

# MACHINE LEARNING

MTRO. ALFONSO GREGORIO RIVERO DUARTE

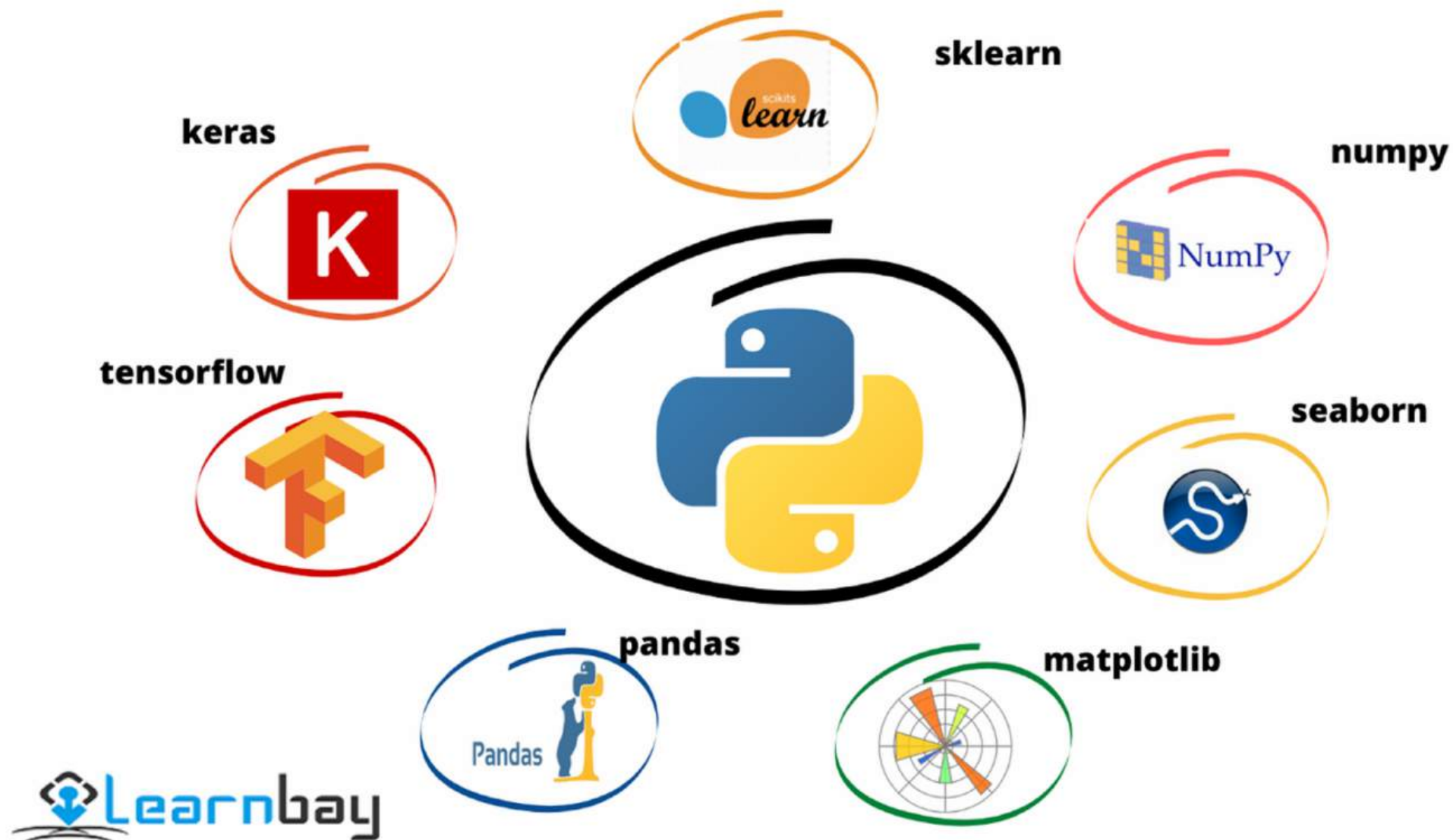


# HERRAMIENTAS DE PROGRAMACIÓN



# HERRAMIENTAS DE PROGRAMACIÓN

La programación es la principal herramienta que podemos utilizar, y para ello el uso de programas enfocados a machine learning facilita su programación. existen numerosos algoritmos que pueden aprender basado en datos, y son fundamentales los siguientes pasos:

