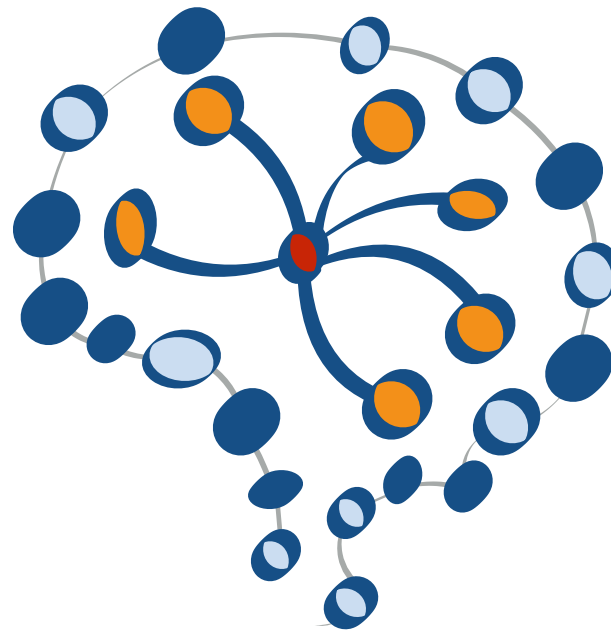


STAT 453: Introduction to Deep Learning and Generative Models

Sebastian Raschka

<http://stat.wisc.edu/~sraschka/teaching>

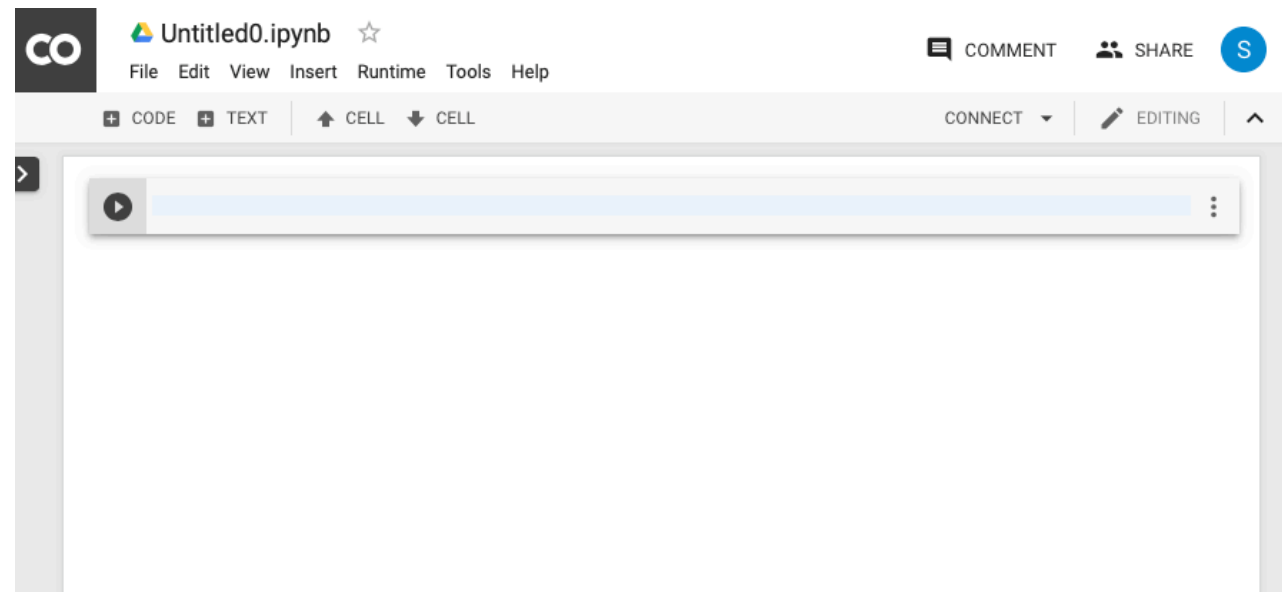


Lecture 07

Cloud Computing Resources

Google Colab

<https://colab.research.google.com>



- Free Google-flavored Jupyter Notebooks in the Cloud
- For each notebook, they spin up a custom (Linux-based) computing instance
- Computations limited to ~12 h though; you won't lose your notebook, but computations will be interrupted
- Maybe useful for quick testing/experimenting/sharing (but maybe tedious as you need to reinstall packages each time)

