El curso de Tuidi sobre Machine Learning

INTRODUCCIÓN A LA IA EN LA CADENA ALIMENTARIA

Este curso tiene como objetivo introducir el concepto de Inteligencia Artificial(IA) respondiendo a tres preguntas fundamentales:

- ¿Qué es la IA?
- ¿Qué tecnologías la sustentan?
- Cosa può fare l'IA per noi?

El propósito es proporcionar los conocimientos y herramientas necesarios para comprender qué se entiende por inteligencia artificial, cómo se procesa la información para desarrollar modelos predictivos mediante aprendizaje automático (machine learning) y qué aplicaciones tiene en los sectores minorista y de producción.
El curso consta de cuatro lecciones:

- **Lección 1** - La Inteligencia Artificial: ¿En qué consiste esta disciplina, cuáles son sus componentes, sus aplicaciones en el sector minorista y cuáles son las principales diferencias con los modelos estadísticos tradicionales utilizados hasta ahora?

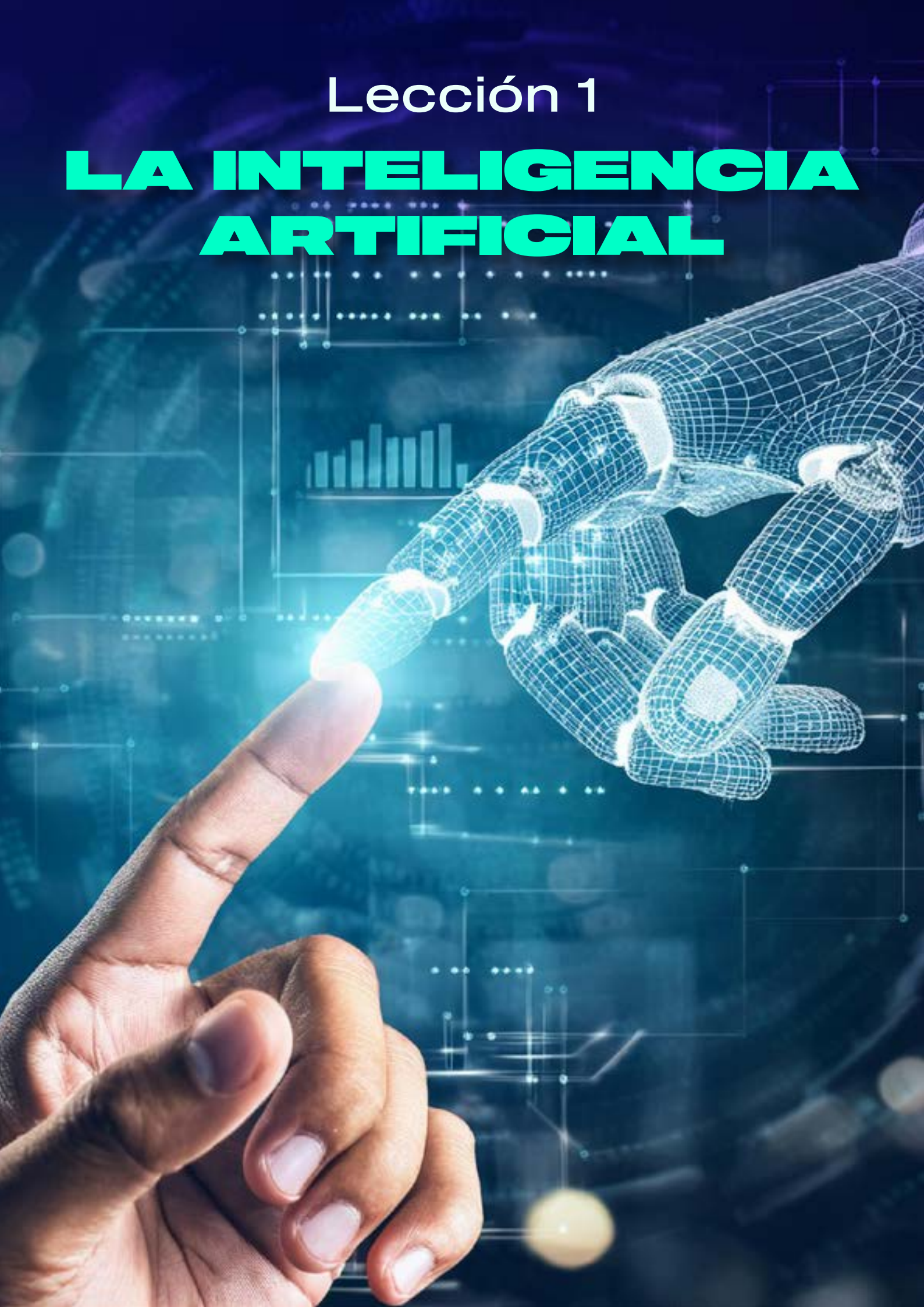
- **Lección 2** - Minería de datos (Data Mining): Se describirá el proceso de extracción de datos, destacando la importancia de contar con información precisa para alcanzar los objetivos establecidos.

- **Lección 3** - Aprendizaje automático (Machine Learning): Introducción al aprendizaje automático y a las técnicas de clasificación, regresión y agrupamiento.

- **Lección 4** - Diferencia entre análisis predictivo y prescriptivo: Se explicará la diferencia entre estos dos enfoques, detallando sus características, ventajas y desventajas.

Lección 1

LA INTELIGENCIA ARTIFICIAL



ÍNDICE

1.1 Introducción a la lección

1.2 ¿Qué es la Inteligencia Artificial?

1.3 Deep Learning (aprendizaje profundo)

1.4 ¿Puede una máquina con Inteligencia Artificial ser consciente?

1.5 El uso de la IA en la cadena alimentaria

1.6 ¿Por qué invertir en Inteligencia Artificial?

1.7 ¿Cómo reconocer una aplicación de Inteligencia Artificial?

1.8 Diferencias con los modelos estadísticos tradicionales

1.1 Introducción a la lección

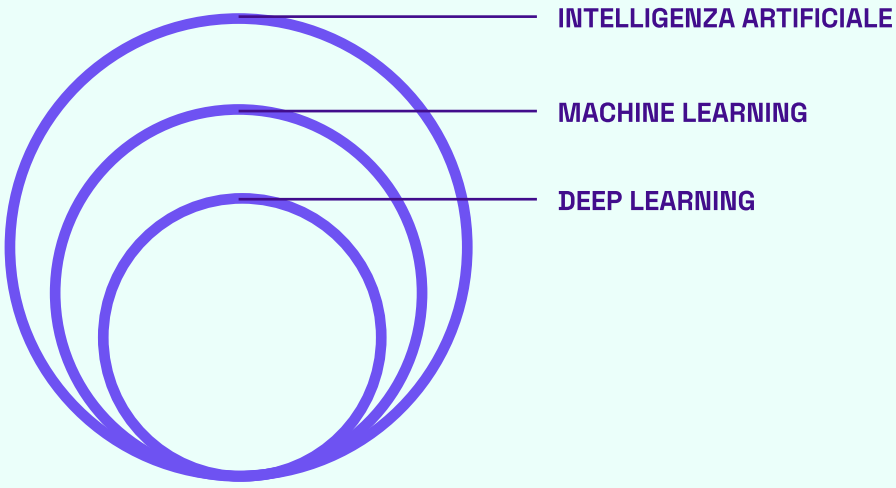
En esta primera lección, nos familiarizaremos con la inteligencia artificial, definiendo el concepto en un sentido amplio y profundizando en algunos aspectos concretos a través de ejemplos. La IA es un campo muy amplio y en constante evolución, caracterizado por una gran variedad de posibles aplicaciones. Para algunos, representa formas de vida artificial capaces de superar la inteligencia humana; otros, en cambio, la consideran simplemente un conjunto de tecnologías para el procesamiento de datos.

A lo largo de esta lección, explicaremos qué es la Inteligencia Artificial, cuáles son sus aplicaciones — con un enfoque especial en el sector alimentario — y qué ventajas ofrece en comparación con los sistemas estadísticos tradicionales.

1.2 ¿Qué es la Inteligencia Artificial?

La Inteligencia Artificial se define como la capacidad de una máquina para mostrar habilidades humanas como el aprendizaje, la planificación y la creatividad. Lo que hace que esta tecnología sea "inteligente" es precisamente su capacidad para aprender a partir de los datos que se le proporcionan manualmente y utilizar la información adquirida para alcanzar objetivos predefinidos.

A menudo, el concepto de IA se confunde con los términos Machine Learning (ML) y Deep Learning (DL), usándolos casi como sinónimos. Para aclarar la diferencia entre estos términos, resulta útil representarlos como círculos concéntricos.



En realidad, ML y DL no son sinónimos, sino dos subcategorías de la Inteligencia Artificial. El Machine Learning, o aprendizaje automático, es un subconjunto de la IA que se basa en el aprendizaje a partir de datos sin necesidad de reglas definidas manualmente. Todos los sistemas de Machine Learning forman parte de la Inteligencia Artificial, pero no todas las Inteligencias Artificiales se basan en el Machine Learning.