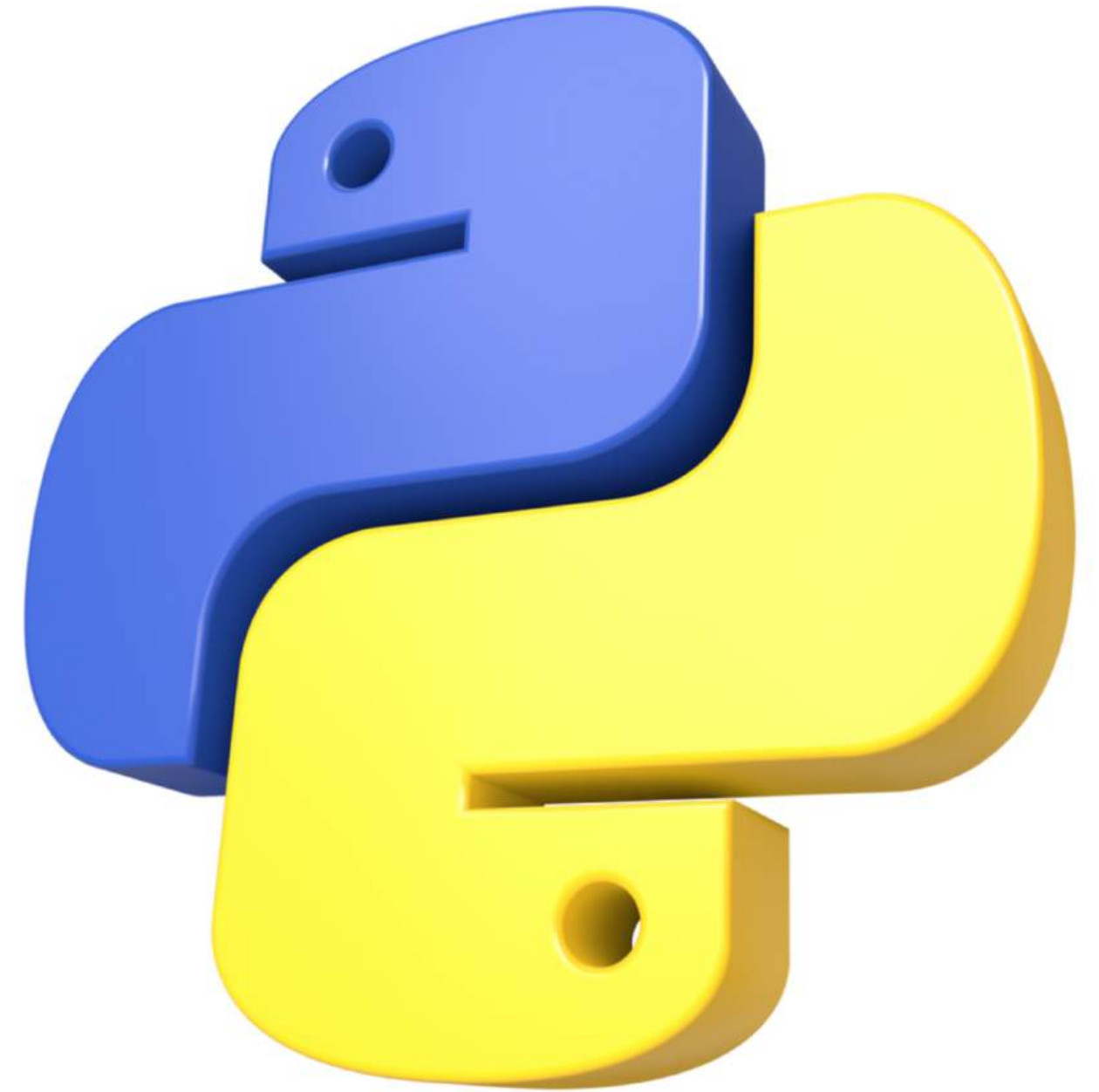


- Que el programador obtenga la colección de datos
- Realizar un pre-proceso
- Alimentar los métodos para su aprendizaje.

# ¿QUÉ ES PYTHON?

Python es un lenguaje de programación ampliamente utilizado en las aplicaciones web, el desarrollo de software, la ciencia de datos y machine learning (ML).

Python es un lenguaje programación fácil de aprender, eficiente, fácil de entender y de utilizar, sobre todo se puede ejecutar en muchas plataformas diferentes.

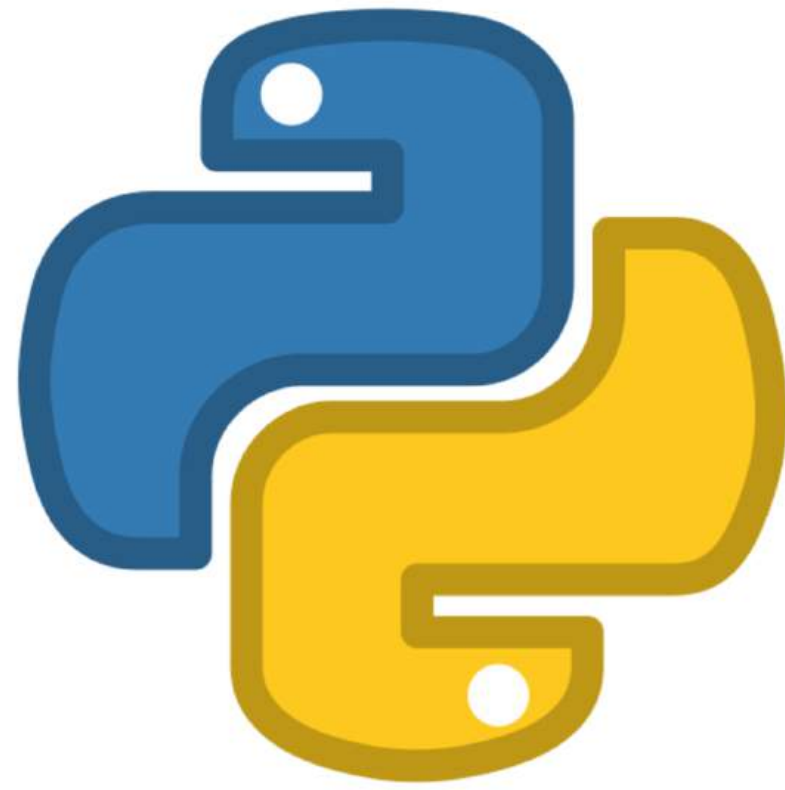


- En el curso usaremos Python 3



# AMBIENTE DE DESARROLLO

Para poder programar, es necesarios que contemos con un ambiente en el cual nos sintamos cómodos, tenga las herramientas necesarias y podamos usarlo de forma rápida



# ¿QUÉ ES COLAB?

Colab es un entorno gratuito propiedad de Google que no requiere configuración y se ejecuta completamente en la nube. Nos permite ejecutar códigos en CPU y GPU



