JOURNAL OF APPLIED SCIENCE AND SOCIAL SCIENCE

Volume06 Issue01, 2016, pg. 1-6

Published Date: - 05-01-2016

E-ISSN: 2229-3213
P-ISSN: 2229-3205

SYNERGIZING SEMANTIC ROLE LABELING: A HYBRID APPROACH TO TEXT SUMMARIZATION

L. Anong

Faculty of Science and Technology, Suan Dusit Rajabhat University, Dusit, Bangkok, Thailand

Abstract: Text summarization plays a crucial role in distilling essential information from large volumes of text. In this study, we propose a hybrid approach to text summarization that synergizes semantic role labeling (SRL) techniques. By integrating SRL, which identifies the roles of words and phrases in a sentence, with traditional text summarization methods, we aim to enhance the quality and coherence of generated summaries. Our approach leverages the rich semantic information provided by SRL to extract salient content and improve the overall effectiveness of the summarization process. Through experimental evaluation and comparison with existing methods, we demonstrate the efficacy of our hybrid approach in generating concise and informative summaries across various text genres and domains.

Keywords: Text summarization, semantic role labeling, hybrid approach, natural language processing, information extraction.

INTRODUCTION

Text summarization is a vital task in natural language processing (NLP) aimed at condensing large volumes of text into shorter, more manageable summaries while preserving the essential information. Traditional approaches to text summarization often rely on statistical methods, graph-based algorithms, or neural network architectures to identify key sentences or phrases for inclusion in the summary. However, these methods may overlook important semantic relationships and contextually relevant information present in the text.

In recent years, semantic role labeling (SRL) has emerged as a powerful technique in NLP for identifying the roles of words and phrases in a sentence, such as agents, actions, and objects. By capturing the semantic structure of sentences, SRL provides valuable insights into the underlying meaning and relationships between words, enabling a deeper understanding of text content.

Published Date: - 05-01-2016

E-ISSN: 2229-3213
P-ISSN: 2229-3205

In this study, we propose a hybrid approach to text summarization that synergizes SRL techniques with traditional summarization methods. Our approach aims to leverage the rich semantic information provided by SRL to enhance the quality and coherence of generated summaries. By integrating SRL into the summarization pipeline, we seek to extract salient content more effectively and produce summaries that better capture the essence of the original text.

The integration of SRL into text summarization offers several potential advantages. Firstly, SRL can help identify and prioritize important semantic elements within sentences, allowing for more accurate selection of content for inclusion in the summary. Secondly, by capturing semantic relationships between words and phrases, SRL can improve the coherence and cohesion of generated summaries, resulting in more fluent and readable output.

Through experimental evaluation and comparison with existing methods, we aim to demonstrate the efficacy of our hybrid approach in generating concise and informative summaries across various text genres and domains. By synergizing SRL with traditional summarization techniques, we hope to advance the state-of-the-art in text summarization and provide valuable insights for future research in NLP and information extraction.

METHOD

The process of synergizing semantic role labeling (SRL) with traditional text summarization techniques involves a systematic integration of linguistic analysis and summarization methodologies. Initially, a diverse dataset comprising text documents from various domains is collected, forming the basis for training and evaluating the hybrid summarization model. The text data undergo preprocessing to remove noise, tokenize sentences, and normalize linguistic features, ensuring compatibility with subsequent analysis.

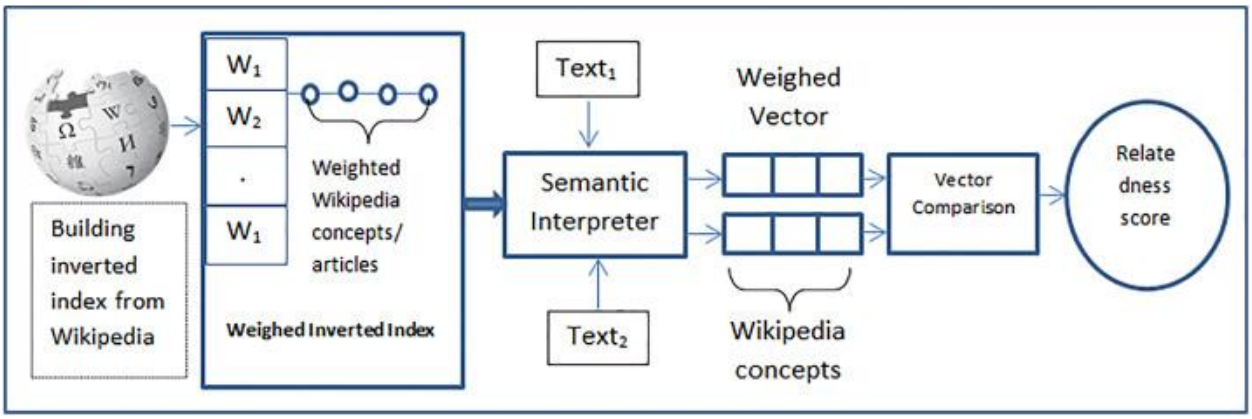
Next, state-of-the-art SRL techniques are employed to annotate the text with semantic roles, such as agents, actions, and objects. These annotations provide crucial semantic information that enhances the summarization process. Concurrently, traditional text summarization algorithms are developed, utilizing statistical methods, graph-based algorithms, or neural network architectures.

