

UNIVERSIDAD MIGUEL HERNÁNDEZ

Facultad de Ciencias Sociales y Jurídicas de Orihuela
Administración y Dirección de Empresas

Machine Learning Y Los Mercados Financieros

Trabajo de Fin de Grado



Autor: Víctor Rafael Rodríguez Manresa
Tutora: Laura Antón Sánchez
Curso: 2020 - 2021

ÍNDICE

I. Introducción.....	5
II. Historia y evolución del Machine Learning.....	7
III. Modelos de Machine Learning.....	11
IV. Algoritmos más utilizados.....	16
V. Big Data y Machine Learning.....	18
VI. Beneficios del Machine Learning en los negocios.....	20
VII. Machine Learning y el mercado financiero.....	21
VII.1 Aplicaciones centradas en el cliente.....	21
VII.2 Aplicaciones centradas en las operaciones.....	23
VIII. Posibles efectos en los mercados financieros.....	28
IX. Ley y ética del Machine Learning.....	30
X. Futuro del Machine Learning.....	33
XI. Conclusiones.....	35
XII. Bibliografía.....	37

Resumen

Los términos Inteligencia Artificial y Machine Learning han ganado un enorme peso durante los últimos años, han sido introducidos en nuestro día a día mientras interactuamos con ellos, muchas veces sin darnos cuenta, por ejemplo, el hecho de desbloquear nuestro teléfono móvil con el reconocimiento facial es gracias a los avances en el Machine Learning. La Inteligencia Artificial y Machine Learning han ido ganando terreno en numerosos campos con aplicaciones en campos como la medicina, economía, negocios o automoción. La popularización de estos sistemas se debe en parte a que vivimos en un mundo en el que se genera una gran cantidad de información y datos debido a Internet y la hiperconectividad. Estas tecnologías son usadas, entre otras cosas, para mejorar la experiencia de los usuarios, por ejemplo, servicios de streaming como Netflix y de música como Spotify, que aprenden a identificar el contenido que le gusta al usuario y presentarle recomendaciones personalizadas gracias al Machine Learning.

En este trabajo se busca estudiar cómo está influenciando el Machine Learning el ámbito financiero. El Machine Learning ha supuesto un gran impacto en los sectores financieros, y estos, a su vez, son un pilar de la sociedad, abarcando una gran gama de negocios desde bancos hasta compañías de seguros pasando por compañías de contabilidad. En el ámbito financiero han aparecido en los últimos años muchos términos relativamente nuevos como Machine Learning, Deep Learning o Big Data, por lo que en este trabajo los identificaremos y diferenciaremos unos de otros. Es importante entender qué se está usando y de dónde viene, por tanto, primero analizaremos la evolución del Machine Learning desde mediados del siglo pasado,; una vez conocida su trayectoria es vital entender qué tipos de Machine Learning existen, como el supervisado o no supervisado, y cómo se pueden utilizar mediante algunos de los algoritmos más comunes. Qué beneficios puede aportar el Machine Learning en las empresas es un punto clave para pensar si quiera en su adopción; analizaremos cómo está relacionado con los mercados financieros y sus aplicaciones en esta área en la actualidad, diferenciando de las centradas en los clientes, como el Credit Scoring, de las centradas en las operaciones,

como el trading y gestión de cartera o los modelos de gestión de riesgo. Una vez conocidos sus aplicaciones y usos en los mercados financieros nos centraremos en sus posibles efectos en el sector financiero y las leyes a las que está sometido o, más bien, a las que podrá estar sometido en un futuro el Machine Learning, debido a que aún no se ha regulado de forma oficial. Para finalizar, hablaremos del futuro del Machine Learning y lo optimista que parece ser de cara a seguir desarrollándose e implantándose en los años próximos, acompañado de un aumento de la inversión en el Machine Learning, crecimiento del sector y creación de puestos de trabajo en este área.



I. Introducción

Machine Learning es un término que se usa comúnmente hoy en día en muchos contextos distintos juntos a otros como Big Data, Inteligencia Artificial o Deep Learning, incluso a veces se confunden unos con otros o no se tiene claro del todo su significado, simplemente se usan indiscriminadamente para un contexto en el que las máquinas, a través de datos, aprenden, evolucionan y ayudan o toman ellas mismas decisiones.

Si bien la definición anterior ayudaría a una persona que no entiende o nunca ha escuchados estos términos, es necesario empezar a definir y dar contexto a cada uno de ellos para poder entender cómo trabajan y se relacionan entre sí.

La base de la que partimos es la Inteligencia Artificial, en 1956 John McCarthy acuñó el término por primera vez y la definió como “la ciencia e ingeniería de crear máquinas inteligentes, especialmente programas de computación inteligentes”. Aunque no hay una definición definitiva, ya que depende de qué tipo de procesos computacionales queramos llamar inteligentes, podemos decir que la Inteligencia Artificial es la habilidad de una máquina para desarrollar o replicar tareas y decisiones igual o mejor que los humanos.

Algunos ejemplos de lo que nos referimos con desarrollar tareas mejor que los humanos serían diagnosticar si una persona tiene cáncer viendo una radiografía al igual que un experto, o AlphaGo, el primer programa informático de Inteligencia Artificial en derrotar al Go a un jugador profesional (Alameda, 2020).

Estos ejemplos podrían considerarse como Inteligencia Artificial débil o estrecha ya que se centran en una simple tarea estando lejos de lo que sería una Inteligencia Artificial fuerte, que busca replicar o mejorar cualquier tarea intelectual de un humano medio, siendo esto algo más cercano a lo que nos imaginamos cuando imaginamos un futuro en el que Inteligencia Artificial esté más desarrollada.

Aunque es incorrecto referirse a la Inteligencia Artificial como Machine Learning

la realidad es que hoy en día prácticamente todos los sistemas de Inteligencia Artificial dependen en gran medida del Machine Learning.

Machine Learning es una rama de la Inteligencia Artificial que permite que las máquinas aprendan a través de datos, identifican patrones para así poder realizar predicciones, todo esto con una mínima intervención humana. El científico informático Tom Mitchell la define como “El estudio de algoritmos de computación que mejoran automáticamente su rendimiento gracias a la experiencia. Se dice que un programa informático aprende sobre un conjunto de tareas, gracias a la experiencia y usando una medida de rendimiento, si su desempeño en estas tareas mejora con la experiencia” (Mitchell,1997).

La base del Machine Learning es la estadística ya que consiste en una serie de algoritmos capaces de analizar grandes cantidades de datos. “Cada paso en un proyecto de Machine Learning requiere el uso de un método estadístico. Tanto para comprender los datos utilizados en el entrenamiento de un modelo de aprendizaje automático como en la interpretación de los resultados obtenidos tras probar diferentes modelos de aprendizaje automático, se requieren métodos estadísticos” (Lopez & Ferraro, 2021).

Hoy en día el Machine Learning está presente en nuestras vidas sin darnos cuenta, desde aplicaciones que están en nuestros móviles como Siri o Alexa, la detección de rostro para desbloquear estos dispositivos hasta ámbitos como la medicina o la detección de fraudes para combatir el blanqueo de dinero.

Los campos más exitosos que se están desarrollando en la actualidad son:

- Big Data, es un término que describe un gran volumen de datos (estructurados y no estructurados) para el que su procesamiento resulta imposible por métodos tradicionales. A principio de la década de los 2000 el analista Doug Laney, articuló la definición actual de Big Data como las tres “V”, aunque hoy en día se consideran cinco “V”. Volumen que hace referencia al tamaño de los datos, por ejemplo, una empresa puede llegar a recolectar PetaBytes (1 millón de Gb) en información. Velocidad, hoy en día los datos llegan a las empresas a una velocidad

sin precedentes, cada segundo se generan millones de datos gracias Internet en gran medida. Variedad, los datos se presentan de muchas formas distintas, desde transacciones bancarias pasando por correos electrónicos hasta videos. Veracidad, la característica que supone un mayor reto ya que con la gran cantidad de datos que se recibe es complicado que muchos no lleguen incompletos o incorrectos. Valor, la oportunidad de sacar el máximo rendimiento a los datos. Combinar técnicas de Big Data como Machine Learning es la mejor manera de procesar estas 5 "V".

- Deep Learning, es un subcampo del Machine Learning que emula el aprendizaje humano al estar compuesto por redes neuronales artificiales entrelazadas para el procesamiento de la información. En lugar de organizar datos a través de ecuaciones predefinidas, el Deep Learning configura parámetros básicos acerca de los datos y entrena a la computadora para que aprenda por cuenta propia reconociendo patrones mediante el uso de muchas capas de procesamiento. Su uso en la actualidad es muy notable en los ámbitos en los que se aplica tales como reconocimiento del habla, reconocimiento de imágenes o conducción autónoma (Kampakis, 2020).

II. Historia y evolución del Machine Learning

Es necesario dar contexto a cómo se ha ido desarrollando y evolucionando para poder entender cómo se usan hoy en día estas herramientas y la importancia que tienen y, sobre todo, tendrán en un futuro.

