Summary of "Automatic Slide Generation for Scientific Papers" [1]

Athar Sefid, Jian Wu, Prasenjit Mitra, C. Lee Giles

SciKnow 2019

Summary by Ioannis Sakiotis
CS 800, Spring 2020

The capability of effectively summarizing text has been heavily studied in the context of natural language processing. On the other hand, producing summaries in alternative formats such as *slide presentations*, has not been investigated with the same rigor. Thus, the authors attempt to bridge that gap and present their work on generating slides that effectively summarize academic papers. This capability can greatly facilitate the development of slide presentations by providing a "good" first draft. Another beneficial application could involve the generation of slides for academic papers that do not have corresponding slides.

According to the proposed method, each sentence is assigned a score highlighting its importance within the entire text. This is accomplished by training a neural network on a corpus of 1200 computer science papers and their corresponding slides; the ground truth consists of pairs of slides and unigram/bigram representations of an academic paper. The input to the neural network encompasses the semantic meaning of each sentence and whether it is suitable to summarize surrounding text, its position in the paper, as well as various other features.

Based on the assigned scores yielded by the trained neural network, the most suitable sentences are extracted in order to be included in the slides. Integer Linear Programming (ILP) can be used to filter sentences that don't match certain criteria such as maximum sentence length. The resulting slides, include first-level bullet points which consist of noun phrases, and second-level bullets that elaborate further and include sentences identified through ILP. Additional levels of bullet-points are not considered in this model as they are not typically encountered.

The performance of the model was evaluated based on the Recall-Oriented Understudy for Gisting Evaluation (ROUGE), a standard metric regarding the quality of generated summaries. In addition, the authors presented precision-oriented scores and the F-measure (harmonic mean of the Precision and Recall oriented scores). The authors also utilized other variants such as ROUGE-2, ROUGE-L, and ROUGE-W, with ROUGE-2 considered to be the most appropriate metric.

Several models were implemented with different types of neural networks (LSTM, CNN, GRU). Furthermore, the authors experimented by utilizing a greedy sentence selection method, according to which each sentence's score was compared against a specific threshold to determine whether it would be selected or not; this greedy method replaced ILP in some of those models. The set of models developed, was compared against AvgTFIDF, TextRank, and the MEAD platform based on their ROUGE scores. The results showed that the model that generates the scores via a convolutional neural network and then selects sentences based on ILP yielded the best results and outperformed the three baseline models in terms of F score. This indicates that the model can successfully be utilized for slide generation, as the high ROUGE scores correspond to significant similarities between the automatically generated slides and the manual ones that were present in the corpus.

The authors were the first to use deep neural networks to represent sentences and their suitability for describing surround text, as features for sentence ranking. Another contribution was the utilization of ILP for sentence selection. Future work will involve the expansion of the model to include figures and tables, which are invaluable elements in slides.

Questions/Comments

1. Would an encoder-decoder network be a suitable/promising candidate for the score generation module?

2. Would it be possible to incorporate information from a knowledge base, in order to improve/expand the information presented on the slides? This could be very beneficial if additional context is desirable. If the presentation is to be tailored for a different audience of greater or more limited expertise, this could be very useful. If I'm writing a paper on particle collision simulation, I'm sure there is enough material in a KB to provide more expansive figures or information that can ease the understanding of non-experts.

3. Coupled with information retrieval methods, could information be extracted from the references as well? This could be very powerful when working with survey papers and we want to get a better idea about a topic without having to read all the references.

References

[1] A. Sefid, J. Wu, P. Mitra, and C. L. Giles, "Automatic Slide Generation for Scientific Papers," in *the 3rd International Workshop on Capturing Scientific Knowledge, Los Angeles, California, USA, November 19th, 2019*, SciKnow 2019, pp. 11–16, 2019.