# CRF Exercise

Yi Ting Ding

#### Environment

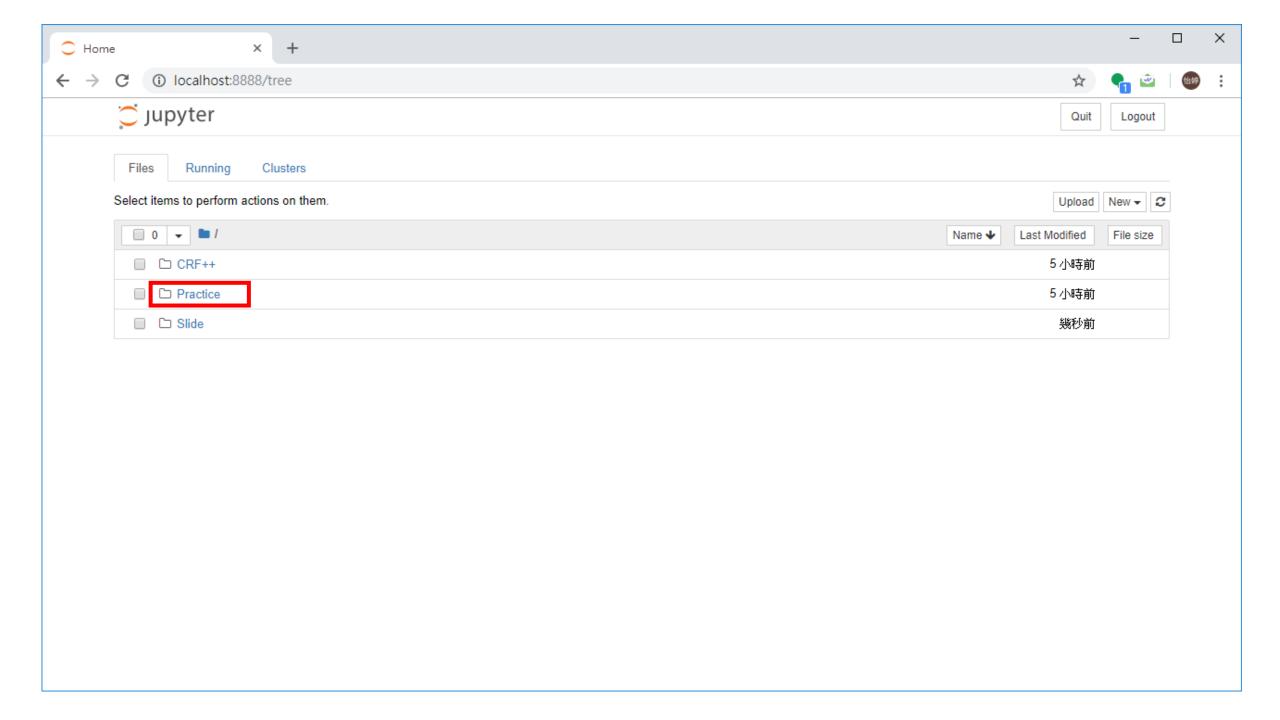
- Python 3.6
- Jupyter Notebook
- CRF++

