**Extractive text summarization of online scientific articles for digital library repository**

Author 1a and Author 2b

aInstitute Address 1

bInstitute Address 2

*corresponding email address*

**Abstract:**

Text Summarization plays a significant role in natural language processing, especially in scientific communities like researchers, students, and so on.

**Keywords:** Keyword 1, keyword 2.

**1. Introduction**

In this study, we investigated two major things which necessary to create an online platform that returns relevant scientific articles with a summary, and evaluated which is a better algorithm based on the content size.

**2. Literature Review**

In the sense of abstractive text summarization, Wei Li et al., [8] proposed an approach to extend the basic neural encoding-decoding framework with an information selection layer. Liwei Hou et al. introduced an approach to Chinese words using Neural Model with Joint Attention [9] to address the problem of the attention encoder-decoder models that has shortcomings to generate repeated words or phrases.

**3. Methodology**

**3.1. Dataset**

We have used two datasets, where one was generated by the SumPubMed dataset [3] and the second one was a manually generated dataset.

Figure 1. Rouge values of Text Rank algorithm

Table 1. ROUGE values for the Text Rank algorithm in Experiment 01

|  |  |  |  |
| --- | --- | --- | --- |
|  | ROUGE-1 | ROUGE-2 | ROUGE-L |
| RECALL | 0.600848 | 0.271301 | 0.537622 |
| PRECISION | 0.223274 | 0.076683 | 0.199196 |
| F-MEASURE | 0.319437 | 0.116757 | 0.285262 |

**5. Conclusion**

In this study, we tried two things, which are finding a better algorithm for generating summaries and checking whether the increment of content size can give better results.

**References**

1. Allahyari, M., Pouriyeh, S., Assefi, M., Safaei, S., Trippe, E. D., Gutierrez, J. B., & Kochut, K. (2017). Text summarization techniques: a brief survey. arXiv preprint arXiv:1707.02268.
2. Meena, S. M., Ramkumar, M. P., Asmitha, R. E., & G SR, E. S. (2020, September). Text summarization using text frequency ranking sentence prediction. In 2020 4th International Conference on Computer, Communication and Signal Processing (ICCCSP), 1-5, IEEE.
3. Gupta, V., Bharti, P., Nokhiz, P., & Karnick, H. (2021, August). SUMPUBMED: Summarization Dataset of PubMed Scientific Articles. In Proceedings of the 59th Annual Meeting of the Association for Computational Linguistics and the 11th International Joint Conference on Natural Language Processing: Student Research Workshop, 292-303.
4. Lin, C. Y. (2004, July). Rouge: A package for automatic evaluation of summaries. In Text summarization branches out, 74-81.
5. Rahimi, S. R., Mozhdehi, A. T., & Abdolahi, M. (2017, December). An overview on extractive text summarization. In 2017 IEEE 4th International Conference on Knowledge-Based Engineering and Innovation (KBEI), 0054-0062, IEEE.
6. Jo, T. (2017, January). K nearest neighbor for text summarization using feature similarity. In 2017 International Conference on Communication, Control, Computing and Electronics Engineering (ICCCCEE), 1-5, IEEE.
7. Mihalcea, R. (2004, July). Graph-based ranking algorithms for sentence extraction, applied to text summarization. In Proceedings of the ACL interactive poster and demonstration sessions, 170-173.
8. Li, W., Xiao, X., Lyu, Y., & Wang, Y. (2018). Improving neural abstractive document summarization with explicit information selection modeling. In Proceedings of the 2018 conference on empirical methods in natural language processing, 1787-1796.
9. Hou, L., Hu, P., & Bei, C. (2017, November). Abstractive document summarization via neural model with joint attention. In National CCF Conference on Natural Language Processing and Chinese Computing (pp. 329-338). Springer, Cham.
10. Li, W., Xiao, X., Lyu, Y., & Wang, Y. (2018). Improving neural abstractive document summarization with structural regularization. In Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing, 4078-4087.
11. Liu, Y., & Lapata, M. (2019). Text summarization with pretrained encoders. arXiv preprint arXiv:1908.08345.
12. You, J., Hu, C., Kamigaito, H., Takamura, H., & Okumura, M. (2021, September). Abstractive Document Summarization with Word Embedding Reconstruction. In Proceedings of the International Conference on Recent Advances in Natural Language Processing (RANLP 2021), 1586-1596.