

Course: CSCI 1460 Computational Linguistics
Faculty-sponsor: Eugene Charniak

Title: Adapting Language Models for Abstractive Text Summarization

Abstract:

Pre-trained language models like BERT and GPT2 have been successfully fine tuned for several natural language understanding tasks, but there is still ongoing research on effectively applying these models for text summarization. Abstractive summarization, specifically, can benefit from language modeling capabilities through pre-training, because new sentences need to be generated for the summary. In this project, I implement abstractive models from 3 papers to investigate the techniques used to adapt these pretrained language models for summarization. The first model [1] fine tunes GPT2 with specialised embeddings and domain adaptive training. The second [2] uses BERT as a document level encoder, and a transformer decoder with separate optimizers. The third model [3] uses a two stage training approach, first generating a draft summary with a BERT encoder and transformer decoder, then refining the summary. The models are trained and tested on the CNN/DailyMail and Gigaword datasets.

References:

- [1] Andrew Hoang, Antoine Bosselut, Asli Celikyilmaz, Yejin Choi. 2019. Efficient Adaptation of Pretrained Transformers for Abstractive Summarization. In *arXiv preprint arXiv:1906.00138v1*.
- [2] Yang Liu, Mirella Lapata. 2019. Text Summarization with Pretrained Encoders. In *arXiv preprint arXiv:1908.08345v2*.
- [3] Haoyu Zhang, Jingjing Cai, Jianjun Xu, Ji Wang. 2019. Pretraining-Based Natural Language Generation for Text Summarization. In *Proceedings of the 23rd Conference on Computational Natural Language Learning*, pages 789–797 Hong Kong, China.