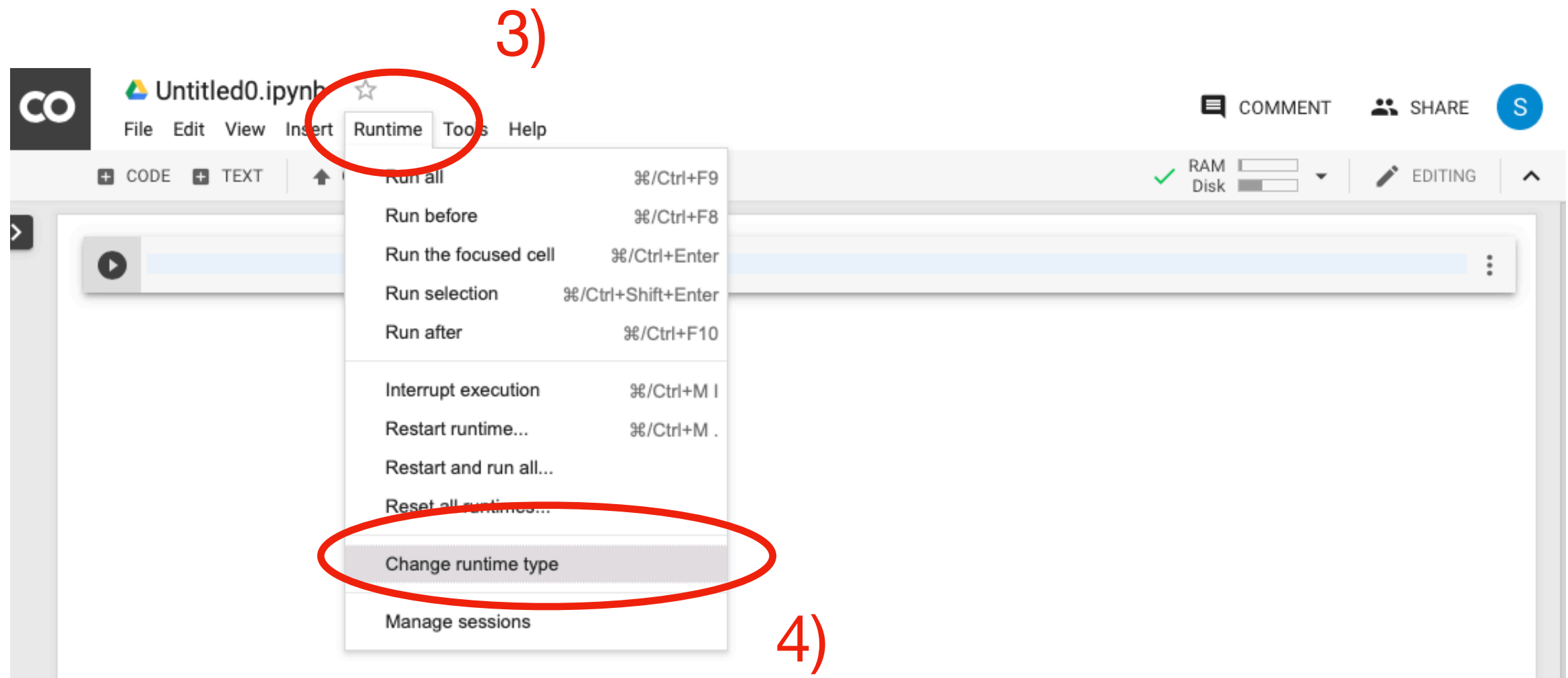
Google Colab

The screenshot shows the Google Colab interface. At the top, there is a navigation bar with tabs: EXAMPLES, RECENT, GOOGLE DRIVE (highlighted with a red circle and labeled 1), GITHUB, and UPLOAD. Below the navigation bar, there is a section titled 'Filter notebooks' with a search icon. A table lists several notebooks, all owned by Sebastian Raschka. The table has columns for Title, Owner, Last modified, and Last opened. At the bottom right, there is a button labeled 'NEW PYTHON 3 NOTEBOOK' (highlighted with a red circle and labeled 2) and a 'CANCEL' button.

Title	Owner	Last modified	Last opened
Untitled	Sebastian Raschka	3 days ago	3 days ago
Untitled7.ipynb	Sebastian Raschka	11 days ago	11 days ago
Untitled6.ipynb	Sebastian Raschka	Jan 12, 2019	Jan 12, 2019
Untitled5.ipynb	Sebastian Raschka	Dec 17, 2018	Dec 17, 2018
Untitled4.ipynb	Sebastian Raschka	Oct 12, 2018	Oct 12, 2018

