

Machine Learning en Economía

Aproximaciones en Perú

Dr(c) Ciro Ivan Machacuay Meza



Huancayo, 2025

Table of Contents

1 Introducción

- Conceptualización
- Generalidades

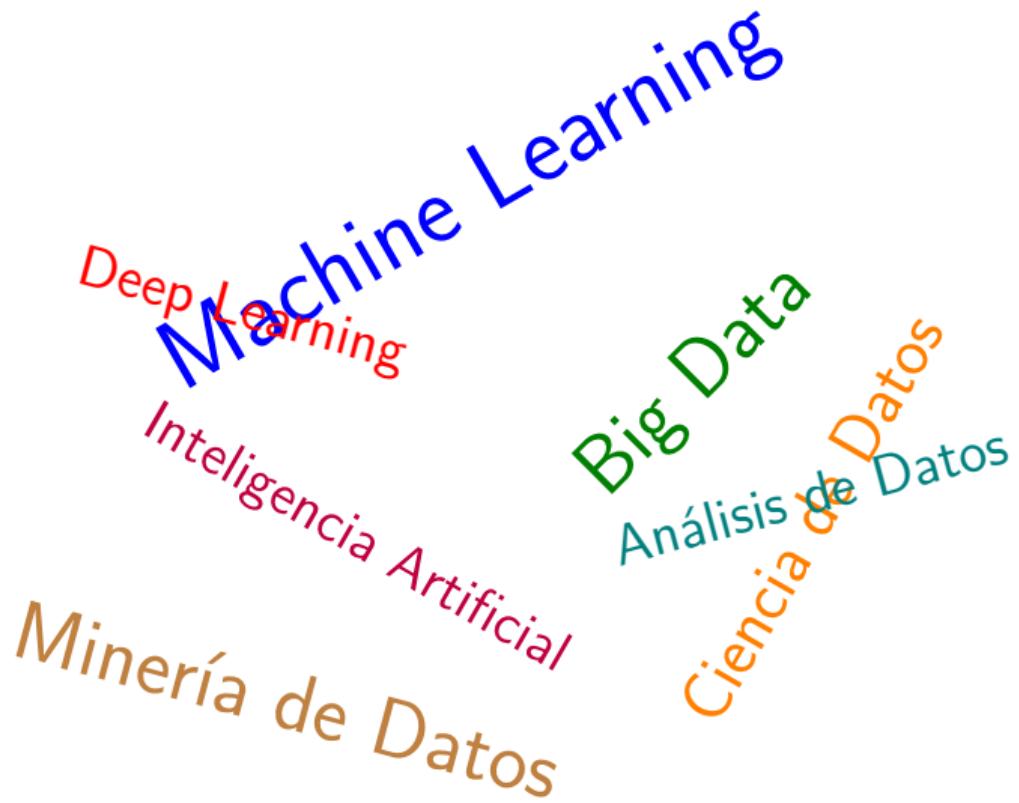
2 Caso Peruano

3 Herramientas

4 Iniciando en Python

- Acerca de
- Instalación
- Cuadernos

Introducción



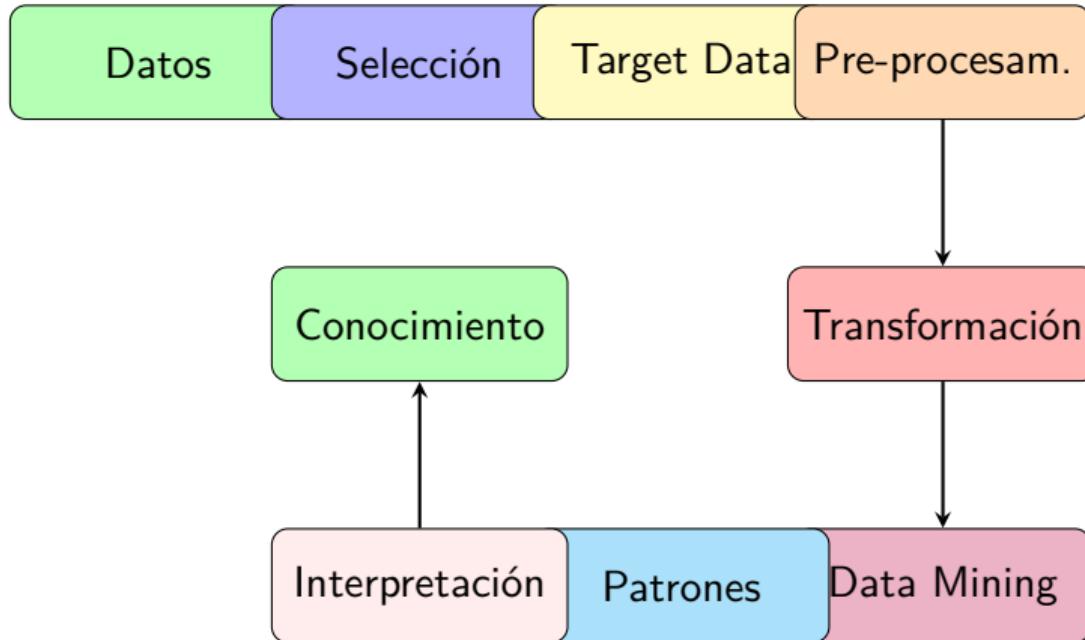
Definición (Ciencia de datos)

Enfoque multidisciplinario que combina las matemáticas, la estadística, la programación y la IA para convertir los datos en conocimiento (IBM).

