

CLOVA AI RUSH 2022

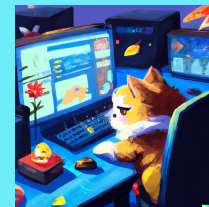
Long Document Summarization 3rd place

Pipelining for T5-Large using Huggingface

Sungho Park

Cipherome

<https://github.com/naem1023>



Solution

- Pipelining using Huggingface for
 - Data Parallel Training → Integration of Pytorch DP
 - Configuration of Training, Model, Data → HfArgumentParser
 - Tokenization, Artifacts saving, etc ...
- Training two model
 - Generator model for whole subject → Inference dialogue, book, journal, note
 - Generator model for journal subject → Inference journal
- Truncate out of length.
 - I tried to extract the back of the document or keywords and SOTA paper(BRIO), but all of them can't get good scores.
- Checkout hyperparameter tendency with small models(Finetuned SKT-KoBART-v2, 124M)
- Checkout scores with big models(paust/pko-t5-large, 800M)
- Tuning hyperparameters
 - max_len of generator for each subject
 - Selection for fp32, fp16, bf16 and quantization optimized level
 - etc ...