Abstract

Paraphrase detection (PD) is a very essential and important task in Natural language processing. The goal of paraphrase detection is to check whether two statements written in natural language have the identical semantic or not. Its importance appears in many fields like plagiarism detection, question answering, document clustering and information retrieval, etc. This paper proposes a hybrid model that combines the text similarity approach with deep learning approach in order to improve paraphrase detection. This model verified results with Microsoft Research Paraphrase Corpus (MSPR) dataset, shows that accuracy measure is about 76.6% and F-measure is about 83.5%.

References

Processing (pp. 891-896).


Index Terms

Computer Science

Information Sciences

Keywords

Paraphrase detection, Deep Learning, Skip thought vector, Text similarity