Text Summary Augmentation for Intelligent Reading Assistant

PRAMOD VADIRAJA, University of Kaiserslautern, Germany
ANDREAS DENGEL, German Research Centre for Artificial Intelligence (DFKI), Germany
SHOYA ISHIMARU, German Research Centre for Artificial Intelligence (DFKI), Germany

This paper presents a technique to assist a reader. We aim to reduce manual efforts of the reader by leveraging the state-of-the-art document summarization techniques and providing summaries about unclear descriptions for each reader. Our system acts as a plug-and-play model that can be modified to support additional methodologies. As a backend of the system, we investigated several text summarization techniques and evaluated three techniques of them: TextRank, LexRank, and Luhn's algorithm.

Additional Key Words and Phrases: Cognitive augmentation, knowledge acquisition, human-document interaction, information retrieval, text summarization

ACM Reference Format:
Pramod Vadiraja, Andreas Dengel, and Shoya Ishimaru. 2021. Text Summary Augmentation for Intelligent Reading Assistant. In *Augmented Humans International Conference 2021 (AHs '21), February 22–24, 2021, Rovaniemi, Finland.* ACM, New York, NY, USA, 4 pages. https://doi.org/10.1145/3458709.3459002

1 INTRODUCTION

Attention can be described as a set of mechanisms that allows the allocation of cognitive resources, which are assumed to be limited [13]. There has been extensive scientific research that studies the attention span of humans [15]. It is found to be very short-lived. Hence it is important to protect the reader's attention, to ensure an immersive cognitive experience while reading a document.

Consider a scenario where a person is reading academic material, different portions of the content will have an impact on cognitive states at different levels. For instance, when a person is reading a chapter on photosynthesis, the reader is able to give his/her undivided attention if he/she is aware of concepts such as carbohydrates, proteins, carbon-dioxide, and chlorophyll. If the reader is unaware of certain aspects of the chapter, the attention is divided and, in some cases, distracted, as this leads to a dip in interest or confidence. This effect leads to a task-switching, which is an unconscious effort of the reader to switch attention from one task to another. Now, the reader tries to read about the unclear concepts. In such cases, due to the reader's limited attention span, it becomes difficult for the reader to maintain consistency while reading a material.

Our text summary augmentation system supports both of a reader and an editor. Although there are rational approaches of preparing a digital supporting material by an editor of the content [2, 5], it is not practical to delegate a lot of work to the experts. We propose a solution to bring about an automated way of supportive content creation and decrease the efforts of experts.

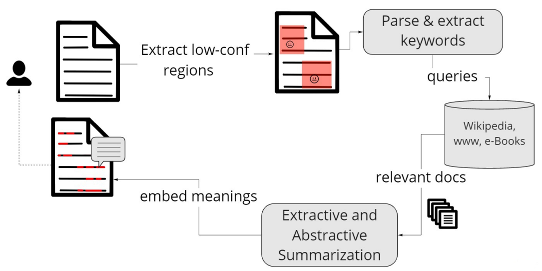

Fig. 1. An overview of the work-flow in a scenario when a reader's cognitive states fluctuate while reading a document.

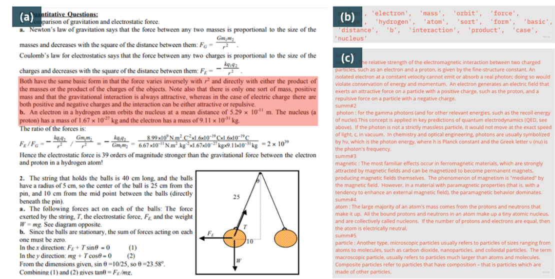

Fig. 2. An example showing (a) a low confidence region, (b) extracted keywords (c), and summaries for keywords.

The contributions of this paper are: (1) demonstrating the concept of our automatic text summary generation as a working system and (2) reporting the results of a literature survey and an evaluation of text summarizations techniques.

2 TECHNICAL BACKGROUND

Figure 1 shows an overview of the work-flow of our proposed pipeline. It consists of the following two stages.

Information Construction – The cognitive states of a reader are quantified by eye-tracking and physiological sensing. We utilize an algorithm proposed by Jacob *et al.* for this step [6]. We extract the regions in the document where the reader has relatively lower confidence using the gaze locations. Then we process the extracted textual content to remove stop-words, punctuations, etc., to filter out only the prominent keywords from these. The keyword set is utilized to query from a previously indexed data store. This can correspond to a database of digital copies of books that are indexed with word entities using powerful indexing software like Apache Lucene [10]. We can also get access to contents using Representational State Transfer (REST) APIs like Wikipedia API.