El proceso de toma de decisiones puede dividirse en tres fases:

1. La máquina analiza y aprende a partir de los datos que se le proporcionan;
2. La máquina toma decisiones de manera autónoma basándose en toda la información que ha logrado procesar;
3. Por último, la máquina optimiza las decisiones anteriores, aprendiendo automáticamente de sus errores.

El Machine Learning necesita grandes volúmenes de datos para funcionar de manera eficaz y adaptarse a distintas situaciones. Un ejemplo de su aplicación en el comercio electrónico es el Dynamic Pricing o fijación dinámica de precios, una estrategia que ajusta los precios de un producto en función de las condiciones actuales del mercado.

El Machine Learning es capaz de analizar datos en tiempo real y sugerir distintos precios según el día de la semana o la franja horaria, con el objetivo de maximizar los beneficios. Por ejemplo, el sistema podría predecir que aumentar el precio de las botellas de vino los viernes por la noche y reducirlo los martes al mediodía optimiza las ganancias (en este caso, nuestra función objetivo) derivadas de su venta. ¿Cómo lo consigue? El Machine Learning analiza datos de los clientes, como sus características demográficas, patrones de compra y frecuencia de acceso al sitio web, para calcular un precio que tenga en cuenta todas las variables que influyen en la maximización de la función objetivo.

1.3 Deep Learning (aprendizaje profundo)

El Deep Learning (DL) se refiere al aprendizaje autónomo. A diferencia del Machine Learning, que se basa en variables introducidas por el ser humano, los algoritmos de Deep Learning pueden identificar por sí mismos características distintivas sin necesidad de categorizaciones externas, es decir, sin intervención humana. Sin embargo, el DL requiere volúmenes de datos aún mayores que el ML y un mayor poder de procesamiento.

Un ejemplo de Deep Learning es el Procesamiento del Lenguaje Natural (NLP, por sus siglas en inglés), un conjunto de algoritmos capaces de analizar, interpretar y comprender textos escritos, considerándolos como secuencias de palabras que transmiten uno o más significados en un idioma. El NLP combina técnicas lingüísticas y computacionales para dotar a las máquinas de habilidades similares a las humanas, como la capacidad de leer y comprender el significado de las palabras.

Esta técnica puede utilizarse en la distribución alimentaria para cuantificar el efecto de las canibalizaciones, ya que el NLP puede asignar automáticamente valores matemáticos a cada producto al analizar la información demográfica y, posteriormente, establecer cuán contextualmente similares o disímiles son los productos. Un producto con la descripción “atún al aceite de oliva” será contextualmente similar a uno con la descripción “atún al natural”, menos similar a uno con la descripción “salsa de atún” y completamente disímil a uno con la descripción “ crema de chocolate “. Al identificar científicamente estas relaciones, es posible establecer el impacto positivo, negativo o neutral que una promoción sobre una referencia puede tener sobre las demás.

1.4 ¿Puede una máquina con Inteligencia Artificial ser consciente?

Para responder a esta pregunta, primero es útil definir el concepto de conciencia, que puede entenderse como “la capacidad del sujeto para ser consciente de sí mismo y del mundo exterior con el que interactúa, de su propia identidad y de sus actividades internas”. Por lo tanto, cuando hablamos de una máquina consciente, no nos referimos a una máquina simplemente inteligente, sino a una que también es consciente de sí misma. Es fundamental distinguir entre conocimiento y conciencia.

La pregunta es, entonces, qué hace que el cerebro humano, compuesto por una red densa de neuronas, sea consciente de que la luz de su automóvil está encendida, y qué hace que el automóvil, un sofisticado sistema electrónico e ingenieril, sea “ignorante”.

Este es un tema ciertamente complejo y muy debatido, porque, si tomamos como ejemplo a Deep Blue, la máquina que derrotó al campeón mundial de ajedrez Kasparov, algunos dirían que, aunque tenga conocimiento de las reglas del ajedrez y sepa cuándo hacer la jugada correcta en el momento adecuado, no es consciente. De hecho, cada jugada de Deep Blue es simplemente la aplicación de reglas y algoritmos proporcionados por los programadores.

La máquina no es consciente porque no sabe de manera autónoma qué es correcto o incorrecto. Se puede ver la conciencia desde una lógica ética. La idea de “correcto” o “incorrecto” que pueda tener proviene de cómo fue programada, y no es un pensamiento desarrollado de manera autónoma.

1.5 El uso de la IA en la cadena alimentaria

El mercado de la Inteligencia Artificial está creciendo cada año a un ritmo cercano al 20% y se está expandiendo a nuevos sectores. Todos conocemos los coches autónomos o los asistentes virtuales como Siri de Apple o Alexa de Amazon, pero existen muchos otros ejemplos menos conocidos.
Uno de ellos es el mundo del retail y la producción. La IA representa una gran oportunidad para las empresas de este sector debido a las múltiples aplicaciones que ofrece.
Una de las aplicaciones fundamentales es la optimización de la gestión de la cadena de suministro, que permite:

- Analizar los datos de clientes y proveedores;
- Optimizar la gestión del almacén;
- Predecir la demanda.

En relación con este último punto, esta tecnología es capaz de considerar una gran cantidad de factores, incluidas variables externas como las condiciones meteorológicas, las festividades y los cambios en las ofertas de los competidores, que pueden tener un impacto significativo en las variaciones de la demanda. Puede parecer intuitivo que las previsiones meteorológicas influyan en las ventas de ciertos productos, pero lo que resulta menos obvio es cuantificar analíticamente el impacto de estas variables externas en todo el surtido de un minorista o en la programación de los pedidos para un productor. Para ello, los algoritmos de Machine Learning pueden identificar tales relaciones. Por ejemplo, es posible prever cuánta cerveza se venderá si la temperatura es de 30 grados y cuánta si la temperatura es de 10 grados. Además, manteniendo el mismo ejemplo, también se podría saber cuál sería el impacto en las ventas de una cerveza específica si un competidor lanza una promoción sobre la misma marca o sobre una diferente.
Minoristas, mayoristas y productores tienen a su disposición grandes volúmenes de datos sobre transacciones e interacciones con los clientes, obtenidos tanto a través de canales offline como online.

Los algoritmos permiten aprovechar toda esta información contenida en los datos, ayudando a las empresas a tomar decisiones rápidas y precisas. Esto representa un factor clave en un sector donde se gestionan y controlan millones de flujos de información sobre productos, tratando de hacer coincidir lo más posible la oferta y la demanda. Una pequeña empresa, en promedio, se enfrenta a más de 400 millones de transacciones diarias, provenientes de diversas fuentes y con diferentes tipos de interpretación, lo que hace imposible para el ser humano realizar un análisis científico adecuado de esos datos.

Según un informe de McKinsey, la Inteligencia Artificial, y más concretamente el Machine Learning, en las previsiones de ventas permite a las empresas reducir los errores entre un 30% y un 50% a lo largo de la cadena de suministro. De hecho, las previsiones se vuelven más precisas gracias a la combinación de cientos de variables que influyen en la venta de un solo producto. Está claro que la calidad del proceso de suministro tiene un impacto directo en los ingresos y, por lo tanto, en las ganancias.

Esta tecnología ofrece beneficios tanto para el retail como para las empresas de producción. En el ámbito del retail, gracias a una predicción precisa de la demanda, es posible:

• Aumentar los ingresos mediante una mejor disponibilidad de productos en los estantes, lo que reduce las rupturas de stock;

• Optimizar la gestión del almacén, lo que permite reducir los costos de manejo de las mercancías en los centros de distribución y en las tiendas;

• Optimizar las existencias de seguridad para productos frescos, que suelen ser muy difíciles de gestionar, reduciendo así el desperdicio alimentario y maximizando las ventas;

• Optimizar el uso del espacio en los almacenes y en los estantes, lo que facilita la introducción de nuevos productos en el surtido, eliminando los productos “bajo-vendientes” que antes se usaban como sustitutos;

• Transferir productos entre puntos de venta, resolviendo el problema de la compra en lotes y la venta por unidades;

• Conciliar los documentos, alineando las facturas recibidas con los pedidos realizados para identificar ineficiencias, como los productos no entregados.

Por su parte, para las empresas de producción, la implantación de un sistema de previsión de la demanda conlleva:

• Aumento de los ingresos: al reducir los pedidos no entregados y el exceso de stock, tanto a corto como a largo plazo;

• Reducir la aleatoriedad en las solicitudes de los clientes: analizando el comportamiento histórico de los productos para entender qué influye en las ventas y predecir picos repentinos de demanda;

• Disponibilidad de materias primas: sugiriendo la cantidad de ingredientes necesarios para satisfacer los pedidos;

• Programar con anticipación la preparación de la mercancía: teniendo en cuenta restricciones comerciales, logísticas y operativas;

• Optimización de la producción: maximizando la eficiencia y organizando mejor las líneas de producción y la mano de obra;

• Optimización de la gestión del almacén: gestionando las existencias para optimizar las entregas;

• Mejora de la rotación del almacén: anticipando los picos y reduciendo los costos de inmovilización de mercancías.

Un ejemplo de cómo la IA en el retail puede ser un factor determinante se observa en el caso de Walmart y su primer Intelligent Retail Lab (IRL), una "tienda inteligente" diseñada para recopilar información sobre lo que ocurre en su interior mediante múltiples sensores, cámaras y procesadores.

