

---

---

# Machine Learning HW3

MLTAs

ntueemlta2021@gmail.com

---

---

# Outline

- HW3 - Face Expression Prediction
  - Dataset and Tasks Description
  - Sample Submission
- Kaggle
- Grading / Assignment Regulation

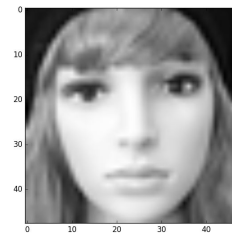
# Task - Face Expression Prediction



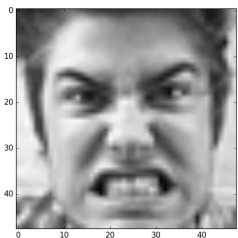
# Task - Face Expression Prediction

本次作業為網路上收集到的人臉表情資料，

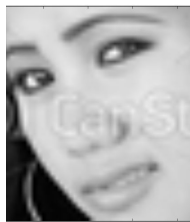
經過特殊處理，每張圖片，均是人臉部份佔大部分



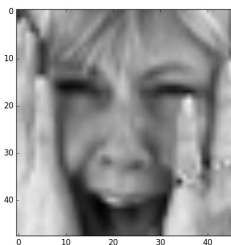
6(中立)



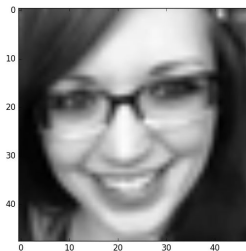
0(生氣)



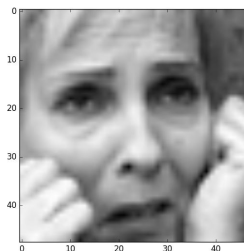
1(厭惡)



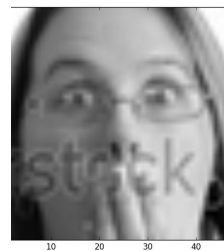
2(恐懼)



3(高興)



4(難過)



5(驚訝)

training data: about 28886 images

testing data: 7000 images (a half will be public test)

# Task and Dataset

- Task : **CNN**

- Build your own model
- Sample code:

<https://drive.google.com/file/d/10Q-K9rrcC9EolQgS-wZdS2byp0XhsLVz/view?usp=sharing>

- Dataset : Facial Expression Recognition Challenge (Fer2013)

Reference : [Challenges in Representation Learning](#)

# Kaggle Info & Deadline

- Link: <https://www.kaggle.com/t/89e76f1cd72e4f608c0394eded88dd00>
- 個人進行、不須組隊
- Team Name:
  - 修課學生: 學號\_任意名稱 (ex: b09901666\_name)
  - 旁聽: 旁聽\_任意名稱
- Maximum Daily Submission: 5 times
- Kaggle Deadline: 11/11/2021 23:59:59 (GMT+8)
- Ceiba Deadline: 11/13/2021 23:59:59 (GMT+8)
- test set的7000筆資料將被分為兩份, 3500筆public, 3500筆private
- Leaderboard上所顯示為public score, 在Kaggle Deadline前可以選擇2份submission作為private score的評分依據。

# Kaggle submission format

請預測test set中七千筆資料並將結果上傳Kaggle

1. 上傳格式為csv。
2. 第一行必須為id,label, 第二行開始為預測結果。
3. 每行分別為id以及預測的label, 請以逗號分隔。
4. Evaluation: Accuracy

1	id, label
2	0,0
3	1,0
4	2,0
5	3,0
6	4,0
7	5,0
8	6,0
9	7,0
10	8,0
11	9,0
12	10,0
13	11,0
14	12,0
15	13,0
16	14,0
17	15,0
18	16,0
19	17,0
20	18,0
21	19,0

# 配分 Grading Criteria-Kaggle(2%)

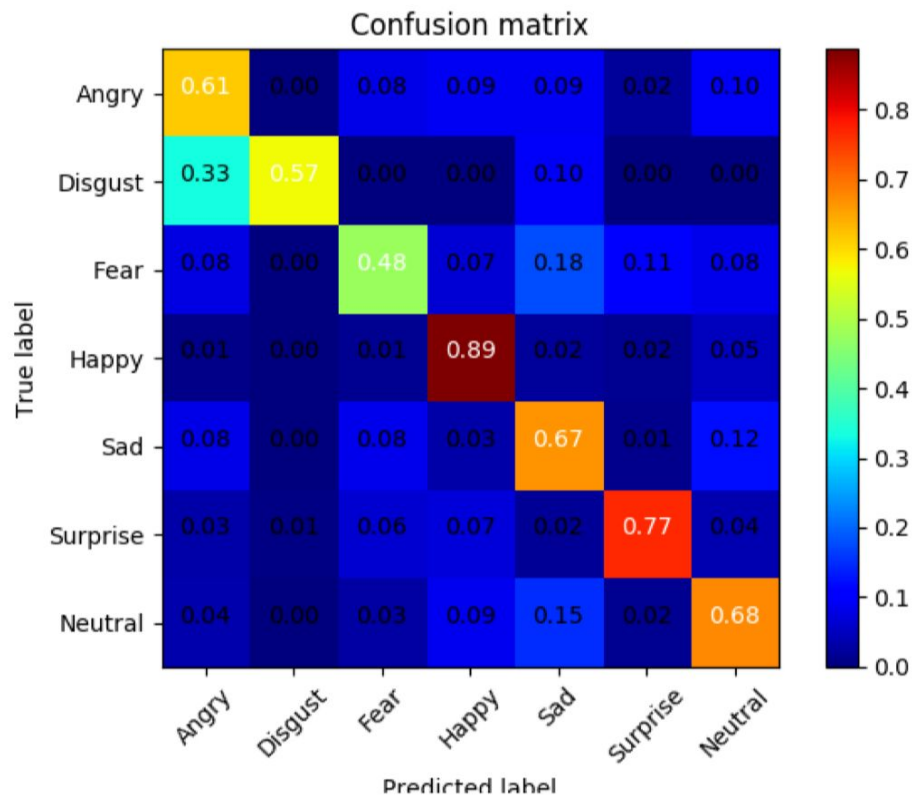
- Kaggle deadline : 11/11/2021 23:59:59 (GMT+8)
- Kaggle - 2%
  - ❑ 超過public leaderboard的simple baseline分數 : **0.5%**
  - ❑ 超過private leaderboard的simple baseline分數 : **0.5%**
  - ❑ 超過public leaderboard的strong baseline分數 : **0.5%**
  - ❑ 超過private leaderboard的strong baseline分數 : **0.5%**
- Bonus - 1%
  - (1.0%) private leaderboard 排名前五名, 並繳交投影片描述實作方法, 另外需錄製一份講解影片(少於三分鐘)作一個簡單的 presentation, 助教將公布給同學們參考



