

# Curso de introducción a Ciencia de Datos

José Luis Cumbreira Sánchez

Julio 2019

## 1 Introducción.

Este curso está orientado para personas con conocimiento en phyton y un nulo conocimiento en ciencia de datos, ya que se partirá desde cero. Para nuestro código en phyton se utilizará un software conocido como Anaconda Navigator, el cual es una distribución de phyton. Dentro de Anaconda usaremos un sistema basado en notebooks llamado Jupiter Notebook.

## 2 Anaconda Navigator.

1. Entrar en la pagina oficial de la distribución anaconda [Anaconda](#).
2. Seleccionar la distribucion 3.7, ya que la 2.7 está desactualizada y su soporte está anunciado que será retirado en unos años, por lo que no es muy recomendable.
3. Descargar la distribución pertinente para nuestro SO, ya sea Windows, Linux o macOS.
4. Una vez descargado e instalado ya tenemos la interfaz gráfica de la distribución.
5. Desde esa interfaz podemos seleccionar Jupiter Notebook, esta es la única herramienta que usaremos ahora, las demás podran ser útiles más adelante.

## 3 Jupiter Notebook.

- Jupiter es un notebook que cuenta con celdas, cada celda se compila-ejecuta individualmente sin afectar al resto hasta una nueva ejecución de las mismas.
- Jupiter cuenta con 2 tipos de celdas:
  - Code: El contenido de las celda se ejecuta como código phyton.
  - Markdown: EL contenido de la celda se basa en html, con soporte Latex, sirve para escribir apuntes,titulos,descripciones...

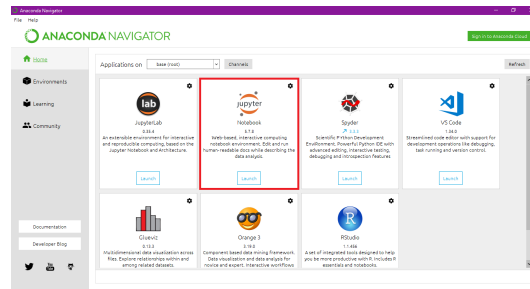


Figure 1: Jupyter Notebook en la distribución Anaconda.

## 4 Anaconda Prompt.

Comandos principales:

- Actualizar conda a la versión mas avanzada: `conda update conda`.
- Instalar cierto paquete phyton: `pip install PACKAGE`.

## 5 Curso.

El curso se divide en 2 bloques:

- Modelos. En esta parte, se explican los diversos y principales métodos en Machine Learning, tanto para regresión como para clasificación. En todos se realiza una breve explicación teórica acompañada de un ejemplo y/o algún caso interesante a explorar.
- Validación-Evaluación. En esta parte, se desarrollan los diversos métodos para la verificación y/o evaluación de nuestros modelos. Se ha tratado de explicar cada método junto con una demostración sobre algún modelo.

Todo el curso va acompañado de esquemas, dibujos y diagramas para que la comprensión sea lo más sencilla posible.