Hall Varian → Chief Economist Google.

Conceptos básicos

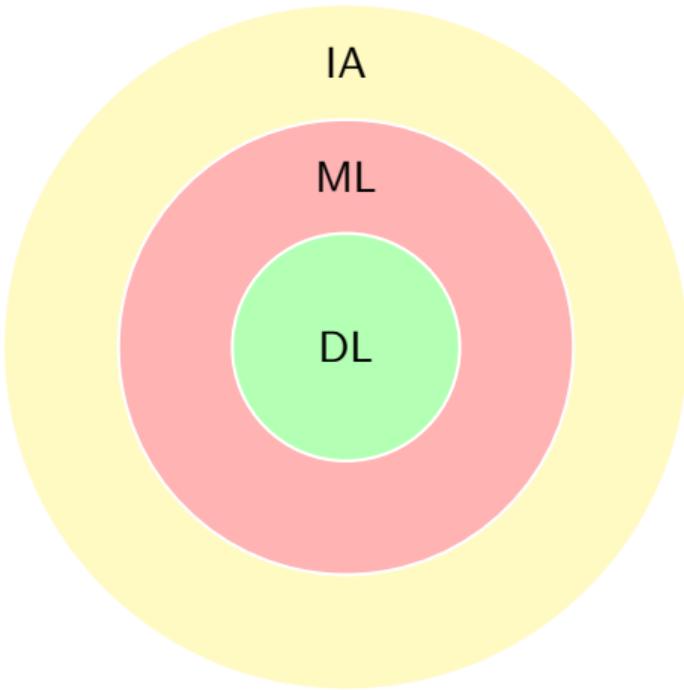
En resumen, el proceso D2K:



Conceptos básicos



Conceptos básicos



Conceptos básicos

Definición (IA)

Toda tecnología que desarrolla la inteligencia de las máquinas para crear sistemas con la finalidad de imitar algunas habilidades humanas.

Definición (ML)

Ciencia que permite a los ordenadores, vía algoritmos y modelos estadísticos, aprender de los datos y realizar tareas automáticamente (sin ser programados previamente).

Definición (DL)

Ciencia que utiliza redes neuronales artificiales para filtrar observaciones (imágenes, textos o sonidos) en capas para aprender a predecir o clasificar información (como el cerebro humano).

Conceptos básicos

Small data:

- ① Maximizar de **pocos** datos.
- ② Solución: estructurar los datos.
- ③ Enfoque: muestreo complejo aproxima muestreo al azar (lento y caro, pero bueno).

Big data:

- ① **Muchos** datos no estructurados.
- ② Extracción inmediata.
- ③ Relativamente baratos.

Conceptos básicos

$$y \leftarrow f(x, u)$$

Estadística clásica:

- ① Interés en estimar el modelo $f(\cdot)$. Efecto causal.
- ② Modelo: teoría o experimento (**out of trouble**).

Machine Learning:

- ① Interés en y . Predecir, medir, clasificar.
- ② Modelo: construirlo y aprenderlo.
- ③ Predicción puntual (no inferencia).

Taxonomía

- ① Lenguaje supervisado: el conjunto de datos contiene tanto entradas (observables) como resultados deseados (target), cuyo objetivo es la predicción.
- ② Lenguaje no supervisado: el conjunto de datos contiene solo entradas, cuyo objetivo es explorar los datos y encontrar patrones o mecanismos de organización.
- ③ Lenguaje por refuerzo: determina automáticamente el comportamiento ideal en un contexto específico.

Taxonomía



Taxonomía

- ① Tareas: predicción o descripción.
- ② Técnicas: algoritmos o procedimientos estadísticos específicos.
- ③ Herramientas: software, paquetes, librerías.

Antecedentes

- ① 1950: Alan Turing crea el “Test de Turing”.
- ② 1952: primer programa capaz de aprender.
- ③ 1967: se crea el algoritmo “Nearest Neighbor”.
- ④ 1997: el ordenador Deep Blue (IBM) vence al campeón mundial de ajedrez Gary Kasparov.
- ⑤ 2006: se acuña deep learning.
- ⑥ 2015: Elon Musk y Sam Altman fundan OpenAI.
- ⑦ 2022: nace chat GPT.

Ejemplo 1 (filtros spam)

El término *spam* se remonta a la serie cómica británica Monty Python, en la cual una pareja en un restaurante recibía sin razón alguna spam (Spiced Pork And Meat). Los correos spam han sido solucionado con filtros (Bayes o Markov).

Ejemplo 2 (reconocimiento facial)

Algoritmos utilizan un gran conjunto de datos de fotografías etiquetadas según tienen un rostro o no para estimar una función $f(x)$ que predice la presencia y de un rostro a partir de los píxeles x .

¿Estos algoritmos sólo aplican técnicas estándar a conjuntos de datos nuevos y grandes?

¿cómo encajan con lo que sabemos? Como economistas, ¿cómo podemos utilizarlos?

Resuelve más problemas, entre ellos el volumen de datos: big data.

¿Porqué utilizar ML?

- ① Logra descubrir patrones generalizables (inferencia).
- ② Ajusta formas funcionales complejas (underfitting).
- ③ Reduce sesgo de selección en evaluación de intervenciones.
- ④ Permite tratar grandes volúmenes de datos.
- ⑤ ...

Implicancias

¿Impacto de ML?