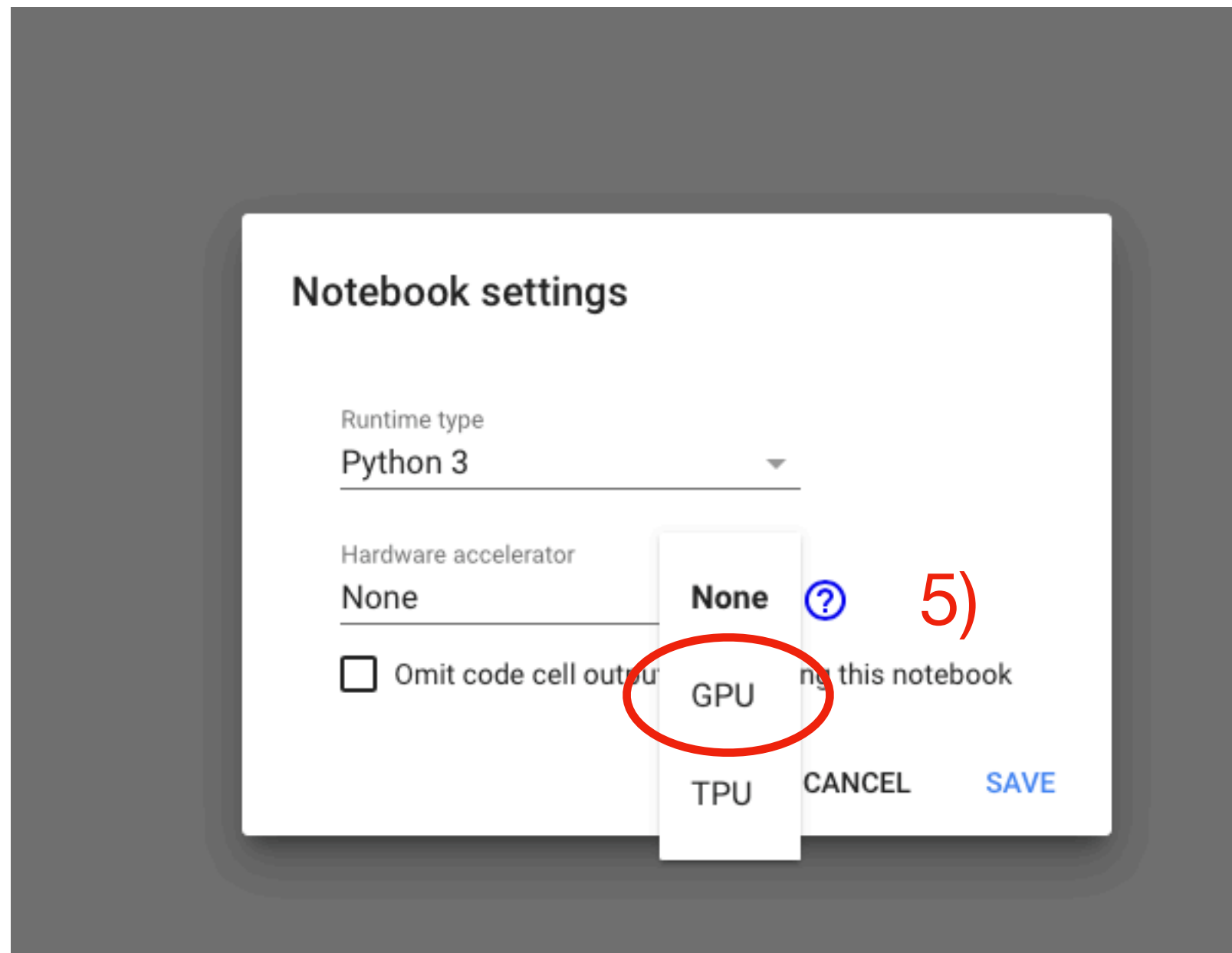
Menu appears if you visit <https://colab.research.google.com>

Google Colab



Follow these steps for running code on GPU later (default is CPU)