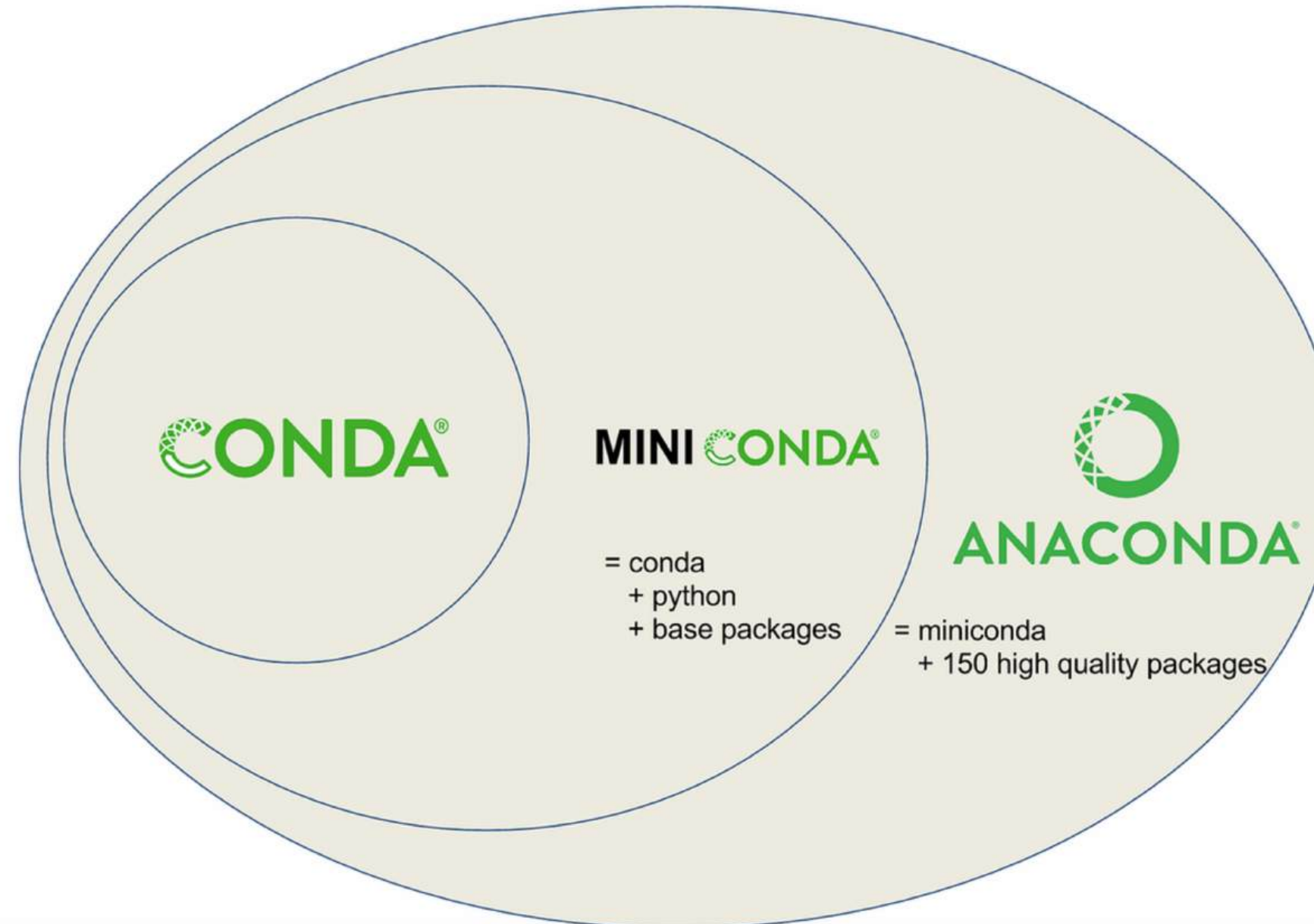
¿QUÉ SERÁ MEJOR UTILIZAR, UN AMBIENTE LOCAL O UN AMBIENTE CLOUD?

<https://colab.research.google.com/>

**Collab:** `00_Introduccion_Python.ipynb`

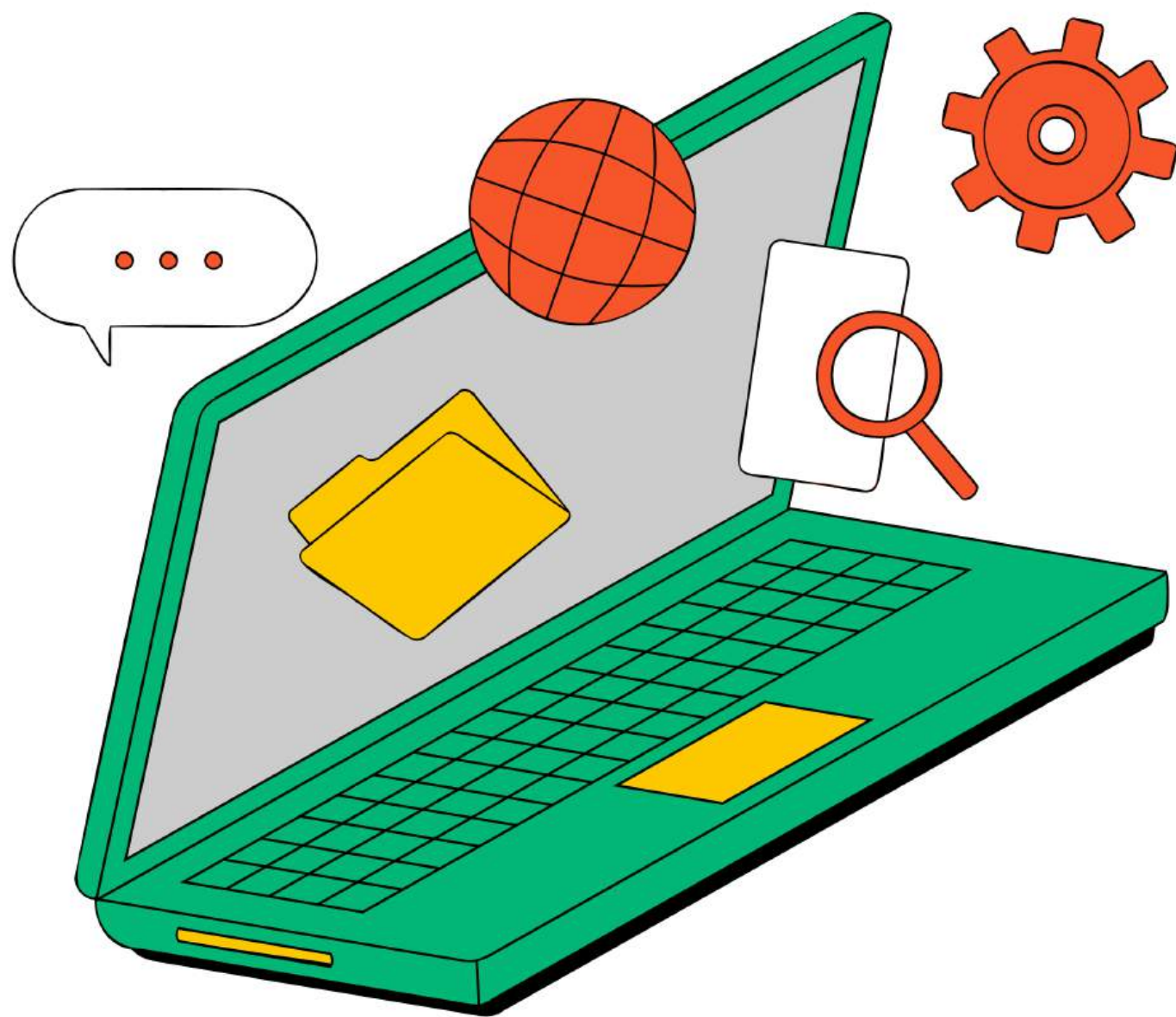
# ¿QUÉ ES MINICONDA?

