

## Modulo 1

---

# Introducción a Machine Learning

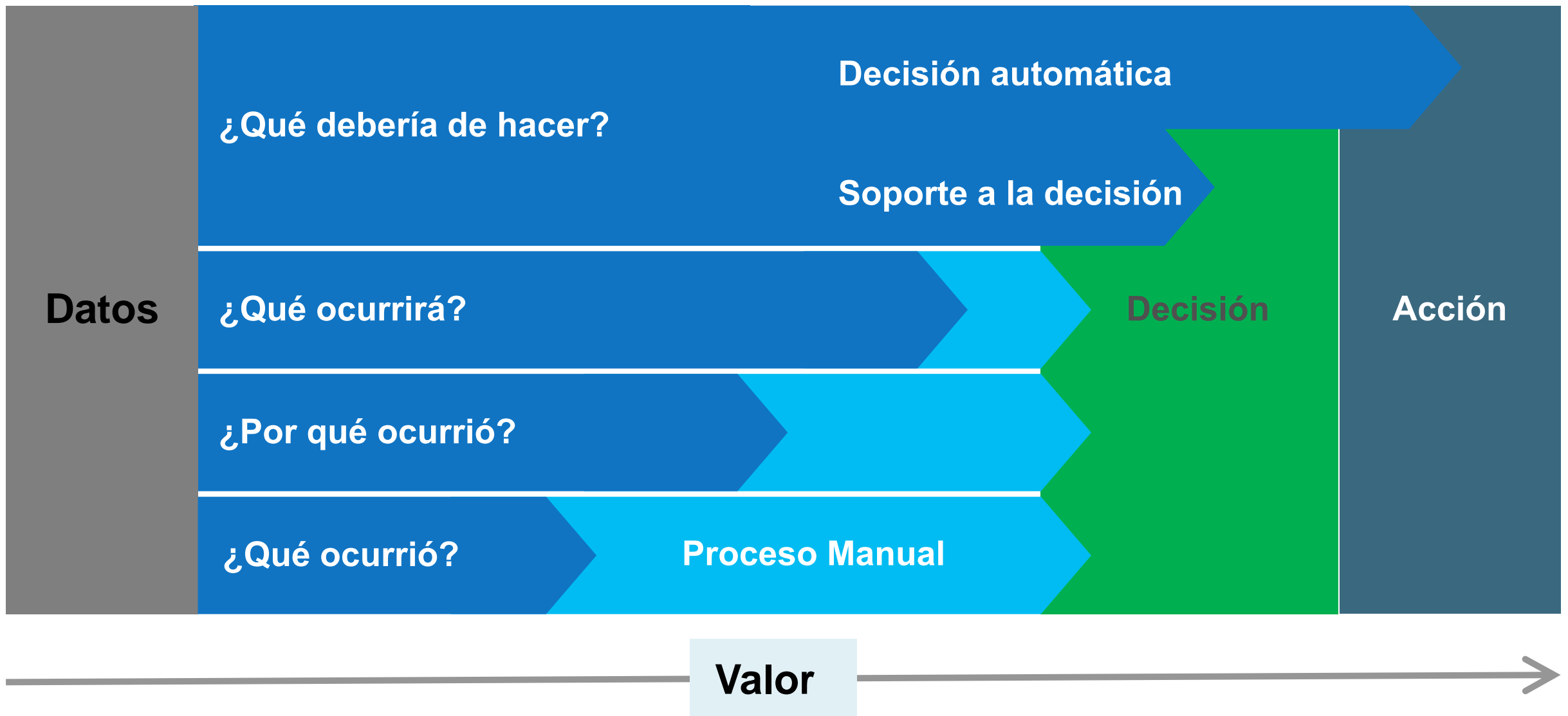
# Agenda

- ¿Qué es Machine Learning?
- Tipos de Problemas que resuelve
- Ciclo de Vida de un proyecto ML
- Notebooks
- La Jerga de Machine Learning

