

# Introducción a la inteligencia artificial y el aprendizaje automático



---

# Agenda

**1**

Introducción y objetivos

**4**

Redes neuronales y  
aprendizaje profundo

**2**

Inteligencia artificial

**5**

Ciencia de datos en la  
actualidad

**3**

Aprendizaje automático

**6**

Ejemplos y casos de  
uso en Alicorp

# Introducción y objetivos de la sesión

---

## Introducción

El concepto de un aprendizaje e inteligencia no humana tiene su origen en la antigüedad. La evolución de este concepto y este conocimiento en conjunto la masificación de herramientas computacionales y ciertas tecnologías en la actualidad han dado origen a una nueva disciplina encargada de analizar información y crear herramientas que aprenden de forma autónoma.

Para entender el rol y la profesión del científico de datos moderno, es necesario empezar por entender el origen de los conceptos que usa.

## Objetivos

- Entender el origen y concepto de aprendizaje automático a través del tiempo
- Conocer a alto nivel los conceptos y teoría enmarcada en el aprendizaje automático
- Conocer las diferentes industrias y casos de uso más comunes para las herramientas de aprendizaje automático
- Entender el porqué del *hype* de las redes neuronales y *Deep Learning* en la actualidad

# El concepto de inteligencia artificial es variado y diverso

## ¿Qué entendemos por IA?

- Tenemos conceptos predefinidos de la inteligencia artificial, de hecho tenemos dos: **IA positiva e IA negativa**
- En concreto la IA es un conjunto de elementos y conceptos en lugar de una disciplina establecida o derivada de las ciencias cognitivas

### IA positiva



Baymax, nuevos héroes



R2D2, Star Wars

### IA negativa



HAL9000, Odisea en el espacio



Terminator

# Pero, ¿qué es realmente inteligencia artificial?

Es un concepto que evoluciona en el tiempo

---

## La inteligencia artificial es la convergencia de dos conceptos

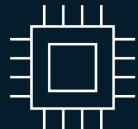


### ¿Qué es inteligencia?

“Inteligencia es el conjunto de procesos que nos permiten comprender, aprender o adaptarnos a nuevas situaciones”.

Nuestro cerebro descompone este proceso en varias habilidades:

- Percepción
- Orientación
- Memoria
- Pronóstico
- Lengua
- Aprendizaje
- Emociones
- Cálculo
- Locomoción



### ¿Por qué artificial?

Es artificial debido a que las acciones y funciones son realizadas por una máquina o programa informático capaz de replicar las habilidades previas.

## ¿Cómo evoluciona?

El concepto de inteligencia artificial **no es estático en el tiempo**, sino que es constantemente redefinido de acuerdo al contexto y el avance de la tecnología.

Hace unos años, jugar al ajedrez con el computador era considerado IA, ahora que un carro se conduzca solo es lo que entendemos por IA.

# Línea del tiempo de la inteligencia artificial

El concepto ha evolucionado a través del tiempo

Boulier  
2500 a. C.



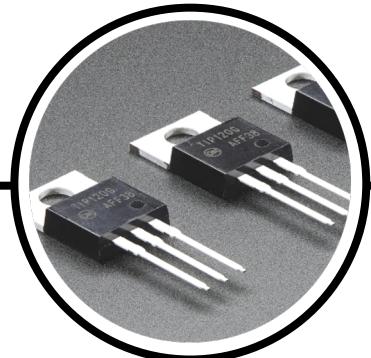
Pascalina  
1642



Cartas perforadas  
1725



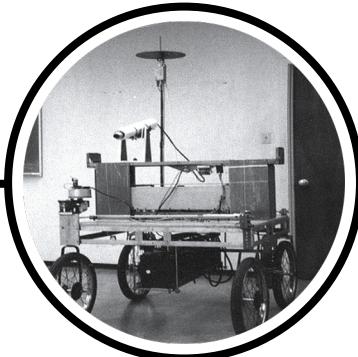
Transistor  
1947



Eliza  
1966



Stanford Cart  
1979



Ajedrez  
1997



Go  
2016



# Existe 4 principales doctrinas para definir el concepto de IA

En realidad “la inteligencia artificial es una combinación de teorías y técnicas utilizadas para crear máquinas capaces de replicar o simular la inteligencia humana.” Es un conjunto de conceptos y tecnologías más que una disciplina por derecho propio.

## Sistemas de pensamiento Like man

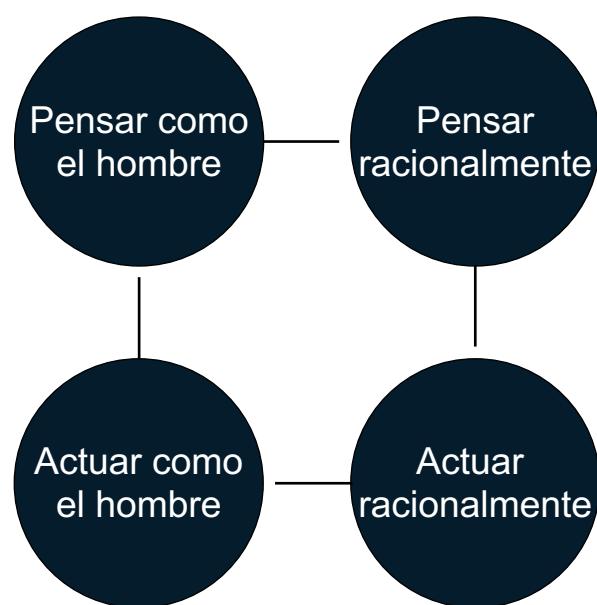
¿Cómo piensa el hombre?  
¿A qué nivel de abstracción desarrolla teorías?

1. Sistemas de apoyo a la toma de decisiones y
2. Resolución de problemas (Bellman 1978)

## Sistemas que actúan Like man

Prueba de Turing (1950): no ser capaz de distinguir a un hombre de una máquina.

1. Computadoras capaces de hacer ciertas manchas mejor que los hombres (Rich - Knight 1991)



## Sistemas de pensamiento racional

Vínculos entre matemáticas, filosofía e IA: codificación del conocimiento.

1. Cálculos computacionales capaces de capturar y razonar (Winston 1992)

## Sistemas que actúan Racionalmente

Hacer lo correcto: de acuerdo con la información disponible, maximizar el logro de una meta.

1. Automatización del comportamiento inteligente (Luger - Stubblefield 1993)

# ¿Cuáles son los campos de estudio y aplicación de la IA?



## Razonamiento

El razonamiento permite a un sistema realizar inferencias basadas en datos.

El razonamiento permite inferir reglas y datos cuando están incompletos.



## Lenguaje Natural

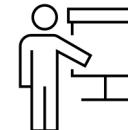
El procesamiento del lenguaje natural (PNL) está interesado en entender el texto escrito y el discurso humano.

Las técnicas de PNL son necesarias para capturar el significado de un texto no estructurado de los documentos o comunicaciones de un usuario.



## Planificación

La planificación automatizada permite que un sistema inteligente actúe de forma autónoma y flexible para tomar medidas para lograr un objetivo final, adaptándose a un entorno determinado.



## Aprendizaje

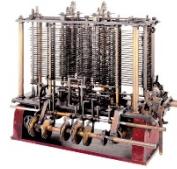
El aprendizaje automático consiste en aprender funciones (previsión, clasificación, acciones) a partir de datos, sin necesidad de programación explícita.

# Evolución de la potencia de una máquina

En los últimos años el avance de la IA se ha impulsado debido al avance de la computación

El cambio acelerado con el crecimiento exponencial en el poder computacional están revolucionando el campo de la IA

La tecnología computacional ha aumentado dramáticamente por potencias de 10, ahora procesa mas información en una hora que lo que hizo en los primeros 90 años



**Motor analítico**  
Nunca totalmente construido, invento de Charles Babbage y diseñado para resolver problemas lógicos y computacionales



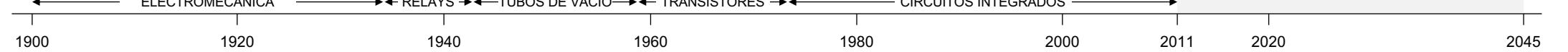
**Colossus**  
Computador electrónico con 1500 tubos de vacío. Ayudó a los ingleses a vencer el código alemán en la WWII



**UNIVAC I**  
El primer computador comercial, usado para tabular el censo de EEUU



**Power Mac G4**  
Primer computador personal con más de 1 billón de operaciones flotantes por segundo



---

# Aprendizaje automático

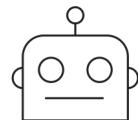
# ¿Qué es el aprendizaje automático?

"Aprendizaje automático es un campo de investigación que da a las computadoras la capacidad de aprender sin ser programadas explícitamente"

(Arthur Samuel, 1959)

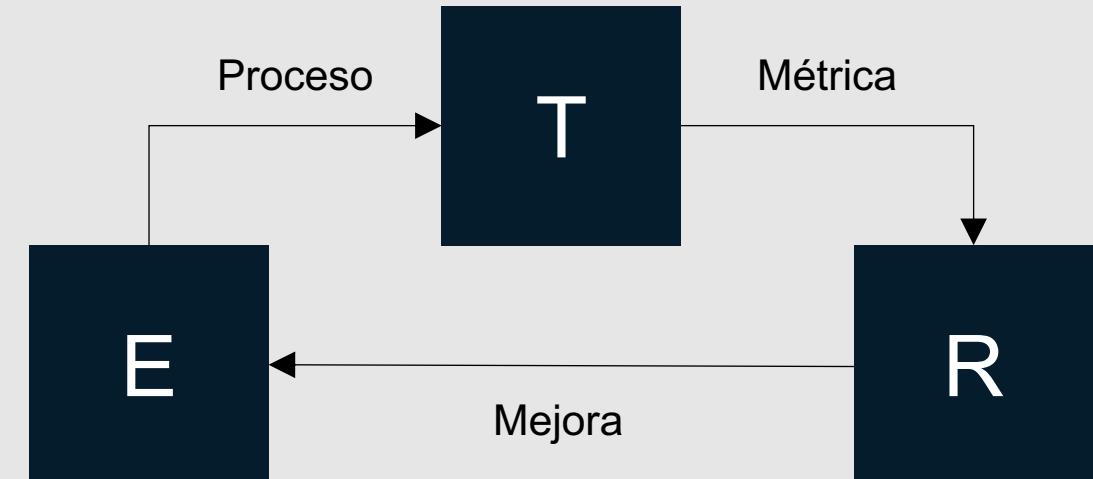
Aprendizaje automático (ML) se refiere a herramientas, modelos y algoritmos, cuyo rendimiento aumenta a medida que se exponen a nuevos datos.

Los algoritmos de ML permiten, en cierta medida, a un sistema impulsado por computadora o un sistema asistido por computadora, **adaptar sus análisis y comportamientos, basados en el análisis de datos empíricos** que se le provee.