**Miniconda** es un ambiente virtual: **Conda** es un sistema de gestión de paquetes de código abierto y un sistema de gestión del entorno que se ejecuta en Windows, macOS y Linux





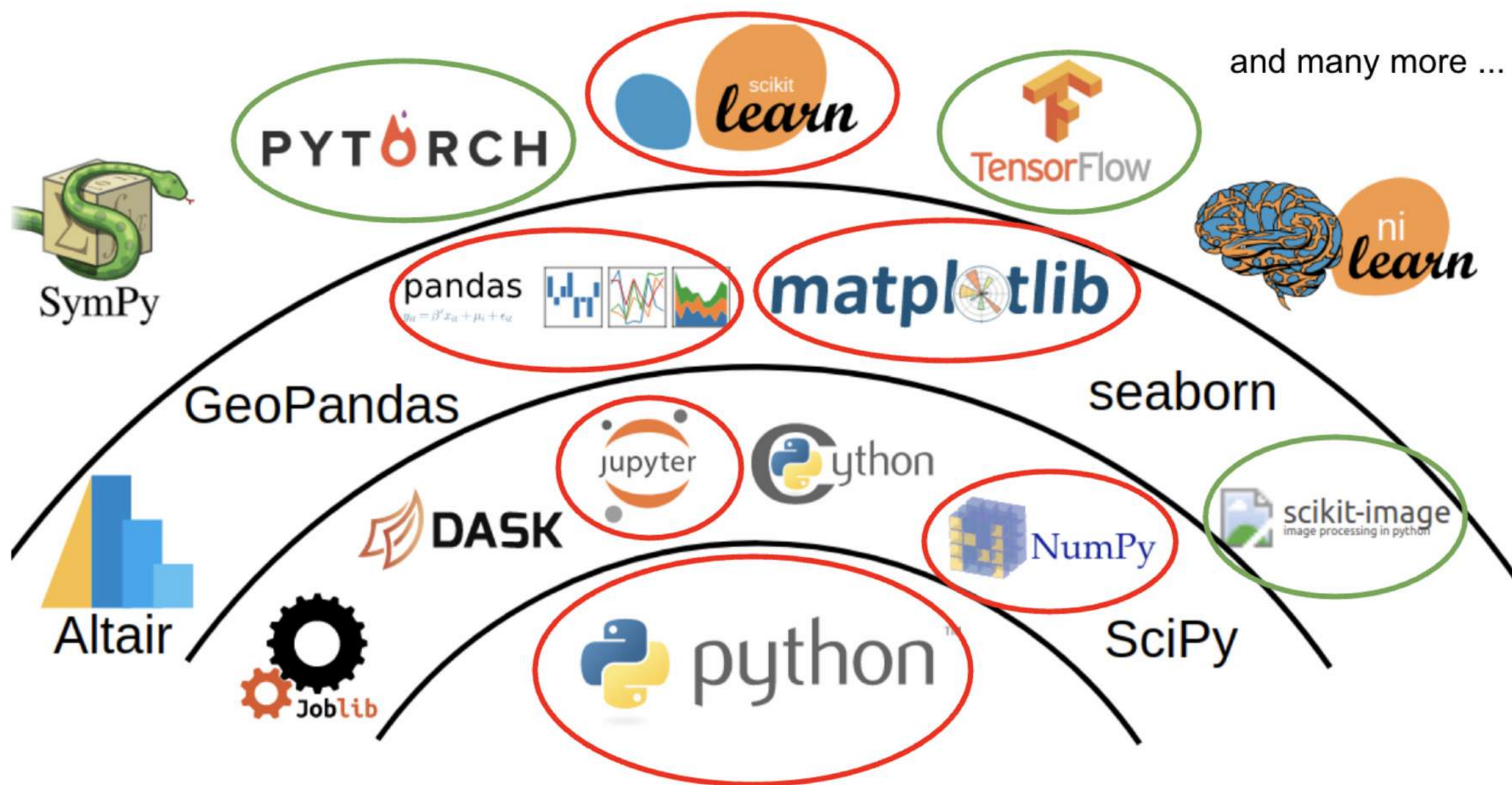




# **LIBRERÍAS BÁSICAS PARA DATA SCIENCE Y MACHINE LEARNING**

Las librerías, son preconfiguraciones de código que nos permite realizar nuestros algoritmos de forma rápida al darnos acceso a funciones específicas para nuestros casos de uso

# LIBRERÍAS





# ¿QUÉ ES NUMPY?

**Numpy** es una biblioteca para el lenguaje de programación **Python** que da soporte para crear **vectores y matrices grandes multidimensionales**, junto con una gran colección de funciones matemáticas de alto nivel para operar con ellas. Es de utilidad práctica para el manejo de **arreglos de objetos N-dimensionales**



# NumPy

**Collab:** [01\\_Introduccion\\_Numpy.ipynb](#)



# ¿QUÉ ES PANDAS?

**Pandas** es un paquete de **Python** que proporciona estructuras de datos rápidas, flexibles y expresivas diseñadas para trabajar con datos etiquetados para hacerlo más fácil e intuitivo

