

---

---

# Machine Learning HW2

MLTAs

ntueemlta2023@gmail.com

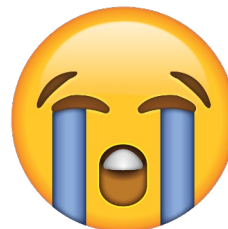
---

---

# Outline

- HW2 - Face Expression Prediction
  - Dataset and Tasks Description
  - Sample Submission
- Kaggle
- Grading / Assignment Regulation

# Task - Face Expression Prediction



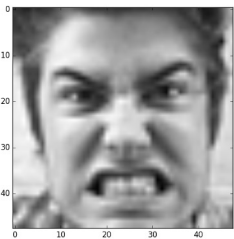
# Task - Face Expression Prediction

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



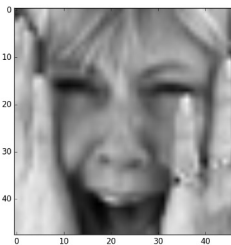
6(驚訝)



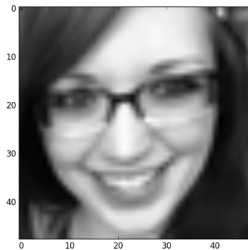
0(生氣)



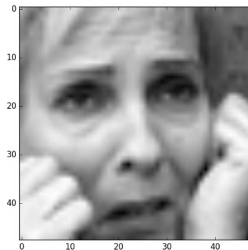
1(厭惡)



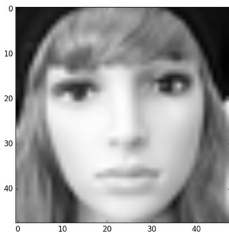
2(恐懼)



3(高興)



4(難過)



5(中立)

# Task and Dataset

- Task : **CNN**

- Build your own model
- Sample code:

<https://colab.research.google.com/drive/1Lxs8RkBkWwlPHqjw20K7urG90bjnyDiv?usp=sharing>

- Dataset : Facial Expression Recognition Challenge (Fer2013)

Reference : [Challenges in Representation Learning](#)

# Kaggle Info & Deadline

- Link: <https://www.kaggle.com/competitions/ml2023-fall-hw2>
- 個人進行、不須組隊
- Team Name:
  - 修課學生：學號\_任意名稱 ( ex: b09901666\_OvErFiTtIng )
  - 旁聽：旁聽\_任意名稱
- Maximum Daily Submission: 5 times
- Kaggle Deadline: 10/27/2023 23:59:59 (GMT+8)
- Cool Deadline: 10/29/2023 23:59:59 (GMT+8)
- 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%)

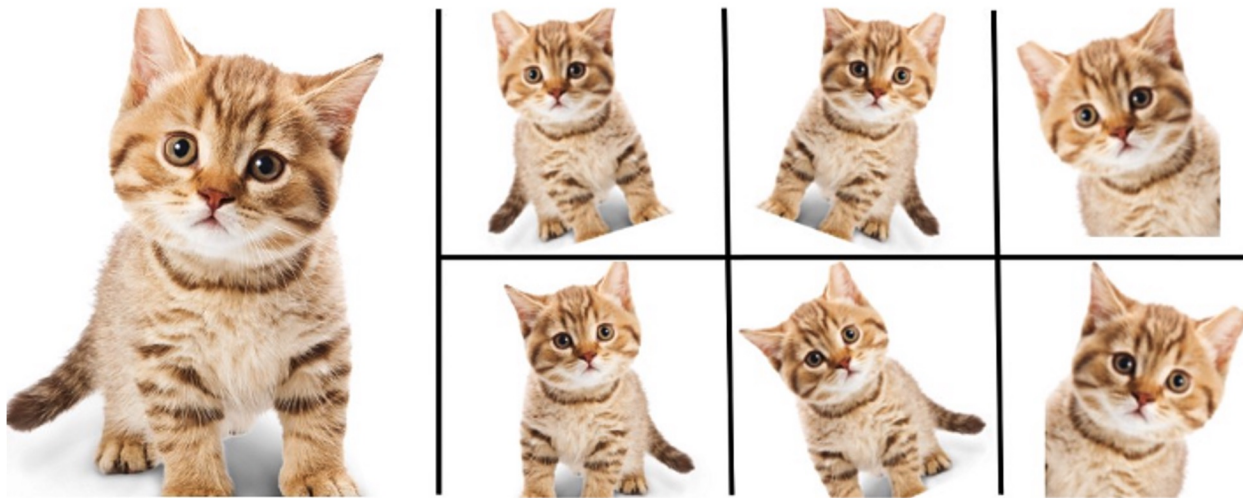
- Kaggle deadline : 10/27/2023 23:59:59 (GMT+8)
- Kaggle - 2%
  - ❑ 超過public leaderboard的simple baseline分數 : **0.5%**
  - ❑ 超過private leaderboard的simple baseline分數 : **0.5%**
  - ❑ 超過public leaderboard的strong baseline分數 : **0.5%**
  - ❑ 超過private leaderboard的strong baseline分數 : **0.5%**