Dentro del IRL, la Inteligencia Artificial, combinada con el Machine Learning y el Deep Learning, ha introducido una nueva modalidad de gestión de inventarios, gracias a una trazabilidad de los productos que garantiza en tiempo real la presencia de todos los artículos en los estantes. La falta de existencias es uno de los principales problemas con los que se enfrentan los minoristas y, en consecuencia, genera una pérdida en términos de ventas no realizadas que puede ocasionar, además de una drástica reducción de la facturación anual, una disminución de clientes habituales. Para lograr este objetivo, el IRL de Walmart dispone de 1500 cámaras colocadas en el techo de los puntos de venta que, mediante reconocimiento de imágenes, aseguran que los productos estén siempre disponibles.

El Image Recognition, o reconocimiento de imágenes, es una técnica de Deep Learning que permite a la máquina reconocer formas específicas mediante el análisis de imágenes. Para que esto sea posible, el sistema se entrena con grandes conjuntos de imágenes que representan los productos presentes en el supermercado, como, por ejemplo, latas de atún. La máquina intentará clasificar el producto visualizado y, después de cometer numerosos errores, será capaz de reconocer con una precisión extremadamente alta qué es una lata de atún y qué no lo es. Así, Walmart y su IRL pueden verificar la presencia de productos en las estanterías.

1.6 ¿Por qué invertir en Inteligencia Artificial?

El enorme avance de la Inteligencia Artificial abre nuevos horizontes hacia áreas que hasta hace poco se consideraban futuristas. Esto se debe, en parte, a:

- Mayor potencia de cálculo en comparación con años anteriores;

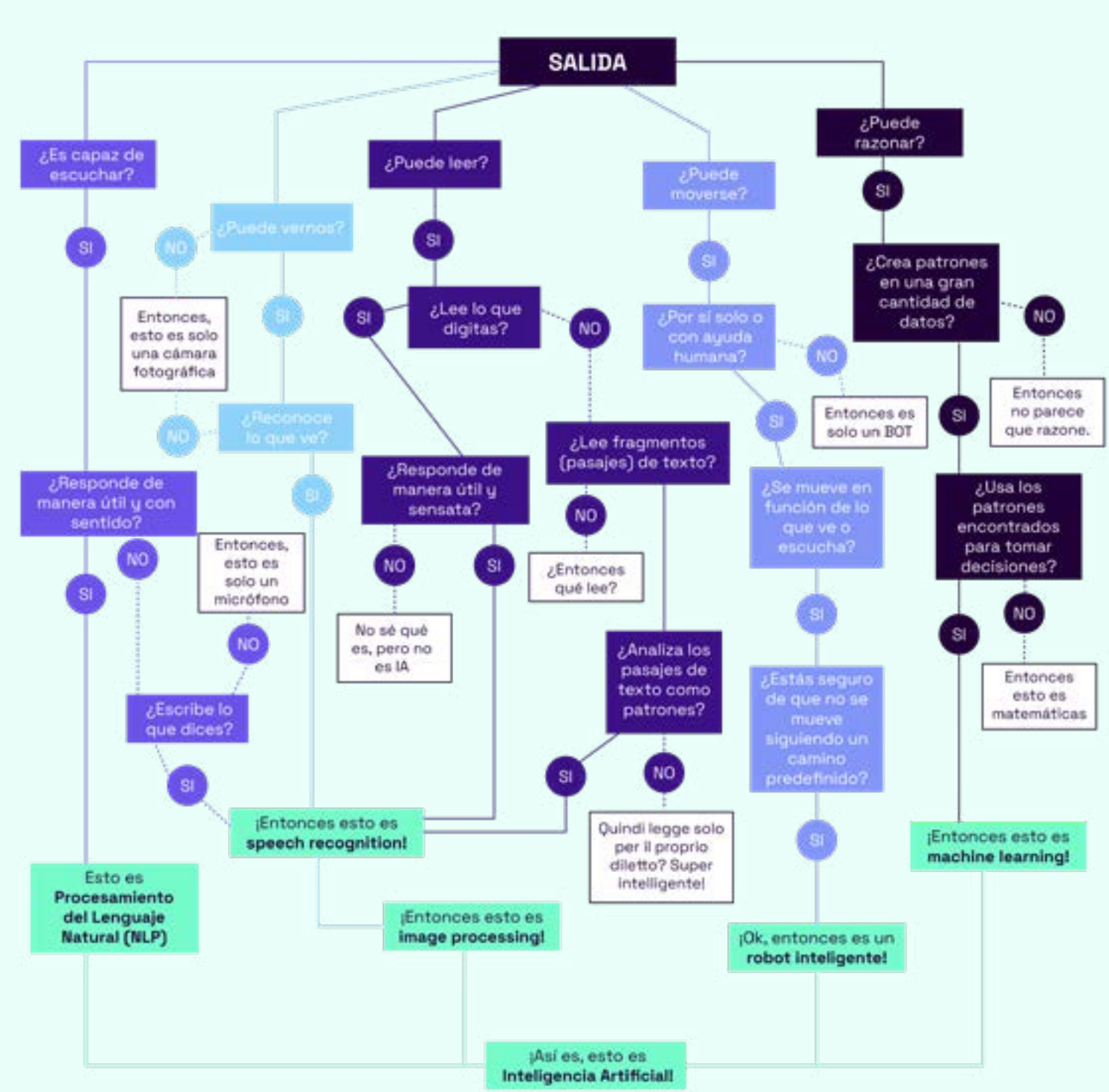
- Acceso a enormes bases de datos, que son fundamentales para el desarrollo de los sistemas.

La IA es capaz de realizar análisis que serían casi imposibles de hacer manualmente, ya que, al gestionar grandes cantidades de datos, logra explorarlos, encontrar correlaciones y extraer información útil sobre la que basar previsiones futuras.

Gracias a los numerosos beneficios que conlleva su uso, las inversiones de las empresas en soluciones de Inteligencia Artificial están creciendo de manera considerable. La International Data Corporation (IDC) ha detallado una previsión sobre el gasto global en IA en su informe “Worldwide AI and Generative AI Spending Guide”. Según este estudio, el gasto mundial en inteligencia artificial alcanzará los 632 mil millones de dólares para el año 2028. Este valor incluye inversiones en aplicaciones habilitadas para IA, infraestructuras y servicios IT y empresariales relacionados.

1.7 ¿Cómo identificar una aplicación de Inteligencia Artificial?

En términos generales, la IA se refiere a máquinas capaces de aprender, razonar y actuar de forma autónoma. Actualmente, la mayoría de las aplicaciones de IA de las que se habla pertenecen a la categoría de Machine Learning. Por lo tanto, la pregunta correcta no es "¿Cómo identificar una aplicación de IA?" (que es la capacidad de una máquina para realizar tareas comúnmente asociadas a seres inteligentes), sino "¿Cómo identificar una aplicación de Machine Learning?"



1.8 Diferencia con los modelos estadísticos clásicos

Es importante destacar las diferencias entre los modelos estadísticos tradicionales y las tecnologías más avanzadas como la Inteligencia Artificial. Ambas herramientas ayudan a comprender mejor los datos, pero lo hacen con enfoques muy distintos.
Los modelos estadísticos tradicionales llevan existiendo desde hace décadas, incluso antes de la aparición de los ordenadores. Se basan en pequeños conjuntos de datos y realizan suposiciones a priori; por ejemplo, la previsión de la demanda se basa en datos históricos de ventas pasadas. En este contexto, se asume que la demanda pasada es un buen indicador para predecir la demanda futura.
Sin embargo, en los últimos años han surgido nuevos factores que deben considerarse al realizar previsiones de demanda de productos. Entre ellos se incluyen:

- La disponibilidad de nuevos datos, como las festividades, datos provenientes del comercio electrónico o información sobre competidores;

- Una mayor potencia de cálculo, más accesible y eficiente económicamente gracias a servicios en la nube ofrecidos por proveedores como Microsoft, Amazon y Google.

Esta combinación permite controlar, mediante previsiones de ventas actualizadas constantemente, la creciente volatilidad de la demanda causada por factores como la introducción de nuevos productos o promociones comerciales. Por ello, el Machine Learning puede desempeñar un papel fundamental en la planificación de la cadena de suministro. En un contexto donde la volatilidad de la demanda es el principal problema, puede ser extremadamente difícil generar una previsión precisa. El ML no solo puede prever la demanda basándose en datos históricos de ventas, sino también teniendo en cuenta numerosos parámetros adicionales, como variables externas (condiciones meteorológicas, estacionalidad y festividades).

Hasta ahora hemos entendido que lo que sustenta una previsión precisa de la demanda son los datos. Es esencial que estos sean precisos y estén correctamente extraídos para poder analizarlos adecuadamente mediante un conjunto de técnicas conocidas bajo un término común: el Data Mining.

