Abstract

Present paper describes the details of the study of the work that has been done in the field of text searching, a sub-division of Natural Language Processing (NLP) till date. The work in this project includes the study and analysis of some of the algorithms devised under this topic, finding the faults or loop-holes and trying to increase the efficiency of these algorithms devised, taking forward the range of work done on it. Experiment is done on the various text search algorithms that have been devised namely Knuth-Morris Pratt Algorithm, Naïve Search Algorithm and Boyer-Moore Algorithm by providing text input of various sizes and analyzing their behavior on these variable inputs. After analyzing and doing the study on these algorithms the results states that Boyer-Moore’s Algorithm worked quite well and efficiently than the
rest of them when dealing with larger data sets. When working on larger alphabets the Knuth-Morris Pratt Algorithm works quite well. These algorithms do have drawbacks as their efficiency depends upon the alphabet/pattern size. And also this paper describes new pattern matching algorithm that uses delimiter for shifting the pattern while matching.

References


Index Terms

Computer Science Algorithms

Keywords

Nlp  Kmp Algorithm Naive Search Bm Algorithm.