas, y en eliminar los prejuicios que pueda aprender el sistema (Jumbo, 2019).

Ha surgido recientemente una polémica respecto a un sistema inteligente desarrollado por Amazon que detecta la productividad de sus trabajadores, llegando a despedirlos. El sistema rastrea la actividad de cada trabajador y lo compara con los estándares de productividad establecidos, generando alertas sin la necesidad de supervisores. El gran número de despidos realizados entre 2017 y 2018 indignó a los trabajadores, quienes defendían que ni siquiera disponían de tiempo para ir al servicio (ABC, 2019).

Por su parte, Amazon respondió que, si no tomaran ese tipo de medidas, sería imposible mantener a sus trabajadores ante un mercado laboral tan competitivo. Sin embargo, afirman que antes de despedir a un empleado se aseguran de que haya recibido un plan de mejora y de formación adicional para que tengan la oportunidad de aumentar su rendimiento (ABC, 2019).

### 3.10. Limitaciones del Machine Learning en la industria

Las principales barreras a las que se enfrentan las empresas en cuanto a su digitalización y adopción de nuevas tecnologías (ver apartado 1.5.3), también engloba las dificultades básicas de la aplicación del *Machine Learning* en la industria. Estas dificultades son la resistencia al cambio por parte tanto de los directivos como de los empleados; el coste asociado a la implantación de estos sistemas digitales y a la reorganización de los procesos; y la ausencia de personal cualificado en herramientas 4.0.

Sin embargo, esas limitaciones se atribuyen a las empresas y no a la tecnología del *Machine Learning* en sí. Por lo tanto, en este apartado se comentan algunas de las limitaciones propias del *Machine Learning* que evitan su completo desarrollo en la industria.

Aunque el *Machine Learning* ya se aplica en cientos de empresas en todo el mundo, se estima que la mayoría de las aplicaciones aún están por venir. No obstante, y al igual que otras tecnologías, el *Machine Learning* genera expectativas irreales debido al uso incorrecto de sus capacidades. Es decir, algunos planes de negocio incluyen la utilización de herramientas de *Machine Learning* para resolver problemas que no pueden resolverse con esta tecnología (Brynjolfsson & McAfee, 2017).

Para los casos en los que el *Machine Learning* se aplica correctamente, surgen dificultades a la hora de comprender por qué el sistema ha tomado unos caminos y no otros. Debido a la complejidad de las redes neuronales, cualquier cambio en alguna de las conexiones puede variar considerablemente el resultado. Por ello, la forma en que los sistemas de *Machine Learning* transmiten información es demasiado clara y simple, pues en ocasiones no pueden llegar a argumentar sus decisiones (Brynjolfsson & McAfee, 2017).

Como consecuencia, es probable que los sistemas de *Machine Learning* aprendan comportamientos sesgados ocultos y que los apliquen sin que los humanos sean conscientes de ello. Un ejemplo claro es la inteligencia artificial creada por Amazon y aplicada en la selección de personal que aprendió a discriminar a las mujeres (ver apartado 3.9).

Otro problema está asociado a la dificultad de demostrar la certeza con la que el sistema de *Machine Learning* tomará una buena decisión en situaciones complejas. Esto incluye situaciones de vida o muerte, como el caso de la conducción autónoma (ver apartado 3.3) o el control de una planta nuclear. Incluso aceptando una tasa de error, es complicado diagnosticar dichos errores y corregirlos (Brynjolfsson & McAfee, 2017).

Popularmente, la inteligencia artificial se caracteriza por su escasa capacidad para detectar y expresar emociones y sentimientos, aunque esta afirmación no es del todo cierta. El *Machine Learning* aprende a detectar y clasificar las emociones de las personas, y a reaccionar ante ellas. Sin embargo, para que una inteligencia artificial tenga emociones humanas habría que recrear el cerebro humano, su cuerpo y su cognición (Bitbrain, 2018).

En las aplicaciones industriales no es un aspecto prioritario. Por ejemplo, la forma que tiene de ver la realidad un *chatbot* se limita a una conversación en formato de texto, y para un sistema de reconocimiento facial se limita a estímulos en formato de vídeo. Por lo tanto, conociendo la realidad desde un único punto de vista es imposible generar una emoción del mismo modo que hacen los humanos. De todas formas, no se sabe hasta qué punto interesa dotar a las máquinas

con inteligencia emocional, pues los sentimientos hacia los humanos podrían ser también negativos (Bitbrain, 2018).

Las labores que un sistema basado en *Machine Learning* puede realizar son variadas y complejas, tal y como se ha demostrado en este capítulo. No obstante, la dependencia total de los datos hace que el sistema resulte inútil si los datos no cumplen unos requisitos de calidad y cantidad. Además, la creatividad y la curiosidad de los científicos, investigadores y emprendedores es algo esencial que las máquinas no pueden alcanzar (Brynjolfsson & McAfee, 2017).

Al fin y al cabo, las máquinas desarrollan las funciones definidas por sus creadores de manera que la generalización de su conocimiento se limita a dichas funciones. Por ello, las máquinas con inteligencia general en distintos dominios aún están lejos de la realidad (Brynjolfsson & McAfee, 2017).

# CONCLUSIONES Y LÍNEAS FUTURAS

En este apartado se exponen las principales conclusiones derivadas de la realización del presente trabajo, así como las dificultades encontradas y las aportaciones ofrecidas por el autor. Por último, se proponen varias líneas futuras de trabajo.

## 1. Conclusiones

La conclusión general a la que se llega con este trabajo es que el *Machine Learning* es una tecnología fundamental en el desarrollo de las empresas debido a su capacidad para extraer información relevante de los datos y a su amplio abanico de aplicaciones en todos los sectores, tanto por sí misma como en combinación con otras tecnologías.

En base a los objetivos establecidos al principio del documento y a lo tratado en los capítulos anteriores, se concluye que dichos objetivos se han alcanzado satisfactoriamente. En el primer capítulo se ha introducido el concepto de Industria 4.0 y las tecnologías que emergen para cubrir las necesidades que han dado lugar a esta Cuarta Revolución Industrial. Comprender estas necesidades es el punto de partida para entender la función que desempeña cada una de las tecnologías.

Además, aunque la Revolución lleve el apellido de “Industrial”, los cambios también se producen fuera de este ámbito. La conectividad hace que las ciudades utilicen estas tecnologías para mejorar la calidad de vida de las personas, quienes demandan productos y servicios cada vez más customizados.

El primer capítulo termina con una reflexión sobre el estado actual de las empresas españolas en su proceso de transformación. En comparación con otros países, España se encuentra en una posición intermedia de digitalización donde, en líneas generales, las grandes empresas tienen planes de digitalización más definidos que las pequeñas y medianas empresas. No obstante, aún queda mucho trabajo por hacer para lograr que los trabajadores se adapten tanto mental como técnicamente a los cambios de la Industria 4.0, constituyendo así el principal reto al que se enfrentan las empresas.

En el capítulo dos se desarrolla el término de *Machine Learning*, desde los orígenes de la inteligencia artificial a mediados del siglo XX hasta nuestros días. A continuación, se exponen los fundamentos teóricos básicos para comprender el funcionamiento de esta poderosa herramienta. En concreto, se explica en qué consiste el aprendizaje automático y la diferencia que presenta frente a la programación tradicional. También se tratan términos como la función de coste, el descenso del gradiente, la varianza, el sesgo, el *overfitting* y el *underfitting*. Se incluye además un apartado con el procedimiento a seguir a la hora de implantar *Machine Learning* para solucionar un problema. Esta exposición se realiza de manera didáctica y evitando un lenguaje excesivamente técnico, de forma que cualquier persona interesada en conocer los principios de esta tecnología pueda hacerlo sin problemas.