# 配分 Grading Criteria - report(10%)

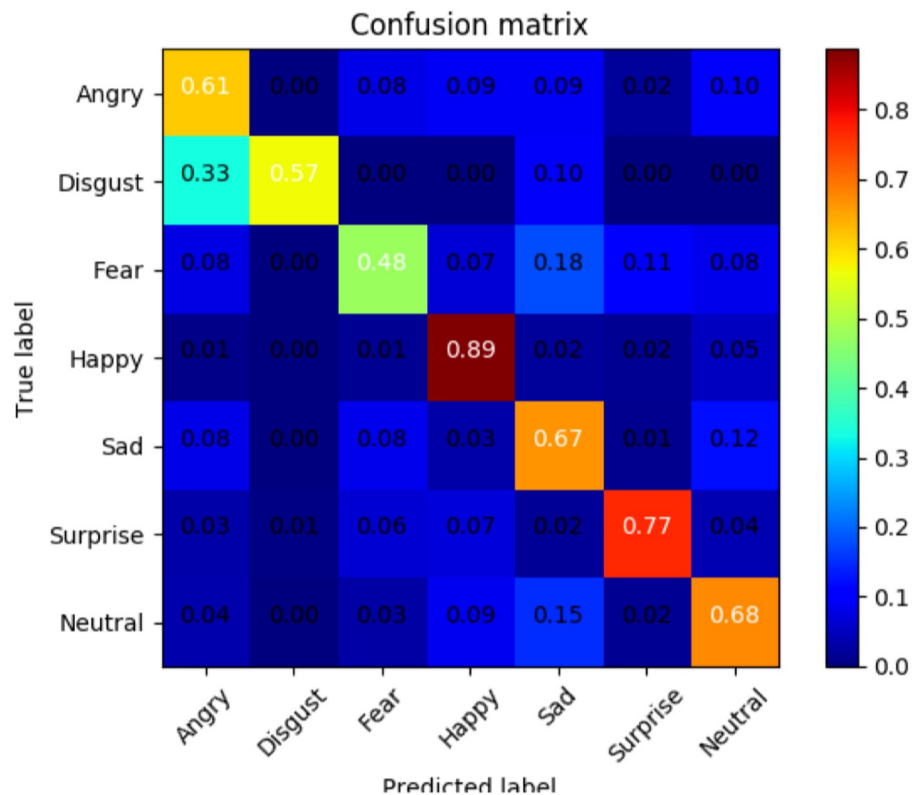
- Programming Report - 4%
  - [https://docs.google.com/document/d/17-\\_3\\_qAvuQMhCGIKns\\_apA8IOlktasso/edit?usp=sharing&oid=112465961449455869485&rtpof=true&sd=true](https://docs.google.com/document/d/17-_3_qAvuQMhCGIKns_apA8IOlktasso/edit?usp=sharing&oid=112465961449455869485&rtpof=true&sd=true)
- Math Problem - 6%
  - <https://ntueemlta2023.github.io/homeworks/hw2/ml-2023fall-hw2-math.pdf>
  - Type in latex(preferable) or take pictures of your handwriting
- Write them in report.pdf

# Data Augmentation



Enlarge your Dataset

# 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
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

# Cool Submissions

你的cool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(最佳model，或其他reproduce必須的檔案)附在report中

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

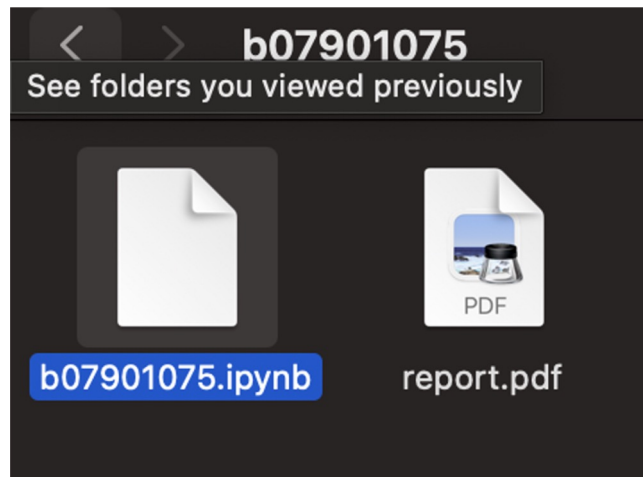
# Hand-in Format

- 一個 zip 檔案，檔案名稱為 **學號\_hw2.zip**，需包含

- 程式碼 ( 任意名稱.ipynb )

- 程式報告 + 數學題 ( **report.pdf** )

( 註：Cool 在繳交第二次以上時會將檔案名稱後綴 (-數字)，這不會影響同學的繳交，請同學放心 )



# 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/17-\\_3\\_qAvuQMhCGIKns\\_apA8IOlktasso/edit?usp=sharing&oid=112465961449455869485&rtpof=true&sd=true](https://docs.google.com/document/d/17-_3_qAvuQMhCGIKns_apA8IOlktasso/edit?usp=sharing&oid=112465961449455869485&rtpof=true&sd=true)

- 截止日期同Cool Deadline: 10/29/2023 23:59:59 (GMT+8)



## 其他規定 Other Policy

- Lateness
  - Cool 遲交一天(不足一天以一天計算)
  - 遲交一天以小時為單位線性遞減至七折，兩天從七折線性遞減零分
  - 不接受程式 or 報告單獨遲交
  - 有特殊原因請找助教。
- 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x0.5。
  - 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