

天氣與人工智慧 期末考說明

2020/12/30

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 LeaderboardPrivate Leaderboard

This leaderboard is calculated with all of the test data.

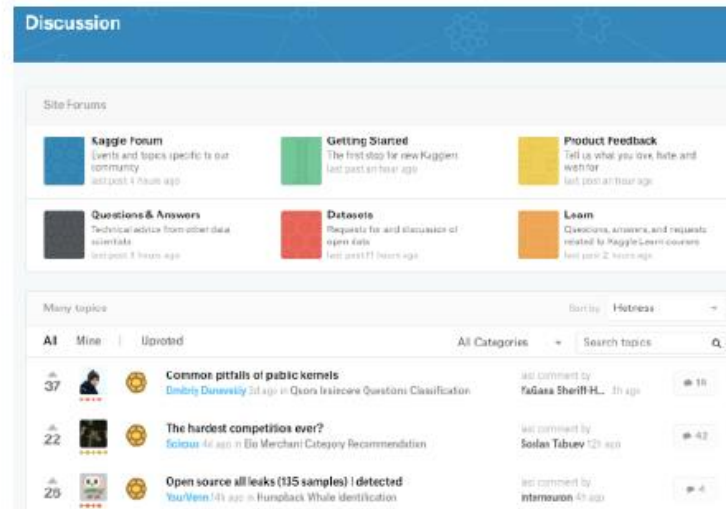
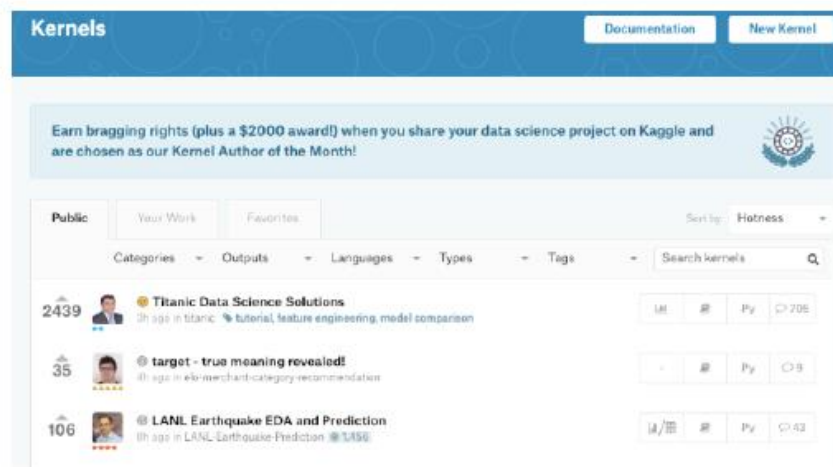
[Raw Data](#)[Refresh](#)

In the moneyGoldSilverBronze

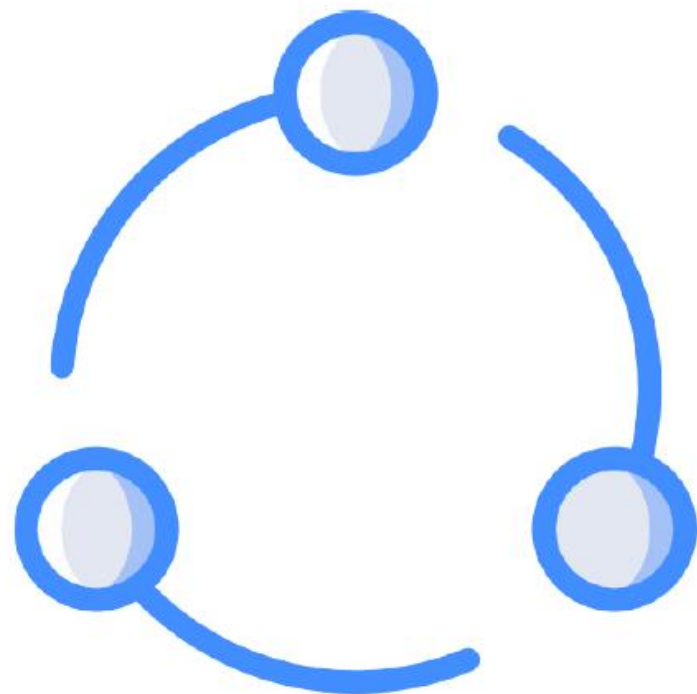
#	Δ1w	Team Name	Kernel	Team Members	Score 🏆	Entries	Last
1	—	Dmytro Poplavskiy [ods.ai]			0.260	71	10h
2	—	Ian Pan & Alexandre Cadrin		 	0.260	189	5h
3	—	thesix		   	0.253	279	7h
4	new	chiyuwu			0.245	21	20h

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



重要知識點



- 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

期末考 - Kaggle 實戰

房價預測 (40%)

鐵達尼號生存預測 (30%)

計程車費預測 (20%)

程式碼完整度與註解說明 (10%)

- 課堂中很多練習都是以這幾題做為範例，請各位結合上課所學到的，包含特徵工程、模型的選用等等，來檢視學習成效。

Kaggle 帳號驗證範例



Che-Wei Chou

Joined 3 years ago · last seen in the past day



Competitions
Novice

成績截圖範例 (My Submissions)

10 submissions for Che-Wei Chou		Sort by Public Score ▼
All Successful Selected		
Submission and Description		Public Score
house_baseline.csv a month ago by Che-Wei Chou add submission details		0.14224

程式碼上傳

請上傳 .ipynb 檔，

程式碼的說明以及整個 notebook 的完整度也會納入成績參考