



Farielberry Lab

Deep Learning for Genomic Prediction



FACULTAD DE
INGENIERÍA



UNIVERSIDAD
DE LA REPÚBLICA
URUGUAY



Agenda & Introduction

Who are we?



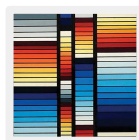
María Inés
Fariello



Federico
Lecumberry



Graciana
Castro



Farielberry Lab



FACULTAD DE
INGENIERÍA



UNIVERSIDAD
DE LA REPÚBLICA
URUGUAY



CICADA.VY
Centro Interdisciplinario en Ciencia de Datos y
Aprendizaje Automático

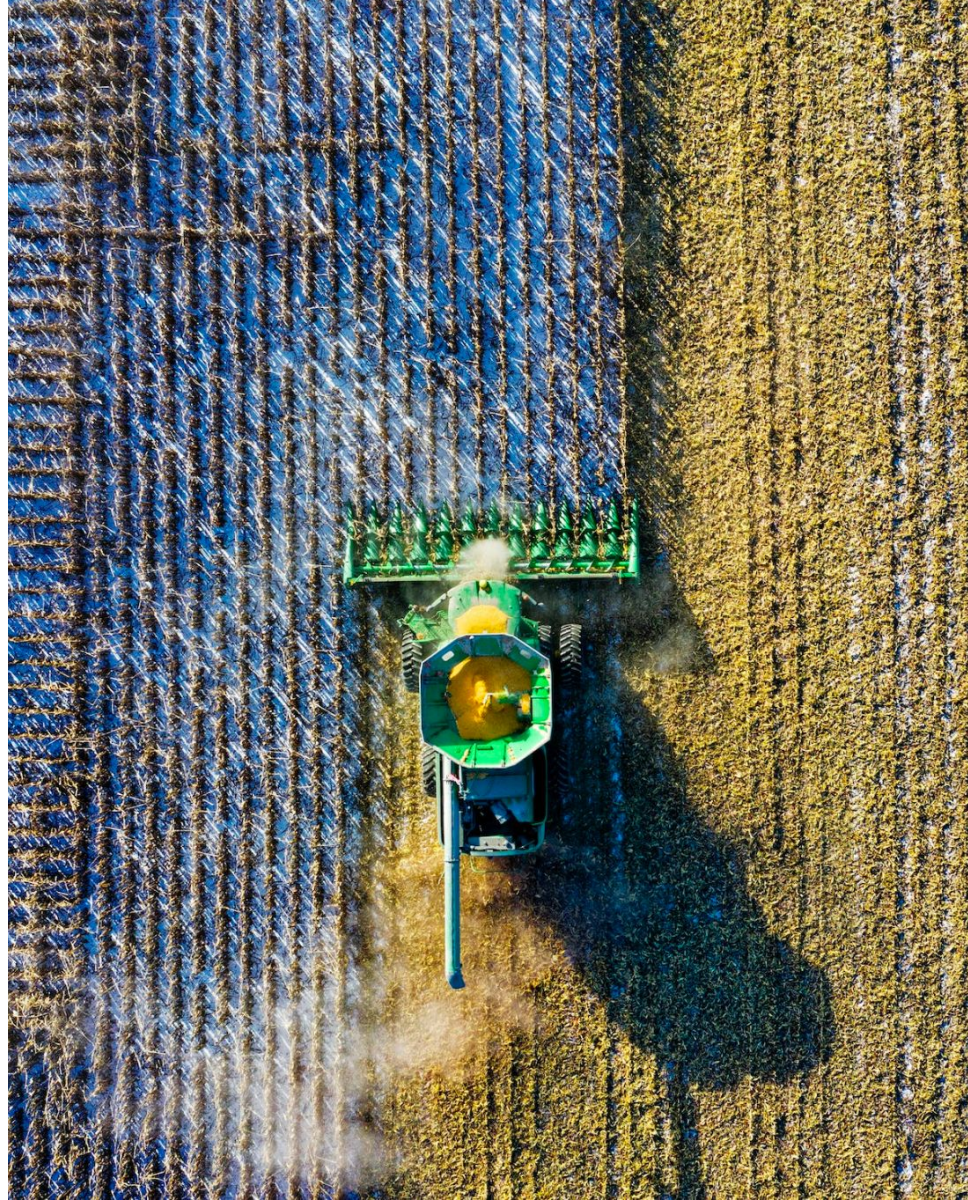


Institut Pasteur
de Montevideo



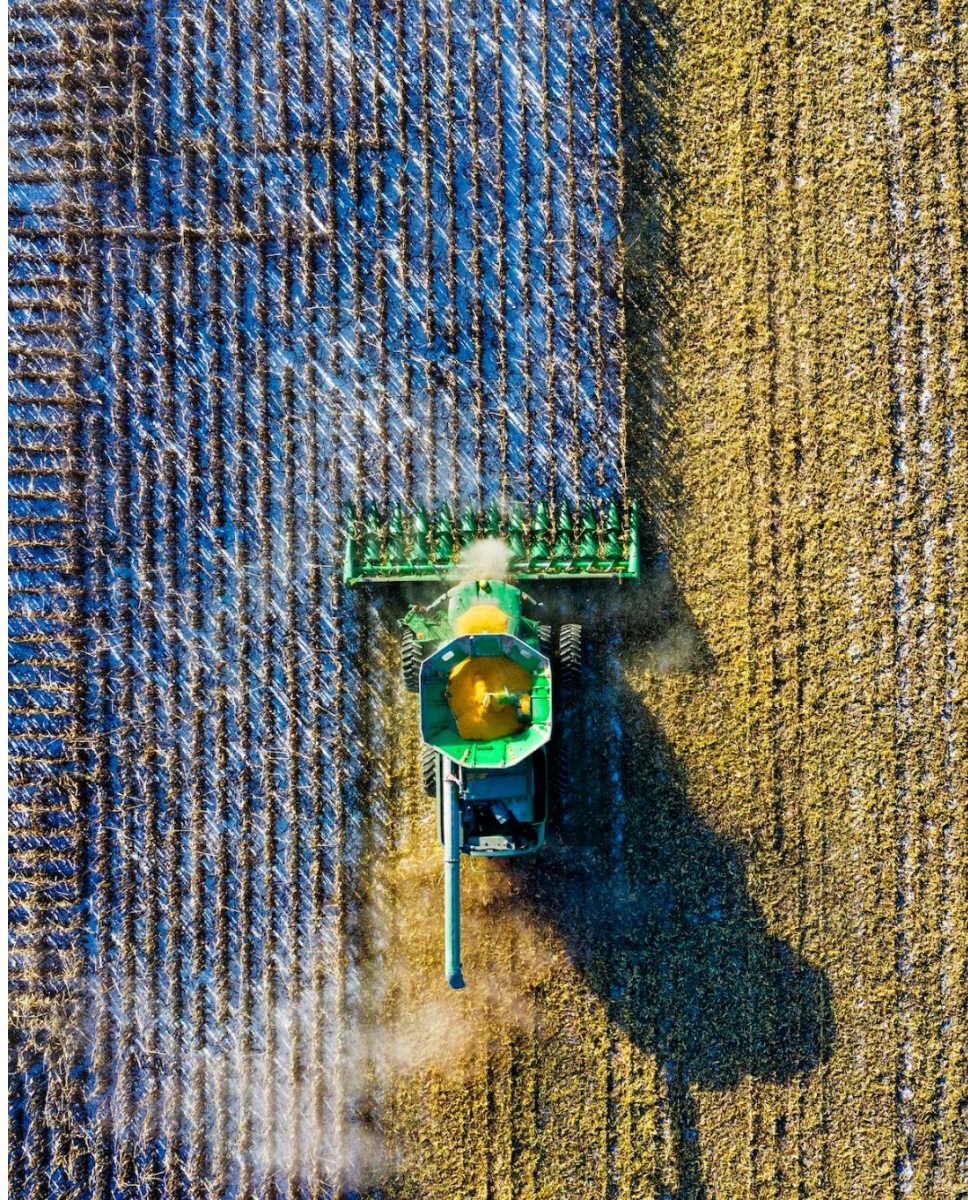
Resumen y objetivos del curso

Se introducirán los conceptos principales del Aprendizaje Profundo (DL, Deep Learning) para lo que se cubrirán