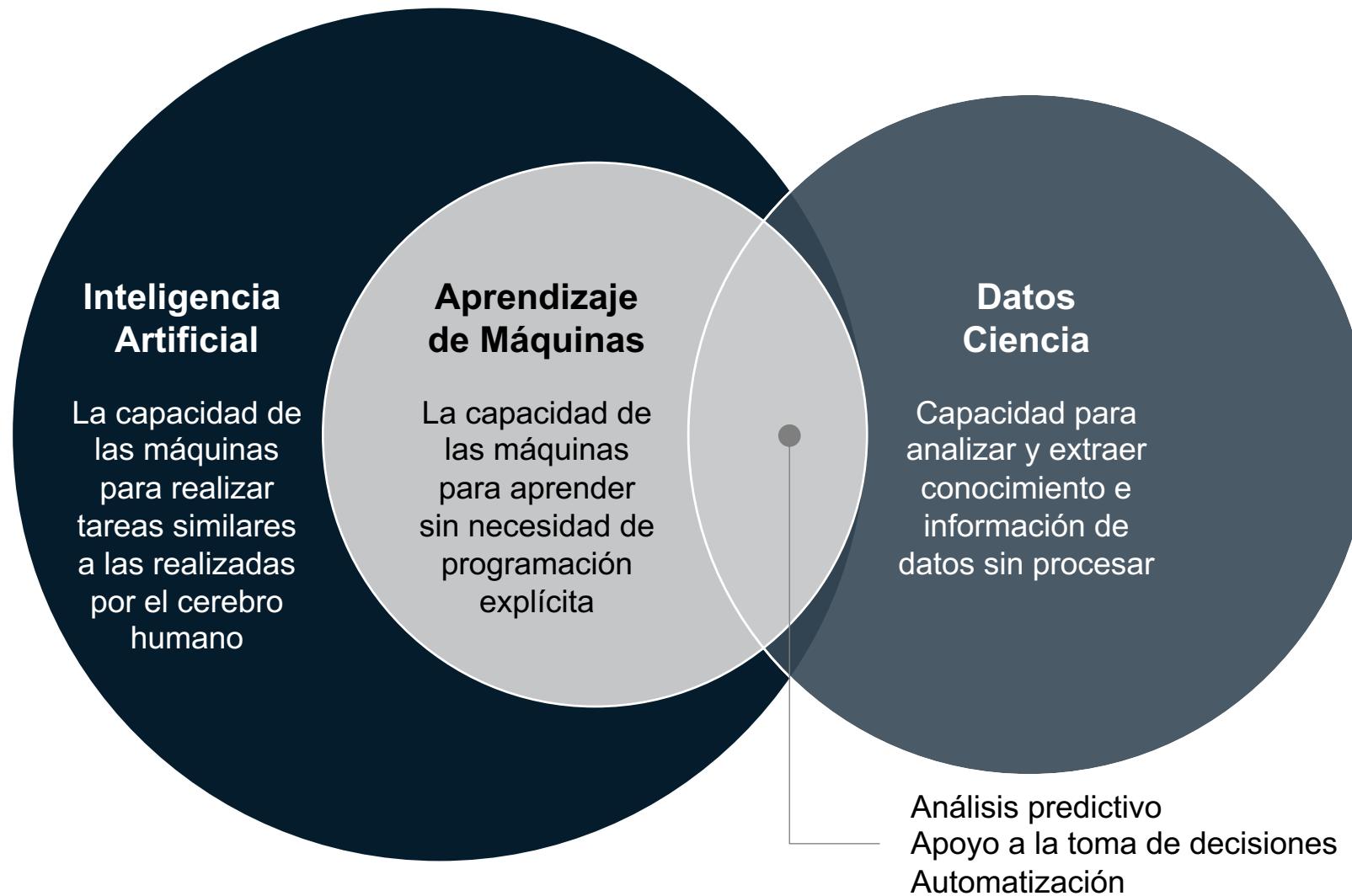


## Concepto

Se dice que un programa informático aprende de la **experiencia E**, con respecto a una **tarea T**, y una cierta métrica de **rendimiento R**; si su desempeño en T, medido por R, mejora con la experiencia E. (*T. Mitchell*)



# ¿Cómo se relaciona la IA y el aprendizaje automático?



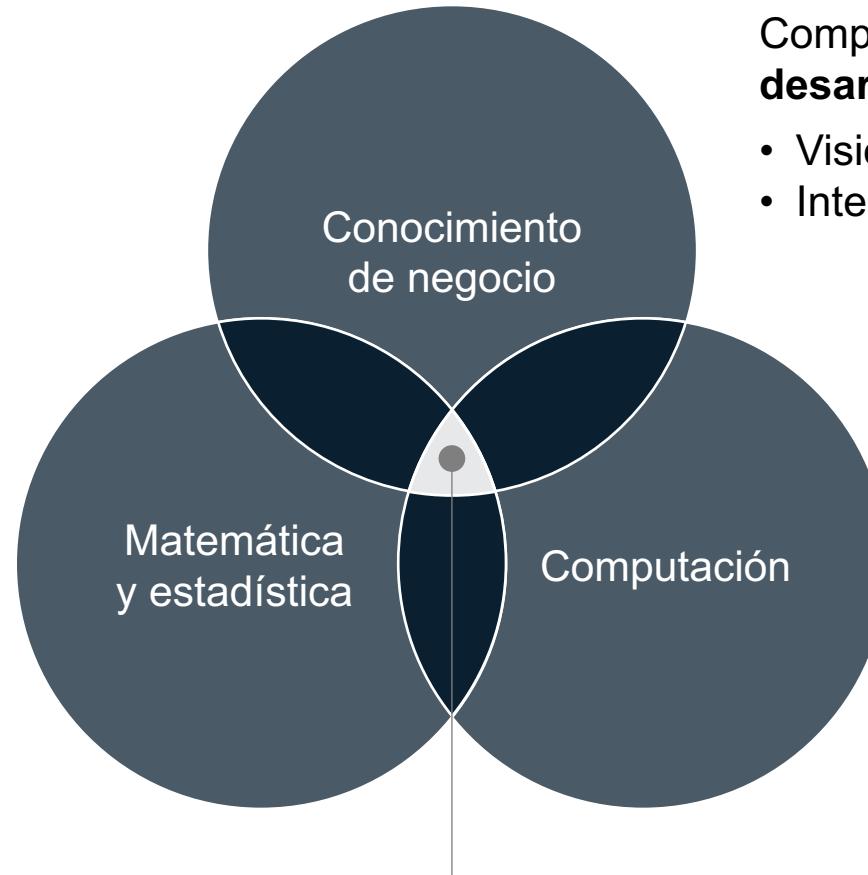
La inteligencia artificial contiene en si mismo los conceptos que aprenderemos en el curso, como aprendizaje de máquinas. Sin embargo como nueva disciplina, la **ciencia de datos** es un conjunto de habilidades y conocimientos que sobrepasan la creación de algoritmos que aprendan automáticamente, ya que por ejemplo incluye **nuevas tecnologías, conceptos estadísticos y experiencia de negocio** según sea la aplicación.

# Luego, ¿cuál es el perfil del científico de datos actual?

---

Comprende la **teórica básica estadística** con conocimientos como:

- Algebra lineal
- Probabilidad
- Calculo
- Lógica
- Optimización



**La ciencia de datos es la intersección de disciplinas previamente separadas**

Comprende la **integración de los desarrollos con el negocio** usando:

- Visión táctica y estratégica del negocio
- Interpretación de los datos desde negocio

Comprende la **tecnología y herramientas para el desarrollo** soluciones con habilidades en:

- Bases de datos
- Programación
- Complejidad
- Algorítmica

# **Sin embargo, si las disciplinas existían antes, ¿por qué ahora es el momento de unirlas?**

Estamos en una época inigualable para el uso de la información

---

## **Fenómeno**

## **Descripción**



**Acceso a datos  
(Internet)**

La recopilación de datos de diversas fuentes y formatos se ha vuelto estándar a través de industrias



**Potencia informática  
(GPU, cloud)**

La tecnología ha avanzado para hacer accesible gran poder computacional a través de tecnologías GPU en ordenadores personales o servicios cloud



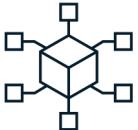
**Acumulación de  
innovaciones  
algorítmicas**

Conceptos tradicionales de estadística han sido potenciados por complejos algoritmos que se apoyan en poder computacional

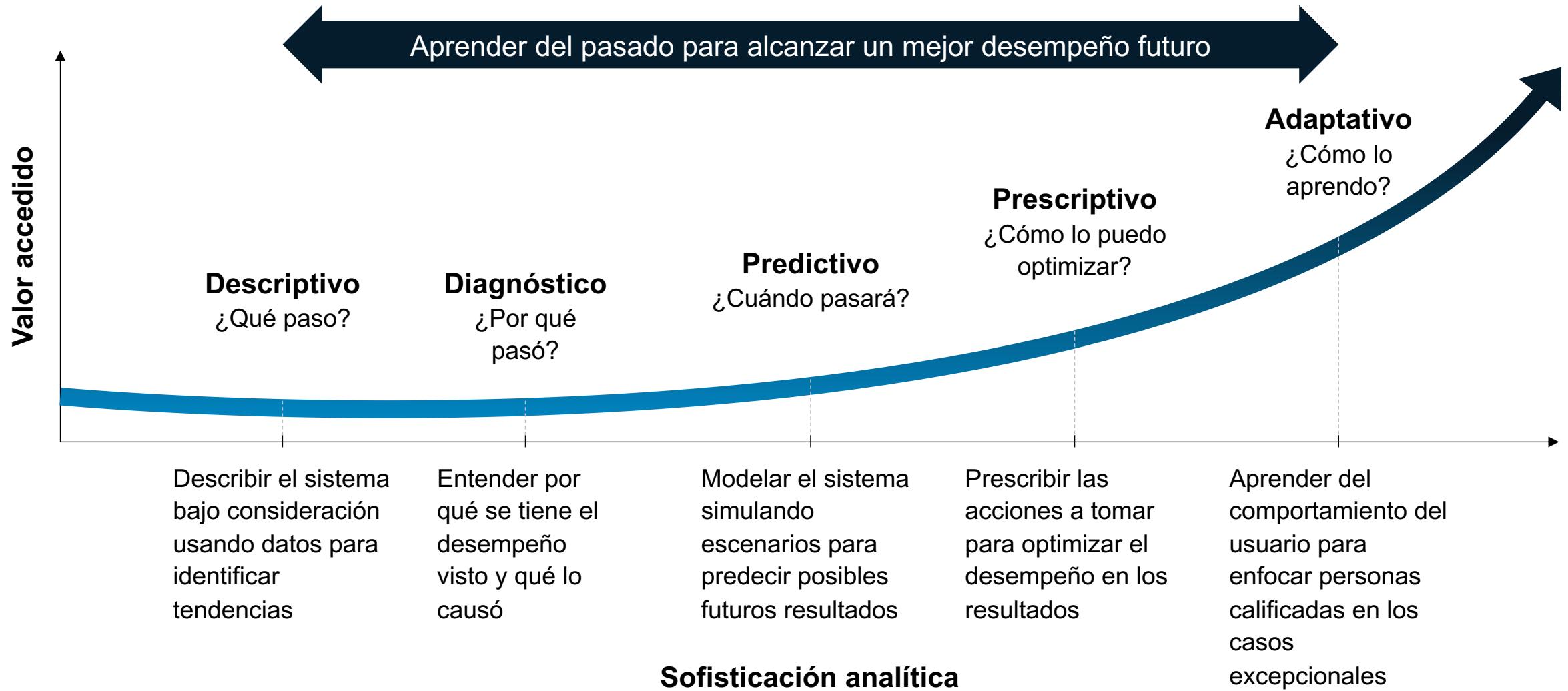
# **La dinámica es tan rápida que tenemos una nueva disciplina para esto: Big Data**

