

Colab Notebooks

Jeff Liu and Gary Chen

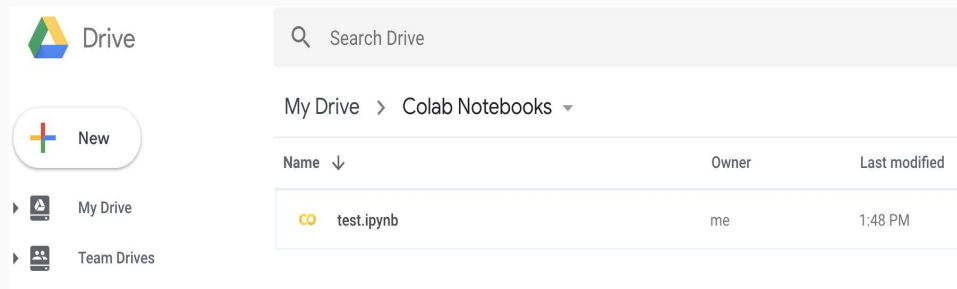
2019/01/12



Introduction

- 類似Jupyter Notebook，互動式介面，及時輸出結果
- 儲存在Google Drive 上，可以共同編輯
- 背後有VM 在運作
- 可連續使用免費的 Tesla K80 GPU 12小時(每次)
- 許多套件已事先安裝，例如：TensorFlow, Scikit-learn, Matplotlib等

→ Text Note: goo.gl/vmJGJ4



下載檔案

從網站下載 code 到本機 -> 網址 : github.com/CloudMile/recommendation_engine_course

CloudMile / **recommendation_engine_course** Watch 5 Star 0 Fork 2

<> Code Issues 0 Pull requests 1 Projects 0 Wiki Insights

No description, website, or topics provided.

45 commits 1 branch 0 releases 1 contributor

Branch: master New pull request Create new file Upload files Find file **Clone or download**

GaryChenCloudMile Update README.md

data	First commit
model	modify jupyter-notebook program structure!
utils	#
.gitignore	First commit

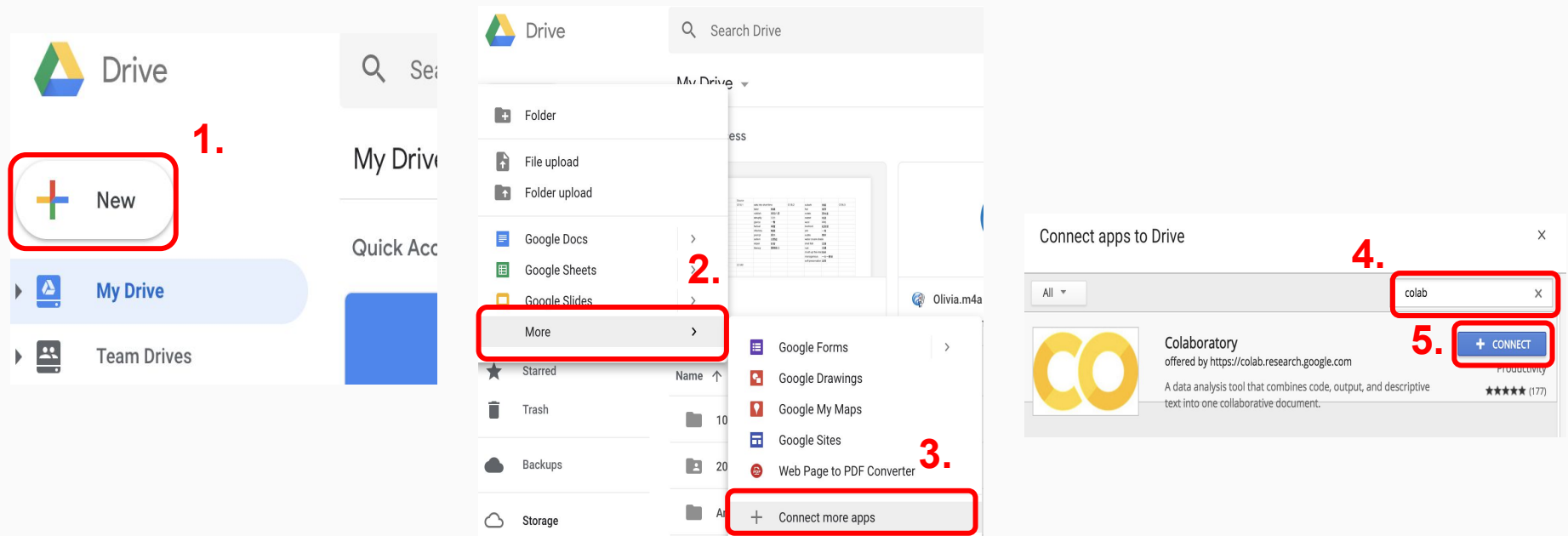
Clone with HTTPS Use Git or checkout with SVN using the web URL.
<https://github.com/CloudMile/recommen>