Su objetivo es ser el bloque de construcción fundamental de alto nivel para realizar **análisis prácticos de datos del mundo real** en Python

# Pandas



**Collab:** [02\\_Introduccion\\_Pandas.ipynb](#)

# Anatomy of a DataFrame

	Column (axis = 1)	Make	Colour	Odometer	Doors	Price	Column name
Index number (starts at 0 by default)	0	Toyota	White	150043	4	\$4,000	
1	Honda	Red	87899	4	\$5,000		
2	Toyota	Blue	32549	3	\$7,000		
Row (axis = 0)	3	BMW	Black	11179	5	\$22,000	
4	Nissan	White	213095	4	\$3,500		



# ¿QUÉ ES MATPLOTLIB?

**Matplotlib** es una biblioteca completa para crear visualizaciones estáticas, animadas e interactivas en Python. Matplotlib hace que las cosas fáciles sean fáciles y las difíciles sean posibles

Por ejemplo:

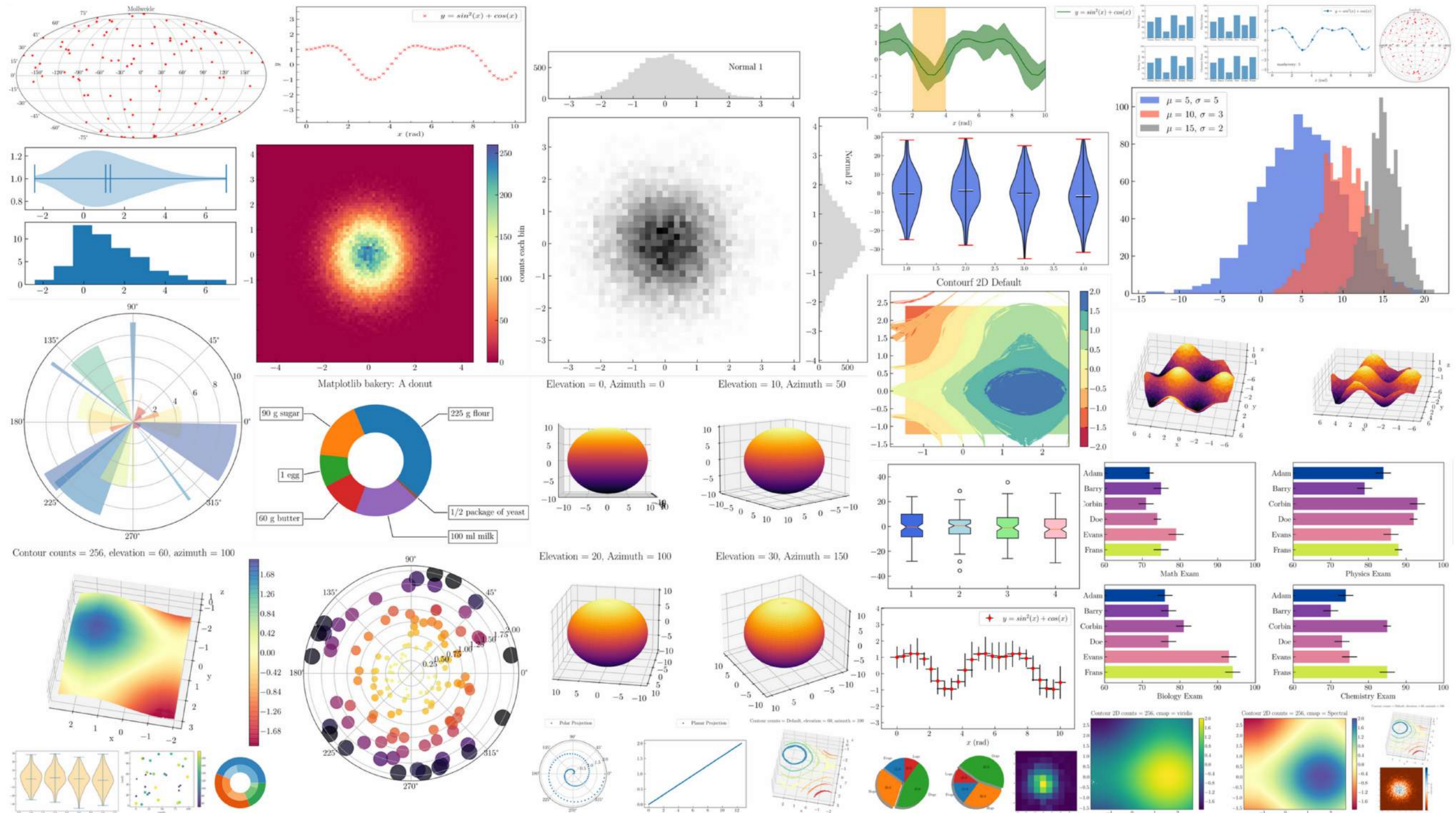
- Hacer figuras interactivas que puedan hacer zoom, desplazarse, actualizar.
- Personalizar el estilo visual y el diseño.
- Exportar a muchos formatos de archivo.
- Visualizar en JupyterLab e interfaces gráficas del usuario.
- Utilizar una gran variedad de paquetes de terceros creados en Matplotlib.



**Collab:** [03\\_Introduccion\\_MatPlotLib.ipynb](#)



# ¿QUÉ ES MATPLOTLIB?





# ¿QUÉ ES SEABORN?

Seaborn es una biblioteca para crear gráficos estadísticos en Python. Está basada en Matplotlib, y se integra con las estructuras de Pandas.

Esta biblioteca es tan potente como Matplotlib, pero aporta simplicidad y funciones inéditas. Permite explorar y comprender rápidamente los datos.

Se pueden capturar marcos de datos enteros, y las funciones internas de mapeo semántico y agregación estadística permiten convertir los datos en visualizaciones gráficas.

Seaborn absorbe toda la complejidad de Matplotlib. Sin embargo, es posible crear gráficos que satisfagan todas tus necesidades y requisitos.



