# COMP8230

## Mining Unstructured Data

Session 1, Special circumstances 2021

*School of Computing*

## Contents





**Notice**

As part of Phase 3 of our return to campus plan, most units will now run tutorials, seminars and other small group activities on campus, and most will keep an online version available to those students unable to return or those who choose to continue their studies online.

To check the availability of face-to-face activities for your unit, please go to timetable viewer. To check detailed information on unit assessments visit your unit's iLearn space or consult your unit convenor.

# General Information

Unit convenor and teaching staff
Unit Convenor/Lecturer
Guanfeng Liu
guanfeng.liu@mq.edu.au
Contact via +61-2-9850-9542

Lecturer
Amin Beheshti
amin.beheshti@mq.edu.au
Contact via +61-2-9850-6344

Lecturer
Rolf Schwitter
rolf.schwitter@mq.edu.au
Contact via +61-2-9850-9533

Credit points
10

Prerequisites
COMP6200 or ITEC657

Corequisites

Co-badged status

Unit description
Unstructured data, like text data, graph data, audios, and videos widely exist in our daily life. Efficiently and effectively mining the unstructured data are significant and acting as the backbone in many real applications, like machine translation, face recognition, and link prediction. This unit will introduce key concepts in unstructured data mining, including specific algorithms and techniques for unstructured data cleaning, pattern mining, knowledge discovery, and the prediction of unstructured data. By taking this unit you will be given a broad view of the general issues surrounding unstructured data and the application of methodologies and algorithms to such a type of data. You will have the chance to explore an assortment of unstructured data mining techniques, which you will apply to solve problems involved in real scenarios.

## Important Academic Dates

Information about important academic dates including deadlines for withdrawing from units are

available at https://www.mq.edu.au/study/calendar-of-dates

# Learning Outcomes

On successful completion of this unit, you will be able to:

**ULO1:** Demonstrate an understanding of basic concepts, techniques, algorithms and modellings in unstructured data mining.

**ULO2:** Identify the appropriate data mining techniques and algorithms for real life unstructured data mining problems.

**ULO3:** Explain how good decision making is supported by descriptive and predictive data mining

**ULO4:** Present and analyse the unstructured data mining results with advanced data mining techniques.

**ULO5:** Communicate clearly and effectively

# General Assessment Information

## Details for each assignment will be available via iLearn

You are encouraged to:

- set your personal deadline earlier than the actual one;
- keep backups of all your important files;
- ensure that no-one else picks up your printouts.

## Late submission

No extensions will be granted without an approved application for Special Consideration. There will be a deduction of 10% of the total available marks made from the total awarded mark for each 24 hour period or part thereof that the submission is late. For example, 25 hours late in submission for an assignment worth 10 marks – 20% penalty or 2 marks deducted from the total. No submission will be accepted after the solutions have been posted.

# Assessment Tasks

| Name | Weighting | Hurdle | Due |
|---|---|---|---|
| Weekly Submission | 10% | No | One week after each lecture |
| Report on Data Mining in Industry | 30% | No | Week 5 |
| Problem Analysis | 30% | No | Week 9 |
| Literature Review | 30% | No | Week 12 |

## Weekly Submission

Assessment Type [1]: Quiz/Test
Indicative Time on Task [2]: 6 hours
Due: **One week after each lecture**
Weighting: **10%**

Students will be marked based on their answers on weekly submissions.

On successful completion you will be able to:

- Demonstrate an understanding of basic concepts, techniques, algorithms and modellings in unstructured data mining.
- Identify the appropriate data mining techniques and algorithms for real life unstructured data mining problems.
- Explain how good decision making is supported by descriptive and predictive data mining
- Present and analyse the unstructured data mining results with advanced data mining techniques.

## Report on Data Mining in Industry

Assessment Type [1]: Report
Indicative Time on Task [2]: 18 hours
Due: **Week 5**
Weighting: **30%**

Students will write a report on an aspect of the application of unstructured data mining in an industry context.

On successful completion you will be able to:

- Demonstrate an understanding of basic concepts, techniques, algorithms and modellings in unstructured data mining.
- Present and analyse the unstructured data mining results with advanced data mining techniques.