Open in Desktop **Download ZIP**

→ Text Note:
goo.gl/vmJGJ4

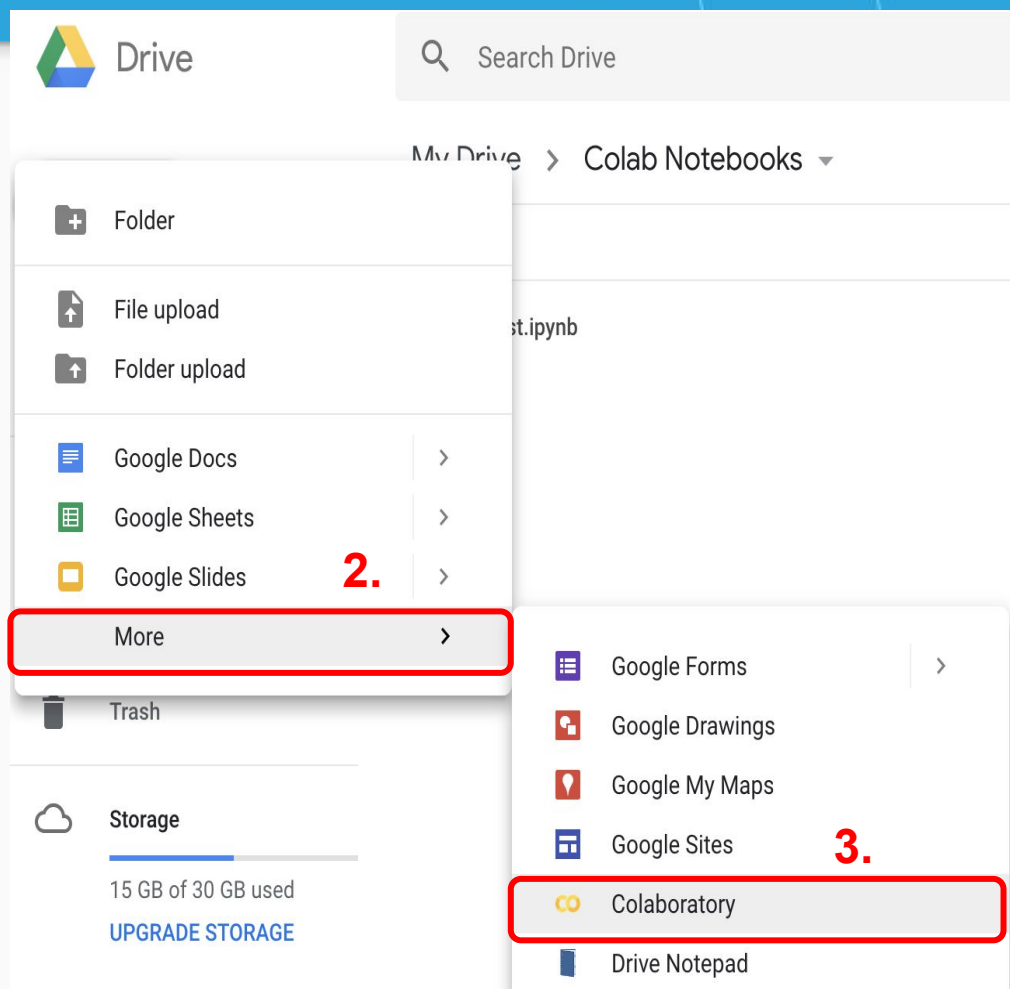
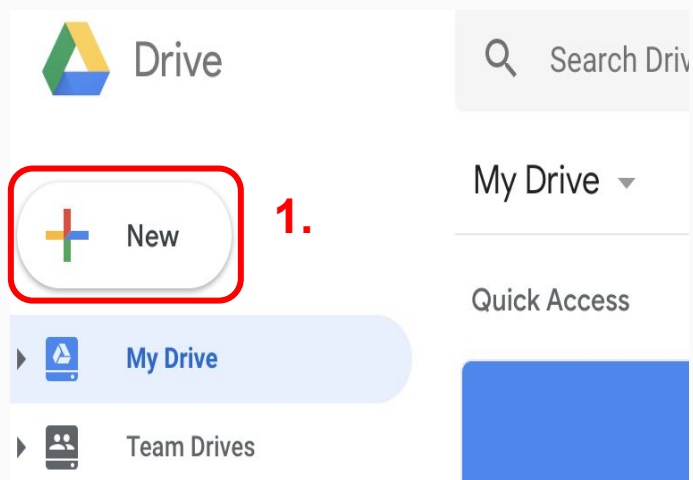
Setup

