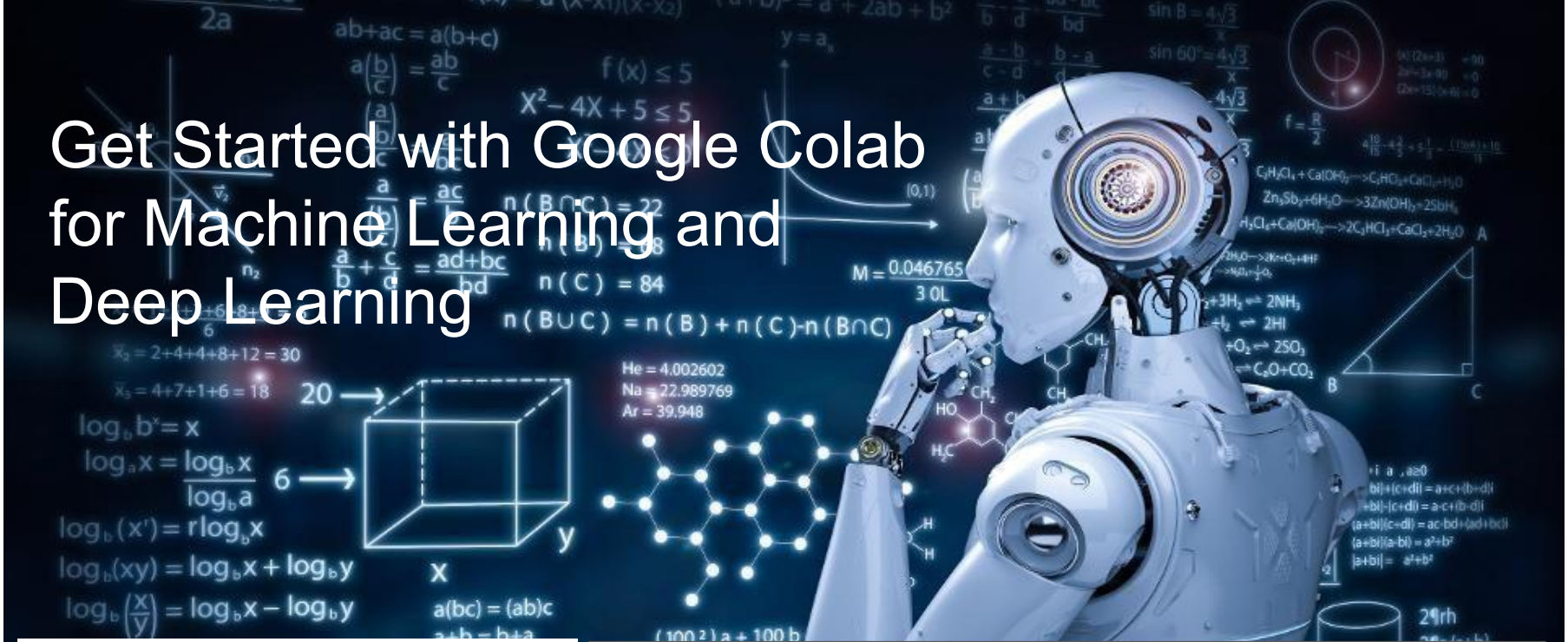
Get Started with Google Colab for Machine Learning and Deep Learning



Asep Muhidin, S.Kom, M.Kom.
asep.muhidin@pelitabangsa.ac.id

Webinar Teknik Informatika
Universitas Pelita bangsa

“Memory Error” – that all too familiar dreaded message in Jupyter notebooks when we try to execute a machine learning or deep learning algorithm on a large dataset. Most of us do not have access to unlimited computational power on our machines.

So how do we build large deep learning models without burning a hole in our pockets?

Google Colab! It's an incredible online browser-based platform that allows us to train our models on machines for free

What is Google Colab?

Google Colaboratory is a free online cloud-based Jupyter notebook environment that allows us to train our machine learning and deep learning models on CPUs, GPUs, and TPUs.