Google Colab



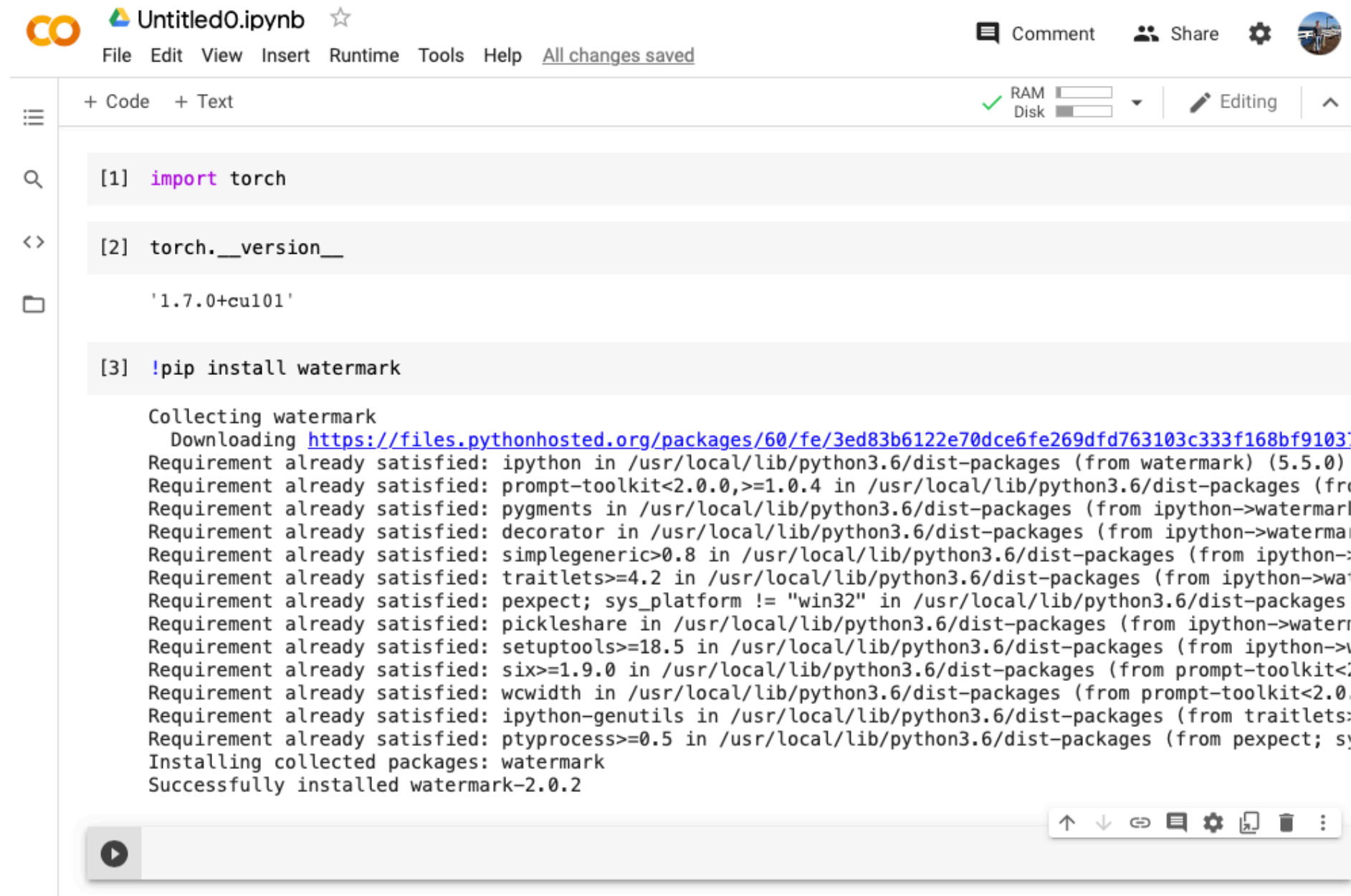
Follow these steps for running code on GPU later (default is CPU)

Google Colab

The screenshot shows the Google Colab web interface. At the top, there's a header with the Colab logo, the file name 'Untitled0.ipynb', and a star icon. Below this is a menu bar with options: File, Edit, View, Insert, Runtime, Tools, Help, and a status message 'All changes saved'. On the right side of the header, there are icons for Comment, Share, and a user profile picture. Below the header, there's a toolbar with '+ Code' and '+ Text' buttons. To the right of the toolbar, there's a status bar showing 'RAM' and 'Disk' usage with progress bars, and a button labeled 'Editing'. The main area contains two code cells. The first cell has the code '[1] import torch'. The second cell has the code '[2] torch.__version__' and its output is displayed as '1.7.0+cu101'. At the bottom of the code cells, there's a toolbar with icons for undo, redo, link, comment, settings, copy, delete, and a menu icon.

- This is nice! It appears that PyTorch is already pre-installed now (it wasn't always the case)

