Team ID: 17

Team Name: OnlyText

Using BART to Summarize Scientific Documents: An Analysis of Graph-Based and Section Summary Methods

UNIVERSITÉ Concordia UNIVERSITY

Abstract

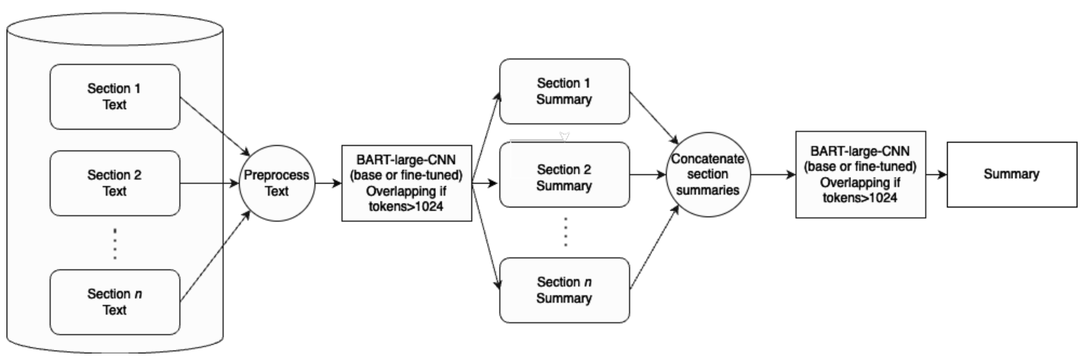
- This project proposes two approaches for condensing long documents: summarizing sections and selecting top sentences using LexRank. We evaluate summaries using various metrics.
- Results show that the ranked sentence method outperforms the section-based approach, with fine-tuned BART improving performance.

Goal of the Project

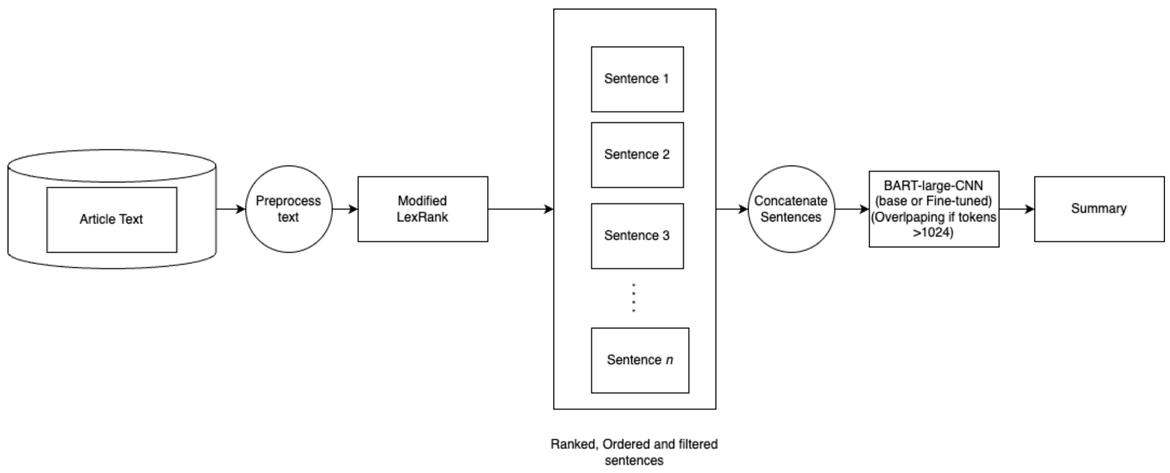
- Address the input size restriction on BART by condensing input text using section summaries and LexRank methods.
- Evaluating and analyzing these approaches using various metrics and their computational requirements.

Methodology

- *BART-large-CNN* [1] model and its fine-tuned version to summarize long documents by condensing input through two approaches.
- Summarizing section summaries to 50-100 tokens (or more for longer sections using overlapping)
- Ranking sentences (LexRank [2]) and selecting the top sentences containing 1024 tokens or less from the document as condensed input to BART.
- The summaries are evaluated against the Abstract of the document. The metrics used for evaluating the summaries are *ROUGE*, *BERTscore*, *BLEU*.
- The datasets used for this project are arXiv and PubMed document datasets [3].