En ese mismo capítulo, se dedica un apartado a cada tipo de aprendizaje con el fin de comprender en qué consiste y qué modelos son los más utilizados en cada caso. El aprendizaje supervisado se emplea cuando se puede alimentar al agente con pares de ejemplos de entrada y salida de manera que, tras el aprendizaje, sea capaz de predecir la salida en función de la entrada. Se parte de los modelos de regresión, como la regresión lineal y la regresión polinomial, y se termina con los de clasificación como el *Support Vector Machine*. Los modelos de regresión predicen una cantidad continua, mientras que los de clasificación predicen una etiqueta. El aprendizaje supervisado es el más utilizado entre las aplicaciones de *Machine Learning* por tratarse de algoritmos sencillos y eficaces, solo que la dificultad reside en encontrar ejemplos de salida para que el algoritmo aprenda.

Por esa misma razón surge el aprendizaje no supervisado, donde el agente detecta los patrones en los datos de entrada sin necesidad de observar ejemplos de salida. Los modelos de *clustering* crean conjuntos de objetos de características similares, mientras que la reducción dimensional simplifica el número de variables. Estos algoritmos se emplean para encontrar estructuras desconocidas en los datos, siendo este un problema muy común en la industria.

El aprendizaje reforzado se basa en un sistema de recompensas y castigos donde el agente busca las mejores decisiones para maximizar las recompensas. Se aplica en robótica y en diversos juegos como el ajedrez. Su impacto en la historia del *Machine Learning* es elevado, dado que muchos de los grandes éxitos de esta tecnología a lo largo de su historia se deben a momentos en los que un sistema de aprendizaje reforzado vencía a los mayores profesionales en determinados juegos de mesa y videojuegos.

Por último, en la guía establecida en el capítulo dos se explican los conceptos de red neuronal y *backpropagation*, así como tres de los tipos de redes de aprendizaje profundo más habituales. Las redes convolucionales (CNN) se especializan en la identificación de imágenes, las redes recurrentes (RNN) se encargan principalmente del procesamiento de cadenas de texto, y las redes generativas adversarias (GAN) se centran en la generación de nuevas imágenes. Las aplicaciones del *Deep Learning* en la industria son de vital importancia, pues van un paso más allá en comparación con otros algoritmos de *Machine Learning*. Son el pilar fundamental del desarrollo del vehículo autónomo y de la visión artificial, prometiendo dar solución a problemas industriales verdaderamente complejos.

Una vez comprendido el contexto industrial y los principios del *Machine Learning*, el capítulo tres reúne algunas de las aplicaciones que actualmente existen en la industria, constituyendo así el corazón del trabajo. Se observa que las empresas son conscientes de la necesidad de esta tecnología en sus negocios, pero actualmente muy pocas han conseguido hacer uso de ella.

El *Machine Learning* se encuentra en constante evolución, habiendo logrado ya un nivel de precisión y autonomía que parecía impensable. En la era de los datos, no basta con llenar las fábricas de sensores y de almacenar los datos de forma masiva esperando que así se solucionen los problemas ligados a la fabricación, a la calidad, al estado de los equipos, etc. En su lugar, es completamente necesario disponer de otras herramientas que hagan la función de intermediario entre los datos procedentes de un entorno industrial y los tomadores de decisiones. Y es aquí donde entra en juego la labor del *Machine Learning*.

La evaluación de las aplicaciones del *Machine Learning* por áreas permite adquirir una idea clara sobre los avances de esta tecnología en la totalidad de una empresa industrial. Con esto se demuestra la flexibilidad, la polivalencia y la potencia que aporta esta herramienta en cualquier situación donde haya datos de los que se quiera extraer cierta información. La mejora de los procesos de fabricación y de control de calidad, el mantenimiento de equipos, la selección de personal, la gestión de almacenes, la atención al cliente, la ciberseguridad, el transporte de mercancías y, en general, el soporte en la toma de decisiones son solo algunas de las aplicaciones tratadas en este documento.

Sin embargo, a pesar de todas las posibilidades que ofrece el *Machine Learning*, aún no se ha explotado la totalidad de su potencial. Por ello, los investigadores siguen trabajando para conseguir que las máquinas sean más eficientes en tareas delicadas que supongan un riesgo para las personas, como por ejemplo la conducción autónoma. También, la gran dependencia de los datos hace que muchos de los esfuerzos por mejorar esta tecnología se dirijan al preprocesamiento de los datos. De todos modos, estas soluciones no podrán hacerse efectivas hasta que las empresas no estén preparadas para aceptar este conjunto de innovaciones tecnológicas.

La mayor dificultad en la realización de este trabajo radica en la gran cantidad de documentación existente. Además, los conocimientos previos sobre este tema eran muy escasos, por lo que prácticamente es un trabajo partido de cero. Sintetizar la información relativa a una tecnología tan completa como el *Machine Learning* ha supuesto un reto personal, sobre todo teniendo en cuenta que la mayoría de la bibliografía consultada era de carácter muy técnico. No obstante, la manera en la que se ha tratado de mostrar esa información es puramente didáctica, con una estructura clara e intuitiva para que el lector pueda seguir los contenidos como si fuera una guía resumen.

Además de la aportación de dicha guía, el repertorio de aplicaciones ha resultado ser bastante completo. Se han tomado ejemplos de aplicación provenientes de grandes empresas como Amazon, Google o DHL; pero no solo eso, sino que también se han expuesto aplicaciones de empresas más pequeñas que han nacido con el objetivo de vender soluciones tecnológicas a las más grandes. De alguna forma, esto intenta decir que no hace falta ser una gran empresa para superar la transformación digital, sino que la clave consiste en realizar un análisis interno para identificar aquellos puntos fuertes y débiles que son susceptibles de ser mejorados ya sea con esta tecnología o con otras.

Se ha observado la combinación de tecnologías a la hora de abordar un problema, por lo que las posibilidades son aún mayores. De hecho, se espera que algunas de estas aplicaciones sirvan de referencia o aporten ideas a aquellos lectores que estén pensando en implantar *Machine Learning* en sus negocios. Aunque mi recomendación en ese caso es que antes de elegir una tecnología, se identifique el objetivo que se pretende conseguir y, a partir de ahí, elegir la tecnología que más se adecúa en función de otros aspectos como por ejemplo el coste de implantación.

## 2. Líneas futuras

Como líneas futuras de este trabajo se propone la realización de un estudio más detallado según la clasificación seguida en este documento. Es decir, ampliar y completar las aplicaciones del *Machine Learning* en producción, en logística, en mantenimiento, en los negocios, en la gestión de clientes, en ergonomía, en seguridad y en recursos humanos. También se propone realizar un estudio similar pero enfocado a un sector en particular. Por ejemplo, el sector de la automoción, el aeronáutico, el textil, el agrícola, el de la construcción, etc. Ya que en cada sector se está aplicando esta tecnología de formas muy diferentes para cubrir las necesidades específicas de dichos sectores.