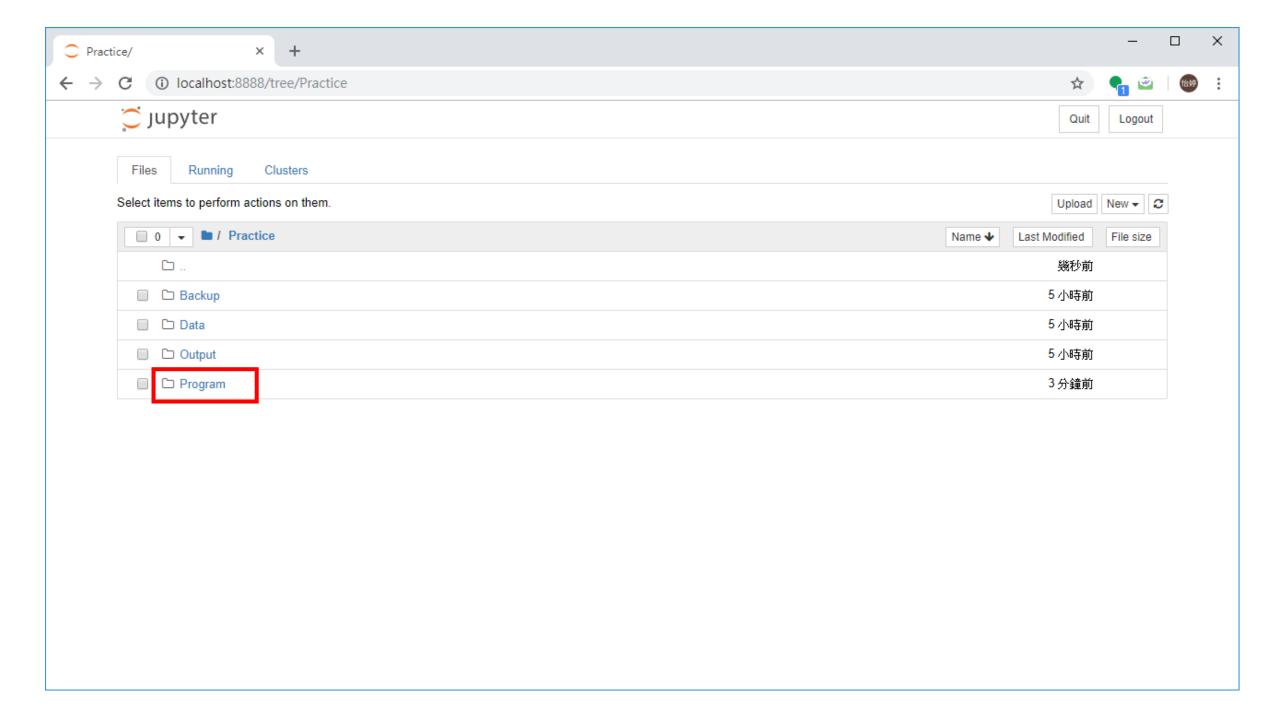
### Download

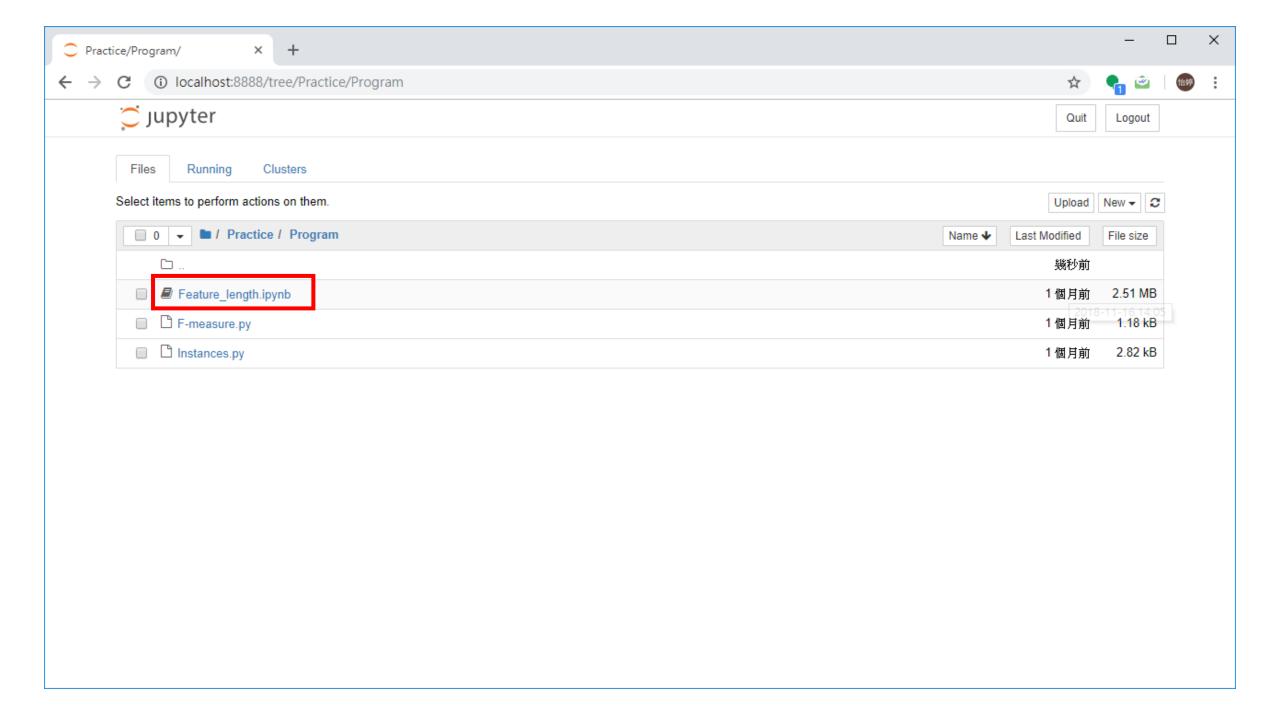
• Link — https://github.com/IKMLab/CRF\_tutorial