Section Summarization Approach



Ranked Sentence Approach

```
{
  'article_id': str,
  'abstract_text': List[str],
  'article_text': List[str],
  'section_names': List[str],
  'sections': List[List[str]]
}
```

Dataset Format

Results

Approach	Dataset	ROUGE-1	BLEU	BERTscore
Ranked Summary	ArXiv	**38.10**	**12.09**	**83.34**
Section-wise Summary		37.33	11.67	83.25
Ranked Summary	PubMed	**37.24**	**12.34**	**84.11**
Section-wise Summary		34.01	11.16	83.18

Metric scores (ROUGE, BLEU, and BERTScore) for summaries generated by BART-large-CNN.

Approach	Dataset	ROUGE-1			BLEU			BERTScore		
		E3	E5	E10	E3	E5	E10	E3	E5	E10
Ranked-summary	ArXiv	**39.12**	**39.92**	**39.96**	**13.40**	**13.75**	**13.75**	84.60	**84.87**	**84.84**
Section-wise Summary		38.64	39.31	39.28	13.34	13.32	13.30	**84.61**	84.74	84.70
Ranked-summary	PubMed	**39.59**	**39.46**	**39.63**	***14.65***	**14.48**	***14.37***	**85.42**	**85.46**	**85.33**
Section-wise Summary		38.02	38.56	37.96	14.29	14.30	13.74	85.06	85.20	84.91

Metric scores for summaries generated by BART-large-CNN fine-tuned across different epochs.

Sample Summaries:

Hearing loss is considered one of the three most common conditions in the elderly. It occurs due to several degenerative and physiologic changes that affect the inner ear. The main goal of an auditory rehabilitation program in elderly patients is to minimize the effects caused by sensory deprivation of hearing. The study was carried out on a group of elderly people aged 75 and older. The results were published in the European Journal of Clinical Psychiatry. Results show evidence of the positive effects of hearingAid use , even if unilateral. The data show that all participants had some degree of social activity constraints before using a unilateral hearing aid. After 30 days of hearing aid use, 12 (93.3%) and 13 (100%) participants showed absence or reduction ofsocial activity constraints.

Summary generated from section summaries

Hearing loss is considered one of the three most common conditions in the elderly. The main goal of an auditory rehabilitation program in elderly patients is to minimize the effects caused by sensory deprivation of hearing. Results show that unilateral hearing aid contributed to the elimination or reduction of depression symptoms and of social activity constraints in elderly participants in this research sample group. After 30 days of hearing aid use, 12 (93.3%) and 13 (100%) participants showed absence or reduction in social activity constraint. Results are consistent with another study that examined the short - term benefits of amplification in new users. It is important to highlight that in the studied group , unilateral prosthesis resulted in substantial benefits not only regarding hearing but also in other matters essential for the well - being and quality of life of individuals.

Summary generated from ranked sentences

Conclusion

- The ranked sentence approach outperforms the section-wise summary method across all datasets, producing more coherent summaries without contextual loss and with greater efficiency.
- Fine-tuning the BART-large-CNN model has further enhanced its summarization performance.

Limitations and Future Work

- Computationally expensive for fine-tuning.
- The dataset has artifacts and placeholders affecting the informativeness of the summary.
- Fine-tuning using ranked sentences will be explored to analyze the performance of the ranked-summary approach.

References

[1] M. Lewis, Y. Liu, N. Goyal, *et al.*, "BART: denoising sequence-to-sequence pre-training for natural language generation, translation, and comprehension," *CoRR*, vol. abs/1910.13461, 2019. arXiv: 1910.13461. [Online]. Available: http://arxiv.org/abs/1910.13461.

[2] G. Erkan and D. R. Radev, "Lexrank: Graph-based lexical centrality as salience in text summarization," *Journal of artificial intelligence research*, vol. 22, pp. 457–479, 2004. [Online]. Available: https://arxiv.org/pdf/1109.2128.

[3] A. Cohan, F. Dernoncourt, D. S. Kim, *et al.*, "A discourse-aware attention model for abstractive summarization of long documents," *Proceedings of the 2018 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 2 (Short Papers)*, 2018. DOI: 10.18653/v1/n18-2097. [Online]. Available: http://dx.doi.org/10.18653/v1/n18-2097.