# seaborn

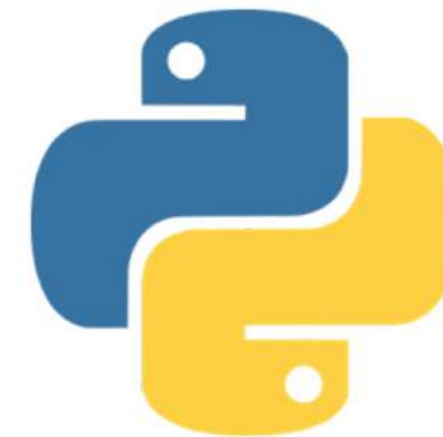
**Collab:** [04\\_Introduccion\\_seaborn.ipynb](#)

# ¿QUÉ ES GITHUB?

GitHub es una plataforma para realizar una colaboración de códigos y un control de versiones



**git**



**Python**

[https://github.com/devil861109/DIP\\_MACHINE\\_LEARNING](https://github.com/devil861109/DIP_MACHINE_LEARNING)



# REQUISITOS NECESARIOS HARDWARE - SOFTWARE - NUBE



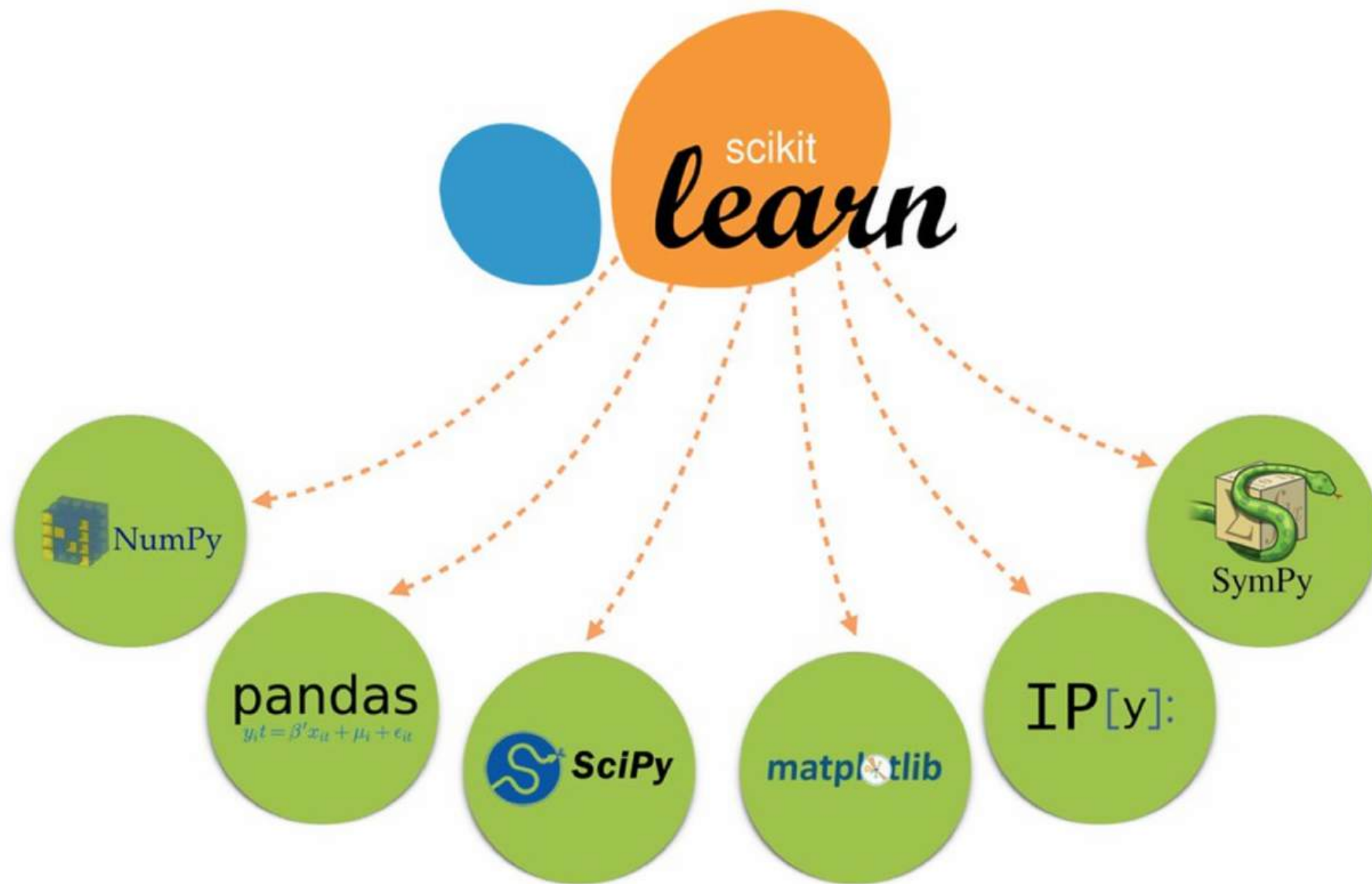
- Seguir las instrucciones del manual de instalación de ambientes para Ubuntu-Mac o Windows, en su material para estudiantes para instalar el ambiente local
- En caso de usar COLAB subir todos los archivos del curso a su Google Drive



# **LIBRERÍAS DE MACHINE LEARNING Y DEEP LEARNING**



# ¿QUÉ ES SCIKIT-LEARN?



Scikit-learn es una biblioteca de aprendizaje automático de software gratuito para el lenguaje de programación Python. En este se implementan algoritmos de Machine Learning eficientemente por lo que es un gran punto de entrada para aprender

A medida que la utilices y la practiques conocerás muchas más características



# ¿QUÉ ES TENSOR FLOW?



# TensorFlow

Tensor Flow es una biblioteca más compleja para **computación numérica** distribuida utilizando grafos de flujo de datos, se puede usar para **crear modelos de aprendizaje profundo** directamente o mediante el uso de otras bibliotecas contenidas en Python

# ¿QUÉ ES KERAS?



Keras es una API diseñada para **reducir la carga cognitiva**, y ¿Por qué elegir Keras?:

- Minimiza la cantidad de acciones del usuario requeridas para casos de uso comunes y proporciona mensajes de error claros y accionables.
- Tiene una extensa documentación y guías para desarrolladores.
- Keras tiene la flexibilidad de bajo nivel para implementar ideas de investigación arbitrarias mientras ofrece funciones opcionales de conveniencia de alto nivel para acelerar los ciclos de experimentación.
- Keras es el marco de aprendizaje profundo más utilizado entre los 5 mejores equipos ganadores en Kaggle. Debido a que Keras facilita la ejecución de nuevos experimentos, le permite probar más ideas que su competencia, más rápido. Y así es como se gana.