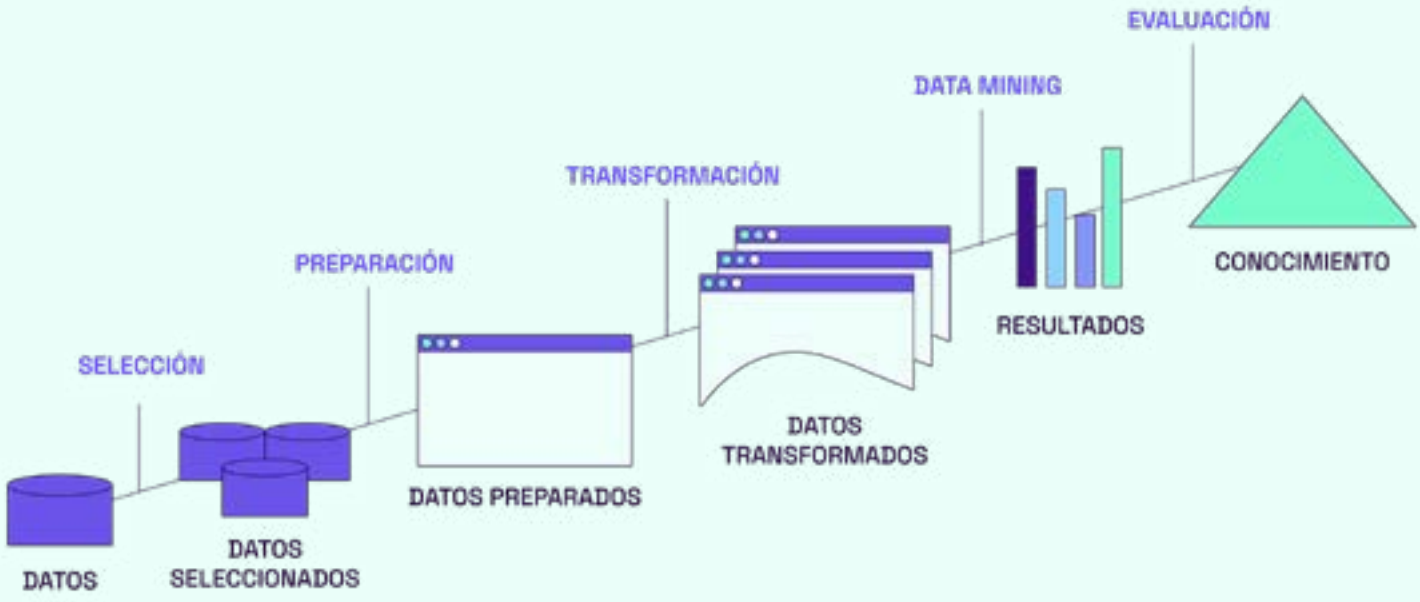
2.1 Data Mining: qué es y por qué es importante

El Data Mining hace referencia al conjunto de técnicas y metodologías que permiten extraer información útil a partir de grandes bases de datos. Los datos extraídos y procesados mediante este proceso se utilizan posteriormente para alcanzar objetivos predefinidos, como la maximización o minimización de una función objetivo.

En el sector del retail, este proceso puede ayudar a los propietarios de supermercados a comprender mejor las elecciones de sus clientes. El historial de compras de uno o varios consumidores, sus características demográficas o el horario en el que adquieren ciertos productos son solo algunas de las fuentes de datos a partir de las cuales es posible encontrar correlaciones y entender las preferencias de compra.

En el ámbito de la producción alimentaria, el Data Mining puede ayudar a los fabricantes a mejorar la eficiencia de sus productos. Por ejemplo, permite personalizar las líneas de producción en función de las preferencias de los clientes, identificando los productos más populares y ajustando la oferta. Si los datos muestran que un determinado tipo de pan es especialmente demandado en una temporada concreta, se puede aumentar su producción durante ese período.

Además, esta técnica permite segmentar a los clientes con características similares en grupos, conocidos como clusters, a los que se pueden dirigir promociones específicas. Por ejemplo, un cliente que compra con poca frecuencia pero en grandes cantidades recibirá una oferta distinta a la de un cliente que realiza compras más pequeñas pero con mayor regularidad.
Un proceso de Data Mining que conduce a la “extracción de conocimiento” puede representarse como se muestra en la siguiente imagen:



La selección, preparación, transformación y evaluación de los datos son fases esenciales para adquirir un conocimiento profundo y poder optar por la mejor solución a un problema. Siguiendo el ejemplo anterior, una posible solución al problema de "promociones más efectivas" sería identificar qué promociones dirigidas aplicar a un determinado cluster de clientes con el objetivo de aumentar su fidelización, lo que en el proceso anterior representaba el objetivo a alcanzar. Cuando hablamos de "datos en bruto", nos referimos a información que aún no ha sido procesada y que, en su estado original, puede carecer de significado. Un ejemplo de ello es el código binario: tomado simplemente como una sucesión de 0 y 1, no tiene ningún sentido para el usuario de un ordenador, pero cuando se procesa, puede transformarse en información comprensible. Por ejemplo, el código binario "01000001" no nos dice nada a simple vista, pero para una máquina representa la letra mayúscula "A".

Por este motivo, los datos en bruto requieren un procesamiento adicional. A menudo, pueden contener información incompleta que necesita ser complementada o datos erróneos que deben corregirse. A lo largo del proceso, el volumen de información se reduce considerablemente, pero su valor aumenta, ya que los datos se vuelven más precisos y menos propensos a errores. De hecho, al principio es necesario seleccionar datos en bruto a partir de enormes bases de datos, que suelen ser muy generales. Sin embargo, tras la fase de Data Mining, se obtienen patrones de información útiles para la toma de decisiones finales, los cuales serían imposibles de identificar sin este proceso.

2.2 La metodología CRISP-DM

Hemos visto que el Data Mining es una fase dentro del proceso de extracción de conocimiento, y existen varias formas de llevarlo a cabo. Uno de los modelos analíticos más utilizados para este propósito es el "Cross-Industry Standard Process for Data Mining" (CRISP-DM).
La metodología CRISP-DM se compone de seis pasos principales, que explicaremos a continuación con ejemplos aplicados a la previsión de demanda en el sector retail:

Business Understanding (Comprensión del negocio): En esta fase inicial, el foco principal es definir el problema empresarial o identificar el tipo de información necesaria. El objetivo es comprender el problema y analizar cómo su resolución impactará en la toma de decisiones estratégicas.

Se responde a la pregunta "¿Cuál es el problema que quiero resolver?"
Por ejemplo: mejorar la previsión de ventas de productos en promoción mediante algoritmos de predicción de demanda.

Data Understanding (Comprensión de los datos): Esta fase es esencial para la recopilación, comprensión y evaluación de la calidad de los datos. En particular, se analiza la fiabilidad de las bases de datos, identificando anomalías (conocidas como outliers) e incompletitudes que requieren ser complementadas con fuentes externas. En esta etapa, también se identifican las relaciones y tendencias entre las variables clave. Por ejemplo, se analiza la relación entre el precio de un producto y la cantidad vendida.

Se responde a la pregunta: "¿Qué datos necesitamos?"
En el caso de una cadena de supermercados, los datos clave pueden provenir de los tickets de compra, donde se registran los productos adquiridos, la hora de la transacción y los planes promocionales actuales y pasados.

Para una empresa de producción alimentaria, los datos relevantes pueden incluir las cantidades de materias primas adquiridas, los tiempos de producción y los comentarios de los clientes. Estos datos permiten analizar la eficiencia productiva y mejorar el control de calidad.

Data Preparation (Preparación de los datos):
Esta fase incluye todas las actividades necesarias para la construcción del dataset final, como la selección y limpieza de datos, la creación de nuevas variables y posibles transformaciones. Es la etapa más laboriosa en términos de tiempo dentro del proceso, ya que la preparación de los datos no es inmediata y requiere un análisis minucioso y una exploración detallada de la información disponible. En muchos casos, las variables originales de la base de datos inicial pueden no ser suficientes. A través de un proceso de Data Engineering, las variables existentes pueden ser combinadas y/o transformadas para generar nuevas variables más elaboradas, proporcionando información adicional al modelo de inteligencia artificial.

Se responde a la pregunta: "¿Los datos son suficientes, coherentes y precisos?" Antes de alimentar el modelo con datos, es necesario asegurarse de que estén correctamente preparados. Si los datos disponibles no son suficientes, es posible generar nuevas variables. Por ejemplo:

- Determinar en qué intervalo de tiempo se suele vender un determinado producto;
- Analizar en qué día de la semana o bajo qué condiciones meteorológicas se realizan ciertas compras;
- Agrupar a los clientes en clusters con características similares.;
- Estudiar la correlación entre productos, como el caso en el que la mayoría de los clientes que compran fresas también adquieren nata montada en spray.

Gracias a este proceso, se pueden preparar los datos para realizar una previsión precisa de la demanda.

Modeling (Modelado de datos): En esta fase se eligen y aplican los modelos de Machine Learning que permitirán alcanzar un objetivo previamente definido. Una vez seleccionada la tipología de modelos a utilizar, es fundamental llevar a cabo un proceso de ajuste de parámetros (tuning) para optimizar su rendimiento.

Durante esta etapa, el modelo se entrena con diferentes configuraciones de parámetros, obteniendo resultados más o menos satisfactorios en función de dichas configuraciones. Si, tras varias modificaciones, los resultados no son óptimos, será necesario volver a la fase de selección y creación de nuevas variables, con el fin de proporcionar al modelo datos más relevantes para su entrenamiento.

