Kaggle 教學

黃于真 韓宏光

Kaggle 簡介

Kaggle是一個資料科學競賽平台，每個人都可在其上發布資料集舉辦比賽，也可以參與其他人舉辦的比賽，期望透過競賽排名，集合眾人力量找出最好的模型。
(https://www.kaggle.com/)
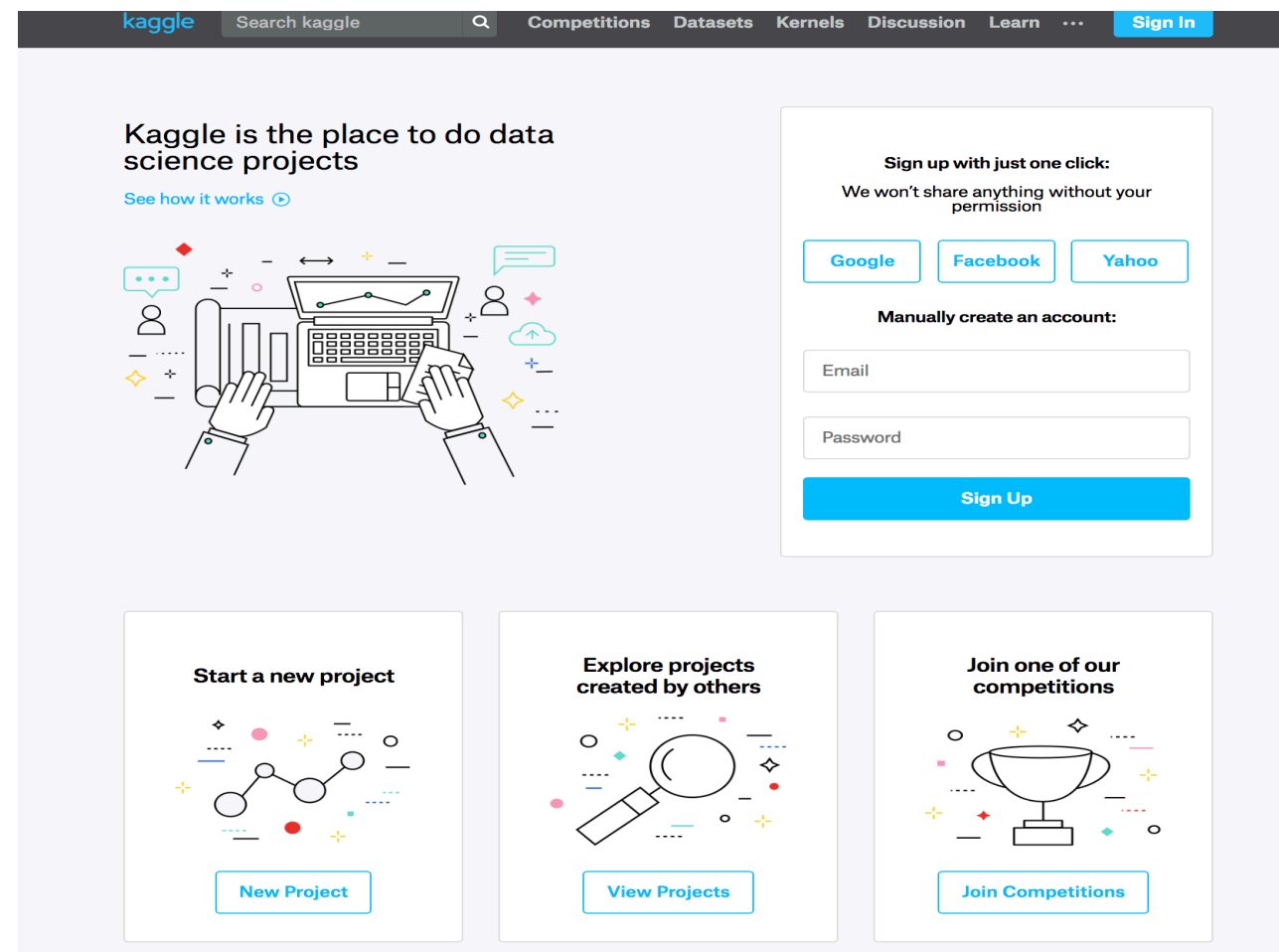
帳號註冊