Summarization – With this kind of setup, if the reader is under-confident in some topics, he/she can just interact with the document, in particular with the words/topics of interest, which in-turn results in the summarised content with respect to the selected topic.

Ideally, the above steps should work on a real-time basis as the reader is involved in reading the document. But there might be problems with respect to the latency of the API requests made.

3 THEORETICAL BACKGROUND

Our system utilizes a plug-and-play model that allows a developer and a user to select text summarization techniques. There are several categories of summarization techniques [1]. We introduce them with categorizing them into two types: extractive and abstractive.

Extractive summarization involves the task of extracting the crucial or important sentences from a huge text, based on the similarity of the sentences with each other. There are several techniques to perform extractive summarization: Neural Networks that classifies a sentence based on its importance, Text Rank based on the page-rank algorithm [11], LexRank [4], and so on.

Abstractive summarization is the task of intelligently generating novel sentences by going through blocks of textual content. Neural language models that have been pre-trained with huge text corpuses like the Transformer [16] based BERT [3], and T5-Transformer [12] are some of the recent developments in the area of Natural Language

Processing. These models can be trained for downstream tasks like text summarization. Although an abstractive summarization is currently an active area of research, there have been several works depicting promising results on human language generations [8]. One advantage of abstractive summarization over the extractive approach is that the abstractive approach can generate human-like sentences as a part of the summary. In contrast, the extractive approach generates few of the main sentences.

4 DEMONSTRATION

We consider a scenario where a learner is reading a textbook in Physics[1]. As the first step, our system discover the regions of lower confidence through analyzing the cognitive states of the reader as shown in Figure 2 (a). We then preprocess the extracted textual content by removing stop-words, punctuations, performing lemmatization, and filtering out the nouns chunks from the text to improve our chances of extracting keywords with associated concepts as seen in Figure 2 (b). Even after filtering, there are still some keywords that were not filtered out. We use Wikipedia's API to extract text with respect to the queries from the previous step. The information from the above step is passed to the TextRank algorithm with minor changes in an attempt to improve results, which can be seen in Figure 2 (c).

5 EVALUATION

Since we expect our system to function in near real-time, we utilized the score of Recall-Oriented Understudy for Gisting Evaluation (ROUGE-1) [7] to evaluate some of the summarization techniques. Essentially ROUGE-1 tells us the number of token overlappings between the generated summary and a reference summary. For this experiment, we compared the three well-known summarization techniques: TextRank, LexRank, and Luhn's algorithm [9].

Summaries are a concise representation that captures the core essence/details of the main content. Hence it is important to be able to determine the optimal length for summaries. The length could be the number of tokens or the number of sentences in summary. Therefore we considered the length as the number of sentences in summary and generated summaries for a range of lengths. It is expected that as the summary length increases, the chances of the summaries capturing the core details increases as well.

Experimental Design – We used Wikipedia Summary dataset [14] that consisted of a mapping between topics and their respective summaries. We fetched a set of Physics-related terms by crawling the physics glossary page of Wikipedia. We considered 50 such terms and used the Wikipedia API to fetch the respective contents. These pages were preprocessed by performing tokenization and punctuations removal. The extracted pages were summarized by the three techniques, and their ROUGE-1 scores were measured. We considered the following two measures: (1) overall average ROUGE-1 scores and (2) average speed-up, i.e., the number of queries per second for different summary lengths.

Results – Figure 3 shows ROUGE-1 scores of each approaches. The number of queries per second for different summary lengths were TextRank: 0.31, LexRank: 0.14, and Luhn: 11.29. We found that although the Luhn's and TextRank algorithms had similar ROUGE-1 scores, the number of queries per second or keywords were generated was relatively higher in the case of Luhn's algorithm. The limitation of this study is that human intervention is often required since summary evaluation is considered subjective.

