# Table of Contents

FIT5142 Advanced data mining - Semester 2, 2015

- Mode of Delivery................................................................. 1
- Workload Requirements....................................................... 1
- Unit Relationships.............................................................. 1
  - Prerequisites................................................................. 1
- Chief Examiner................................................................. 1
- Campus Lecturer............................................................... 1
  - Caulfield....................................................................... 1
- Tutors................................................................................. 2
  - Caulfield....................................................................... 2
- Your feedback to Us........................................................... 2
- Previous Student Evaluations of this Unit............................. 2

**Academic Overview**.................................................................. 3
  - Learning Outcomes.......................................................... 3

**Unit Schedule**......................................................................... 4
  - Teaching Approach.......................................................... 4
  - Assessment Summary...................................................... 5

**Assessment Requirements**...................................................... 6
  - Assessment Policy.......................................................... 6
  - Assessment Tasks........................................................... 6
  - Participation................................................................. 6
  - Examinations................................................................. 7
    - Examination 1........................................................... 7
  - Learning resources......................................................... 8
  - Feedback to you............................................................ 8
  - Extensions and penalties.................................................. 8
  - Returning assignments................................................... 8
  - Assignment submission................................................ 8
  - Online submission....................................................... 8

**Other Information**................................................................. 9
  - Policies............................................................................. 9
  - Faculty resources and policies......................................... 9
    - Graduate Attributes Policy......................................... 9
  - Student Charter............................................................ 9
  - Student services.......................................................... 9
  - Monash University Library............................................. 9
  - Disability Liaison Unit.................................................... 9
FIT5142 Advanced data mining - Semester 2, 2015

Advanced methods of discovering patterns in large-scale multi-dimensional databases are discussed. Solving classification, clustering, association rules analysis and regression problems on different kinds of data are covered. Data pre-processing methods for dealing with noisy and missing data in the context of Big Data are reviewed. Evaluation and analysis of data mining models are emphasised. Hands-on case studies in building data mining models are performed using popular modern software packages.

Mode of Delivery

Caulfield (Day)

Workload Requirements

Minimum total expected workload equals 12 hours per week comprising:

(a.) Contact hours for on-campus students:

• Two hours of lectures
• One 2-hour laboratory

(b.) Additional requirements (all students):

• A minimum of 8 hours independent study per week for completing lab and project work, private study and revision.

See also Unit timetable information

Unit Relationships

Prerequisites

FIT5047 or FIT5045 or equivalent
Sound fundamental knowledge in maths and statistics; database and computer programming knowledge.

Chief Examiner

Dr Grace Rumantir

Campus Lecturer

Caulfield

Grace Rumantir

Consultation hours: Monday 4pm-5pm, Thursday 1-2pm or by appointment
Tutors

Caulfield

Ashish Singh

Consultation hours: TBA

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 the Student Evaluation of Teaching and Units (SETU) survey. 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, see:

www.monash.edu.au/about/monash-directions/ and on student evaluations, see:
www.policy.monash.edu/policy-bank/academic/education/quality/student-evaluation-policy.html

Previous Student Evaluations of this Unit

Based on previous student feedback this unit is well structured and no major changes have been made for this semester. Further refinements of the materials on the topics covered will be made on a weekly basis and made available on Moodle.

If you wish to view how previous students rated this unit, please go to
Academic Overview

Learning Outcomes

On successful completion of this unit, students should be able to:

- explain the kinds of data from which knowledge can be mined, the way each data type can be presented to a data mining algorithm, the kinds of patterns that can be mined from each data type;
- evaluate the quality of data mining models;
- perform pre-processing of large-scale multi-dimensional datasets in preparation for data mining experiments;
- perform data pre-processing for data with outliers, incomplete and noisy data;
- compare the various learning algorithms and the ability to effectively apply suitable algorithms to mine frequent patterns and associations from data, to perform data classification, data clustering and regression analysis;
- use modern data mining tools to solve non-trivial data mining problems;
- research the current trends in data mining applications;
- work in a team to extract knowledge from a common dataset using various data mining methods and techniques.
# Unit Schedule