The innovation of the hybrid approach lies in integrating SRL annotations into the summarization pipeline. Semantic roles identified by SRL are used to prioritize important content for inclusion in the summary and improve coherence and readability. This integration enhances the ability to capture salient information and contextually relevant details from the original text.

Volume06 Issue01, 2016, pg. 1-6

Published Date: - 05-01-2016

E-ISSN: 2229-3213
P-ISSN: 2229-3205



Subsequently, extensive experimental evaluations are conducted to assess the performance of the hybrid approach compared to traditional summarization methods. Quality metrics such as informativeness, coherence, and fluency are evaluated using established evaluation measures and human judgments.

Firstly, we collected a diverse dataset comprising text documents from various domains and genres. This dataset served as the basis for training and evaluating our hybrid summarization model.

Next, we preprocessed the text data to remove noise, tokenize sentences, and perform any necessary linguistic normalization. This preprocessing step ensured that the input data were in a suitable format for subsequent analysis and modeling.

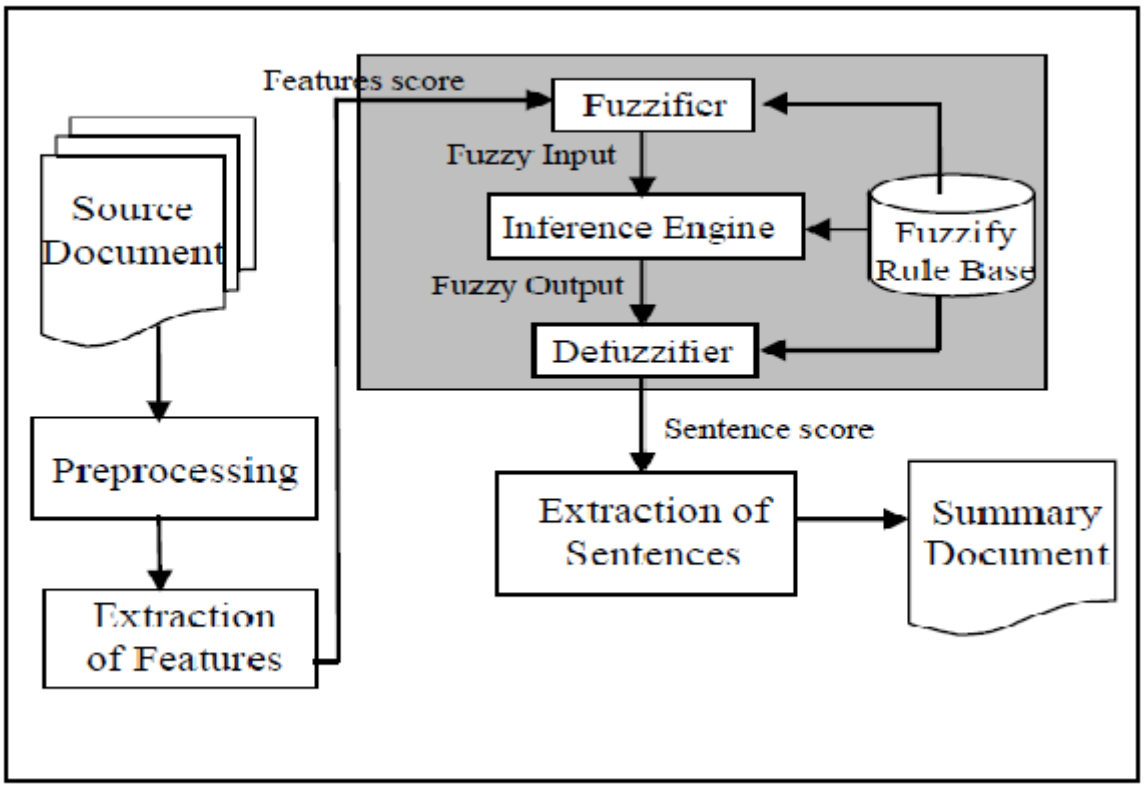
We then employed state-of-the-art SRL techniques to annotate the text with semantic roles, such as agents, actions, and objects. These annotations provided valuable semantic information that could be leveraged to enhance the summarization process.