Estas tendencias se formalizan en el concepto de las 3 Vs

---

| <b>Concepto</b>   | <b>Descripción</b>  |
|---|---|
|    | <b>Volumen</b><br>Grandes cantidades de información se acumulan de forma continua, proveniente de diversas fuentes, por ejemplo redes sociales o IoT  |
|    | <b>Variedad</b><br>Datos heterogéneos, estructurados, semi-estructurados, no-estructurados, estáticos, dinámicos, distribuidos, ...   |
|  | <b>Velocidad</b><br>La información es generada y acumulada a gran velocidad, lo cual exige gran velocidad de procesamiento y análisis para que la información sea fundamento de la operación en muchas industrias |

# En el espectro de soluciones analíticas, el aprendizaje automático encuentra más potencial en etapas pos-predictivas



# Tipos de aprendizaje automático

Existen 3 tipos de teorías principales en la actualidad

## Aprendizaje Supervisado

### “Aprendemos de ejemplos”

La máquina recibe entradas de ejemplos y resultados deseados. El objetivo es aprender una regla general que permita correlacionar la entrada con las salidas



Dado

$$X_i, Y_i$$

el algoritmo busca el mejor  $f$  tal que:

$$Y_i = f(X_i)$$

## Aprendizaje no supervisado

### “Aprendemos explorando los datos”

La máquina recibe ejemplos sin los resultados deseados. Debe descubrir por sí mismo la estructura de los datos (sus relaciones y distribución)



Dado

$$X_i$$

el algoritmo busca el mejor  $P$  tal que:  
 $P(X_i)$

Aprenda la estructura de datos de  $X$

## Aprendizaje reforzado

### “Aprendemos experimentando”

La máquina experimenta en un entorno. Debe aprender una acción para los estados posibles. La acción produce una recompensa que guía el aprendizaje



Un sistema con estado  $S(X_i)$  y acciones  $X_i$  y recompensa  $R(X_i, S_i)$   
el algoritmo busca el mejor  $F$  tal que:

$$X_i = F(S_i)$$

maximize  $R(X_i, S_i)$

# Modelos más comunes

Existen múltiples modelos, siendo los más comunes

---

Aprendizaje  
Supervisado

Clasificación

Regresión

Logística

Regresión lineal

Árbol de decisión  
- Bosque Rnd

Árbol de decisión  
- Bosque Rnd

Aprendizaje  
No supervisado

Clustering

Reducción  
dimensional

K-means

Pca

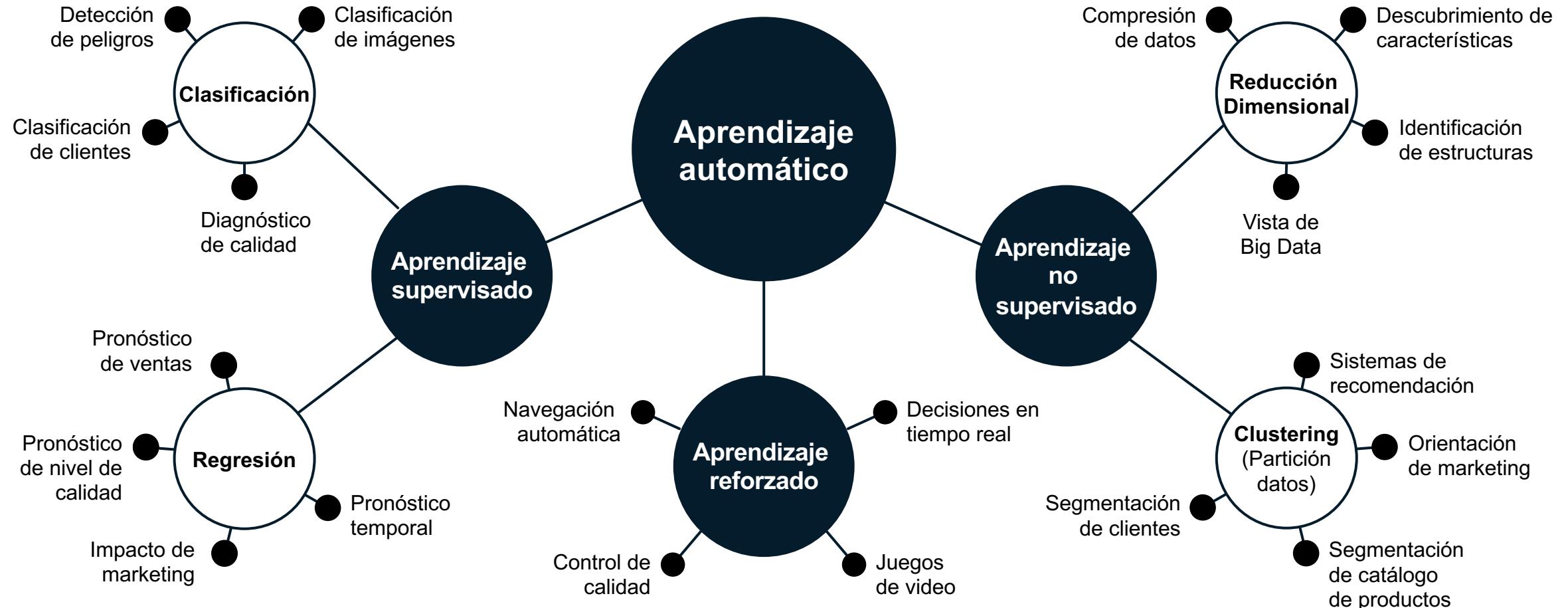
Basado en la  
densidad

Análisis de  
discriminación  
lineal

Aprendizaje  
Reforzado

Q-learning

# ¿Cómo se mapean estas herramientas en las aplicaciones y casos de uso?



---

# Redes neuronales y aprendizaje profundo

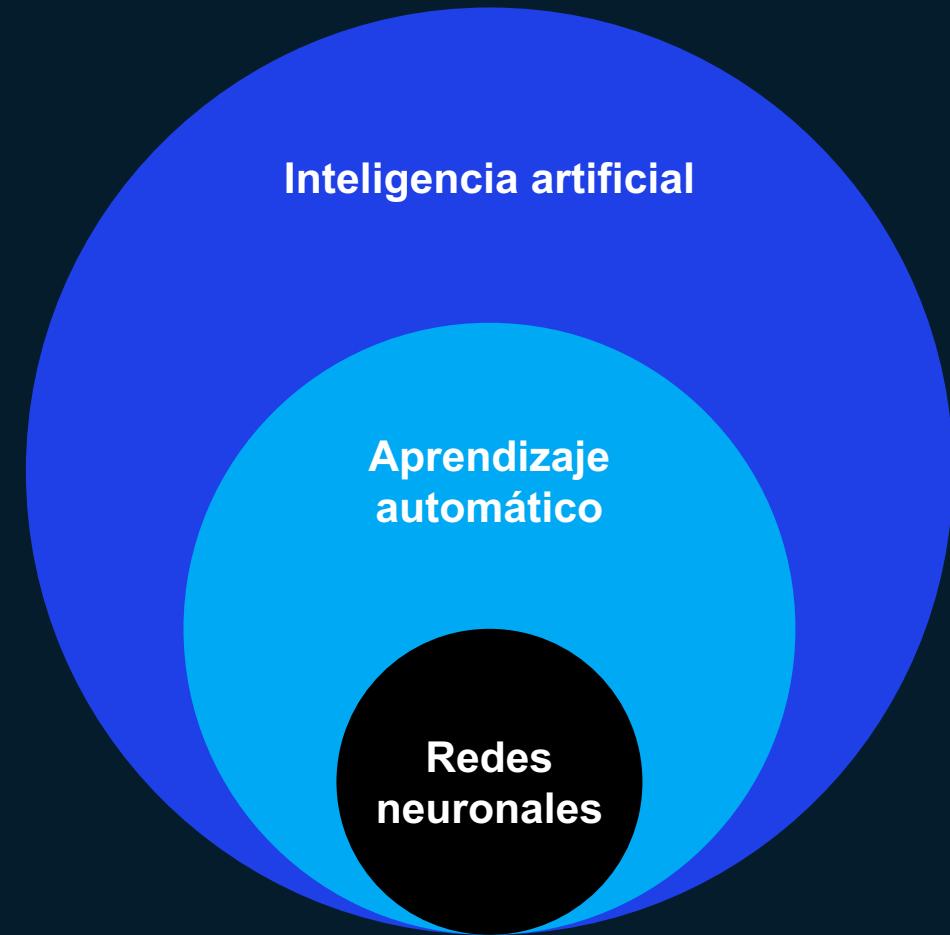
# Redes neuronales y aprendizaje automático

---

Dentro del aprendizaje automático encontramos las redes neuronales artificiales, que son hoy las herramientas más utilizadas en aplicaciones de inteligencia artificial. Se originaron en la década de los 50s **inspiradas por los sistemas biológicos y el funcionamiento del cerebro.**

Se caracterizan por:

- Ser sistemas capaces de aprender e implementar el principio de inducción, es decir, aprenden por experiencia
- Son modelos tipo caja negra, ya que no sabemos exactamente cómo funcionan al interior y por ende pueden ser indeseables para algunas aplicaciones
- Es necesario entrenar con muchos ejemplos para tener buen rendimiento



# ¿Qué es una neurona artificial?

La neurona es el concepto fundamental para el funcionamiento del algoritmo

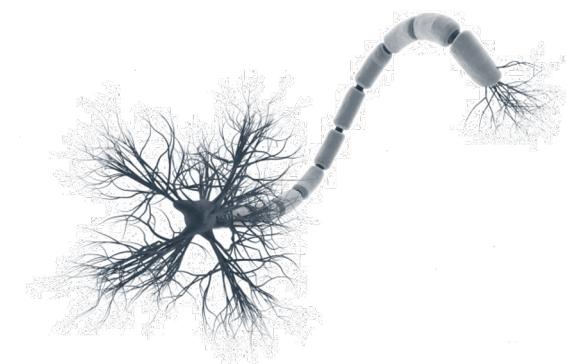
## Concepto

Las redes neuronales artificiales basan su funcionamiento en la replicación del funcionamiento de las neuronas en el cerebro humano.