Google Colab



The screenshot shows a Google Colab notebook interface. At the top, there's a header with the Colab logo, the file name 'Untitled0.ipynb', and a star icon. Below this is a menu bar with 'File', 'Edit', 'View', 'Insert', 'Runtime', 'Tools', 'Help', and 'All changes saved'. On the right side of the header, there are icons for 'Comment', 'Share', and a user profile. Below the header, there's a toolbar with '+ Code' and '+ Text' buttons. On the right side of the toolbar, there are indicators for 'RAM' and 'Disk' usage, a 'Editing' mode button, and a scroll-up arrow. The main area of the notebook contains three code cells. The first cell has the code '[1] import torch'. The second cell has the code '[2] torch.__version__' followed by the output '1.7.0+cu101'. The third cell has the code '[3] !pip install watermark'. Below this code, the output shows the process of collecting and installing the watermark package, including a list of requirements that are already satisfied and the final installation of watermark-2.0.2.

```
[1] import torch

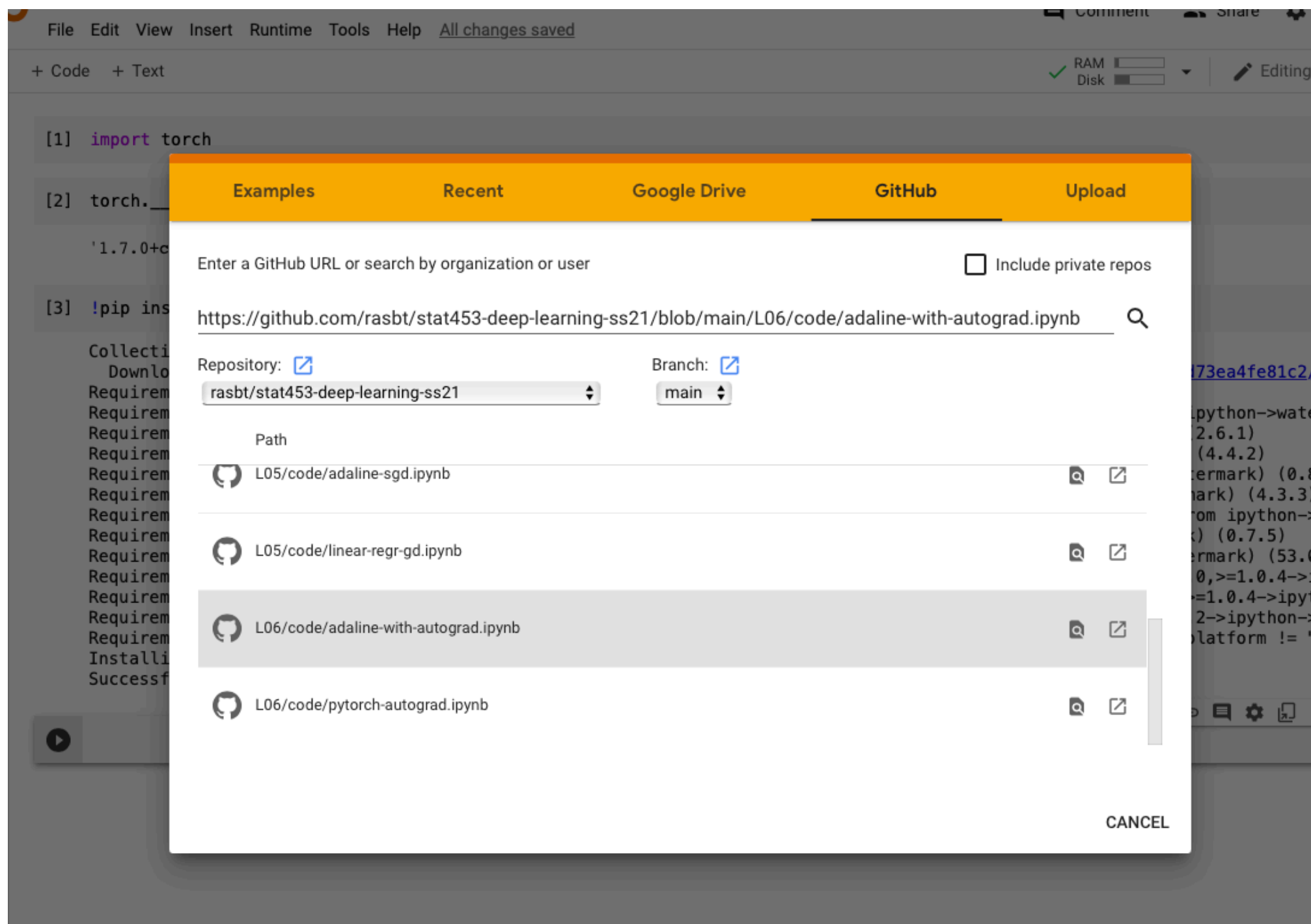
[2] torch.__version__
'1.7.0+cu101'

[3] !pip install watermark

Collecting watermark
  Downloading https://files.pythonhosted.org/packages/60/fe/3ed83b6122e70dce6fe269dfd763103c333f168bf9103/
Requirement already satisfied: ipython in /usr/local/lib/python3.6/dist-packages (from watermark) (5.5.0)
Requirement already satisfied: prompt-toolkit<2.0.0,>=1.0.4 in /usr/local/lib/python3.6/dist-packages (from ipython->watermark) (1.0.3)
Requirement already satisfied: pygments in /usr/local/lib/python3.6/dist-packages (from ipython->watermark) (2.4.2)
Requirement already satisfied: decorator in /usr/local/lib/python3.6/dist-packages (from ipython->watermark) (4.4.2)
Requirement already satisfied: simplegeneric>0.8 in /usr/local/lib/python3.6/dist-packages (from ipython->watermark) (0.1.1)
Requirement already satisfied: traitlets>=4.2 in /usr/local/lib/python3.6/dist-packages (from ipython->watermark) (4.2.0)
Requirement already satisfied: pexpect; sys_platform != "win32" in /usr/local/lib/python3.6/dist-packages (from ipython->watermark) (4.7.0)
Requirement already satisfied: pickleshare in /usr/local/lib/python3.6/dist-packages (from ipython->watermark) (0.7.5)
Requirement already satisfied: setuptools>=18.5 in /usr/local/lib/python3.6/dist-packages (from ipython->watermark) (44.0.0)
Requirement already satisfied: six>=1.9.0 in /usr/local/lib/python3.6/dist-packages (from prompt-toolkit<2.0.0,>=1.0.4->watermark) (1.11.0)
Requirement already satisfied: wcwidth in /usr/local/lib/python3.6/dist-packages (from prompt-toolkit<2.0.0,>=1.0.4->watermark) (0.1.7)
Requirement already satisfied: ipython-genutils in /usr/local/lib/python3.6/dist-packages (from traitlets->watermark) (0.2.0)
Requirement already satisfied: ptyprocess>=0.5 in /usr/local/lib/python3.6/dist-packages (from pexpect; sys_platform != "win32"->watermark) (0.6.0)
Installing collected packages: watermark
Successfully installed watermark-2.0.2
```