conceptos de Aprendizaje Automático (ML, Machine Learning) con énfasis en su aplicación a la Predicción Genómica (PG). También se utilizarán otros dominios como la visión artificial o el lenguaje natural para presentar ejemplos ilustrativos.



Resumen y objetivos del curso

Al finalizar el curso el estudiante comprenderá los fundamentos matemáticos del Aprendizaje Profundo, tendrá experiencia en la implementación de redes neuronales en Python y Pytorch, y del análisis de sus resultados.

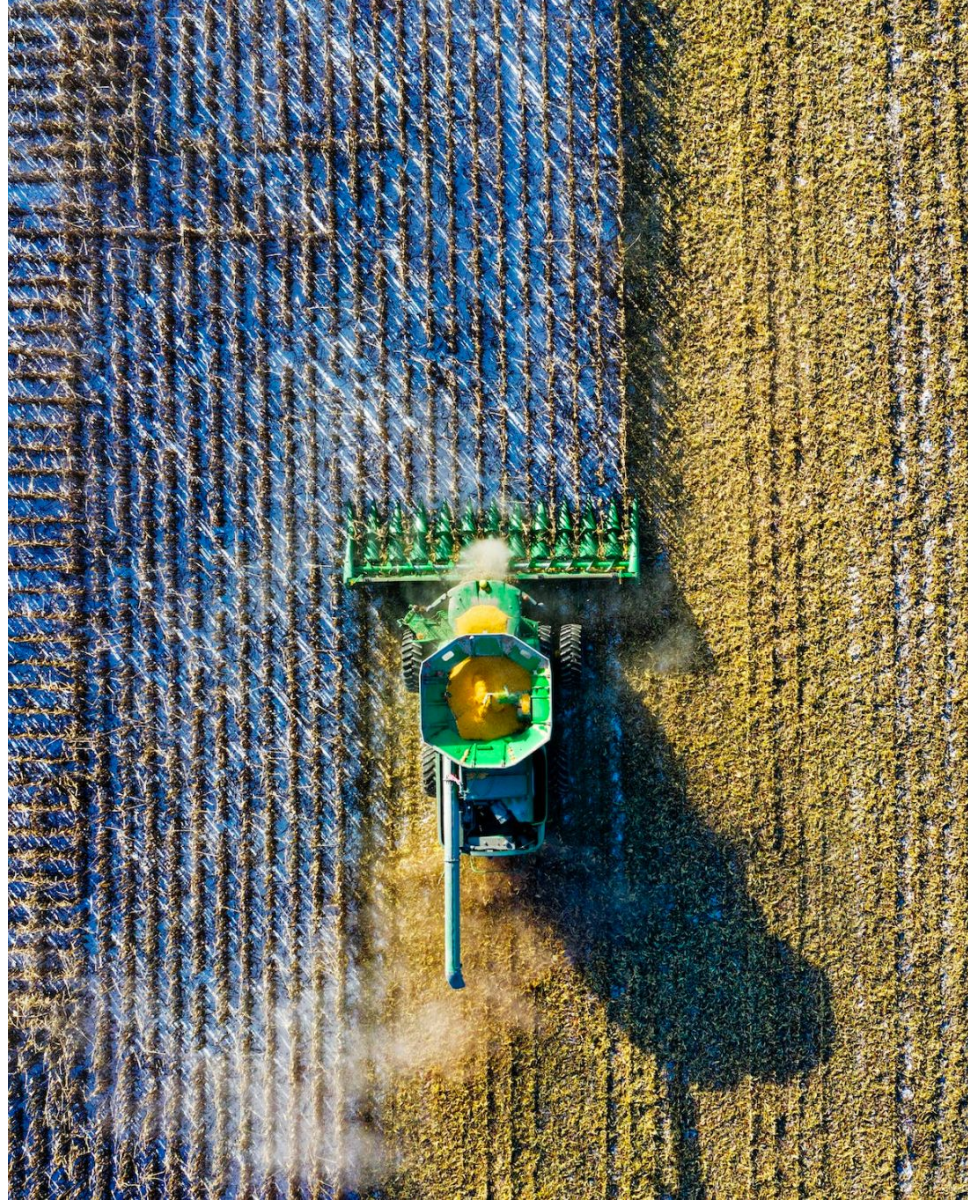


Temario



1. Intro to Genomic Prediction
2. Intro to Machine/Deep Learning
3. Convolutional Neural Networks
4. Graph Neural Networks
5. Transformers