GPUs and TPUs on Google Colab

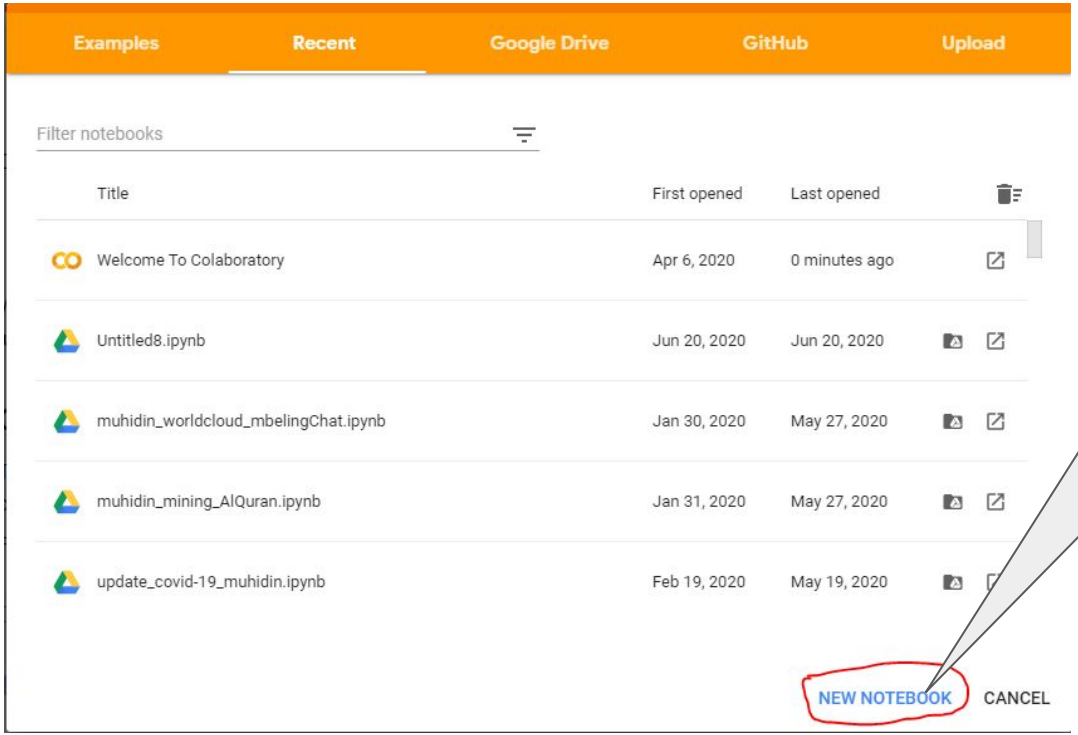
Google Colab gives us three types of runtime for our notebooks: CPUs,GPUs, and TPUs

CPU	GPU	TPU
Intel Xeon Processor with two cores @ 2.30 GHz and 13 GB RAM	Up to Tesla K80 with 12 GB of GDDR5 VRAM, Intel Xeon Processor with two cores @ 2.20 GHz and 13 GB RAM	Cloud TPU with 180 teraflops of computation, Intel Xeon Processor with two cores @ 2.30 GHz and 13 GB RAM

Colab gives us 12 hours of continuous execution time. After that, the whole virtual machine is cleared and we have to start again