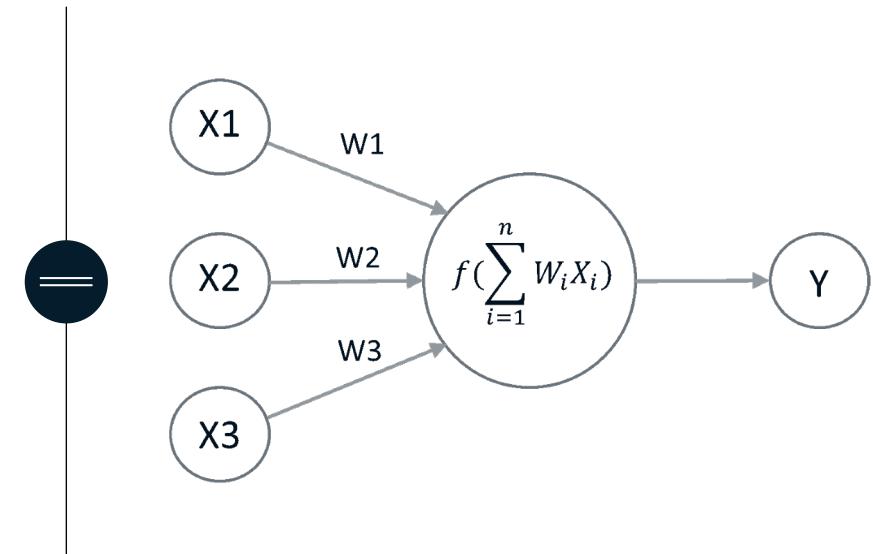
En el cerebro cada neurona recibe información que es procesada y retransmitida a otras neuronas en el proceso de sinapsis.

Una neurona artificial funciona como una función matemática que recibe entradas de información y los procesa de forma no-lineal para producir una salida usada para la siguiente unidad de procesamiento.

## Neurona del cerebro humano



## Pitts y McCullough neurona “lógica”, 1943



En la actualidad no solo la neurona ha sido usada como concepto para inspirar algoritmos de aprendizaje automático, sino que conceptos de **visión humana** también han sido apalancados

# ¿Cómo funcionan las redes neuronales artificiales?

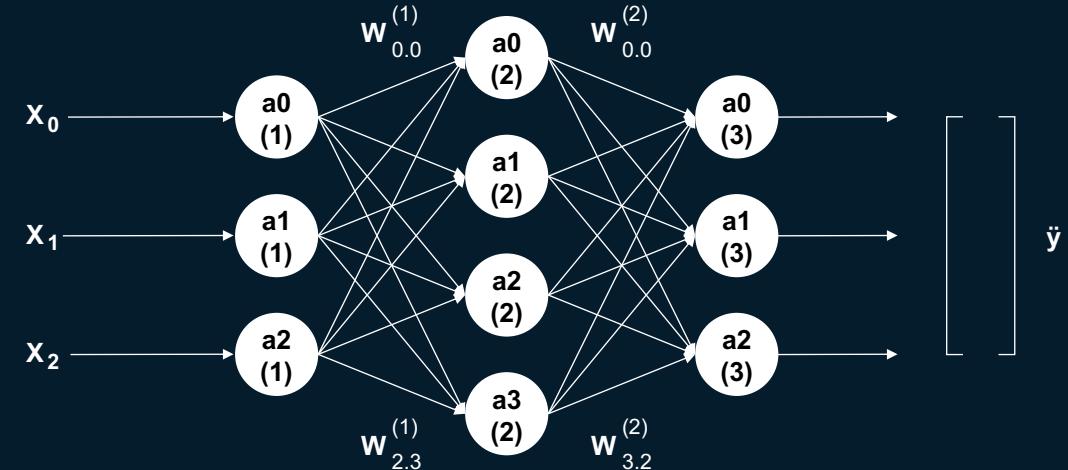
## Arquitectura del algoritmo

- A través de la conexión de neuronas en redes de gran tamaño la información de entrada  $X$  es transformada para predecir la salida  $Y$
- La arquitectura de las redes se da a modo de construcción de capas. Existen 3 tipos de capas:
  - Capa de entrada: neuronas que reciben la información pura de entrada  $X$
  - Capas escondidas: Múltiples capas intermedias que transmiten transformaciones de la información original a las siguientes capas
  - Capa de salida: neuronas finales de la arquitectura que se encarga de ajustar la información transmitida a través del algoritmo para replicar la salida  $Y$

**Perceptrón**  
Frank Rosenblatt  
1956



## Capa de entrada      Capa oculta      Capa de salida



El **perceptrón** fue el primer gran experimento en poner en implementación y replicar la arquitectura propuesta por el algoritmo

# ¿Qué es aprendizaje profundo?

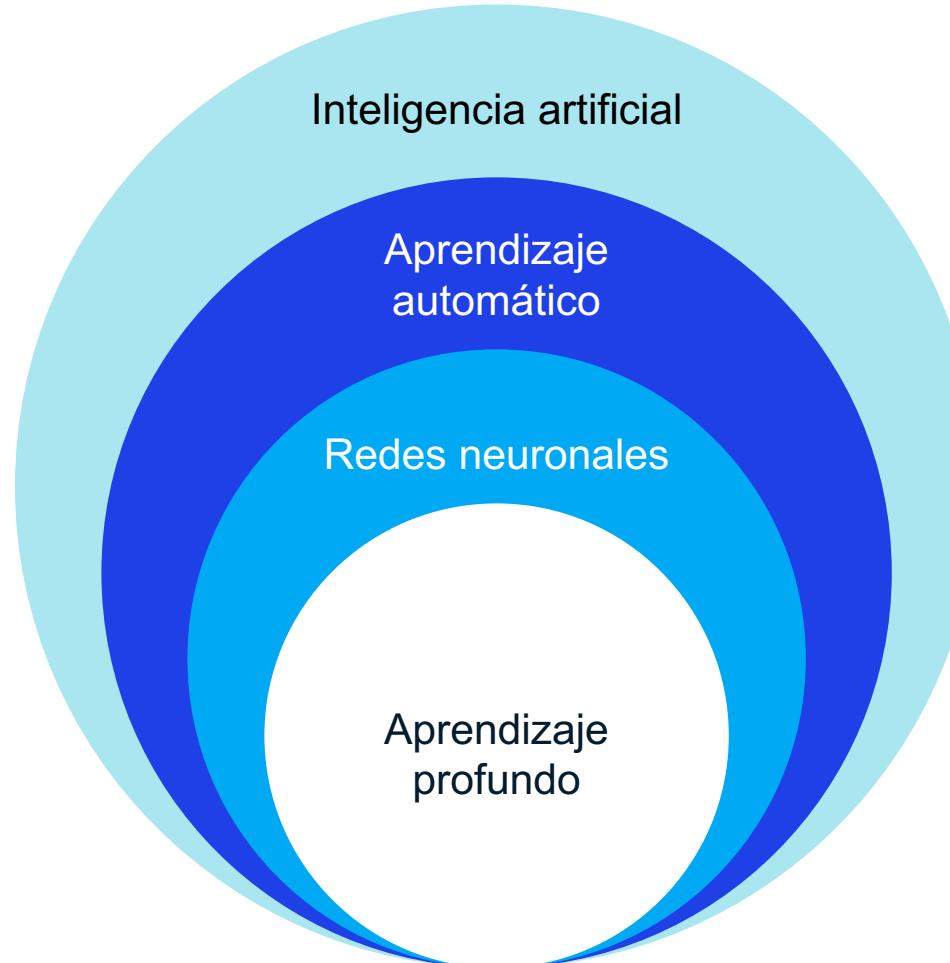
Partiendo de las redes neuronales surge la teoría de aprendizaje profundo

---

El aprendizaje profundo (Deep Learning) es un método de análisis de información basado en el algoritmo de redes neuronales artificiales multicapa o también conocidas como "redes profundas" (de ahí el término "profundas").

Deep Learning se basa en el aprendizaje de conceptos simples, luego combinando estos conceptos para descubrir conceptos más complejos y nociones abstractas.

Por lo tanto, una de las principales ventajas del aprendizaje profundo se debe al aprendizaje de "características" que reduce el trabajo de ingeniería de características



## En la práctica

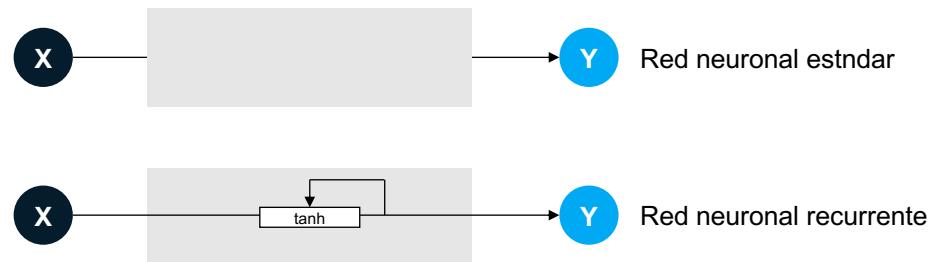
- El aprendizaje profundo es potente pero aún limitado en su aplicación a casos particulares
- Todas las herramientas de ML diferentes al aprendizaje profundo se denominan aprendizaje superficial (shallow learning)
- ~80% de las tareas usuales de un científico de datos recaen en el shallow learning, siendo

# Dos tipos de redes neuronales: recurrentes y convolucionales

Hay dos tipos principales de aprendizaje profundo

## Redes neuronales recurrentes

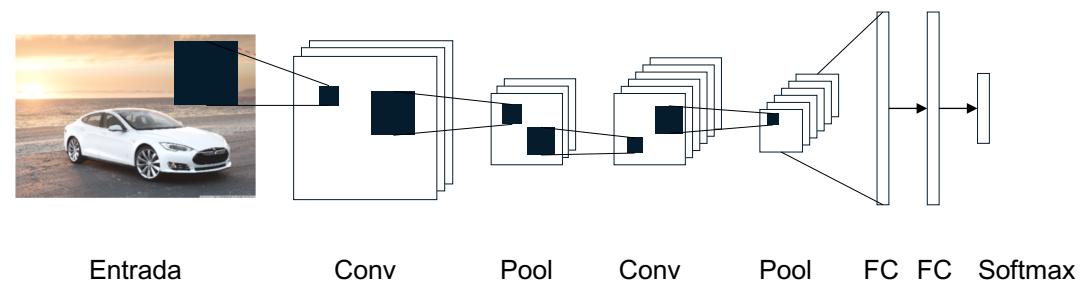
Son arquitecturas que integran bucles de realimentación, permitiendo que la información persista durante un periodo mayor en cada neurona. Este proceso se da a través conexiones entre la salida de la neurona, que se ingresan nuevamente como información de entrada.