Dado que este trabajo se centra únicamente en el sector industrial, otra propuesta interesante consiste en la aplicación del *Machine Learning* fuera de la industria. Por ejemplo, estudiar la influencia del *Machine Learning* en el día a día de las personas, su impacto en la evolución de las ciudades, del transporte, etc. Aunque sin duda, donde más se está investigando el uso de esta tecnología es en medicina. Ya se habla de sistemas de clasificación de células cancerígenas a partir de imágenes y del diagnóstico de otras enfermedades a partir de los síntomas. Esto supone una revolución en ese ámbito, pues estas máquinas llegan a ser en algunos casos más eficientes que los propios médicos.

En este trabajo se han tratado las aplicaciones tanto del *Machine Learning* como del *Deep Learning*. Sin embargo, puesto que el *Deep Learning* es el subcampo que más interés ha despertado en los últimos años, se propone realizar una investigación más detallada sobre su uso en el sector industrial. En particular, la logística presenta grandes retos para el *Deep Learning* en la forma de gestionar la *supply chain*. Por ejemplo, la optimización y planificación de la última milla, pues es el punto más crítico de toda la cadena de suministro debido a la amplia concentración de problemas tales como el tráfico, el impacto ambiental, el coste elevado, la satisfacción del cliente, etc. Otros ejemplos incluyen el papel del *Deep Learning* en el desarrollo de vehículos autónomos, o en la mejora de la trazabilidad de los productos junto con el *Blockchain*.

Los informes consultados sobre el estado de la transformación digital en las empresas ofrecen una idea general sobre la importancia que otorgan a las tecnologías habilitadoras y su grado de implantación. Podría resultar interesante realizar un informe similar enfocado al *Machine Learning* donde se muestre la postura que adoptan tanto las pequeñas como las grandes empresas a la hora de utilizar esta herramienta para solucionar sus problemas. De esta forma se podrían comparar las tendencias que presentan las empresas en función de su tamaño y sector, y ver las ventajas o inconvenientes que presenta el *Machine Learning* frente a otras tecnologías para un mismo problema industrial. También se podrían realizar estudios económicos sobre casos reales para tener una idea del beneficio que aporta esta tecnología en comparación con su coste de implantación.

Por último, se propone realizar una guía más completa sobre el *Machine Learning* y los algoritmos existentes, puesto que aquí solo se han explicado brevemente los algoritmos más conocidos y empleados, dejando atrás otros algoritmos que pueden resultar de interés para ciertas aplicaciones.



# BIBLIOGRAFÍA

- ABC. (30 de Abril de 2019). Amazon despide a cientos de trabajadores mediante un algoritmo. *ABC*, págs. [https://www.abc.es/economia/abci-amazon-despide-cientos-trabajadores-mediante-algoritmo-201904301229\\_noticia.html](https://www.abc.es/economia/abci-amazon-despide-cientos-trabajadores-mediante-algoritmo-201904301229_noticia.html).
- Abdi, H. (2003). *Least Squares*. Richardson: The University of Texas and Dallas.
- Abdullatif, H. (6 de Febrero de 2019). *Understanding SVD (Singular Value Decomposition)*. Obtenido de Towards Data Science: <https://towardsdatascience.com/svd-8c2f72e264f>
- ACAN. (2018). *La Industria 4.0 Tecnologías habilitadoras*.
- AEC. (2019). *Smart Grid*. Obtenido de Asociación Española para la Calidad: <https://www.aec.es/web/guest/centro-conocimiento/smart-grid>
- Aldecoa, I., Carné, N., & Monte, E. (2003). Separación ciega de fuentes. *BURAN. Ingeniería de Telecomunicaciones de Barcelona. UPC, Nº19*, pág 63-72.
- Álogos. (9 de julio de 2018). *Introducción a Machine Learning*. Obtenido de Álogos: <http://alogos.es/introduccion-machine-learning/>
- Aluja, T. (2001). La Minería de Datos, entre la estadística y la inteligencia artificial. *Universitat Politècnica de Catalunya*, Vol. 25, 3, p.479-498.
- Apte, C. (2010). The Role of Machine Learning in Business Optimization.
- Aydogan, R., Marsa-Maestre, I., Klein, M., & Jonker, C. (2018). A Machine Learning Approach for Mechanism Selection in Complex Negotiations. *Journal of Systems Science and Systems Engineering*.
- Azad, C., & Kumar, V. (2015). Genetic algorithm to solve the problem of small disjunct in the decision tree based intrusion detection system. *International Journal Computer Network and Information Security*, vol 8, pags 56-71.
- Bagnato, J. I. (11 de Septiembre de 2017). *7 pasos del Machine Learning para construir tu máquina*. Obtenido de Aprende Machine Learning: <http://www.aprendemachinelearning.com/7-pasos-machine-learning-construir-maquina/>
- BBC. (2019). *The history of Machine Learning*. Obtenido de BBC Academy: <https://www.bbc.com/timelines/zydp97h>
- Bitbrain. (25 de Mayo de 2018). *¿Es posible una inteligencia artificial con emociones y sentimientos?* Obtenido de Bitbrain: <https://www.bitbrain.com/es/blog/inteligencia-artificial-emociones>
- Blanco, R., Fontrodona, J., & Poveda, C. (2017). La industria 4.0: el estado de la cuestión. *Economía Industrial*, 151-164.
- Bortolini, M., Faccio, M., Gamberi, M., & Pilati, F. (2018). Motion Analysis System (MAS) for production and ergonomics assessment in the manufacturing processes. *Computers & Industrial Engineering*.

- Branco, A. (27 de Octubre de 2018). ¿Qué es la Inteligencia Artificial y cuáles son sus diferentes tipos? *Omicrono El Español*.
- Breunig, M., Kelly, R., Mathis, R., & Wee, D. (2016). *Industry 4.0 after the initial hype*. McKinsey Digital.
- Brynjolfsson, E., & McAfee, A. (2017). The Business of Artificial Intelligence. How AI Fits into Your Science Team. *Harvard Business Review*.
- Caminero, G. (27 de Junio de 2018). *Reinforcement learning como reacción frente a anomalías en la red*. Obtenido de Trabajo Fin de Máster en Ingeniería de Telecomunicaciones: <http://uvadoc.uva.es/bitstream/handle/10324/33081/TFM-G934.pdf?sequence=1&isAllowed=y>
- Carbonneau, R., Kersten, G., & Vahidov, R. (2008). Predicting opponent's moves in electronic negotiations using neural networks. *Expert Systems with Applications*, vol 34, pages 1266-1273.
- Çaydas, U., & Hasçalik, A. (2008). A study on surface roughness in abrasive waterjet machining process using artificial neural networks and regression analysis method. *Journal of Materials Processing Technology*, Vol 202, pages 574-582.
- Chandupatla, A. (9 de Febrero de 2019). *Accelerating Neural Networks using PCA*. Obtenido de Medium.
- Chandupatla, A. (8 de Febrero de 2019). *Clustering - Unsupervised Machine Learning*. Obtenido de Towards Data Science: <https://towardsdatascience.com/clustering-unsupervised-machine-learning-8ebe089a1673>
- Chaves Palacios, J. (2004). Desarrollo tecnológico en la primera revolución industrial. *Norba*, 93-109.
- Chen, S., & Weiss, G. (2015). An approach to complex agent-based negotiations via effectively modeling unknown opponents. *Expert Systems with Applications*, vol 42, pages 2287-2304.
- Cho, H., Kim, D. P., Roh, K., & Hwang, W. (2018). 2D Barcode Detection using Images for Drone-assisted Inventory Management. *Software and Computer electronics, Ajou University, South Korea*.
- Chollet, F. (2017). *Deep Learning with Python*. MANNING. Obtenido de <https://www.manning.com/books/deep-learning-with-python>
- Comisión Europea. (2018). *2018 International Digital Economy and Society Index*. Luxemburgo.
- Comon, P. (1994). Independent component analysis, A new concept? *Signal Processing*, 287-314.
- Dar, P. (27 de Junio de 2018). 'AI Guardman' – A Machine Learning Application that uses Pose Estimation to Detect Shoplifters. Obtenido de Analytics Vidhya: <https://www.analyticsvidhya.com/blog/2018/06/ai-guardman-machine-learning-application-estimates-poses-detect-shoplifters/>
- Dastin, J. (10 <https://www.reuters.com/article/us-amazon-com-jobs-automation-insight/amazon-scrap-secret-ai-recruiting-tool-that-showed-bias-against-women-idUSKCN1MK08G> de Octubre de 2018). Amazon scraps secret AI recruiting tool that showed bias against women. *REUTERS. Business News*.

