Python
for
Machine Learning
&
Data Science

Suman Samui

Beginner's Guide to Python for Machine Learning

- Basics of Python Programming → A good starting point
- 'You have to be very good at programming concepts before starting to analysing data using Python.'

  → A wrong idea
- As a data scientist or ML engineer, you will use Python for retrieving, cleaning, visualizing and building models; and not for developing software applications.
- Therefore, you have to be more focussed on most of your time in learning the modules and libraries in Python to perform these tasks.

Learn the basics of python:

- **Variables and Types**
- **Lists**
- **Basic Operators**
- **String Formatting**
- **Basic String Operations**
- **Conditions**
- **Loops**
- **Functions**
- **Classes and Objects**
- **Dictionaries**
- **Modules and Packages**

Two important packages:

- Numpy → Array
- Pandas → Data-frames

Basic coding suggestions:

- **Indentation in Python is very important.**
- **Body of the Main Program should not be unnecessarily long**
- **You should use modularity as much as possible.**
- **Your code should be always ready for deployment and sharing.**

What framework to use in writing the python code?

- It is a good starting point to install the newest Python version with the latest versions of the packages.
- Avoid the temptation to pip install <packages> into your default Python environment.
- **Configure your programming environment**

A better way — the correct way — is to create a project-specific environment

Python Programming Environment

- The **Jupyter Notebook** is a powerful programming environment

(A best way to start your python journey)

- An installation of **Anaconda** is also a common way to install Jupyter Notebooks.

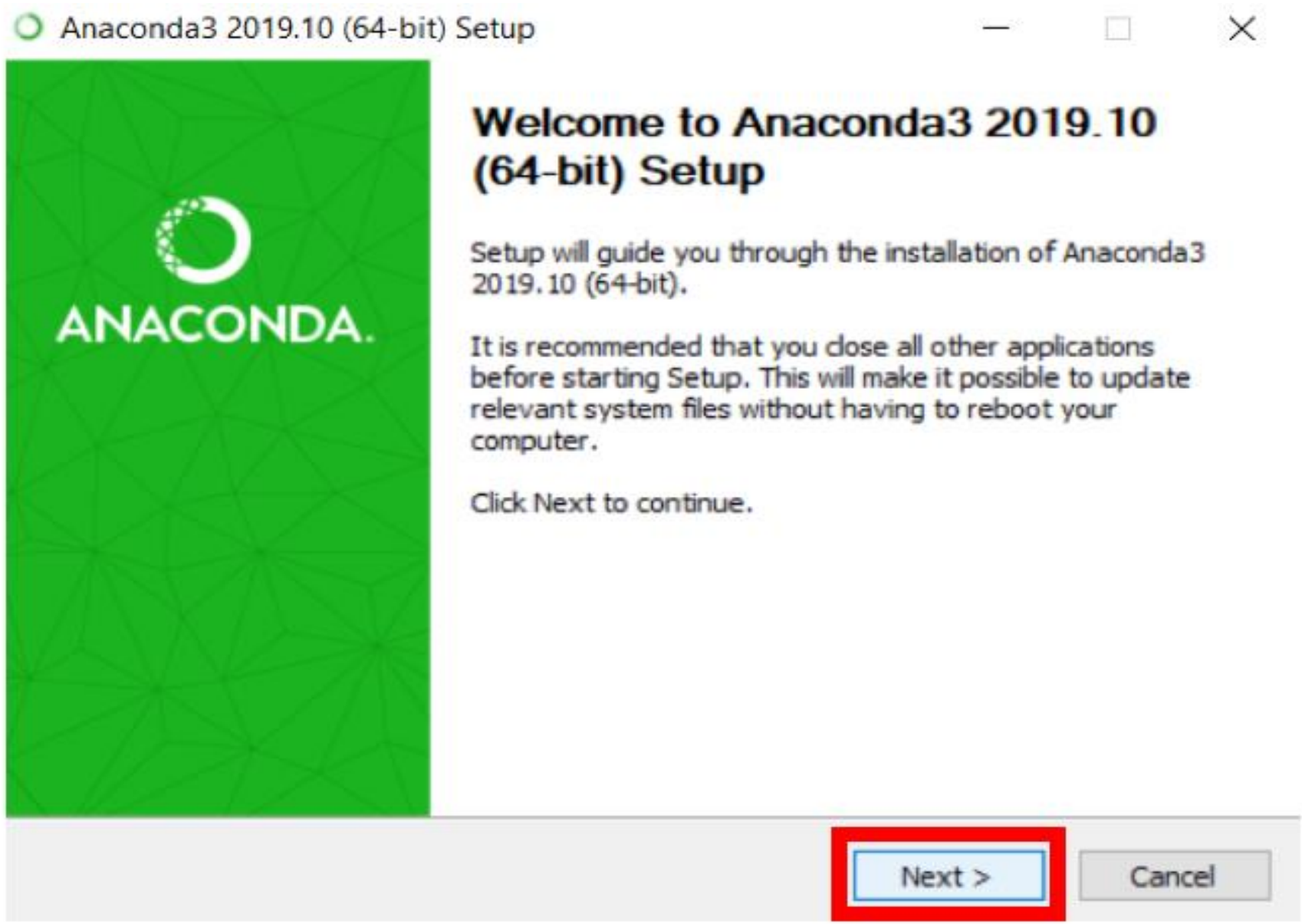
Anaconda is a package manager and Python distribution that contains a collection of many open source packages (**numpy, scikit-learn, scipy, pandas to name a few**).