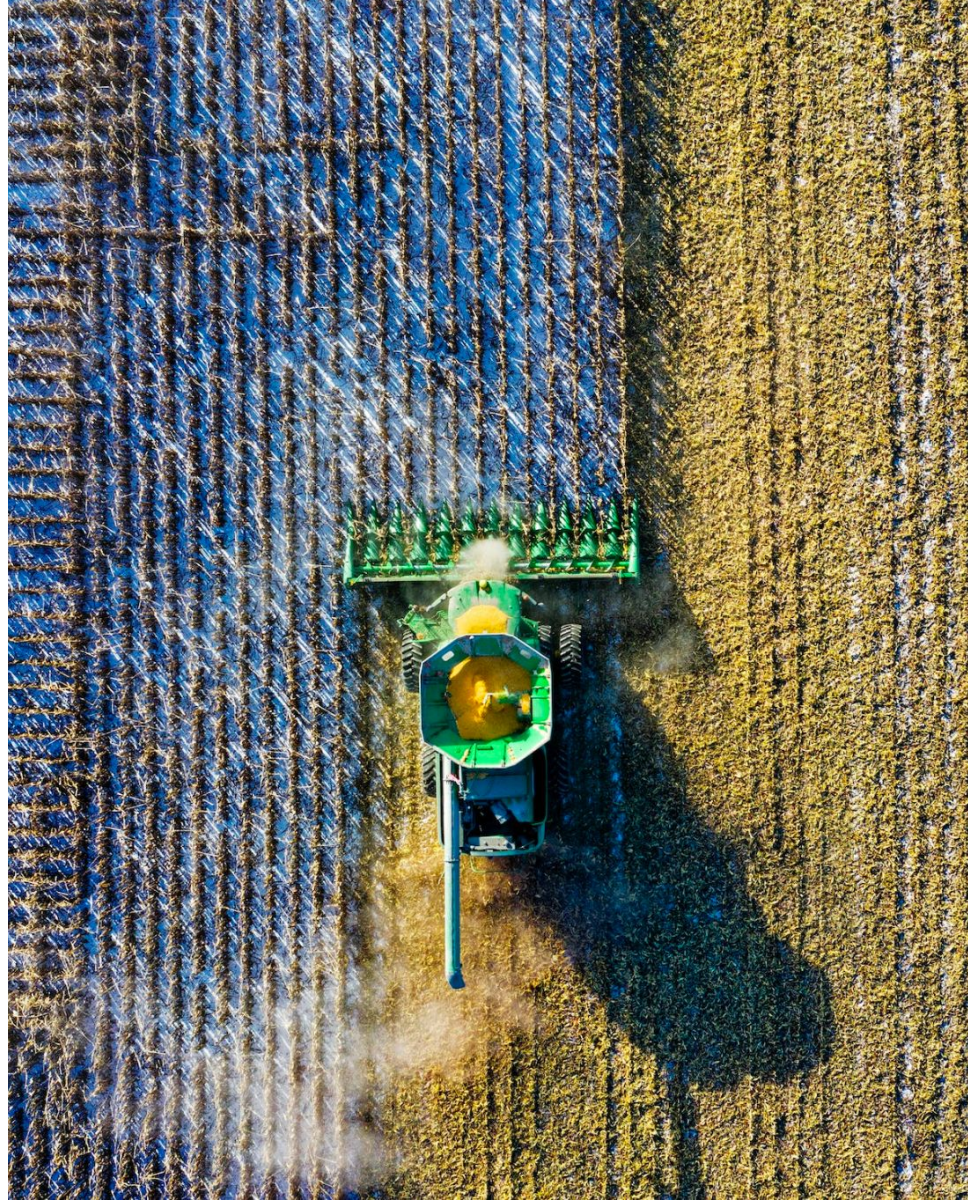
Lunes a viernes de 9:00 a 13:00.



Evaluación



A conversar...

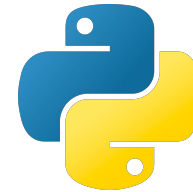




Environment setup for practicals

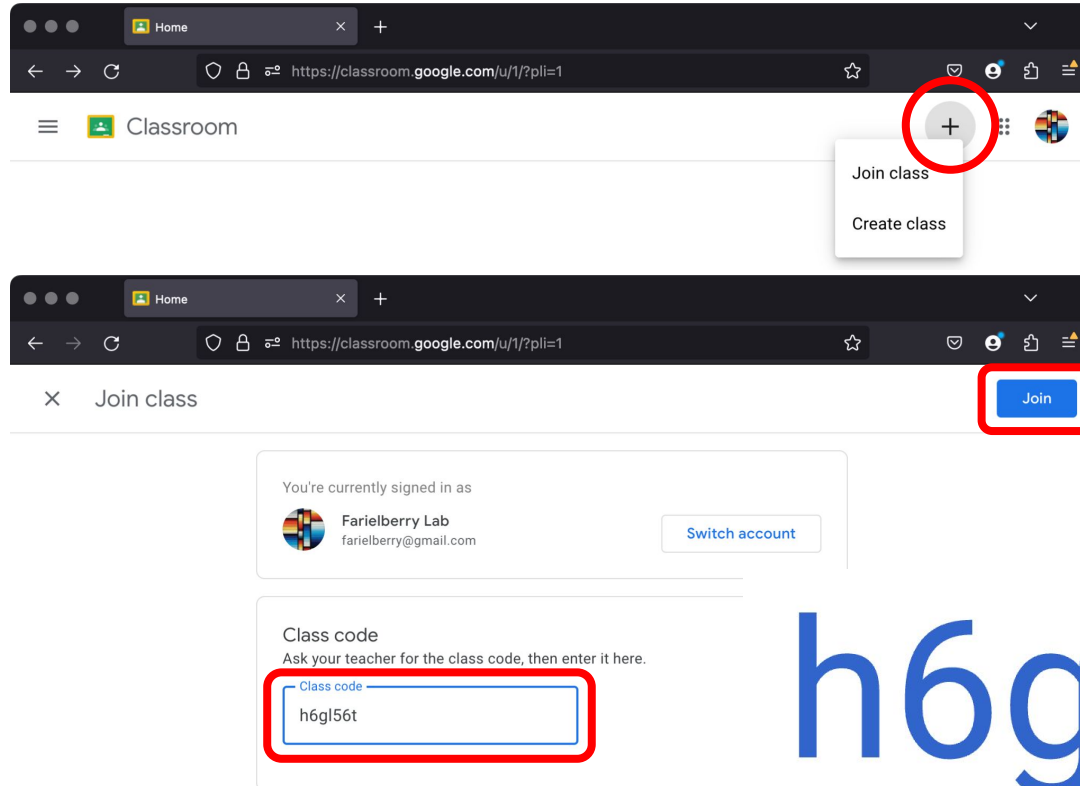
Environment setup

- Jupyter Notebook
- Google Colaboratory
- Python (numpy, matplotlib, altair, ...)
- PyTorch