En el año 1943 el matemático Walter Pitts y el neurólogo Warren McCulloch, dieron a conocer su trabajo en el que proponían analizar el cerebro como un organismo computacional y la creación de computadoras que funcionaran igual o mejor que nuestra red neuronal, esto es lo que hoy en día conocemos como Inteligencia Artificial (Council of Europe [COE], 2021).

El término Inteligencia Artificial nace en 1956 de la mano de John McCarthy pero hay que retroceder unos cuantos años, donde Alan Turing, considerado

uno de los padres de la ciencias de la computación y percursor de la informática moderna, en su artículo “Computing machinery and intelligence” publicado en 1950, ya se planteaba la posibilidad de crear máquinas capaces de pensar. Este matemático inglés ingenió una prueba conocida hoy en día como Test de Turing, un test creado para poder conocer si una máquina es inteligente. La Figura 1 muestra cómo funciona el Test de Turing.

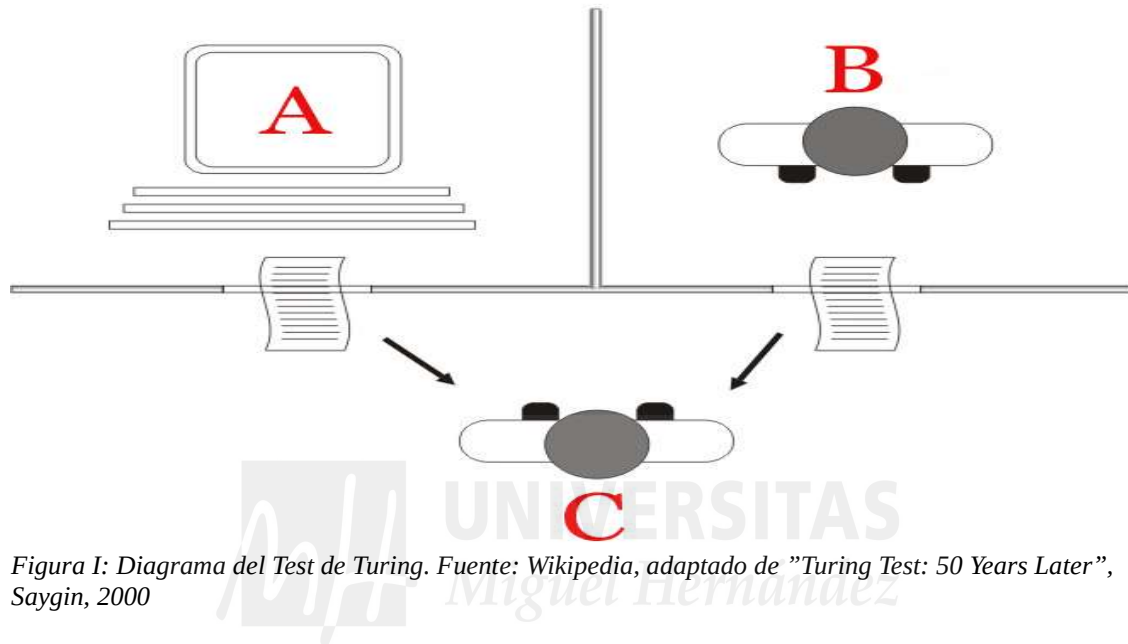


Figura 1: Diagrama del Test de Turing. Fuente: Wikipedia, adaptado de "Turing Test: 50 Years Later", Saygin, 2000

El Test de Turing consiste en tres participantes, una máquina, un humano y un interrogador humano mostrados en la Figura 1 como A, B y C, respectivamente. El interrogador C debe ser capaz de distinguir mediante una serie de preguntas cuál de los dos participantes, A o B, es una máquina. Si el interrogador no es capaz de distinguir cual es la máquina, esta última sería considerada inteligente al ser sus respuestas indistinguibles de las de un humano (Encyclopedia Britannica, s.f).

Arthur Samuel un desarrollador que trabajaba para IBM desarrolló en 1952 un programa para jugar a las damas. Este programa incorporaba una función de puntuación donde intentaba medir las posibilidades de que cada bando ganara, eligiendo el programa su siguiente movimiento como una estrategia llamada minimax, que más tarde se convirtió en el algoritmo minimax.

La primera persona que inventó el término “Machine Learning” también fue

Samuel, ya que su programa incluía un mecanismo por el que mejoraba jugando a las damas. El programa guardaba y recordaba todas las posiciones que ya había visto con anterioridad y, junto con los valores de una función de recompensa surgió este término.

En la conferencia de Darmouth en 1956 nace el término Inteligencia Artificial por John McCarthy con la ayuda de Mavin Minsky y Claude Shannon. Desde esta fecha hasta los 70 se avanza en la Inteligencia Artificial con aportaciones que ayudarán a su desarrollo, como el LISP por John McCarthy, el primer lenguaje que logró un procesamiento simbólico y no numérico, o el diseño de la primera red neuronal artificial por Frank Rosenbalt (Caparrini, s.f).

Frank Rosenbalt construyó el Perceptron, originalmente planteado como una máquina, este software se instaló en una máquina hecha a medida que fue construida para el reconocimiento de imágenes. Fue descrito como el primer neuroordenador y, aunque en principio se podría calificar de éxito, el Perceptron no conseguía reconocer muchos tipos de patrones visuales como las caras, provocando una ruptura de las expectativas. Este hecho hizo que la investigación sobre Machine Learning fuese más lenta hasta que logró resurgir un par de décadas después.

A mediados de los 70 llegó un parón en los avances debido a la limitada capacidad computacional y potencia de procesamiento de las computadoras de la época. El no cumplirse las expectativas depositadas en este campo conllevó una pérdida de financiación. Además, acompañado de esta ralentización en los avances, empezaron a surgir críticas de otros campos científicos hacia la Inteligencia Artificial, críticas y dudas que han perdurado a lo largo de los años hasta la actualidad.

Por suerte, en 1967 surgió el algoritmo "Nearest Neighbor", que supuso el comienzo del reconocimiento básico de patrones. En esta misma década se descubrió el uso de las multicapas. Al utilizar múltiples capas, programas como el Perceptron se volvían mucho más precisos en su tarea al ofrecer una mayor potencia de procesamiento que con el uso de una sola capa. El uso de estas múltiples capas dio lugar a las redes neuronales feedforward y a la

retropropagación.

A principio de los 80 ocurrió lo que podría denominarse como una ruptura parcial entre la Inteligencia Artificial y el Machine Learning. La Inteligencia Artificial, centrada en esa época en el uso de la lógica, utilizaba el Machine Learning como una forma de entrenar los programas. La industria del Machine Learning cambió de rumbo, dejó de buscar ser un simple método de entrenar programas para la Inteligencia Artificial y se centró en buscar soluciones a problemas prácticos, ganando importancia teorías de la probabilidad y la estadística como herramienta para solucionar estos problemas.

En la década de los 90 tuvo un resurgimiento gracias a la llegada de Internet, beneficiándose de todos los datos digitales disponibles y de la capacidad de compartir a través de Internet. En este contexto surgió el Boosting, necesario para la evolución del Machine Learning. El Boosting es un meta-algoritmo que reduce la varianza y el sesgo en un contexto de aprendizaje supervisado combinando los resultados de varios modelos “débiles” creando un modelo “fuerte”. Un ejemplo claro de como el Boosting funciona sería coger un modelo de regresión lineal, un modelo de árbol de decisiones y un clasificador de Naive Bayes, siendo estos los modelos “débiles” y usando una combinación de estos modelos nos daría el resultado final denominado modelo “fuerte”, como se ilustra en la Figura 2. También ayudó a popularizar este campo el ordenador Deep Blue venciendo al campeón de ajedrez del mundo Gary Kaspárov en 1997.

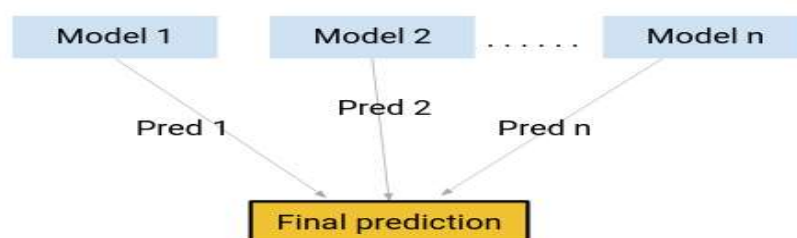


Figura II: Esquema del funcionamiento del Boosting (2020). Fuente: <https://www.analyticsvidhya.com/blog/2020/02/4-boosting-algorithms-machine-learning/>

En 2006 el Machine Learning consiguió empezar a expandirse globalmente para el manejo y procesamiento de datos gracias a IBM y Microsoft. Durante

este tiempo se desarrollaron métodos capaces de identificar el habla o el reconocimiento facial, métodos que hacía tan solo unas décadas habían frustrado a muchos investigadores con el Perceptron.

En la actualidad el Machine Learning es responsable de grandes avances y no solo en el ámbito informático o en el ámbito empresarial, en los cuales es imprescindible por su función de manejar gran cantidad de información entre otras cosas, también en ámbitos como el de la conducción de coches autónomos, nuevos algoritmos para robots, herramientas de análisis o incluso en medicina ayudando en los diagnósticos.