<table>
<thead>
<tr>
<th>Week</th>
<th>Activities</th>
<th>Assessment</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td></td>
<td>No formal assessment or activities are undertaken in week 0</td>
</tr>
<tr>
<td>1</td>
<td>Introduction</td>
<td>There is a self-assessed test (not marked) on basic maths and statistics and the fundamentals of Data Mining on Moodle that will be discussed in the Week 1 tutorial. Please complete this to see if you need to do further study prior to completing this unit.</td>
</tr>
<tr>
<td>2</td>
<td>Data Preprocessing</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>Data Warehousing and Data Mining</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>Classification and Prediction</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>Cluster Analysis</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>Mining Stream, Time-Series and Sequential Data</td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>Graph Mining, Social Network Analysis and Multirelational Data Mining</td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>Unit Test (during the lecture timeslot, tutorials are still on)</td>
<td>Unit Test during Week 8 lecture (Monday 14 September 2015)</td>
</tr>
<tr>
<td>9</td>
<td>Ensemble Methods in Data Mining</td>
<td>Assignment Stage 1 due online Monday 21 September 2015 at 11am</td>
</tr>
<tr>
<td>10</td>
<td>Mining Object, Spatial, Multimedia, Text and Web Data (Part 1)</td>
<td></td>
</tr>
<tr>
<td>11</td>
<td>Mining Object, Spatial, Multimedia, Text and Web Data (Part 2)</td>
<td>Assignment Stage 2 due on online Monday 12 October 2015 at 11am</td>
</tr>
<tr>
<td>12</td>
<td>Application &amp; Trends in Data Mining and Revision</td>
<td></td>
</tr>
<tr>
<td></td>
<td>SWOT VAC</td>
<td>No formal assessment is undertaken in SWOT VAC</td>
</tr>
</tbody>
</table>

*Unit Schedule details will be maintained and communicated to you via your learning system.

# Teaching Approach

**Lecture and tutorials or problem classes**

This teaching and learning approach helps students to initially encounter information at lectures, discuss and explore the information during tutorials, and practice in a hands-on lab environment.
## Assessment Summary

Examination (3 hours): 60%; In-semester assessment: 40%

<table>
<thead>
<tr>
<th>Assessment Task</th>
<th>Value</th>
<th>Due Date</th>
</tr>
</thead>
<tbody>
<tr>
<td>Unit Test</td>
<td>17%</td>
<td>Unit Test during Week 8 lecture (Monday 14 September 2015)</td>
</tr>
<tr>
<td>Report on Implementations of Advanced Data Mining Techniques</td>
<td>17%</td>
<td>Assignment Stage 1 due online Monday 21 September 2015. Assignment Stage 2 due online Monday 12 October 2015)</td>
</tr>
<tr>
<td>9 Weekly Revision Quizzes</td>
<td>3%</td>
<td>Each weekly quiz will open at the end of the last tutorial in the week and will close at midnight on the Sunday of the week.</td>
</tr>
<tr>
<td>Participation in the clicker sessions in the lecture from Week 2 to Week 12</td>
<td>3%</td>
<td>Weekly (during the lectures from Week 2 to Week 12)</td>
</tr>
<tr>
<td>Examination 1</td>
<td>60%</td>
<td>To be advised</td>
</tr>
</tbody>
</table>
Assessment Requirements

Assessment Policy

Faculty Policy - Unit Assessment Hurdles

Academic Integrity - Please see resources and tutorials at
http://www.monash.edu/library/skills/resources/tutorials/academic-integrity/

Assessment Tasks

Participation

• Assessment task 1

Title:
Unit Test

Description:
Learning Outcomes: 1, 2, 3, 4

Closed-book unit test to be conducted in the lecture time slot in Week 8.

Weighting:
17%

Criteria for assessment:
Correct answers to questions, and quality of solutions to problems, which demonstrates understanding of the learning materials. Further detail of the format and coverage of the unit test will be made available on Moodle.