Getting Started with Google Colab

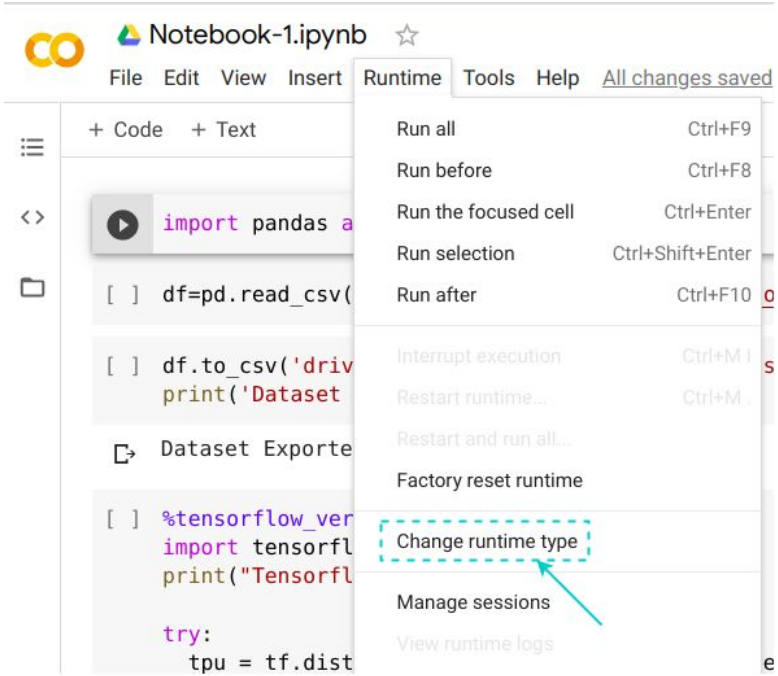
You can go to Google Colab using this link (https://colab.research.google.com/) . This is the screen you'll get when you open Colab:

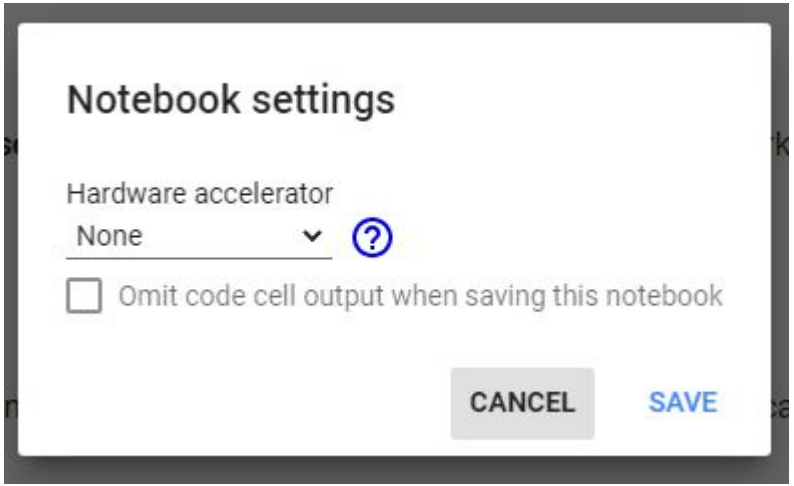


Click on the NEW NOTEBOOK button to create a new Colab notebook. You can also upload your local notebook to Colab by clicking the upload button:

Google Colab Runtimes – Choosing the GPU or TPU Option

The ability to choose different types of runtimes is what makes Colab so popular and powerful. Here are the steps to change the runtime of your notebook:





Using Terminal Commands on Google Colab

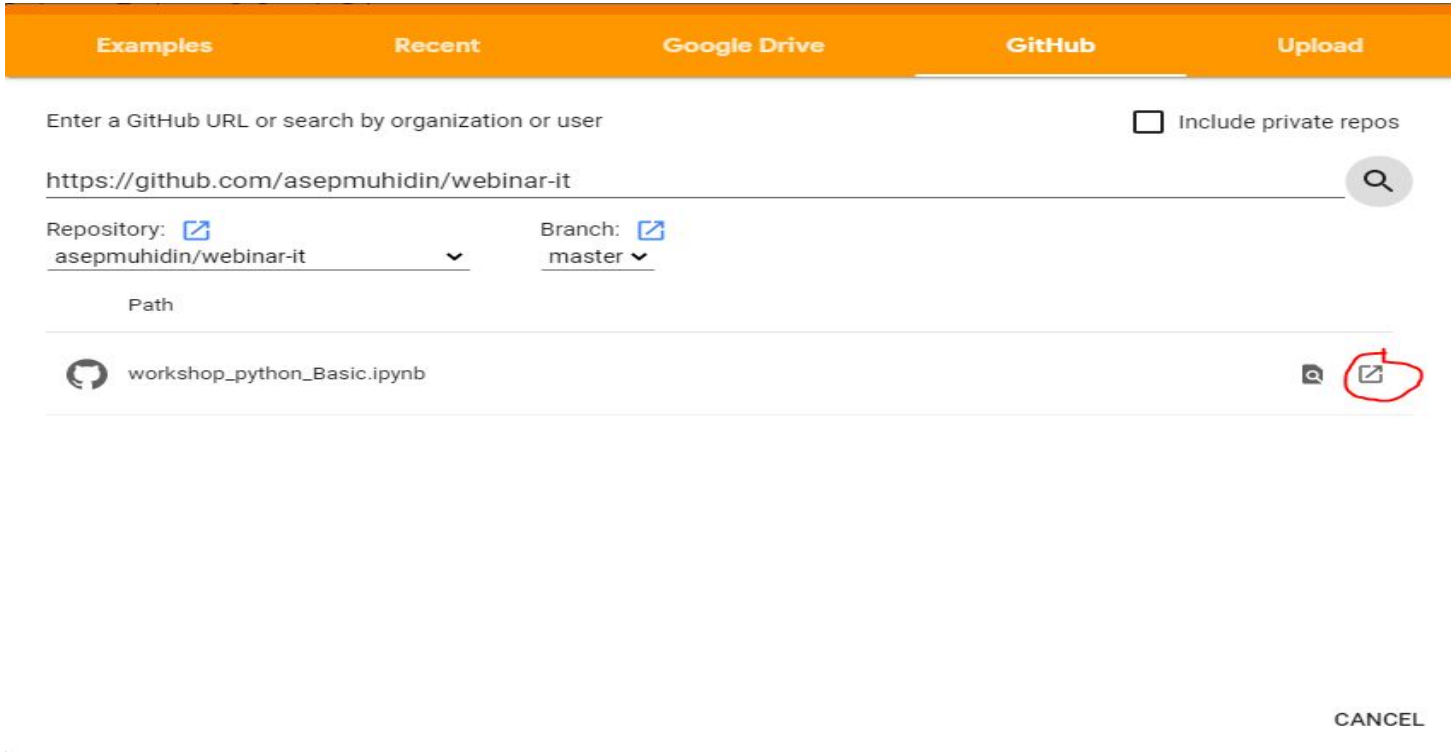
You can use the Colab cell for running terminal commands. Most of the popular libraries come installed by default on Google Colab. Yes, Python libraries like Pandas, NumPy, scikit-learn are all pre-installed.

If you want to run a different Python library, you can always install it inside your Colab notebook like this:

```
!pip install library_name
```

