Abstract

Whenever the text contains multiple ways of saying “the same thing,” but the application requires the same treatment of those various alternatives, an automated paraphrase recognition mechanism would be useful. One reason why paraphrase recognition systems have been difficult to build is because paraphrases are hard to define. Although the strict interpretation of the term “paraphrase” is quite narrow because it requires exactly identical meaning, in linguistics literature paraphrases are most often characterized by an approximate equivalence of meaning across sentences or phrases. This paper presents a survey of paraphrase detection techniques for Indian regional languages.

References

Penang Resort.


16. Cordeiro, Dias, Brazdil A Metric for Paraphrase Detection Proceedings of the International Multi-Conference on Computing in the Global Information Technology (ICCGI'07) IEEE.

17. Kozareva and Montoyo, Paraphrase identification on the basis of supervised machine learning techniques, Advances in Natural Language Processing: 5th International Conference on NLP (FinTAL 2006), Turku, Finland, 524-533.


Index Terms

Computer Science Pattern Recognition

Keywords

Paraphrase detection, Textual Similarity metrics, String similarity, Semantic relatedness, Statistical and semantic analysis, Bi-CNN-MI.