# 配分 Grading Criteria - report(8%)

- Programming Report - 4%
  - <https://docs.google.com/document/d/1KoGQML8xyo9pXN6HEfGgxgwOjXcK3v4/edit?usp=sharing&oid=102528209626787107677&rtpof=true&sd=true>
- Math Problem - 4%
  - [https://hackmd.io/@GfOkB4kgS66YhhM7j6TJew/Sly\\_akYUK](https://hackmd.io/@GfOkB4kgS66YhhM7j6TJew/Sly_akYUK)
  - Type in latex(preferable) or take pictures of your handwriting
- Write them in report.pdf

# Example - confusion matrix



# Assignment Regulation

- Only Python 3.7 available !!!!
- 開放使用套件(或是你可以直接下載我們當初的[環境yaml檔案](#))
  - numpy == 1.19
  - pandas == 1.1.3
  - python standard library
  - pytorch == 1.10.0 (torchvision == 0.11.1)
  - tensorflow == 2.1.0
  - keras == 2.2.4
  - cv2
  - pillow >= 6.1.0
- 若需使用其他套件，請儘早寄信至助教信箱詢問，並請闡明原因。

# Requirements

1. 請使用CNN[【機器學習 2021】卷積神經網路 \(Convolutional Neural Networks, CNN\) - YouTube](#)
  - 補充資料：[【機器學習 2021】類神經網路訓練不起來怎麼辦 \(三\):自動調整學習速率 \(Learning Rate\) - YouTube](#)
2. 不能使用額外data
3. 如果你的code不只一個檔案(或有多個參數)請附上readme或shell script
4. testing process要在10分鐘內跑完

# Ceiba Submissions

你的ceiba上請至少包含：

1. **report.pdf** : Please refer to report template
2. your python (or ipynb) files
3. 請將參數連結附在report中

請不要上傳dataset, 請不要上傳dataset, 請不要上傳dataset

# Report 格式

- 限制
  - 檔名必須為 report.pdf !!!
  - 檔名必須為 report.pdf !!!
  - 檔名必須為 report.pdf !!!
  - 請標明系級、學號、姓名，並按照report模板回答問題，切勿隨意更動題號順序
  - 若有和其他修課同學討論，請務必於題號前標明 collaborator (含姓名、學號)
- Report模板連結
  - 連結  
: <https://docs.google.com/document/d/1KoGQMQL8xyo9pXN6HEfGgxgwOjXcK3v4/edit?usp=sharing&ouid=102528209626787107677&rtpof=true&sd=true>
- 截止日期同 Ceiba Deadline: 11/13/2021 23:59:59 (GMT+8)

## 其他規定 Other Policy

- Lateness
  - Ceiba 每遲交一天(不足一天以一天計算) hw3 所得總分將  $\times 0.7$
  - 不接受程式 or 報告單獨遲交
  - 不得遲交超過一天, 若有特殊原因請儘速聯絡助教
- Runtime Error
  - 當 程式錯誤, 造成助教無法順利執行, 請在公告時間 內寄信向助教說明, 修好之後重新執行所得kaggle部分分數將  $\times 0.5$ 。
  - 可以更改的部分僅限 syntax 及 io 的部分, 不得改程式邏輯或是演算法, 至於其他部分由助教認定為主。

# 其他規定 Other Policy



- Cheating

- 抄 code、抄 report (含之前修課同學)
- 開設 kaggle 多重分身帳號註冊 competition
- 於訓練過程以任何不限定形式接觸到 testing data 的正確答案
- 不得上傳之前的 kaggle 競賽
- 教授與助教群保留請同學到辦公室解釋 coding 作業的權利, 請同學務必自愛



# TA Hour

- 11/3, 11/10(三) 14:15~15:15@ google meet
- ML2021 TA Hour
- 連結: <https://meet.google.com/zyi-gfgj-tdu>