- Deng, L., & Yu, D. (2013). Deep Learning: Methods and Applications. *Foundations and Trends in Signal Processing*, Vol 7. Pags 197-387.
- DHL. (2014). *Self-driving vehicles in logistics*. Troisdorf, Germany: DHL Customer Solutions & Innovation.
- Diario Expansión. (Marzo de 2019). *Índice de Competitividad Global*. Obtenido de Datosmacro.com: <https://datosmacro.expansion.com/estado/indice-competitividad-global/alemania>
- Dormehl, L. (5 de Enero de 2019). *What is an artificial neural network?* Obtenido de DIGITAL TRENDS: <https://www.digitaltrends.com/cool-tech/what-is-an-artificial-neural-network/>
- Droniter. (22 de Julio de 2019). *Termografía desde drones*. Obtenido de Droniter: <http://www.droniter.com/blog/termograf-a-desde-drones>
- Duhalde, F. (10 de Febrero de 2017). *Audi usará drones dentro de la fábrica*. Obtenido de Motor Trend: [https://motortrend.cienradios.com/test\\_drive/audio-usara-drones-dentro-de-la-fabrica/](https://motortrend.cienradios.com/test_drive/audio-usara-drones-dentro-de-la-fabrica/)
- Endesa Educa. (2014). *Smart Grids*. Obtenido de Endesa Educa: [https://www.endesaeduca.com/Endesa\\_educa/recursos-interactivos/smart-city/smart-grid](https://www.endesaeduca.com/Endesa_educa/recursos-interactivos/smart-city/smart-grid)
- Endesa Educa. (2014). *Smart Metering*. Obtenido de Endesa Educa: [https://www.endesaeduca.com/Endesa\\_educa/recursos-interactivos/smart-city/smart-metering](https://www.endesaeduca.com/Endesa_educa/recursos-interactivos/smart-city/smart-metering)
- eSMARTCITY. (7 de Febrero de 2014). *Innolid 2020, la propuesta de Valladolid como Smart City*. Obtenido de esmartcity.es: <https://www.esmartcity.es/2014/02/07/innolid-2020-la-propuesta-de-valladolid-como-smart-city>
- Esquivel, E. (29 de Agosto de 2018). *El consumidor 4.0*. Obtenido de <http://eduesquivel.com/el-consumidor-4-0/>
- Esteve, J. (11 de Julio de 2017). *Introducción al Machine Learning*. Obtenido de LinkedIn SlideShare: <https://es.slideshare.net/JavierEsteveMeli/introduccion-al-machine-learning-77766343>
- everis Spain. (2018). *Estudio Smart Industry 4.0*.
- Fernández, Ó. (2018). Sistema de detección de señales y vehículos mediante Redes Neuronales Convolucionales. *Trabajo Fin de Grado en Ingeniería Informática. Universidad de Valladolid. Valladolid*.
- Friedrich-Ebert-Stiftung. (2016). *La estrategia alemana Industria 4.0: el capitalismo renano en la era de la digitalización*. Madrid.
- Fuiza, M., & Rodríguez, J. (2000). La regresión logística: una herramienta versátil. *Revista Nefrología*, Vol. 20. Núm. 6 págs 477-565.
- Gandhi, R. (7 de Junio de 2018). *Support Vector Machine-Introduction to Machine Learning Algorithms*. Obtenido de Towards Data Science: <https://towardsdatascience.com/support-vector-machine-introduction-to-machine-learning-algorithms-934a444fca47>
- Garbisu, H. (2016). Utilización de Redes Neuronales Recurrentes para el análisis de firmas. *Trabajo Fin de Grado en Ingeniería Informática. Universidad de Las Palmas de Gran Canaria*.

- García Samartino, P. (2 de Mayo de 2018). *Logística 4.0: La revolución tecnológica de la cadena de suministro*. Obtenido de Izertis: <https://transformaciondigital.izertis.com/blog/logistica-4.0-la-revolucion-tecnologica-en-la-cadena-de-suministro>
- García, I. (26 de Enero de 2013). Ciudades inteligentes: gobierno participativo. *Nuevatribuna.es*.
- Gartner. (18 de Julio de 2017). *Gartner Says AI Technologies Will Be in Almost Every New Software Product by 2020*. Obtenido de Gartner: <https://www.gartner.com/en/newsroom/press-releases/2017-07-18-gartner-says-ai-technologies-will-be-in-almost-every-new-software-product-by-2020>
- Germany Trade & Invest. (2014). *Industrie 4.0 - Smart Manufacturing for the Future*. Berlin.
- Gesing, B., Peterson, S., & Michelsen, D. (2018). *Artificial Intelligence in Logistics*. Troisdorf, Germany: DHL Customer Solutions & Innovation.
- Glavas, H., Vidakovic, D., Jersek, Z., & Kraus, Z. (2018). Infrared Thermography in Maintenance of Building Applied Photovoltaics. *Journal of Energy*, volume 67, numer 4.
- Golob, D., & Osterman, D. (2008). Determination of pigment combinations for textile printing using artificial neural networks. *Fibres and Textiles in Eastern Europe*, vol 16, pags 93-98.
- González, L. (23 de Marzo de 2018). *Aprendizaje No Supervisado: K-Means Clustering*. Obtenido de Aprende tofo sobre inteligencia artificial: <http://ligdigonzalez.com/aprendizaje-no-supervisado-k-means-clustering/>
- González, L. (22 de Marzo de 2018). *Aprendizaje Supervisado: Linear Regression*. Obtenido de <http://ligdigonzalez.com/aprendizaje-supervisado-linear-regression/#comment-5873>
- González, L. (22 de Marzo de 2018). *Aprendizaje Supervisado: Polynomial Regression*. Obtenido de Aprende todo sobre inteligencia artificial: <http://ligdigonzalez.com/aprendizaje-supervisado-polynomial-regression/>
- González, L. (22 de Marzo de 2018). *Historia de Machine Learning*. Obtenido de Ligdi González. Aprende todo sobre inteligencia artificial: <http://ligdigonzalez.com/historia-de-machine-learning/>
- González, L. (9 de Noviembre de 2018). *Ligdi González. Aprende todo sobre inteligencia artificial*. Obtenido de Introducción a Bias y Varianza: <http://ligdigonzalez.com/bias-y-varianza-en-machine-learning/>
- Goodfellow, I. (2014). *Generative Adversarial Nets*. *Université de Montréal*.
- Goodfellow, I., Bengio, Y., & Courville, A. (2016). *Deep Learning*. MIT Press.
- GreyOrange. (2019). *GreyMatter*. Obtenido de Decision Science Driven Robotic Goods-to-Person System: <https://www.greyorange.com/butler-goods-to-person-system>
- Guo, J., Wu, J., Sun, Z., Long, J., & Zhang, S. (2019). Fault Diagnosis of Delta 3D Printers Using Transfer Support Vector Machine With Attitude Signals. *SPECIAL SECTION ON ADVANCES IN PROGNOSTICS AND SYSTEM HEALTH MANAGEMENT*, vol 7, pags 40359-40368.
- Gutiérrez, D. (17 de Junio de 2019). *Volvo Vera: el camión eléctrico y autónomo de Volvo ya trabaja en el puerto de Gotemburgo*. Obtenido de Híbridos y eléctricos:

- <https://www.hibridosyelectricos.com/articulo/actualidad/camion-autonomo-volvo-vera-trabajando-puerto-gotemburgo/20190617181146028326.html>
- Handa, A., Sharma, A., & Shukla, S. (2019). Machine Learning in cybersecurity: A review. *WIREs Data Mining and knowledge discovery*.
- Hayashi, C. (1998). What is Data Science? Fundamental Concepts and a Heuristic Example. *The Institute of Statistical Mathematics*.
- Hayes, A. (30 de Abril de 2019). *Bayes' Theorem Definition*. Obtenido de Investopedia: <https://www.investopedia.com/terms/b/bayes-theorem.asp>
- HBM. (2019). *El internet industrial de las cosas*. Obtenido de <https://www.hbm.com/es/6264/internet-industrial-de-las-cosas/>
- Hommax Sistemas. (19 de Julio de 2019). *Deep Learning, el siguiente paso en seguridad*. Obtenido de <https://www.hommaxsistemas.com/deep-learning-seguridad-inteligente/>
- Huang, S. (28 de Enero de 2018). *Introduction to Recommender System. Part 1 (Collaborative Filtering, Singular Value Decomposition)*. Obtenido de Hackernoon: <https://hackernoon.com/introduction-to-recommender-system-part-1-collaborative-filtering-singular-value-decomposition-44c9659c5e75>
- Huang, S. (12 de Enero de 2018). *Introduction to Various Reinforcement Learning Algorithms. Part I (Q-Learning, SARSA, DQN, DDPG)*. Obtenido de Towards Data Science: <https://towardsdatascience.com/introduction-to-various-reinforcement-learning-algorithms-i-q-learning-sarsa-dqn-ddpg-72a5e0cb6287>
- Hyvärinen, A., & Karhunen, J. (2000). Independent component analysis: algorithms and applications. *Neural Networks. Helsinki University of Technology*, Nº13 pág 411-430.
- Hyvärinen, A., Karhunen, J., & Oja, E. (2004). *Independent Component Analysis*. John Wiley & Sons.
- IBM. (2016). *IBM Watson. La tecnología cognitiva que abre una nueva era de la computación*.
- IBM. (2019). *Cloud Computing; guía completa*. Obtenido de <https://www.ibm.com/es-es/cloud/learn/what-is-cloud-computing>
- IBM. (2019). *Speeding Customer Service by 99% with Watson*. Obtenido de IBM: <https://www.ibm.com/watson/autodesk/>
- Infinite Analytics. (2019). *Infinite Analytics Home Page*.
- i-SCOOP. (Marzo de 2019). *Industry 4.0: the fourth industrial revolution - guide to Industrie 4.0*. Obtenido de [https://www.i-scoop.eu/industry-4-0/#Industry\\_40\\_and\\_technologies](https://www.i-scoop.eu/industry-4-0/#Industry_40_and_technologies)
- Izertis. (22 de mayo de 2018). *¿Qué esperar de Blockchain en la industria 4.0?* Obtenido de <https://transformaciondigital.izertis.com/blog/que-esperar-de-blockchain-en-la-industria-4.0>
- Jumbo, G. N. (Febrero de 2019). *La inteligencia artificial y su impacto en los procesos de selección*. Loja, Ecuador: Universidad Internacional de La Rioja.
- Kapur, R. (22 de Enero de 2016). *Logistic Regression case study - diagnosing cancer*. Obtenido de A Year Of AI: <https://ayearofai.com/rohan-1-when-would-i-even-use-a-quadratic-equation-in-the-real-world-13f379edab3b>

- Karras, T., Aila, T., Laine, S., & Lehtinen, J. (2018). Progressive Growing of GANs for Improved Quality, Stability and Variation. *Published as a conference paper at ICLR 2018*.
- Kavousi-Fard, A., Samet, H., & Marzbani, F. (2014). A new modified firefly algorithm and support vector regression model for accurate short term load forecasting. *Expert System with Applications*, Vol 41, pags 6047-6056.
- Kite-Powell, J. (4 de Abril de 2017). *This Recycling Robot Uses Artificial Intelligence To Sort Your Recyclables*. Obtenido de Forbes: <https://www.forbes.com/sites/jenniferhicks/2017/04/04/this-recycling-robot-uses-artificial-intelligence-to-sort-your-recyclables/#aa3ad8e2d353>
- KLM Royal Deutch Airlines. (2019). *Meet BB. KLM's smart assistant*. Obtenido de KLM: <https://bb.klm.com/en>
- Kobie, N. (8 de Diciembre de 2018). Amazon's HR proves artificial intelligence is truly dumb. *ITPRO. IT Analysis, Business Insight*, págs. <https://www.itpro.co.uk/machine-learning/32505/amazon-s-hr-proves-artificial-intelligence-is-truly-dumb>.
- Kozyrkov, C. (22 de Diciembre de 2018). *¿Qué diablos es Ciencia de Datos? En búsqueda de una definición útil*. Obtenido de Ciencia & Datos: <https://medium.com/datos-y-ciencia/qu%C3%A9-diablos-es-ciencia-de-datos-f1c8c7add107>
- Krizhevsky, A., Sutskever, I., & Hinton, G. (2012). ImageNet Classification with Deep Convolutional Neural Networks. *University of Toronto*.
- Kumar, A., & Gelin, R. (2018). A Mass-Produced Sociable Humanoid Robot. *IEEE Robotics & Automation Magazine*, 40-48.
- La Razón. (20 de Marzo de 2019). IA y machine learning para optimizar el transporte de mercancías por carretera. *La Razón*.
- Laguens, J. L. (12 de Enero de 2018). *La industria ajustada 4.0 puede reducir los costes del sector hasta un 40%*. Obtenido de El empresario: [https://www.elpresario.com/noticias/economia/2018/01/12/la\\_integracion\\_fabricacion\\_ajustada\\_con\\_industria\\_puede\\_reducir\\_los\\_costes\\_hasta\\_40\\_69984\\_1098.html](https://www.elpresario.com/noticias/economia/2018/01/12/la_integracion_fabricacion_ajustada_con_industria_puede_reducir_los_costes_hasta_40_69984_1098.html)
- Laha, D., Ren, Y., & Suganthan, P. (2015). Modeling of steelmaking process with effective machine learning techniques. *Expert Systems with Applications*, vol 42, pags 4687-4696.
- Lariviere, B., & Van den Poel, D. (2005). Predicting customer retention and profitability by random forests and regression forests techniques. *Expert Systems with Applications*, Vol 29, pags 472-484.
- Lee. (29 de Abril de 2016). *John Deere and IBM Pilot to Create a Futuristic Model for Manufacturing*. Obtenido de World Industrial Reporter: <https://worldindustrialreporter.com/john-deere-ibm-pilot-create-futuristic-model-manufacturing/>
- Lee, C., Lv, Y., Ng, K., Ho, W., & Choy, K. (2017). Design and application of Internet of thingsbased warehouse management system for smart. *International Journal of Production Research*.
- Li, K., Zhang, R., Li, F., Su, L., Wang, H., & Chen, P. (2019). A New Rotation Machinery Fault Diagnosis Method Based on Deep Structure and Sparse Least Squared Support Vector Machine.

