Abstract

In this paper, a plagiarism detection tool for plagiarism detection in Malayalam documents is presented. Many language-sensitive tools for detecting plagiarism in natural language documents have been developed, particularly for English. Detecting plagiarism in Malayalam documents is particularly a challenging task because of the complex linguistic structure of Malayalam. The plagiarism detection tool presented here has the mechanism of detecting similarity beyond exact words match of Malayalam documents. The tool is based on a new comparison algorithm that uses some NLP techniques to compare suspect documents which may not be identified using existing methods for Malayalam document plagiarism detection.

References

- Paul Clough 2000 Plagiarism in natural and programming languages: an overview of current tools and technologies, Department of Computer Science, University of Sheffield, technical report


Index Terms

 Computer Science                      Pattern Recognition

Keywords

Plagiarism Detection    Malayalam    Natural Language Processing    Lemmatization