

Day 48

機器學習

Kaggle 介紹

Home of Data Science



Kaggle

- Kaggle 為全球資料科學競賽的網站，許多資料科學的競賽均會在此舉辦，吸引全球優秀的資料科學家參加
- Kaggle 上的競賽者都非常厲害，對機器學習技術沒有一定的掌握是沒辦法拿到好名次的
- 名列前茅的參賽者都會分享自己的作法與程式碼供參考，非常具有參考價值

排行榜

- 主辦單位通常會把測試資料分為 public set與 private set，參賽者上傳預測結果可以看到 public set 的成績，但比賽最終會使用 private set 的成績作為排名

Public Leaderboard

Private Leaderboard

This leaderboard is calculated with all of the test data.

[Raw Data](#)









[Refresh](#)

In the money

Gold

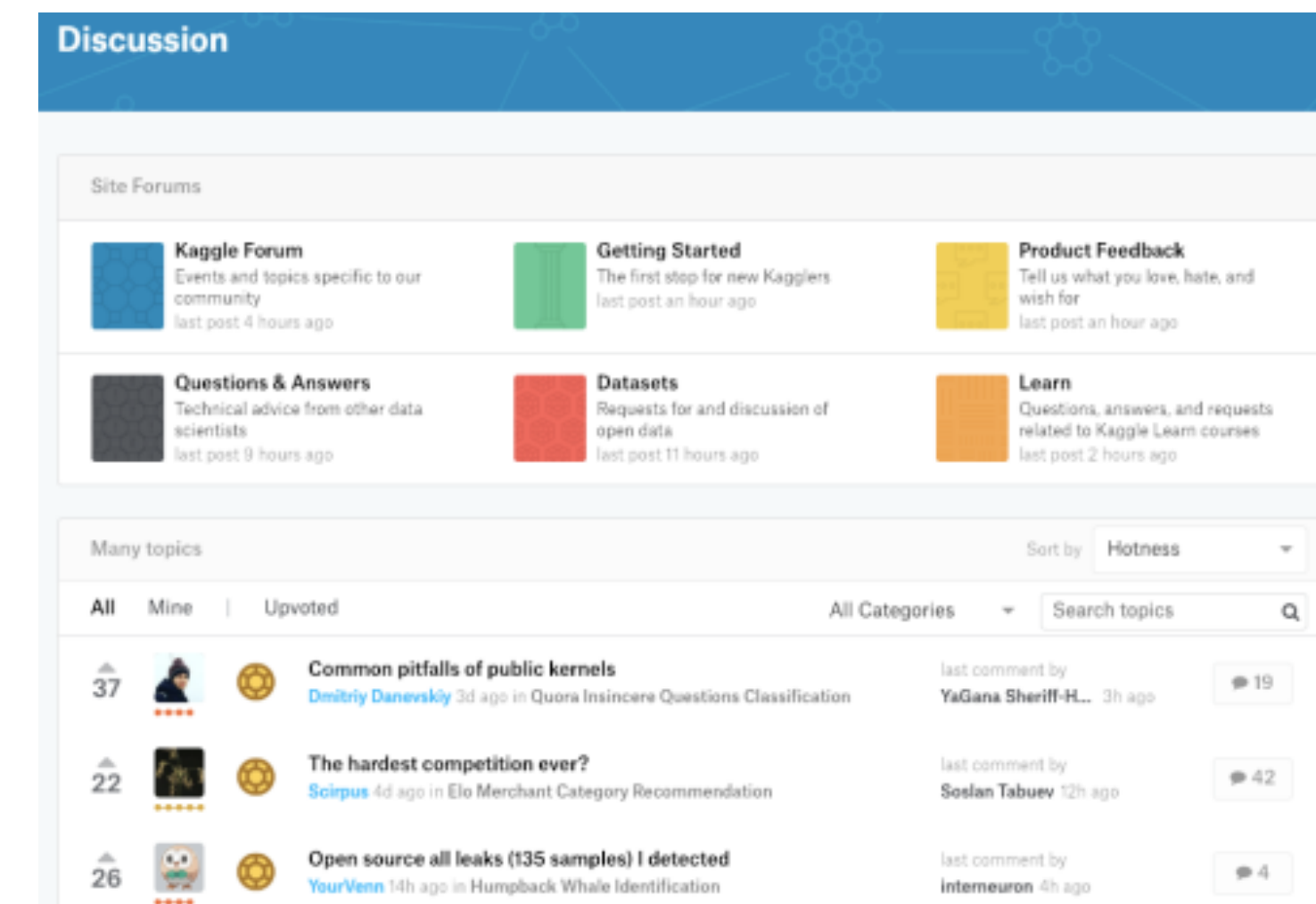
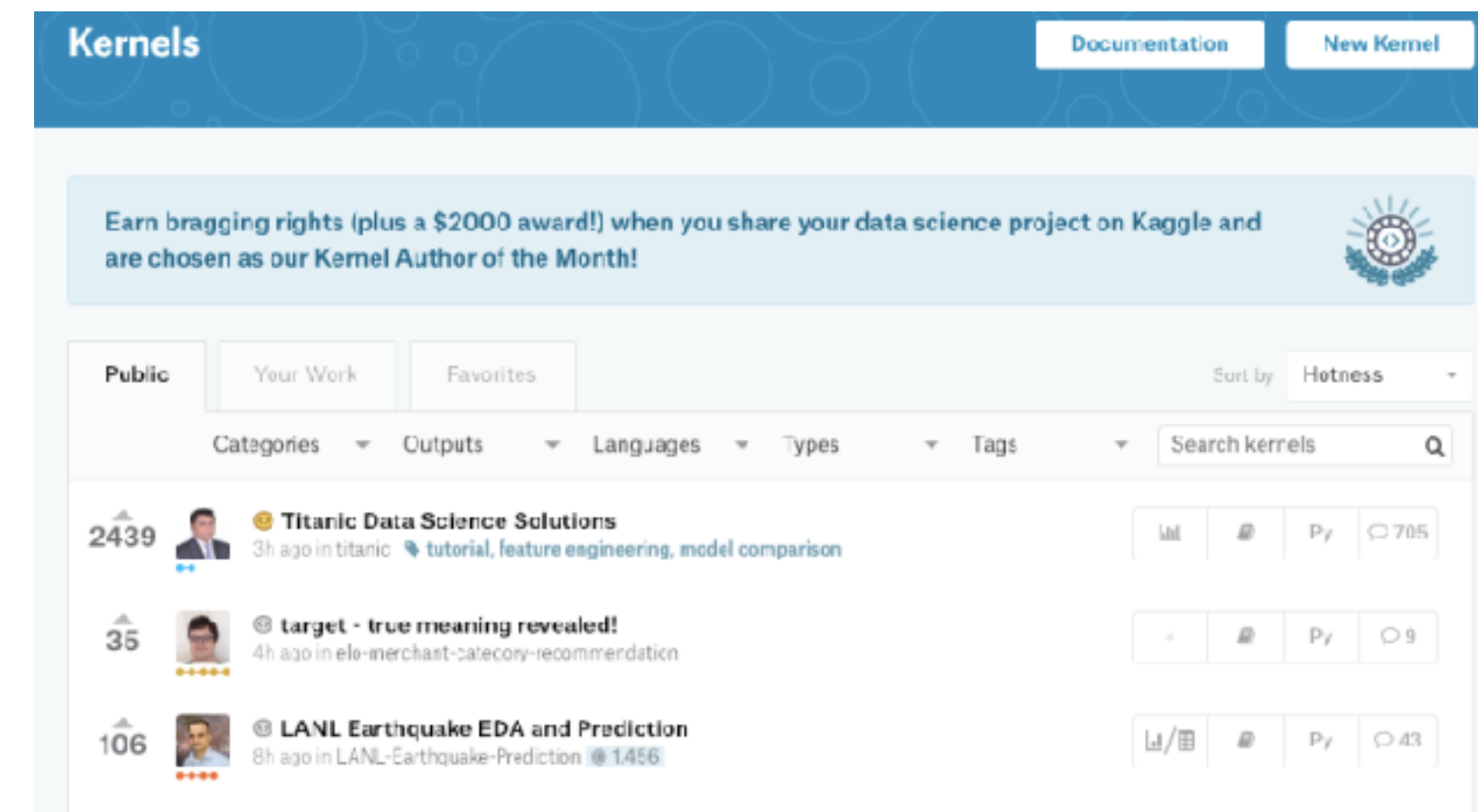
Silver

Bronze

#	△1w	Team Name	Kernel	Team Members	Score 🏆	Entries	Last
1	—	Dmytro Poplavskiy [ods.ai]			0.260	71	10h
2	—	Ian Pan & Alexandre Cadrin		 	0.260	189	5h
3	—	thesix		    +7	0.253	279	7h
4	new	chiyuwu			0.245	21	20h

Kernels 與 Discussion

- Kernels 可以看見許多高手們分享的程式碼與結果，多半會以 jupyter notebook 呈現
- Discussion 可以看到高手們互相討論可能的做法，或是資料中是否存在某些問題



復現結果

- 機器學習的專案，常常需要參考別人的論文或 Github，然後重現結果，其中可能會因為環境、套件或是理解不對而產生不對的結果
- 學習如何復現別人的成果是機器學習中非常重要的技能

解題時間 It's Your Turn

請跳出PDF至官網Sample Code & 作業
開始解題

