DOI: <https://doi.org/10.48009/1_iis_2021_xx-xx> (Note: The DOI will be completed by the editor)

**Text summarization using the relationship structure of a document**

**George Ray,** *Shepherd University,* [*gray@shepherd.edu*](mailto:gray@shepherd.edu)

**Abstract**

|  |
| --- |
| This article introduces a new approach for textual analytics that creates a relationship structure for a document by using the relationship words in the system of basic English. Relationship words interlock concepts in sentences. The author of a document proposes certain relationships between concepts and uses relationship words in sentences to present a pattern of those relationships. In this paper, the concept of a document’s relationship structure is used for automating summaries for documents. In addition to developing an automated approach to analyze, understand and present knowledge in documents, this article also considers how to validate the correctness of an artificial intelligence algorithm. |

**Keywords**: artificial intelligence, text summarizer, Rouge scoring, validating AI, thinking machines

**Addressing Reviewer Comments**

I thank the reviewers for their insightful comments. With respect to the comments, I have put in significant effort and made the following changes to the paper.

* Rewrote paragraphs in the methodology section describing the use of Rouge in validating the relationship word summarizer through comparison with a published summarizer. Emphasized that Rouge is used to test the null hypothesis that there is no statistical difference in the Rouge performance between the two summarizer approaches.
* Changed the results chart in Figure 1 to a visual that better compares the Rouge performances.
* In Table 1, repositioned *recall precision and fmeasure* so it is visually apparent each of the three words applies to both relWord and lexRankr columns
* Capitalized each instance of Rouge
* Added paragraphs comparing extractive and abstractive summarizers, including abstractive summarizers that use LLM for understanding and paraphrasing a document. Included a discussion on the performance of extractive and abstractive summarizers.
* Reworded *this* and *they* at beginning of paragraphs. Also, removed repetitions.
* Added a paragraph on impact of the paper on the field of artificial intelligence and rearranged other paragraphs to form a contribution to the field section
* Combined and reworded some short paragraphs to improve readability.

The reviewer comments follow.

\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*  
  
AUTHOR COMMENTS: The article is very interesting and the literature  
review clearly defines what the article is about and the current state  
of the topic. I would recommend revising how the methodology is  
presented, because is not very clear how each of the tools listed in  
the results chart was set or used to compare results. In other words,  
it would be good to organize the information in a way where we can see  
the similarities in the tools so we can compare differences in the  
results.  
  
  
  
\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*  
  
AUTHOR COMMENTS: This paper presents an approach for text  
summarization and compares it to lexRankr. It is well established  
within historical literature. Comments on performance and relation to  
recent LLM developments could have been discussed. The proposed model  
does not seem to be an improvement on the numerous existing text  
summarizers. The contributions to the field, readers, and attendees  
should be made more explicit and impactful.  
  
The text requires further revision. The style is better suited for a  
verbal discussion and consists of very short disorganized paragraphs.  
These should be combined to promote readablity. Repetition should be  
removed (KS is defined twice, consequative paragraphs state that A  
three-sentence summary is selected.) Paragraphs start with sentences  
using undefinded references to "they" and "this". The capitalizaiton  
of Rogue should be standardized.