# ¿Qué es Machine Learning?

- Proceso de utilizar datos para tomar decisiones
- Determinar que tipo de algoritmos utilizar basándose en los datos
  - ¿tenemos un conjunto finito de respuestas?
  - ¿Estamos buscando respuestas de algún tipo basadas en ciertos datos?
- Debemos de definir claramente el problema
  - Debe de responder a una pregunta específica

# Datos -> decisiones -> acciones

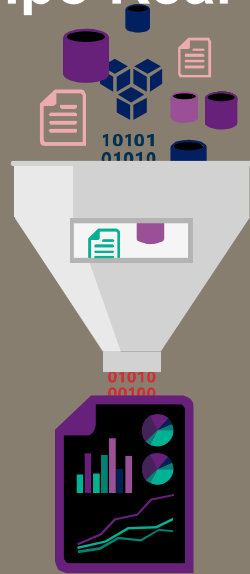


# ¿Qué tipos de Analítica de datos?

## Analítica retrospectiva



## Analítica en Tiempo Real



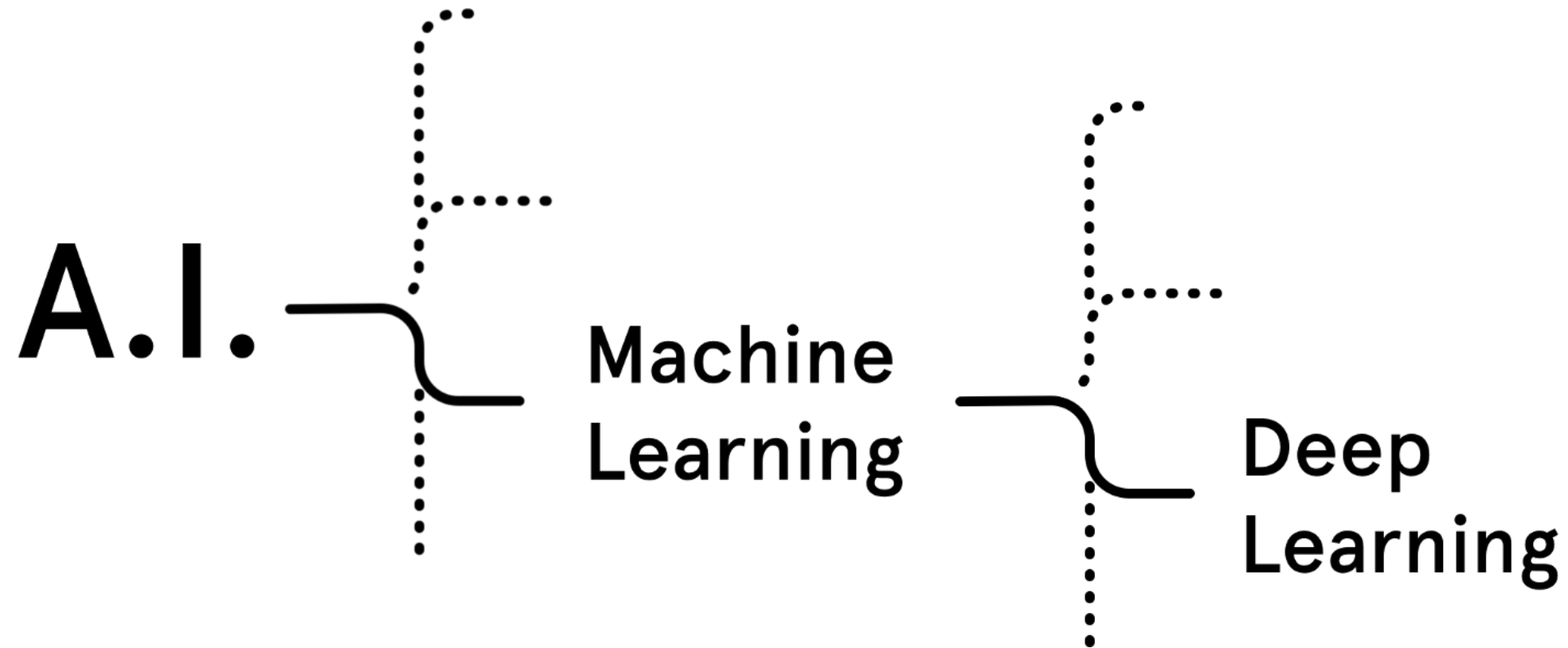