- 進入Google Drive 連接到Colaboratory

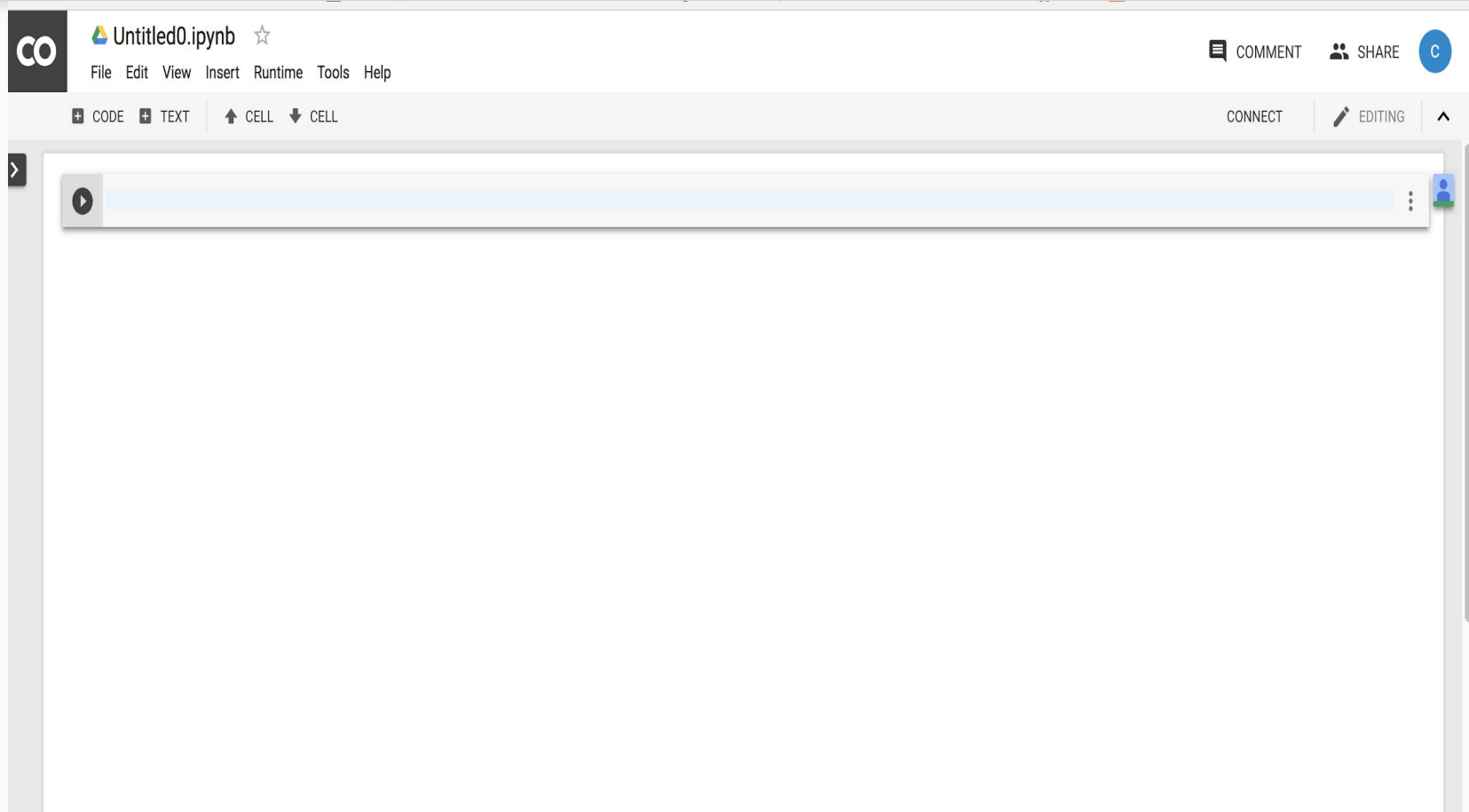


Setup

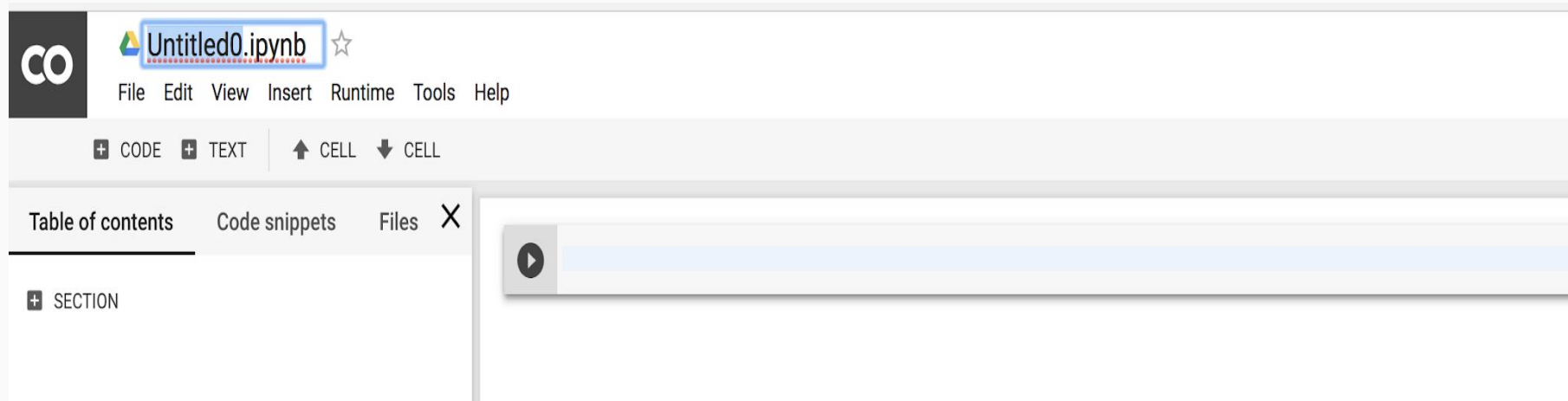
- 建立一個新的Colaboratory



建立完成

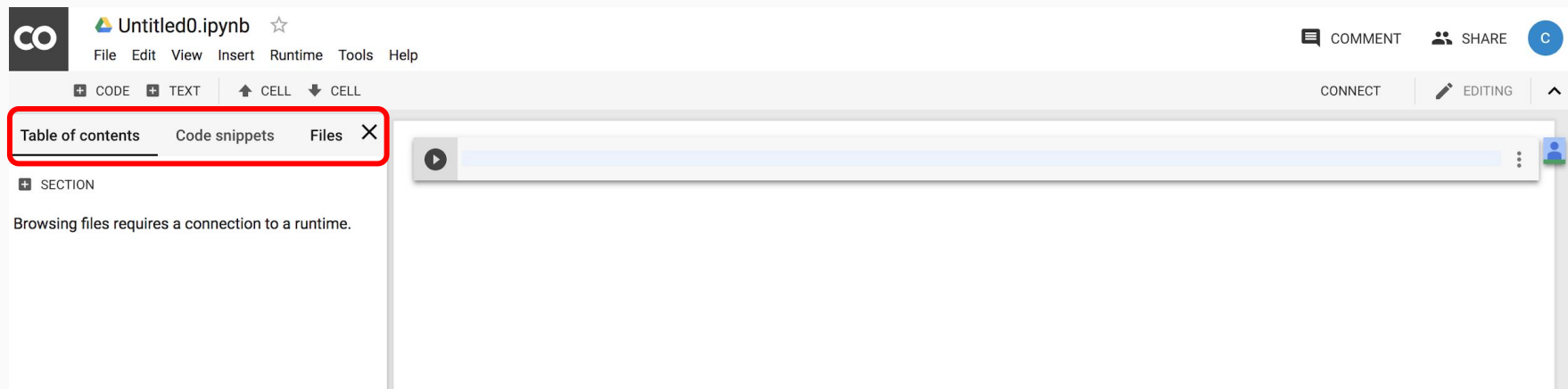


修改名稱



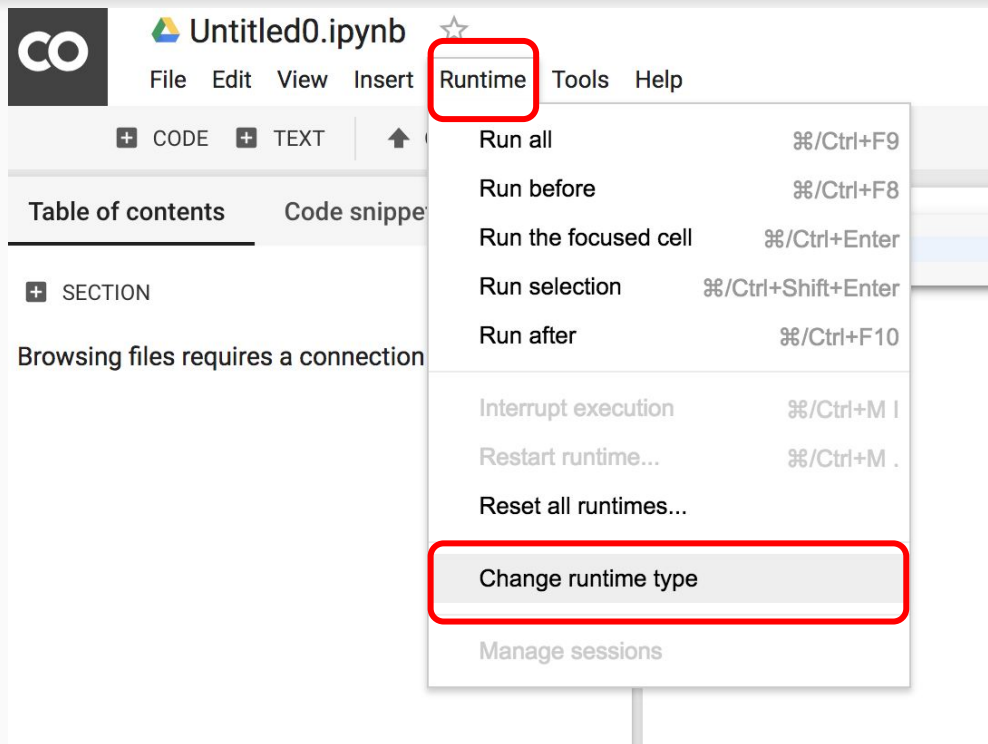
The screenshot displays the JupyterLab user interface. At the top, the file name 'Untitled0.ipynb' is highlighted in a blue box. Below the title bar, a menu bar contains 'File', 'Edit', 'View', 'Insert', 'Runtime', 'Tools', and 'Help'. A secondary bar features buttons for '+ CODE', '+ TEXT', '↑ CELL', and '↓ CELL'. On the left, a sidebar with a close button 'X' contains 'Table of contents', 'Code snippets', and 'Files'. Under 'Table of contents', there is a '+ SECTION' button. The main workspace area on the right is currently empty, showing a light blue horizontal bar and a play button icon.

