**ABSTRACT**

The web is the huge and most extravagant wellspring of data. To recover the information from the World Wide Web, Search Engines are commonly utilized. Search engines provide a simple interface for searching for user query and displaying results in the form of the web address of the relevant web page, but using traditional search engines has become very challenging to obtain suitable information. This paper proposed a search engine using Machine Learning technique that will give more relevant web pages at top for user queries.

**EXISTING SYSTEM:**

To create a new Programmable Search Engine, all you have to do is choose which sites to search and give your search engine a name. In the Sites to search box, type one or more sites you want to include in the search results. You can include any sites on the web, even sites you don't own. By our calculations, for mid to large size retailer to build its own high-quality, Solr-based site search engine would take 30 to 40 engineers as long as two years.140+ Search Engines and Directories. Search, the holy grail that pushed Google into global Internet domination, is still coveted by many. The fact that most users don't even consider switching Google for anything else doesn't mean that there's no innovation going on in the field of search.

**DISADVANTAGES OF SEARCH ENGINE**

* Sometimes the search engine takes too much time to display relevant, valuable, and informative content.
* Search engines, especially Google, frequently update their algorithm, and it is very difficult to find the algorithm in which Google runs.

**PROPOSED SYSTEM:**

To build a search engine which gives web address of the most relevant web page at the top of the search result, according to user queries. The main focus of our system is to build a search engine using machine learning technique for increasing accuracy compare to available search engine.

Following is the step by step procedure for building the search engine:

1) Collect data from WWW using web crawler.

2) Perform data cleaning using NLP.

3) Study and compare the existing page ranking algorithm.

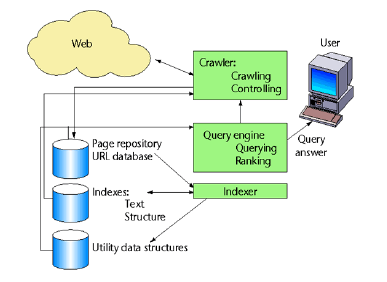
4) Merge the selected page rank algorithm with current technologies in machine learning.

5) Implement query engine to display the efficient results for user query.

**ADVANTAGES:**

* Time Savings. A **search engine** saves you time in two ways: by eliminating the need to find information manually, and by performing **searches** at high speeds.
* Relevance. When a **search engine** scans a website, it scores the content for relevance to particular **search** words.
* Free Access.
* Comprehensive.
* Advanced **Search**.

**SYSTEM ARCHITECTURE:**

****

**SYSTEM SPECIFICATION:**

**HARDWARE REQUIREMENTS:**

* **System :** Pentium IV 2.4 GHz.
* **Hard Disk :** 40 GB.
* **Floppy Drive :** 1.44 Mb.
* **Monitor** : 14’ Colour Monitor.
* **Mouse :** Optical Mouse.
* **Ram :** 512 Mb.

**SOFTWARE REQUIREMENTS:**

* **Operating system :** Windows 7 Ultimate.
* **Coding Language :** Python.
* **Front-End :** Python.
* **Designing :** Html,css,javascript.
* **Data Base :** MySQL.