## Analítica Prescriptiva



## Aplicaciones SaaS Inteligentes



# ¿Pero entonces que es la Inteligencia Artificial?



# Inteligencia Artificial



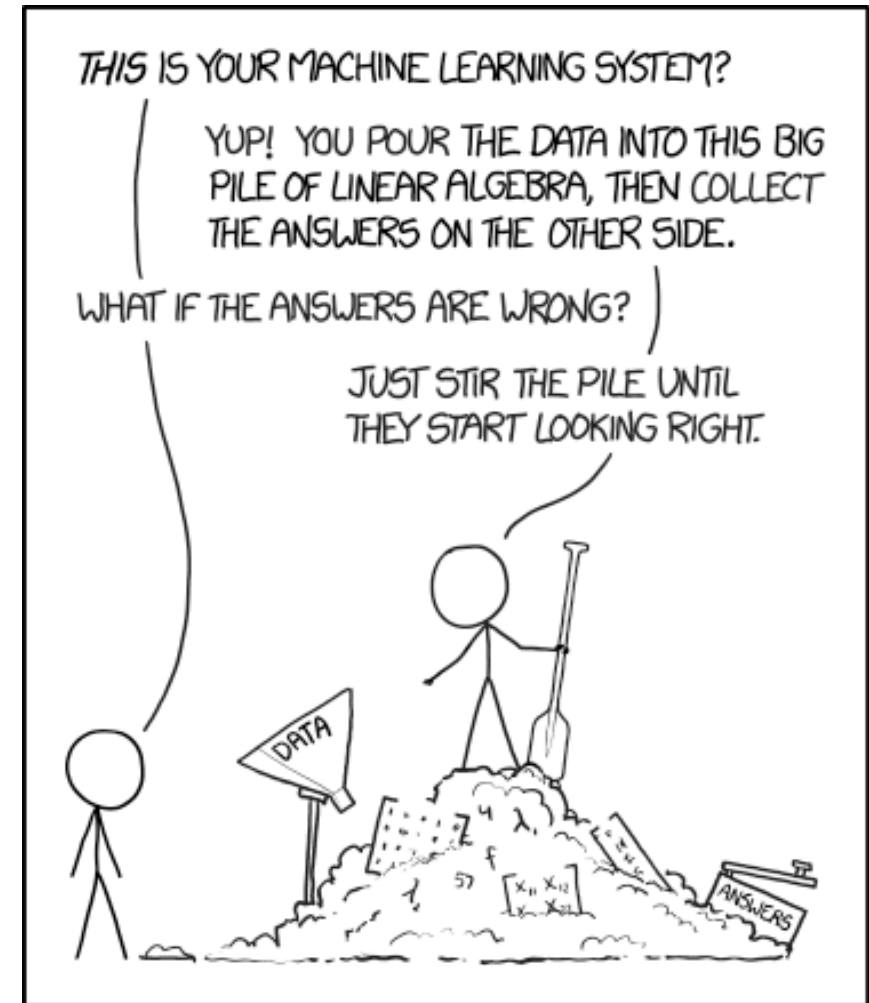
# El sesgo de la IA





# La importancia del ML interpretable

- ¿Cómo se que puedo confiar en el modelo?
- ¿Cómo toma sus decisiones?
- Balance entre rendimiento e interpretabilidad
- Interpretaciones
  - Global ¿Cómo hace las predicciones? ¿Cómo influyen los subconjuntos de datos?
  - Local ¿Por qué ha tomado una decisión en concreto para un caso en concreto?