Encuentran gran aplicación en procesamiento de texto o series temporales

## Redes neuronales convolucionales

Son redes donde las neuronas representan campos receptivos de una manera muy similar a las neuronas en la corteza visual primaria del cerebro. Su originalidad son los filtros convolutivos inspirados en biología, donde la red analiza la imagen buscando formas o patrones.



En la actualidad se usan para tareas de clasificación de imágenes, análisis de movimiento en tiempo real o autos que se conducen solos.

# Ejemplo de redes convolucionales

## Transferencia de estilo

En las aplicaciones actuales es muy conocido la transferencia de estilo. En este ejercicio, se desea enseñar a un algoritmo a identificar los patrones y características que definen el estilo de un artista particular.

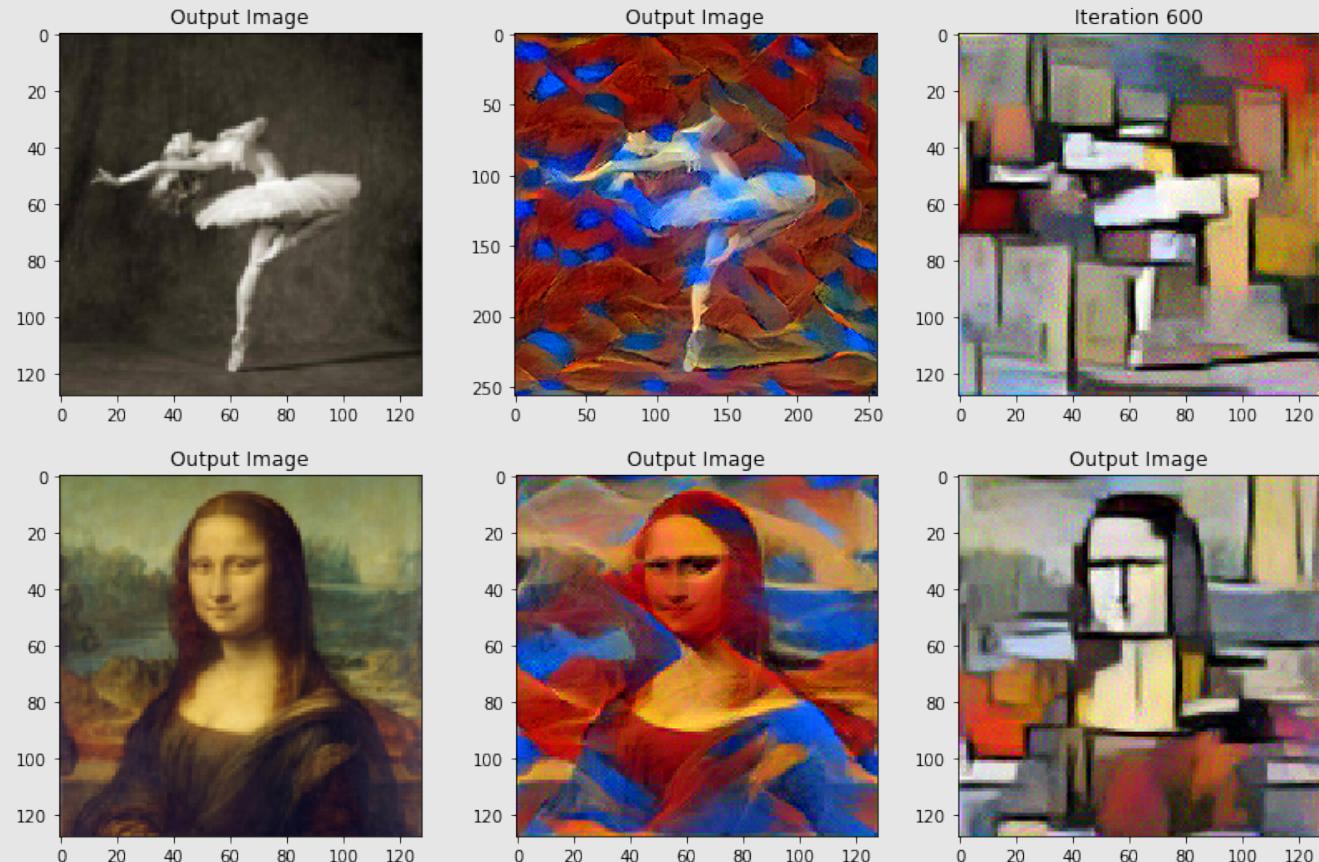
Una vez esto es aprendido por el algoritmo, se puede usar para aplicar transformaciones de imágenes diferentes y aplicar estos patrones estilísticos a la nueva obra.

**Veámoslo en acción !**

Tarea: clonar el repositorio y correr el código

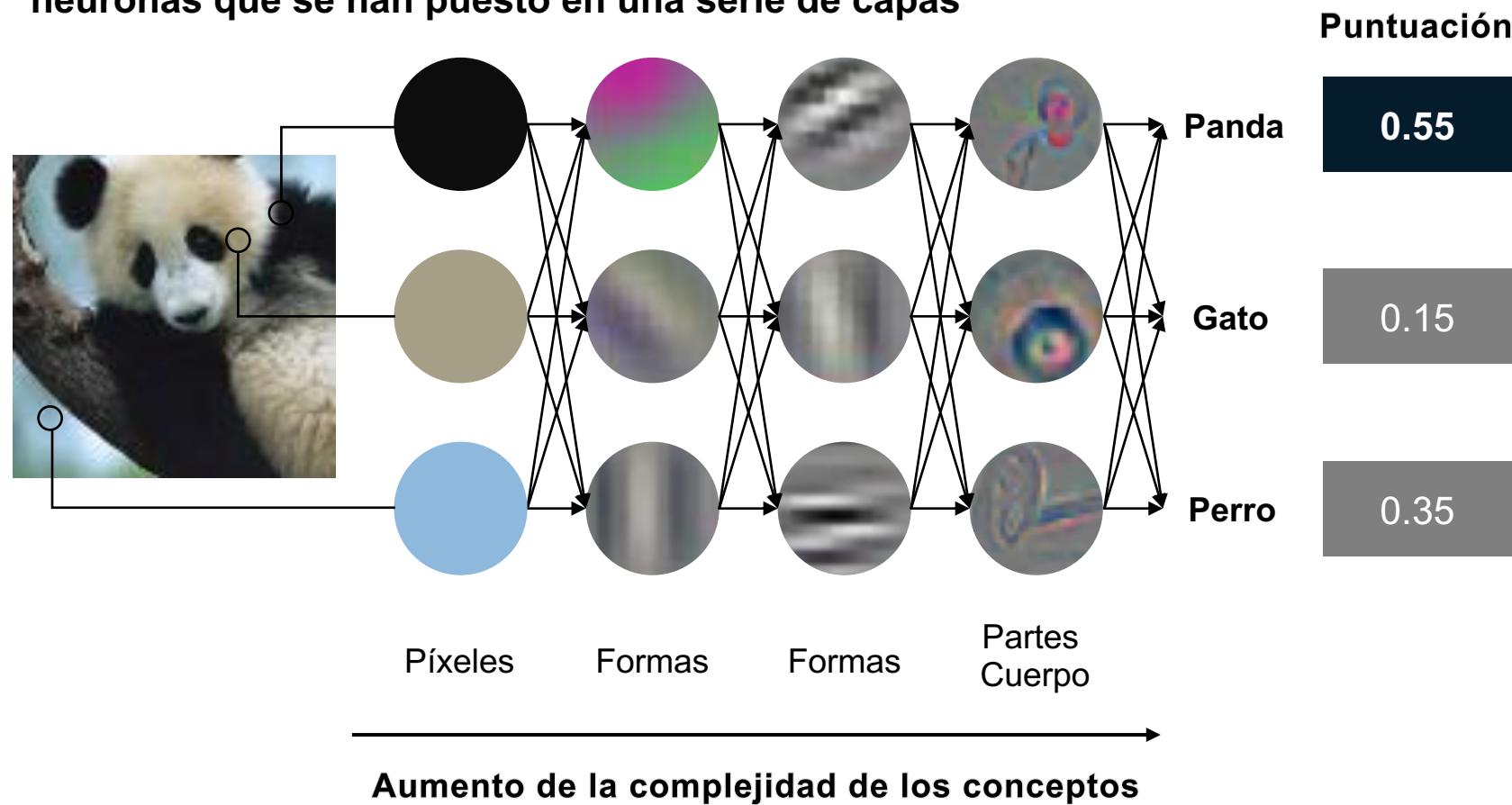
[https://github.com/simontamayo/demo\\_artistic\\_neural\\_networks](https://github.com/simontamayo/demo_artistic_neural_networks)

Y para la próxima session, tener un avatar de zoom artístico!



# ¿Por qué es aprendizaje profundo?

La profundidad de la red se asocia con el número de neuronas que se han puesto en una serie de capas



El aprendizaje profundo o Deep Learning ha adquirido gran popularidad en los últimos años debido a sus increíbles aplicaciones como:

- Autos que se conducen solos
- Detección eficiente y precisa de enfermedades

Sin embargo requieren grandes insumos para funcionar correctamente, por ejemplo:

- Grandes cantidades de información
- Alto poder computacional

---

# Ciencia de datos en la actualidad

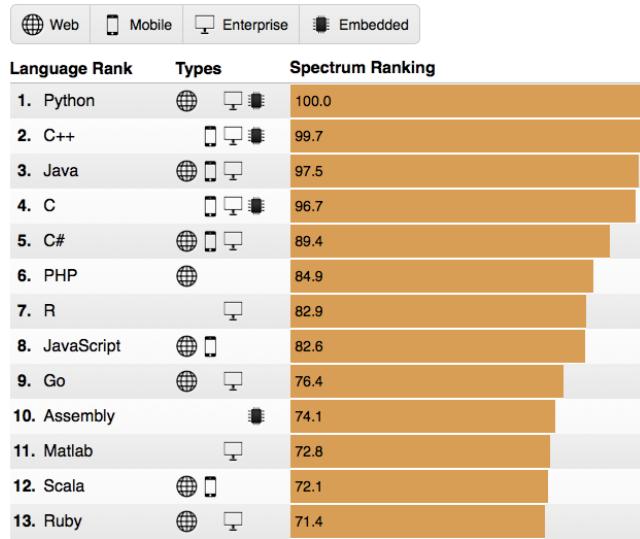
# Ciencia de datos en la actualidad

¿Cómo podemos empezar?