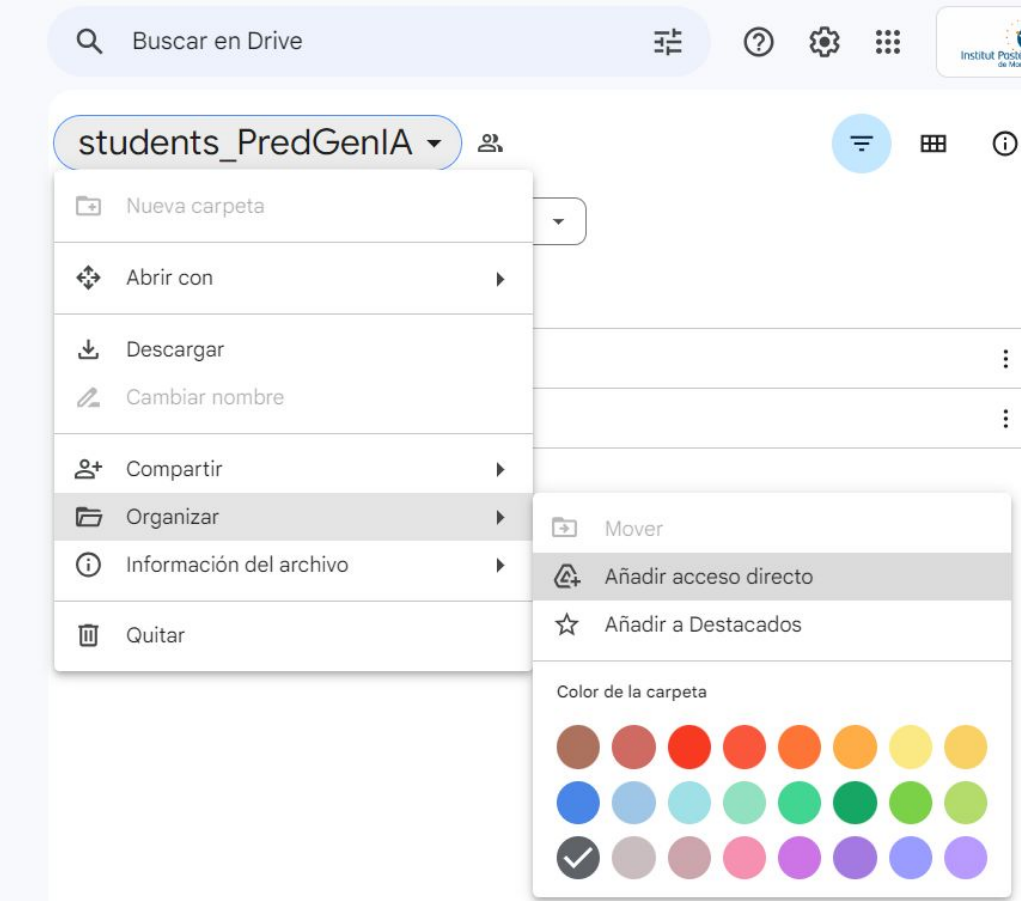
Environment setup

1. Go to <https://classroom.google.com/>
2. Join with code **h6gl56t** clicking the +

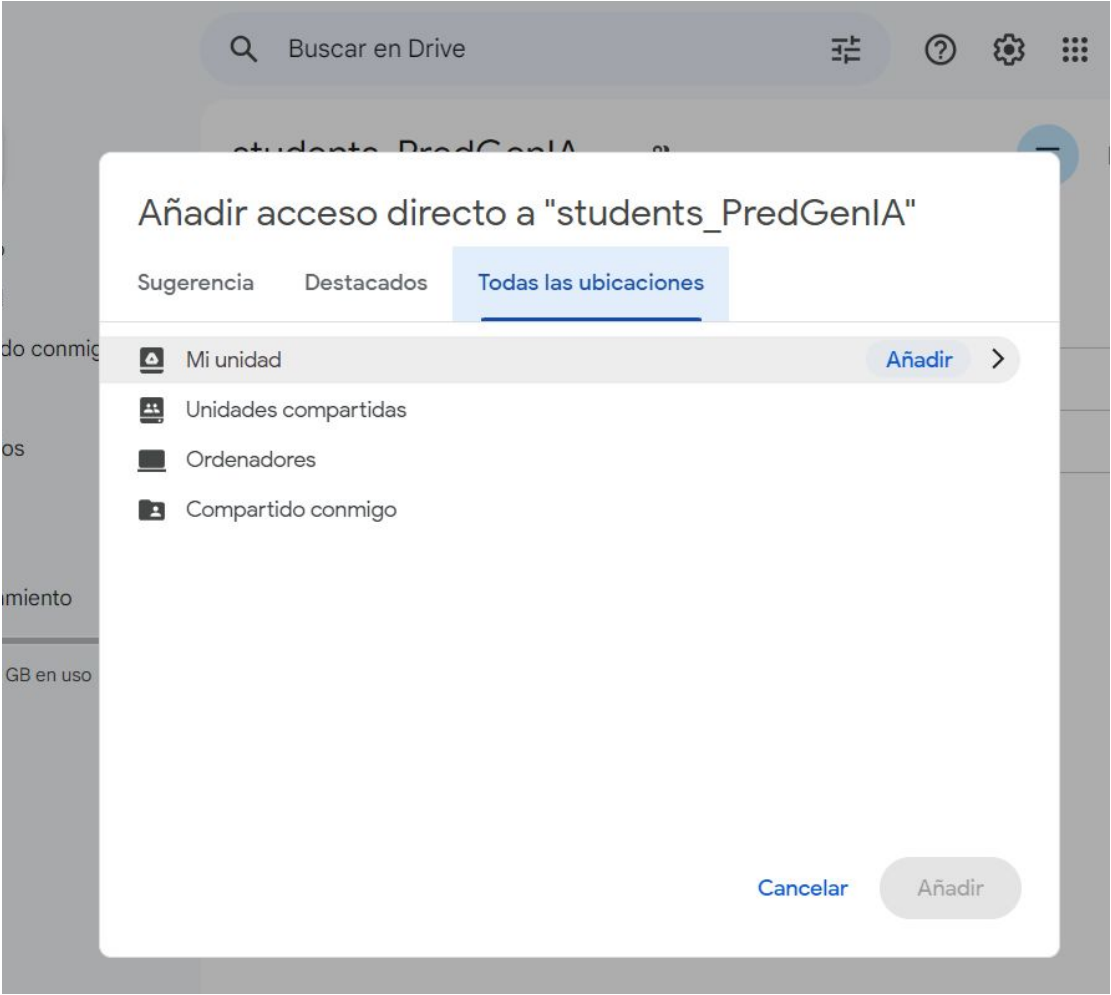


h6gl56t

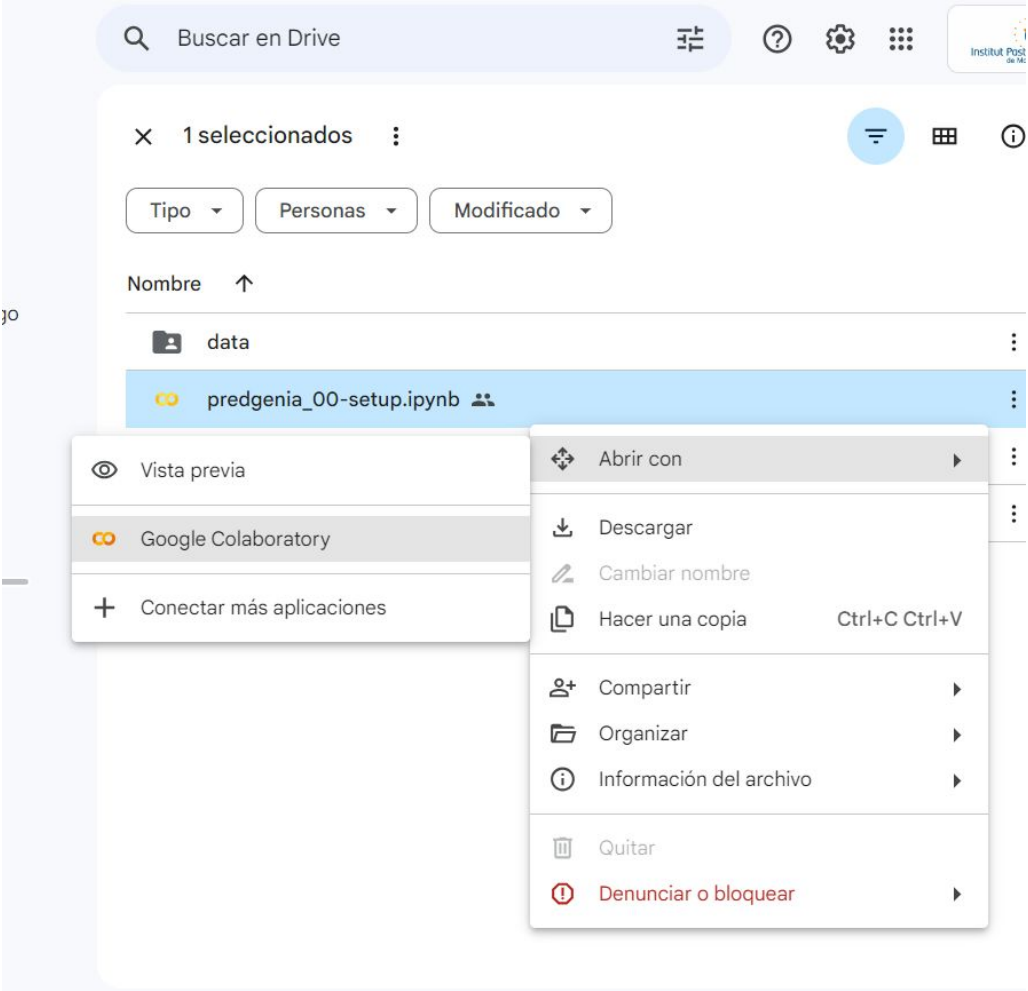
Environment setup



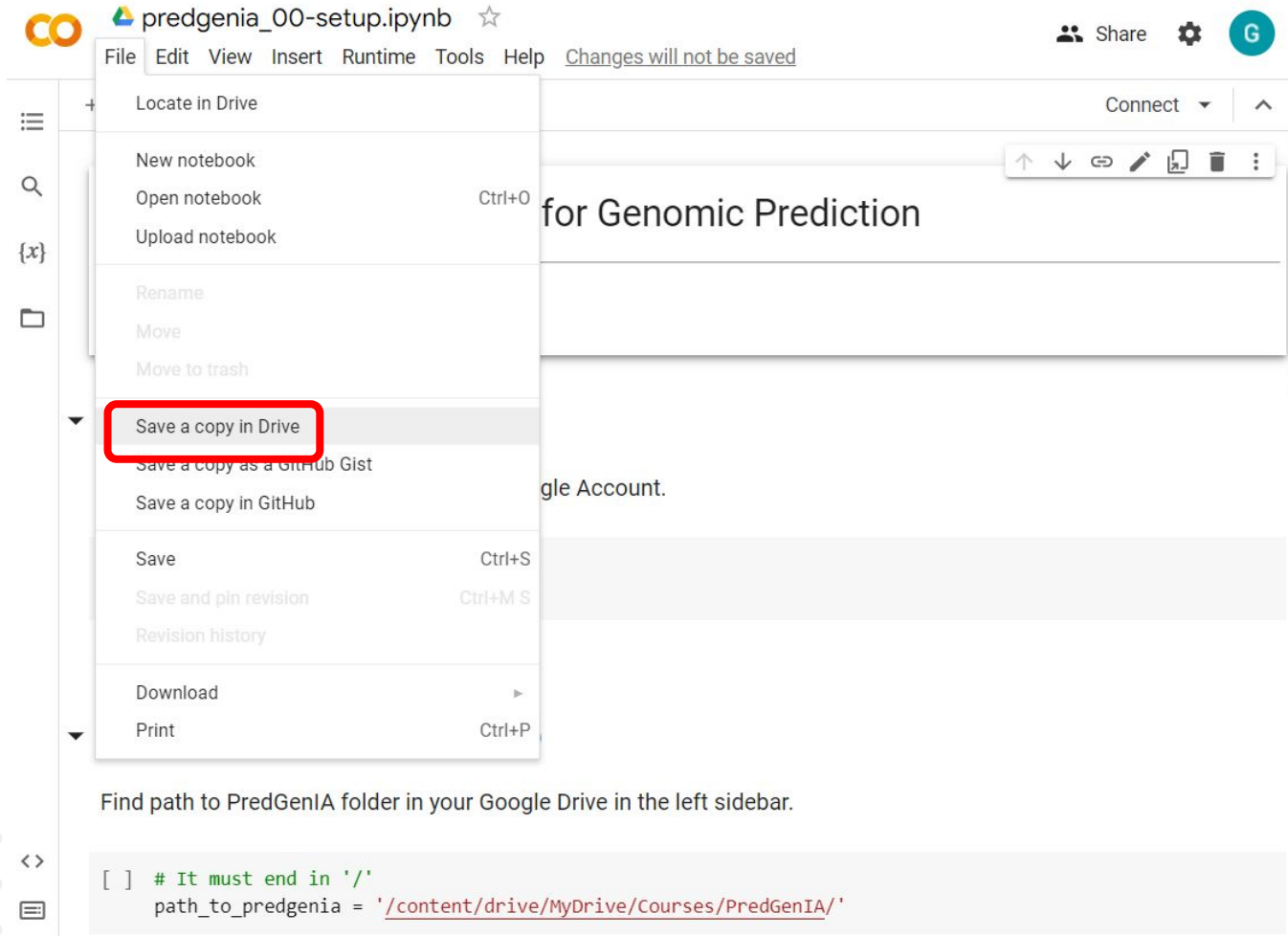
