Abstract

In the past couple of years multi-topic summarization is a research investigation that has expanded much attention. There has been a variety of effort on generating natural language summaries for variety of topics, but this is feasible only for a very small number of topics. In this research paper the method trying to provide automatic detection of topics to be summarized that is can determine how many topics should be chosen for automatic summarization. This is effectively done through the combined efficient framework of ontology based and non ontology based systems will be optimistic for the excellence of topic summary. To achieve this propose to apply latent Dirichlet allocation (LDA) model for capturing the semantic information on topic transcription. LDA is a generative model and defines a probabilistic method for generating a new document. The LDA model is utilized for estimating topic sharing in queries and word recorded topic documents, and the identical is performed at the topic level. Concept based topic matching between query words and topic documents are performed using ontology based and non ontology based matching algorithm. The results of topic level matching methods are evaluated between the automatic topic detection method and predefined topic detection mechanism along with the experimental results shown to be complementary.
Using Latent Dirichlet Allocation to Incorporate Domain Knowledge with Concept based Approach for Automatic Topic Detection

References

- James Allan, Victor Lavrenko, Hubert Jin, First story detection in TDT is hard, In CIKM'00.
- T. Brants, F. Chen and A. Farahat, "A system for New Event Detection", In SIGIR'03.
- Qi He, Kuiyu Chang and Ee-Peng Lim, "A Model for Anticipatory Event Detection", In ER’06.
- G. Kumaran and J. Allan, "Text classification and named entities for new event detection", In SIGIR’04.
- Nicola Stokes and Joe Carthy, "Combining semantic and syntactic document classifiers to improve first story detection", In SIGIR’01.

Index Terms

Computer Science

Algorithms

Keywords

Latent Dirichlet allocation  Ontology based matching algorithm  Non ontology based matching algorithm

Topic summarization

and Automatic topic detection