○ 點選首頁右上角的sign in，選擇登入或註冊新帳號

Step 1: 連結google、fb、yahoo帳號 or 使用email註冊

Step 2: 填寫基本資料

Step 3: 同意隱私政策

Kaggle 競賽

- kaggle競賽分為兩類：general、inclass

- 作業三採inclass方式，即只有知道連結的人才能進入比賽

- 希望同學不要任意把連結給其他非修課相關的人！

- 連結如下：

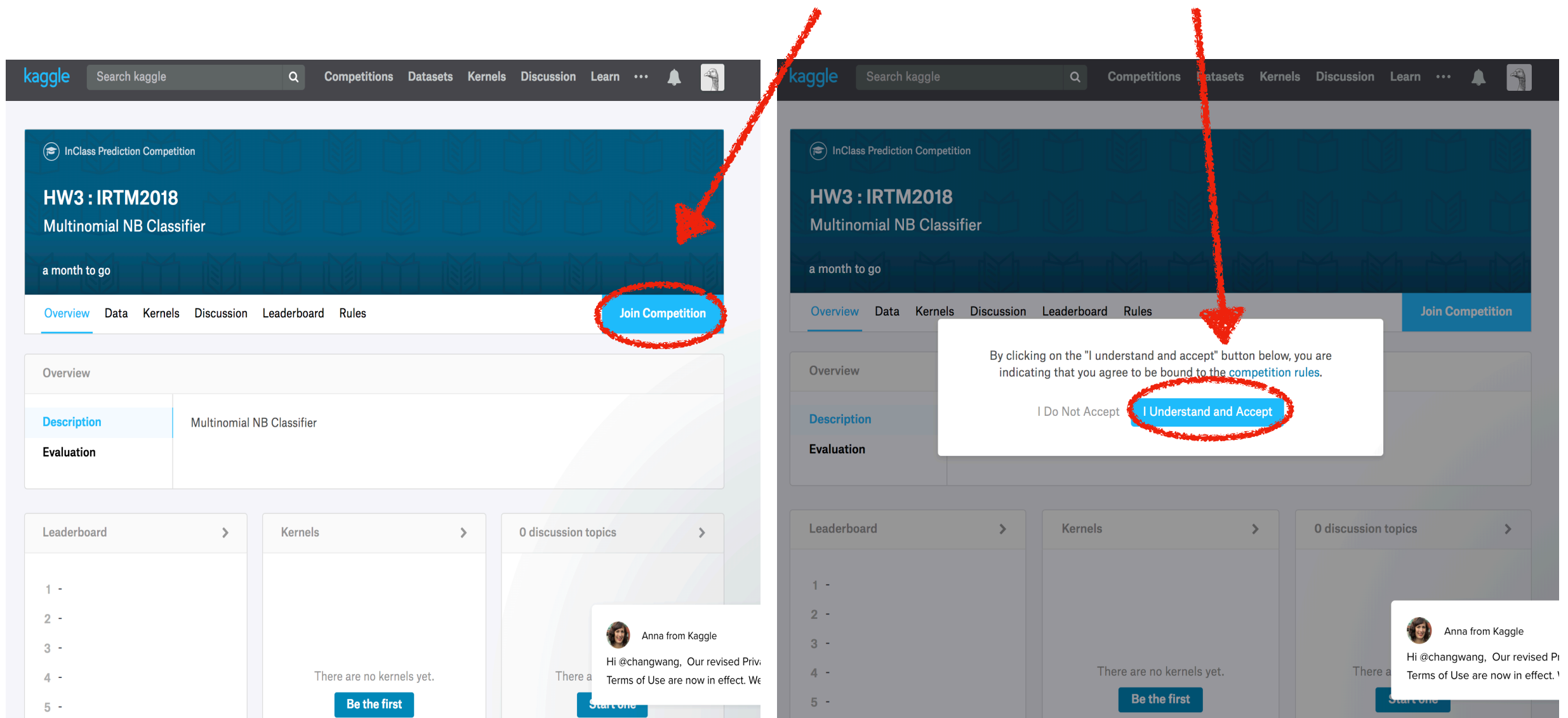 *https://www.kaggle.com/t/9982faa24bdb44c5a2b0d68e6944b607*