# Agenda

- ¿Qué es Machine Learning?
- **Tipos de Problemas que resuelve**
- Ciclo de Vida de un proyecto ML
- Notebooks
- La jerga de Machine Learning

# Tipos de problema que puede resolver

- Identificar un código postal escrito a mano en un sobre
- Determinar si un tumor es benigno basado en imagen médica
- Detectar actividad fraudulenta en transacciones de tarjeta de crédito
- Identificar temas de un conjunto de entradas de blog
- Segmentar clientes in grupos con preferencias similares
- Detectar patrones de acceso anómalos a un sitio web

# Tipos de Machine Learning

- Aprendizaje Supervisado tiene un conjunto definido de entradas y salidas
  - Datos etiquetados
  - Feedback directo
  - Predice salida / futuro
- Aprendizaje No Supervisado tiene entradas pero las salidas son desconocidas
  - No tenemos etiquetas
  - No feedback
  - Busca estructuras ocultas en los datos
- Aprendizaje Reforzado
  - Proceso de decisión
  - Sistema de recompensa
  - Aprende una serie de acciones

# Que tipo de análisis necesitamos

- ¿Va una persona a dejar de pagar un préstamo?

Supervisado. Solo dos respuestas, si o no

- ¿Cuáles son las características de los clientes que gastarán más dinero?

No supervisado – varias respuestas posibles

# ¿Existe un conjunto de datos relevante?

- ¿Qué pregunta estoy intentando resolver?
- ¿Cuál es el mejor modo de expresar la pregunta como un problema de Machine Learning?
  - ¿Podemos adivinar si un cliente nos va a abandonar?
- ¿Disponemos de datos detallados o tenemos resúmenes?
- ¿Cómo mediremos el éxito?
- ¿Cómo interactuará con otros componentes?

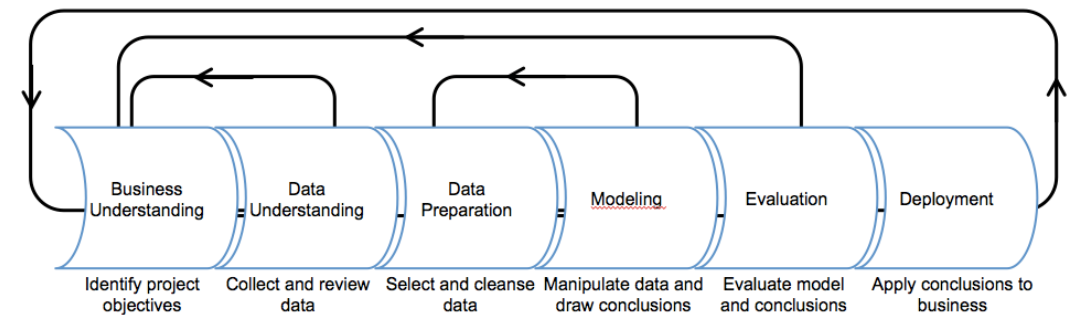
# Agenda

- ¿Qué es Machine Learning?
- Tipos de Problemas que resuelve
- **Ciclo de Vida de un proyecto ML**
- Notebooks
- La jerga de Machine Learning