III. Modelos de Machine Learning

Aunque el Machine Learning no para de evolucionar con el tiempo y van surgiendo nuevos modelos y métodos, hoy por hoy podemos hablar de tres grandes modelos de aprendizaje en el Machine Learning, representados en la Figura 3. “El aprendizaje supervisado implica retroalimentación para indicar cuándo una predicción es correcta o incorrecta, mientras que el aprendizaje no supervisado no implica ninguna respuesta: el algoritmo simplemente intenta categorizar los datos en función de su estructura oculta. El aprendizaje por refuerzo es similar al aprendizaje supervisado en el sentido de que recibe retroalimentación, pero no es necesariamente para cada entrada o estado” (Jones, 2017).

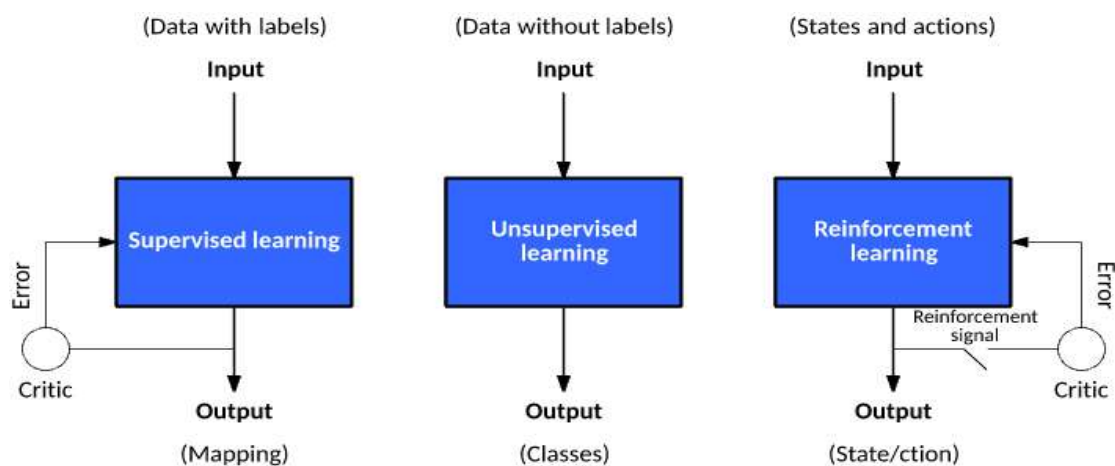


Figura III: Tres modelos de aprendizaje de algoritmos, por M. Tim Jones (2017). Fuente: <https://developer.ibm.com/articles/cc-models-machine-learning/>

- Aprendizaje Supervisado

El aprendizaje supervisado necesita empezar con un conjunto establecido de datos que deben poder identificarse bajo una cierta etiqueta. Este conjunto de datos se introduce con el objetivo de que pueda reconocer patrones y que estos puedan aplicarse a procesos de análisis. La representación del proceso del aprendizaje supervisado se muestra en la Figura 4.

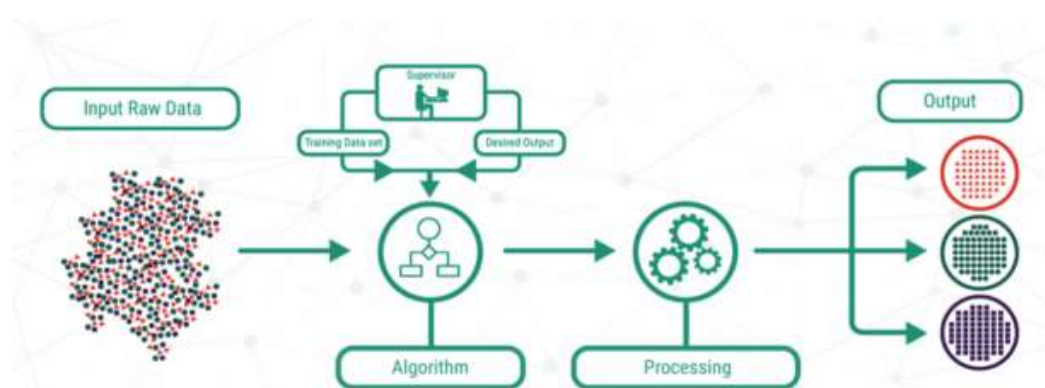


Figura IV: Aprendizaje Supervisado. Fuente: <https://www.ironhack.com/es/data-analytics/que-es-machine-learning>

Un ejemplo de cómo funcionaría este modelo sería el siguiente: si al programa le introducimos una gran cantidad de imágenes de perros y una gran cantidad de imágenes de gatos, cada una etiquetada correctamente, después de haber entrenado al modelo con estos datos, si ahora introducimos una imagen de un perro sin haberlo etiquetado, el modelo podrá predecir correctamente que es un perro.

Cuando la variable a predecir es una entre distintas categorías, conocidas como clases, se llamaría a este problema clasificación (estas clases pueden ser cuantitativas con valores discretos o cualitativas).

Por otro lado, dentro del aprendizaje supervisado cuando la etiqueta es continua tenemos los problemas de regresión. La regresión utilizada para el aprendizaje supervisado ayuda a comprender la correlación entre las variables. Un ejemplo sería predecir el valor de una casa conociendo el número de cuartos, tamaño, antigüedad, etc.

La formación supervisada de los modelos tiene una amplia aplicabilidad a una gran variedad de problemas empresariales, como la detección de fraudes, las soluciones de recomendación, el reconocimiento de voz o el análisis de riesgos.

- Aprendizaje No Supervisado

El aprendizaje no supervisado, al contrario que el supervisado, se utiliza cuando los datos introducidos no van a ser etiquetados, por lo que el algoritmo debe encontrar puntos comunes entre sus datos de entrada o descubrir los patrones intrínsecos que subyacen a los datos necesitando solo los datos que se le introduzcan, como se ilustra en la Figura 5.

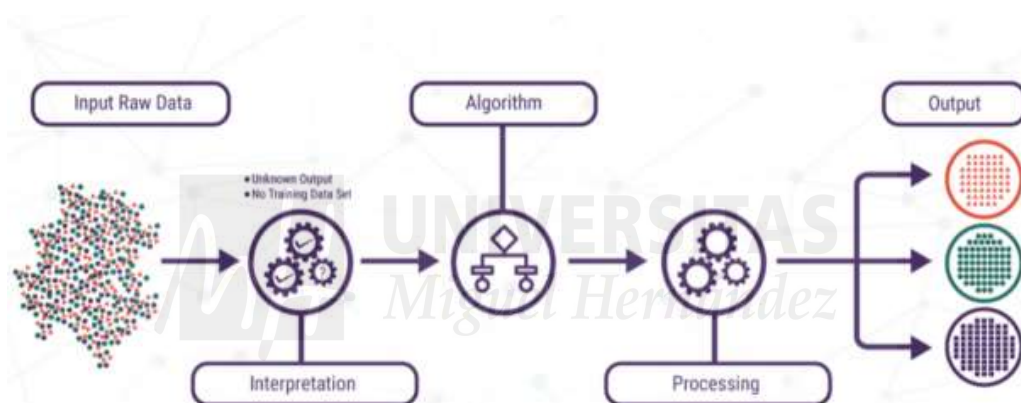


Figura V: Aprendizaje no supervisado. Fuente: <https://www.ironhack.com/es/data-analytics/que-es-machine-learning>

La subdivisión que se hace en el aprendizaje no supervisado es: el análisis de conglomerados, estimación de la densidad y reducción de la dimensión. En el análisis de conglomerados los datos son agrupados según sus similitudes o diferencias entre ellos, en la estimación de densidad los patrones y datos son representados por distribuciones o formas definidas, y en la reducción de la dimensión, datos duplicados o innecesarios son eliminados de la muestra de datos para producir una más pequeña.

Un ejemplo del aprendizaje no supervisado, podría ser determinar los segmentos de clientes en datos de marketing. Ser capaz de determinar diferentes segmentos de clientes ayuda a los equipos de marketing a abordar estos segmentos de clientes de manera única.

- Aprendizaje Semisupervisado

El aprendizaje semisupervisado es un término medio entre el supervisado y el no supervisado que se utiliza cuando no se tiene suficientes datos etiquetados para que llegue a ser supervisado. Se utiliza este pequeño conjunto de datos etiquetados para ayudar y guiar en procesos con datos mucho más grandes que no tiene ninguna etiqueta (IBM, 2020).

- Aprendizaje por refuerzo

En el aprendizaje por refuerzo el sistema aprende de la experiencia, se podría decir que es gran parte parecido a la forma de actuar de los humanos. En cierta manera es parecido al aprendizaje supervisado, pero no solo se entrena con el conjunto de los datos si no que necesita ir aprendiendo a base de aciertos y errores para determinar cuál es el resultado óptimo para un problema dado (Cheng, 2020). La Figura 6 ilustra el aprendizaje por refuerzo.



Figura VI: Aprendizaje por refuerzo, por Chieh-Feng Cheng (2020). Fuente: https://www.inwinstack.com/blog-en/blog_ai-en/6262/

Un ejemplo del aprendizaje por refuerzo estaría en el mundo de las finanzas; los modelos supervisados pueden determinar precios de acciones en el futuro pero estos modelos no determinan la acción que hay que tomar, en el aprendizaje por refuerzo se entrena para que decida si tiene que mantener, comprar o vender. IBM, por ejemplo, tiene una sofisticada plataforma basada en el aprendizaje por refuerzo que tiene la capacidad de realizar operaciones financieras. Calcula la función de recompensa basándose en la pérdida o el beneficio de cada operación financiera.

