

Introduction to TensorFlow

Setup

郭耀仁

大綱

- 開發環境
- Kaggle Datasets
- 隨堂練習

開發環境

兩種開發環境

- 瀏覽器：Google Colaboratory
- 本機端

建立瀏覽器開發環境

[illegible]

建立本機端開發環境

Miniconda (<https://docs.conda.io/en/latest/miniconda.html>).

安裝 jupyter

run in command line

```
pip install jupyter
```

檢視可用虛擬環境

run in command line

```
conda env list
```


建立虛擬環境

run in command line

```
conda create --name tensorflow python=3.6
```

啟動虛擬環境

run in command line

```
conda activate tensorflow
```

填寫需要的模組

- 在專案資料夾中新增一個 requirements.txt 文字檔作為安裝模組清單

```
tensorflow  
scipy  
scikit-learn  
matplotlib  
ipykernel  
pandas  
numpy  
kaggle
```

安裝

run in command line

```
pip install -r requirements.txt
```

建立新的 Jupyter Notebook Kernel

run in command line

```
python -m ipykernel install --user --name tensorflow --display-name "TensorFlow"
```

檢視可用的 Jupyter Notebook Kernel

run in command line
jupyter kernelspec list

開啟以 TensorFlow 為 Kernel 的 Notebook

The screenshot shows the Google Colab interface. At the top, there are three tabs: 'Files', 'Running', and 'Clusters'. Below the tabs, a message says 'Select items to perform actions on them.' To the right of this message are three buttons: 'Upload', 'New', and a refresh icon. The 'New' button is clicked, opening a dropdown menu. The menu has a section titled 'Notebook:' followed by a list of options: 'Julia 1.0.0', 'Python 3', 'R', 'TensorFlow', and 'Create a new notebook with TensorFlow'. Below this section are three more options: 'Text File', 'Folder', and 'Terminal'. The 'TensorFlow' option is highlighted with a grey background.

Files Running Clusters

Select items to perform actions on them.

Upload New ↕

Notebook:

- Julia 1.0.0
- Python 3
- R
- TensorFlow
- Create a new notebook with TensorFlow

Text File

Folder

Terminal

Hello TensorFlow

```
In [1]: # Run in Jupyter Notebook  
import tensorflow as tf  
  
hello = tf.constant('Hello, TensorFlow!')  
with tf.Session() as sess:  
    print(sess.run(hello))
```

```
b'Hello, TensorFlow!'
```


Kaggle Datasets

註冊一個 Kaggle (<https://www.kaggle.com/>) 帳號

使用 pip 安裝 kaggle 套件 (Done)

run in command line

```
pip install kaggle
```

參與三個 Getting Started 競賽

- [Titanic \(https://www.kaggle.com/c/titanic\)](https://www.kaggle.com/c/titanic)
- [House Prices: Advanced Regression Techniques \(https://www.kaggle.com/c/house-prices-advanced-regression-techniques\)](https://www.kaggle.com/c/house-prices-advanced-regression-techniques)
- [Digit Recognizer \(https://www.kaggle.com/c/digit-recognizer\)](https://www.kaggle.com/c/digit-recognizer)

建立新的 API 憑證

- My Account
- Create New API Token

將 kaggle.json 置放在

- UNIX: `/Users/USERNAME/.kaggle/kaggle.json`
- Windows: `C:\Users\USERNAME\.kaggle\kaggle.json`