# CRISP

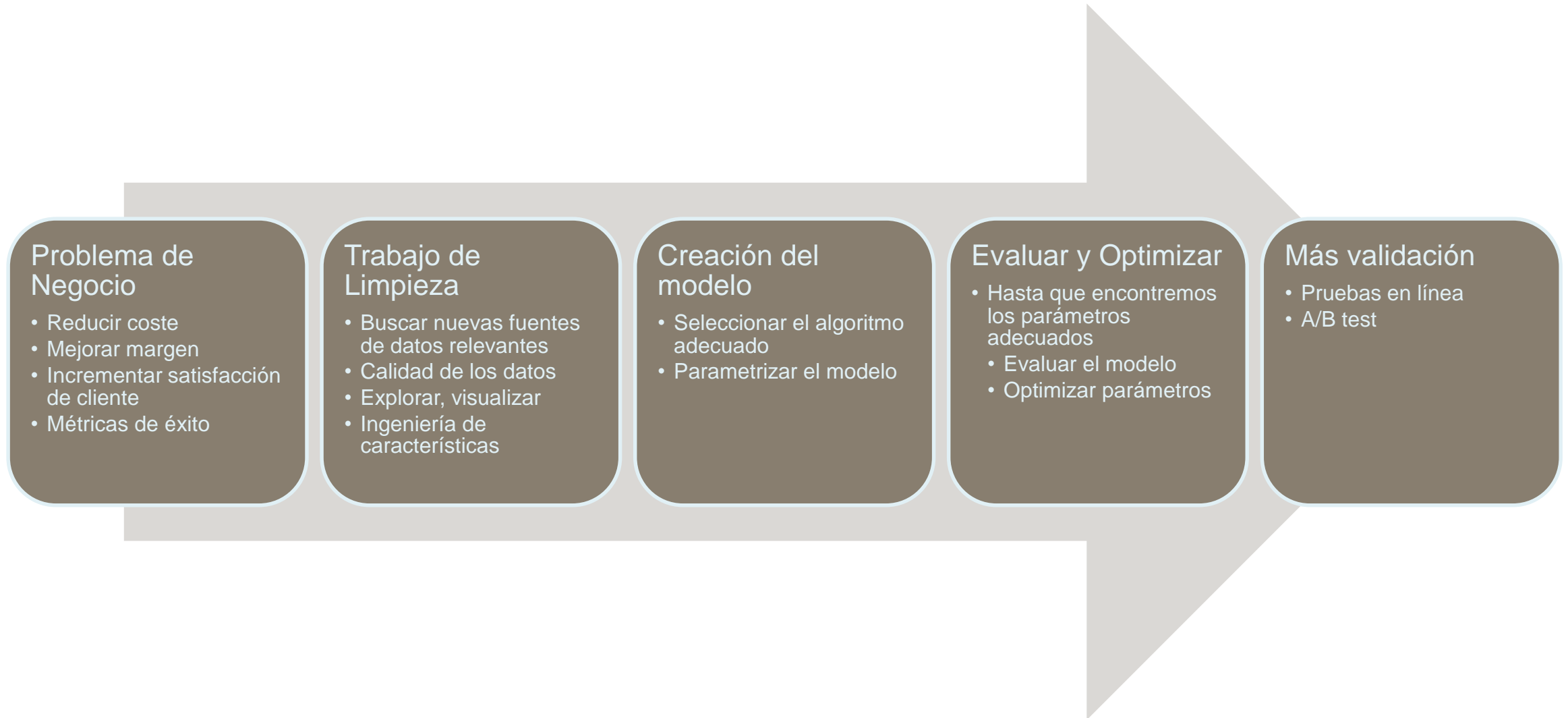


## Cross Industry Standard Process

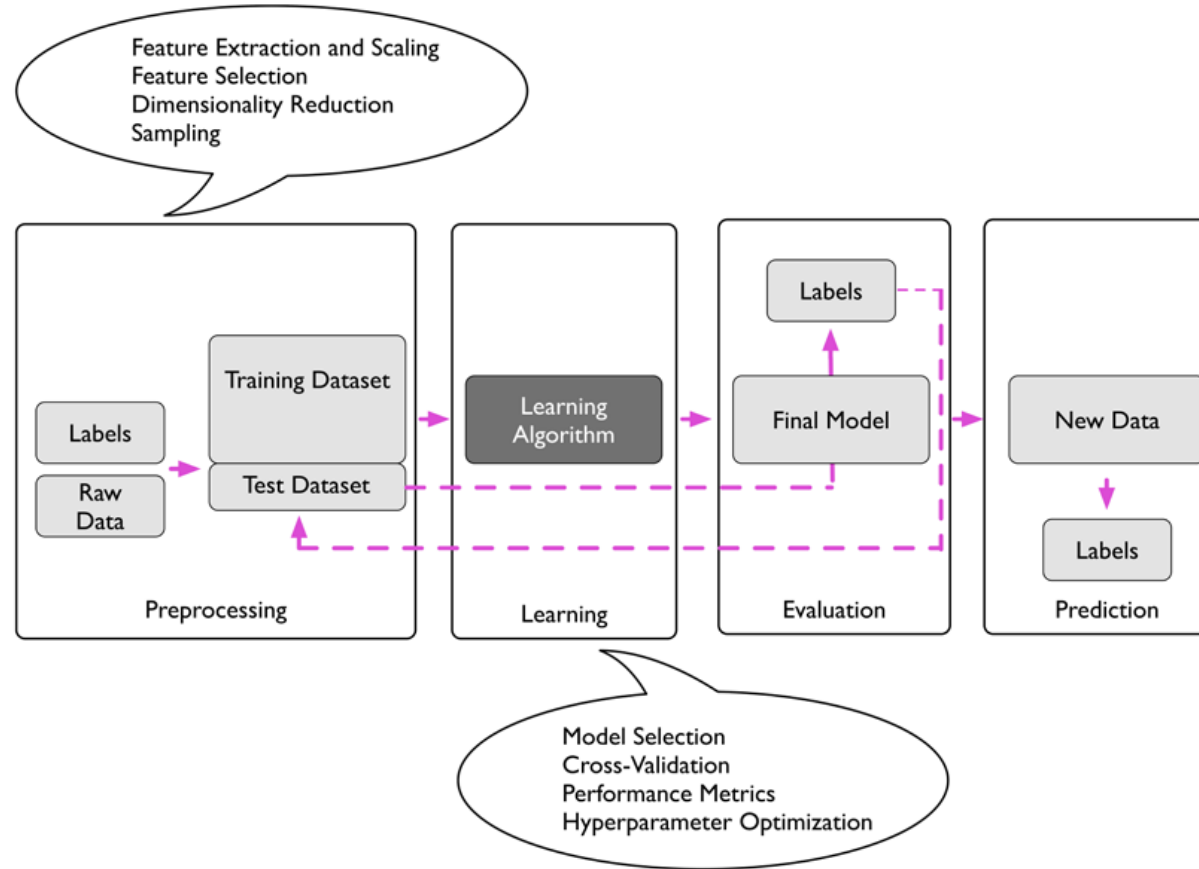




# Ciclo de vida de un modelo de ML



# Flujo de trabajo para un modelo