FIT5166
Information retrieval systems

Unit Guide

Semester 2, 2012

The information contained in this unit guide is correct at time of publication. The University has the right to change any of the elements contained in this document at any time.

Last updated: 18 Jul 2012
Table of Contents
FIT5166 Information retrieval systems - Semester 2, 2012 ................................................................. 1

Mode of Delivery ............................................................................................................................... 1
Contact Hours ................................................................................................................................. 1
Workload ........................................................................................................................................ 1
Chief Examiner ................................................................................................................................. 1
Campus Lecturer .............................................................................................................................. 1
Caulfield ........................................................................................................................................... 1
Tutors ................................................................................................................................................ 1
Caulfield ........................................................................................................................................... 1

Academic Overview ........................................................................................................................ 2
Outcomes .......................................................................................................................................... 2
Graduate Attributes .......................................................................................................................... 2
Assessment Summary ...................................................................................................................... 2
Teaching Approach .......................................................................................................................... 3
Feedback .......................................................................................................................................... 3
  Our feedback to You ................................................................................................................... 3
  Your feedback to Us ................................................................................................................... 3
Previous Student Evaluations of this unit ....................................................................................... 3
Recommended Resources .............................................................................................................. 4

Unit Schedule ................................................................................................................................... 5

Assessment Requirements .............................................................................................................. 6
  Assessment Policy ....................................................................................................................... 6
  Assessment Tasks ....................................................................................................................... 6
    Participation .............................................................................................................................. 6
Examinations .................................................................................................................................... 8
  Examination 1 ............................................................................................................................ 8
Assignment submission ................................................................................................................... 8
  Online submission ...................................................................................................................... 8
Extensions and penalties .................................................................................................................. 8
Returning assignments ..................................................................................................................... 8
Resubmission of assignments ......................................................................................................... 8
Referencing requirements .............................................................................................................. 9

Other Information .......................................................................................................................... 10
  Policies ........................................................................................................................................ 10
  Student services ......................................................................................................................... 10
  Other .......................................................................................................................................... 11
FIT5166 Information retrieval systems - Semester 2, 2012

This unit presents students with the theory and practice underlying computerised information retrieval. Topics covered include: history and context of information retrieval systems, retrieval models, Boolean, vector space and probabilistic, evaluation strategies, test collections, web search engines, indexing, content-based multimedia retrieval and relevance feedback.

Mode of Delivery

Caulfield (Day)

Contact Hours

2 hrs lectures/wk, 2 hrs laboratories/wk

Workload

Students will be expected to spend up to a total of 12 hours per week during semester on this unit as follows:

- Watching video presentations - up to 1 hour/week
- Discussion and participation in class - 2 hours/week
- Tutorial Sessions: up to 2 hours per week per tutorial at discretion of student
- Assignment work

Chief Examiner

Dr Campbell Wilson

Campus Lecturer

Caulfield

Campbell Wilson

Tutors

Caulfield

Michael Niemann

Viranga Ratnaike

Shan He (Dora)
Academic Overview

Outcomes

At the completion of this unit students will:

- appreciate the context and application of information retrieval systems;
- understand the different models of information retrieval and their comparative advantages and disadvantages;
- understand how objective and subjective evaluation strategies are used with information retrieval systems;
- be conversant with the issues and challenges of managing very large collections of heterogeneous data for information retrieval;
- understand how web search engines and search algorithms are constructed, utilised and deployed;
- appreciate how to index different media;
- understand how information retrieval systems can be used for the retrieval of audio-visual information;
- appreciate the theoretical and practical underpinnings of relevance feedback in information retrieval systems.

Graduate Attributes

Monash prepares its graduates to be:

1. responsible and effective global citizens who:
   a. engage in an internationalised world
   b. exhibit cross-cultural competence
   c. demonstrate ethical values

2. critical and creative scholars who:
   a. produce innovative solutions to problems
   b. apply research skills to a range of challenges
   c. communicate perceptively and effectively

Assessment Summary

Examination (2 hours): 50%; In-semester assessment: 50%

<table>
<thead>
<tr>
<th>Assessment Task</th>
<th>Value</th>
<th>Due Date</th>
</tr>
</thead>
<tbody>
<tr>
<td>Assignment 1 - Hello World</td>
<td>1%</td>
<td>Week 1</td>
</tr>
<tr>
<td>Assignment 2 - Tokenization</td>
<td>4%</td>
<td>Week 3</td>
</tr>
<tr>
<td>Assignment 3 - Constructing an inverted index</td>
<td>6%</td>
<td>Week 5</td>
</tr>
<tr>
<td>Assignment 4 - Vector space retrieval</td>
<td>6%</td>
<td>Week 7</td>
</tr>
<tr>
<td>Report</td>
<td>23%</td>
<td>Week 10</td>
</tr>
<tr>
<td>Presentation</td>
<td>10%</td>
<td>Week 11</td>
</tr>
</tbody>
</table>
Teaching Approach

Lecture and tutorials or problem classes

We are hoping to supplement lectures with video presentations and subsequent discussion.