查看檔案

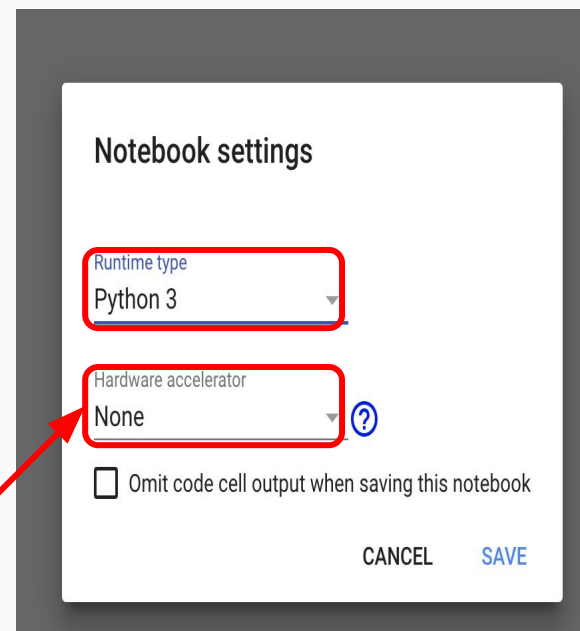


The screenshot displays the Google Colab web interface. At the top left is the Colab logo. The main header shows 'Untitled0.ipynb' with a star icon and a menu bar containing 'File', 'Edit', 'View', 'Insert', 'Runtime', 'Tools', and 'Help'. On the right side of the header are 'COMMENT', 'SHARE', and a user profile icon labeled 'C'. Below the header is a toolbar with buttons for '+ CODE', '+ TEXT', 'CELL', and 'CELL'. A sidebar on the left is titled 'Table of contents' (highlighted with a red box), 'Code snippets', and 'Files' (with a close button 'X'). Under 'Table of contents' is a '+ SECTION' button and a message: 'Browsing files requires a connection to a runtime.' The main workspace area is mostly empty, with a play button icon on the left and a user profile icon on the right.

修改Runtime

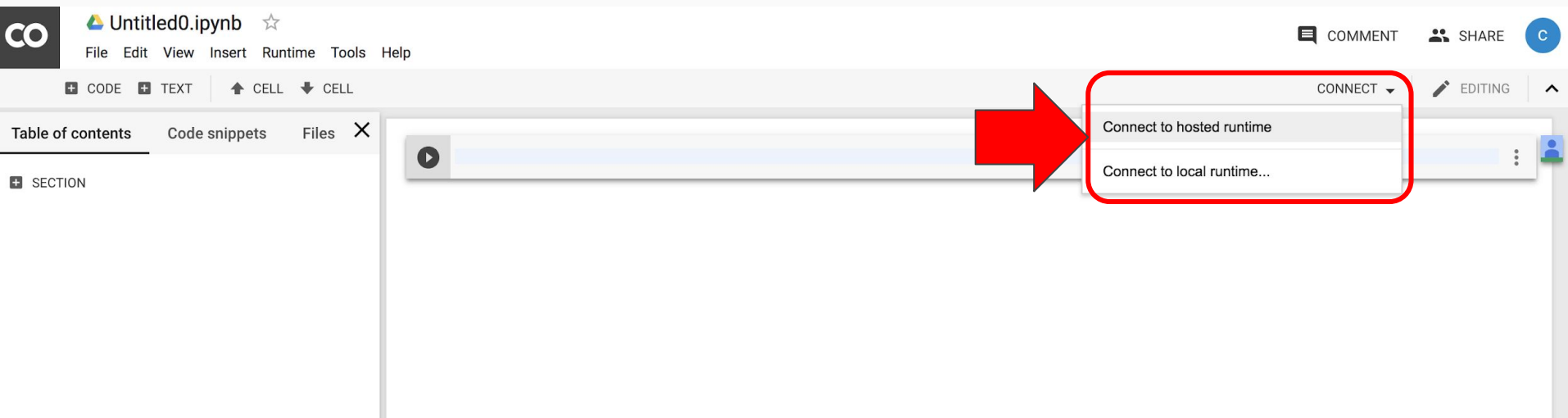


● 使用 Python 3



可選GPU

連接VM



The screenshot displays the Google Colab web interface. At the top, the title bar shows 'Untitled0.ipynb' with a star icon. Below it are menu items: File, Edit, View, Insert, Runtime, Tools, and Help. On the right side of the title bar are links for COMMENT, SHARE, and a user profile icon. The main toolbar contains buttons for CODE, TEXT, and two CELL buttons (one with an up arrow, one with a down arrow). On the left, there is a sidebar with 'Table of contents', 'Code snippets', and 'Files' sections. The central area is a code editor with a play button icon. A red arrow points from the code editor to a 'CONNECT' dropdown menu that is open, showing two options: 'Connect to hosted runtime' and 'Connect to local runtime...'. The dropdown menu is highlighted with a red border.

co Untitled0.ipynb ☆

File Edit View Insert Runtime Tools Help

COMMENT SHARE C

CODE TEXT CELL CELL

CONNECT

- Connect to hosted runtime
- Connect to local runtime...