[1]An online demo: https://shoya.io/projects/hypermind#ahs2021demo

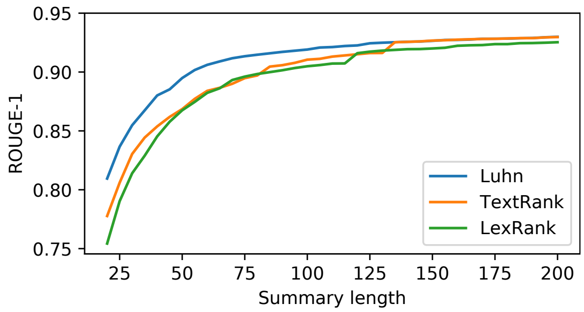

Fig. 3. Text summarization performances

6 CONCLUSION

This paper presented a concept that can potentially improve the cognitive capability of a person reading a digital document. We proposed an intelligent reading assistant system using a text summarization technique. In the future, we aim to improve our system by incorporating a stronger text preprocessing pipeline to eliminate errors. We intend to explore several other summarization techniques, including abstractive ones.

REFERENCES

[1] Mehdi Allahyari, Seyedamin Pouriyeh, Mehdi Assefi, Saeid Safaei, Elizabeth D Trippe, Juan B Gutierrez, and Krys Kochut. 2017. Text summarization techniques: a brief survey. *arXiv preprint arXiv:1707.02268* (2017).

[2] Ralf Biedert, Georg Buscher, Sven Schwarz, Jörn Hees, and Andreas Dengel. 2010. Text 2.0. In *CHI'10 Extended Abstracts on Human Factors in Computing Systems*. 4003–4008.

[3] Jacob Devlin, Ming-Wei Chang, Kenton Lee, and Kristina Toutanova. 2018. Bert: Pre-training of deep bidirectional transformers for language understanding. *arXiv preprint arXiv:1810.04805* (2018).

[4] Günes Erkan and Dragomir R Radev. 2004. Lexrank: Graph-based lexical centrality as salience in text summarization. *Journal of artificial intelligence research* 22 (2004), 457–479.

[5] Shoya Ishimaru, Ko Watanabe, Nicolas Großmann, Carina Heisel, Pascal Klein, Yutaka Arakawa, Jochen Kuhn, and Andreas Dengel. 2018. Demonstration of HyperMind Builder: Pervasive User Interface to Create Intelligent Interactive Documents.

[6] Soumy Jacob, Shoya Ishimaru, Syed Saqib Bukhari, and Andreas Dengel. 2018. Gaze-based interest detection on newspaper articles. In *Proceedings of the 7th Workshop on Pervasive Eye Tracking and Mobile Eye-Based Interaction*. 1–7.

[7] Chin-Yew Lin. 2004. Rouge: A package for automatic evaluation of summaries. In *Text summarization branches out*. 74–81.

[8] Hui Lin and Vincent Ng. 2019. Abstractive summarization: A survey of the state of the art. In *Proceedings of the AAAI Conference on Artificial Intelligence*, Vol. 33. 9815–9822.

[9] Hans Peter Luhn. 1957. A statistical approach to mechanized encoding and searching of literary information. *IBM Journal of research and development* 1, 4 (1957), 309–317.

[10] Michael McCandless, Erik Hatcher, Otis Gospodnetić, and O Gospodnetić. 2010. *Lucene in action*. Vol. 2. Manning Greenwich.

[11] Rada Mihalcea and Paul Tarau. 2004. Textrank: Bringing order into text. In *Proceedings of the 2004 conference on empirical methods in natural language processing*. 404–411.

[12] Colin Raffel, Noam Shazeer, Adam Roberts, Katherine Lee, Sharan Narang, Michael Matena, Yanqi Zhou, Wei Li, and Peter J Liu. 2019. Exploring the limits of transfer learning with a unified text-to-text transformer. *arXiv preprint arXiv:1910.10683* (2019).

[13] Claudia Roda. 2011. *Human attention and its implications for human–computer interaction*. 11–62.

[14] Thijs Scheepers. 2017. *Improving the Compositionality of Word Embeddings*. Master's thesis. Universiteit van Amsterdam, Science Park 904, Amsterdam, Netherlands.

[15] Kalpathy Subramanian. 2017. Product Promotion in an Era of Shrinking Attention Span. *International Journal of Engineering and Management Research* Volume 7 (03 2017), Pages 85 –91.

[16] Ashish Vaswani, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N Gomez, Łukasz Kaiser, and Illia Polosukhin. 2017. Attention is all you need. In *Advances in neural information processing systems*. 5998–6008.