- ① Nuevos modelos econométricos.
- ② Nuevos métodos de identificación de efecto causal.
- ③ Investigaciones interdisciplinarias.
- ④ Automatización de procesos.
- ⑤ ...

Adds: <https://n9.cl/dfa62>

Definición (entrenamiento de modelo)

Proceso que consiste en proporcionar datos a un algoritmo para que aprenda determinados patrones. Puede realizar predicciones, evidenciar estructuras entre otros.

Implicancias

Underfitting y overfitting:



Figure: Extraído de Aprende ML

Caso Peruano

Introducción

The screenshot shows a web browser displaying the website ia.up.edu.pe/nosotros/. The header includes the logo of the Centro de Investigación of the Universidad del Pacífico. The main navigation menu has tabs for NOSOTROS (highlighted in green), PROYECTOS, PUBLICACIONES, ENSEÑANZA, and NOTICIAS. Below the menu, there are three main content boxes:

- INVESTIGACIÓN**: Features a magnifying glass icon. Text: "La investigación científica es un pilar fundamental del laboratorio. En ese sentido, los investigadores del laboratorio se dedican a estudiar y desarrollar algoritmos y modelos de toma de decisiones orientados a la IA. Estas investigaciones contribuyen al conocimiento general en el campo de la IA y también en sostener proyectos específicos."
- DESARROLLO DE PROYECTOS**: Features a gear icon. Text: "El desarrollo de proyectos diseñados para contribuir significativamente al avance del conocimiento en IA. A través de estas iniciativas, buscamos impulsar la tecnología, desarrollar soluciones novedosas y comprender cómo la IA puede transformar diversos sectores."
- CAPACITACIÓN ESPECIALIZADA**: Features a person icon. Text: "En tercer lugar, nuestro enfoque incluye la capacitación especializada en inteligencia artificial. Ofreciendo talleres destinados a la comunidad UP y organizaciones con habilidades especializadas en IA."

Figure: Lab-IA UP

<https://ia.up.edu.pe/>

Caso peruano

Nº	Institución	Aplicación	ML
1	IGP	Pronóstico metereológico de precipitación	Downscaling
2	MIDIS y UPCH	Detección de anemia en niños	Deep Learning
3	Banca Comercial	Modelos de riesgo crediticio y consumo	Deep Learning
4	ESSALUD	Predicción de sobrepeso y obesidad	Redes reuronales
5	BCRP	Proyecciones de Inflación	Random Forest

Table: Algunos tratados en Perú

Herramientas

Plataformas

- Microsoft Azure → Azure Machine Learning.
- Anaconda Navigator → Anaconda, Jupiter, PyCharm, R.
- IBM Watson → watsonx.
- Amazon Web Services (AWS) → Amazon Sagemaker.
- Apache spark.
- Vertex AI Platform (Google).

Plataformas

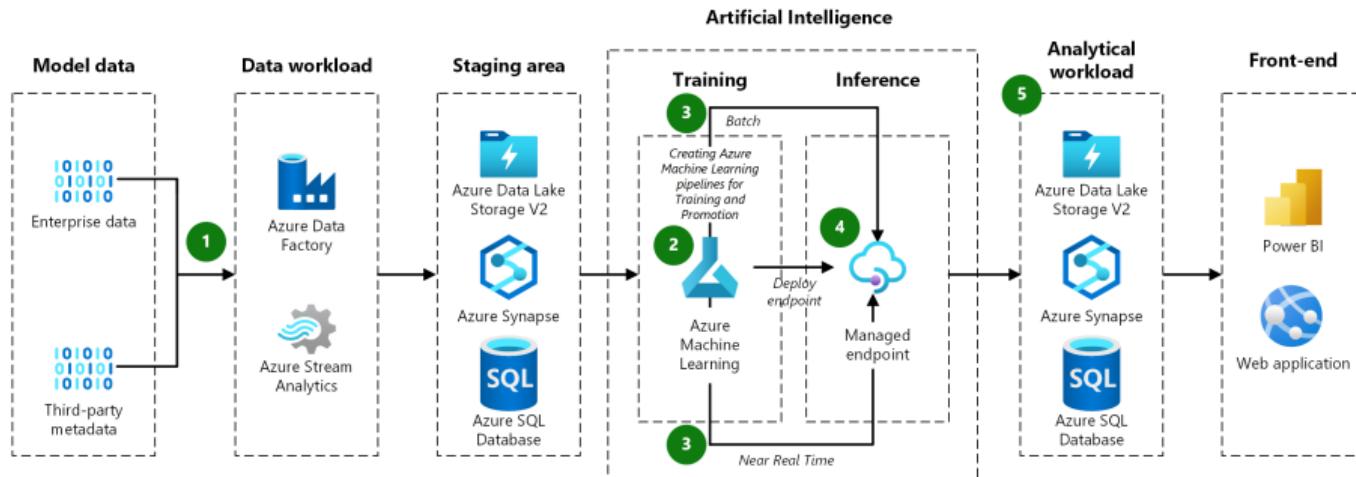


Figure: Azure

Plataformas

The screenshot shows the IBM Watson website at ibm.com/es-es/watson. The navigation bar includes links for Products, Solutions, Consultoría, Soporte, and Más. A sidebar on the left lists various product categories under 'Destacados'. The main content area displays a grid of 'Destacados' products:

