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.
  - Josef Steinberger, Karel Jezek, Update Summarization Based on Novel Topic

  - Eric Wehrli , Pierre-Etienne Genest, Guy Lapalme Luka Nerima, A Symbolic Summarizer
for the Update Task of TAC 2008, In Proceedings of the First Text Analysis Conference,
Gaithersburg, Maryland, USA, 2008.
  - Praveen Bysani, Vijay Bharat, Vasudeva Varma, Modeling Novelty and Feature
Combination using Support Vector Regression for Update summarization, In Proceedings of 7th
International Conference on Natural Language Processing, 2009.
  - Ravindranath Chowdary C. , Sreenivasa Kumar P. , Update Summarizer using MMR

gov/javanumerics/jama/
  - Kogilavani A. , Balasubramanie P. , Clustering and Feature Specific Sentence Extraction
Based Summarization of Multiple Documents, International Journal of Computer Science and
  - Pierre-Etienne Genest, Guy Lapalme, Luka Nerima, Eric Wehrli, A Symbolic
  - Yihong Gong, Xin Liu, Generic Text Summarization Using Relevance Measure and
Latent Semantic Analysis, SIGIR 2001. USA.
  - Lin C. Y. , ROUGE: A package for automatic evaluation of summaries. In proceedings of
the workshop on Text Summarization, Barcelona. ACL, 2004.

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