

Recent Advances in Sequence-To-Sequence Learning: Project Guidelines

Tsz Kin Lam

February 25, 2020

Deadline: 11th June 2020

Requirements: Submit both 1) Code and its documentation and 2) a paper summarising the findings.

1 Experiments

Each project contains two components: 1) re-implementation of the selected paper and 2) the proposed variants. In the re-implementation part, one or two language pairs to confirm your implementation is sufficient. If the data size involved is too huge to be trained, sub-sampling can be a choice. In the proposed variants, you can just run on the same language pair and make your comparisons. In replying your proposals, I also provided some possible variants but you can ignore them and work on your ideas. Improvement is not the focus but your motivations.

2 Coding and Documentation

- Please use PyTorch or TensorFlow as your deep learning framework.
- Please comment your code when necessary, e.g., the size or shape of your tensors before and after complicated transformations.
- Upload your code to GitHub or GitLab
- Describe steps to generate experimental results reported on your paper. As an example,
 1. Download the data by `”bash download_data.sh”`
 2. Train the model by `”python train.py”`
 3. Get inference on test set by `”python generate.py”`
- Please report the rough **Training time** of your models
- Please provide the trained model(s).

3 Paper summary

- Please follow the ACL submission guidelines and styles (See 2. Paper Submissions and Templates): <https://acl2020.org/calls/papers/>
- Maintain the basic structure of 1) Abstract, 2) Introduction, 3) Methods, 4) Results and Analysis and 5) Conclusion.
- A maximum of 8 pages is applied to all people
- A minimum of 2 pages if you work individually but 4 if you work in pairs.
- If you work in pairs, both people receive the same grade for the project part.