Nombre del Producto	Descripción
API Connect	Software para crear, proteger, socializar y gestionar API rápidamente
IBM Cloud	Plataforma de computación en la nube bajo demanda y APIs
Robotic Process Automation (RPA)	Software para automatizar flujos de trabajo y procesos empresariales
Concert	IA para analizar y gestionar la infraestructura de TI mediante lenguaje natural
IBM Z	Sistema con IA integrado en chip y seguridad cuántica
Storage Defender	Software de resiliencia de datos para la detección de amenazas y la recuperación de datos
Environmental Intelligence	Software como servicio para predecir y responder a eventos meteorológicos y climáticos
Instana	Software para la supervisión y automatización del rendimiento de las aplicaciones
Turbonomic	Software para gestionar y optimizar el uso de recursos de TI
Envizi	Software como servicio para analizar indicadores ambientales, sociales y de gobierno
MaaS360	Software de gestión unificada de puntos finales para muchos tipos de dispositivos
Watson Assistant	Plataforma de datos e IA
FlashSystem	Matriz All-Flash con rendimiento y capacidad para cualquier carga de trabajo
Maximo	Software para la gestión de activos y flujos de trabajo asociados
Planning Analytics	Software para automatizar la planificación
watsonx	Plataforma de datos e IA
watsonx Assistant	Plataforma de datos e IA
Watson Assistant	Sistema de asistente personal que automatiza tareas repetitivas

A blue callout box from the 'watsonx Assistant' section contains the text: '¡Hola! ¿Cómo podemos ayudarte?'. The bottom of the page features a footer with navigation icons and the text 'javascript:void(0) para Watsonx'.

Figure: IBM

Plataformas

The screenshot shows the AWS SageMaker landing page in Spanish. The top navigation bar includes links for Acerca de AWS, Contacto, Soporte, Español, Mi cuenta, Iniciar sesión, and Cree una cuenta AWS. Below the navigation is a secondary menu with tabs: re:Invent, Productos, Soluciones, Precios, Documentación, Aprender, Red de socios, AWS Marketplace, Habilitación para clientes, Eventos, Explorar más, and a search icon. The main content area features a large heading "Amazon SageMaker" and a sub-headline: "Cree, entrene e implemente modelos de machine learning para cualquier caso de uso con infraestructura y flujos de trabajo completamente administrados". Below this are two buttons: "Comience a utilizar SageMaker" and "Pruebe un tutorial práctico". To the right of the main content is a sidebar titled "Herramientas de ML" which lists various services: Todas las características, Piloto automático, Clarify, Data Wrangler, Despliegue, Administrador de periféricos, Experimentos, Almacén de características, HyperPod, Gobernanza de ML, MLOps, Notebooks, Ground Truth, JumpStart, Canalizaciones, Studio Lab, and Formación. A dark overlay box contains the text "Puedo ponerlo en contacto con un representante de AWS." At the bottom left is the URL "https://aws.amazon.com/es/sagemaker/#".

Figure: AWS

Plataformas

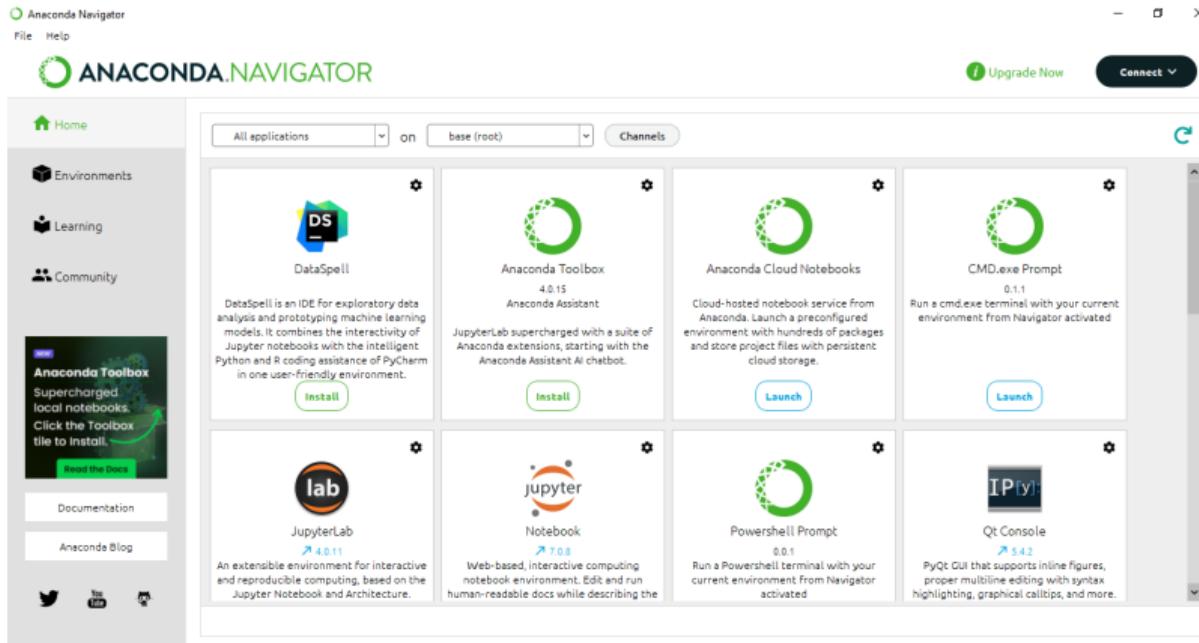


Figure: Anaconda Navigator

Lenguajes de Programación

- Python.
- R.
- Julia.
- Scala.
- MATLAB ML.
- C++.
- ...