- Your path/Aldea CRF
  - Windows cmd -> /Your path/CRF Data -> jupyter notebook

```
📆 命令提示字元 - jupyter notebook
                                                                                                               Microsoft Windows [版本 10.0.17134.407]
(c) 2018 Microsoft Corporation. 著作權所有,並保留一切權利。
C:\Users\Ding>cd ../
C:\Users>cd ../
C:\>cd AI Cup/AIdea CRF
C:\AI Cup\AIdea CRF>jupyter notebook
[W 15:21:08.593 NotebookApp] Terminals not available (error was No module named 'winpty.cywinpty')
  16:21:08.695 NotebookApp] Serving notebooks from local directory: C:\AI Cup\AIdea CRF
   16:21:08.695 NotebookAppl The Jupyter Notebook is running at:
   16:21:08.695 NotebookApp] http://localhost:8888/?token=1bd894bc1755fde0146017e058a77dae95eb91344c9bd3f7
  16:21:08.695 NotebookApp] Use Control-C to stop this server and shut down all kernels (twice to skip confirmation).
[C 16:21:08.725 NotebookApp]
   Copy/paste this URL into your browser when you connect for the first time,
    to login with a token:
       http://localhost:8888/?token=1bd894bc1755fde0146017e058a77dae95eb91344c9bd3f7
  16:21:08.913 NotebookAppl Accepting one-time-token-authenticated connection from ::1
```







#### **Data**

```
Sequential→0
 2 involvement **O
 3 of → 0
   and*0
   Egr B-Gene
   transcription—⊸I-Gene
8 factors I-Gene
  in⊸o
10 FasL──B-Gene
   regulation⊸0
12 . → 0
13
14 The * 0
15 | critical → 0
16 function—⊸0
   of⊸i0
18 NFAT → B-Gene
19 | proteins── ×I-Gene
20 | in → 0
21 maintaining*0
22 lymphoid → 0
23 homeostasis ** O
24 was∗0
25 revealed—⊸0
26 | in→0
   mice---→0
```

```
Sequential→10→0
2 involvement 11 - 0
3 of→2---+0
4 NFAT → 4 → B-Gene
   and∗3—×0
   Egr∗3-----B-Gene
7 transcription—>13—>I-Gene
8 factors ≈ 7 → I-Gene
9 in→2—×0
   FasL → 4 → B-Gene
   regulation→10→0
12 .─>1 -> 0
13
14 The 3 → 0
15 critical—×8—×0
16 function——×8——×0
   of → 2 — → 0
18 NFAT → 4 → B-Gene
19 proteins → 8 → I-Gene
20 in→2---×0
21 maintaining 11-0
22 lymphoid ×8 ×0
23 homeostasis 11 → 0
24 was∗3—×0
25 revealed → 8 → 0
26 in→2 → 0
   mice → 4 → 0
```

### **CRF Command**

New Windows cmd

cd /your path/Aldea CRF/CRF++

×

#### **CRF Command**

- Train
  - crf\_learn.exe template <Path/Train Data> <Path/Model>
  - crf\_learn.exe template ../Practice/Data/Train\_len.data ../Practice/Output/Train\_model
- Test
  - crf\_test.exe -m <Path/Model> <Path/Test Data> >> <Path/Output Data>
  - crf\_test.exe -m ../Practice/Output/Train\_model ../Practice/Data/Test\_len.data >> ../Practice/Output/Result.txt

#### Evaluation

• cd ../Practice/Program

- python Entities.py <Path/Output Data> <Path/New Filename>
  - python Entities.py ../Output/Result.txt ../Output/Entity.txt
- Python F-measure.py <Path/New Filename>
  - python F-measure.py ../Output/Entity.txt

## Goal

#### Baseline

- Precision = 0.5364
- Recall = 0.2430
- F-score = 0.3343