- Redes Neuronales Artificiales

Las redes neuronales artificiales son un modelo que busca replicar la forma de trabajar del cerebro humano. La red neuronal consta de múltiples capas de nodos interconectados, donde podemos diferenciar: una capa de entrada, una o varias capas ocultas y una capa de salida. La capa de entrada recibe la información desde el exterior. A continuación, la información se transmite de un nodo a otro en la capa oculta. Cada nodo posee a su vez un peso, es decir, un valor numérico, con el que modifica la entrada recibida y al llegar a la capa de salida nos devuelve un resultado según el peso aplicado por los nodos. Una red neuronal típica puede estar formada por miles o incluso millones de nodos de procesamiento simples que están densamente interconectados. Estos sistemas aprenden y se forman a sí mismos, en lugar de ser programados de forma explícita.

El proceso descrito es el que utiliza el Deep Learning. Podemos diferenciar el Deep Learning de un modelo de redes neuronales artificiales normales porque este último puede realizar predicciones aproximadas con una sola capa oculta mientras que el Deep Learning utiliza tres o más capas ocultas que le ayudan a optimizar y mejorar sus resultados como se muestra en la Figura 7.

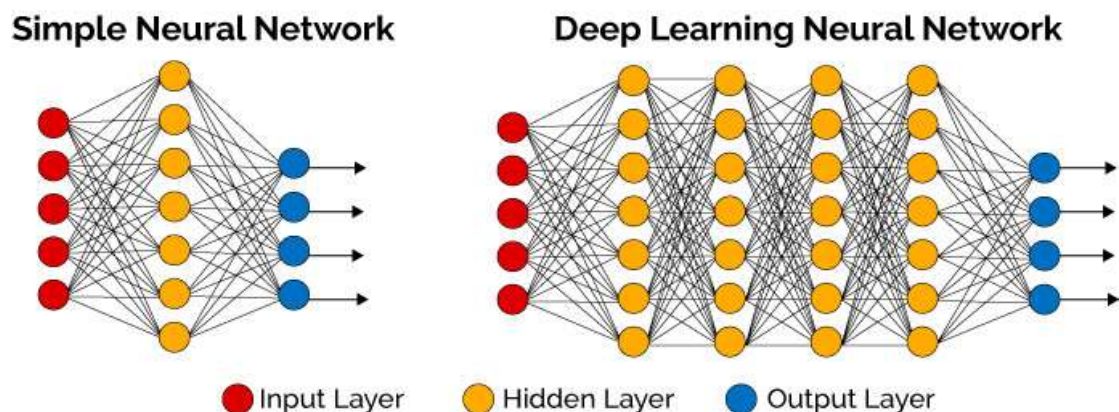


Figura VII: Redes neuronales artificiales y Deep Learning, por Dr Stylianos Kampakis. Fuente: <https://thedata scientist.com/what-deep-learning-is-and-isnt/>

Algunos usos de este método son: reconocimiento de voz, localización e identificación de caras y emociones faciales, detección de fraudes, recomendaciones a clientes, gestión de relaciones con los clientes, entre otros.

IV. Algoritmos de Machine Learning

Los algoritmos son una serie de instrucciones que se le incluyen en un programa para decirle cómo actuar con la información que recibe. Los algoritmos abarcan un amplio abanico de posibilidades, y su finalidad puede consistir, por ejemplo, en identificar qué número va detrás de otro en una serie o identificar la cara de alguien en un video.

Existen muchos tipos de algoritmos, y elegir cuál de ellos es el que mejor resultados nos daría depende de varios factores: el tipo de problema, capacidad de computación, cantidad de información, naturaleza de la información, etc. Dos científicos de datos pueden enfrentarse a un mismo problema empresarial y elegir distintos tipos de algoritmos según el tipo de visión que tengan a la hora de resolver el problema (Mueller & Massaron, 2021).

Es importante entender cómo funcionan estos algoritmos para poder conseguir la mayor precisión en el resultado con la información disponible. Entre todos los algoritmos que pueden ser utilizados en el Machine Learning estos serían los más usados y relevantes:

1. Regresión lineal: normalmente este algoritmo es usado para estimar valores reales basado en variables continuas. Un ejemplo sería el valor de una casa o el total de ventas de una empresa. La regresión lineal se basa en encontrar la relación lineal entre dos o más variables ajustando una ecuación lineal a los datos. En el caso de la regresión lineal simple esta ecuación es $y=a+bx$.
2. Regresión logística: de forma contraria a lo que sucede en la regresión lineal que se utiliza para predecir valores continuos, la regresión logística estima valores discretos. Predice la probabilidad de que ocurra un evento ajustando la información a una función logística. La regresión logística es una técnica de aprendizaje automático relativamente sencilla y gracias a sus resultados interpretables hace que sea muy usada en la industria del Machine Learning.

3. **Árbol de decisión:** el árbol de decisión es uno de los algoritmos más populares en Machine Learning hoy en día, es un algoritmo de aprendizaje supervisado que se usa en problemas de clasificación, aunque también puede ser usado en problemas de regresión. Este algoritmo funciona tanto para variables dependientes categóricas como para continuas. Los árboles de decisión utilizan una estructura de ramificación para ilustrar los resultados de una decisión, los árboles están compuestos por nodos a los que se les asignan porcentajes en función de la probabilidad de que se produzca un resultado. Se suelen utilizar en campañas de marketing para predecir resultados pudiendo dividir los clientes en segmentos (Ray, 2017).
4. **Naïve-Bayes:** es una técnica de clasificación basada en el teorema de Bayes, donde se asume que la presencia de una característica particular en una clase no está relacionada con la presencia de otra característica, aunque esta característica sí esté relacionada con las otras. El algoritmo considera todas las características independientes para calcular la probabilidad. Este es un algoritmo fácil de construir y que trabaja muy bien para conjuntos muy grandes de información.
5. **K-Means:** es un algoritmo de aprendizaje no supervisado que resuelve los problemas de agrupamiento. Los datos son clasificados en un número determinado de grupos de forma que todos los puntos de datos de un mismo grupo sean homogéneos y con los de otros grupos sean heterogéneos.
6. **Reducción de dimensionalidad:** ayuda a los sistemas a eliminar datos que no son útiles para el análisis. La cantidad de información que se genera hoy en día es inmensa, por eso algoritmos que eliminen datos redundantes, valores atípicos y datos no útiles son necesarios para poder realizar predicciones más precisas (Brownlee, 2019).

V. Big Data y Machine Learning

Desde hace unos años la palabra Big Data ha ido ganando popularidad sobre todo en el mundo de los negocios desde el auge de los smartphones, las redes sociales y el e-commerce, que aportan información muy valiosa a las organizaciones sobre los clientes actuales y los potenciales clientes.

Big Data, como ya hemos explicado anteriormente, se refiere a grandes volúmenes de información y se puede definir mediante las 5 "V". Una característica del Big Data, es que incorpora algunos tipos de información que no se incluían con anterioridad en los análisis de clientes. Podemos diferenciar la información que se incluye en el Big Data en:

- **Texto:** es uno de los elementos más fáciles de procesar, por ejemplo, todo lo que la gente escribe puede ser analizado e identificado, nos indica productos que están recibiendo buenas o malas críticas o las tendencias del momento, esto incluso lo hace de manera pública Twitter con un apartado sobre qué es lo más hablado en el momento, diferenciando por zonas geográficas.
- **Imágenes:** incluye tanto fotos como videos, YouTube y Facebook o Instagram han tenido un gran impacto popularizando ambos formatos. El Machine Learning puede analizar tanto fotos como videos y encontrar patrones en los mismos. Hoy en día, buscadores como el de Google incluyen la opción de buscar por imagen en vez de añadir un término escrito. YouTube, por ejemplo, tiene programas entrenados para detectar el Copyright en los videos.
- **Información de redes sociales:** incluye la información extraída de tus conexiones con otra gente, cosas que sigues o tus gustos. Esta información es muy útil para las empresas a la hora de generar anuncios personalizados.
- **Información geográfica:** uno de las informaciones que más relevancia ha ganado gracias a los smartphones es la información geográfica, dónde

se encuentra la gente y sus movimientos.

- Información generada por los productos: este tipo de información es relativamente nuevo. La mayoría de nuestros productos hoy en día, televisores, frigoríficos, sistemas de sonidos, etc, están diseñados para compartir información entre ellos y por Internet. Este concepto se denomina “Internet de las cosas” y hace referencia a la interconexión digital de objetos cotidianos.

Los ordenadores modernos especializados en Big Data han evolucionado para poder almacenar capacidades masivas de información, muchas veces usan sistemas híbridos que almacenan parte de esta información en la nube, procesan información a gran velocidad y limpian la información, estandarizando esta para que pueda ser procesada de una manera más eficiente.