## ¿DÓNDE UTILIZAN KERAS?

Keras es utilizado por el [CERN](#), [la NASA](#), [los NIH](#) y muchas más organizaciones científicas de todo el mundo (se utiliza en el [LHC](#)).



La Organización Europea para la Investigación Nuclear, comúnmente conocida por la sigla CERN, es una organización de investigación europea que opera el laboratorio de física de partículas más grande del mundo.



La Administración Nacional de Aeronáutica y el Espacio, más conocida como NASA, es la agencia del gobierno estadounidense responsable del programa espacial civil, así como de las investigaciones aeronáuticas y aeroespaciales.



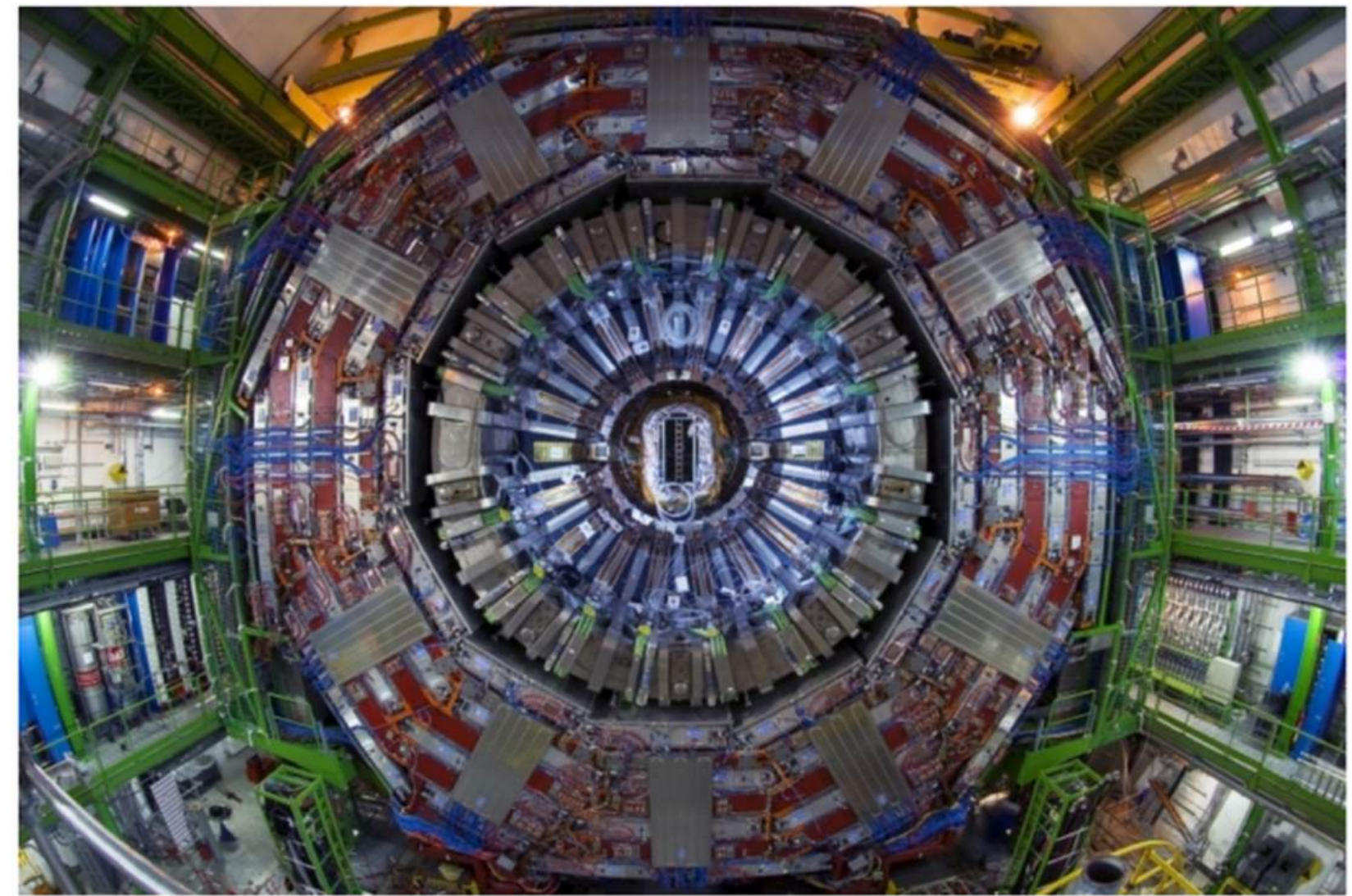
National Institutes  
of Health

NIH es el Instituto Nacional de la Salud, que apoya estudios científicos para el descubrimiento en proyectos de salud.



# ¿QUÉ ES EL LHC?

El LHC – El Gran Colisionador de Hadrones es el acelerador de partículas más grande y de mayor energía que existe y la máquina más grande construida por el ser humano en el mundo.