設定憑證讀寫權限

run in command line

```
chmod 600 /Users/USERNAME/.kaggle/kaggle.json
```

在 Google Colab 中執行

```
!mkdir /root/.kaggle
import json
token = {"username": "YOUR-USERNAME", "key": "YOUR-KEY"}
with open('/root/.kaggle/kaggle.json', 'w') as file:
    json.dump(token, file)
!chmod 600 /root/.kaggle/kaggle.json
```



```
In [3]: !kaggle competitions download -c house-prices-advanced-regression-techniques --force
```

```
Downloading sample_submission.csv to /Users/kuoyaojen/intro-to-tensorflow
 0%|          | 0.00/31.2k [00:00<?, ?B/s]
100%|          | 31.2k/31.2k [00:00<00:00, 336kB/s]
```

```
Downloading test.csv to /Users/kuoyaojen/intro-to-tensorflow
100%|          | 441k/441k [00:02<00:00, 158kB/s]
```

```
Downloading train.csv to /Users/kuoyaojen/intro-to-tensorflow
100%|          | 450k/450k [00:01<00:00, 314kB/s]
```

```
Downloading data_description.txt to /Users/kuoyaojen/intro-to-tensorflow
 0%|          | 0.00/13.1k [00:00<?, ?B/s]
100%|          | 13.1k/13.1k [00:00<00:00, 844kB/s]
```

```
In [4]: !kaggle competitions download -c digit-recognizer --force
```

Downloading train.csv to /Users/kuoyaojen/intro-to-tensorflow

[illegible]

Downloading test.csv to /Users/kuoyaojen/intro-to-tensorflow

```
100% |██████████████████████████████████████████████████████████████████████████| 48.8M/4  
8.8M [00:37<00:00, 1.51MB/s]
```

Downloading sample submission.csv to /Users/kuoyaojen/intro-to-tensorflow

```
0% |          0.00/235k [00:00<?, ?B/
s ]
```

```
k/235k [00:00<00:00, 3.16MB/s]
```

```
!kaggle datasets download -d zalando-research/fashionmnist
```

Downloading fashionmnist.zip to /Users/kuoyaojen/intro-to-tensorflow

[illegible]

如果碰到 403-Forbidden 呢？

檢查是否有加入 Competition

隨堂練習

建立瀏覽器的開發環境

在 Google Colab 完成 Hello TensorFlow

```
import tensorflow as tf

hello = tf.constant('Hello, TensorFlow!')
with tf.Session() as sess:
    print(sess.run(hello))
```

建立本機端的開發環境

在 TensorFlow 環境完成 Hello TensorFlow

```
import tensorflow as tf

hello = tf.constant('Hello, TensorFlow!')
with tf.Session() as sess:
    print(sess.run(hello))
```


建立本機端的開發環境

在 TensorFlow 環境完成 Hello TensorBoard]

```
In [6]: import tensorflow as tf

x = tf.constant(8, name="x")
y = tf.constant(7, name="y")
tf_add = tf.add(x, y)

with tf.Session() as sess:
    writer = tf.summary.FileWriter('./graphs/tf_add', sess.graph)
    print(sess.run(tf_add))
writer.close()
```

啟動 TensorBoard

```
tensorboard --logdir=./graphs/tf_add --host=127.0.0.1
```

```
## Starting TensorBoard b'41' on port 6006
```

```
## (You can navigate to http://127.0.0.1:6006)
```


打開 TensorBoard


- 打開瀏覽器前往 <http://127.0.0.1:6006> (<http://127.0.0.1:6006>).
- 點選 GRAPHS 頁籤

We are all set!

TensorBoard

SCALARSIMAGESAUDIO**GRAPHS**DISTRIBUTIONSHISTOGRAMSEMBEDDINGS

 Fit to screen

 Download PNG

Run
(1)

Session
runs (0)


Upload

Trace inputs ☐

Color ☒ Structure
☐ Device

colors ☐ same substructure
☐ unique substructure

Main GraphAuxiliary Nodes



以 Kaggle Dataset API 讀入 titanic 並顯示前五列


```
In [8]: import pandas as pd
```

```
df = pd.read_csv("train.csv")  
df.head()
```

Out[8]:

	PassengerId	Survived	Pclass	Name	Sex	Age	SibSp	Parch	Ticket	Fare	Cabin	Embarked
0	1	0	3	Braund, Mr. Owen Harris	male	22.0	1	0	A/5 21171	7.2500	NaN	S
1	2	1	1	Cumings, Mrs. John Bradley (Florence Briggs Th...	female	38.0	1	0	PC 17599	71.2833	C85	C
2	3	1	3	Heikkinen, Miss. Laina	female	26.0	0	0	STON/O2. 3101282	7.9250	NaN	S
3	4	1	1	Futrelle, Mrs. Jacques Heath (Lily May Peel)	female	35.0	1	0	113803	53.1000	C123	S
4	5	0	3	Allen, Mr. William Henry	male	35.0	0	0	373450	8.0500	NaN	S

以 Kaggle Dataset API 讀入 house-prices-advanced-regression-techniques 並顯示前五列

```
In [9]: !kaggle competitions download -c house-prices-advanced-regression-techniques --for
ce
```

[illegible][illegible][illegible]

```
Downloading data_description.txt to /Users/kuoyaojen/intro-to-tensorflow  
0%|  
s] | 0.00/13.1k [00:00<?, ?B/  
100%|███████████████████████████████████████████████████| 13.1k/  
13.1k [00:00<00:00, 863kB/s]
```

```
In [10]: df = pd.read_csv("train.csv")
df.head()
```

Out[10]:

	Id	MSSubClass	MSZoning	LotFrontage	LotArea	Street	Alley	LotShape	LandContour	Utilities	...	PoolArea	PoolQC	Fence
0	1	60	RL	65.0	8450	Pave	NaN	Reg	Lvl	AllPub	...	0	NaN	NaN
1	2	20	RL	80.0	9600	Pave	NaN	Reg	Lvl	AllPub	...	0	NaN	NaN
2	3	60	RL	68.0	11250	Pave	NaN	IR1	Lvl	AllPub	...	0	NaN	NaN
3	4	70	RL	60.0	9550	Pave	NaN	IR1	Lvl	AllPub	...	0	NaN	NaN
4	5	60	RL	84.0	14260	Pave	NaN	IR1	Lvl	AllPub	...	0	NaN	NaN

5 rows × 81 columns

以 Kaggle Dataset API 讀入 digit-recognizer 並視覺化前十張
圖片

```
In [11]: !kaggle competitions download -c digit-recognizer --force
```

Downloading train.csv to /Users/kuoyaojen/intro-to-tensorflow

```
100% |██████████████████████████████████████████████████████████████████████████| 73.2M/7  
3.2M [01:50<00:00, 1.43MB/s]
```

Downloading test.csv to /Users/kuoyaojen/intro-to-tensorflow

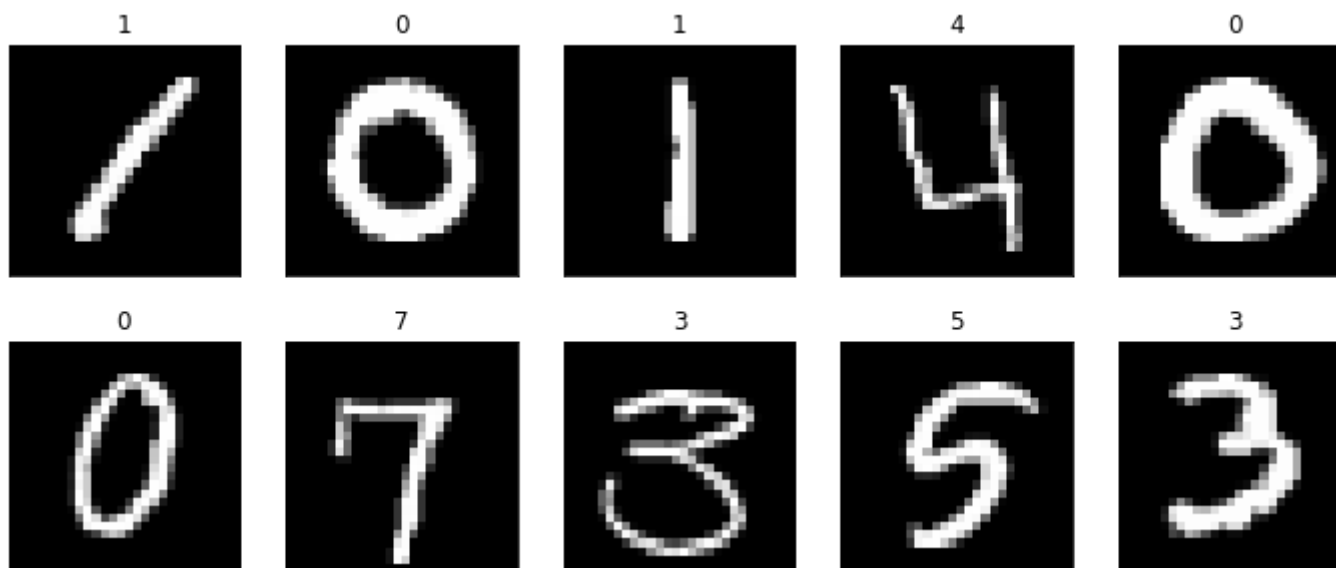
[illegible]