---

1

Identifiquemos la tecnología correcta



Python es en el actualidad el lenguaje más usado en diversas aplicaciones pero especialmente en ciencia de datos

2

Aprendamos sobre la teoría y herramientas necesarias:



3

Finalmente, practiquemos para poner a prueba el conocimiento:



---

# Ejemplos y casos de uso en aprendizaje automático

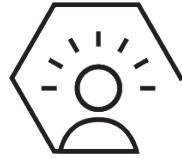
# Un proyecto de analítica avanzada implica una transformación que estructuramos según "5i"

---



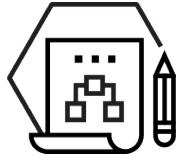
## 1 Ideation

- Mapa de proceso
- Explorar los desafíos y medir la capacidad
- Identificar oportunidades
- Priorizar como cartera



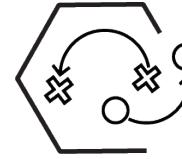
## 2 Intelligence

- Definir meta de rendimiento
- Desarrollar hipótesis detalladas
- Navegar por los requisitos de InfoSec
- Definir los datos y evaluar la calidad
- Validar métodos analíticos
- Cuantificar el valor potencial en juego
- Priorizar oportunidades de alto valor



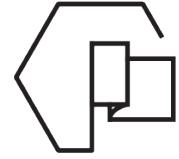
## 3 Inception

- Extraer, limpiar y vincular datos
- Construir modelos explicativos y predictivos; Restringir adecuadamente
- Refinar el valor en juego
- Cuantificar palancas de rendimiento y abordabilidad
- Priorizar las palancas de intervención
- Definir plan de acción (intervenciones)



## 4 Intervention

- Implementar las intervenciones necesarias para captar el rendimiento
- Equipos de coaching sobre la incorporación de intervenciones
- Monitorear el impacto de las intervenciones
- Actualizar plantillas según corresponde

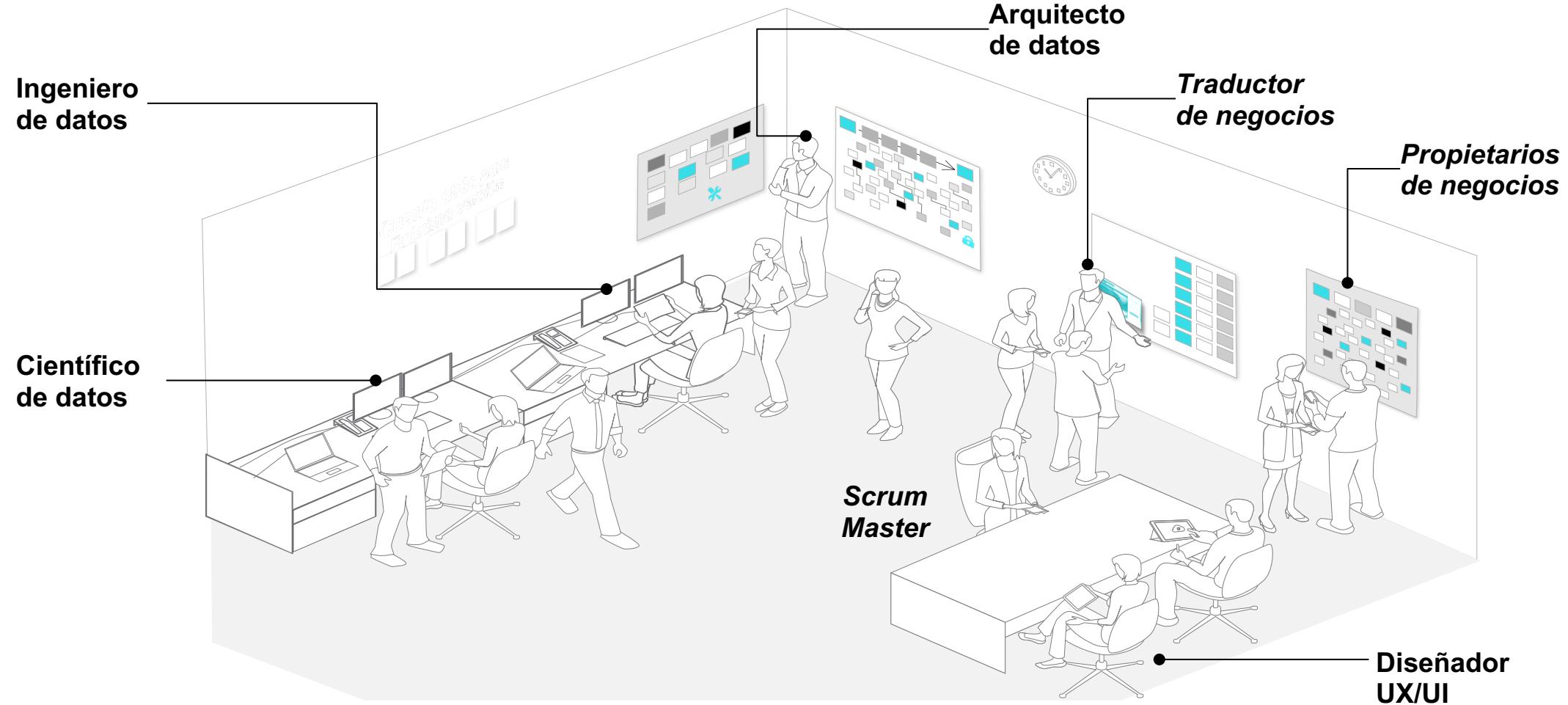


## 5 Independence

- Implementar y transferir los recursos tecnológicos necesarios para mantener mejoras en el rendimiento
- Proporcionar capacitación para desarrollar e incorporar las habilidades necesarias para apoyar nuevas características

# La entrega de un proyecto de Analytics exitoso requiere un esfuerzo conjunto de equipos profesionales con diferentes perfiles

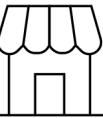
---



# Alicorp ha desarrollado de forma exitosa casos de uso aplicando AA

---

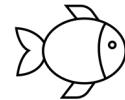
**1**



## Ejecución de puntos de venta

- Enfocado en desarrollar el mejor modelo de segmentación de clientes bodegas
- Se identificaron palancas específicas que podrían utilizarse para crear impacto en cada segmento
- Se diseñó un plan para ejecutar palancas a nivel de cliente

**2**



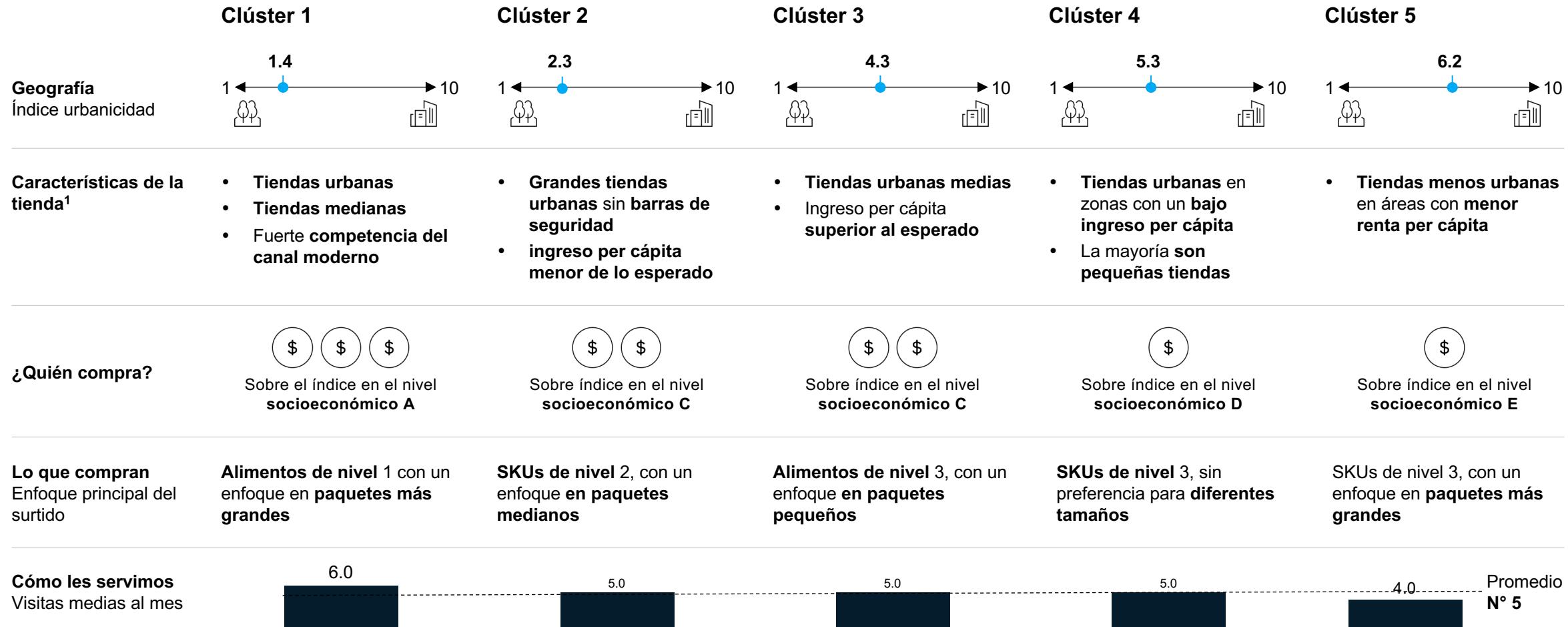
## Modelos predictivos en la industria camaronera

- Se trabajó con Vitapro para desarrollar modelos predictivos utilizando AA
- El objetivo fue transformar la estrategia comercial llevando el modelo de asesoramiento técnico al siguiente nivel

# 1. El caso de uso de la ejecución de PDV se centró en proporcionar un modelo de segmentación de clientes y aprovecharlo para identificar otras palancas para agregar valor