Due date:
Unit Test during Week 8 lecture (Monday 14 September 2015)

• Assessment task 2

Title:
Report on Implementations of Advanced Data Mining Techniques

Description:
Learning Outcomes: 5, 6, 7, 8

A pair of students work in a group to come up with a report on an approved research topic on advances in Data Mining and the students’ implementations of the relevant advanced data mining methods (maximum 20 page long)

Weighting:
17%

Criteria for assessment:
The report will be assessed on the usual criteria, namely: breadth of literature survey, quality of analysis of literature, topicality, the use of a data set to illustrate the implementations of the relevant advanced data mining methods.

There are 2 stages of the assignment:

Stage 1: Write up of the structure of the report and the aspects to be covered in the report (non-assessable)
Stage 2: Submission (17%).

Due date:
Assignment Stage 1 due online Monday 21 September 2015. Assignment Stage 2 due online Monday 12 October 2015

• Assessment task 3

Title:
9 Weekly Revision Quizzes

Description:
Learning Outcomes:1 to 7

Weekly online revision quizzes will be conducted from Week 2. Each weekly quiz covers the material covered in the respective week.

Weighting:
3%

Criteria for assessment:
Each weekly online quiz will have 10 multiple choice questions. There is no negative mark for wrong answers. The mark of each quiz will be recorded based on student's submission of his/her quiz attempt prior to the closing date of each quiz.

Due date:
Each weekly quiz will open at the end of the last tutorial in the week and will close at midnight on the Sunday of the week.

• Assessment task 4

Title:
Participation in the clicker sessions in the lecture from Week 2 to Week 12

Description:
Participation in the clicker sessions will be marked during the lectures

Weighting:
3%

Criteria for assessment:
Active participation in lecture sessions using the response systems is expected. 3% of the final marks is allocated for this participation. The mark will be allocated based on the percentage of recorded participation in the clicker sessions during lectures.

Due date:
Weekly (during the lectures from Week 2 to Week 12)

Examinations

• Examination 1

Weighting:
60%

Length:
3 hours

Type (open/closed book):
Closed book

Electronic devices allowed in the exam:
Scientific Calculators
Learning resources

Monash Library Unit Reading List (if applicable to the unit)
http://readinglists.lib.monash.edu/index.html

Feedback to you

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

- Informal feedback on progress in labs/tutes
- Graded assignments with comments
- Test results and feedback
- Quiz results

Extensions and penalties

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

You must negotiate any extensions formally with your campus unit leader via the in-semester special consideration process: http://www.monash.edu.au/exams/special-consideration.html

Returning assignments

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

Assignment submission

It is a University requirement 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 electronic submission). Please note that it is your responsibility to retain copies of your assessments.

Online submission

If Electronic Submission has been approved for your unit, please submit your work via the learning system for this unit, which you can access via links in the my.monash portal.
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: www.policy.monash.edu.au/policy-bank/academic/education/index.html

Faculty resources and policies

Important student resources including Faculty policies are located at http://intranet.monash.edu.au/infotech/resources/students/

Graduate Attributes Policy

http://www.policy.monash.edu/policy-bank/academic/education/management/monash-graduate-attributes-policy.html

Student Charter


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 http://www.monash.edu.au/students. For Malaysia see http://www.monash.edu.my/Student-services, and for South Africa see http://www.monash.ac.za/current/.

Monash University Library

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

Disability Liaison Unit

Students who have a disability or medical condition are welcome to contact the Disability Liaison Unit to discuss academic support services. Disability Liaison Officers (DLOs) visit all Victorian campuses on a regular basis.

- Website: http://www.monash.edu/equity-diversity/disability/index.html
- Telephone: 03 9905 5704 to book an appointment with a DLO; or contact the Student Advisor, Student Community Services at 03 55146018 at Malaysia
- Email: dlu@monash.edu
- Drop In: Equity and Diversity Centre, Level 1, Building 55, Clayton Campus, or Student Community Services Department, Level 2, Building 2, Monash University, Malaysia Campus