# Piscine Pro AI / Machine Learning

## Training tools

*Summary:*  *In this document, you will discover the tools you are about to use during your training.*

*Version: 1.00*

# Contents

# Chapter I

# A word about this Piscine Pro

Welcome !

You will start this Machine Learning Piscine Pro. We both want you to discover what the artificial intelligence is and to experience peer-learning, the educational model of 42.

Rather than giving you a course with only one solution for each problem, and which will probably be outdated in a few years, we have chosen to put you in a peer-learning situation. You are going to look for the elements that could serve you for your challenge, sort out those that are actually interesting by testing and manipulating them. To do this, discuss with others, exchange your points of view, find new ideas together, and finally test by yourself.

Peer-evaluation is a key moment to discover other ways of doing things, as well as special cases that you have not thought of and that could undermine your program. Like different clients who don't pay attention to the same things, each reviewer will be different from the last. And who knows, you might have made new acquaintances for later collaborations.

At the end of this cursus, you will not have done the same things as the other participants, you will not have validated the same projects, you will have chosen to do one challenge rather than another... .and that's normal! It's both a collective and a personal experience. Everyone will benefit from what he or she experiences during this time.

Good luck to all.

# Chapter II

# Tools

For this course we'll be using [colab](), it's an online tool with everything already pre-installed.

You may already have Anaconda installed on your local computer and be using the Jupyter Notebook. If so, you'll quickly learn how to use Google Colab. Colab is a free Jupyter Notebook environment that runs entirely in the cloud. There's no configuration process. The notepads you create can be edited simultaneously by your team members, just as you edit documents in Google Docs. Colab supports many popular machine learning libraries.

Google Colab already has pre-installed libraries such as pandas, numpy, Tensorflow, Keras, OpenCV, so you don't need to run "pip install" yourself.
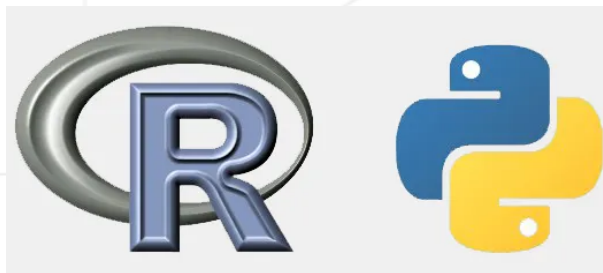
You can start by following the tutorials on the home page, which will help you understand how the tools we'll be using throughout the course work.

> ⚠️ You need a google mail to use colab

# Chapter III

# Programming languages

In the field of data science and machine learning, python3 and R are the most popular languages.

It's not a question of pigeonholing either language into a single category: Python can be used effectively as a data analysis tool, and R is flexible enough to do good work in machine learning. There are plenty of packages for both languages that seek to replicate each other's functionality. Python has libraries to enhance its statistical inference capability, and R has packages to improve its predictive accuracy.

For which you can take the course in the language of your choice.

In the edit tab, then in the parameters, you can choose your programming language, python3 or R.
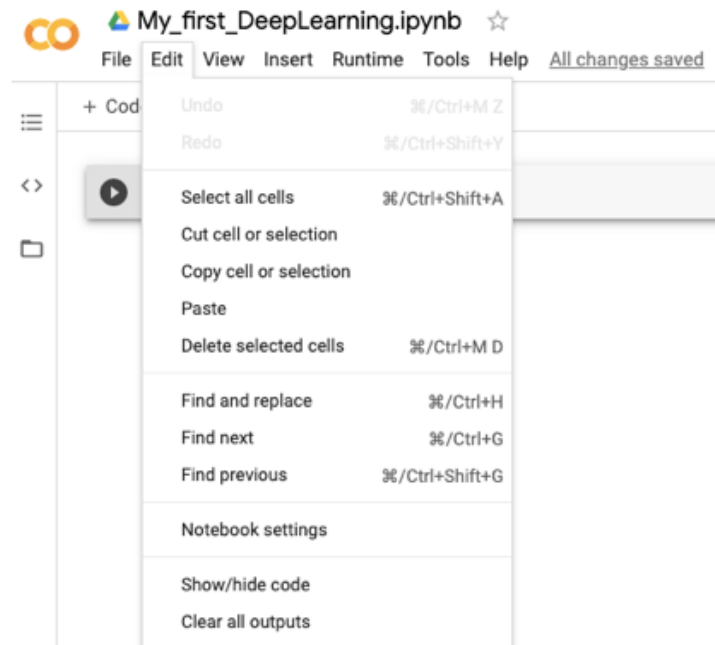
# Chapter IV