Environment setup



Environment setup



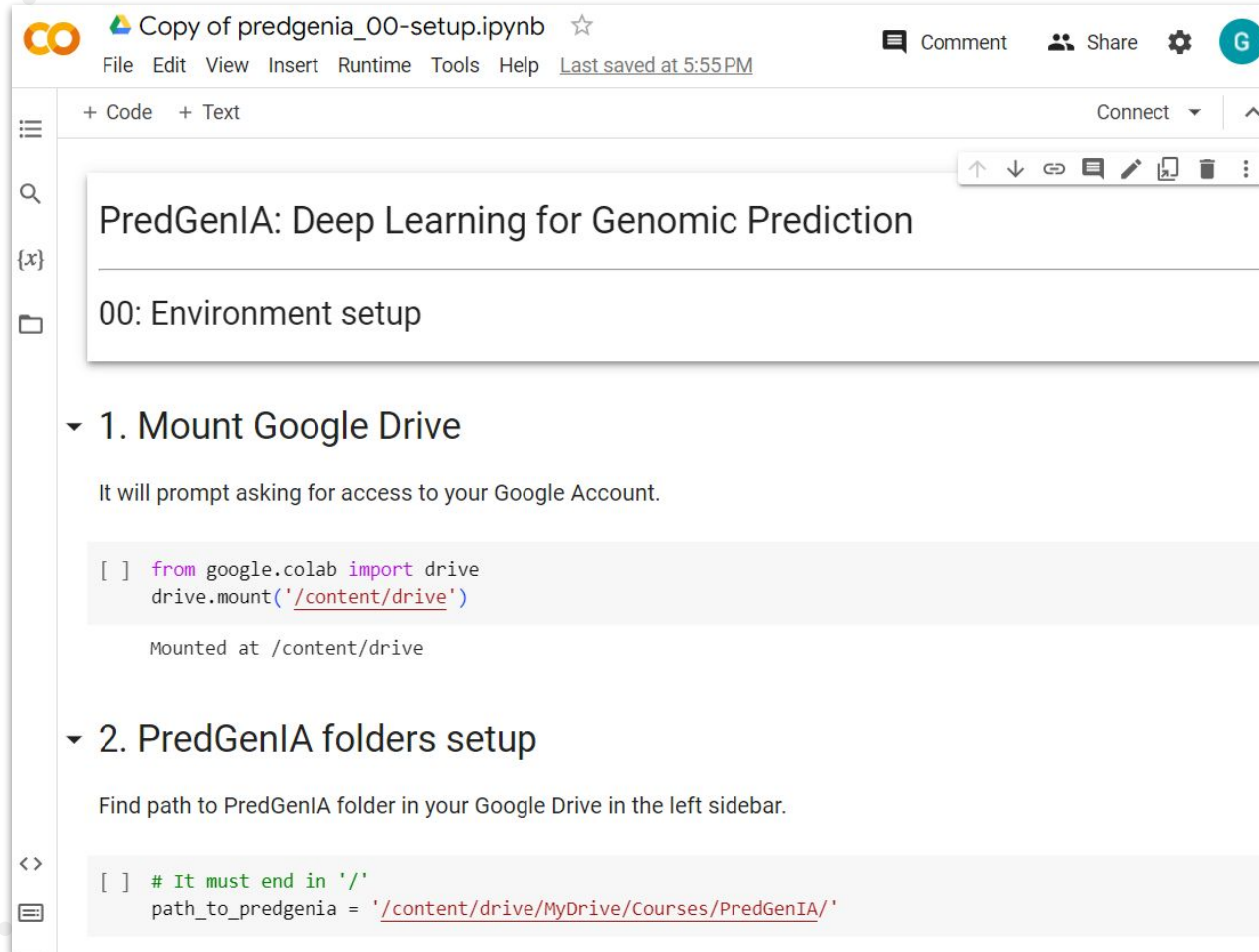
Environment setup



The screenshot shows a Jupyter Notebook titled 'predgenia_00-setup.ipynb'. The 'File' menu is open, and the option 'Save a copy in Drive' is highlighted with a red rectangle. The notebook content area displays the title 'for Genomic Prediction' and a partially visible instruction 'Find path to PredGenIA folder in your Google Drive in the left sidebar.' Below this, a code cell contains the following Python code:

```
[ ] # It must end in '/'  
path_to_predgenia = '/content/drive/MyDrive/Courses/PredGenIA/'
```