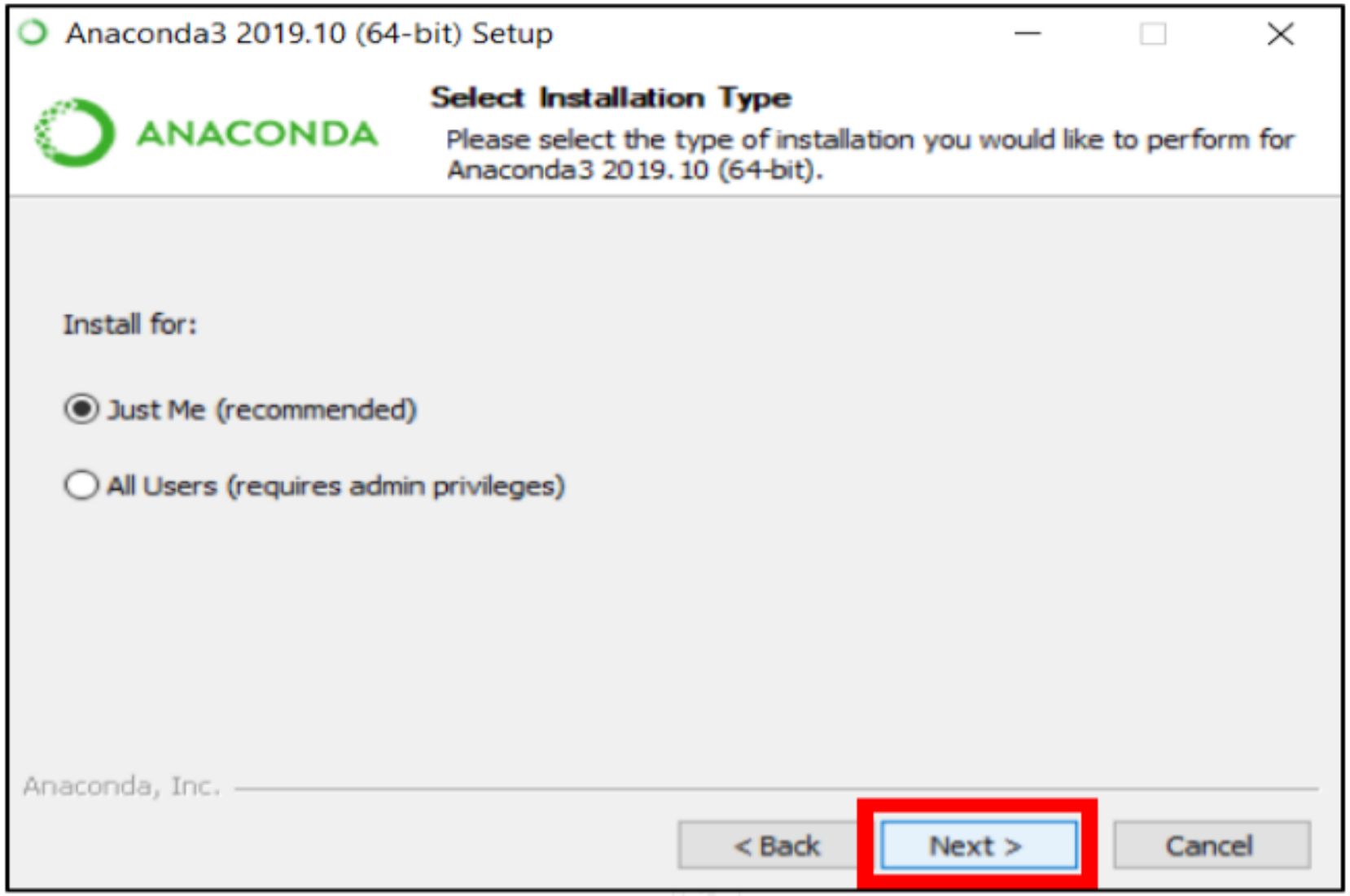
Installation steps (Windows):