Lenguajes de Programación

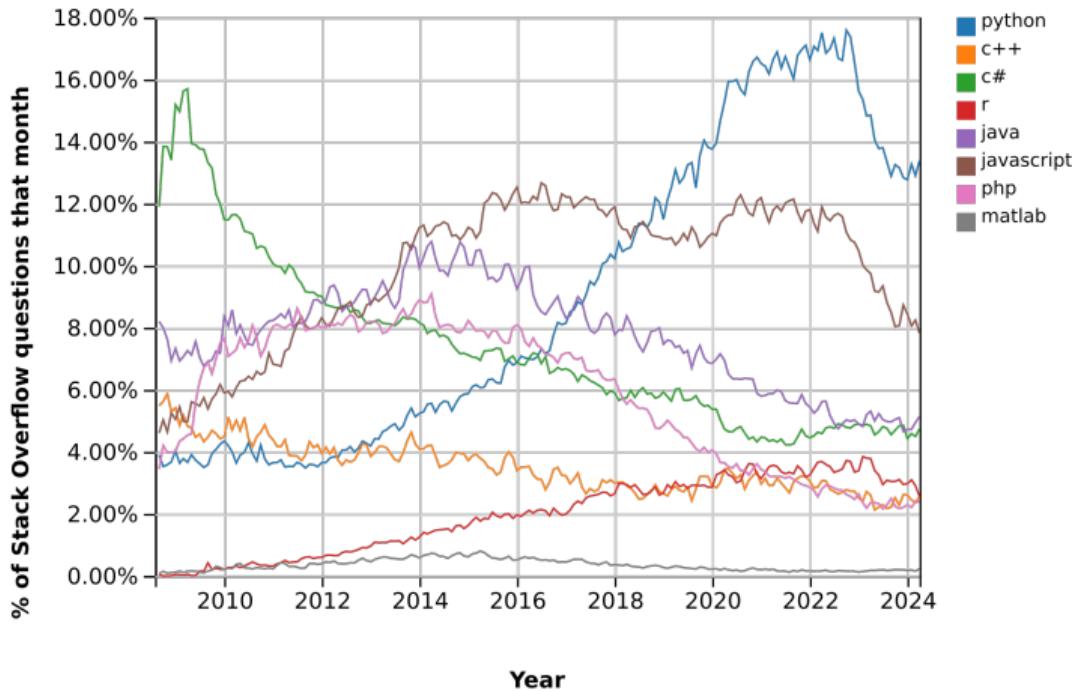


Figure: Tendencia actual (ver)

Paquetes

- TensorFlow.
- Keras.
- Scikit-learn.
- MLib
- PyTorch
- ...

Iniciando en Python

Fundado por Guido Van Rossum en 1989. Inspirado en el programa de televisión de la BBC “Monty Python’s Flying Circus”.

Definición (Python)

Es un lenguaje de programación orientado al desarrollo de software, ciencia de datos y ML.

Características

- Lenguaje interpretado.
- Lenguaje sencillo.
- Orientado a objetos (POO).
- Bibliotecas dinámicas.

- Jupiter: <https://jupyter.org/>.
- PyCharm: <https://www.jetbrains.com/es-es/pycharm/>.
- IDLE: madre.
- Sublime Text: <https://www.sublimetext.com/>.
- Spyder: <https://www.spyder-ide.org/>.
- Atom: <https://atom-editor.cc/>.
- Google Colab: <https://colab.research.google.com/>.