# GPU Activation

Enable GPU computing (for PyTorch or Tensorflow for exemple):
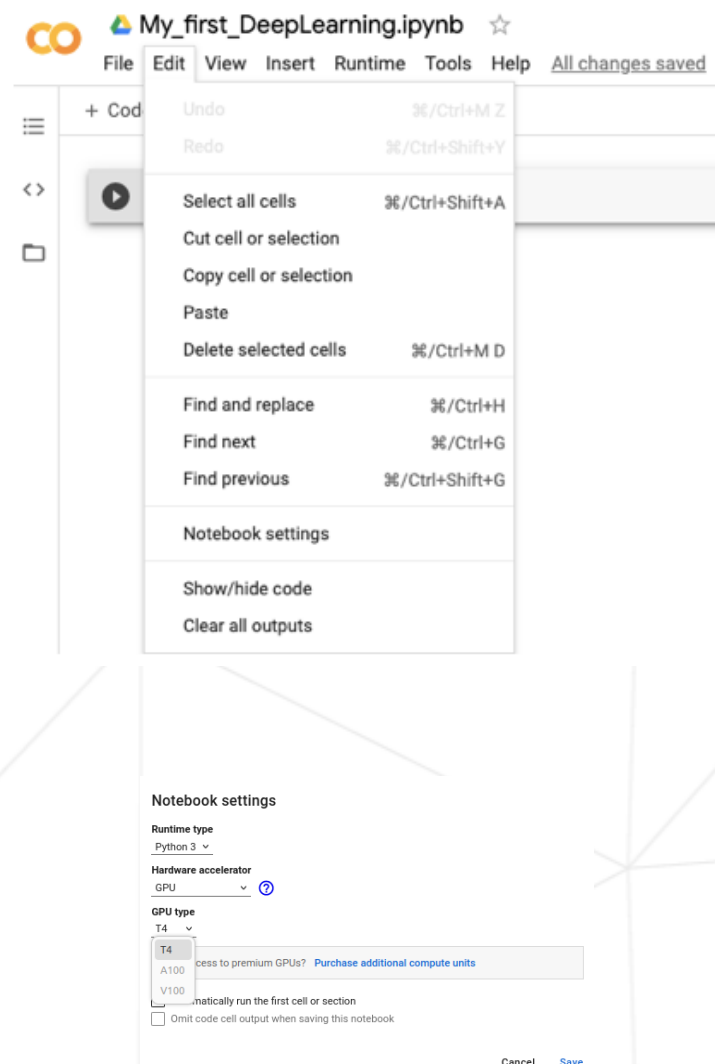


In the same Notebook Settings, you can activate the GPU or TPU activation, which gives you free calculation units, which you can then upgrade to a paid version (not in this training) if you need more calculation power.