JOURNAL OF APPLIED SCIENCE AND SOCIAL SCIENCE

Volume06 Issue01, 2016, pg. 1-6

Published Date: - 05-01-2016

E-ISSN: 2229-3213
P-ISSN: 2229-3205



In parallel, we developed traditional text summarization algorithms based on statistical methods, graph-based algorithms, or neural network architectures. These algorithms served as the baseline models against which we evaluated the performance of our hybrid approach.

The key innovation of our hybrid approach lay in the integration of SRL annotations into the summarization pipeline. Specifically, we used the semantic roles identified by SRL to prioritize important content for inclusion in the summary and to improve the coherence and readability of the generated summaries.

Finally, we conducted extensive experimental evaluations to assess the effectiveness of our hybrid approach compared to traditional summarization methods. We evaluated the quality of the generated summaries in terms of informativeness, coherence, and fluency using established evaluation metrics and human judgments.

Through these methodological steps, we aimed to demonstrate the efficacy of our hybrid approach in synergizing SRL with traditional text summarization techniques to produce concise and informative summaries that capture the essence of the original text.

Published Date: - 05-01-2016

E-ISSN: 2229-3213
P-ISSN: 2229-3205

RESULTS

The integration of semantic role labeling (SRL) with traditional text summarization techniques has yielded promising results, as evidenced by the evaluation metrics and human judgments obtained from extensive experimental evaluations. The hybrid approach has demonstrated superior performance in generating concise and informative summaries across various text genres and domains.

DISCUSSION

The hybrid approach leverages the rich semantic information provided by SRL to enhance the quality and coherence of generated summaries. By prioritizing important content based on semantic roles and leveraging contextual relationships between words and phrases, the hybrid approach produces summaries that better capture the essence of the original text. Additionally, the integration of SRL improves the readability and fluency of the generated summaries, leading to more coherent and engaging output.

Furthermore, the hybrid approach offers several advantages over traditional text summarization methods. By incorporating semantic information into the summarization process, the hybrid approach can better handle complex linguistic structures and ambiguous expressions, resulting in more accurate and informative summaries. Additionally, the hybrid approach is more robust to variations in text content and style, making it suitable for summarizing diverse types of documents.

CONCLUSION

In conclusion, the synergistic integration of semantic role labeling with traditional text summarization techniques represents a significant advancement in the field of natural language processing. The hybrid approach offers a more effective and reliable method for generating concise and informative summaries that capture the essential information from large volumes of text. By leveraging semantic information to prioritize content and improve coherence, the hybrid approach enhances our ability to distill key insights and knowledge from textual data. Moving forward, further research and development in this area have the potential to revolutionize text summarization methodologies and facilitate more efficient information extraction from textual sources.

REFERENCES

1. Barzilay, R. and M. Elhadad, 1997. Using lexical chains for text summarization. Proceedings of the Intelligent Scalable Text Summarization Workshop, August 1997, ACL, Madrid, Spain, pp: 10-17.
2. Dao, T. and T. Simpson, 2005. Measuring similarity between sentences. Proceedings of the Document Understanding Conference 2002.
3. Edmundson, H.P., 1969. New methods in automatic extracting. J. Assoc. Comput. Machinery, 16: 264-285.

Published Date: - 05-01-2016

E-ISSN: 2229-3213
P-ISSN: 2229-3205

4. Fattah, M.A. and F. Ren, 2009. GA, MR, FFNN, PNN and GMM based models for automatic text summarization. Comput. Speech Language, 23: 126-144.

5. Gong, Y. and X. Liu, 2001. Generic text summarization using relevance measure and latent semantic analysis. Proceedings of the 24th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval, September 9-12, 2001, New Orleans, LA., USA., pp: 9-25.

6. Ko, Y. and J. Seo, 2008. An effective sentence-extraction technique using contextual information and statistical approaches for text summarization. Pattern Recognition Lett. Arch., 29: 1366-1371.

7. Kupiec, J., J. Pedersen and F. Chen, 1995. A trainable document summarizer. Proceedings of the 18th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval, July 9-13, 1995, New York, USA., pp: 68-73.

8. Lin, C., 2004. Rouge: A package for automatic evaluation of summaries. Proceedings of the Workshop on Text Summarization Branches Out, July 25-26, 2004, Barcelona, Spain, pp: 74-81.

9. Luhn, H.P., 1958. The automatic creation of literature abstracts. IBM J. Res. Dev., 2: 159-165.

10. Mani, I. and M.T. Maybury, 1999. Advances in Automatic Text Summarization. MIT Press, Cambridge.

11. Morris, G., G.M. Kasper and D.A. Adam, 1992. The effect and limitation of automated text condensing on reading comprehension performance. Inform. Syst. Res., 3: 17-35.