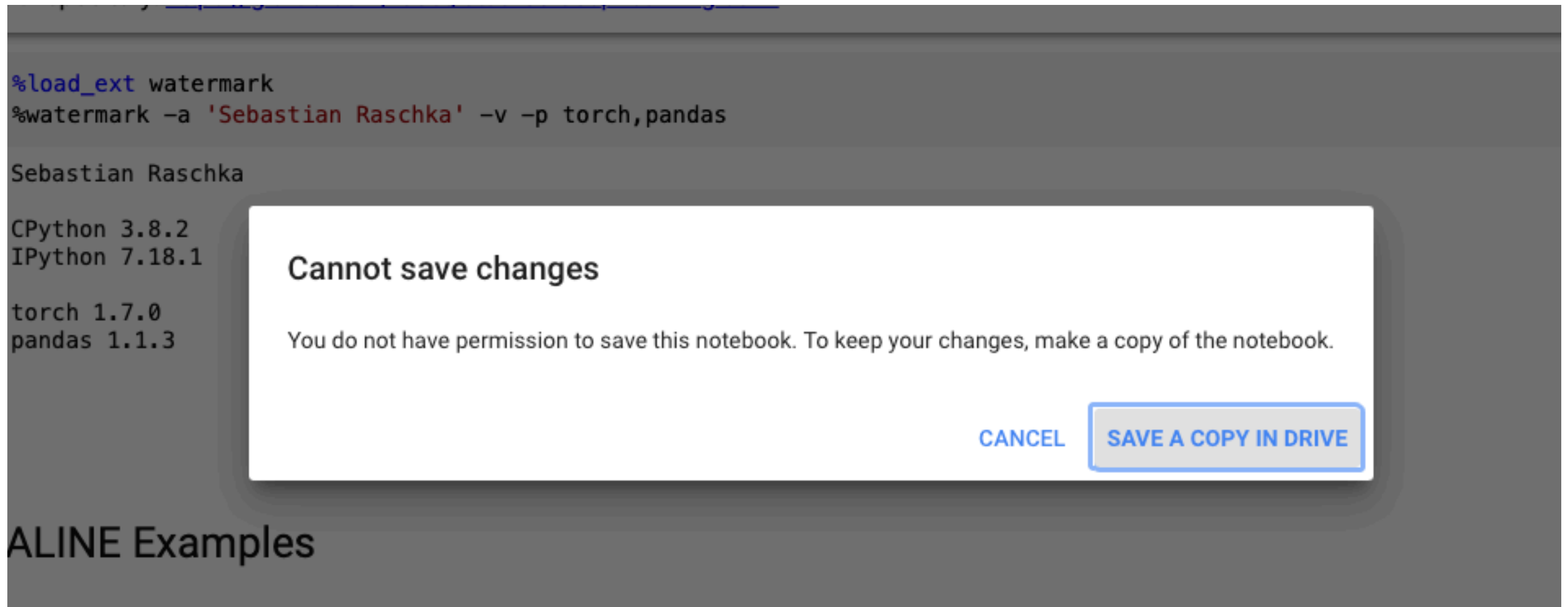
- In any case, if you'd like/need to install packages, you can do it as shown in the example above
- Note that in Jupyter Notebooks, the "!" indicates that what follows on that line is a "shell command" (you can think of a "shell" as the Linux & macOS command-line terminal, e.g., a Bash Shell)

Google Colab



- You can also upload Notebooks or directly paste GitHub links to notebooks

Google Colab



When you import a Notebook from a GitHub link, make sure to save it in your Google Drive if you plan to make edits, otherwise it will be gone later

Google Colab



FileNotFoundError Traceback (most recent call last)