# Chapter V

# Load sources in colab

## V.1   Upload Notebook

In each mold, you'll have an ipynb file to guide you. (and possibly CSV data files) You'll need to download it from intra then load it into colab.
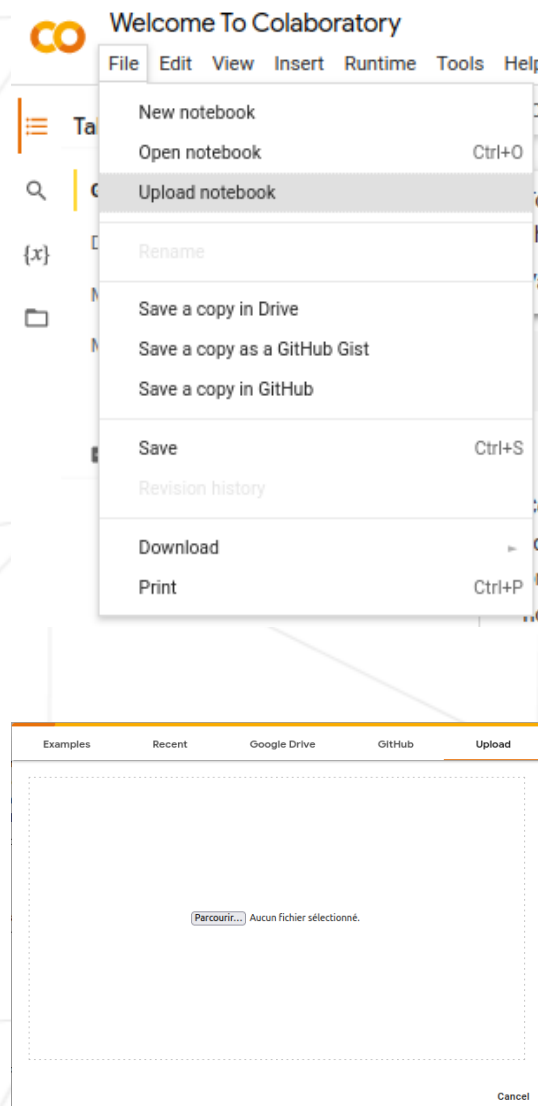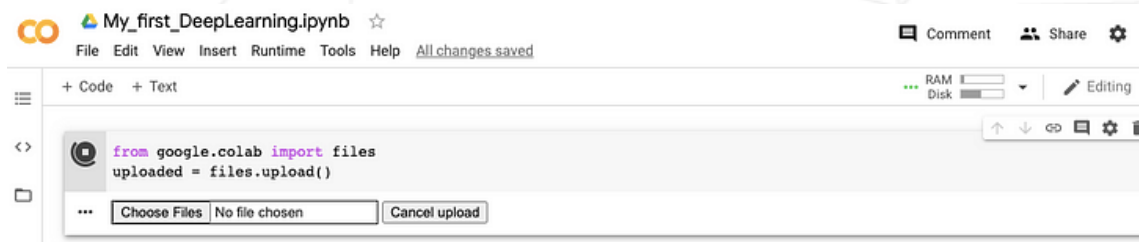
## V.2   Upload Sources

If you have CSV files to import, you need to copy the code below and put it in a colab cell.

```
from google.colab import files
uploaded = files.upload()
```





Don't forget to click on "Choose files", otherwise the cell will load indefinitely
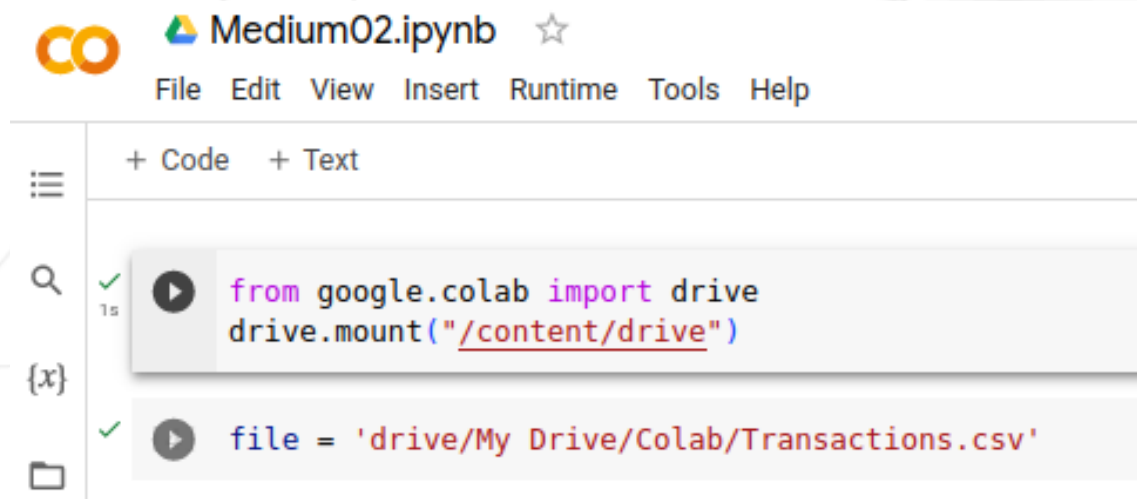


You can use shift + enter to execute the cells

# V.3  Upload Heavy Sources

To speed up the process with large CSV files, you can upload your file to Google Drive and then access it directly in Colab using the following method:

```
from google.colab import drive
drive.mount("/content/drive")
```





You need to establish a connection between your Google Drive and Colab.