Environment setup



The screenshot shows a Google Colab notebook interface. At the top, the title bar reads "Copy of predgenia_00-setup.ipynb" with a star icon and a "Last saved at 5:55 PM" timestamp. Below the title bar is a menu bar with "File", "Edit", "View", "Insert", "Runtime", "Tools", and "Help". On the right side of the title bar are icons for "Comment", "Share", "Settings", and a user profile icon labeled "G". Below the menu bar is a toolbar with "+ Code" and "+ Text" buttons, and a "Connect" button with a dropdown arrow. The notebook content is displayed in a white box with a light gray border. It starts with a title "PredGenIA: Deep Learning for Genomic Prediction" followed by a subtitle "00: Environment setup". The first section is "1. Mount Google Drive", which includes a text prompt "It will prompt asking for access to your Google Account." and a code cell containing the following Python code:

```
[ ] from google.colab import drive
drive.mount('/content/drive')
```

 Below the code cell, the output "Mounted at /content/drive" is displayed. The second section is "2. PredGenIA folders setup", which includes a text prompt "Find path to PredGenIA folder in your Google Drive in the left sidebar." and a code cell containing the following Python code:

```
[ ] # It must end in '/'
path_to_predgenia = '/content/drive/MyDrive/Courses/PredGenIA/'
```

Copy of predgenia_00-setup.ipynb ☆
File Edit View Insert Runtime Tools Help Last saved at 5:55 PM

+ Code + Text Connect ^

PredGenIA: Deep Learning for Genomic Prediction

00: Environment setup

▼ 1. Mount Google Drive

It will prompt asking for access to your Google Account.

```
[ ] from google.colab import drive
drive.mount('/content/drive')
```

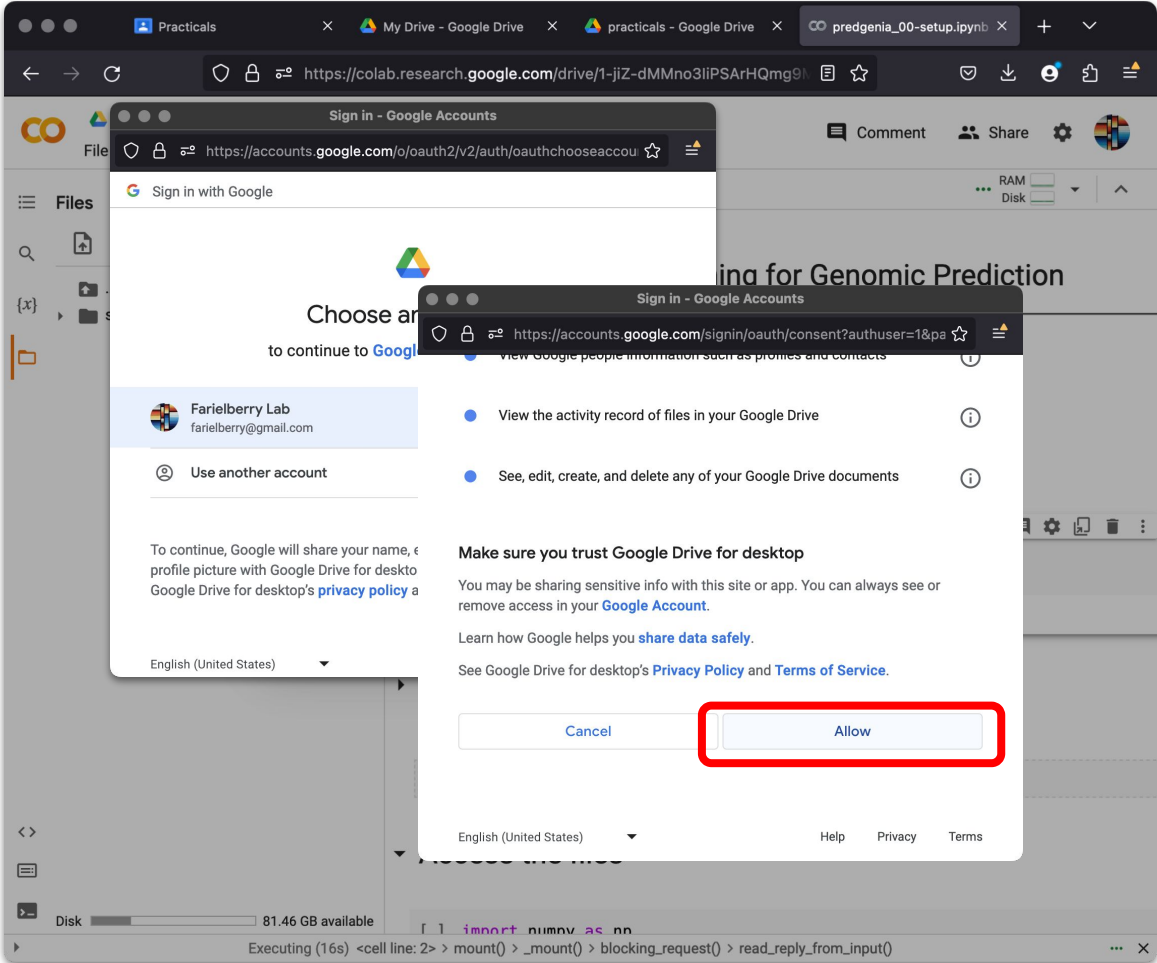
Mounted at /content/drive

▼ 2. PredGenIA folders setup

Find path to PredGenIA folder in your Google Drive in the left sidebar.

```
[ ] # It must end in '/'
path_to_predgenia = '/content/drive/MyDrive/Courses/PredGenIA/'
```

Environment setup



Environment setup

Copy of predgenia_00-setup.ipynb

File Edit View Insert Runtime Tools Help All changes saved

Comment Share

Files

- bin
- boot
- content
 - drive
 - MyDrive
 - Colab Notebooks
 - students_PredGenIA**
 - practicals
 - slides
 - 2021.09.19.460980v1.f...
 - Info útil.gdoc
 - Notebooks en Nahual....
 - Pruebas semana 28 8....
 - URU1_fgene-12-73319...
- Shareddrives
 - sample_data
 - datalab
 - dev
 - etc
 - home
 - lib
 - lib32

Disk 81.46 GB available

+ Code + Text

PredGenIA: Deep Learning for Genomic Prediction

00: Environment setup

1. Mount Google Drive

It will prompt asking for access to your Google Account.

```
[1] from google.colab import drive
drive.mount('/content/drive')
```