Uno de los problemas que presenta la inmensa capacidad de creación de información en la actualidad es que las empresas no tienen forma de controlar toda la información que se crea ni el formato en el que se emite. Esta información que se genera muchas veces no está completa o es correcta, un ejemplo de esto sería el review bombing o bombardeo de reseñas en el que un grupo de personas coordinadas deja de golpe opiniones negativas sobre un producto, por circunstancias ajenas al producto, mostrando una realidad distorsionada, como en el estreno de la película Capitana Marvel que fue bombardeada con críticas negativas por opiniones de la actriz principal a favor del feminismo.

Toda esta información que se genera puede ser analizada para extraer los mejores insights gracias al Machine Learning, y cuanto mayor sea la cantidad de información mejor será la precisión del resultado. Es importante diferenciar que el Big Data necesita técnicas de Machine Learning para producir unos buenos resultados pero, una empresa no necesita Big Data para poder utilizar Machine Learning, aunque pueda ayudar a la hora de obtener los resultados.

VI. Beneficios del Machine Learning en los negocios

Cuando se habla de la introducción y uso del Machine Learning en el ámbito empresarial es normal que surjan dudas como qué beneficios puede aportar el Machine Learning sobre expertos ya establecidos en el campo que se quiera usar el Machine Learning. La respuesta a esta cuestión sobre qué valor añade el Machine Learning podría centrarse sobre 4 puntos:

- **Rapidez:** uno de los beneficios más significativos es la rapidez. El Machine Learning puede realizar tareas repetitivas y que requieren de gran cantidad de recursos y tiempo si son desarrolladas por personas, aumentando la eficacia de las mismas.
- **Precisión:** los modelos predictivos tienen una mayor precisión a la hora de identificar casos o tener menos error que su contraparte humana, rondando entre un 20-30% más de precisión, minimizando casi por completo la posibilidad de error.
- **Objetividad:** los modelos predictivos no tienen prejuicios por lo que el resultado será completamente imparcial.
- **Económico:** normalmente, una vez desarrollados, utilizar estos modelos predictivos es más barato que contratar personas para realizar el mismo trabajo. Además, el Machine Learning ofrece la oportunidad de no tener que desarrollar tú el programa ahorrándote el coste inicial, simplemente contratando servicios como Amazon Machine Learning o Azure Machine Learning de Microsoft para usos específicos.

Estos beneficios, entre otros, nos dejan claro la importancia que tiene y tendrá el Machine Learning en las empresas que lo adopten. Hoy en día no hay un sistema que pueda competir contra este por lo que, si nada cambia, en un futuro la mayoría de empresas acabarían adoptando el Machine Learning en su día a día.

Aún con el Machine Learning aportando todas estas ventajas y mejorando a

pasos agigantados los sistemas y su eficacia, es importante no dejar de lado el trabajo humano a la hora de analizar situaciones o fomentar la creatividad. El Machine Learning, al basar su aprendizaje y comportamiento en datos dados, sus repuestas pueden no adaptarse a las necesidades reales en caso de un cambio sustancial que altere la realidad de las personas como, por ejemplo, una crisis o una pandemia.

VII. Machine Learning y el mercado financiero

El Machine Learning ha revolucionado muchas áreas y ámbitos diversos que nos afectan día a día, por ejemplo, a quién no le salen anuncios personalizados cada vez que se encuentra navegando, todo esto gracias al Machine Learning que analiza nuestra actividad y la compara con millones de usuarios. En algunos ámbitos el Machine Learning supondrá una revolución en el futuro, por ejemplo, todos estamos esperando a que se generalice que los coches nos lleven a los sitios sin tener que conducir nosotros.

Sin duda, una de las áreas que ha revolucionado el Machine Learning ha sido el mundo de las finanzas, donde el impacto del Machine Learning ha supuesto una revolución en la forma de hacer las cosas, tomar decisiones en base a grandes cantidades de datos, tantos que un humano no podría procesarlos, y a una gran velocidad, que supone para las empresas la posibilidad de generar grandes beneficios. Los usos y aplicaciones son muy diversos pero si los agrupáramos los podríamos dividir en aplicaciones centradas en el cliente y aplicaciones centradas en las operaciones.

VII.1 Aplicaciones centradas en los clientes

- **Credit Scoring**

El Credit Scoring, o evaluación de crédito en español, se utiliza para determinar la aprobación o negación de un préstamo en base a unos parámetros. El uso del Machine Learning para la evaluación de crédito es el más común en el ámbito de las finanzas (Institute of International Finance, 2019).

La capacidad de evaluar el potencial riesgo de un cliente a la hora de devolver el préstamo es necesario para las instituciones encargadas de prestar dinero. El Machine Learning utiliza modelos de regresión, arboles de decisión y análisis estadístico para evaluar el crédito utilizando datos estructurados como, por ejemplo, transacciones o historial de deudas. Además, cada vez más bancos añaden datos adicionales, no estructurados, para mejorar la precisión en la evaluación de créditos como, por ejemplo, la actividad en redes sociales o el uso del teléfono móvil (Financial Stability Board [FSB], 2017).

El Machine Learning tiene la capacidad de procesar todos estos datos de una manera más rápida y barata, proporcionando una evaluación de crédito mucho más precisa que la que pudiera realizar una persona. Esto no solo beneficia a las entidades encargadas de conceder préstamos, también a la personas o empresas que no dispongan de un historial de transacciones suficiente y que con métodos antiguos sería lo único a evaluar, en cambio una evaluación más amplia y general proporciona una visión más clara que solo un historial de transacciones.

Aunque estos modelos son completamente objetivos, ya que simplemente analizan y procesan datos, algunos expertos apuntan a la posibilidad de la introducción de sesgos debido a la falta de transparencia por parte de las organizaciones. Otro desafío al que se enfrentan estos modelos es la calidad de la información, ya que se procesa muchísima información y no toda es útil para la evaluación de crédito.

- **Chatbots**

Los Chatbots es una tecnología que cada vez implementan más las instituciones financieras, sobre todo en aplicaciones móviles. Son programas que pueden mantener una conversación con el usuario, se utilizan para llevar a cabo tareas y funciones de atención al cliente, ayudan y aconsejan al cliente. Presentan una serie de ventajas como la inmediatez, ya que no tienes que esperar a que te atiendan, ahorro de costes en formación de personal, accesibilidad y eficiencia.

Estos programas utilizan el procesamiento del lenguaje natural (NLP), una rama del Machine Learning para interactuar con los clientes por texto o voz. Uno de los referentes en el uso de estos programas es el mercado financiero asiático donde, respecto a los occidentales, los Chatbots son mucho más sofisticados.

- **Servicios de seguros**

El sector de los seguros trabaja con grandes cantidades de información, es un caso parecido al de la evaluación de crédito, pero en los seguros existen muchos factores distintos que le añaden complejidad ya que, por ejemplo, no es lo mismo un seguro de vida que un seguro del hogar. El sector de los seguros utiliza el Machine Learning para analizar la información a un menor coste mientras que aumenta su beneficio.

El sector de los seguros se basa en establecer tarifas basadas en los pagos esperados para que la mayoría de veces terminen con ingresos positivos. El Machine Learning ayuda a establecer esas tarifas y comprender el pago posible para mantener la rentabilidad.

Centrándonos en los seguros de salud y vida, hay muchos factores que aumentan o disminuyen el riesgo de una persona para contraer enfermedades, condiciones graves o tener una vida larga. Mientras que, por ejemplo, para un seguro de coche se tienen que tener en cuenta factores como las opciones de reparación, disponibilidad de piezas, etc. El Machine Learning es capaz de procesar los distintos tipos de información y agruparla para conseguir el resultado adecuado según la situación que presente el cliente.

VII.2 Centradas en las operaciones

- **Optimización del capital**

Todos los bancos y empresas buscan maximizar su capital, es decir. sacar el máximo rendimiento en sus operaciones, lo cual recae en gran medida en el uso de funciones matemáticas. Los bancos y empresas utilizan programas para

optimizar el riesgo-peso de los activos y también en el margen de ajuste de valoración.

El Machine Learning se utiliza para encontrar la mejor combinación entre:

- Ejecución de pares de operaciones de derivados que se compensan.
- Ejecución de estrategias de compensación con el mismo distribuidor.
- Novación de operaciones de la cartera de un distribuidor a otra.

- **Modelo de gestión de riesgo**

Una de las áreas donde sin duda el Machine Learning destaca en el mercado financiero es en la gestión de riesgo. Los modelos de gestión de riesgo son claves para los bancos, estos modelos utilizan una gran variedad de opciones financieras para realizar pruebas considerando cómo actuará el mercado y serán las tendencias, con la esperanza de reducir el riesgo en los escenarios posibles.

Gracias a la potencia del Machine Learning los bancos pueden analizar grandes cantidades de datos, tanto estructurados como no estructurados, en poco tiempo. Los algoritmos analizan el historial de casos de riesgo e identifican si hay signos de posibles riesgos en el futuro. Además, ayudan a la hora analizar información en tiempo real para detectar problemas.

La Inteligencia Artificial y el Machine Learning también son usadas para los test de estrés realizados por los bancos, pruebas que determinan la capacidad de estabilidad de un banco.