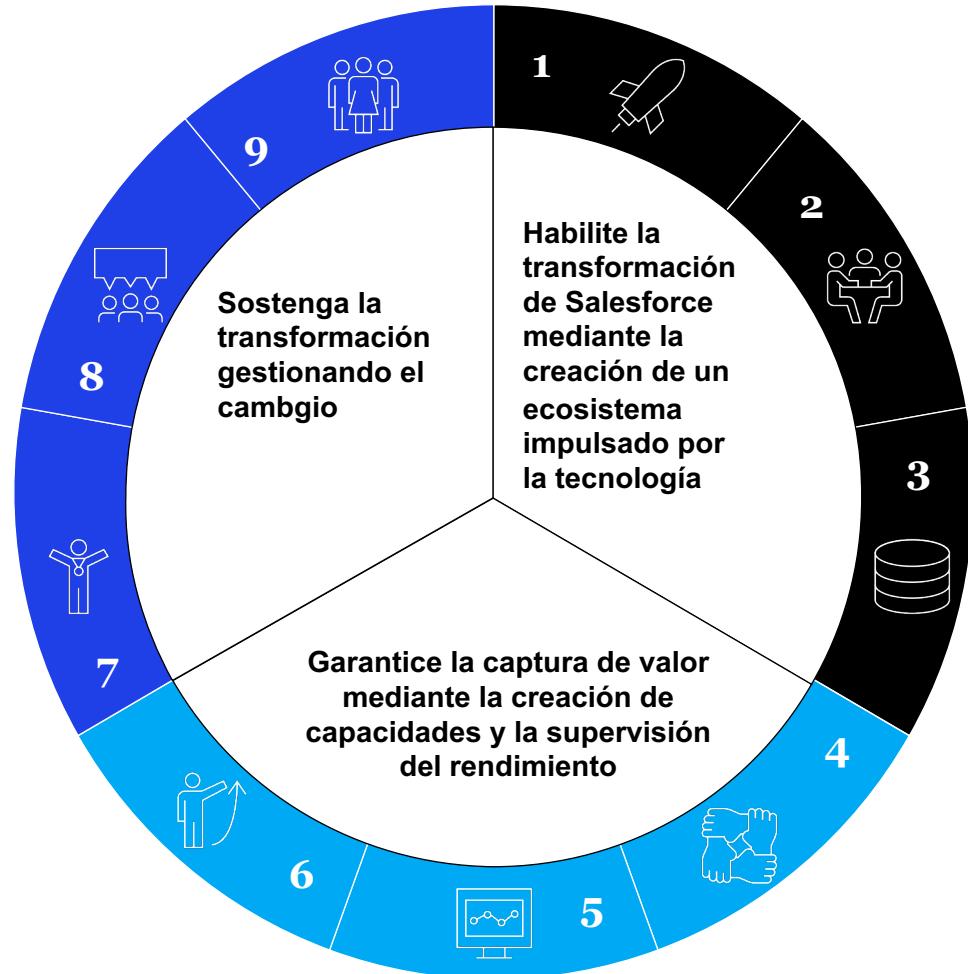
|                         | Definición de segmentos   | Impacto potencial  | Palancas de ejecución   |
|-------------------------|---|--|---|
| Preguntas para resolver | <p>¿Qué nos dice la cartera sobre el consumidor promedio en el PoS?</p> <ul style="list-style-type: none"> <li>• ¿Cómo afecta el entorno sociodemográfico/competitivo a las ventas en cada PoS?</li> <li>• ¿Qué nos dice la dispersión de ventas en un segmento sobre el potencial sin restricciones?</li> </ul>                                      | <p>¿Cuál es el impacto potencial de la ejecución de las palancas?</p>  | <p>¿Qué palancas de ejecución pueden crear valor en cada segmento?</p>  |
| Análisis clave          | <p>Atributos de cartera (precio, productos, promociones, SKUs, etc.)</p> <ul style="list-style-type: none"> <li>• Atributos de entorno (ubicación, tráfico, situación socioeconómica, densidad, población, etc.)</li> <li>• Variables de competencia (tiendas, farmacias, etc.)</li> </ul> <p>Definición de clúster con datos internos y externos</p> | <ul style="list-style-type: none"> <li>• Análisis de la dispersión de ventas entre PoS de un mismo clúster</li> <li>• Definición de impacto potencial por PoS</li> </ul> | <ul style="list-style-type: none"> <li>• Evaluar las palancas de ejecución para cada segmento</li> <li>• Determinar la importancia variable para incrementar el valor en cada segmento</li> </ul> |
| Resultado               | <p>5-6 segmentos PoS con atributos significativamente diferenciados y potencial sin restricciones</p>   | <p>Aumento de ventas potencial identificado por PoS</p>  | <p>Palancas de ejecución identificadas para cada segmento</p>   |

# 1. Se usó modelos de Analítica Avanzada para identificar diferentes segmentos de clientes



1. Grande: más de 6 pasos de puerta a mostrador, medio: entre 3 y 6 pasos de puerta a mostrador, pequeño: menos de 3 pasos de puerta en mostrador

# 1. Para garantizar una adopción exitosa de herramientas aplicamos 9 elementos de gestión del cambio centrados en las transformaciones comerciales de DnA



1. Crear una visión convincente y pilares estratégicos de la transformación, enfocada en impulsar el crecimiento a través de DnA y la excelencia comercial
2. Construir equipos ágiles de científicos de datos, UX/UI, traductores y líderes empresariales para impulsar conocimientos y desarrollar nuevas soluciones digitales como MVP con un enfoque de mejora continua, fomentando la prueba y el aprendizaje
3. Garantizar la infraestructura tecnológica y los datos para desarrollar la solución digital, aprovechando asociaciones y alianzas con proveedores de datos y tecnología de 3a parte
4. Desarrolle una cultura y capacidad operativa de la fuerza de ventas basada en el uso de los análisis para capturar valor, incorporando información en el recorrido de ventas a través de rutinas y scripts e impulsando al equipo de análisis a mejorar continuamente las soluciones
5. Esbozar incentivos, métricas y procesos para permitir la gestión del rendimiento en tiempo real y el cambio de cable dentro de salesforce
6. Diseñar sistemas de incentivos y promociones para los clientes, impulsados por análisis y datos avanzados para maximizar el valor en cada TPV
7. Cambio de modelo de rol desde la parte superior, alineando las formas de trabajo del líder senior
8. Diseñar una historia de cambio convincente y un plan de comunicación, incorporando una nueva identidad cultural canalizada a través de la organización de ventas
9. Inspirar y dar rienda suelta a los agentes de cambio para garantizar que las personas más influyentes impulsen el cambio de abajo hacia arriba y crea la capacidad de escalar el cambio

## 2. Problema de negocio: recomendar la mejor alimentación posible para cada día de un ciclo de crecimiento de camarones

---

### Nuestro objetivo:

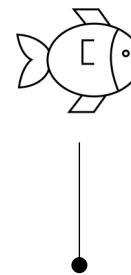
Adaptar las **recomendaciones de alimentación** a cada cliente, a cada **tipo de ciclo** productivo y las **condiciones de cada día**.

Para nuestros clientes, esto se traduce en **mayor crecimiento y menor consumo de alimento**.

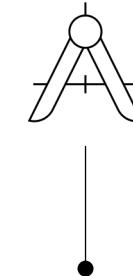
### Enfoque de analítica avanzada

10101  
01010  
10101  
01010

**1**  
**Procesamiento de datos:**  
colecta, limpieza, llenado, y creación de variables



**2**  
**Predictión del peso**  
de los animales, basada en el crecimiento mas probable para el estado del ciclo



**3**  
**Recomendación de alimentación,**  
basada en las mejores prácticas de la industria para el estado del ciclo

## 2. Implementamos una metodología de analítica avanzada basada en 3 grandes etapas

---

### 1 Procesamiento de datos

El procesamiento de datos **transforma los datos de entrada** (históricos de ciclos productivos) **en DataSets de aprendizaje estructurados**, utilizados en el entrenamiento de los modelos.

Para esto creamos un “pipeline” y utilizamos la **estructura de procesamiento de QuantumBlack**:  
Raw → Intermedite → Primary → Feature → Model Input → Model Output.

### 2 Predicción de crecimiento

La predicción de crecimiento se hace con un **modelo de Gradient Boosting**. El modelo de crecimiento **predice semanalmente el peso esperado de los animales en condiciones normales** para un ciclo productivo. Es decir, en la próxima semana, cuál es el peso mas probable que se puede lograr dadas condiciones actuales del ciclo (peso actual, densidad, temperatura, oxígeno, ...).

### 3 Recomendación de alimento

La recomendación de alimentación se hace con un **modelo de K-nearest neighbors**.

El modelo de alimentación **hace recomendaciones de alimentación diaria, replicando las mejores prácticas observadas en la industria** – en condiciones similares- teniendo como criterios el **crecimiento y el factor de conversión de alimento**.

### Toolkit técnico



Python



Scikit Learn



Pandas



Kedro



GCP



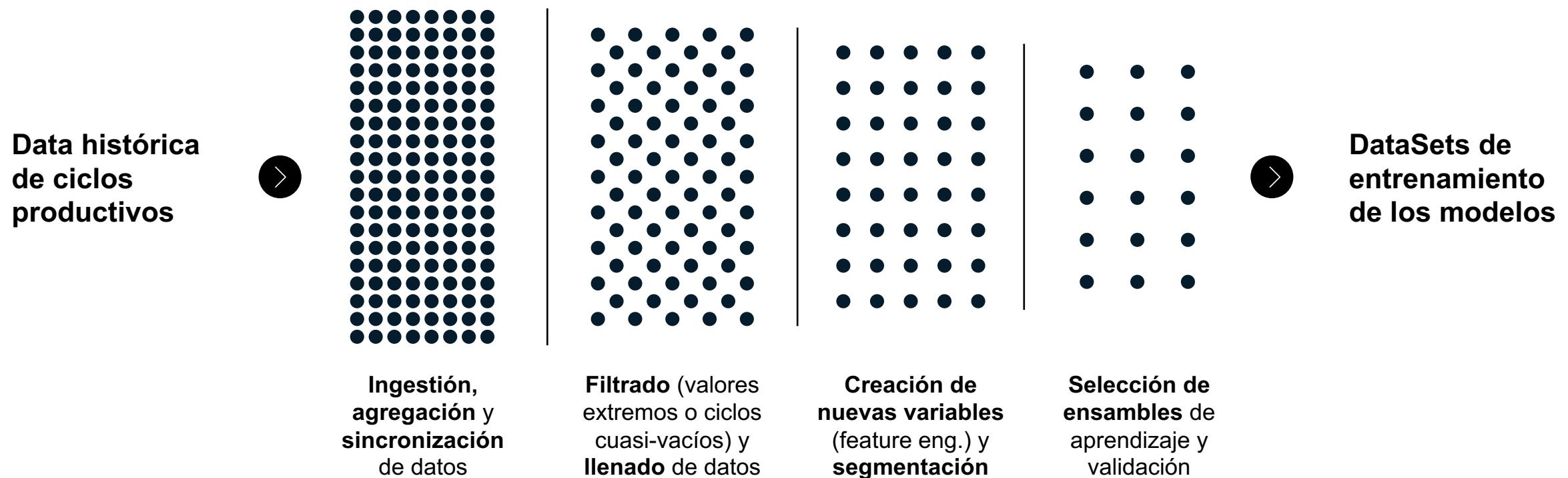
GitHub

## 2.1. Procesamiento de datos

---

Hemos creado un pipeline de datos que permite un **flujo automatizado en la transformación de datos** de entrada hasta el entrenamiento y la utilización de los modelos analíticos.

