



Summaformers by iREL, Systems to Generate Summaries of Scholarly Documents

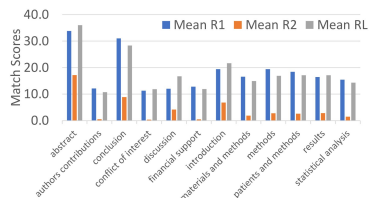
Conspectus

Automatic text summarization is a widely studied task in Natural Language Processing. It is cognitively challenging, in that, extracting summary worthy sentences is laborious, and expressing semantics in brief when doing abstractive summarization is time consuming. In our work, we specifically look at the problem of summarizing scientific research papers from multiple domains. We differentiate between two types of summaries, namely, (1) LaySumm: A very short summary that captures the essence of the research paper in layman terms restricting overly specific technical jargon, and (2) LongSumm: A much longer and detailed summary aimed at providing specific insights into various ideas touched upon in the paper. While leveraging latest Transformer based models, our systems are simple, intuitive and based on the manner in which specific paper sections contribute to human summaries of the two types described above. Evaluations against gold standard summaries using ROUGE metrics prove the effectiveness of our approach. On blind test corpora, our systems rank 1st and 3rd globally for the LongSumm and LaySumm tasks respectively.

Challenges

Scientific research papers are fairly structured documents containing standard sections like ‘abstract’, ‘introduction’, ‘background’, ‘related work’, ‘conclusion’, and ‘results’. Architectures meant for processing such scholarly documents should therefore be aware of such sectional structure. A simple way of achieving that would be to pick a few sentences from each of the sections to be a part of the summary. However these questions would have to be answered:

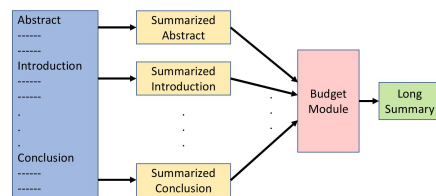
- How do we calculate the number of sentences to pick from each particular section?
- Which are the sentences to be picked from a chosen section?
- Can we rewrite sentences so as to obtain a concise and coherent abstractive summary?



— LaySumm Section Contribution

LongSumm

For our summary generation pipeline, we consider one section at a time with no global context, based on existing scientific evidence. We found that ‘introduction’ was the most important section closely followed by ‘related work’ and ‘results’. We utilized SummaRunner, a simple neural extractive text summarization architecture as our section-level summarizer. We pre-trained SummaRuNner on the PubMed dataset to generate paper abstracts as closely as possible from various paper sections before fine-tuning it on the Long Summarization data. Based on section contribution evaluations, we constructed a budget module to calculate how much weight to assign each section for the purpose of combining section summaries into the final long-summary.



LaySumm

From section-contribution studies, we saw that the ‘abstract’ was the most significant section followed by ‘conclusion’. High ROUGE-L overlap indicates some degree of verbatim copying from ‘abstract’ onto the lay-summary. With the ‘conclusion’ section, we observe a high ROUGE overlap plus a relatively shorter section length indicating that it contains a great deal of useful information in a condensed fashion.

We fine-tuned pre-trained Transformer-based encoder-decoder models for a given set of individual sections as the source. Our results indicate that using ‘abstract’ as the only sequence for conditional generation is a better choice as compared to utilizing more sections. Therefore, the problem at hand is one of capturing salient information as one would expect from a summarization task, with the additional flavor of text style transfer.

Supplementary Material

- [Associated Publication](#)
- [Slide Deck](#)
- [Repository](#)
- [Technical Poster](#)