Course code: IT362
Course Name: Information Retrieval
L-T-P - Credits: 3-0-0-3
Year of Introduction: 2017

Pre-requisites: CS205 Data structures

Course Objectives:
- To provide with foundation knowledge in information retrieval.
- To equip with sound skills to solve computational search problems.

Syllabus
Introduction to the Concepts of Information Retrieval, Retrieval models, Searching the web and Parallel and Distributed Information Retrieval systems.

Expected outcome:
The students will be able to
1. use different information retrieval techniques in various application areas
2. apply IR principles to locate relevant information large collections of data and analyse performance of retrieval systems when dealing with unmanaged data sources
3. implement retrieval systems for web search tasks.

Text Books:
2. C.J. Van Rijsbergen, Information Retrieval; http://www.dcs.gla.ac.uk/Keith/Preface.html

References:

Course Plan

<table>
<thead>
<tr>
<th>Module</th>
<th>Contents</th>
<th>Hours</th>
<th>Sem. Exam Marks</th>
</tr>
</thead>
<tbody>
<tr>
<td>II</td>
<td>Classification, Measures of Association, Cluster Hypothesis, Single Link Clusters, File Structures, Inverted Files, Index Sequential Files, Ring Structures, Doubly Chained Trees, Hash Addressing.</td>
<td>7</td>
<td>15%</td>
</tr>
</tbody>
</table>

FIRST INTERNAL EXAMINATION

| III    | Evaluation, Relevance, Precision and Recall, Interpolation, Averaging techniques, The Swets Model. | 7 | 15% |
| IV     | Search Engines, Boolean Search, Matching Functions, Serial Search, Cluster Representatives, Cluster based retrieval. | 7 | 15% |

SECOND INTERNAL EXAMINATION

<p>| V      | Web search basics – Web characteristics - crawling and indexes – Features of a crawler – Crawler architecture – DNS | 7 | 20% |</p>
<table>
<thead>
<tr>
<th>Resolution – The URL frontier – Distributing indexes – Connectivity servers.</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>VI</strong> Link Analysis – The Web as a graph – Anchor text and the web graph, PageRank – Markov chains, Page Rank computation, Topic-specific Page Rank, Hubs and authorities.</td>
</tr>
<tr>
<td>7</td>
</tr>
</tbody>
</table>