- **Análisis de impacto de mercado**

El análisis de impacto de mercado consiste en evaluar el efecto que tiene sobre los precios de mercado las acciones que compra y vende una empresa. Las empresas buscan minimizar su impacto en grandes compras/ventas de acciones encontrando el momento adecuado y minimizando los costes de la ejecución. Las técnicas de Machine Learning complementan los modelos

tradicionales para el análisis de impacto de mercado, y las empresas utilizan estas técnicas para recabar más información de modelos creados anteriormente o ayudar a identificar relaciones no lineales.

El efecto que una empresa puede tener a la hora de la compra/venta de acciones en el mercado es difícil de analizar para ella misma. “En el caso de los fondos sistemáticos más activos, se calcula que hasta dos tercios de las ganancias de las operaciones se pierden por los costes de impacto del mercado” (Day, 2017).

El Machine Learning es utilizado en este área sobre todo para reforzar modelos ya en uso donde la información comparable de intercambios pasados es escasa. Se utiliza, por ejemplo, para identificar grupos de bonos que actúan de forma similar para así poder compararlos o para encontrar el momento adecuado para la compra/venta de acciones que minimice su impacto en el mercado comparándolo con otros. “Los bancos también están probando el aprendizaje por refuerzo para enseñar a las herramientas de Inteligencia Artificial a reaccionar ante el desequilibrio de las órdenes y la posición de cola en el libro de órdenes al límite” (FSB, 2017).

- **Trading y gestión de cartera**

La Inteligencia Artificial y el Machine Learning han cambiado tanto el mundo del trading y la gestión de carteras, que las empresas diseñan sus estrategias en base al Machine Learning, aunque el grado de libertad que le otorgan a este depende de la empresa. Las inversiones impulsadas por los datos procesados han ido aumentando y en 2018 se acercaron al trillón de dólares (Sharpe, 2018).

El trading genera una gran cantidad de datos necesarios para que el Machine Learning sea aún más preciso. Además, el Machine Learning tiene en cuenta datos no estructurados de noticias, redes sociales o incluso foros, con lo que se consigue predecir con mayor precisión. Un ejemplo de la importancia de los datos no estructurados es el caso del foro inglés Reddit en el que los usuarios se pusieron de acuerdo y consiguieron que las acciones de la empresa AMC

subieran en 72 horas un 467% apoyados por tweets de personalidades como Elon Musk.

Las grandes empresas de trading y bancos usan, entre otras cosas, técnicas para analizar el riesgo de las transacciones basándose en el análisis Big Data, lo que ha permitido que estas empresas minimicen el riesgo a la vez que optimizan el uso de su capital.

La velocidad a la que se analiza la información es también uno de los puntos fundamentales en esta área. En el trading los precios pueden variar significativamente en minutos por lo que analizar grandes cantidades de información en periodos muy cortos de tiempo, cosa que sería imposible para un humano, hace que el Machine Learning esté provocando que los mercados se estén adaptando a lo que llamamos trading de alta frecuencia. Algoritmos de trading automáticos están constantemente analizando el mercado y mandando acciones de compra o venta en frecuencias muy elevadas (Conerly, 2014). La Figura 8 muestra una comparación simplificada del cambio de precios en las acciones con trading de alta frecuencia y sin trading de alta frecuencia.



Figura VIII: Trading de alta frecuencia, por Bill Conerly (2014). Fuente: <https://www.forbes.com/sites/billconerly/2014/04/14/high-frequency-trading-explained-simply/?sh=ef7931c3da8d>

En la gestión de cartera la Inteligencia Artificial y el Machine Learning están siendo usados para descubrir señales que ayuden a predecir movimientos futuros en los precios y aumentar la eficacia con la que se procesa la información comparado con los antiguos modelos.

Este tipo de programas para gestionar la cartera es utilizado mayormente por fondos de cobertura, y aunque ayudan a crear las carteras que más beneficios pueden ofrecer en un futuro gracias a las previsiones, también se tiene en cuenta la opinión de expertos en el sector, buscando siempre entender por qué el Machine Learning ofrece esos resultados y no dejándolo actuar autónomamente.

Otras empresas también utilizan el Machine Learning para proporcionar a sus clientes estrategias de inversión totalmente personalizadas y automatizadas. También existen servicios como, por ejemplo, AlpacaForecast, una aplicación para predecir los precios en el mercado a corto plazo impulsada por la Inteligencia Artificial de Bloomberg disponible para todo tipo de clientes.

- **Detección de fraude**

El fraude es un gran problema para las empresas, informes en 2014 mostraban que los negocios perdían aproximadamente un 5% de sus ganancias al año debido al fraude, lo que supone 3,7 trillones de dólares en todo el mundo debido al fraude (FCASE, 2018).

La Inteligencia Artificial y el Machine Learning son capaces de ayudar en la detección de fraude. Estos métodos son especialmente efectivos contra el fraude de tarjetas de créditos que se ha popularizado gracias al aumento de las compras a través de Internet. Los sistemas contra el fraude analizan los comportamientos anteriores de los clientes, su localización y hábitos de compras, haciendo que a cualquier indicio sospechoso estos programas den la voz de alarma para que se bloqueen las transacciones hasta que sean investigados por las autoridades correspondientes.

Por otra parte, son muy útiles para la detección de otro tipo de fraude como es el del lavado de dinero, un área que se estudia con mucha detención y se espera mejorar en un futuro ya que no solo implica una pérdida de dinero si no que previene de actividades criminales como el tráfico humano, distribución de drogas o terrorismo. Entre 1 y 2 trillones de dólares en transacciones anuales son hechas con dinero lavado (Ayasdi, 2018).

Muchos bancos y autoridades ya están implementando estos sistemas de detección de lavado de dinero como, por ejemplo, la Autoridad Monetaria de Singapur, que utiliza el Machine Learning para el análisis de transacciones sospechosas, ya que analizarlas personalmente requiere de una gran cantidad de tiempo y en muchas ocasiones da lugar a falsos positivos. El Machine Learning identifica patrones en las transacciones, analiza el perfil de los clientes y distintos tipos de datos no estructurados, claves para llamar la atención y realizar una investigación más detenida en estos casos, haciendo un uso más eficaz del tiempo disponible.

Organismos reguladores como la SEC (Securities and Exchange Comissions) en Estados Unidos utilizan estos métodos para detectar los fraudes y conductas en contra de la normativa establecida. El personal de la SEC afirma que estos métodos son cinco veces mejores que realizar controles aleatorios aunque, en ocasiones, produzcan falsos positivos.

VIII. Posibles Efectos en los mercados financieros

La Inteligencia Artificial y el Machine Learning han supuesto grandes cambios en los mercados financieros. Se ha cambiado la forma de trabajar en distintas áreas, también ha abierto la puerta a más posibilidades, y todo esto ha afectado para bien o para mal al sector financiero desde un punto de vista micro económico y macro económico.

A medida que más empresas vayan mejorando sus sistemas de Machine Learning y procesen información a un ritmo parecido, se podrán reducir asimetrías en la información que son capaces de manejar y crear mercados más estables.

El Machine Learning, como hemos dicho con anterioridad, una de las cosas que busca es reducir costes de las transacciones, por lo que la propagación de estos métodos puede abrir los mercados a una mayor cantidad de gente. Una mayor cantidad de participantes en el mercado mejoraría la determinación de precios.

Fomentar la colaboración entre grandes empresas e instituciones a través del desarrollo de la Inteligencia Artificial y el Machine Learning es algo que ya estamos viendo como, por ejemplo, en Partnership on AI, organización dedicada a compartir mejores prácticas, investigaciones y charlas sobre la Inteligencia Artificial y que cuenta entre sus muchos socios con empresas tan diversas e importantes como McKinsey & Company, Apple, Google y Amazon.

Los clientes e inversores también pueden beneficiarse de la aplicación de la Inteligencia Artificial y el Machine Learning en los mercados financieros con una posible reducción de costes de servicios. Por ejemplo, si el riesgo de conceder un préstamo es menos gracias al Machine Learning y el Credit Scoring puede repercutir en una reducción de intereses. Además, los clientes podrán tener un servicio y recomendaciones más personalizadas en cuanto a servicios financieros, gracias al análisis de su información y comportamientos, adecuándose a sus necesidades.

Estos efectos positivos en los mercados financieros pueden trasladarse a otros mercados y ámbitos, incluso afectar a la economía de forma global. La reducción del riesgo gracias al Machine Learning en la prestación de crédito puede beneficiar a todos los ámbitos. Los beneficios que aporta el Machine Learning en el área financiera pueden hacer que también se adopte en otros sectores aumentando la inversión en Inteligencia Artificial y Machine Learning repercutiendo en un crecimiento económico en todos los ámbitos.

El Machine Learning no solo puede traer beneficios para los mercados financieros, sino que también puede exponer sus vulnerabilidades y permitir, por tanto, el abusar de ellas. Uno de los casos que preocupan a las autoridades es el uso del Machine Learning a la hora del trading y el trading de alta frecuencia, ya que podrían ser fuentes de nuevas vulnerabilidades. Un ejemplo del trading de alta frecuencia es el aumento de la volatilidad del mercado si se amplía el uso del Machine Learning generando grandes ventas y compras ejecutadas simultáneamente al determinar todos los modelos que es lo correcto.

