Effective Term
Semester A 2022/23

Part I Course Overview

Course Title
Information Retrieval

Subject Code
CS - Computer Science

Course Number
4485

Academic Unit
Computer Science (CS)

College/School
College of Engineering (EG)

Course Duration
One Semester

Credit Units
3

Level
B1, B2, B3, B4 - Bachelor's Degree
R8 - Research Degree

Medium of Instruction
English

Medium of Assessment
English

Prerequisites
CS3402 Database Systems

Precursors
Nil

Equivalent Courses
Nil

Exclusive Courses
Nil
Part II Course Details

Abstract
This course aims to provide a broad view and detailed knowledge of all key topics in modern information retrieval (IR). Basic concepts such as retrieval evaluation, query languages, query operations, indexing and searching are introduced. Some advanced topics including parallel and distributed IR, and multimedia IR are discussed.

Course Intended Learning Outcomes (CILOs)

<table>
<thead>
<tr>
<th>CILOs</th>
<th>Weighting (if app.)</th>
<th>DEC-A1</th>
<th>DEC-A2</th>
<th>DEC-A3</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Describe the basic concepts and models in information retrieval.</td>
<td>x</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>Apply basic techniques in user relevance feedback, indexing and searching.</td>
<td>x</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>Design and build a simple text retrieval systems. Such a system includes (1) storage of documents (converting documents into vectors); (2) ranking algorithms and relevance feedback algorithms.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>Read research papers in this area.</td>
<td>x</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

A1: Attitude
Develop an attitude of discovery/innovation/creativity, as demonstrated by students possessing a strong sense of curiosity, asking questions actively, challenging assumptions or engaging in inquiry together with teachers.

A2: Ability
Develop the ability/skill needed to discover/innovate/create, as demonstrated by students possessing critical thinking skills to assess ideas, acquiring research skills, synthesizing knowledge across disciplines or applying academic knowledge to real-life problems.

A3: Accomplishments
Demonstrate accomplishment of discovery/innovation/creativity through producing /constructing creative works/new artefacts, effective solutions to real-life problems or new processes.

Teaching and Learning Activities (TLAs)

<table>
<thead>
<tr>
<th>TLAs</th>
<th>Brief Description</th>
<th>CILO No.</th>
<th>Hours/week (if applicable)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Lectures will be given to introduce the basic concepts, information retrieval models, basic techniques in user relevance feedback, indexing and searching; (2) some exercises (using concept examples) will be given in the tutorial to illustrate the concepts, methods and models. For the first few weeks in the semester, small assignments will be given.</td>
<td>1, 2, 3, 4</td>
<td>3 hours/week</td>
</tr>
</tbody>
</table>
### Assessment Tasks / Activities (ATs)

<table>
<thead>
<tr>
<th>ATs</th>
<th>CILO No.</th>
<th>Weighting (%)</th>
<th>Remarks (e.g. Parameter for GenAI use)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Assignment</td>
<td>1, 2</td>
<td>10</td>
<td></td>
</tr>
<tr>
<td>Midterm test</td>
<td>1, 2</td>
<td>10</td>
<td></td>
</tr>
<tr>
<td>Mini-project</td>
<td>3, 4</td>
<td>30</td>
<td></td>
</tr>
</tbody>
</table>

**Continuous Assessment (%)**

50
Examination (%)  
50  
Examination Duration (Hours)  
2  
Additional Information for ATs  
For a student to pass the course, at least 30% of the maximum mark for the examination must be obtained.

Assessment Rubrics (AR)  

Assessment Task  
Assignments  

Criterion  
Question has a score  

Excellent (A+, A, A-)  
High  

Good (B+, B, B-)  
Significant  

Fair (C+, C, C-)  
Moderate  

Marginal (D)  
Basic  

Failure (F)  
Not even reaching marginal levels  

Assessment Task  
Mini-project  

Criterion  
Ability to complete the system, creating new ideas/methods  

Excellent (A+, A, A-)  
High  

Good (B+, B, B-)  
Significant  

Fair (C+, C, C-)  
Moderate  

Marginal (D)  
Basic  

Failure (F)  
Not even reaching marginal levels
Assessment Task
Midterm

Criterion
Question has a score

Excellent (A+, A, A-)
High

Good (B+, B, B-)
Significant

Fair (C+, C, C-)
Moderate

Marginal (D)
Basic

Failure (F)
Not even reaching marginal levels

---

Assessment Task
Final exam

Criterion
Question has a score

Excellent (A+, A, A-)
High

Good (B+, B, B-)
Significant

Fair (C+, C, C-)
Moderate

Marginal (D)
Basic

Failure (F)
Not even reaching marginal levels

---

Part III Other Information

Keyword Syllabus

Syllabus
• Retrieval evaluation
  Recall, precision, and alternative measures for the quality of queries.
• Query language
• Query operations
• Indexing and searching
  Indexing methods include inverted files, suffix trees, and signature files. Searching: exact matching methods including KMP, Boyer-Moore algorithm, etc. Approximate pattern matching including string matching allowing errors, regular expressions and extended patterns.
• Parallel and Distributed IR
  Parallel IR: MIMD and SIMD architectures. Distributed IR: Collection partitioning, source selection, query processing and web issues.
• Multimedia IR
  Models: Multimedia data support in commercial DBMSs, MULTOS data model. Query languages: request specification, conditions on multimedia data, uncertainty, proximity and weights in query expressions. Indexing and searching: spatial access methods, a generic multimedia indexing approach, one-dimensional time series, two-dimensional color images, and automatic feature extraction.
• Others
  Searching the web. Libraries and bibliographical systems. Digital libraries
  (This part is optional.)

Reading List

Compulsory Readings

<table>
<thead>
<tr>
<th>Title</th>
<th>1</th>
<th>Nil</th>
</tr>
</thead>
</table>

Additional Readings

<table>
<thead>
<tr>
<th>Title</th>
<th>1</th>
<th>Nil</th>
</tr>
</thead>
</table>