

# Instalación de Python y un entorno de trabajo para MachineLearning

La forma más rápida y eficiente de instalar Python, manejar sus librerías y evitar problemas de dependencia (por actualizaciones y versiones) es mediante el uso de **Miniconda** que además de la instalación de Python, instala conda[1] un sistema gestor de paquetes y de entornos virtuales<sup>1</sup> [2].

Una forma alternativa es utilizar Anaconda, que al igual que Miniconda, es soportada por la misma compañía Anaconda. Ambas alternativas instalan la misma versión de Python y de conda. La razón por la que se recomienda Miniconda es porque permite instalar solamente lo que se requiere para este curso, con la opción de descargar otros paquetes cuando se requieran. Anaconda por el contrario instala por defecto, una cantidad excesiva de paquetes que raramente son utilizados, lo cual requiere demasiado espacio y tiempo.

## Instalación de Miniconda [3]

1. [Descarga aquí](#) el archivo correspondiente a tu sistema operativo, ya sea que sea que MacOS, Windows y Linux, eligiendo la versión última versión de Python.

### *Instalación para Windows*

2. Da doble clic en el archivo descargado.
3. Sigue las instrucciones que se muestren aceptando las opciones que se proponen por defecto (si es necesario se pueden cambiar después).
4. Desde el menú de Inicio de Windows abre el programa Anaconda Prompt.

### *Instalación para MacOS*

2. En la Terminal de tu Mac, navega hasta el directorio en donde descargaste el instalador y corre la siguiente línea :  

```
bash Miniconda3-latest-MacOSX-x86_64.sh
```
3. Sigue las instrucciones, aceptando las opciones por defecto (si es necesario se pueden cambiar después)
4. Cierra la Terminal y vuélvela a abrir, para que los cambios sean actualizados.

---

1 Los entornos virtuales permiten controlar las versiones de software (en nuestro caso Python y sus librerías) usado para análisis o aplicaciones

### ***Instalación para Linux***

2. En la Terminal de Linux, navega hasta el directorio en donde descargaste el instalador y corre la siguiente línea :

```
bash Miniconda3-latest-Linux-x86_64.sh
```

3. Sigue las instrucciones, aceptando las opciones por defecto (si es necesario se pueden cambiar después)
4. Cierra la Terminal y vuélvela a abrir, para que los cambios sean actualizados.

### ***Para todos los sistemas operativos:***

5. Para ver todos los paquetes que instalados por defecto en la distribución de Miniconda, corre la siguiente línea:

```
conda list
```

## **Creación de un entorno virtual para análisis de datos**

Un entorno virtual es un ambiente de trabajo aislado, lo que permite instalar determinadas librerías o versiones de librerías sin que afecte al resto del sistema principal o de otros ambientes.

Por defecto, Miniconda crea un entorno llamado **base** que estará activo cuando abramos la terminal. Este entorno contiene todos los programas enlistados cuando se utilizó el comando **conda list**.

Aunque podríamos usar este entorno, una mejor práctica es crear un entorno nuevo. Para ello es conveniente cambiar la configuración de conda para que no abra automáticamente el entorno **base**.

### ***Desactivar el entorno base***

- En **Mac y Linux**, se debe correr la línea que se muestra a continuación y enseguida cerrar y volver a abrir la terminal:

```
conda config --set auto_activate_base false
```

- Los usuarios de **Windows** deberán desactivar manualmente el entorno base utilizando
- ```
conda deactivate
```

## ***Creación de un entorno específico para análisis de datos mediante conda-forge***

El entorno que vamos a crear lo titularemos `MachineLearning`, en éste instalaremos las librerías que necesitaremos para nuestros análisis.

1. Crea el ambiente de trabajo mediante la siguiente línea

```
conda create -n MachineLearning
```

2. Actívalo con la siguiente instrucción:

```
conda activate MachineLearning
```

3. Agrega el canal<sup>2</sup> `conda-forge` mediante:

```
conda config --env --add channels conda-forge
```

Se recomienda utilizar el canal `conda-forge`, ya que es resultado de un esfuerzo colectivo con una gran variedad de librerías y paquetes que son cuidadosamente actualizados y donde se asegura tener versiones compatibles para macOS, Linux y Windows[4].

Puedes ver los canales instalados mediante la siguiente línea:

```
conda config --show channels
```

Cada canal tiene al menos una dirección URL asociada donde se localizan los repositorios de paquetes y librerías. Puedes ver estas direcciones utilizando:

```
conda info
```

---

<sup>2</sup> Los repositorios desde donde podemos descargar las librerías de Python se llaman canales, Anconda, Inc provee por defecto el canal llamado `default`, sin embargo permite a cualquier usuario crear su propio canal si necesita otros paquetes que no están incluidos en el canal `default`, por esta razón existe una gran variedad de canales disponibles en la web desde donde se pueden descargar librerías de Python.

4. Instala las librerías Pandas<sup>3</sup>, Scikit-learn<sup>4</sup>, Scikit-image<sup>5</sup>, Matplotlib<sup>6</sup> y Jupyter Notebooks<sup>7</sup> copiando el siguiente comando (es importante copiarlo en una sola línea, separando mediante un espacio, cada librería a ser instalada):

```
conda install pandas matplotlib notebook jupyter_contrib_nbextensions  
scikit-image scikit-learn
```

## Referencias

- [1] «Conda — conda 4.8.3.post5+125413ca documentation». <https://docs.conda.io/projects/conda/en/latest/> (accedido mar. 26, 2020).
- [2] *Anaconda is Bloated - Set up a Lean, Robust Data Science Environment with Miniconda and Conda-Forge*. .

---

3 Pandas es la principal librería de python utilizada Analisis de datos. Al instalar Pandas, por defecto se instala Numpy, sobre la cual funciona Pandas. Numpy es ampliamente usada en Ciencia de Datos porque permite el fácil manejo de matrices y sus operaciones.

4 Scikit-learn es una importante librería para machine learning

5 Es una librería para el procesamiento de imagenes

6 Matplotlib es las principales librerías de Python usadas para graficar y visualizar información

7 Jupyter notebook es una aplicación que nos ayudará a hacer nuestros análisis paso a paso, crear visualizaciones e incluir comentarios, como si se tratara de nuestro cuaderno de apuntes.

- [3] «Miniconda — Conda documentation». <https://docs.conda.io/en/latest/miniconda.html> (accedido mar. 26, 2020).
- [4] «A brief introduction — conda-forge 2019.01 documentation». <https://conda-forge.org/docs/user/introduction.html> (accedido mar. 26, 2020).