Otro aspecto preocupante del uso del Machine Learning en los mercados

financieros es su uso para maximizar el capital y minimizar el riesgo, lo que puede traer consigo un estancamiento de los mercados. Todos los mercados financieros han tenido un crecimiento desde su creación, con sus altos y bajos, pero la media es positiva. El uso del Machine Learning puede hacer peligrar esto lo que conllevaría a un estancamiento.

IX. Ley y ética del Machine Learning

El Machine Learning, aun teniendo sus orígenes a mitad del siglo pasado, ha tenido un estallido en popularidad reciente y, por tanto, no existen leyes internacionales específicas que regulen su uso o aplicaciones. Esto sucede también en otras áreas ya que con la era de Internet y la información surgen nuevos sistemas, métodos y proyectos a un ritmo sin precedentes. Un ejemplo en cuanto a regularización son las criptomonedas como el Bitcoin; al surgir hace tan poco tiempo no existe ninguna regularización, pero ya se empiezan a ver leyes y controles sobre esto, como Hacienda exigiendo información a los contribuyentes sobre estas inversiones.

Con las leyes para regular el Machine Learning pasa exactamente lo mismo, aunque hasta hace poco no existían, ya se pueden ver esfuerzos para poder controlar los riesgos que este supone por parte de organismos internacionales de regularización y foros internacionales.

Como apunta el FSB (Financial Stability Board) “varios organismos internacionales de normalización han considerado los riesgos asociados a la negociación algorítmica, ya que se ha convertido en una característica omnipresente de los mercados que puede entre otras cosas, amplificar el riesgo sistémico”.

Entre los organismos que han advertido de los peligros de usar el Machine Learning para el trading se encuentran la Organización Internacional de Comisiones de Valores (IOSCO) o el Grupo de Supervisores Senior (SSG) para el control de las actividades de negociación de los bancos. Para estas organizaciones, entre otras, es importante evitar que el Machine Learning acabe usándose para aprovecharse y abusar del mercado y fomentar el

descontrol de este.

La Comisión Europea también ha tomado partido en la regularización de la Inteligencia Artificial y del Machine Learning, el 21 de abril de 2021 lanzaba una propuesta de reglamento para establecer una armonización de las normas sobre la Inteligencia Artificial con la siguiente justificación: “ Para fomentar el desarrollo de la Inteligencia Artificial y hacer frente por igual a los elevados riesgos que plantea para la seguridad y los derechos fundamentales, la Comisión presenta tanto una propuesta de marco normativo sobre la Inteligencia Artificial como un plan coordinado revisado sobre la misma” (Comisión Europea, 2021).

En esta propuesta presentada por la Comisión Europea se incluyen puntos que afectarían al uso del Machine Learning en los negocios y en el ámbito financiero, tales como:

- Prohibir el uso de los sistemas de Inteligencia Artificial para la creación de sistemas de crédito social, según su comportamiento social o rasgos de personalidad.
- Prohibir la utilización de la Inteligencia Artificial para la vigilancia indiscriminada.
- Notificar a la persona que esté interactuando con un sistema de Inteligencia Artificial.
- Una evaluación de los sistemas considerados de alto riesgo antes de ponerlos en marcha, que incluirían los usos para contratar personas, sentencias judiciales y evaluación de crédito.

Aunque esto es solo una propuesta por parte de la Comisión Europea nos da una pista de que en los siguientes años podemos esperar una regulación de los sistemas de Inteligencia Artificial y Machine Learning que nos hará tener que replantear algunos de sus usos actuales.

Otro de los temas en los que aún no hay consenso al haber tenido un desarrollo tan rápido, es la ética en el uso de estos sistemas. La ética, aunque afecta al ámbito de los negocios y financiero, va mucho más allá de eso, el uso de sistemas de Machine Learning nos proporciona un beneficio obvio, pero también afecta a personas pudiéndoles causar un gran perjuicio y que estos sistemas simplemente ven como datos.

Grandes nombres en el mundo de la ciencia y los negocios como Bill Gates y Elon Musk hablan abiertamente de los riesgos que supone para la humanidad un avance de la Inteligencia Artificial sin control. Elon Musk se pronunció en 2018 hablando sobre la Inteligencia Artificial: “Tiene que ser un organismo público el que tenga una visión y luego una supervisión para confirmar que todo el mundo está desarrollando la Inteligencia Artificial de forma segura. Esto es extremadamente importante. Creo que el peligro de la Inteligencia Artificial es mucho mayor que el de las cabezas nucleares, y nadie sugeriría que permitiéramos a cualquiera construir cabezas nucleares si lo desea. Eso sería una locura” (Musk, 2018).

Uno de los problemas éticos que presenta el Machine Learning es la utilización de sesgos. Por ejemplo, el Machine Learning puede procesar información sobre el género, edad, religión, si se está casado, si se tiene hijos, etc. y a partir de ahí definir, por ejemplo, qué candidato es el mejor para un puesto específico, pero en el caso de que usar esta información fuese legal, no está tan claro hasta qué punto es ético.

Es importante que las empresas se conciencien de que, aunque la opción que presenten estos sistemas es precisa, a veces no es ética. Que las empresas sean éticas es más importante ahora que nunca: construir una buena reputación, una imagen positiva, un compromiso con los clientes, que los trabajadores se sientan identificados con la empresa, etc.

Steven Finlay en su libro “Inteligencia Artificial y Machine Learning para los negocios” (Finlay, 2018), nos da tres aspectos importantes a tener en cuenta en la evaluación del riesgo ético en los sistemas de toma de decisiones:

- Beneficiario: cuanto más nos beneficiemos a costa de otra persona el riesgo ético que supone es más alto.
- Inmutabilidad de los datos: hay características con las que las personas nacen y no se puede cambiar como, por ejemplo, la edad, la raza o el lugar de nacimiento y tomar decisiones en base a estas características se percibirá de una manera poco ética de cara al público general.
- Impacto: según el peso que tenga la decisión tomada sobre la vida de la persona a la que afecta, es decir, decisiones que cambien la vida por completo tendrán un mayor impacto que otras que afecten solo a pequeños aspectos de las personas.

La ética en este campo está muy presente para las grandes empresas. En 2016 un grupo de grandes empresas, entre las que podemos encontrar a Amazon, Facebook, Google, IBM y Microsoft, fundaron la Asociación para la Inteligencia Artificial entre cuyos objetivos se encuentra identificar y fomentar la Inteligencia Artificial para fines socialmente beneficiosos y entre sus pilares el que la Inteligencia Artificial sea justa, transparente y responsable.

X. Futuro del Machine Learning

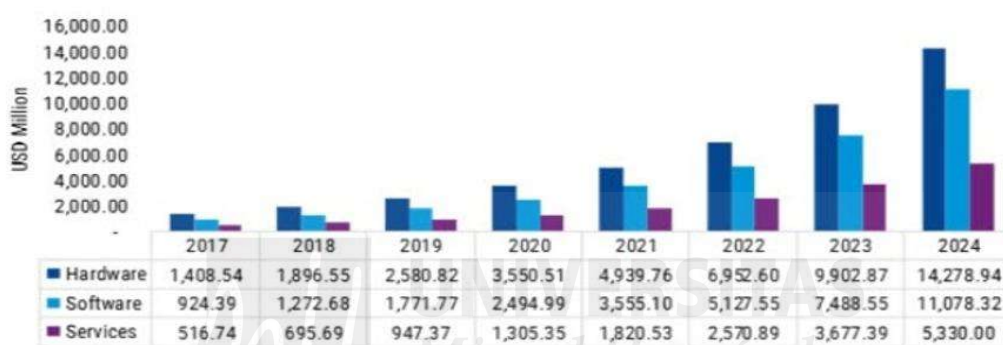
El Machine Learning es uno de los campos que más relevancia ha ganado en los últimos años y se espera que siga ganando aún mayor importancia según se vaya desarrollando y entrando cada vez más en distintos sectores. En estos momentos, el Machine Learning aporta un gran valor a las empresas que necesitan procesar grandes cantidades de información para entenderla y, con el ritmo al que se genera información, en un futuro no solo aportará valor si no que será necesario su uso.

Según Gartner, una empresa líder, consultora y de investigación de las tecnologías de la información, alrededor del 37% de todas las empresas analizadas están utilizando algún tipo de Machine Learning y prevé que alrededor del 80% de los avances modernos en 2022 estén fundamentado con Machine Learning o Inteligencia Artificial. También comenta que para 2023 los

recursos computacionales utilizados para la Inteligencia Artificial se multiplicarán por cinco desde 2018, lo que convertía la Inteligencia Artificial en la principal categoría de cargas de trabajo que impulsan las decisiones de infraestructura en las empresas (Gartner, 2020).

El mercado del Machine Learning, que fue evaluado en 2017 en 1,58 billones de dólares, ha crecido hasta alcanzar los 7,3 billones de dólares en 2020 y se espera que en 2024 alcance la cifra de 30,6 billones de dólares, como podemos ver en la evolución de la Figura 9.