- Se responde a la pregunta: "¿Qué técnica se adapta mejor a los datos disponibles y conducirá a los resultados esperados?"

En el ámbito del retail y la producción, las variables y datos obtenidos en la fase previa se utilizan como input para un algoritmo de predicción de la demanda, que no solo se basa en datos históricos, sino que también incorpora factores externos como condiciones meteorológicas, festividades y estacionalidad.

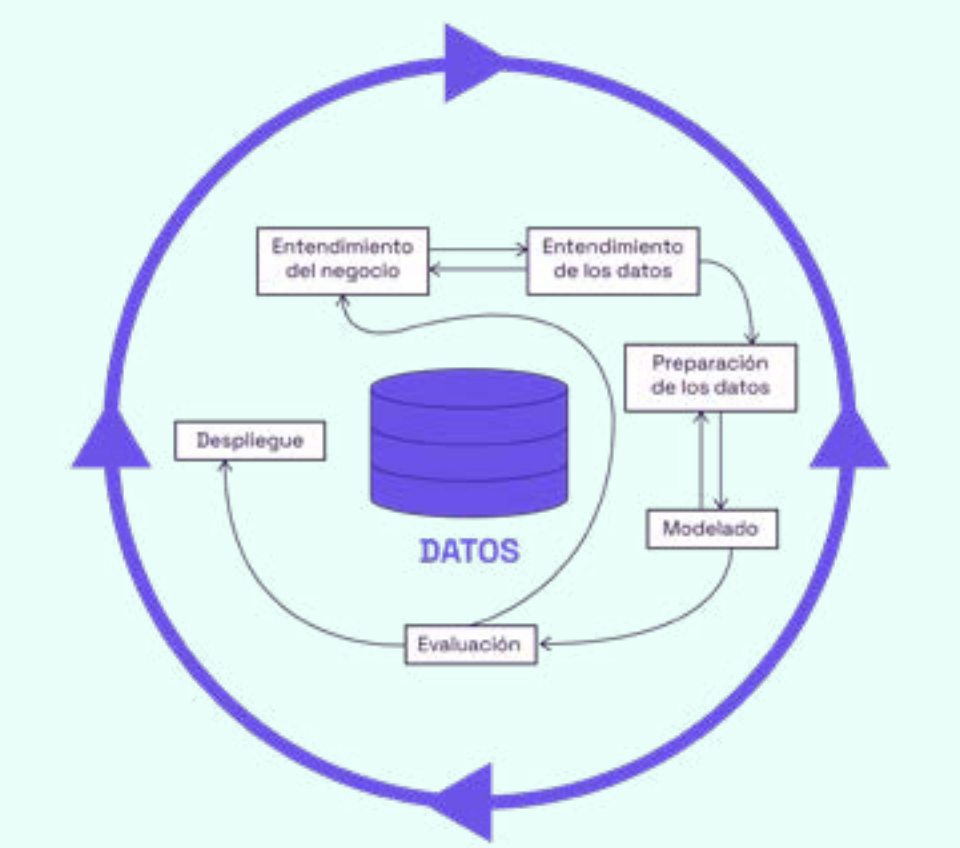
Evaluation (Evaluación del modelo): En esta fase, los modelos son analizados y comparados con el objetivo de seleccionar el que mejor se ajuste a los objetivos de negocio y determinar si existen puntos débiles que puedan comprometer su eficacia.

La pregunta clave en esta etapa es: "¿Estoy obteniendo una respuesta significativa a mi problema inicial?"
Para comprobar si el modelo es adecuado, es necesario analizar en qué medida los resultados obtenidos a partir del data mining cumplen con los criterios de éxito empresarial.

Deployment (Implementación del modelo): La correcta implementación del modelo es un paso clave para garantizar su integración efectiva en los procesos empresariales.

Se responde a la pregunta: "¿Cómo se pondrá en producción el modelo?"En esta fase, primero se lleva a cabo un proceso de planificación y monitorización para supervisar la aplicación de los resultados, seguido de la elaboración de un informe final para la revisión del proyecto.

Una de las principales diferencias entre este modelo y otros procesos típicamente lineales y unidireccionales es que las fases que componen CRISP-DM están diseñadas para ser iterativas y repetitivas.



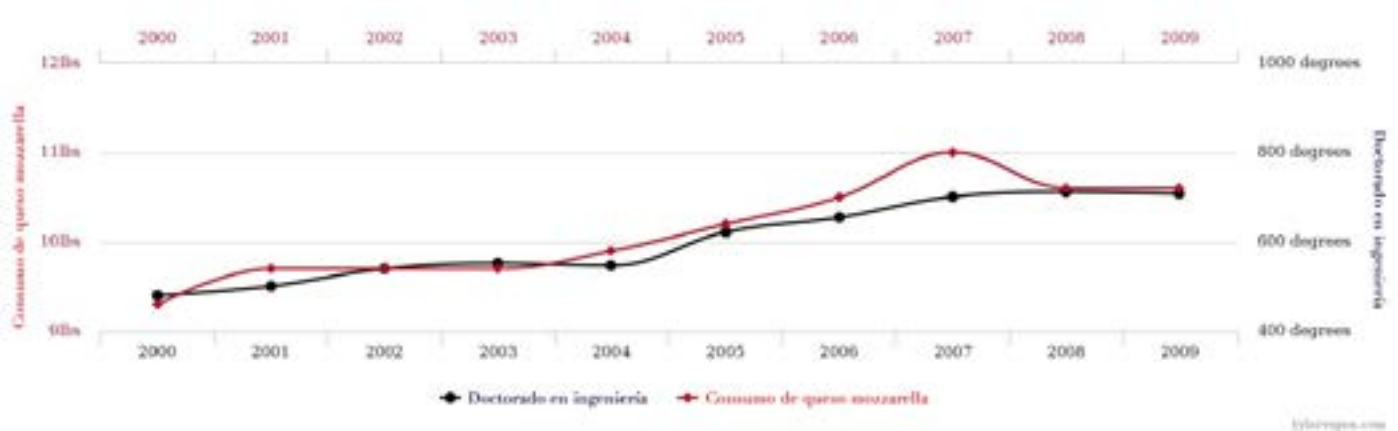
Por ejemplo, si durante la fase de Modelado (Modeling) se detecta que los modelos no ofrecen resultados válidos, es posible volver atrás a la fase de Preparación de Datos (Data Preparation) para revisar si es necesario seleccionar un subconjunto de datos diferente o si hay información faltante que deba ser incorporada. En un modelo lineal o unidireccional, en cambio, las fases avanzan en una única dirección sin posibilidad de retroceso. Un buen ejemplo sería la construcción de un edificio: los pasos están claramente definidos y no es posible construir la estructura y después rehacer los cimientos.

Por naturaleza, los proyectos de minería de datos son inciertos, ya que dependen en gran medida de la calidad y disponibilidad de los datos. Revisar, corregir e integrar nueva información son pasos esenciales en un proceso de minería de datos, lo que lo diferencia por completo de un modelo de desarrollo en cascada, típico en proyectos de ingeniería.

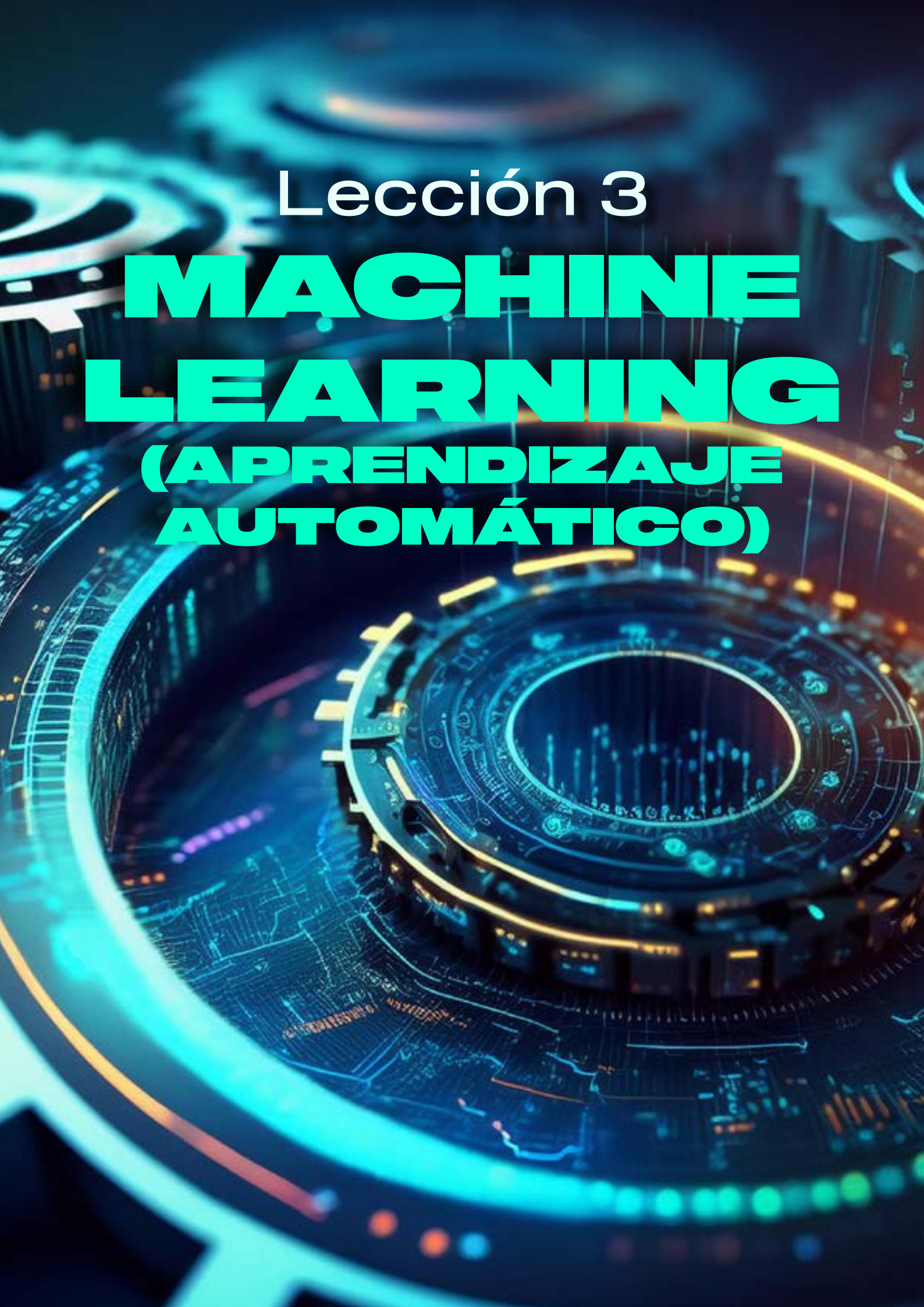
2.3 Data Mining e Inteligencia Artificial