  (如果尚未登入，點選連結會導向登入頁面，但登入後有可能直接跳回首頁，此時建議先登入再點選連結，就可順利進入比賽頁面)
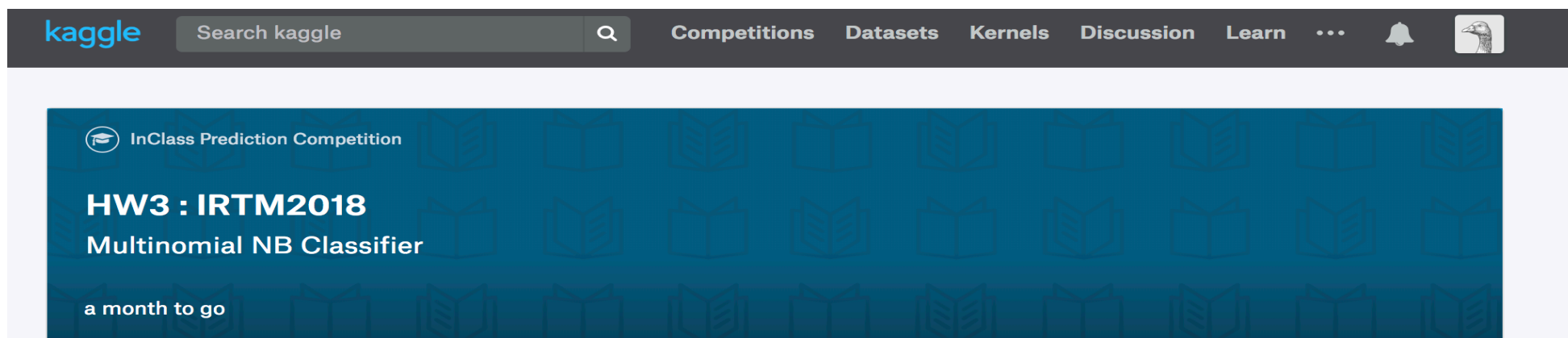
參加競賽

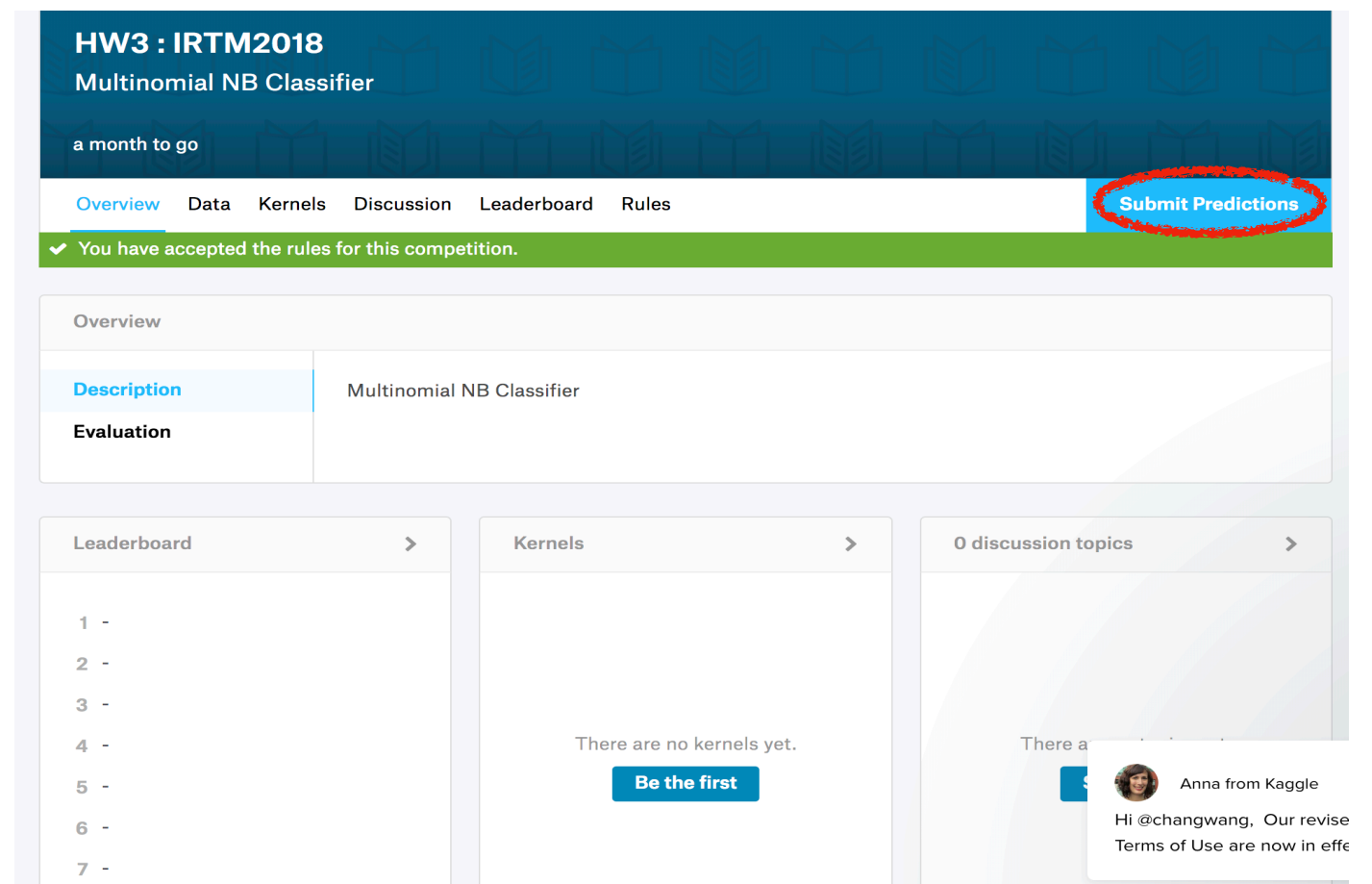○ 成功進入競賽畫面後，點選右上角的Join Competition，再同意競賽規則即可