Table of contents Code snippets Files

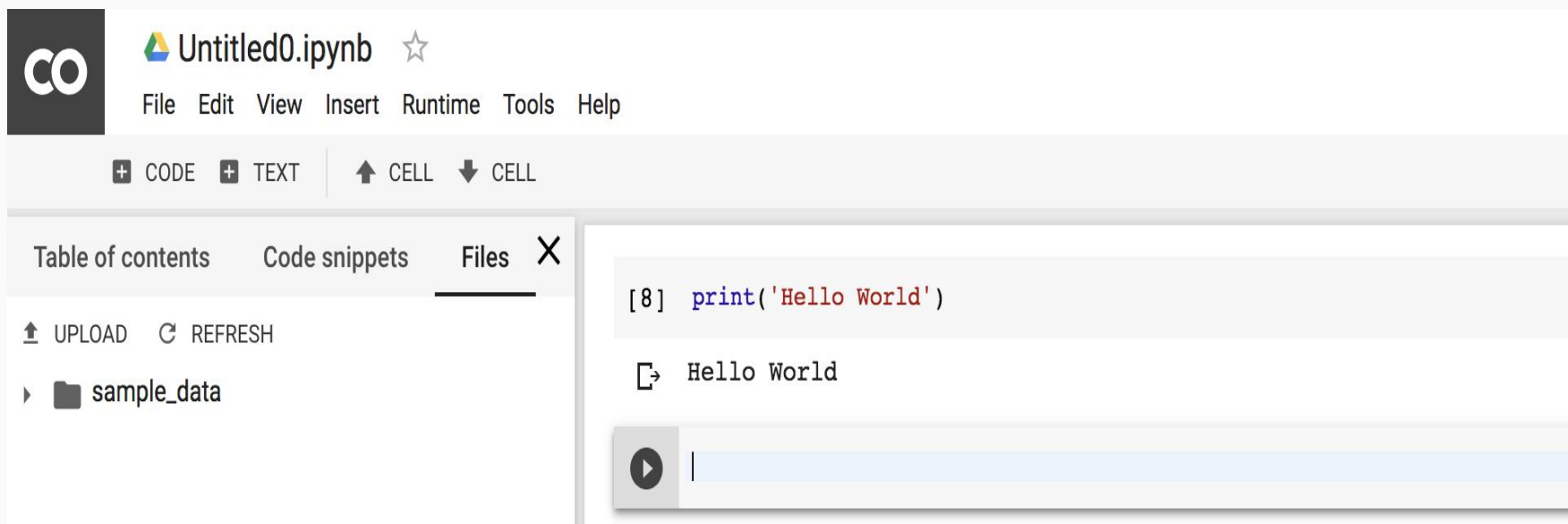
SECTION

連接完成

The screenshot displays a Jupyter Notebook environment. At the top, the title bar shows 'Untitled0.ipynb' with a star icon. Below it, a menu bar includes 'File', 'Edit', 'View', 'Insert', 'Runtime', 'Tools', and 'Help'. On the right side of the title bar, there are buttons for 'COMMENT', 'SHARE', and a user profile icon. Below the title bar, a toolbar contains tabs for '+ CODE', '+ TEXT', '↑ CELL', and '↓ CELL'. To the right of the toolbar, there are status indicators: a green checkmark labeled 'CONNECTED', a pencil icon labeled 'EDITING', and an upward arrow icon. The main interface is divided into two panels. The left panel, titled 'Files', shows a file explorer with a folder named 'sample_data' expanded, revealing several files: 'README.md', 'anscombe.json', 'california_housing_test.csv', 'california_housing_train.csv', 'mnist_test.csv', and 'mnist_train_small.csv'. The right panel is a large code editor area, currently empty, with a play button icon in the top left corner. A user profile icon is visible in the top right corner of the code editor area.

基本語法 - Python

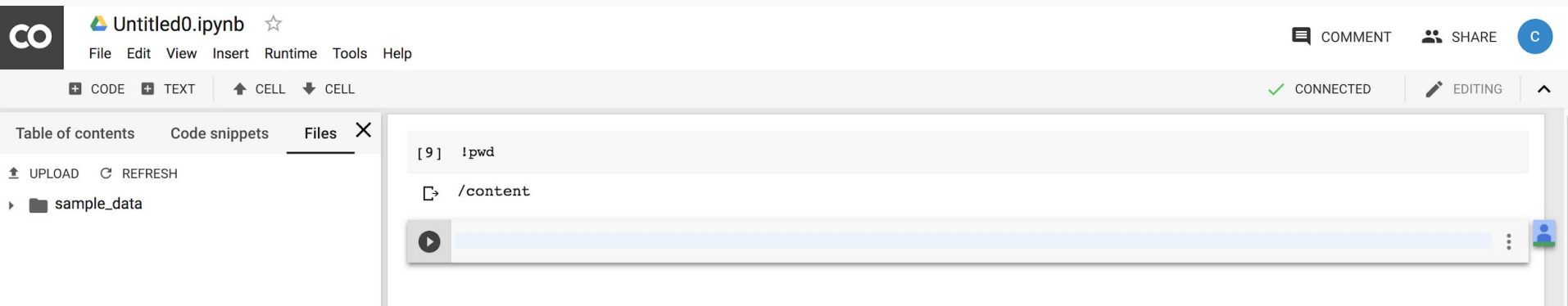
- 輸入 python 程式碼，點擊「Run」或是「Ctrl + Enter」執行程式碼區塊。



The screenshot displays a Jupyter Notebook interface. At the top, the title bar shows 'Untitled0.ipynb' with a star icon. Below it is a menu bar with 'File', 'Edit', 'View', 'Insert', 'Runtime', 'Tools', and 'Help'. A toolbar contains icons for '+ CODE', '+ TEXT', '↑ CELL', and '↓ CELL'. On the left, a sidebar with a 'Table of contents' and 'Code snippets' tab shows a 'Files' section with an 'UPLOAD' button, a 'REFRESH' button, and a folder icon labeled 'sample_data'. The main area features a code cell with the input '[8] print('Hello World')' and the output 'Hello World' displayed below it. At the bottom of the code cell is a play button icon and a horizontal progress bar.

基本語法 - Shell

- 前面加上！



The screenshot displays a Jupyter Notebook interface. At the top, the title bar shows 'Untitled0.ipynb' with a star icon. Below it is a menu bar with 'File', 'Edit', 'View', 'Insert', 'Runtime', 'Tools', and 'Help'. To the right of the menu bar are buttons for 'COMMENT', 'SHARE', and a user profile icon. Below the menu bar is a toolbar with '+ CODE', '+ TEXT', '↑ CELL', and '↓ CELL'. On the right side of the toolbar, there is a green checkmark and the text 'CONNECTED', followed by a pencil icon and the text 'EDITING', and an upward arrow icon. On the left side, there is a sidebar with 'Table of contents', 'Code snippets', and 'Files'. Under 'Files', there is a folder icon and the text 'sample_data'. The main area of the notebook shows a code cell with the following content:

```
[9] !pwd
```

```
/content
```

 Below the code cell is a play button icon and a blue progress bar. On the far right of the play button bar is a user profile icon.

特殊語法 - 查看

- 在function 前面加上? 或是用help()可以查看documentation

The screenshot shows a Jupyter Notebook interface with a code cell containing the following code:

```
[1] import numpy as np

np.asarray?
# help(np.asarray)
```

The notebook is titled "Untitled0.ipynb" and has a menu bar with options: File, Edit, View, Insert, Runtime, Tools, Help. The status bar at the bottom indicates "CONNECTED" and "EDITING".

