

COMP550 Natural Language Processing

Assignment 3

Jonathan Guymont

November 21, 2018

1 Reading Assignment - Multi-document Summarization

1.1 Brief summary

Three main approach for sentence selection: (1) use a binary classifier that predict wether a sentence should be kept, (2) use a markov model (3) assign weights to sentence and pick the sentence with highest weight The method describe in this paper is based on the third approach. The term frequency is often use as a features to compute sentence importance. The author showed that there is a strong relation between the frequency of a content (e.g. a word, a sentence, etc.) and its likeliness of being in a summary. They created a summarizer that uses only the frequency of content to predict wether or not a content should be extracted. They discuss the impact of frequency in the summary in the input at the word level, and then look at frequency at a semantic level, using semantic content units. At the time, the frequency of top words in the original text that where in summary produce by state of the art machine summarizer was lower then the frequency of top words in the original text that were present in the human summary. Also, machine summarizer were using words that were not in the human summary. The authors found that the words that were used by all human summarizers in their summaries include the ones with high frequency in the original text and the words that appear in only one human summary tend to be the words with low frequency in the original text. The model the sentence that maximize the likelihood of the selected set of words under a Binomial distribution. They used an heuristic since an exhaustive search is intractable. There result shows that there algorithm produce better match when comparing the top words frequency in the machine summary vs the human summary. The performance of sumBasic on the multi documents summarization task was average compared to other algorithms.

1.2 Limitation of the approach

When human summarize, they generally do not limit there summary to selecting most important sentence. They combine sentence together to compress information or they paraphrase the original content to improve readability or to modify the content according to there style. If the goal is to replicate human summary, then summary produce by algorithm that select most relevant sentences according to some score will always have some level of difference with human summary.

1.3 ROUGE evaluation metric

Advantage. Summary that have high ROUGE-1 score are seen as good summary by humans (Lin and Hovy, 2003; Lin, 2004). It is a simple metric that can easily be used to compare the performance of different algorithm.

Disadvantage. The ROUGE evaluation metric tells how a summary is close to the human's. This assume that humans produce good summary. This is definitely arguable since humans sometimes use non optimal strategy like paraphrasing sentence for style reason (e.g. do not repeat the same word to often) or just because we are use to do it to show some level of comprehension. Most of the time, fancying sentence with style will not improve comprehension.

1.4 Three questions

- (1) Some sentence are very long and are not likely to be in a human summary. Is there any approach that include a sentence simplification module?
- (2) Could using semantic features like POS tags or word position improve the ROUGE score?
- (3) How RNN based extractive summarizer are performing?

2 Experiment

Below are reported all the summary for one of the cluster as an example to support the observations. Their is not much to say about the **leading** algorithm except that it is very dependent on the order in which it sees the content. The **original** and the **best-average** algorithm gives very similar result - though the last 2 sentence selected by the best-average algorithm seem to be less relevant, supporting the usefulness of step 3 in the sumBasic method. The summary produced by the simplified algorithm have a lot of redundancy - in both words and sentence meaning - showing that updating the probabilities of the words have the expected effect on the sentence selection.

Leading.

- About 8,000 light-years from Earth lies a star system unlike any astronomers have ever seen.
- **Gamma-ray bursts have been observed in other galaxies, but never in our own.**
- The objects responsible for this poorly understood phenomenon are just as interesting as the gamma-ray bursts themselves.
- In a paper published Monday in the journal Nature Astronomy, an international team of researchers reveals their findings on this new object, dubbed Apep.
- Apep had been seen in X-ray and radio observations more than 20 years ago, but had never been studied in-depth.
- In 2012, astronomer Joe Callingham, then working on his Ph.

Original.

- **The two stars in the binary system are known as Wolf-Rayet stars.**
- **Gamma-ray bursts have been observed in other galaxies, but never in our own.**
- **The researchers theorize that one star is rapidly rotating, producing fast winds at poles, but not at the equator where the dust would be.**
- **It will go supernova, probably in 100,000 years, Callingham said.**
- It's almost difficult to convey just how weird Apep truly is.
- But whether or not the conditions will remain, astronomers can't say for certain.
- **It could happen at any time in the. It left him scratching his head.**

Best-Average.

- **The two stars in the binary system are known as Wolf-Rayet stars.**
- **Gamma-ray bursts have been observed in other galaxies, but never in our own.**
- **The researchers theorize that one star is rapidly rotating, producing fast winds at poles, but not at the equator where the dust would be.**
- **It will go supernova, probably in 100,000 years, Callingham said.**
- **It could happen at any time in the. It left him scratching his head.**
- That's when it became clear we had found something really special, Tuthill said.
- It left him scratching his head.

Simplified.

- **The two stars in the binary system are known as Wolf-Rayet stars.**
- One of stars is an unusually massive sun known as a Wolf-Rayet star.
- Wolf-Rayet stars are massive, more than 20 times that of our sun.

- It's a rare type of nebula called a pinwheel nebula, found in multiple star systems in which at least one of the stars is a Wolf-Rayet.
- The more the researchers studied these stars, the stranger it became.
- But this rapidly rotating Wolf-Rayet star is turning that theory on its head, as our young galaxy holds stars with high metallicity.
- The dust barely moved.