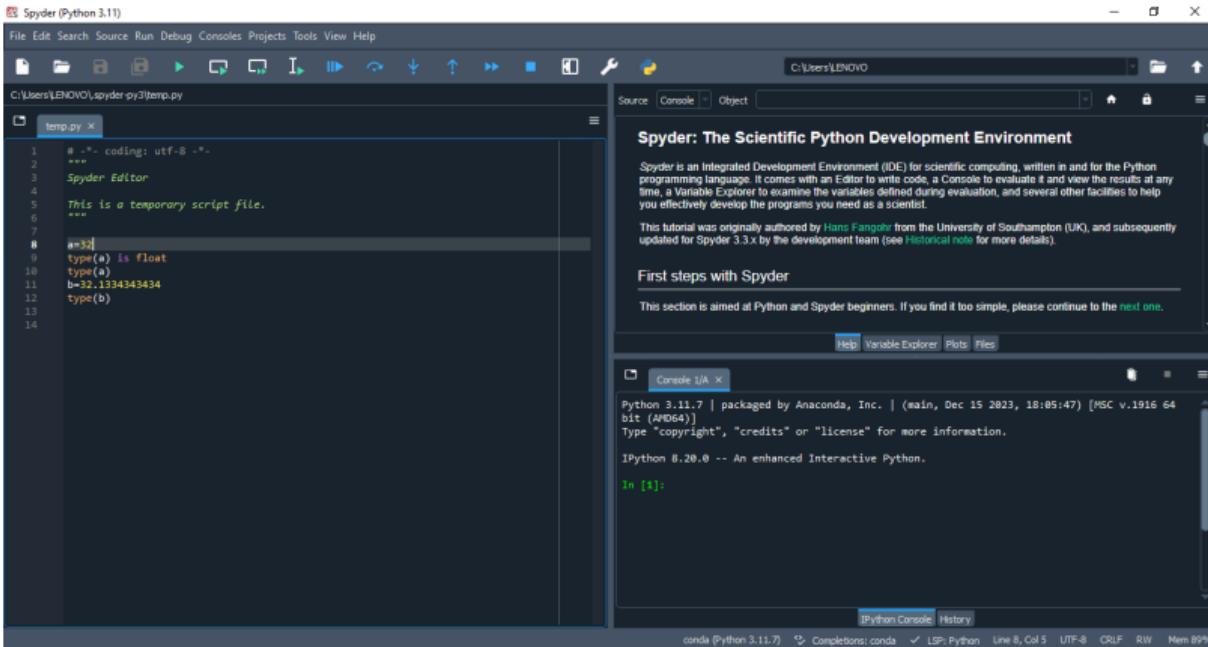


Figure: Spyder

IDE

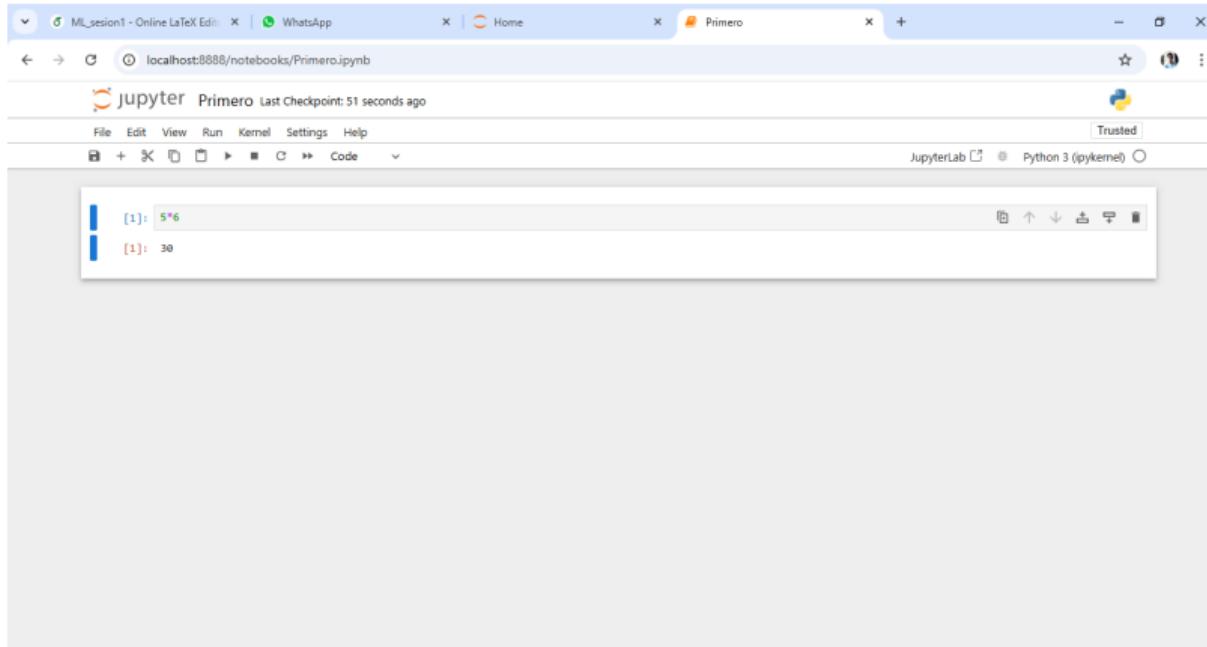


Figure: Jupiter

Instalando Colab

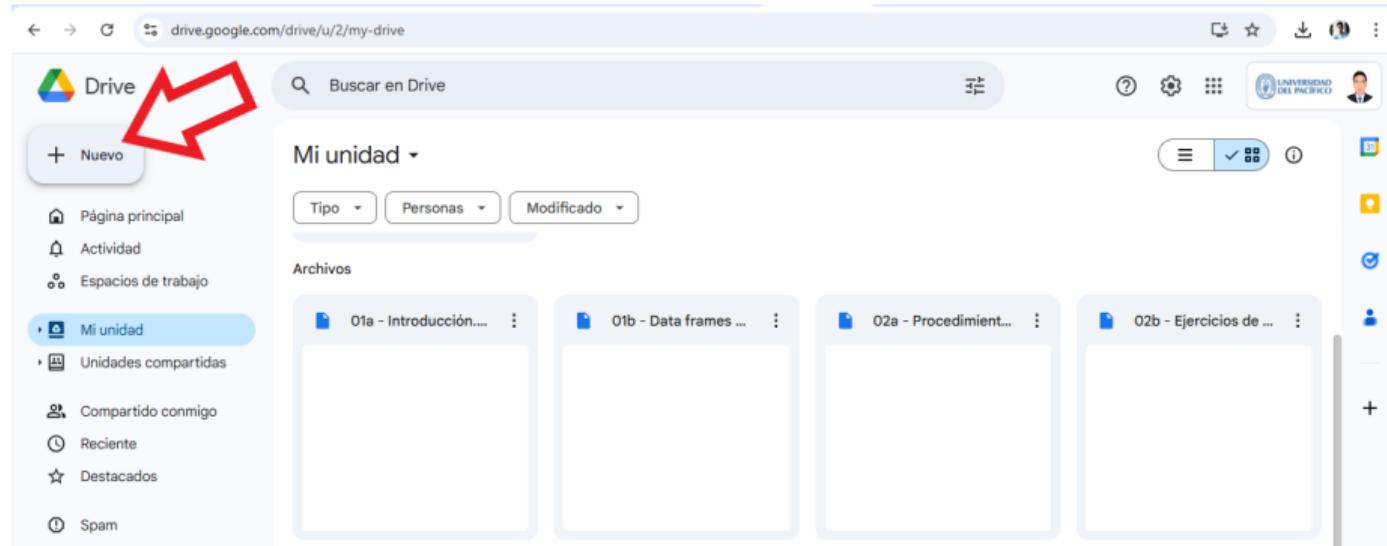


Figure: Paso 1

Instalando Colab

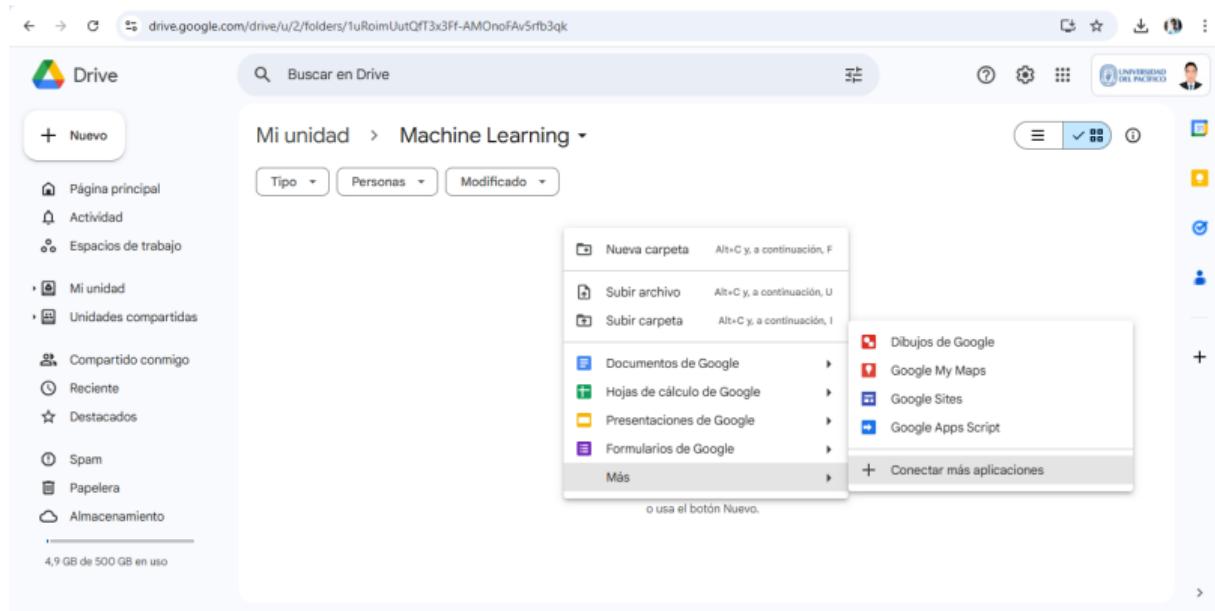


Figure: Paso 2

Instalando Colab

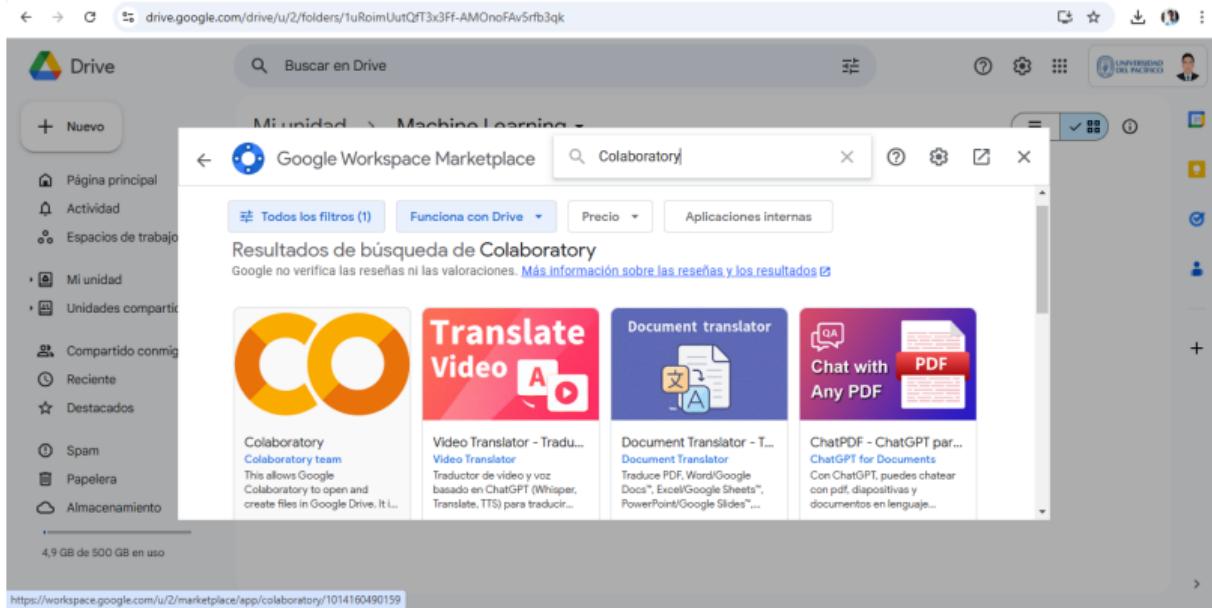