*SPECIAL SECTION ON ADVANCES IN PROGNOSTICS AND SYSTEM HEALTH MANAGEMENT*, vol 7, pag 26571-26580.

- LKS. (2019). *La Industria 4.0 se basará en las personas y en la tecnología*. Obtenido de Think Up LKS: <http://www.thinkuplks.com/la-industria-4-0-se-basara-en-las-personas-y-en-la-tecnologia/>
- Locus Robotics. (22 de Julio de 2019). *Collaborative, autonomous robots that just work*. Obtenido de Locus Robotics: <https://www.locusrobotics.com/features/autonomous-robots/>
- López. (30 de Abril de 2017). *Este es el primer almacén robotizado de Amazon en España*. Obtenido de Xataka: <https://www.xataka.com/robotica-e-ia/este-es-el-primer-almacen-robotizado-de-amazon-en-espana>
- López, M., Restrepo, L., & López, G. (2013). Resistencia al cambio en organizaciones modernas. *Scientia et Technica*, Vol 18. No 1.
- Lowe's Companies. (30 de Agosto de 2016). *Lowe's introduces LoweBot - the next generation robot to enhance the home improvement shopping experience in the bay area*. Obtenido de PR Newswire: <https://www.prnewswire.com/news-releases/lowes-introduces-lowebot---the-next-generation-robot-to-enhance-the-home-improvement-shopping-experience-in-the-bay-area-300319497.html>
- Luna, J. (8 de Febrero de 2018). *Tipos de aprendizaje automático*. Obtenido de SoldAI: <https://medium.com/soldai/tipos-de-aprendizaje-automático-6413e3c615e2>
- Management Solutions España. (2018). *Machine Learning, una pieza clave en la transformación de los modelos de negocio*.
- Marín, J. M. (2013). *Tema 3: Análisis de Componentes*. Obtenido de Universidad Carlos Tercero de Madrid: <http://halweb.uc3m.es/esp/Personal/personas/jmmarin/esp/AMult/tema3am.pdf>
- Márquez, N. (2016). *Minería de Datos vs Grandes Datos*. Obtenido de TUATARA TECH: <http://www.tuataratech.com/2015/06/mineria-de-datos-data-mining-vs-grandes.html>
- Marr, B. (2016). A short history of Machine Learning. *Forbes*.
- Martín, E. (28 de Noviembre de 2017). Por qué "machine learning" será la tecnología más importante en 2018. *El País*.
- Martínez, J. (21 de Diciembre de 2018). *Fases del Proceso de Machine Learning*. Obtenido de IArtificial.net: <https://iartificial.net/fases-del-proceso-de-machine-learning/>
- Martínez, J. (28 de Marzo de 2019). *Redes Neuronales Generativas Adversarias (GANs)*. Obtenido de IArtificial.net: <https://iartificial.net/redes-neuronales-generativas-adversarias-gans/>
- McClendon, L., & Meghanathan, N. (2015). Using Machine Learning Algorithms to analyze crime data. *Machine Learning and Applications: an international journal*, Vol 2, No 1.
- Minsky, M., & Papert, S. (1969). *Perceptrons: an introduction to computational geometry*. Cambridge MA: MIT Press.
- Mirapeix, J., García-Allende, P., Cobo, A., Conde, O., & López-Higuera, J. (2007). Real-time arc-welding defect detection and classification with principal component analysis and artificial neural networks. *NDT&E International*, vol 40, pags 315-323.

- Mitchell, T. (1997). *Machine Learning*. Portland: McGraw-Hill.
- Monostori, L. (2003). AI and machine learning techniques for managing complexity. *Engineering Applications of Artificial Intelligence*, Vol 16, pags 277-291.
- Mozer, M. (1989). A Focused Backpropagation Algorithm. *Complex Systems* 3, 349-381.
- Nath, N., Chaspari, T., & Behzadan, A. (2018). Automated ergonomic risk monitoring using body-mounted sensors and machine learning. *Advanced Engineering Informatics*, 514-526.
- Negnevitsky, M. (2002). *Artificial Intelligence: a guide to intelligent systems*. Hobart: Pearson Education Limited.
- Neosentec. (29 de Noviembre de 2017). *7 ventajas de la realidad aumentada en empresas*. Obtenido de Neosentec blog: <https://www.neosentec.com/7-ventajas-realidad-aumentada-empresas/>
- Ng, A. (2012). *K-Means Algorithm*. Obtenido de Machine Learning. Coursera. Stanford University: <https://www.coursera.org/learn/machine-learning/lecture/93VPG/k-means-algorithm>
- Ng, A. (2012). *The problem of Overfitting*. Obtenido de Machine Learning. Coursera. Stanford University.: <https://www.coursera.org/learn/machine-learning/lecture/ACpTQ/the-problem-of-overfitting>
- OBS Business School. (2019). *Almacenes 4.0: La automatización*. Obtenido de OBS Blogs investigación: <https://www.obs-edu.com/es/blog-investigacion/logistica/almacenes-40-la-automatizacion>
- Observatorio Vodafone de la Empresa. (2017). *Estudio sobre el estado de digitalización de las empresas y Administraciones públicas españolas*.
- Oliva, E. D. (Enero de 2018). *Industria 4.0: Retos y Oportunidades en las factorías de Automoción*. Valladolid.
- Olivares, V., Cordova, F., Sepúlveda, J., & Derpich, I. (2015). MODELING INTERNAL LOGISTICS BY USING DRONES ON THE STAGE OF ASSEMBLY OF PRODUCTS. *Information Technology and Quantitative Management*, 1240-1249.
- Oliver, J. (1996). A Machine Learning Approach to Automated Negotiation and Prospects for Electronic Commerce. *Journal of Management Information Systems*.
- Orozco, J. A. (30 de agosto de 2018). *Machine Learning y su importancia en la actualidad*. Obtenido de IPADE Business School: <https://www.ipade.mx/2018/08/30/machine-learning-y-su-importancia-en-la-actualidad/>
- Oxford Dictionaries. (2019). *Definition of Markov chain in Us English by Oxford Dictionaries*. Obtenido de [https://en.oxforddictionaries.com/definition/us/markov\\_chain](https://en.oxforddictionaries.com/definition/us/markov_chain)
- Peres-Neto, P., Jackson, D., & Somers, K. (2004). How many principal components? stopping rules fordetermining the number of non-trivial axes revisited. *Computational Statistics & Data Analysis*, 974-977.
- Pérez, E. (9 de Abril de 2019). *Los drones mensajeros ya son una realidad: Google y Project Wing inician en Australia el primer servicio comercial del mundo*. Obtenido de XATAKA: <https://www.xataka.com/drones/drones-mensajeros-realidad-google-project-wing-inician-australia-primer-servicio-comercial-mundo>