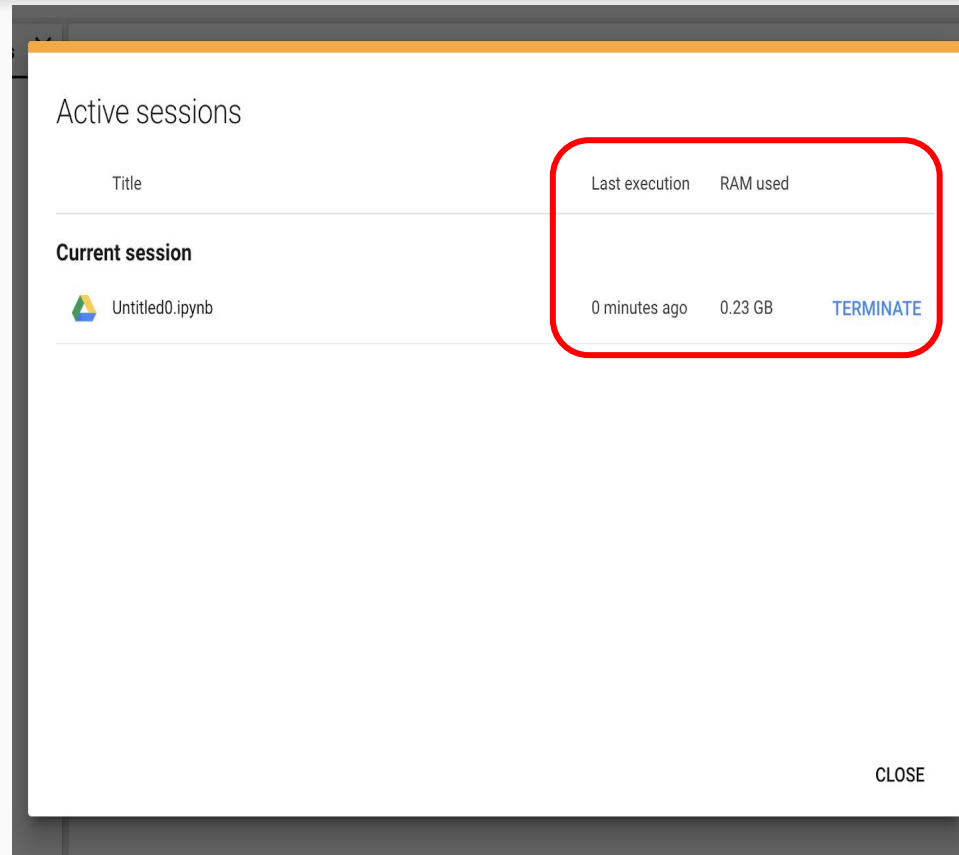
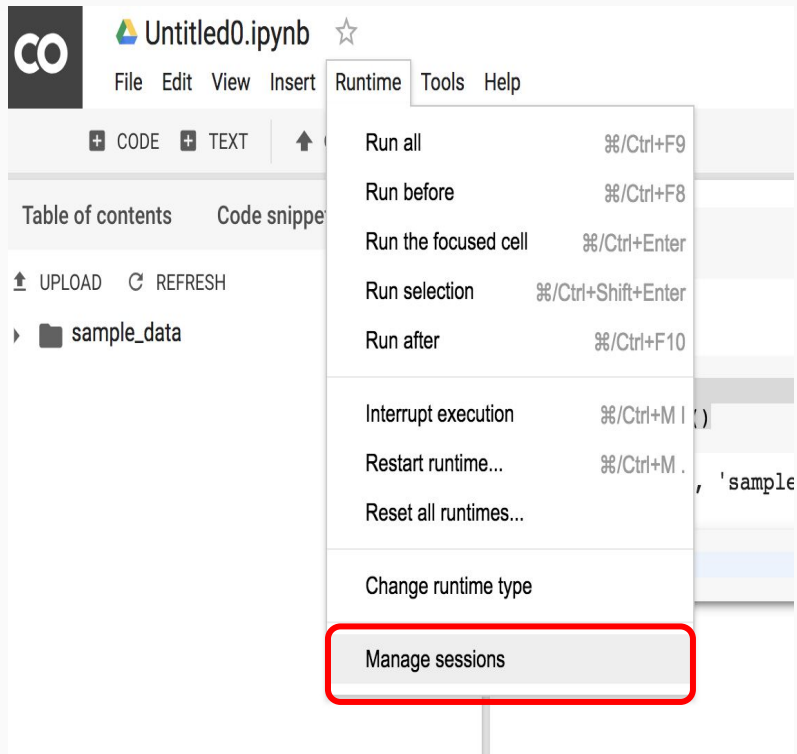
The help documentation for `np.asarray` is displayed below the code cell:

Signature: `np.asarray(a, dtype=None, order=None)`
Docstring:
Convert the input to an array.

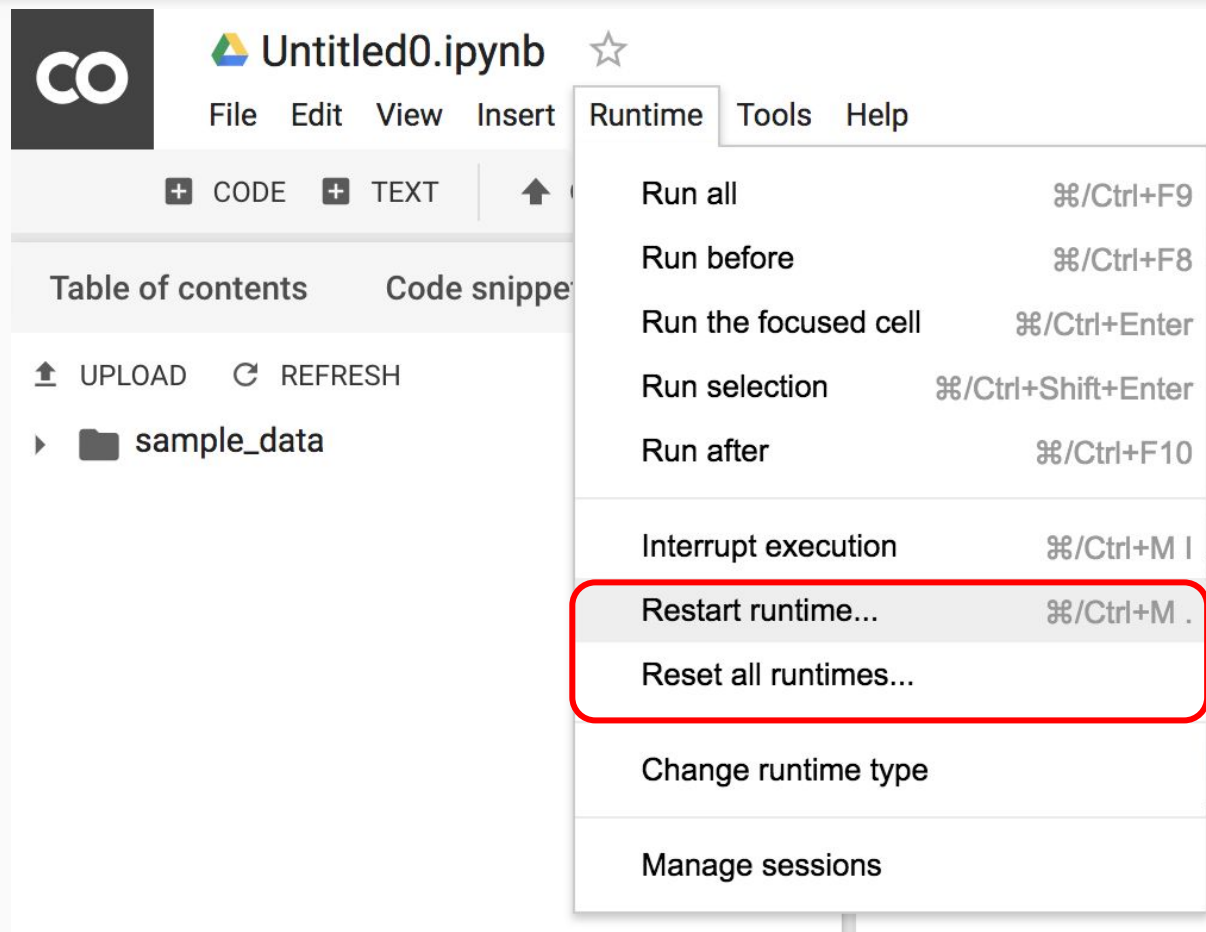
Parameters

a : array_like
Input data, in any form that can be converted to an array. This includes lists, lists of tuples, tuples, tuples of tuples, tuples of lists and ndarrays.