<ipython-input-5-56bd975cla02> in <module>()

```
1 import pandas as pd
2
----> 3 df = pd.read_csv('./datasets/iris.data', index_col=None, header=None)
4 df.columns = ['x1', 'x2', 'x3', 'x4', 'y']
5 df = df.iloc[50:150]
```

4 frames

/usr/local/lib/python3.6/dist-packages/pandas/io/parsers.py in __init__(self, src, **kwargs)

```
2008     kwargs["usecols"] = self.usecols
2009
-> 2010     self._reader = parsers.TextReader(src, **kwargs)
2011     self.unnamed_cols = self._reader.unnamed_cols
2012
```

pandas/_libs/parsers.pyx in pandas._libs.parsers.TextReader.__cinit__()

pandas/_libs/parsers.pyx in pandas._libs.parsers.TextReader._setup_parser_source()

FileNotFoundError: [Errno 2] No such file or directory: './datasets/iris.data'

SEARCH STACK OVERFLOW

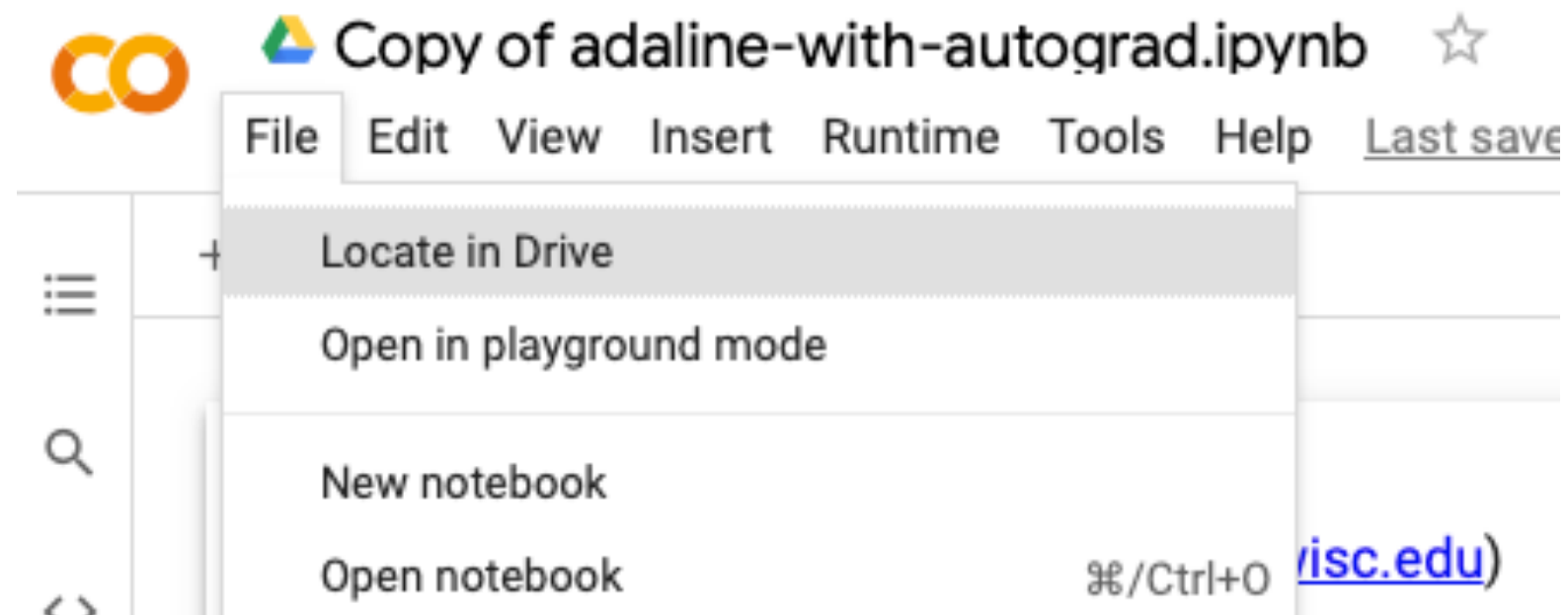


If you'd run the HW3 notebook, you'd likely encounter this error.
This is because it can't find the dataset via the specified, relative link ...

