

What is Information Retrieval System

It is concerned with representing, searching, manipulating large collection of electronic text and other human language data.

Finding document (data) of an unstructured nature

Corpus:

Large repository of document stored in computer.

Information Need:

A topic about which we want to get information.

Relevance:

Some of the document that may contain what I want to search.

Types of Searches:

1.Desktop Search:

Inside a system search. Example - file

2.Search Engine Search:

Do an Search engines. Example - Google , chrome etc.

3.Social Search:

Search of Social networking sites.

4.Enterprise Search:

GLA student search on GLA App.

Methods to develop IRS:

1.Classical

binary Retrieval model

(i) TDIM

(ii) Inverted Index

2.Non Classical

Binary Retrival Model:

This model work on 3 expression

AND, OR and NOT

Term Document Incidence Matrix (TDIM):

In which document term is present mark as 1 and in absense mark as 0.

Example:

Doc: 1 Break through drug for covid.

Doc: 2 New Covid drug.

Doc: 3 New approach for treatment of covid.

Doc: 4 New hope for covid patient.

Queries:

(i) Covid AND Drug

(ii) New OR Hope

(iii) New AND (NOT) Drug

Step:1

Create matrix of all document having term as word pesent in it without repeating.

Step:2

Perform the Query as we want.