管理Session



Restart runtime/ Reset runtime



Lab 1: Memory-Based CF

→ Note: goo.gl/vmJGJ4

Colab Links

1.

Lab 1: Memory base CF
[reco_memory_base.ipynb](#)

Lab 2: DNN practice:
[tutorial_dnn_practice.ipynb](#)
[lab_tutorial_dnn_practice.ipynb](#)

Lab 3: Basic MF
[reco_model_mf.ipynb](#)
[lab_reco_model_mf.ipynb](#)

Lab 4: MF with user histories
[reco_model_mf_with_history.ipynb](#)
[lab_reco_model_mf_with_history.ipynb](#)

Lab 5: MF with DNN
[reco_model_mf_dnn.ipynb](#)
[lab_reco_model_mf_dnn.ipynb](#)

reco_memory_base.ipynb

File Edit View Insert Runtime Tools Help

CODE TEXT CELL COPY TO DRIVE

```
1 # Clear
2 !rm -fr data model utils recomm
3
4 # Clone source code from git
5 !git clone https://github.com/CloudMile/recommendation_engine_course recomm
6
7 # Move data and utils to parent folder that notebook could access
8 !mv recomm/data .
9 !mv recomm/utils .
10 # Copy the model checkpoints (optional)
11 !mv recomm/model .
12 !rm -fr recomm
```

Cloning into 'recomm'...

```
remote: Enumerating objects: 3, done.
remote: Counting objects: 100% (3/3), done.
remote: Compressing objects: 100% (3/3), done.
remote: Total 354 (delta 0), reused 1 (delta 0), pack-reused 351
Receiving objects: 100% (354/354), 55.62 MiB | 9.59 MiB/s, done.
Resolving deltas: 100% (185/185), done.
```

```
1 """
2 Import some required packages.
3 """
4
5 # from __future__ import division, print_function, with_statement, absolute_import, unicode_literals
```

2.

從Colab打開Jupyter Notebook

從網站下載 code 到本機 -> 網址 : github.com/CloudMile/recommendation_engine_course

CloudMile / [recommendation_engine_course](#)

Watch 5

★ Star 0

Fork 2

Code

Issues 0

Pull requests 1

Projects 0

Wiki

Insights

No description, website, or topics provided.

45 commits

1 branch

0 releases

1 contributor

Branch: master

New pull request

Create new file

Upload files

Find file

Clone or download

GaryChenCloudMile Update README.md

data First commit

model modify jupyter-notebook program structure!

utils #

.gitignore First commit

Clone with HTTPS

Use SSH

Use Git or checkout with SVN using the web URL.

<https://github.com/CloudMile/recommen>

Open in Desktop

Download ZIP

從Colab打開Jupyter Notebook

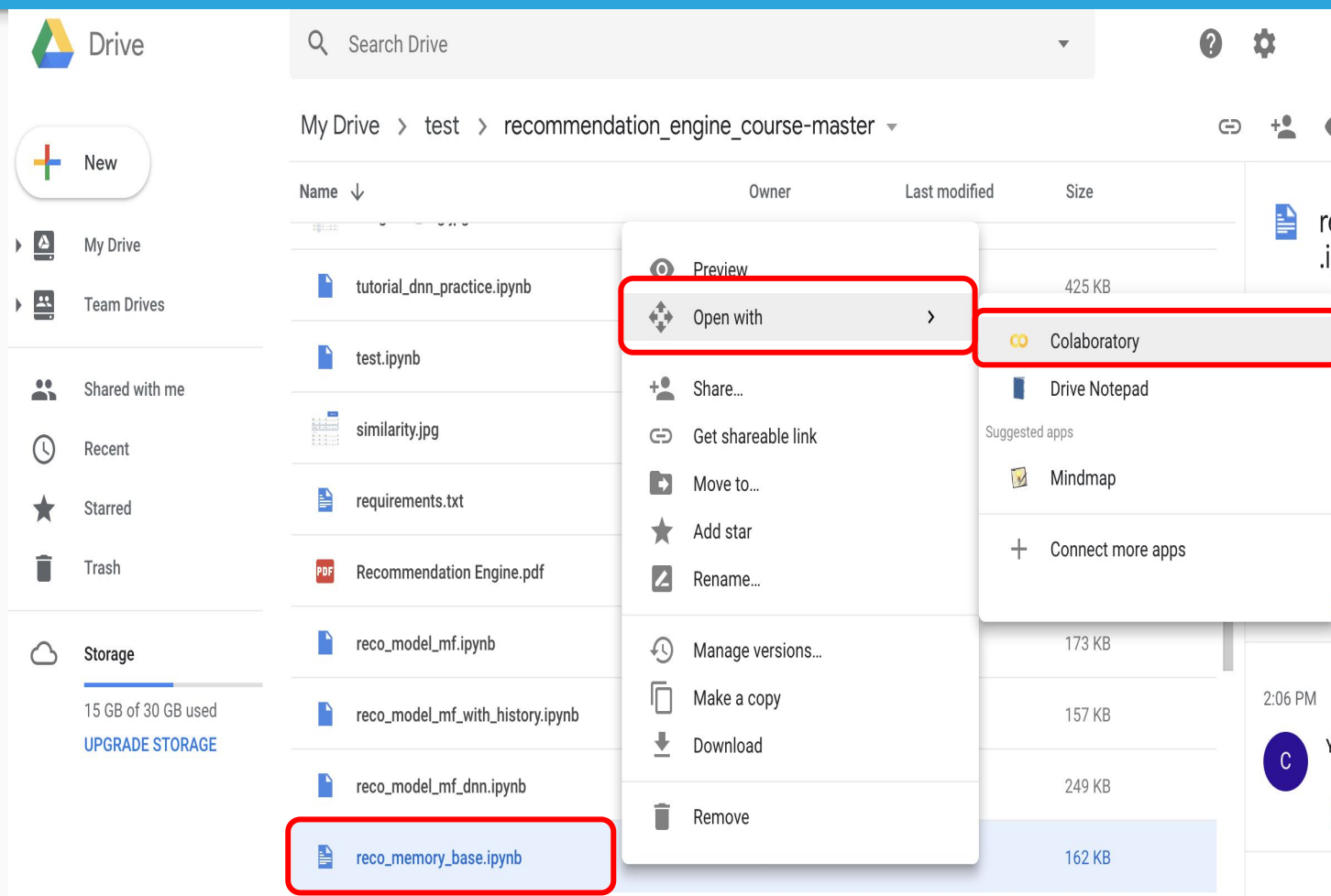
- 解壓縮後，上傳整個資料夾到Google Drive

The screenshot shows the Google Drive web interface. On the left sidebar, there's a 'New' button and a list of locations: My Drive, Team Drives, Shared with me, Recent, Starred, and Trash. At the bottom of the sidebar, there's a 'Storage' section showing '15 GB of 30 GB used' and a link to 'UPGRADE STORAGE'. The main area shows the 'My Drive' view with a search bar and a breadcrumb 'My Drive > test'. Below this is a table of files:

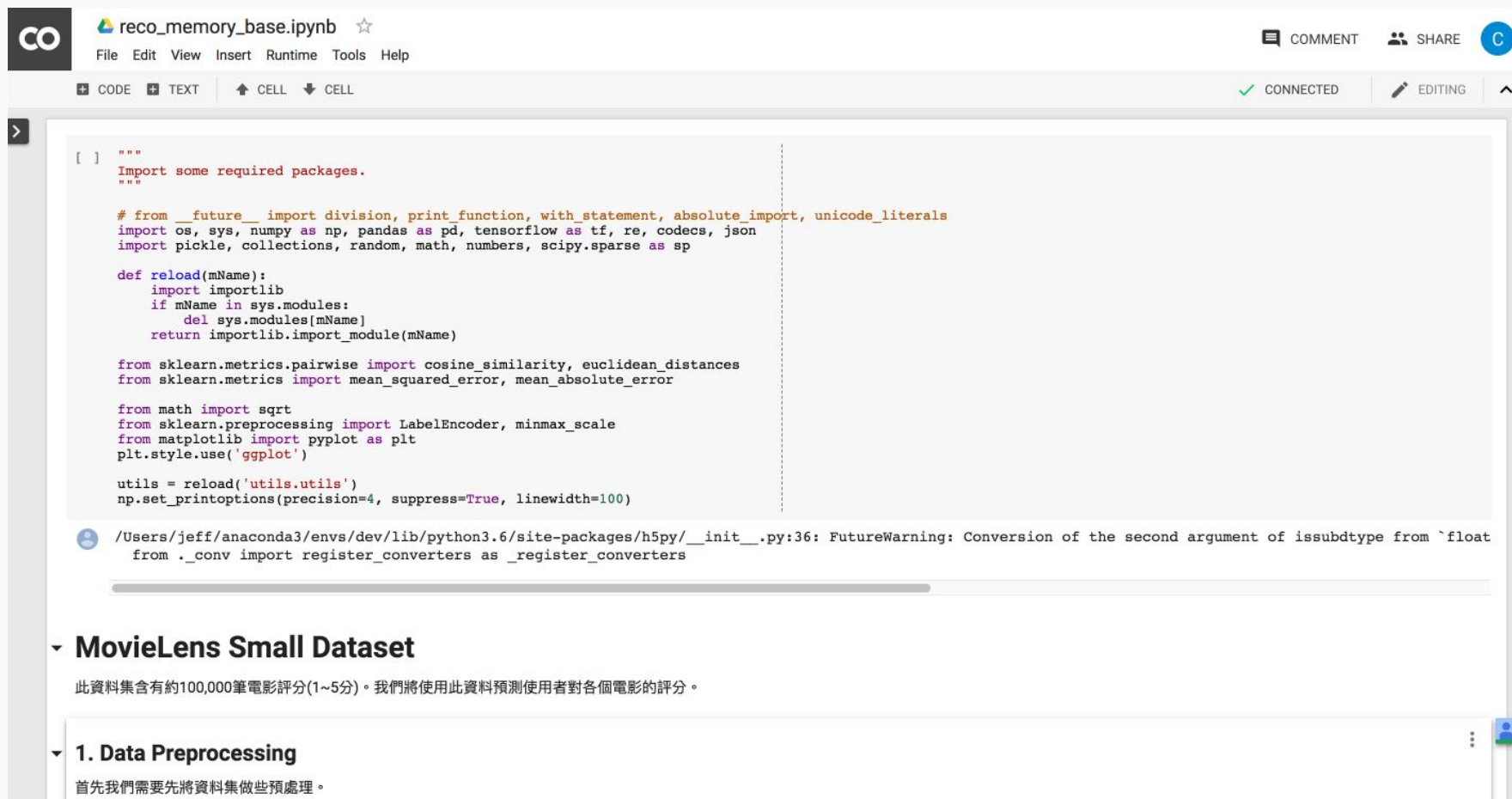
Name	Owner	Last modified	Size
recommendation_engine_course-master	me	1:45 PM	—

On the right side, there's a panel for the 'test' folder, showing 'Details' and 'Activity' tabs. The 'Activity' tab is active, showing a log of actions. The first entry is 'You created an item in' followed by 'My Drive' and the 'test' folder.

從Colab打開Jupyter Notebook



從Colab打開Jupyter Notebook



reco_memory_base.ipynb

File Edit View Insert Runtime Tools Help

CODE TEXT CELL CELL

CONNECTED EDITING

```
[ ] """
Import some required packages.
"""

# from __future__ import division, print_function, with_statement, absolute_import, unicode_literals
import os, sys, numpy as np, pandas as pd, tensorflow as tf, re, codecs, json
import pickle, collections, random, math, numbers, scipy.sparse as sp

def reload(mName):
    import importlib
    if mName in sys.modules:
        del sys.modules[mName]
    return importlib.import_module(mName)

from sklearn.metrics.pairwise import cosine_similarity, euclidean_distances
from sklearn.metrics import mean_squared_error, mean_absolute_error

from math import sqrt
from sklearn.preprocessing import LabelEncoder, minmax_scale
from matplotlib import pyplot as plt
plt.style.use('ggplot')

utils = reload('utils.utils')
np.set_printoptions(precision=4, suppress=True, linewidth=100)
```

/Users/jeff/anaconda3/envs/dev/lib/python3.6/site-packages/h5py/_init_.py:36: FutureWarning: Conversion of the second argument of issubdtype from `float` from ._conv import register_converters as _register_converters

MovieLens Small Dataset

此資料集含有約100,000筆電影評分(1~5分)。我們將使用此資料預測使用者對各個電影的評分。

1. Data Preprocessing

首先我們需要將資料集做些預處理。

Reference

- https://colab.research.google.com/drive/1jt6Dk2MoUaBZu0nUuQpNAzVfMpm08Vps?authuser=1#scrollTo=zmVS_blsW2Cq
- <https://research.google.com/colaboratory/faq.html#gpu-availability>
- <https://research.google.com/colaboratory/local-runtimes.html>