Downloading sample submission.csv to /Users/kuoyaojen/intro-to-tensorflow

100% | ██████████ 235
k/235k [00:00<00:00, 482kB/s]

```
In [14]: import matplotlib.pyplot as plt

train = pd.read_csv("train.csv")
fig, axes = plt.subplots(2, 5, figsize=(12, 5))
for i in range(10):
    obs = train.loc[i, "pixel0":].values.reshape(28, -1)
    row_i, col_i = i//5, i%5
    subplot = axes[row_i, col_i]
    subplot.imshow(obs, cmap="gray")
    subplot.set_xticks([])
    subplot.set_yticks([])
    subplot.set_title(train.loc[i, "label"])
plt.show()
```



以 Kaggle Dataset API 讀入 fashionmnist 並視覺化前十張圖片

```
In [ ]: !kaggle datasets download -d zalando-research/fashionmnist  
        !unzip fashionmnist.zip
```



```
In [16]: import pandas as pd
import matplotlib.pyplot as plt

ref_dict = {
    0: "T-shirt/top",
    1: "Trouser",
    2: "Pullover",
    3: "Dress",
    4: "Coat",
    5: "Sandal",
    6: "Shirt",
    7: "Sneaker",
    8: "Bag",
    9: "Ankle boot"
}
train = pd.read_csv("fashion-mnist_train.csv")
fig, axes = plt.subplots(2, 5, figsize=(12, 5))
for i in range(10):
    obs = train.loc[i, "pixel1":].values.reshape(28, -1)
    row_i, col_i = i//5, i%5
    subplot = axes[row_i, col_i]
    subplot.imshow(obs, cmap="gray")
    subplot.set_xticks([])
    subplot.set_yticks([])
    lab = train.loc[i, "label"]
    item = ref_dict[lab]
    subplot.set_title(item)
plt.show()
```

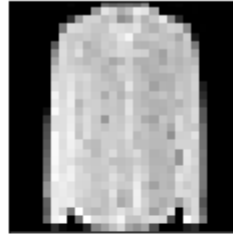
Pullover



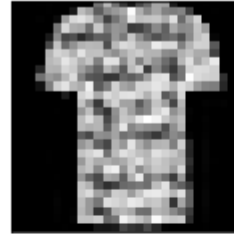
Ankle boot



Shirt



T-shirt/top



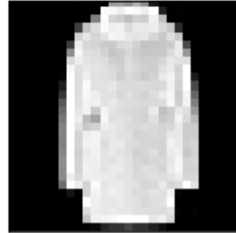
Dress



Coat



Coat



Sandal



Coat



Bag

