HW1 Supplemental

Yi-Lun Wu 吳易倫
yilun.ee08@nycu.edu.tw

Download Kaggle Dataset to Google Drive on Colab

- Kaggle $\rightarrow$ your profile $\rightarrow$ Account $\rightarrow$ Create New API token $\rightarrow$ upload token to google drive `./.kaggle/kaggle.json`.
- Mount google drive to `/content/drive`.

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

- Use `kaggle` package to download dataset.

```
! KAGGLE_CONFIG_DIR=/content/drive/MyDrive/.kaggle \
  kaggle competitions download \
    -c ccbda-2022-hw1 \
    -p ./drive/MyDrive
```

Download Kaggle Dataset to Google Drive on Colab

- Copy dataset to VM storage and unzip it.

```
! cp ./drive/MyDrive/ccbda-2022-hw1.zip .
! unzip ./ccbda-2022-hw1.zip -d ./data
```

- Check the number of samples.

```
! find ./data/train -name '*.mp4' | wc -l
! find ./data/test -name '*.mp4' | wc -l
```

```
30000
10000
```

- Start to read data from path `./data/train` and `./data/test`.

What is `requirements.txt`

- Export installed package names to `requirements.txt`.

```
pip freeze > requirements.txt
```

- Install package from `requirements.txt`.

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

- Manually create `requirements.txt`

```
pandas
tqdm
tensorboard
torch==1.11.0
torchvision==0.12.0
```

Better Practice

- Create virtual environment for a project and install the needed packages only.
  - virtualenv

```
python -m venv project_env_name
```

  - conda

```
conda create -name project_env_name python=3.9
```

  - pyenv

```
pyenv virtualenv 3.9.13 project_env_name
```