- Pilati, F., & Regattieri, A. (2018). The impact of digital technologies and artificial intelligence on production systems in today Industry 4.0 environment. *Network Industries Quarterly*, Vol 20.
- PINC. (22 de julio de 2019). *Warehouse drones: real-time inventory tracking by air*. Obtenido de PINC: <https://www.pinc.com/warehouse-drone-inventory-management>
- Pinto, A., Rocha, L., & Moreira, P. (2013). Object recognition using laser range finder and machine learning techniques. *Robotics and Computer-Integrated Manufacturing*, vol 29, pags 12-22.
- Polacco, A., & Backes, K. (2018). The Amazon Go Concept: Implications, Applications, and Sustainability. *Journal of Business and Management*, vol 24, pags 79-92.
- Pontes, F., de Paiva, A., Belestrassi, P., Ferreira, J., & da Silva, M. (2012). Optimization of Radial Basis Function neural network employed for prediction of surface roughness in hard turning process using Taguchi's orthogonal arrays. *Expert Systems and Applications*, vol 39, pags 7776-7787.
- Power Data. (2019). *Big Data: ¿En qué consiste? Su importancia, desafíos y gobernabilidad*.
- Prevent Security Systems. (23 de Julio de 2019). *Seguridad y videovigilancia con Drones*. Obtenido de Prevent Security Systems: <https://www.prevent.es/servicios-de-seguridad/camaras-de-seguridad/empresas/seguridad-y-videovigilancia-con-drones>
- Prometeus Global Solutions. (19 de Febrero de 2019). *Volumen, Variedad, Velocidad, Veracidad y Valor, las 5 dimensiones del Big Data*. Obtenido de <https://prometeusgs.com/volumen-variedad-velocidad-veracidad-y-valor-las-5-dimensiones-del-big-data-la/>
- Redondo, R., Melchor, N., & Quintela, F. (20 de Septiembre de 2012). *Conceptos de gradiente y de derivada direccional*. Obtenido de Universidad de Salamanca: <http://electricidad.usal.es/Principal/Circuitos/Comentarios/Temas/ConceptoGradiente.pdf>
- Rejoiner. (2019). *The Amazon Recommendations Secret to Selling More Online*. Obtenido de Rejoiner: <http://rejoiner.com/resources/amazon-recommendations-secret-selling-online/>
- Riquelme, M. (23 de Septiembre de 2018). *Tercera Revolución Industrial*. Obtenido de Web y Empresas: <http://www.escuelpedia.com/primerasegunda-ytercera-revolucion-industrial/>
- Rodríguez, D. (17 de Agosto de 2018). *Seis aplicaciones del Aprendizaje Automático en la Industria 4.0*. Obtenido de Analytics Lane: <https://www.analyticslane.com/2018/08/17/seis-aplicaciones-del-aprendizaje-automatico-en-la-industria-4-0/>
- Rojko, A. (2017). Industry 4.0 Concept: Background and Overview. *International Journal of Interactive Mobile Technologies (IJIM)*, Vol 11, No 5.
- Roland Berger. (2016). *España 4.0. El reto de la transformación digital de la economía*. Madrid.
- Román, V. (6 de Febrero de 2019). *Introducción al Machine Learning: Una Guía Desde Cero*. Obtenido de CIENCIA & DATOS: <https://medium.com/datos-y-ciencia/introduccion-al-machine-learning-una-gu%C3%ADa-desde-cero-b696a2ead359>
- Román, V. (31 de Enero de 2019). *Supervised Learning: Basics of Classification and Main Algorithms*. Obtenido de Towards Data Science: <https://towardsdatascience.com/supervised-learning-basics-of-classification-and-main-algorithms-c16b06806cd3>

- Ruiz, S. (20 de Julio de 2017). *El algoritmo K-NN y su importancia en el modelado de datos*. Obtenido de AD TECH & ANALYTICS: <https://www.analiticaweb.es/algoritmo-knn-modelado-datos/>
- Rumelhart, D., Hinton, G., & Williams, R. (1986). Learning representations by back-propagatin errors. *Letters to nature*, Vol. 3, 533-536.
- Russell, S. J., & Norvig, P. (2010). *Artificial Intelligence: A Modern Approach*. Prentice Hall.
- Saha, S. (15 de Diciembre de 2018). *A Comprehensive Guide to Convolutional Neural Networks*. Obtenido de Towards Data Science: <https://towardsdatascience.com/a-comprehensive-guide-to-convolutional-neural-networks-the-eli5-way-3bd2b1164a53>
- Salas, M. (1996). La regresión logística. Una aplicación a la demanda de estudios universitarios. *Departamento de Economía Aplicada. Universidad de Granada*, Vol. 38, Núm 141, págs. 193 a 217.
- Samaniego, J. (30 de Mayo de 2019). *El protagonismo decisivo de los cobots en la industria 4.0*. Obtenido de Hablemos de empresas: <https://hablemosdeempresas.com/grandes-empresas/que-son-los-cobots/>
- Sánchez, M. (24 de Enero de 2019). *¿Cómo hacer que tu proyecto de Machine Learning tenga éxito?* Obtenido de mc.ai: <https://mc.ai/como-hacer-que-tu-proyecto-de-machine-learning-tenga-exito/>
- Santana, C. (16 de Diciembre de 2017). *Regresión Lineal y Mínimos Cuadrados Ordinarios*. Obtenido de Canal de YouTube de DotCSV: [https://www.youtube.com/watch?v=k964\\_uNn3I0](https://www.youtube.com/watch?v=k964_uNn3I0)
- Santana, C. (4 de Febrero de 2018). *¿Qué es el Descenso del Gradiente? Algoritmo de Inteligencia Artificial*. Obtenido de Canal de YouTube de DotCSV: [https://www.youtube.com/watch?v=A6FiCDoz8\\_4](https://www.youtube.com/watch?v=A6FiCDoz8_4)
- Santana, C. (3 de Octubre de 2018). *¿Qué es una Red Neuronal? Parte 3: Backpropagation*. Obtenido de Canal de YouTube de DotCSV: [https://www.youtube.com/watch?v=eNIqz\\_noix8](https://www.youtube.com/watch?v=eNIqz_noix8)
- Santana, C. (1 de Marzo de 2019). *¿Qué veía Claude Monet mientras pintaba en 1873? - CycleGAN* . Obtenido de Canal de YouTube de DotCSV: <https://www.youtube.com/watch?v=wHLshgKr8ms>
- Santana, C. (19 de Enero de 2019). *Overfitting y Underfitting*. Obtenido de Canal de YouTube de DotCSV: <https://www.youtube.com/watch?v=7-6X3DTt3R8>
- Saxena, H., & Richariya, V. (2014). Intrusion detection in KDD99 dataset using SVM-PSO and feature reduction with information gain. *International Journal of Computer Applications*, vol 98, pags 25-29.
- Schilperoort, L. (13 de Septiembre de 2018). *First Smart Item Robotics (SIR) application in Finland*. Obtenido de Smart Robotics: <https://www.smart-robotics.nl/first-smart-item-robotics-sir-application-in-finland/?lang=es>
- Seebo. (2019). *Machine Learning and AI in manufacturing*. Obtenido de Seebo.com: <https://www.seebo.com/machine-learning-ai-manufacturing/>
- Shalev-Shwartz, S., & Ben-David, S. (2014). *Understanding machine learning: from theory to algorithms*. Nueva York: Cambridge University press.