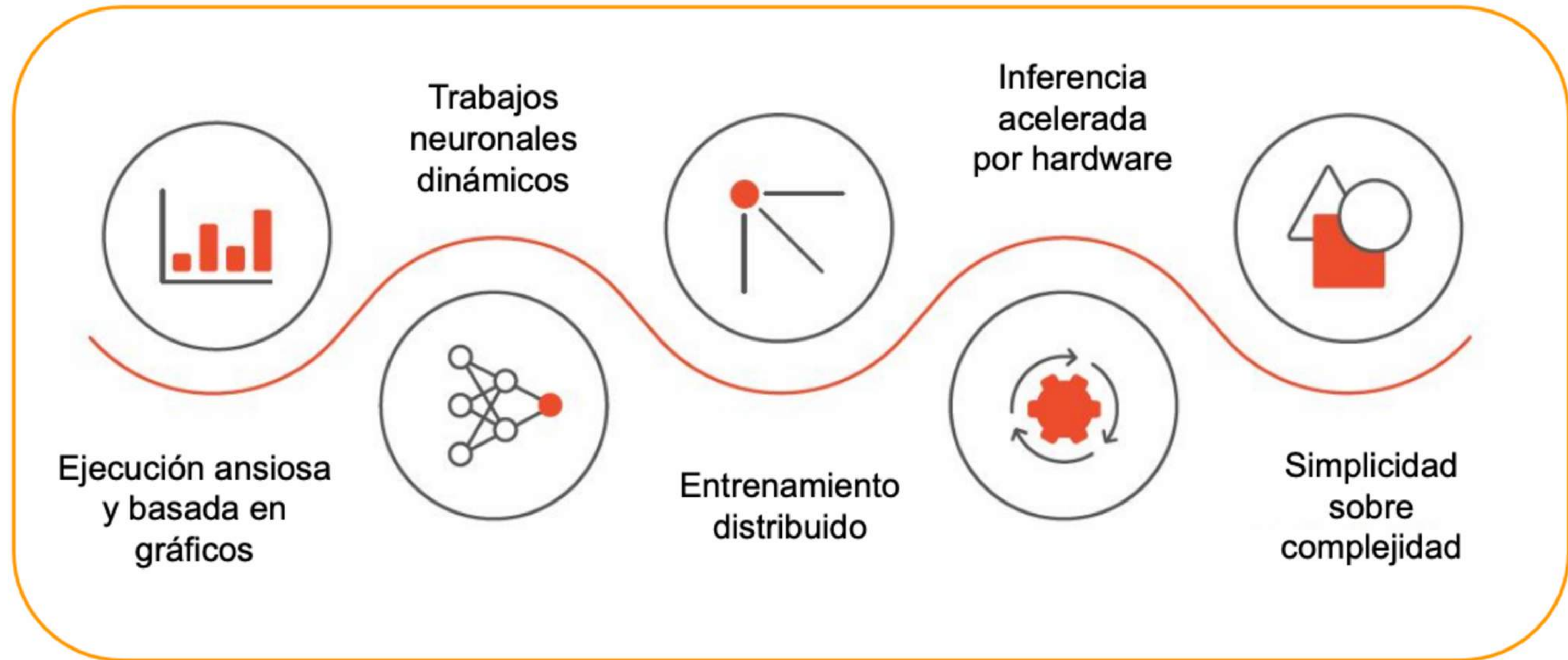
# ¿QUÉ ES PYTORCH?

Pytorch es una biblioteca de **aprendizaje automático** de código abierto, basado en el lenguaje de programación Python y en la biblioteca de Torch

- Es una de las plataformas preferidas para la investigación de aprendizaje profundo
- El marco está diseñado para acelerar el proceso entre la creación de prototipos de investigación y la implementación
- Utilizado para aplicaciones que implementan cosas como visión artificial y procesamiento de lenguaje natural, principalmente
- Fue desarrollado por el Laboratorio de Inteligencia Artificial de FaceBook.



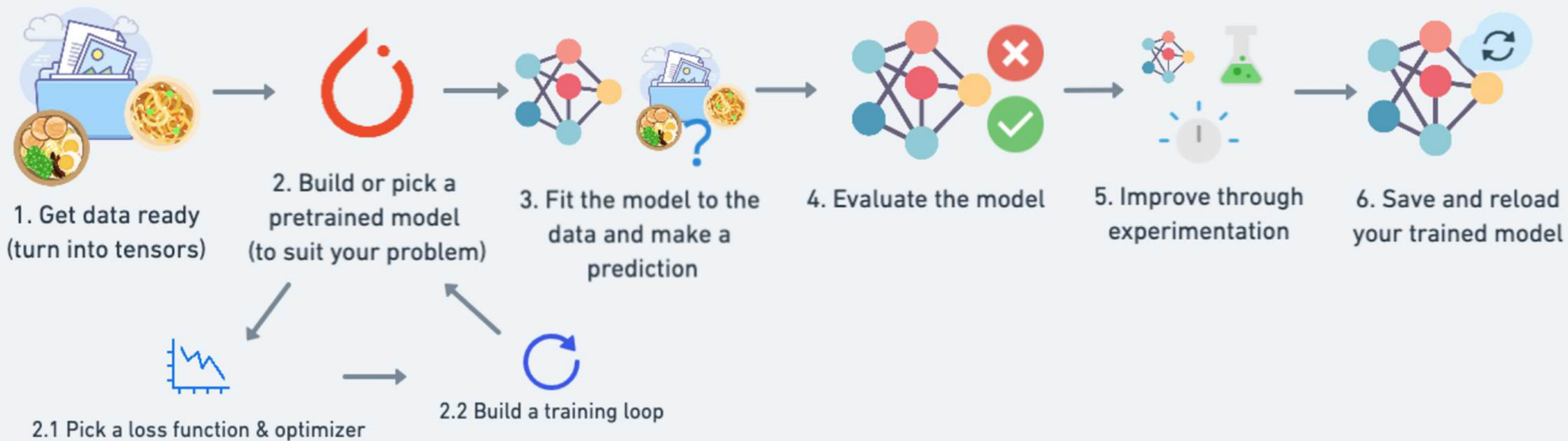
# ¿QUÉ ES PYTORCH?





# PYTORCH WORKFLOW

## A PyTorch Workflow



# PREGUNTAS Y RESPUESTAS

Mtro. Alfonso Gregorio Rivero Duarte

Senior Data Manager – CBRE

**(+52) 5528997069**

**devil861109@gmail.com**

**<https://www.linkedin.com/in/alfonso-gregorio-rivero-duarte-139a9225/>**