Google Colab

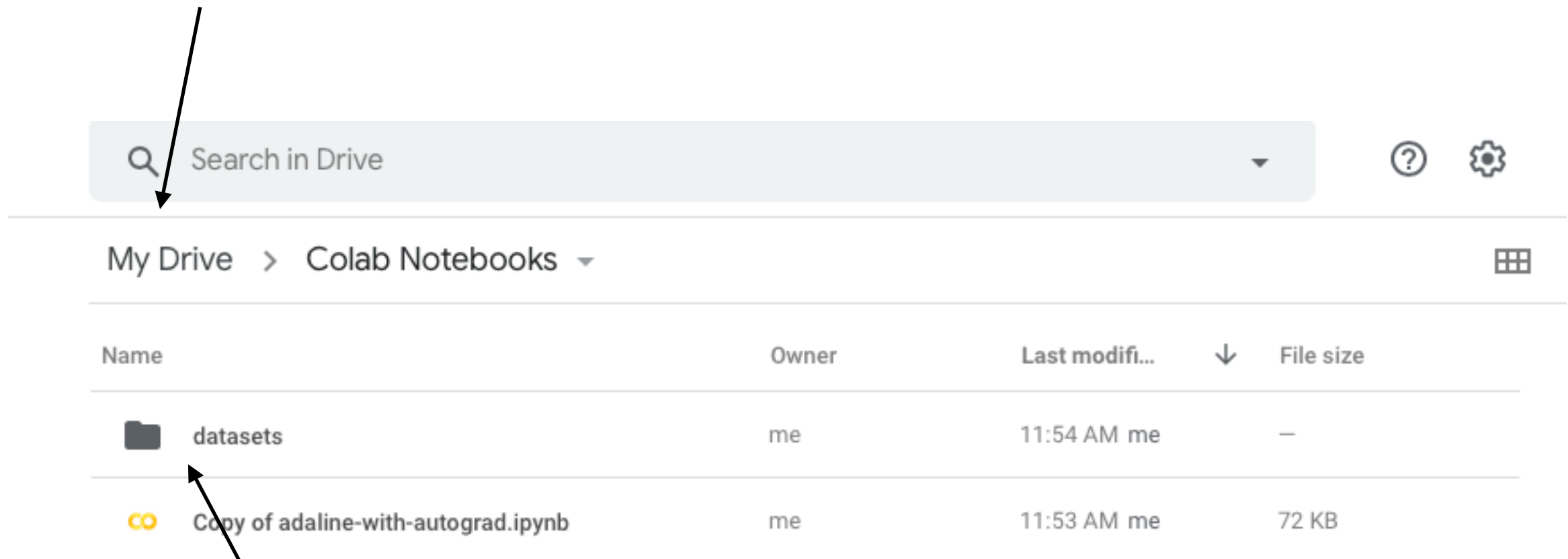
... you'd need to get the datafile into the same location as the notebook (or provide an absolute file path).

First, locate the position of the saved Notebook in your Google Drive:



Google Colab

Notebooks are usually in a directory called "M



I recommend uploading the whole folder though (simply drag&drop it from your computer into this window)

Google Colab

Unfortunately, there's some extra step required: mounting your Google Drive to the computer that now runs the Notebook. You need to execute the following code:

1)

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

... Go to this URL in a browser: https://accounts.google.com/o/oauth2/auth?client_id=9473189

Enter your authorization code:

Then, click on the link and enter it in the field above

2)

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

... Go to this URL in a browser: <https://accounts.google.co>

Enter your authorization code:
.....

3)

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

Go to this URL in a browser: <https://ac>

Enter your authorization code:
.....
Mounted at /content/drive

Your Google Drive should now be finally mounted:

Google Colab

Now, you simply need to provide the correct address to the dataset inside the Notebook and it should work:

▼ Dataset Prep

```
import pandas as pd

df = pd.read_csv('/content/drive/My Drive/Colab Notebooks/datasets/iris.data', index_col=None, header=None)
df.columns = ['x1', 'x2', 'x3', 'x4', 'y']
df = df.iloc[50:150]
df['y'] = df['y'].apply(lambda x: 0 if x == 'Iris-versicolor' else 1)

# Assign features and target
X_train = df[['x1', 'x2', 'x3', 'x4']].values
y_train = df['y'].values
```

Google Colab

However, for performance reasons during model training when working with larger datasets, I recommend making a copy of the Google Drive dataset to the local machine that is running the notebook:

▼ Dataset Prep

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

Mounted at /content/drive

```
▶ import shutil  
  
shutil.copytree('/content/drive/My Drive/Colab Notebooks/datasets/', 'datasets')
```

📁 'datasets'

```
▶ import pandas as pd  
  
df = pd.read_csv('datasets/iris.data', index_col=None, header=None)  
df.columns = ['x1', 'x2', 'x3', 'x4', 'y']  
df = df.iloc[50:150]  
df['y'] = df['y'].apply(lambda x: 0 if x == 'Iris-versicolor' else 1)
```

Google Colab

In practice, it might be even better to do that with .zip files because it will be faster to download a zip file from Google Drive and then unzip it

▼ Dataset Prep

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

Mounted at /content/drive

```
[15] import shutil  
  
shutil.copy('/content/drive/My Drive/Colab Notebooks/datasets.zip', './')  
  
 './datasets.zip'
```

```
!unzip datasets.zip
```

```
... Archive: datasets.zip  
  inflating: __MACOSX/._datasets  
replace datasets/iris.data? [y]es, [n]o, [A]ll, [N]one, [r]ename: yes|
```