

Machine Learning Final Project

ml2016ta@gmail.com

Outline

- Projects
- Schedule
- Scoring
- Policy
- Report
- Github
- Bonus

Projects

- Cyber Security Attack Defender
- Outbrain Click Prediction
- Transfer Learning on Stack Exchange Tags

Schedule

DATE	TODO
11/25	Final Project Releases
12/02 21:00:00	Team Makeup (Stage One) Release Simple Baseline
12/09	停修截止
12/16 21:00:00	Beyond Simple Baseline (Stage Two) Release Strong Baseline
12/23	Keep on moving moving.
12/30 21:00:00	Beyond Strong Baseline (Stage Three)
01/06 21:00:00	Top 10% Teams are Chosen
01/13	Top 10% Teams Presentation
01/18 (Wed) 21:00:00	Kaggle Ranking Deadline
01/20 21:00:00	Github (Report, src) Deadline

Scoring (65/60)

- Report : 30
- Team Makeup before **Stage One**: 5
- Beyond Simple Baseline before **Stage Two**: 4
- Beyond Strong Baseline before **Stage Three**: 4
- Beyond Simple Baseline before Kaggle Rank Deadline: 4
- Beyond Strong Baseline before Kaggle Rank Deadline: 4
- Ranking: 4
- Github: 5
- 組內互評: 5

Scoring details

- Team makeup (5%)
 - 填好組隊表單 (2~4人)
 - 在Kaggle的比賽上Merge成一隊 (包含所有的組員)
 - Before Stage one 12/02 21:00:00
- 如果你 “Beyond Simple Baseline before **Stage Two**” 代表你也會 “Beyond Simple Baseline before Kaggle Rank Deadline” 所以你會得到 4+4 分。同理 Strong baseline 。
- 可同時進行多個題目，但助教只會依照最後繳交報告的**一個**題目評分。
 - 如果在題目A “Beyond Simple Baseline before Stage Two” 但是你最後繳交的報告是題目B，則你不會拿到題目A的任何分數，一切以題目B為準。
 - 中途換題目會使你拿不到前面Stage的分數，建議大家盡早決定一個題目。
 - 組隊的分數要在kaggle中完成merge才算成功，因此每個比賽的組隊是獨立計分。
- 提醒：可以先在每一個題目都組隊完成，如果等到公布Simple baseline才組隊選擇題目會害你拿不到組隊的分數。

Scoring details (cont)

- Ranking (4%)
 - beyond 30% : **4**
 - beyond 50% : **3**
 - beyond 70% : **2**
 - beyond sample submission : **1**
- 所有的Kaggle分數都是依據Public set ◦ (beyond baseline , ranking判定)

Policy

- Kaggle Team Name: " NTU_隊長學號_隊伍名稱 "
- Team merge before stage one
- Feel free to use any toolkit.
- Abide by the rules of the competition.
- No cheating.
- Team sheet: <https://goo.gl/mSYNIs>

Report -- Rules

- 12 號字
- 單行間距
- 至少 6 頁，至多 12 頁 (含圖表)
- 報告中不須附上 code

Report -- Contents

- **請清楚標示你選的題目 (* 1/0)**
 - Team name, members and your work division (5)
 - Preprocessing/Feature Engineering (5)
 - Model Description (At least two different models) (5+5)
 - Experiments and Discussion (5+5)

Github -- what it should contain?

- Team leader's github account !
- Only master, no branch.
- Directory "final" should contain:
a directory "src" with the source codes you write
a PDF file "Report.pdf" which is your final project report

ML2016/final

```
|
|----- /src
|         |----- codes
|         |_____ ...
|
|----- Report.pdf
|
|_____ any other files
```

BONUS a.k.a 大天使的祝福

- If you are the top 10% team in the kaggle competition, you are welcomed to give a presentation on 2017/01/13
- Presentation time: 12 minutes
- Bonus score for presentation: 2 point
- The presentation would be voted.
The top three presentation teams can get additional bonus score:
1st -> 3
2nd -> 2
3rd -> 1

Deadline

- Stage ONE deadline: 2016/12/02 21:00:00
- Stage TWO deadline: 2016/12/16 21:00:00
- Stage THREE deadline: 2016/12/30 21:00:00
- Kaggle deadline: 2017/01/18 (三) 21:00:00
- Github deadline: 2017/01/20 (五) 21:00:00

Kaggle score會依照每一個stage及此投影片所公布的時間為準，不需理會Kaggle上設定之deadline。

Late : <https://goo.gl/mufC6C> (70% discount per day. At most 48 hours late)

QAQ

Question and Questions?