El pipeline es responsable de la estructura y la integridad de los datos que alimentan los modelos, para esto implementa las funciones que describimos aquí:



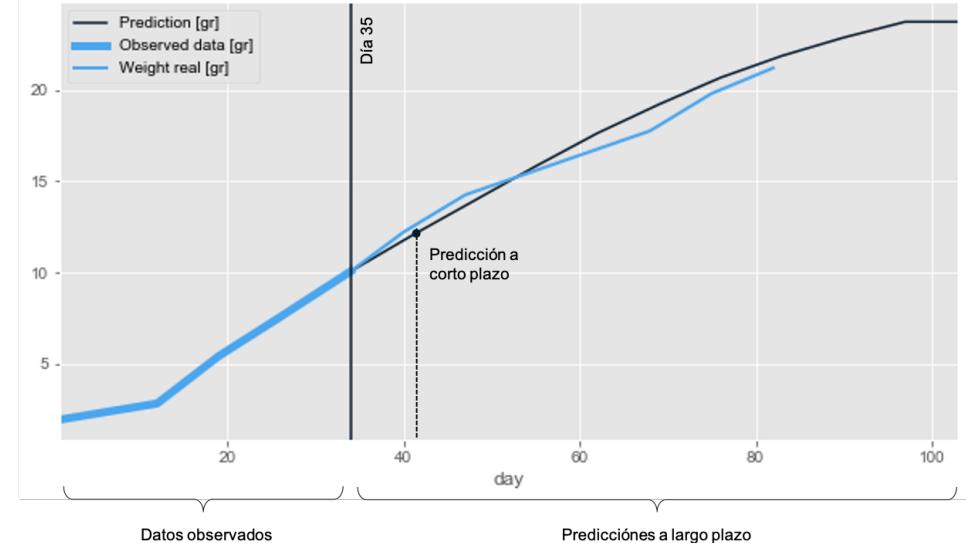
## 2.2. Predicción de crecimiento

### ¿Qué hace el modelo?

El modelo **recibe información sobre el ciclo productivo** en un momento dado (teniendo en cuenta el peso actual de los animales, las variables ambientales y las variables de configuración del ciclo) **y predice el peso más probable a la siguiente semana**.

### ¿Cuáles son sus casos de uso?

- **Predicción semanal:** el modelo debe poder proponer, para un estado dado del sistema productivo, la predicción más aproximada del peso para la próxima semana.
- **Predicción de todo un ciclo:** el modelo debe poder proponer, para un estado dado del sistema productivo al comienzo del ciclo, la curva de peso más precisa hasta el final del ciclo. En este caso el modelo debe obtener predicciones de las variables ambientales.



### ¿Cómo funciona el modelo?

Es un modelo de aprendizaje supervisado en regresión, es decir, que aprende a partir de ejemplos, la función que relaciona unas entradas X con una salida Y, en nuestro caso las entradas son las variables de estado del ciclo, y la salida es el peso de la próxima semana. El modelo de regresión implementado es el **GradientBoostingRegressor**, que se basa en el principio general de GradientBoosting.

Se trata de un modelo de ensamble donde el resultado de predicción es obtenido combinando varios modelos. En nuestro caso, el Gradient-Boosting combina muchos árboles de decisión. El nombre de “Gradient Boosting” (en español “aumento del gradiente”) implica una la noción de optimización de la función de error. La idea clave es que se puede añadir cada nuevo modelo al ensamble buscando que éste tenga mejores resultados de error en las partes donde el ensamble todavía debe mejorar.

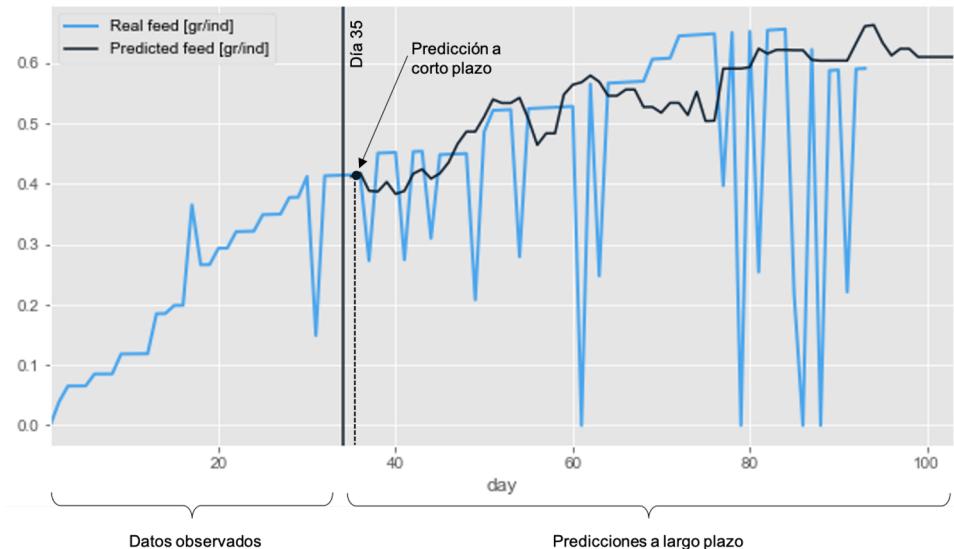
## 2.3. Recomendación de alimento

### ¿Qué hace el modelo?

El modelo recibe información sobre el ciclo productivo en un momento dado y **propone la recomendación de alimento más adecuada para cada día**. Concretamente el modelo **replica las “mejores” dosis de alimento que existen en la data de aprendizaje** para situaciones similares. Estas “mejores prácticas” resultan de un **muestreo estratificado**.

### ¿Cuáles son sus casos de uso?

- **Predicción diaria:** El modelo debe poder proponer, para un estado dado del sistema productivo, la alimentación más adecuada del próximo día.
- **Predicción de todo un ciclo:** El modelo debe poder proponer, para una predicción de peso, la alimentación correspondiente durante todo el ciclo.



### ¿Cómo funciona el modelo?

Es un modelo de aprendizaje supervisado en regresión, donde las entradas son las variables de estado del ciclo, y la salida es la dosis de alimento en gramos por individuo. En este caso implementamos un modelo de **K-nearest neighbors KNN** ó “método de los K vecinos más cercanos”.

Este algoritmo funciona realizando una predicción para una nueva observación en función de sus vecinos. Es decir, las observaciones en la data de entrenamiento más cercanas a la nueva observación. Para esto es necesario disponer de una medida de “distancia” entre la nueva observación y todos los datos disponibles. Técnicamente, el algoritmo predice un elemento desconocido al observar los k elementos vecinos más cercanos. En el caso de una regresión, la predicción puede ser el promedio de los k vecinos, o un promedio ponderado en función de la distancia.

## 2. ¿Cómo llegamos a estos resultados?

---

¿Cómo **identificar el mejor modelo** y sus hiper-parámetros?

Para elegir el mejor modelo y el conjunto de hiper-parámetros óptimos para el problema de ML evaluamos muchos modelos y muchísimas combinaciones de hiper-parámetros, utilizando un **algoritmo de exploración exhaustiva “Grid Search”**. Es un cálculo computacionalmente costoso, pero que nos permitió llegar a la mejor configuración para cada modelo predictivo.

¿Cómo **escoger las variables de entrada** más relevantes?

Inicialmente validamos hipótesis con expertos del negocio. Luego evaluamos la contribución de las variables a los modelos utilizando 3 métodos: (1) Selección de los **K-mejores features** en función de **p-values y f-scores**; (2) Selección en función de las “**importancias**” en un **algoritmo de ensamble**; y (3) **Eliminación recursiva de features con validación cruzada**, utilizando el algoritmo “**RFECV**”.

¿Cómo **entrenar los modelos** evitando el sobre-ajuste y el data leakage?

Utilizamos **validación cruzada en k-folds**, que consiste en entrenar y evaluar los modelos sobre diferentes particiones de datos. En nuestro caso utilizamos grupos personalizados por cliente. Este enfoque (1) **evita el leakage**, pues ciclos productivos asociados a un mismo cliente poseen estructuras similares; y (2) **evalúa la capacidad de generalización** del modelo (limitando el sobre ajuste), ya que replica el escenario (de producción).

## 2. Conclusiones

---

### 1 Podemos aplicar analítica avanzada en acuicultura

Aun con pocos datos (relativamente) y sin necesidad de grandes infraestructuras, podemos crear modelos sofisticados que aportan valor. En este caso implementamos algoritmos de vanguardia -como Gradient boosting- en un sector que se está digitalizando rápidamente.

### 2 Podemos realizar predicciones dinámicas y multi-variadas

Las herramientas que implementamos permiten integrar los cambios en muchas variables endógenas y exógenas para predecir el crecimiento y la alimentación. Esto permite reaccionar ante cambios en el medio ambiente y asegurar un control del rendimiento productivo.

### 3 Podemos identificar y replicar las mejores prácticas de la industria

A través de una metodología de selección de mejores prácticas y un modelo de regresión, identificamos y replicamos la mejor acción para cada estado del ciclo productivo. El enfoque de regresión -y no de optimización- nos permite replicar acciones óptimas aún con pocos datos.

### 4 Los modelos seguirán aprendiendo y se adaptarán

El principio de base del ML es que los algoritmos aprenden a medida que son expuestos a nuevos datos, en este proyecto, el pipeline de procesamiento seguirá entrenando los modelos cada vez que llegan nuevos datos, adaptándolos a las nuevas prácticas productivas.

### 5 Podemos extrapolar a nuevos clientes y a nuevos escenarios

Aunque nuestra base de datos proviene de un grupo reducido de clientes, tenemos una muestra importante de estrategias productivas, esto, acompañado de la capacidad de generalización de los modelos, hace que podamos extrapolar su utilización a nuevos clientes y escenarios.

# Creación de contenido y referencias

---

## Autores



**Iván Torroledo**

Data science fellow

Ivan\_Torroledo@mckinsey.com

Bogotá



**Simon Tamayo**

Data science specialist

Simon\_Tamayo@mckinsey.com

Bogotá

## Referencias

El contenido ha sido desarrollado gracias a la contribución de los autores y las siguientes fuentes externas:

- Simon Tamayo Ph.D, Introducción a la inteligencia artificial y el aprendizaje automático, MINES ParisTech Robotics Center – PSL, 2018
- Alain Schnapper, Simon Tamayo. Aprendizaje automático y cadena de suministro: ¿revolución o efecto de moda? (Economía y Gestión), 2019
- Simon Tamayo, Ronan Cardin, Sebastián Echeverri, Daniel Pino. Investigación de operaciones aplicada a la gestión industrial: Aprendizaje por ejemplo basado en el uso de software de hoja de cálculo, 2018