



Abstract

Information retrieval (IR) is finding text documents material of unstructured nature that satisfies information need from within large collections. The need for effective methods of IR has grown in importance because of the tremendous explosion in the amount of unstructured data, and growing number of document sources on the Internet. This work is focussed on document ranking system by a search engine in information retrieval . The objective of this project is to formulate unstructured queries, indexing (including term weighting) of document collections, methods for computing the similarity of queries and documents, the scoring of a document for a query leading to the idea of a list of documents that are ranked ordered for a query. In this project I would like to point out the few drawbacks of existence syntax based traditional retrieval system and will give ideas on how to improve their performance.

Keywords: unstructured document and query, indexing, scoring, ranking, Syntax Based Search, Context Based Search.