Es fundamental destacar la importancia del Data Mining cuando hablamos de Inteligencia Artificial. De hecho, los datos obtenidos a través de este proceso sirven como entrada para la máquina, permitiéndole alcanzar los objetivos predefinidos. Por ello, es crucial que los datos seleccionados sean lo más precisos y fiables posible, ya que datos incorrectos solo pueden generar resultados erróneos. Este fenómeno, que debe evitarse, se conoce como el principio de "garbage in, garbage out": si los datos de entrada son incorrectos (garbage in), los resultados que se obtendrán también lo serán (garbage out).
En algunos casos, es importante diferenciar entre causalidad y casualidad. No siempre una información que parece estar correlacionada con un problema significa que realmente exista una relación lógica entre ambas.

Un claro ejemplo de este error son las correlaciones espurias, es decir, situaciones en las que un modelo identifica relaciones entre variables cuando, en realidad, no existe ninguna conexión lógica.



En el gráfico mostrado, se identifica una aparente correlación entre el número de personas con un doctorado en ingeniería y el consumo de mozzarella.
A simple vista, esta relación carece de sentido. Sin embargo, si el modelo de IA no cuenta con datos de calidad, podría detectar este tipo de correlaciones falsas y utilizarlas para tomar decisiones erróneas. Si el objetivo fuera predecir la demanda de un producto, basarse en una correlación incorrecta podría llevar a decisiones comerciales equivocadas, como aumentar innecesariamente el stock de mozzarella en base a datos sin fundamento.
Para que la Inteligencia Artificial pueda ser aplicada correctamente a modelos predictivos, es esencial contar con datos precisos y bien estructurados.
En este sentido, la calidad de los datos es más importante que la cantidad.
Hemos visto que, en sectores como el retail y la producción, una correcta previsión de la demanda requiere del uso del Machine Learning.
Pero, exactamente, ¿qué es el Machine Learning?


Lección 3
MACHINE LEARNING
(APRENDIZAJE AUTOMÁTICO)

ÍNDICE

3.1 Aprendizaje automático

3.2 Clasificación, regresión y agrupación

3.1 Machine Learning

El Machine Learning (ML) es una rama de la Inteligencia Artificial (IA) que se encarga de desarrollar sistemas capaces de aprender y mejorar su rendimiento en función de los datos que utilizan.

Cada día interactuamos con aplicaciones basadas en Machine Learning sin darnos cuenta. Un claro ejemplo es el reconocimiento de voz presente en muchos smartphones, que permite activar comandos mediante órdenes habladas.

Otra aplicación común es la publicidad personalizada: las empresas utilizan el ML para analizar el comportamiento de los usuarios en Internet y mostrar anuncios relevantes. Por ejemplo, si un usuario pasa mucho tiempo navegando por páginas sobre fútbol, es muy probable que los anuncios que verá en sus redes sociales estén relacionados con botas de fútbol o artículos deportivos.

El Data Mining y el Machine Learning están estrechamente relacionados. Mientras que el Data Mining se encarga de extraer y procesar información relevante, el Machine Learning utiliza estos datos como entrada para generar conocimiento más profundo y realizar predicciones más precisas.

Dependiendo del tipo de algoritmo utilizado para entrenar un modelo, el Machine Learning se puede clasificar en diferentes categorías. Entre las principales, encontramos:

• **El aprendizaje supervisado:** es una técnica que permite hacer predicciones a partir de modelos entrenados con datos supervisados. Esto significa que cada entrada de datos está asociada a un resultado esperado (output), el cual el modelo debe aprender a predecir. En este caso, el algoritmo se entrena con un conjunto de datos etiquetados, lo que significa que cada dato de entrada tiene asignada una etiqueta con información relevante. Supongamos que queremos predecir si un cliente dejará de comprar en una tienda o seguirá siendo fiel. Los datos de entrada pueden incluir: frecuencia de compra, número de productos adquiridos, nivel de satisfacción, etc.
A partir de estos datos, el modelo aprende a identificar patrones de comportamiento y genera una predicción. Cada caso de entrenamiento incluye un resultado ya etiquetado manualmente (por ejemplo, “cliente fiel” o “cliente que deja de comprar”). Este tipo de ML es ampliamente utilizado en escenarios donde hay datos históricos con resultados conocidos, como predicciones de ventas, diagnósticos médicos o reconocimiento de imágenes.

• **El aprendizaje no supervisado:** utiliza un enfoque más autónomo, en el que un sistema aprende a identificar patrones y estructuras en los datos sin intervención humana directa. A diferencia del aprendizaje supervisado, en este caso el modelo trabaja con datos sin etiquetas y no cuenta con una salida predefinida.

Por ejemplo, si proporcionamos a un sistema un conjunto de datos con las características de cinco especies de plantas sin indicar de qué especie se trata, el algoritmo analizará los datos y agrupará las plantas basándose en las similitudes que encuentre, sin que una persona defina previamente los grupos.

3.2 Clasificación, regresión y agrupación

Dentro del aprendizaje supervisado, los problemas que pueden resolverse se dividen en clasificación y regresión.

La clasificación se aplica a problemas en los que el resultado pertenece a un conjunto de opciones limitado y definido. Algunos ejemplos de problemas de clasificación son:

• Predecir si un cliente será insolvente o no basándose en su historial bancario.

• Determinar si una planta es carnívora o no a partir de sus características físicas.

• Identificar el continente de residencia de una persona en función del clima, el hábitat y otros factores.

Un método común de clasificación es el Decision Tree (Árbol de Decisión), ampliamente utilizado en Machine Learning supervisado por su rapidez y facilidad de interpretación. Este algoritmo funciona organizando los datos en una estructura de árbol en la que cada decisión se toma en base a ciertos criterios.

Los elementos clave de un árbol de decisión son:

• **Nodos:** puntos donde se realiza una división de los datos en función de una condición específica.
• **Hojas:** resultados finales o intermedios en los que se agrupan los datos después de pasar por las divisiones del árbol.

El objetivo es dividir el conjunto inicial de datos en grupos que deberían ser lo más homogéneos posible internamente (con características similares entre los componentes de un mismo grupo) y heterogéneos entre sí (con ca-

racterísticas diferentes entre los distintos grupos). Para comprender mejor este proceso, es útil ilustrarlo con un ejemplo. Supongamos que tenemos que decidir si jugar al tenis o no en función del clima y la temperatura exterior. Este problema, por supuesto, solo tiene dos posibles resultados, pero es necesario evaluar las combinaciones de las variables consideradas para determinar si jugar al tenis o no.

Los nodos son los lugares donde ocurren las divisiones, por lo que se considera la situación de si está lloviendo o no, mientras que las hojas son los resultados, los lugares donde encontramos los datos después de una división, como vemos en el último nivel. Considerando la situación de la izquierda, podemos observar que si el tiempo está soleado y la temperatura es inferior a 30 grados, es posible jugar al tenis. Si la temperatura supera los 30 grados, no se recomienda jugar al tenis. Lo que se muestra es un ejemplo muy simple de un árbol de decisión, mientras que los procedimientos basados en Machine Learning consideran cientos de variables y producen resultados derivados de miles de divisiones del árbol y diferentes niveles de profundidad. Se vuelve entonces comprensible que tomar decisiones precisas en problemas mucho más complejos sea imposible solo con razonamiento humano, por lo que es útil confiar en tecnologías avanzadas.

