Abstract

This paper proposes an approach of personalizing the vector space model with dependency parse relations and applying Latent Semantic Analysis on it to generate update summary from multiple documents. The purpose of the update summary is to inform the reader of new information about the topic. The main task was to produce two concise summaries from two related sets of documents, where the second summary was an update summary of the first one. In the proposed system individual word weight is calculated using tsf-isf and dependency parse of the document has been used to modify the tsf-isf weight of words by incorporating the dependency between each pair of words. To preserve important semantic information in the text LSA is performed and to select relevant sentences basic features, advanced features and update specific features are used. The experiment result shows that low overlap between initial summary and its update summary.

References

- Bysani P., Bharat V., Varma V., Modeling Novelty and Feature combination using Support Vector Regression for Update Summarization, In Proceedings of the 7th International
Conference on Natural Language Processing, India, 2007.


Index Terms

Computer Science
Artificial Intelligence

Keywords

Preprocessing Pos Tagging Similarity Matrix Dependency Parsing Semantic Similarity Matrix
Feature
Specific Sentence Ranking Strategy
Initial Summary
Update Summary