predictivo



# Agenda

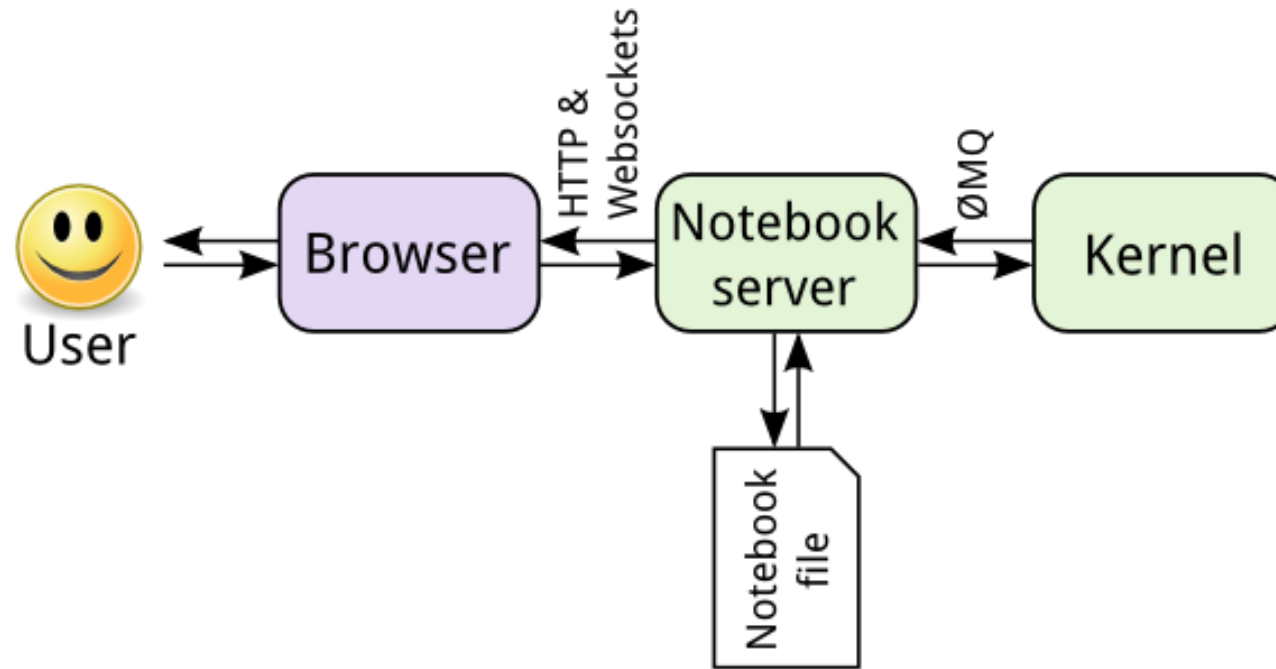
- ¿Qué es Machine Learning?
- Tipos de Problemas que resuelve
- Ciclo de Vida de un proyecto ML
- **Notebooks**
- La jerga de Machine Learning