## Problem Analysis

Assessment Type [1]: Case study/analysis
Indicative Time on Task [2]: 18 hours
Due: **Week 9**
Weighting: **30%**

Students will be given a sample problem and will discuss the relevant data mining techniques and develop a plan to explore the problem.

On successful completion you will be able to:

- Demonstrate an understanding of basic concepts, techniques, algorithms and modellings in unstructured data mining.
- Explain how good decision making is supported by descriptive and predictive data mining
- Communicate clearly and effectively

## Literature Review

Assessment Type [1]: Literature review
Indicative Time on Task [2]: 18 hours
Due: **Week 12**
Weighting: **30%**

Review of work relevant to one of the topics presented in the unit.

On successful completion you will be able to:

- Demonstrate an understanding of basic concepts, techniques, algorithms and modellings in unstructured data mining.
- Identify the appropriate data mining techniques and algorithms for real life unstructured data mining problems.
- Communicate clearly and effectively

---

[1] If you need help with your assignment, please contact:

- the academic teaching staff in your unit for guidance in understanding or completing this

type of assessment

- the Writing Centre for academic skills support.

[2] Indicative time-on-task is an estimate of the time required for completion of the assessment task and is subject to individual variation

# Delivery and Resources

## Classes

Each week has two hours of lectures, and all the lectures will be delivered online. For details of days, times and rooms consult the timetables webpage. There is no workshop/practical class for this unit.

## Required and Recommended Texts

All required and recommended readings will be provided as part of the lecture material.

## Unit Web Page

The unit web page will be hosted in iLearn, where you will need to log in using your Student One ID and password. The unit will make extensive use of discussion boards also hosted in iLearn. Please post questions there, they will be monitored by the staff on the unit.

# Policies and Procedures

Macquarie University policies and procedures are accessible from Policy Central (https://policies.mq.edu.au). Students should be aware of the following policies in particular with regard to Learning and Teaching:

- Academic Appeals Policy
- Academic Integrity Policy
- Academic Progression Policy
- Assessment Policy
- Fitness to Practice Procedure
- Grade Appeal Policy
- Complaint Management Procedure for Students and Members of the Public
- Special Consideration Policy

Students seeking more policy resources can visit Student Policies (https://students.mq.edu.au/support/study/policies). It is your one-stop-shop for the key policies you need to know about throughout your undergraduate student journey.

To find other policies relating to Teaching and Learning, visit Policy Central (https://policies.mq.edu.au) and use the search tool.

## Student Code of Conduct

Macquarie University students have a responsibility to be familiar with the Student Code of Conduct: https://students.mq.edu.au/admin/other-resources/student-conduct

## Results

Results published on platform other than eStudent, (eg. iLearn, Coursera etc.) or released directly by your Unit Convenor, are not confirmed as they are subject to final approval by the University. Once approved, final results will be sent to your student email address and will be made available in eStudent. For more information visit ask.mq.edu.au or if you are a Global MBA student contact globalmba.support@mq.edu.au

# Student Support

Macquarie University provides a range of support services for students. For details, visit http://students.mq.edu.au/support/

## Learning Skills

Learning Skills (mq.edu.au/learningskills) provides academic writing resources and study strategies to help you improve your marks and take control of your study.

- Getting help with your assignment
- Workshops
- StudyWise
- Academic Integrity Module

The Library provides online and face to face support to help you find and use relevant information resources.

- Subject and Research Guides
- Ask a Librarian

# Student Services and Support

Students with a disability are encouraged to contact the Disability Service who can provide appropriate help with any issues that arise during their studies.

# Student Enquiries

For all student enquiries, visit Student Connect at ask.mq.edu.au

If you are a Global MBA student contact globalmba.support@mq.edu.au

# IT Help

For help with University computer systems and technology, visit http://www.mq.edu.au/about_us/offices_and_units/information_technology/help/.

When using the University's IT, you must adhere to the Acceptable Use of IT Resources Policy.

The policy applies to all who connect to the MQ network including students.