- Shapoorifard, H., & Shamsinejad, P. (2017). Intrusion Detection using a Novel Hybrid Method Incorporating an Improved KNN. *International Journal of Computer Applications*.
- Siemens España. (17 de Junio de 2015). *La salud 4.0 en la era de la digitalización*. Obtenido de Ciudades del futuro: <https://ciudadesdelfuturo.es/la-salud-4-0-en-la-era-de-la-digitalizacion.php>
- Siemens España. (Febrero de 2019). *El Futuro de la Industria: Digitalización Industrial*. Obtenido de [https://w5.siemens.com/spain/web/es/el-futuro-de-la-industria/pages/el\\_futuro\\_de\\_la\\_industria.aspx](https://w5.siemens.com/spain/web/es/el-futuro-de-la-industria/pages/el_futuro_de_la_industria.aspx)
- Simard, P., Steinkraus, D., & Platt, J. (2003). Best Practices for Convolutional Neural Networks Applied to Visual Document Analysis. *Microsoft Research*.
- Simbe Robotics. (2019). *Simbe Robotics*. Obtenido de Say Hello to Tally: <https://www.simberobotics.com/platform/intelligence/>
- Song, S. Y. (2017). Modeling the customer experience of retail service robots. *University of Tennessee*, PhD.
- Standard & Poor's. (2018). The Road Ahead For Autonomous Vehicles. *RatingsDirect*.
- Sumesh, A., Rameshkumar, K., Mohandas, K., & Shyam Babu, R. (2015). Use of Machine Learning Algorithms for Weld Quality Monitoring using Acoustic Signature. *Procedia Computer Science*, Vol 50, pags 316-322.
- Sutton, R., & Barto, A. (2018). *Reinforcement Learning: An Introduction*. Cambridge, Massachusetts: The MIT Press.
- Transportes Callizo. (8 de Noviembre de 2017). *Logística inteligente - Un nuevo modelo basado en la conectividad y el Big Data*. Obtenido de Transportes Callizo Web: <http://www.transportescallizo.com/logistica-inteligente-conectividad-bigdata/>
- Tsai, M., Li, C., & Chen, C. (2008). Optimal laser-cutting parameters for QFN packages by utilizing artificial neural networks and genetic algorithm. *Journal of Materials Processing Technology*, vol 208, pags 270-283.
- Tüfekci, P. (2014). Prediction of full load electrical power output of a base load operated combined cycle power plant using machine learning methods. *Electrical Power and Energy Systems*, vol 60, pags 126-140.
- Tusell, F. (2011). *Análisis de Regresión. Introducción teórica y práctica basada en R*. Bilbao.
- Underwood, C. (19 de Junio de 2019). *Robots in Retail - Examples of Real Industry Applications*. Obtenido de Emerj: <https://emerj.com/ai-sector-overviews/robots-in-retail-examples/>
- Universia. (7 de Marzo de 2017). *5 beneficios de usar la realidad virtual en el aula*. Obtenido de Universia Noticias: <http://noticias.universia.es/ciencia-tecnologia/noticia/2017/03/07/1150198/5-beneficios-usar-realidad-virtual-aula.html>
- Uriarte, J. M. (27 de Septiembre de 2017). *Segunda Revolución Industrial*. Obtenido de Caracteristicas.co: <https://www.caracteristicas.co/segunda-revolucion-industrial/>
- Van Pay, B. (30 de Septiembre de 2018). How Artificial Intelligence Is Reinventing Human Resources. *Entrepreneur*, pág. <https://www.entrepreneur.com/article/320763>.

- Velogig. (28 de Septiembre de 2018). *¿Qué es el Machine Learning y cómo es su proceso?* Obtenido de VELOGIG Marketing Científico: <https://velogig.com/que-es-el-machine-learning-y-como-es-su-proceso/>
- Venkatesh, N. (2018). Industry 4.0: Reimagining the Future of Workplace (Five Business Case Applications of Artificial Intelligence, Machine Learning, Robots, Virtual Reality in Five Different Industries). *International Journal of Engineering, Business and Enterprise Applications (IJEBA)*, págs 5-8.
- Villanueva, J. D. (2 de Febrero de 2019). *Redes Neuronales desde cero - Introducción*. Obtenido de IArtificial.net. Inteligencia Artificial y Machine Learning en Español: <https://iartificial.net/redes-neuronales-desde-cero-i-introduccion/>
- Villén, M. (14 de Abril de 2019). *Big Data Analytics y la inteligencia Artificial*. Obtenido de CAMINOS MADRID. Colegio de Ingenieros de Caminos, Canales y Puertos de Madrid.: <https://www.caminosmadrid.es/9938-2>
- Vincent, J. (5 de Junio de 2019). *Here's Amazon's new transforming Prime Air delivery drone*. Obtenido de The Verge: <https://www.theverge.com/2019/6/5/18654044/amazon-prime-air-delivery-drone-new-design-safety-transforming-flight-video>
- Violante, A. (18 de Marzo de 2019). *Simple Reinforcement Learning: Q-learning*. Obtenido de Towards Data Science: <https://towardsdatascience.com/simple-reinforcement-learning-q-learning-fcddc4b6fe56>
- Wang, Y.-H., Li, T.-H., & Lin, C.-J. (2013). Backward Q-learning: The combination of Sarsa algorithm and Q-learning. *Engineering Applications of Artificial Intelligence*, Volume 26, Pages 2184-2193.
- Williams, B. (7 de Agosto de 2017). *There are now over 3 billion social media users in the world — about 40 percent of the global population*. Obtenido de Mashable: <https://mashable.com/2017/08/07/3-billion-global-social-media-users/?europe=true>
- X Company. (2019). *Wing. Transforming the way goods are transported*. Obtenido de X Company. Project Wing: <https://x.company/projects/wing/>
- Xin, B., Li, Y., Qiu, J., & Liu, Y. (2012). Texture modelling of fabrics appearance evaluation based on image analysis. *Fibres Text East Eur*, Vol 20, págs 48-52.
- Xin, Y., Kong, L., Liu, Z., & Chen, Y. (2018). Machine Learning and Deep Learning Methods for Cybersecurity. *IEEE Access*, Vol 6, págs 35365-35381.
- Yiakopoulos, C., Gryllias, K., & Antoniadis, I. (2011). Rolling element bearing fault detection in industrial environments based on a K-means clustering approach. *Expert Systems with Applications*, vol 38 págs 2888-2911.
- Yildirim, P., Birant, D., & Alpyildiz, T. (2018). Data mining and machine learning in textile industry.
- Yu, Y., Long, J., & Cai, Z. (2017). Network Intrusion Detection through Stacking Dilated Convolutional Autoencoders. *Security and Communication Networks*, 1-10.
- Yupcharge. (2019). *Smart Consumer, nuevo modelo de consumidor*. Obtenido de yupcharge.com: <https://www.yupcharge.com/es/blog/smartconsumer-nuevo-concepto/>

- Zaremba, W., Sutskever, I., & Vinyals, O. (2015). RECURRENT NEURAL NETWORK REGULARIZATION.
- Zhang, J., & Yang, C. (2014). Evaluation model of color difference for dyed fabrics based on the support vector machine. vol 84, pags 2184-2197.
- Zhu, J.-Y., Park, T., Isola, P., & Efros, A. (2018). Unpaired Image-to-Image Translation using Cycle-Consistent Adversarial Networks. *Berkeley AI Research*.
- Zychlinski, S. (23 de Febrero de 2019). *The Complete Reinforcement Learning Dictionary*. Obtenido de Towards Data Science: <https://towardsdatascience.com/the-complete-reinforcement-learning-dictionary-e16230b7d24e>

