Paper Commentary Exercise

PoHao CHOU

November 7, 2014

1 A Bayesian Method to Incorporate Background Knowledge during Automatic Text Summarization – Rating: 4/5

Louis [1] proposed another aspect in summarization, which focuses on extracting new text deviating from previous knowledge on the topic. They proposed an algorithm that exploits Bayesian and KL-divergence to extract new and important sentences.

The paper proposed an interesting issue on summarization. Rather than extracting general importances for each document, extracting novel information is more helpful in some applications.

This inspired me that summarization can be variable, here the summarization is varied with knowledge corpus. Maybe there are others, like time or weather.

The setting of experiment and dataset is clear, but I wish they can provide the comparison between small and big background corpus for topic-based and surprise-based method.

I wish the paper can show the reason or comparison to use Dirichlet distribution to model hypothesis probability.

2 Summarizing Sporting Events Using Twitter – Rating: 3/5

Nichols *et al.* [**?**] proposed an method for implicitly crowdsourcing summaries of events. They used word-frequency based method to analyse twitter updates when there is an extreme changes in updates volume. According to the score above, the result of their algorithm is top N sentences that score highly and each include different information.

The summarization using twitter updates is interesting.

In twitter updates, the connections between them are weak. The difference of methods between analysing weak and strong connections is worth noting.

I think the performance of this method is highly dependent on the topics. For example, the events about soccer game may be extracted more easily. So it will be better if the evaluation

data covers more topics.

In phrase diagram, directly using the longest sentence is easy to be affected by noise. Using the sentence which is long and covers most keywords in other sentences may be better.

References

[1] A. Louis. A bayesian method to incorporate background knowledge during automatic text summarization. In *Proceedings of the 52nd Annual Meeting of the Association for Computational Linguistics*, volume 2 of *ACL14*, pages 333–338. Association for Computational Linguistics, June 2014.