Feedback

Our feedback to You

Types of feedback you can expect to receive in this unit are:

- Informal feedback on progress in labs/tutes
- Graded assignments without comments
- Interviews
- Quiz results

Your feedback to Us

Monash is committed to excellence in education and regularly seeks feedback from students, employers and staff. One of the key formal ways students have to provide feedback is through SETU, Student Evaluation of Teacher and Unit. The University's student evaluation policy requires that every unit is evaluated each year. Students are strongly encouraged to complete the surveys. The feedback is anonymous and provides the Faculty with evidence of aspects that students are satisfied and areas for improvement.

For more information on Monash's educational strategy, and on student evaluations, see:
http://www.policy.monash.edu/policy-bank/academic/education/quality/student-evaluation-policy.html

Previous Student Evaluations of this unit

The following lists changes that have been made to the unit. Some of these have been made in response to student feedback, however due to the University's requirements for confidentiality, this feedback is not allowed to be included in the unit guide. Hence some of these improvements will lack context:

- Demonstration videos will be provided so students can look at material at their own pace.
- Code templates provided for programming assignments and marking criteria modified.
- Unit is available in the Intelligent Systems specialisation
- Assignment deadlines do not co-occur
- Revision classes have been removed
Academic Overview

- Modified material on multimedia IR

- Programming assignment specifications simplified

If you wish to view how previous students rated this unit, please go to https://emuapps.monash.edu.au/unitevaluations/index.jsp

Recommended Resources

Recommended reading will be provided during the lectures as appropriate.
### Unit Schedule

<table>
<thead>
<tr>
<th>Week</th>
<th>Activities</th>
<th>Assessment</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>This schedule is subject to change, please refer to the Lecturer for updates</td>
<td>Assessment due dates may be subject to change. Any changes will be communicated to you by the Lecturer.</td>
</tr>
<tr>
<td>1</td>
<td>Introduction</td>
<td>Assignment 1 (Hello World) due</td>
</tr>
<tr>
<td>2</td>
<td>Indexing, lexical analyses and the Boolean model</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>Vector space model</td>
<td>Assignment 2 (Tokenizing) due</td>
</tr>
<tr>
<td>4</td>
<td>Content-based image retrieval (CBIR) I</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>Vector space model II</td>
<td>Assignment 3 (Indexing) due</td>
</tr>
<tr>
<td>6</td>
<td>Web search</td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>Probabilistic models: Language and topic models</td>
<td>Assignment 4 (Vector Space) due</td>
</tr>
<tr>
<td>8</td>
<td>Relevance feedback and text classification</td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>CBIR II and multimedia retrieval</td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>Advanced topics</td>
<td>Report due</td>
</tr>
<tr>
<td>11</td>
<td>Discussion class</td>
<td>Presentations</td>
</tr>
<tr>
<td>12</td>
<td>Evaluation</td>
<td>Presentations</td>
</tr>
<tr>
<td></td>
<td>SWOT VAC</td>
<td></td>
</tr>
</tbody>
</table>

*Unit Schedule details will be maintained and communicated to you via your MUSO (Blackboard or Moodle) learning system.*
Assessment Requirements

Assessment Policy

Faculty Policy - Unit Assessment Hurdles

Academic Integrity - Please see the Demystifying Citing and Referencing tutorial at http://lib.monash.edu/tutorials/citing/

Assessment Tasks

Participation

Please ask lots of questions during lectures and tutorials. You won't be assessed on the quality of your questions, but you will get a lot more out of the unit!

• **Assessment task 1**
  
  **Title:**
  Assignment 1 - Hello World
  
  **Description:**
  Can I follow instructions? :)
  
  **Weighting:**
  1%
  
  **Criteria for assessment:**
  Assignment is graded by a computer program, which measures correlation between student answers and expected results. The point of this assignment is for students to familiarize students with the submission system used in the unit. They therefore are expected to complete this task in order to proceed in the unit. The submission system will be set up as such.
  
  Note: special exemptions for people who join the subject late (within the normal first two week enrolment period) will be put in place.
  
  **Due date:**
  Week 1

• **Assessment task 2**
  
  **Title:**
  Assignment 2 - Tokenization
  
  **Description:**
  Tokenization of text.
  
  **Weighting:**
  4%
  
  **Criteria for assessment:**
  Assignment is graded by a computer program, which measures correlation between student answers and expected results.
  