Figure: Paso 3

Instalando Colab

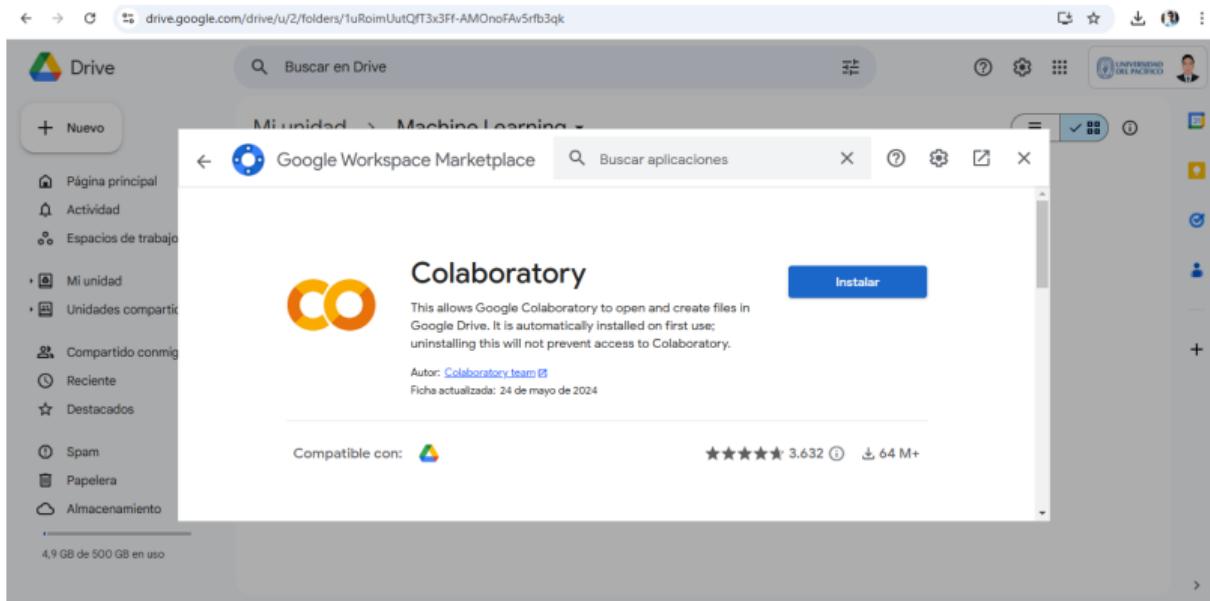


Figure: Paso 4

Instalando Colab

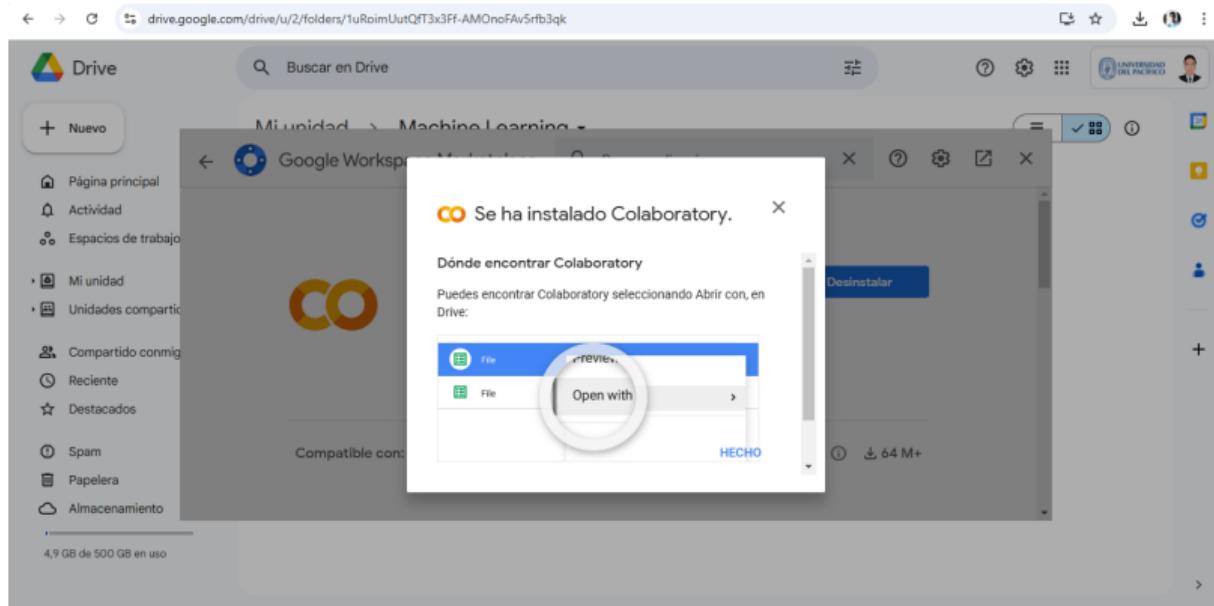


Figure: Paso 6

Cuadernos

Los ficheros con extensión *.py* corresponden al script de python escritos en formato de texto plano (código fuente), mientras que *.ipynb* corresponden a un **python notebook** escrito en un IDE (Integrated Development Environment).

```
# Primer código en Python
print((2+5)*3)
21
```

A trabajar...

Referencias

- Géron, A. (2023). Hands-On Machine Learning with Scikit-Learn, Keras & TensorFlow. O'Reilly.
- Sosa, W. (2019). Big Data. Siglo XXI.
- <https://quantecon.org/>.

¡Gracias!