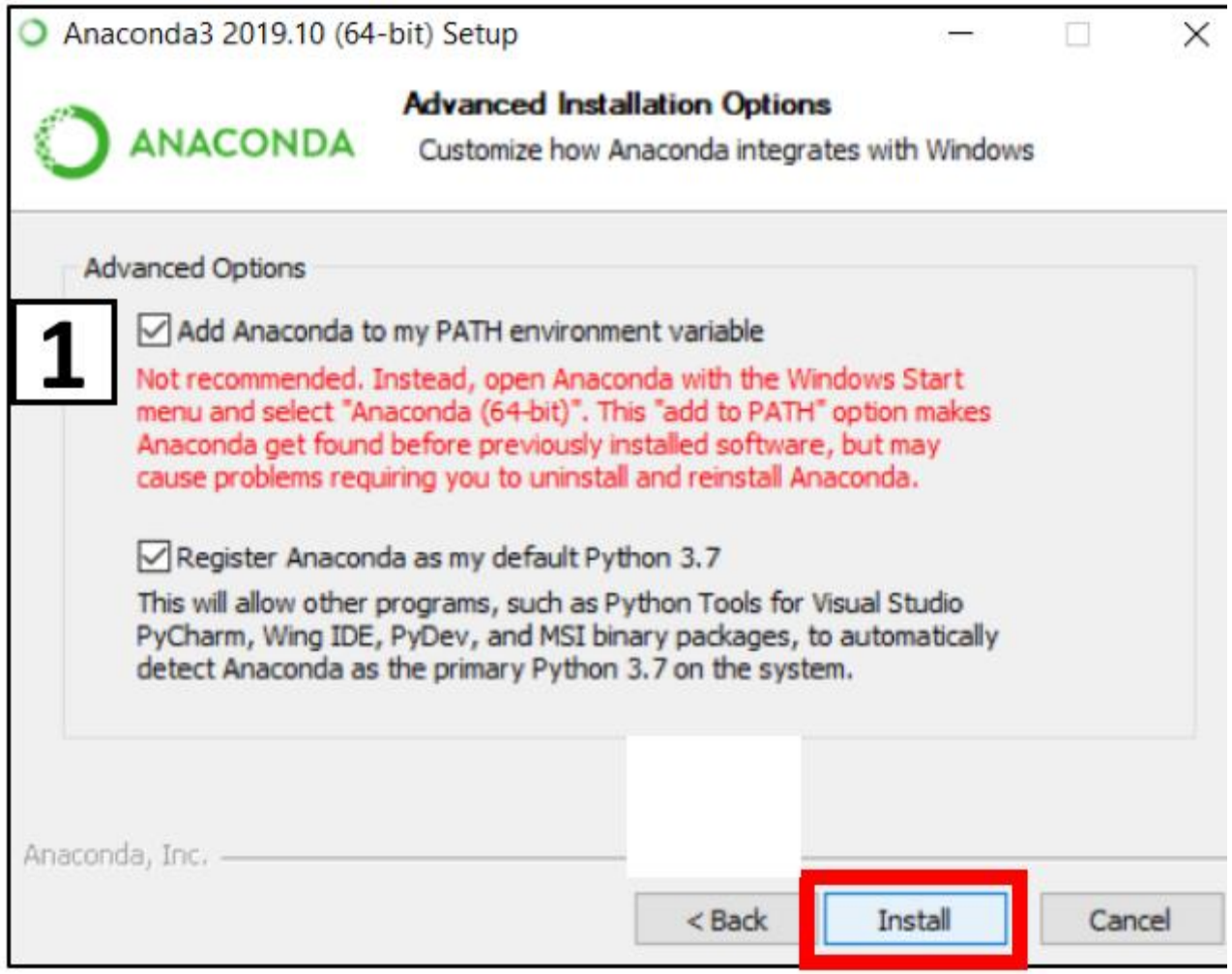
- Go to the Anaconda Website (https://www.anaconda.com/products/individual#windows)

- **Select →** 64-Bit Graphical Installer (466 MB)

- Run the setup file as an administrator


Anaconda3 2019.10 (64-bit) Setup
ANACONDA.
Welcome to Anaconda3 2019.10 (64-bit) Setup
Setup will guide you through the installation of Anaconda3 2019.10 (64-bit).
It is recommended that you close all other applications before starting Setup. This will make it possible to update relevant system files without having to reboot your computer.
Click Next to continue.
Next >
Cancel


Anaconda3 2019.10 (64-bit) Setup
ANACONDA
Select Installation Type
Please select the type of installation you would like to perform for Anaconda3 2019.10 (64-bit).
Install for:
Just Me (recommended)
All Users (requires admin privileges)
Anaconda, Inc.
< Back
Next >
Cancel


Anaconda3 2019.10 (64-bit) Setup
ANACONDA
Advanced Installation Options
Customize how Anaconda integrates with Windows
Advanced Options
1
Add Anaconda to my PATH environment variable
Not recommended. Instead, open Anaconda with the Windows Start menu and select "Anaconda (64-bit)". This "add to PATH" option makes Anaconda get found before previously installed software, but may cause problems requiring you to uninstall and reinstall Anaconda.
Register Anaconda as my default Python 3.7
This will allow other programs, such as Python Tools for Visual Studio PyCharm, Wing IDE, PyDev, and MSI binary packages, to automatically detect Anaconda as the primary Python 3.7 on the system.
Anaconda, Inc.
< Back
Install
Cancel

Launch Jupyter Notebook