# Notebooks

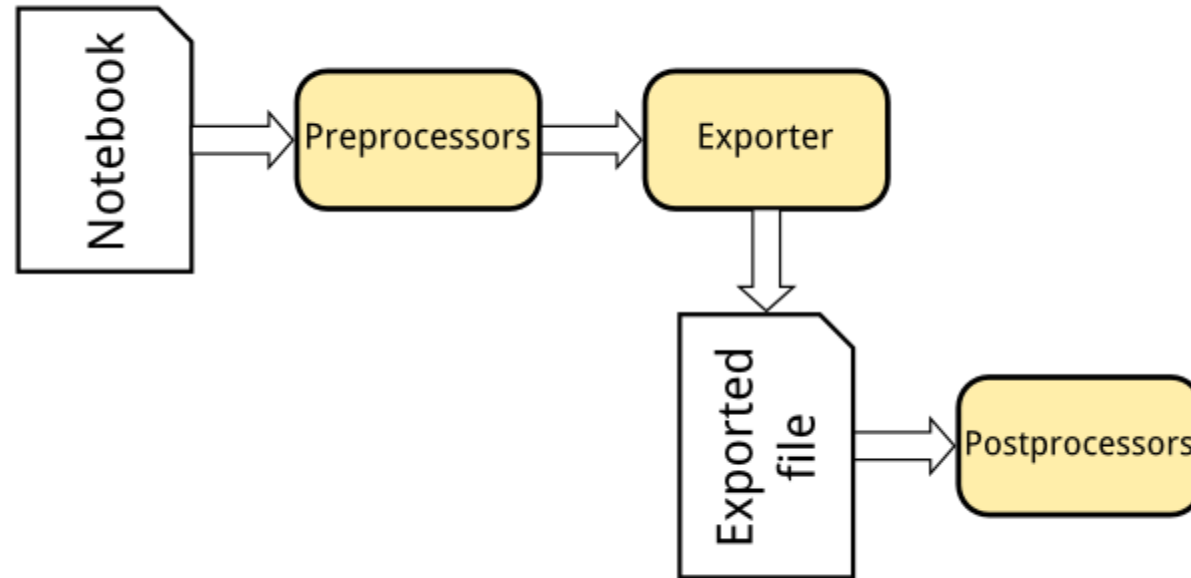


- REPL – Read-Evaluate-Print-Loop
- Prototipo, desarrollo rápido, exploración,...
- Colaboración en equipos

