Curriculum Learning for Domain Adaptation in Neural Machine Translation

Why can't I stand in the center of the stage?

Sigh... high-quality in-domain data is so limited

Try curriculum learning!

DATA ORDERED BY LEVELS OF DOMAIN SIMILARITY

Now let me start learning from the most relevant data and learn them more often

Data ordered by levels of domain similarity given by data selection methods!

Unlabeled-domain data is easy to get, but not all of them are useful.

I can order them using the domain similarity scores given by data selection methods!

Unlabeled-domain data