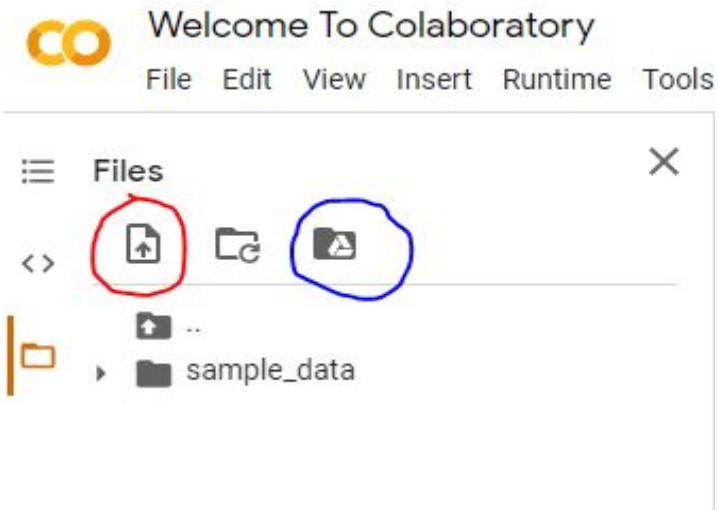
Cloning Repositories in Google Colab

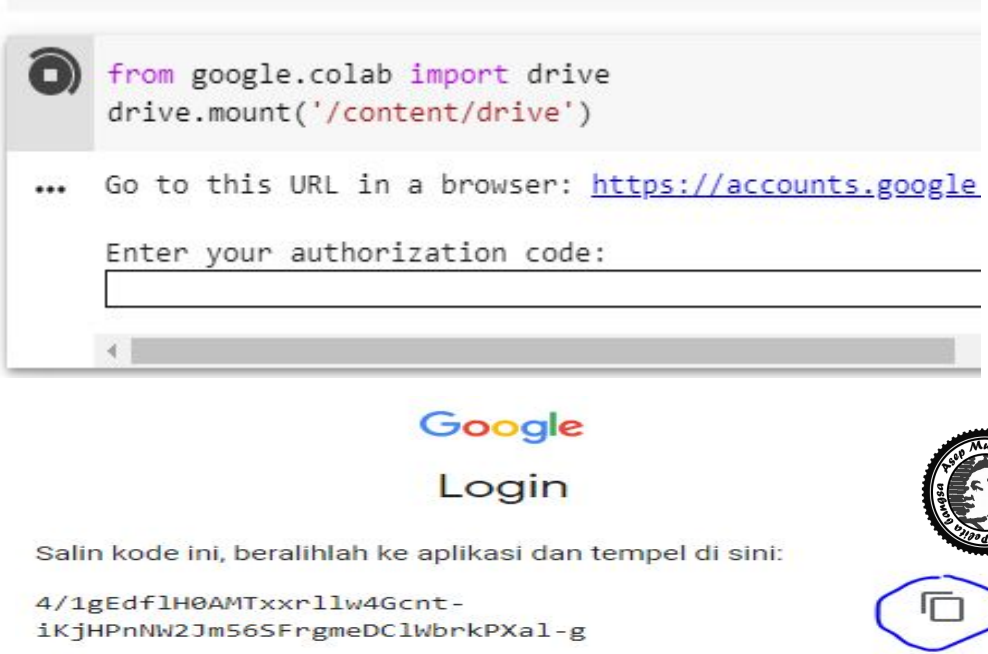
You can also clone a Git repo inside Google Colaboratory. Just go to your GitHub repository and copy the clone link of the repository:



Uploading Files and Datasets

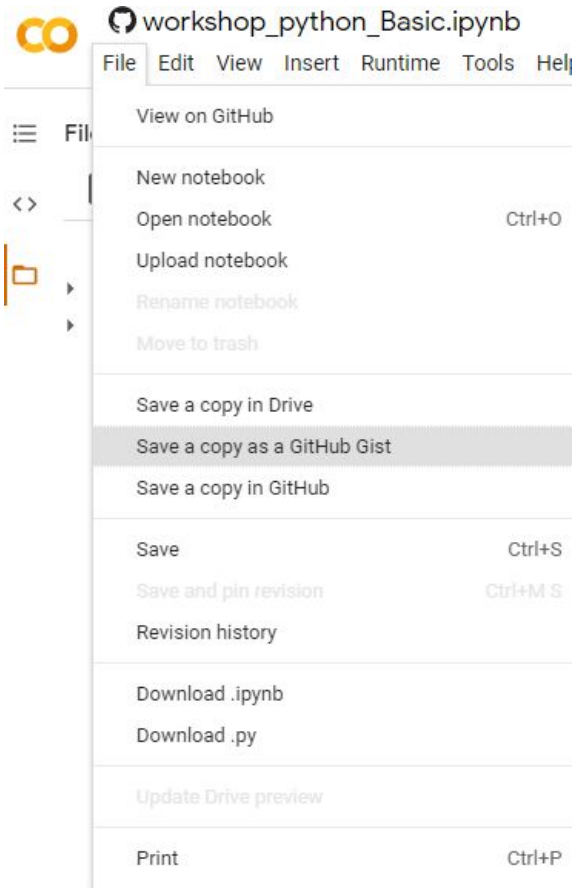
Here's a must-know aspect for any data scientist. The ability to import your dataset into Colab is the first step in your data analysis journey. The most basic approach is to upload your dataset to Colab directly. You can use this approach if your dataset or file is very small because the upload speed in this method is quite low. Another approach that I recommend is to upload your dataset to Google Drive and mount your drive on Colab.