# Arquitectura



# Exportación




# Azure Notebooks

Microsoft Azure Notebooks Preview

Solidq ▾

Libraries What's New Status Help

SolidQ

Solidq > Libraries > SolidQ


▶ Run + New ⚙ Settings 🔗 Share 📄 Clone 0 Clones ☆ Star (0) 📄 Terminal ■ Shutdown 🖨 Preview 📄 Edit File ⬇ Download 🗑 Delete

🔍

FILE NAME ▾	FILE TYPE	MODIFIED
📄 Motor Recomendacion R.ipynb	Notebook	Jul 4, 2018
📄 README.md	Markdown	Mar 29, 2018
📄 Series de tiempo con Pandas.ipynb	Notebook	Mar 31, 2018
📄 u.data	DATA	Jul 4, 2018

Showing 4 search results (1 hidden) < 1 >

# Google Colaboratory



Hola, Colaboratory

Archivo Editar Ver Insertar Entorno de ejecución Herramientas Ayuda

+ CÓDIGO

+ TEXTO

↑ CELDA

↓ CELDA

COPIAR EN DRIVE

COMPARTIR 

Iniciar sesión

CONECTAR EDICIÓN

Índice

Fragmentos de código

Archivos

Te damos la bienvenida a Colaboratory

Introducción

Funciones destacadas


Ejecución de TensorFlow

GitHub

Visualización

Compatibilidad con tiempos de ejecución locales

SECCIÓN



Te damos la bienvenida a Colaboratory

Colaboratory es un entorno gratuito de Jupyter Notebook que no requiere configuración y que se ejecuta completamente en la nube. Puedes consultar más información en la sección de [preguntas frecuentes](#).

Introducción

- [Descripción general de Colaboratory](#)
- [Cargar y guardar datos: archivos locales, Drive, Hojas de cálculo y Google Cloud Storage](#)
- [Importar bibliotecas e instalar dependencias](#)
- [Usar Google Cloud BigQuery](#)
- [Formularios, Gráficos, Markdown y Widgets](#)
- [TensorFlow con GPU](#)
- [Curso intensivo de aprendizaje automático: Introducción a Pandas y Primeros pasos con TensorFlow](#)

Funciones destacadas

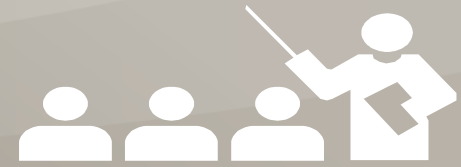
Ejecución de TensorFlow

Colaboratory permite ejecutar código de TensorFlow en el navegador con un solo clic. En el siguiente ejemplo se añaden dos matrices.

$$\begin{bmatrix} 1. & 1. & 1. \\ 1. & 1. & 1. \end{bmatrix} + \begin{bmatrix} 1. & 2. & 3. \\ 4. & 5. & 6. \end{bmatrix} = \begin{bmatrix} 2. & 3. & 4. \\ 5. & 6. & 7. \end{bmatrix}$$



# Demo 01 A Notebooks



Creación de Notebooks

Ejecución de comandos

# Agenda

- ¿Qué es Machine Learning?
- Tipos de Problemas que resuelve
- Ciclo de Vida de un proyecto ML
- Notebooks
- **La jerga de Machine Learning**

# La jerga de ML

- Característica (Feature)
  - Variable, columna, atributo, campo...
- Instancia (Instance)
  - Fila, observación, punto de datos, valor, caso,...
- Objetivo (Target)
  - Variable dependiente, predictant
- Datos (Data)
  - Conjunto de variables predictores, predictor

# La jerga de ML(II)

- Algoritmo
  - Aprender una función objetivo ( $f$ ) que mejor mapee las variables de entrada( $x$ ) a una variable de salida( $y$ )  $y=f(x)$
- Modelo
  - Datos + Algoritmo = Algoritmo entrenado y listo para generar salidas

# La jerga de ML (III)

- Conjuntos de datos
  - Entrenamiento
    - Para entrenar el modelo
    - 2/3
  - Prueba
    - Para probar como funciona el modelo
    - 1/3



[www.solidq.com](http://www.solidq.com)

[info@solidq.com](mailto:info@solidq.com)