Los algoritmos de **regresión** responden a preguntas que no tienen un valor finito como respuesta. Algunos ejemplos de problemas de regresión son:

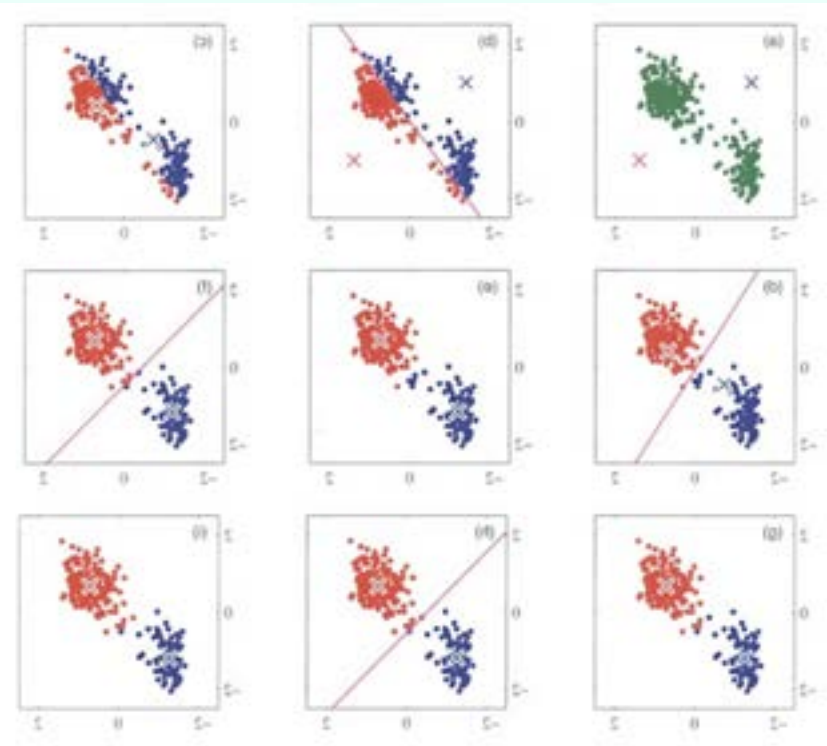
• En relación con la inflación y el comportamiento de los precios del trigo en el mercado, ¿cuánto costará la pasta por kilogramo el próximo mes? → (1€, 2€, 3€, etc.)
• ¿Cuál es la velocidad máxima de un coche con determinadas características? → (220 km/h, 220.5 km/h, 220.566 km/h, etc.)

En este sentido, la predicción de ventas también es un problema de regresión y se puede formular el problema de la siguiente manera:
• En relación con los precios, las ventas históricas, los precios de los competidores, eventos exógenos y las relaciones con otros productos, ¿cuál es la **predicción de ventas** para mañana de un determinado producto? → (1 unidad, 2 unidades, 3 unidades, etc.).

Algunas de las técnicas no supervisadas más populares utilizadas dentro del retail son, sin duda, los algoritmos de **clustering**, que identifican de forma autónoma las similitudes entre las observaciones y las dividen en función de características comunes. Estos grupos se denominan “clusters”.

Por ejemplo, consideremos a los clientes en línea de un comercio electrónico específico de artículos deportivos. Cada uno de ellos presenta características personales y comportamientos de compra, como sexo, edad, frecuencia de compra, etc. Utilizando un algoritmo de clustering, es posible agrupar a los clientes con el fin de ofrecerles promociones especiales relacionadas con su comportamiento de compra. A priori no se conoce el comportamiento de compra de los clientes ni se sabe si son deportistas a nivel profesional o aficionado, ni si son clientes ocasionales o habituales. Será tarea del algoritmo mismo descubrir si entre los comportamientos de compra de los clientes hay características que permitan definir grupos de consumidores.

K-means es uno de los algoritmos de clustering más utilizados y eficientes. El objetivo de este algoritmo es agrupar las observaciones similares, descubriendo patrones ocultos y dividiendo el conjunto de datos en un número k de clusters. Un cluster es, por lo tanto, un conjunto de observaciones que comparten características similares.

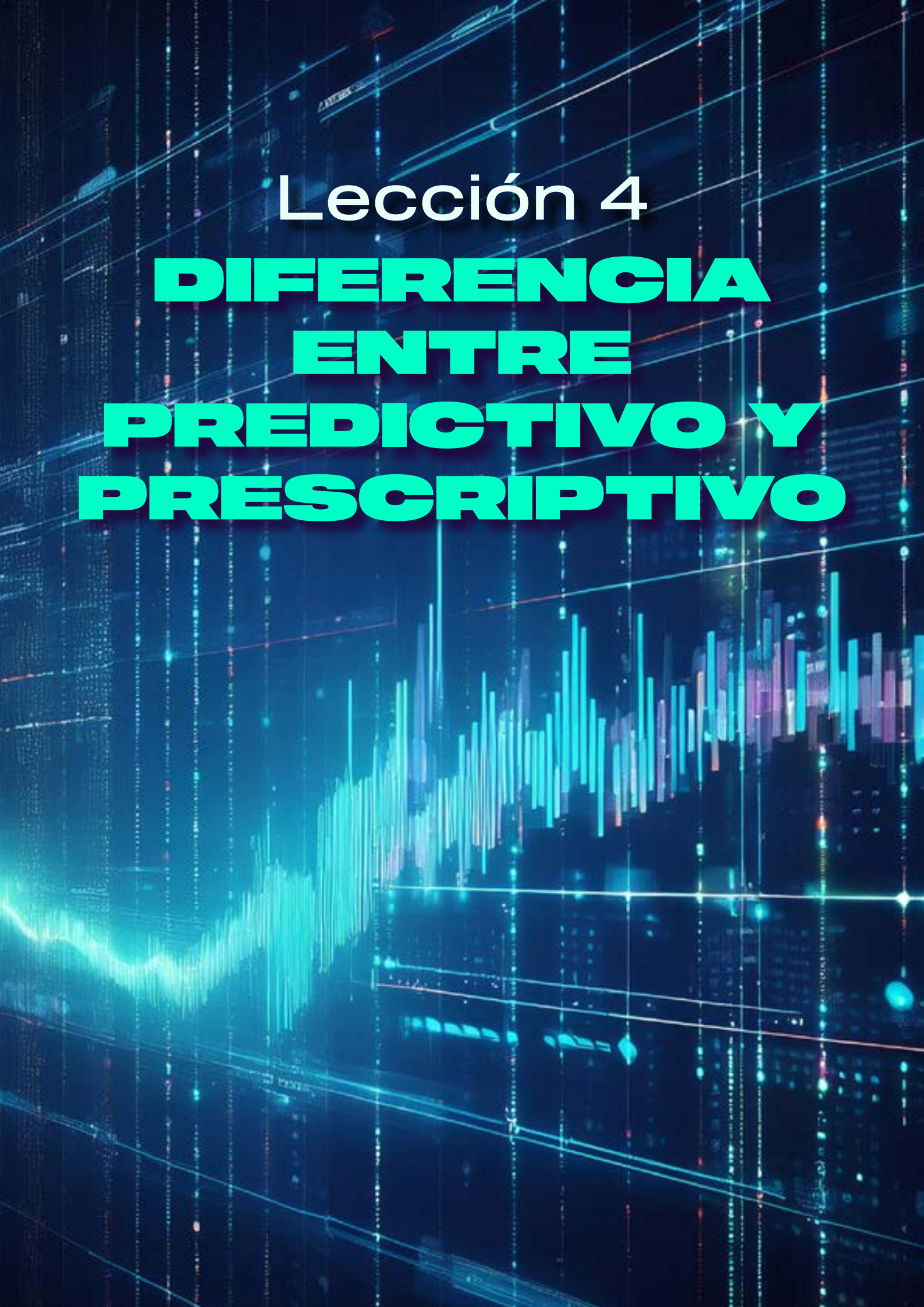


Una vez definido el número de clusters K deseado, el algoritmo se encargará de agrupar las observaciones sobre las que se entrene.

En la imagen se puede observar cómo, durante el proceso de aprendizaje, el algoritmo trata de dividir progresivamente un conjunto de observaciones de manera cada vez más clara. En la novena iteración (en la parte inferior derecha), el algoritmo ha delineado dos grupos distintos de observaciones que se pueden diferenciar claramente.

Una característica extraordinaria del Machine Learning es su capacidad predictiva. En el pasado, las decisiones empresariales a menudo se tomaban basándose en resultados históricos, mientras que hoy, como hemos visto gracias a esta tecnología, se pueden utilizar análisis mucho más avanzados que tienen en cuenta muchos más factores, como por ejemplo, las variables externas.

Como se analizó en la segunda lección, en la que se presentaba la metodología CRISP-DM, el problema consistía en predecir las ventas en caso de promoción (pág. 8), y lo que se buscaba era, por tanto, un análisis predictivo. Sin embargo, si tuviéramos que resolver un problema relacionado con la decisión de compra o el reabastecimiento de productos, este análisis predictivo no sería suficiente. De hecho, para alcanzar este objetivo es necesario introducir el concepto de análisis prescriptivo.


