

ML2023Spring HW5 Bonus Report

資工所 林家禾

R11922048

Public Score	Private Score
36.92	36.56

The methods I used to pass the boss baselines include:

1. Train a back-translation transformer model, with parameters:
batchsize=2048*8, layer=8, dim_ff=1024, dim_embed=512,
head_attention=4, dropout=0.25, epoch=50, others same as sample code.
2. Use mono dataset to generate additional data, then create new dataset.
3. Use new dataset to train a transformer model, with parameters:
batchsize=4096*4, dim_ff=4096, dim_embed=1024, epoch=78,
others same as transformer base parameters.