上傳注意事項

- 成功參加比賽後，頁面右上角即會變成Submit Predictions

- 關於上傳結果有以下事項需要注意：

  - 注意上傳格式！

  - 注意每日上傳次數上限5次！

  - 請更改隊名為學號！

上傳格式

- 請使用csv檔，其他檔會無法上傳

- 需有兩欄，分別為Id和Value

- 兩欄資料型態請用文字！！！

  - kaggle採用比對方式評分，型態不同可能導致比對失敗拿不到分數

  - id部分須是連續的

sub

Id	Value
1	0
2	0
3	0
4	0
5	0

計分方式

- 排名採用micro-averaged F1-score

  - 右圖結果為0.6

- 由於kaggle只能提供單一評分方式，
  因此作業截止後，會再計算其他
  micro/macro-averaging precision、
  recall，公布在ceiba評語中

sub	
Id	**Value**
1	1
2	2
3	3
4	0
5	0

sol	
Id	**Value**
1	1
2	2
3	3
4	4
5	5

上傳預測結果

- 點選右上角Submit Predictions，就進入上傳頁面

- 若上傳等很久，可以重開頁面再上傳，未上傳成功不會計入上傳次數

- 設次數限制是怕大家無限上傳去試出答案

點選可以更改隊名

拖入檔案

剩餘上傳次數和reset時間

等待上傳完成

Make a submission for TA_test

You have 5 submissions remaining today. This resets 14 hours from now (00: 00 UTC).

Step 1
Upload submission file

Upload Submission File

ans.csv (33 B)

Complete   100%   33 B

File Format
Your submission should be in CSV format.
You can upload this in a zip/gz/rar/7z
archive, if you prefer.

Number of Predictions
We expect the solution file to have 5 prediction rows. This file
should have a header row. Please see sample submission file on
the data page.

Step 2
Describe submission

B *I*   🔗 ❝ </> 🖼   ☰ ☰ H ↦   ↺ ↻   M↓ Styling with Markdown supported

Briefly describe your submission.

Make Submission

預測結果評分

有時候系統負荷過大，可能評分很久，可以嘗試重開頁面，看看是否有評分成功，即有顯示出新的分數，若失敗的話不會有紀錄，也就不會用到上傳次數，重新再上傳即可

更改隊名

- kaggle提供組隊功能，但這次有限制只能一人一隊

- 為了方便掌握大家的上傳結果，隊名開頭請一律用學號，底線後則不限制

- 例：r06725123_xxxx

- 輸入新隊名後按Save Team Name即可