Global Machine Learning Market, by Component, 2017-2024 (USD Million)



Source: MRFR Analysis

Figura IX: Informe de previsión del mercado del Machine Learning, por MRFR Analysis (2020). Fuente: <https://www.forbes.com/sites/louiscolumbus/2020/01/19/roundup-of-machine-learning-forecasts-and-market-estimates-2020/>

Uno de los aspectos que conlleva el brillante futuro del Machine Learning es sin duda la necesidad de la creación de puestos de trabajos relacionados con la Inteligencia Artificial y el Machine Learning. Solo en los últimos 4 años han aumentado las contrataciones en esta área en un 75% (Terra, 2021). Ingeniero en Machine Learning, Data Scientist, Research Scientist, R&D Engineer o Business Intelligence Developer son algunos de los puestos de trabajo que se buscarán en los próximos años para poder seguir desarrollando y utilizando la Inteligencia Artificial y el Machine Learning. Un informe de la empresa IDC predice que en el año 2023 habrá disponibles más de 20 millones de trabajos en el área de la Inteligencia Artificial (Eddy, 2020).

Por último, uno de los aspectos que cambiará el Machine Learning en un futuro será la transparencia. La transparencia de cara a los clientes y usuarios es un elemento clave que se ha propuesto la Unión Europea que los desarrolladores

de Machine Learning incorporen. El futuro será que la Inteligencia Artificial sea capaz de razonar sus decisiones, por ejemplo, actualmente a la hora de conceder un préstamo nos da un “sí” o “no” sin explicación, cosa que cambiará en el futuro cuando se añada el razonamiento de su decisión.

XI. Conclusión

El uso de la Inteligencia Artificial y el Machine Learning ha ido evolucionando desde mediados del siglo pasado hasta hoy en día, cuando sus modelos son usados en muchas áreas distintas, desde la medicina para diagnósticos prematuros, la conducción de vehículos autónomos o la detección de rostro muy utilizado en los teléfonos móviles.

La gran capacidad de procesamiento de información, tanto estructurada como no estructurada, su objetividad, precisión y su velocidad, sin duda han marcado un antes y un después en los mercados financieros. Destaca su uso sobre todo en casos como la evaluación de crédito, servicios de seguros, optimización de capital, modelos de gestión de riesgo, análisis de impacto de mercado, trading y gestión de cartera y detección de fraude.

Sus diferentes modelos como pueden ser el supervisado, no supervisado o de refuerzo, permiten que el Machine Learning se adapte a muchas situaciones distintas, otorgándole mucha versatilidad y permitiéndole adaptarse a situaciones muy diversas en las que no siempre tiene que depender de una persona que le introduzca datos etiquetados.

El Machine Learning ha traído al mercado financiero grandes beneficios, no solo para las empresas, sino también para los clientes. Por ejemplo, una mejor evaluación de crédito supondrá que las empresas o bancos dedicados a conceder préstamos reduzcan sus pérdidas lo que puede conllevar a menores intereses para los clientes. El aumento de detección de fraude, mejora en el trading y gestión de cartera o modelos de gestión de riesgo, entre otros, aportan beneficios y reducen costes.

Estabilidad es otro elemento importante que aporta la Inteligencia Artificial y el

Machine Learning a los mercados financieros. Estas herramientas accesibles hacen posible que se procese una mayor cantidad de datos en un espacio muy corto de tiempo ayudando a un mejor entendimiento de lo que está sucediendo en los mercados financieros. Aunque también hay que ser conscientes de los elementos negativos que pueden traer el uso indiscriminado de estos métodos a los mercados financieros.

La Inteligencia Artificial y el Machine Learning, aunque aportan muchos beneficios, su crecimiento es relativamente nuevo por lo que no están regulados como otras herramientas. Algunos usos del Machine Learning es posible que en un futuro, más a corto plazo que a largo, tengan que ser ajustados para poder entrar dentro de la legalidad que buscan imponer desde organizaciones como la Comisión Europea, en muchos casos para proteger a los consumidores evitando, por ejemplo, sistemas de crédito social.

En un futuro el Machine Learning deberá adaptarse a estas leyes, pero el futuro también se presenta muy esperanzador. Los beneficios que aporta vendrán con un aumento en la inversión en Inteligencia Artificial y Machine Learning y un aumento en la creación de puestos de trabajo para estas áreas, mientras sigue desarrollándose para mejorar su rendimiento y crear nuevas aplicaciones.

XII. Bibliografía

Alameda, T. (2020, 9 septiembre). «Machine learning»: ¿qué es y cómo funciona? BBVA NOTICIAS. <https://www.bbva.com/es/machine-learning-que-es-y-como-funciona/>

Ayasdi. (2018). Anti-Money Laundering Solution Deep Dive. https://s3.amazonaws.com/cdn.ayasdi.com/wp-content/uploads/2018/04/22170635/AML_Solutions_Deep_Dive_WP_051617v01.pdf

Brownlee, J. (2019, 12 agosto). A Tour of Machine Learning Algorithms. Machine Learning Mastery. <https://machinelearningmastery.com/a-tour-of-machine-learning-algorithms/>

Caparrini, F. S. (s. f.). Breve Historia de la Inteligencia Artificial. <http://www.cs.us.es/%7Efsancho/?e=221>

Cheng, C. (2020, 25 diciembre). Reinforcement Learning - inwinSTACK | Technical Blog. InwinSTACK. https://www.inwinstack.com/blog-en/blog_ai-en/6262

Comisión Europea. (2021). Communication on Fostering a European approach to Artificial Intelligence | Shaping Europe's digital future. <https://digital-strategy.ec.europa.eu/en/library/communication-fostering-european-approach-artificial-intelligence>

Conerly, B. (2014, 14 abril). High Frequency Trading Explained Simply. Forbes. <https://www.forbes.com/sites/billconerly/2014/04/14/high-frequency-trading-explained-simply/?sh=ef7931c3da8d>

Council of Europe. (2021). History of Artificial Intelligence. Artificial Intelligence. Recuperado de <https://www.coe.int/en/web/artificial-intelligence/history-of-ai>

Day, S. (2017, 16 agosto). Quants turn to machine learning to model market impact. Risk.Net. <https://www.risk.net/asset-management/4644191/quants-turn-to-machine-learning-to-model-market-impact>

Eddy, N. (2020, 16 noviembre). Artificial Intelligence (A.I.) Job Trends Important to Watch in 2021. Dice Insights. <https://insights.dice.com/2020/11/16/artificial-intelligence-a-i-job-trends-important-to-watch-in-2021/>

Encyclopedia Britannica. Artificial intelligence - Methods and goals in AI. (s. f.). <https://www.britannica.com/technology/artificial-intelligence/Methods-and-goals-in-AI>

FCASE. (2018, 18 octubre). The Actual Cost of Fraud. <https://fcase.io/the-actual-cost-of-fraud/>

Financial Stability Board. (2017). Artificial intelligence and machine learning in financial services Market developments and financial stability implications. <https://www.fsb.org/wp-content/uploads/P011117.pdf>

Finlay, S. (2018). Artificial Intelligence and Machine Learning for Business: A No-Nonsense Guide to Data Driven Technologies (3.a ed.)

Gartner. (2020). Data & Analytics Conference 2020 in Geneva, Switzerland. <https://www.gartner.com/en/conferences/emea/data-analytics-switzerland/featured-topics/topic-ai-machine-learning>

IBM (2020, 15 julio). Machine Learning. IBM. https://www.ibm.com/cloud/learn/machine-learning?mhsrc=ibmlearning_l&mhq=machine%20learning

Institute of International Finance. (2019). Machine Learning in credit risk (N.o 2). https://www.iif.com/Portals/0/Files/content/Research/iif_mlcr_2nd_8_15_19.pdf

Jones, M. T. (2017). Models for machine learning. IBM Developer. <https://developer.ibm.com/articles/cc-models-machine-learning/>

Lopez, J. L., & Ferraro, R. (2021, 16 septiembre). El papel de la Estadística en el Machine Learning. Máxima Formación. <https://www.maximaformacion.es/blog-dat/el-papel-de-la-estadistica-en-el->

machine-learning/

Kampakis, S. (2020, 26 noviembre). What deep learning is and isn't. The Data Scientist. <https://thedata scientist.com/what-deep-learning-is-and-isnt/>

Mitchell, T. M. (1997). Machine Learning (1.a ed.). McGraw-Hill Education

Mueller, J. P., & Massaron, L. (2021). Machine Learning for Dummies (2nd ed.)

Musk, E. (2018, 12 marzo). Elon Musk Answers Your Questions! | SXSW 2018 [Video]. YouTube. <https://www.youtube.com/watch?v=kzIUyrcbos>

Ray, S. (2017). Commonly used Machine Learning Algorithms (with Python and R Codes). Analytics Vidhya.

<https://www.analyticsvidhya.com/blog/2017/09/common-machine-learning-algorithms/>

Sharpe, S. (2018, 2 octubre). Quant Hedge Funds closing in on \$1 trillion in AUM. Inventure Recruitment.

<https://inventurerecruitment.com/news/2018/10/2/quant-hedge-funds-closing-in-on-1-trillion-in-aum>

Terra, J. (2021, 3 junio). The Rise of Artificial Intelligence and Machine Learning Job Trends in 2021. Simplilearn.Com. <https://www.simplilearn.com/rise-of-ai-and-machine-learning-job-trends-article>