Saving Your Notebook

All the notebooks on Colab are stored on your Google Drive. The best thing about Colab is that your notebook is automatically saved after a certain time period and you don’t lose your progress.



Exporting Data/Files from Google Colab

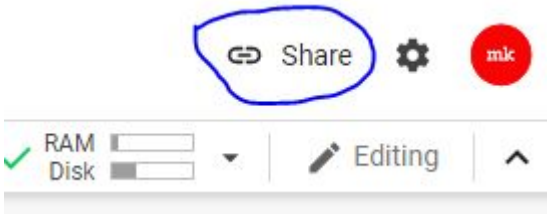
You can export your files directly to Google Drive, or you can export it to the VM instance and download it.

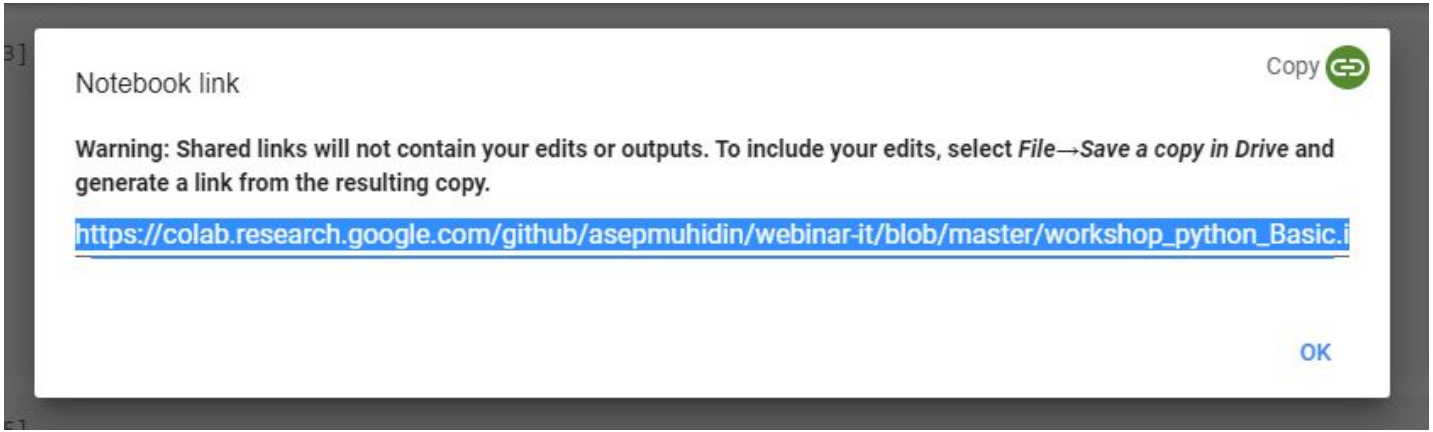
```
import pandas as pd
## nama     umur  email
#  Ahmad    24    ahmad@gmail.com
nama=['Ahmad','Susi','Joko','Cindy']
umur=[34,23,20,22]
email=['ahmad@gmail.com','sus@gmail.com','joko@gmail.com','cind@gmail.com']
participants={'nama':nama,'umur':umur,'email':email}

participants=pd.DataFrame(participants, index=None);
participants.to_excel('participans.xls')
```

Sharing Your Notebook

Google Colab also gives us an easy way of sharing our work with others.





THE END

Sources :

1. https://www.analyticsvidhya.com/blog/2020/03/google-colab-machine-learning-deep-learning/
2. https://github.com/asepmuhidin/webinar-it
3. https://towardsdatascience.com/getting-started-with-google-colab-f2fff97f594c

to be continued.. إن شاء الله