Mounted at /content/drive

2. PredGenIA folders setup

Find path to PredGenIA folder in your Google Drive in the left sidebar.

```
[ ] # It must end in '/'
path_to_predgenia = '/content/drive/MyDrive/Courses/PredGenIA/'
```

Environment setup

Copy of predgenia_00-setup.ipynb

File Edit View Insert Runtime Tools Help All changes saved

Comment Share

RAM Disk

Files

- bin
- boot
- content
 - drive
 - MyDrive
 - Colab Notebooks
- datalab
- dev
- etc
- home
- lib
- lib32

Upload

New file

New folder

Rename folder

Delete folder

Copy path

Refresh

PredGenIA: Deep Learning for Genomic Prediction

00: Environment setup

1. Mount Google Drive

It will prompt asking for access to your Google Account.

```
[1] from google.colab import drive
    drive.mount('/content/drive')
```

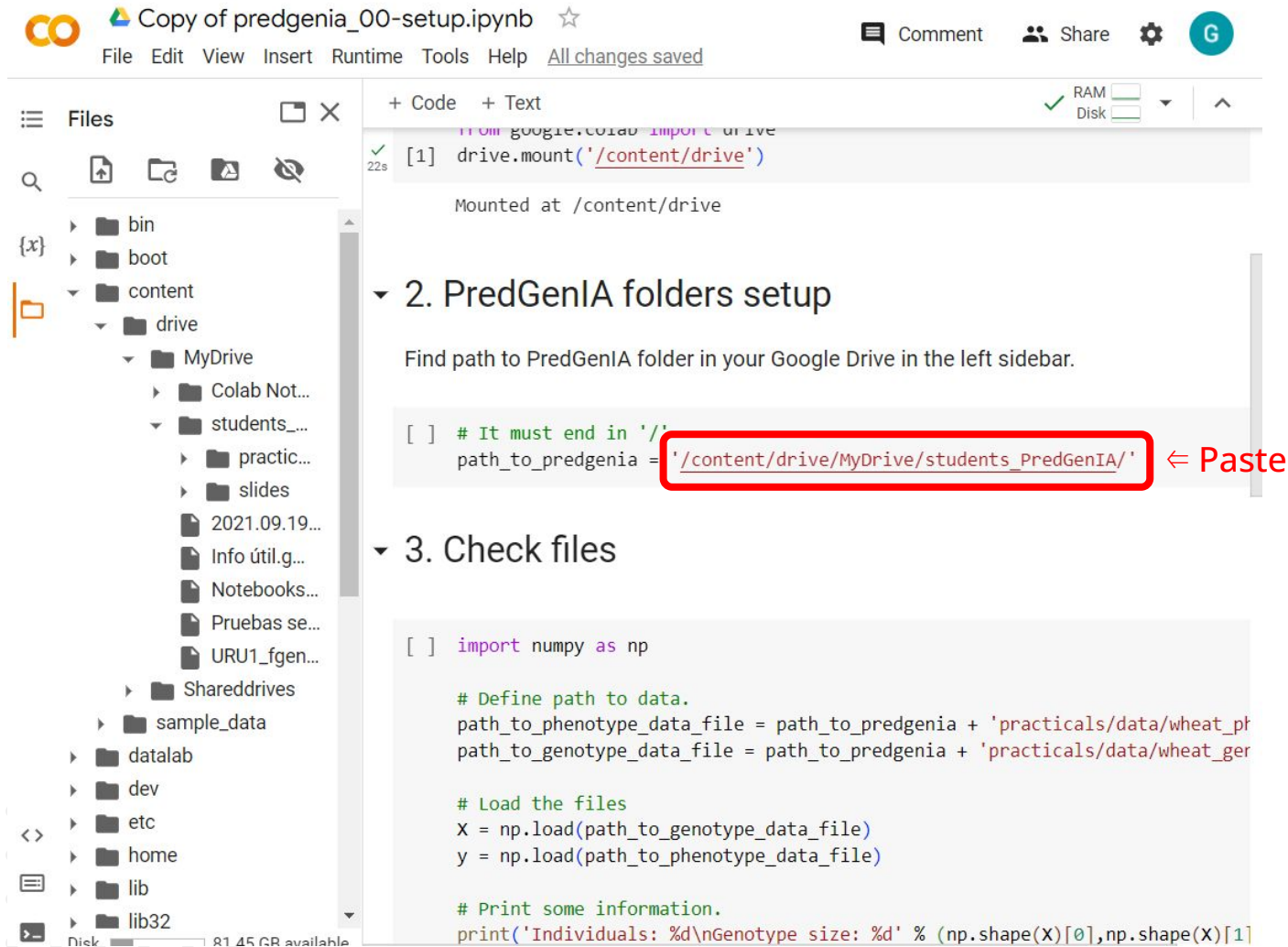
Mounted at /content/drive

2. PredGenIA folders setup

Find path to PredGenIA folder in your Google Drive in the left sidebar.

```
[ ] # It must end in '/'
    path_to_predgenia = '/content/drive/MyDrive/Courses/PredGenIA/'
```

Environment setup



Copy of predgenia_00-setup.ipynb

File Edit View Insert Runtime Tools Help All changes saved

RAM Disk

Files

- bin
- boot
- content
 - drive
 - MyDrive
 - Colab Not...
 - students_...
 - practic...
 - slides
 - 2021.09.19...
 - Info útil.g...
 - Notebooks...
 - Pruebas se...
 - URU1_fgen...
 - Shareddrives
 - sample_data
 - datalab
 - dev
 - etc
 - home
 - lib
 - lib32

81.45 GB available

2. PredGenIA folders setup

Find path to PredGenIA folder in your Google Drive in the left sidebar.

```
[ ] # It must end in '/'
path_to_predgenia = '/content/drive/MyDrive/students_PredGenIA/'
```

← Paste

3. Check files

```
[ ] import numpy as np

# Define path to data.
path_to_phenotype_data_file = path_to_predgenia + 'practicals/data/wheat_ph'
path_to_genotype_data_file = path_to_predgenia + 'practicals/data/wheat_gen'

# Load the files
X = np.load(path_to_genotype_data_file)
y = np.load(path_to_phenotype_data_file)

# Print some information.
print('Individuals: %d\nGenotype size: %d' % (np.shape(X)[0], np.shape(X)[1]))
```