  **Due date:**
  Week 3
Assessment Requirements

• **Assessment task 3**

**Title:**
Assignment 3 - Constructing an inverted index

**Description:**
Constructing an inverted index.

**Weighting:**
6%

**Criteria for assessment:**
Assignment is graded by a computer program, which measures correlation between student answers and expected results.

**Due date:**
Week 5

• **Assessment task 4**

**Title:**
Assignment 4 - Vector space retrieval

**Description:**
Building a vector space retrieval engine.

**Weighting:**
6%

**Criteria for assessment:**
Assignment is graded by a computer program, which measures correlation between student answers and expected results.

**Due date:**
Week 7

• **Assessment task 5**

**Title:**
Report

**Description:**
A 3,000 word report on a research topic related to information retrieval.

**Weighting:**
23%

**Criteria for assessment:**
The report will be assessed on the usual criteria, namely: breadth of literature survey, quality of analysis of literature, and topicality. Appropriate use of the English language will be required.

**Due date:**
Week 10

• **Assessment task 6**

**Title:**
Presentation

**Description:**
Presentation of research report and taking of questions.

**Weighting:**
10%

**Criteria for assessment:**
Quality of presentation and quality of response to questions

**Due date:**
Examinations

- Examination 1
  Weighting: 50%
  Length: 2 hours
  Type (open/closed book): Closed book
  Electronic devices allowed in the exam: None

Assignment submission

It is a University requirement (http://www.policy.monash.edu/policy-bank/academic/education/conduct/plagiarism-procedures.html) for students to submit an assignment coversheet for each assessment item. Faculty Assignment coversheets can be found at http://www.infotech.monash.edu.au/resources/student/forms/. Please check with your Lecturer on the submission method for your assignment coversheet (e.g. attach a file to the online assignment submission, hand-in a hard copy, or use an online quiz).

Online submission

If Electronic Submission has been approved for your unit, please submit your work via the VLE site for this unit, which you can access via links in the my.monash portal.

Extensions and penalties

Submission must be made by the due date otherwise penalties will be enforced.


Returning assignments

Students can expect assignments to be returned within two weeks of the submission date or after receipt, whichever is later.

Resubmission of assignments

Students may resubmit assignments 1, 2, 3 and 4 any number of times before the due date.
Referencing requirements

Referencing should follow the IEEE style.
Other Information

Policies

Monash has educational policies, procedures and guidelines, which are designed to ensure that staff and students are aware of the University’s academic standards, and to provide advice on how they might uphold them. You can find Monash's Education Policies at: http://policy.monash.edu.au/policy-bank/academic/education/index.html

Key educational policies include:

- Plagiarism (http://www.policy.monash.edu/policy-bank/academic/education/conduct/plagiarism-policy.html)
- Special Consideration (http://www.policy.monash.edu/policy-bank/academic/education/assessment/special-consideration-policy.html)
- Grading Scale (http://www.policy.monash.edu/policy-bank/academic/education/assessment/grading-scale-policy.html)
- Discipline: Student Policy (http://www.policy.monash.edu/policy-bank/academic/education/conduct/student-discipline-policy.html)
- Academic Calendar and Semesters (http://www.monash.edu.au/students/key-dates/);
- Orientation and Transition (http://www.infotech.monash.edu.au/resources/student/orientation/); and

Student services

The University provides many different kinds of support services for you. Contact your tutor if you need advice and see the range of services available at www.monash.edu.au/students. For Sunway see http://www.monash.edu.my/Student-services, and for South Africa see http://www.monash.ac.za/current/.

The Monash University Library provides a range of services and resources that enable you to save time and be more effective in your learning and research. Go to http://www.lib.monash.edu.au or the library tab in my.monash portal for more information. At Sunway, visit the Library and Learning Commons at http://www.lib.monash.edu.my/. At South Africa visit http://www.lib.monash.ac.za/.

Academic support services may be available for students who have a disability or medical condition. Registration with the Disability Liaison Unit is required. Further information is available as follows:

- Website: http://monash.edu/equity-diversity/disability/index.html;
- Email: dlu@monash.edu
- Drop In: Equity and Diversity Centre, Level 1 Gallery Building (Building 55), Monash University, Clayton Campus, or Student Community Services Department, Level 2, Building 2, Monash University, Sunway Campus
- Telephone: 03 9905 5704, or contact the Student Advisor, Student Community Services at 03 55146018 at Sunway
Other

Note: We are trialling a new approach in this unit, whereby lectures may be supplemented by video presentations, which will be discussed in the class time.

Note: Damocles (a plagiarism detection system) will be employed for the research report submission.