Lección 4
DIFERENCIA ENTRE PREDICTIVO Y PRESCRIPTIVO

INDICE

4.1 Diferencia entre análisis predictivo y prescriptivo

LEC. 4: DIFERENCIA ENTRE ANÁLISIS PREDICTIVO Y PRESCRIPTIVO

4.1 Diferencia entre análisis predictivo y prescriptivo

Un modelo predictivo analiza datos actuales e históricos para hacer predicciones, como la probabilidad de que un evento se repita en una situación determinada. Por ejemplo, es posible predecir, basándose en datos históricos y variables externas, cuántos kilos de plátanos se venderán la próxima semana en un punto de venta específico según las condiciones meteorológicas que haya. Pero, ¿es esto suficiente para maximizar las ganancias de los minoristas? No, porque hay otras series de variables que deben ser consideradas para comprender la cantidad correcta de producto que se debe volver a pedir, como los inventarios y las restricciones del proveedor.

Por esta razón, los minoristas deberían adoptar un análisis prescriptivo, mediante el cual, partiendo de los resultados de predicción, la máquina es capaz de ofrecer recomendaciones detalladas para tomar decisiones que maximicen el rendimiento económico de la empresa. Un ejemplo en el área de producción: **análisis prescriptivo frente a análisis predictivo.**

A través del Machine Learning, como ya hemos visto, es posible acceder a una amplia gama de datos internos y externos, revelar patrones y correlaciones previamente desconocidas a nivel de tienda, producto o estante. Luego, es posible optimizar los resultados obtenidos para proporcionar recomendaciones semanales o incluso diarias para ajustar los precios, las promociones o el surtido de cada punto de venta con el objetivo de maximizar las ganancias. Según datos de McKinsey, el análisis prescriptivo puede aumentar las ventas entre un 2% y un 5%. Pero, ¿qué le falta al ML para lograr este resultado? Es necesario tener en cuenta muchas otras restricciones, como en el caso de un minorista o mayorista, el tiempo de entrega, la cantidad mínima que se debe reordenar, el stock de cobertura, las restricciones de pallets, los descuentos por valor, etc. Para tener en cuenta todas estas variables simultáneamente, es necesario utilizar algoritmos matemáticos de optimización lineal, que cruzan las previsiones de ventas con las restricciones mencionadas, eligiendo así la opción que maximice la diferencia entre ingresos y costos, es decir, el margen.

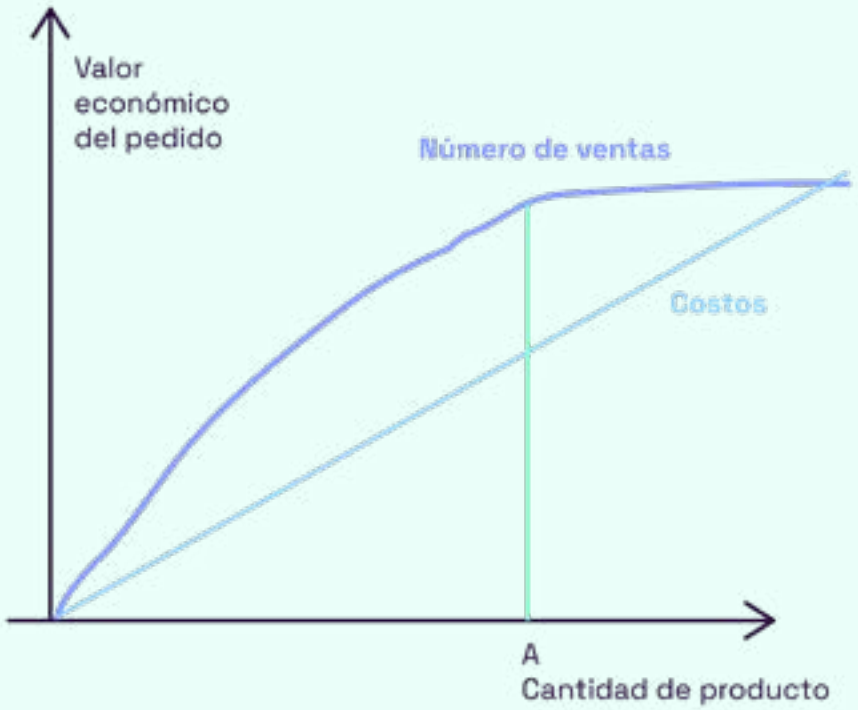
Este es el objetivo del análisis prescriptivo: analizar diversas situaciones y sugerir la que sea económicamente más rentable. El principal resultado del análisis prescriptivo en el sector del retail y la producción es, sin duda, tomar decisiones más acertadas económicamente. Sin embargo, las consecuencias de adoptar este enfoque son muchas.

LEC. 4: DIFERENCIA ENTRE ANÁLISIS PREDICTIVO Y PRESCRIPTIVO

Por ejemplo, en el sector retail, es posible determinar si mantener en el estante productos de baja rotación que atraen a los mejores compradores o si es mejor retirarlos para introducir nuevos productos que puedan aumentar la rentabilidad de áreas específicas del punto de venta. Esto es algo que el análisis predictivo por sí solo no podría hacer, ya que se limita a predecir únicamente las futuras ventas de los productos.

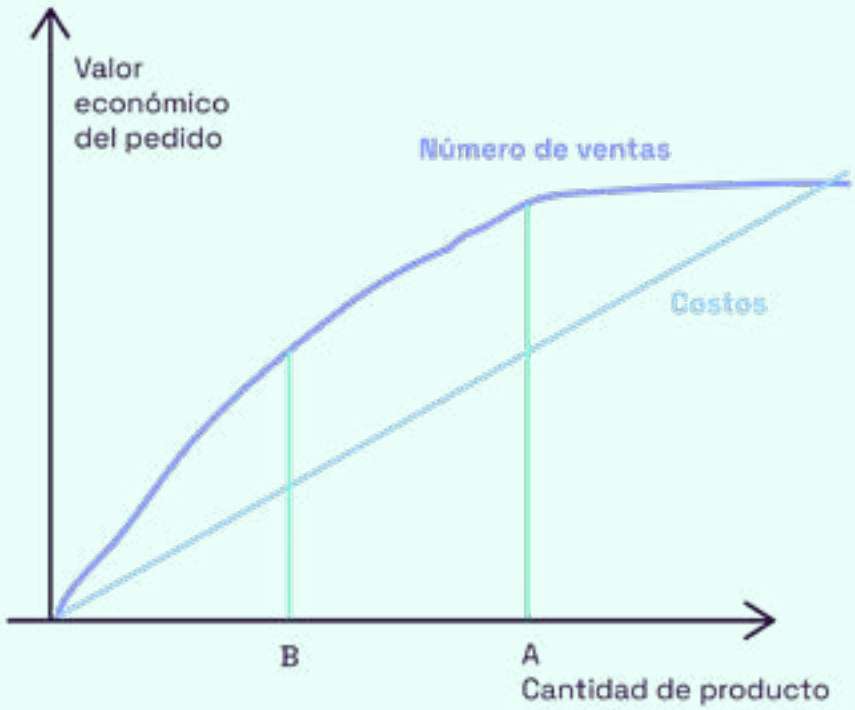
Otro ejemplo de la aplicación de este enfoque es el reconocimiento de patrones para recomendar el momento adecuado para promociones y ajustes de precios de productos. Sin embargo, estas correlaciones no son fáciles de identificar, ya que las variables a considerar son siempre muy complejas y diversas. Por ejemplo, en algunas ciudades, las ventas de aguacates y patatas fritas podrían aumentar un 5% el viernes antes de un evento deportivo importante, pero la herramienta prescriptiva podría recomendar almacenar un 20% más de estos dos productos en las tiendas donde es más probable que los equipos locales lleguen a tiempo extra.

Este análisis puede entonces recomendar cambios específicos en el surtido y estimar el valor monetario de dichos cambios. En el ámbito de los aprovisionamientos, podemos entender la importancia del análisis prescriptivo para las empresas del sector alimentario, ya que la recomendación que ofrece es la cantidad que se debe reordenar, buscando maximizar las ganancias al encontrar un equilibrio entre el máximo nivel de ventas alcanzables y el mínimo nivel de costos sostenibles. Para explicar mejor este concepto, puede ser útil observar este gráfico:



LEC. 4: DIFERENCIA ENTRE ANÁLISIS PREDICTIVO Y PRESCRIPTIVO

Podría pensarse que la cantidad óptima de productos a ordenar/producir corresponde al punto A, ya que representa un máximo de ventas. Sin embargo, un modelo basado en análisis prescriptivos podría sugerir que la cantidad ideal se encuentra en el punto B, ya que representa cantidades más pequeñas de productos, pero también costos proporcionalmente más bajos. Esto significa que lo que se busca no es la cantidad máxima de ventas, sino la cantidad de productos que maximice la diferencia entre ingresos y costos.



Esto ocurre porque las ventas no son el único factor determinante; hay otros aspectos a considerar, como el aumento de la cantidad de productos, que no es directamente proporcional al aumento de las ventas y puede generar un mayor riesgo de sobrestock. Por lo tanto, el beneficio que las empresas deben buscar no es tanto mejorar la precisión de las previsiones, sino tomar decisiones más rentables desde el punto de vista económico. Algunos de los indicadores de desempeño objetivo que se pueden obtener son:

- Reducción de los excedentes de inventario hasta un 46%;
- Incremento de la eficiencia de ventas en un 18%;
- Disminución de los costos de inmovilización de inventario en un 8%.

Porque, como siempre se dice, ¡el futuro es de quienes saben anticiparlo!