ML2021Spring HW5 Report

Department of Physics 林展慶 B0520244

Public Score	Private Score
30.92	29.07

The methods I used to pass the strong baselines include:

1. clean data
2. unigram subword tokenization
3. transformer
4. back translation on mono dataset
5. label smoothed
6. learning rate scheduling

I build a transformer with embed_dim=512, ffn_embed_dim=2048, layers=6, dropout=0.3 for both backward and forward transformers.