Final Project for NLP

Reading Comprehension

Machine Reading Comprehension

- "A machine **comprehends** a passage of **text** if, for any **question** regarding that text that can be **answered** correctly by a majority of native speakers, that machine can provide a string which those speakers would agree both answers that question, and does not contain information irrelevant to that question." (Christopher J.C. Burges, 2013)

- When you ask your smartphone a specific question like "when's the next flight to Atlanta?" or even "what's President Obama's current approval rating?" You'll be able to receive direct and concise answers on your device.

Task

- A MRC system should be built to read documents and answer questions about their content.
- **The task is to answer the question with a noun/noun phrase to conform to the meaning of the document.**

- Document:

  Ovaries are the female part of the flower, which produces eggs that are needed for making seeds.

- The question:

  XXXXX produces the seeds.

- The answer choices:

  the flower   the leaves   the stem   the roots

Recommended

- Deep Learning Tools
  - Theano, Torch, TensorFlow, Keras, Numpy and so on
- You can try：
  - RNN, CNN, LSTM and so on
  - word embedding: word2vec, GloVe and so on
- Search Paper:
  - reading comprehension, question answering

Dataset-train/valid/test

- The document given consists of 20 consecutive sentences.

- The question is given in line 21 by replacing a single word with an anonymous placeholder token.

- There are 10 candidates for the missing word. Only one of them is the real answer.

```
1 I must go ! ''
2 the girl cried feverishly .
3 She was afraid Mrs. Cameron would try to prevent her going , and all at
once she knew that she could not bear that .
4 `` Must go ?
5 Where ?
6 Dinner is almost ready , and -- '' `` Oh , I do n't want any dinner .
7 I 'm going home -- I will sail over . ''
8 `` My dear child , do n't be foolish .
9 It 's too late to go over the harbour tonight .
10 They wo n't be expecting you .
11 Wait until the morning . ''
12 `` No -- oh , you do n't understand .
13 I must go -- I must !
14 My mother is over there . ''
15 Something in the girl 's last sentence or the tone in which it was
uttered brought a look of pain to Mrs. Cameron 's face .
16 But she made no further attempt to dissuade her .
17 `` Well , if you must .
18 But you can not go alone -- no , Nora , I can not allow it .
19 The wind is too high and it is too late for you to go over by yourself
.
20 Clark Bryant will take you . ''
21 XXXXX would have protested but she knew it would be in vain . Nora
Bryant|Cameron|Dinner|Mrs. |Nora|Something|harbour|morning|sentence|tonight
```

Report

- Experimental Result
  - Evaluation will be the accuracy on the test set. (Baseline: 50%)
  - You need to produce a file of your answers (one for each query with the original order).
- Report
  - Your method, including setting.
  - Your experiment result.
  - Your idea about future work and improvement.
- Grade
  - Your codes and report.

Submission

- Your submission file should be uploaded to
    - ftp://bcmi.sjtu.edu.cn:2122/FinalProject   (nlp_public/nlp_public)
    - or e-mail to ezonne0628@gmail.com
- Pack all your submission files, including solution report and source code, in one .rar (or .zip) file with the file name   "Name_StudentID_fp.rar" .
- **Deadline: June 11, 2017**
- Attention: Please carefully check the final version of your homework. **Once you submit it, you can not modify it anymore**.

- TA : Huang Yafang ezonne0628@gmail.com