Abstract

Keywords play a vital role in extracting the correct information as per user requirements. Keywords are like index terms that contain the most important information about the content of the document. Keyword Extraction is the task of identifying a keyword or keyphrase from a document that can help users easily to understand the documents. Meeting transcripts is significantly different from document or other speech domains. This paper aims to extract keywords and keyphrases from meeting transcripts and also to add some additional features for improving the keyword and keyphrase extraction method. Here, this method is performed by both human transcripts and ASR transcripts and the keywords are extracted through MaxEnt and SVM classifier and Extraction of bigram and trigram keywords retrieval using N-gram based approach efficiently and also to identify the low frequency keywords using LDA (Latent Dirichlet Approach). Finally, the quality of the Extracted keywords is improved using pattern features through sequential pattern mining.

References

- Jasmeen Kaur and Vishal Gupta. 2010. Effective Approaches For Extraction Of

- Eibe Frank and Gordon W. Paynter. Domain specific Keyphrase extraction.
- Jiajia Feng. 2011. Keyword Extraction Based on Sequential Pattern Mining, ICIMCS August 5-7, china, ACM.

Index Terms

Computer Science Computations

Keywords

Keywords Keyphrase Meeting Transcripts LDA SVM MaxEnt